



Lemus Villafuerte, Fabiola (2025) *Estimation of the independent causal effect of T2DM on cancer incidence*. PhD thesis.

<https://theses.gla.ac.uk/85506/>

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses

<https://theses.gla.ac.uk/>

research-enlighten@glasgow.ac.uk

Estimation of the independent causal effect of T2DM on cancer incidence

Fabiola Lemus Villafuerte

MSc

**Submitted in fulfilment of the requirements for the degree of Doctor of
Philosophy**

Health Economics and Health Technology Assessment (HEHTA)

School of Health and Wellbeing

College of Medical, Veterinary and Life Sciences

University of Glasgow

April 2025

Abstract

Background

Diabetes mellitus (DM) is a chronic metabolic disease whose prevalence has risen dramatically over the past two decades, increasing from 108 million cases in 1980 to 537 million in 2021, with projections estimating 783 million cases by 2045 (International Diabetes Federation, 2021). Type 2 diabetes mellitus (T2DM) is the most prevalent form, accounting for over 90% of cases globally. In addition to its micro and macrovascular complications, emerging evidence links T2DM to an elevated cancer risk. While the association between T2DM and cancer has been widely studied, no causally interpretable meta-analyses have been published on the relationship between T2DM and cancer risk across a range of cancer types.

Aim

This thesis aims to assess the independent causal effect of T2DM on cancer risk across multiple cancer types. In this context, independent causal effect refers to a relationship that is not influenced by confounding factors.

Methods

A comprehensive and updated systematic review (SR) was conducted to identify evidence on the association between T2DM and cancer risk across various cancer types.

Directed Acyclic Graphs (DAGs) were used to identify the minimum sufficient adjustment set (MSAS) to control for confounders.

Studies that accounted for the MSAS in their adjusted analyses were included in the adjusted (causal effect) meta-analyses, which used the DerSimonian and Laird inverse variance method for random-effects models to synthesize data from multiple studies. Subgroup analyses were performed by sex, race, study design, menopausal status (for breast cancer), time since T2DM diagnosis, and antidiabetic drug classes. Additionally, unadjusted (association) meta-analyses were conducted to examine the relationship between T2DM and cancer risk, with corresponding subgroup analyses. Statistical heterogeneity was assessed using the I^2 statistic, and potential publication bias was examined through funnel plot inspection and Egger's test.

A burden of disease analysis was also conducted to estimate the cancer burden due to T2DM for the cancer types found to be causally associated with T2DM in the meta-analyses.

Findings

A total of 18,086 publications were screened, and 250 studies were included in the SR, covering 30 countries and 63 cancer types. Of these, at least one effect estimate meeting the MSAS criteria was reported for 31 types. Meta-analyses for 17 cancer types revealed statistically significant causal associations for breast (relative risk [RR] 1.11; 95% confidence interval [CI]: 1.04-1.18), colon (RR 1.30; 95% CI: 1.12-1.50), endometrial (RR 1.94; 95% CI: 1.43-2.63), kidney (RR 1.67; 95% CI: 1.40-1.98), pancreatic (RR 1.73; 95% CI: 1.46-2.05), and prostate cancers (RR 0.81; 95% CI: 0.72-0.91). Subgroup analyses showed no statistically significant differences in cancer risk by sex, geographic region (proxy for race), or study design. For breast cancer, the risk due to T2DM was comparable in premenopausal and postmenopausal women.

In the global burden analysis, 2% of cancer cases (105,084 cases) in 2025 are expected to be attributable to T2DM prevalence in 2021 for the cancer types found to be causally associated with T2DM. Endometrial cancer will have the highest burden, with 1 in 20 cases (24,072) attributable to T2DM, followed by pancreatic cancer, with 1 in 25 cases (23,530). Pancreatic cancer is also projected to account for the highest disability-adjusted life years (DALY) and economic burden (INT \$21,844.70 million). Regionally, the Middle East and North Africa is projected to experience the highest burden (2.8%), while Africa will have the lowest (0.7%).

Conclusion

This thesis identifies specific cancer types for which causality with T2DM is interpretable. Recognizing these causally interpretable associations can inform the development of more targeted public health interventions, particularly in cancer screening and prevention strategies for individuals with T2DM.

Table of contents

Abstract	I
Acknowledgements	VIII
Author's Declaration	IX
Presentations	X
Abbreviations	XI
Chapter 1 Introduction	1
1.1 Background	1
1.2 Research question and aim	5
1.2.1 Thesis layout	5
Chapter 2 Context of Evidence on T2DM and Cancer Risk	10
2.1 Summary	10
2.2 Introduction	11
2.3 Methods	13
2.4 Results	13
2.5 Discussion	20
Chapter 3 Systematic review	25
3.1 Summary	25
3.2 Introduction	26
3.3 Methods	27
3.3.1 Eligibility criteria	27
3.3.2 Search strategy for identification of relevant studies	31
3.3.3 Process for study selection and risk of bias assessment	32
3.4 Results	34
3.4.1 Characteristics of included studies	37
3.4.2 Findings related to the study design	40
3.4.3 Sociodemographic characteristics	42
3.4.4 Lifestyle variables	43
3.4.5 Variables included in the adjusted analyses	45
3.4.6 Types of cancer	47
3.4.7 Risk of bias assessment	51
3.5 Discussion	52
Chapter 4 Causal association	60
4.1 Summary	60
4.2 Introduction	60
4.3 Methods	64
4.3.1 Directed Acyclic Graphs (DAGs) development	64
4.3.2 Minimum sufficient adjustment set (MSAS) and software	66
4.3.3 Cancer Types for Which a DAG Was Created	66
4.3.4 Example of a DAG	68

4.4	Results	71
4.4.1	MSAS for the different types of cancer	71
4.4.2	Proportion of effect estimates with the MSAS for each type of cancer ..	76
4.4.3	Cancer types with no effect estimates meeting MSAS criteria.....	78
4.4.4	Proportion of confounders omitted in Non-MSAS Estimates by Cancer type	79
4.5	Discussion.....	83
Chapter 5	Assessing the quality of covariable adjustment in studies making causal claims	92
5.1	Summary	92
5.2	Introduction	92
5.3	Methods	94
5.3.1	Identification of studies claiming causality	94
5.3.2	Assessing Justification for Covariable Selection	100
5.3.3	Assessing the Appropriateness of Covariable Selection in Relation to Study Objectives	101
5.4	Results	101
5.4.1	Classification of Studies by Causal Claims, Covariate Selection, and Justification.....	101
5.4.2	MSAS in studies making explicit (or implicit) causal claims according to cancer type	103
5.5	Discussion.....	108
Chapter 6	Meta-analyses	113
6.1	Summary	113
6.2	Introduction	114
6.3	Methods	116
6.3.1	Meta-analysis software	117
6.3.2	Data synthesis	118
6.3.3	Heterogeneity assessment.....	121
6.3.4	Subgroup analyses	122
6.3.5	Publication bias assessment.....	124
6.4	Results	125
6.4.1	Meta-analyses of adjusted (causal effect) estimates: Overall results	125
6.4.2	Meta-analyses of adjusted (causal effect) estimates: Subgroup analyses	128
6.4.3	Publication bias assessment.....	133
6.4.4	Adjusted (causal effect) versus unadjusted (association) meta- analyses	133
6.5	Discussion.....	134
Chapter 7	Burden of disease.....	145
7.1	Summary	145
7.2	Introduction	146
7.3	Methods	148

7.3.1	Measures of potential impact	149
7.3.2	Number of cancer cases due to T2DM	150
7.3.3	Number Disability-Adjusted Life Years (DALYs) Due to T2DM	151
7.3.4	Monetised DALYs (economic burden)	152
7.3.5	Data sources	153
7.3.6	Further analysis: Number of Cancer Cases Attributable to Confirmed T2DM.....	157
7.4	Results	157
7.4.1	Cancers attributable to T2DM.....	158
7.4.2	Cancers averted to T2DM.....	169
7.4.3	Further analyses: Number of Cancer Cases Attributable to Confirmed T2DM.....	172
7.4.4	Summary of findings and recommendations	175
7.5	Discussion.....	178
Chapter 8	Discussion	185
8.1	Introduction	185
8.2	Main findings	185
8.2.1	Context (Chapter 2) and Systematic Literature Review (Chapter 3)	185
8.2.2	Causal Association (Chapter 4)	188
8.2.3	Assessing the quality of covariable adjustment in studies making causal claims (Chapter 5)	189
8.2.4	Meta-analyses (Chapter 6).....	191
8.2.5	Burden of disease (Chapter 7)	192
8.3	Strengths and Limitations	194
8.3.1	Strengths.....	194
8.3.2	Limitations.....	195
8.4	Future Research and Policy Recommendations	197
8.5	Contribution to knowledge.....	200
8.6	Conclusion	201
Appendices		203
Appendix A. Systematic Review		204
A1.	Scoping search to identify the types of cancer associated with T2DM	204
A1.1	Objective	204
A1.2	Materials and methods	204
A1.3	Results	205
A1.4	Conclusions.....	207
A2.	Systematic Review. Terms and search strategies.....	208
A2.1	CINAHL search strategy	209
A2.2	COCHRANE search strategy	211
A2.3	EMBASE search strategy	214
A2.4	PUBMED Search strategy	216
A2.5	Web of Science search strategy	219

A3. Criteria for screening selection	221
A4. Newcastle-Ottawa Scale (NOS)	223
A5. Reasons for Title and Abstract Screening Exclusions	226
A6. Searching other sources	226
A7. Selected studies	228
A8. Number of Studies by Cancer Type and Study Design	272
Appendix B. Causal association	274
B1. Risk factors	274
B1.1 T2DM	274
B1.2 Types of cancer.....	274
B1.3 Associations between risk factors.....	283
B2. Directed Acyclic graphs.....	285
Lymphoma.....	286
Leukemia	289
Myeloma	292
Sarcoma.....	293
Carcinoma	298
Appendix C. Meta-analyses	342
C1. Unadjusted (association) meta-analyses	342
C2. Forest plots.....	350
C2.1 Adjusted (causal effect) meta-analyses	351
Lymphoma.....	351
Carcinoma	351
Sarcoma.....	355
C2.2 Unadjusted (association) meta-analyses.....	356
Lymphoma.....	356
Leukemia	357
Carcinoma	357
Sarcoma.....	370
C3. Publication bias (funnel plots and Egger tests)	370
C3.1 Unadjusted (association) meta-analyses.....	371
Appendix D. Burden of disease	377
References	386

List of tables

Table 1-1 Overview of the thesis structure (Next: Chapter 2).....	9
Table 2-1 Summary of (Noto et al., 2011)	14
Table 2-2 Summary of (Starup-Linde et al., 2013)	15
Table 2-3 Summary of (Ohkuma et al., 2018).....	16
Table 2-4 Summary of (Ling et al., 2020).....	17
Table 2-5 Overview of the thesis structure (Next: Chapter 3).....	24
Table 3-1 Data Collected for the Main Components of the PICOS Framework	34
Table 3-2 Number of studies on cancer types (including subsites).....	48
Table 3-3 Number of studies on additional cancer types (including subsites) identified beyond the original search strategy, with respective references.	49
Table 3-4 Overview of the thesis structure (Next: Chapter 4).....	59
Table 4-1 Minimum sufficient adjustment sets (MSAS) identified through DAGs for different cancer types.....	72
Table 4-2 Confounders from the MSAS Omitted Across All Adjusted Effect Estimates, by Cancer Type.....	79
Table 4-3 Proportion of Each Confounder Omitted from Effect Estimates That Did Not Meet the MSAS Criteria, by Cancer Type (Color-Coded)	80
Table 4-4 Overview of the thesis structure (Next: Chapter 5).....	91
Table 5-1 Confounders omitted in in Studies Making Explicit Causal Claims That Did Not Meet the MSAS Criteria, by Cancer Type.....	106
Table 5-2 Overview of the thesis structure (Next: Chapter 6).....	112
Table 6-1 Pooled adjusted RR (causal effect) estimates of T2DM on cancer risk, based on studies adjusting for MSAS.	126
Table 6-2 Adjusted RR estimates of the causal effect of T2DM on cancer risk, based on a single study per cancer type, that accounted for MSAS variables.....	127
Table 6-3 Subgroup analyses by study design (cohort and case control) from the adjusted (causal effect) meta-analyses of T2DM on cancer risk.....	128
Table 6-4 Subgroup analyses by sex from the adjusted (causal effect) meta-analyses of T2DM on cancer risk	129
Table 6-5 Subgroup analyses by geographic region (America, Asia, Europe) from the adjusted (causal effect) meta-analyses of T2DM on cancer risk	130
Table 6-6 Subgroup analyses by menopausal status (premenopausal and postmenopausal) from the adjusted (causal effect) meta-analyses of T2DM on cancer risk	131
Table 6-7 Subgroup analyses by T2DM duration (time [years] since diagnosis) from the adjusted (causal effect) meta-analyses of T2DM on cancer risk	131
Table 6-8 Subgroup analyses of antidiabetic drug usage from the adjusted (causal effect) meta-analyses of T2DM on cancer risk	132
Table 6-9 Summary comparing findings from adjusted (causal effect) versus unadjusted (association) meta-analyses of T2DM and cancer risk.....	134
Table 6-10 Overview of the thesis structure (Next: Chapter 7)	144
Table 7-1 PIF, projected number of cancer cases, DALYs loss and economic value of DALYs loss (monetised DALYs) attributable to T2DM in 2025	159
Table 7-2 PIF, expected number of attributable cancer cases, DALYs loss and economic value of DALYs loss (monetised DALYs) attributable to T2DM in 2025 by region.....	162
Table 7-3 PFP, expected number of averted cancer cases, DALYs averted and economic value of DALYs averted (monetised DALYs) to T2DM in 2025 by region.	170
Table 7-4 PIF and projected number of cancer cases attributable to T2DM in 2025 among individuals with a confirmed T2DM diagnosis against those derived in the base case ..	173
Table 7-5 Results of the different measures of disease burden by region and type of cancer: PIF, expected number of attributable cancer cases, DALYs loss and economic value of DALYs loss (monetised DALYs) attributable to T2DM in 2025.....	175

List of figures

Figure 3-1 PRISMA Study flow diagram	36
Figure 3-2 Study Period by Study Design. MRS: Mendelian Randomisation Study.	37
Figure 3-3 Number of Publications by Region and Country.....	39
Figure 3-4 Distribution of Studies by Duration of Follow-up.	40
Figure 3-5 Number of Studies by Exposure Assessment Method.	41
Figure 3-6 Number of Studies by Outcome Assessment Method.	42
Figure 3-7 Variables included in the adjusted analyses of studies in the SR.	46
Figure 3-8 Risk of bias classification by study design..	51
Figure 4-1 Elements of a DAG..	63
Figure 4-2 DAG of the association between T2DM and risk of pancreatic cancer, including potential confounders..	68
Figure 4-3 DAG of the association between T2DM and risk of pancreatic cancer, including potential confounders..	70
Figure 4-4 Proportion of Effect Estimates that included all variables in the MSAS.	77
Figure 5-1 Flowchart to identify studies claiming causality.	95
Figure 5-2 Number and proportion of studies categorized as Explicit Causal Claim and No Explicit Causal Claim, with at least one estimate accounting for the MSAS and providing a justification for variable selection.	102
Figure 5-3 Number of effect estimates by cancer type in studies making explicit causal claims, categorized by whether the MSAS was accounted for (MSAS vs. No MSAS).	105
Figure 6-1 Funnel plot from the adjusted (causal effect) meta-analysis of T2DM and prostate cancer risk.....	133
Figure 7-1 A) Projected number of cancer cases by sex and cancer type, with percentages indicating the share of each cancer type relative to the total cases within each sex. B) Projected cancer-related DALYs by sex and cancer type, with percentages showing the contribution of each cancer type to the total DALYs within each sex.	161
Figure 7-2 A) Projected number of cancer cases by region and cancer type. B) Projected cancer-related DALYs by region and cancer type.....	164
Figure 7-3 Population Impact Fraction (PIF) of All Cancer Incidence in 2025.	165
Figure 7-4 Projected Number of Cancer Cases Attributable to T2DM in 2025..	165
Figure 7-5 Projected Cancer-Related DALY Loss Due to T2DM in 2025.	166
Figure 7-6 Projected Monetised DALY Loss (in Millions) in 2025.....	166
Figure 7-7 Expected PIF by region and type of cancer.....	167
Figure 7-8 PIF and projected number of cancer cases attributable to T2DM in 2025 among individuals with a confirmed T2DM diagnosis compared to those in the base case, by region.....	174

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Professor Neil Hawkins and Professor Kathleen Boyd, for their invaluable guidance, insightful advice, and continuous encouragement throughout my PhD journey. Your shared expertise and thoughtful support have been key in developing this work.

I am also thankful to my colleagues at HEHTA and the PhD students whose support and motivation made my time in Glasgow a memorable experience.

Finally, I am deeply grateful to my family for always being there for me, even from the other side of the world.

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Printed Name: Fabiola Lemus Villafuerte

Signature:

Presentations

Title: Causal Association between Type 2 Diabetes Mellitus and Risk of Cancer,
June 2023, Adelaide, Australia

(shortlisted in the category **Oral Presentations Student**)

Abbreviations

BMI	Body Mass Index
Brain/NS	Brain/Nervous System
BTC	Biliary Tract Cancer
CC	Colon Cancer
CCI	Charlson Comorbidity Index
CI	Confidence Interval
CRC	Colorectal Cancer
CUP	Cancer of Unknown Primary
DAG	Directed Acyclic Graph
DALYs	Disability-Adjusted Life Years
DM	Diabetes Mellitus
ECC	Extrahepatic Cholangiocarcinoma
GBD	Global Burden of Disease
GDP	Gross Domestic Product
GI	Gastrointestinal Cancer
H&N	Head and Neck
HCC	Hepatocellular Cancer
HDI	Human Development Index
HR	Hazard Ratio
ICC	Intrahepatic Cholangiocarcinoma
ICD	International Classification of Diseases
IDF	International Diabetes Federation
INT \$	International Dollars
MENA	Middle East and North Africa
MRS	Mendelian Randomization Study
MSAS	Minimum Sufficient Adjustment Set
N/A	Not Applicable
NHL	Non-Hodgkin Lymphoma
NOS	Newcastle Ottawa Scale
OR	Odds Ratio
PFP	Prevented Fraction of the Population
PIF	Potential Impact Fraction
RC	Rectal Cancer
RCT	Randomised Controlled Trial
RR	Relative Risk
SES	Socioeconomic Status
SIR	Standard Incidence Ratio
SR	Systematic Review
T2DM	Type 2 Diabetes Mellitus
UK	United Kingdom
USA	United States of America
UTC	Urinary Tract Cancer

Chapter 1 Introduction

The aim of this thesis is to explore the causal effect of type 2 diabetes mellitus (T2DM) on cancer risk.

This chapter aims to outline the background behind the thesis, introduce the research question, define its aim, and provide an overview of the thesis structure.

1.1 Background

Diabetes mellitus (DM) is a chronic metabolic disease whose prevalence has risen dramatically over the past two decades, increasing from 108 million cases in 1980 to 537 million in 2021, with projections estimating 783 million cases by 2045 (International Diabetes Federation, 2021). Among its types, type 2 diabetes mellitus (T2DM) is the most prevalent, accounting for over 90% of cases worldwide (International Diabetes Federation, 2021). Beyond its rising prevalence, T2DM increases morbidity and mortality. Individuals with T2DM face a 15% higher risk of premature death and lose an average of 20 years in life expectancy compared to the general population (J. Ye et al., 2023). The impact of T2DM extends beyond mortality, affecting quality of life and ranking as the seventh leading cause of disability-adjusted life years (DALYs) in 2017, up from nineteenth in 1990 (M. A. B. Khan et al., 2020).

The increased morbidity and mortality associated with T2DM is primarily due to microvascular and macrovascular complications (Hossain et al., 2022; Reed et al., 2021). However, emerging evidence also links T2DM to an elevated cancer risk. Major health organizations, including the Centers for Disease Control and Prevention (CDC), American Cancer Society, and American Diabetes Association, report associations between T2DM and colorectal, liver, pancreatic, bladder, breast, and uterine cancer (American Cancer Society, 2021; American Diabetes Association, 2024; Centers for Disease Control and Prevention (CDC), 2023).

In its 9th edition report, the International Diabetes Federation (IDF), a leading authority in global diabetes care, refers to an umbrella review of meta-analyses conducted up to 2013 (International Diabetes Federation, 2019). This review

found statistically significant associations between T2DM and breast cancer (relative risk [RR] 1.20; 95% confidence interval [CI]: 1.12-1.28), endometrial cancer (RR 1.20; 95% CI: 1.12-1.28), colorectal cancer (RR 1.27; 95% CI: 1.21-1.34), and intrahepatic cholangiocarcinoma (RR 1.97; 95% CI: 1.57-2.46). Since 2013, research into the link between T2DM and cancer has expanded, reinforcing and refining these findings (Tsilidis, Kasimis, et al., 2015).

In an early scoping search conducted for this thesis, recent studies were identified that further support the associations between T2DM and cancer risk. For example, (Hossain et al., 2022) found that T2DM increases the risk of breast cancer across multiple molecular subtypes, including Luminal A (RR 1.35; 95% CI: 1.16-1.56), Luminal B (RR 1.43; 95% CI: 0.97-2.11), triple-negative breast cancer (TNBBC) (RR 1.80; 95% CI: 1.36-2.37), and HER2-positive breast cancer (RR 1.69; 95% CI: 1.04-2.76). However, other studies identified in this search found no statistically significant associations, such as the Reykjavik cohort for breast cancer (hazard ratio [HR] 0.95; 95% CI: 0.57-1.56) and a case-control study from the Kaiser Permanente Southern California (KPSC) health system, which found no association between T2DM and colorectal cancer (odds ratio [OR] 0.80; 95% CI: 0.56-1.14) (Maskarinec, Haraldsdóttir, et al., 2019; Schumacher et al., 2021). Additionally, some research suggests that T2DM may have a protective effect against prostate cancer (RR 0.86; 95% CI: 0.80-0.92) (Fernandez et al., 2021).

The scoping search also identified studies investigating the role of antidiabetic medications and T2DM duration in cancer risk. For example, some studies suggest that certain medications, such as pioglitazone, may be associated with an elevated cancer risk (odds ratio [OR] 1.20; 95% CI: 1.05-1.38), while others, such as metformin, show no significant association (OR 1.00; 95% CI: 0.93-1.07) or even exhibit a protective effect (hazard ratio [HR] 0.50; 95% CI: 0.32-0.83) (H. J. Kim et al., 2018; Y.-C. Liu et al., 2019). Moreover, epidemiological studies indicate that the risk of cancer associated with T2DM may vary with disease duration. Higher risks are observed for individuals with T2DM for 4.1-6 years (HR 1.37; 95% CI: 1.25-1.50), while the risk is less pronounced after 15 years (HR 1.04; 95% CI: 0.95-1.14) (Y. Hu, Wu, et al., 2021).

Some studies show an association between T2DM and certain cancers, while others find no significant link. This inconsistency may be due to shared risk factors, such as age, sex, smoking, alcohol consumption, and obesity, which can confound the observed associations (American Diabetes Association, 2024; Garg et al., 2014; Giovannucci et al., 2010; Maddatu et al., 2017; Malhotra et al., 2016). For example, Ling et al. (2020) conducted a systematic review and cohort-level meta-analyses with bias analysis to assess the strength of causal evidence. They hypothesized that unmeasured confounders might still bias the estimated associations, despite the studies adjusting for certain variables. Their analysis suggested a very likely causal association between T2DM and liver, pancreatic, and endometrial cancer incidence, as well as a likely causal association with gallbladder cancer. However, the associations with kidney, colorectal (CRC), and thyroid cancer incidence were less robust to unmeasured confounding (Ling et al., 2020).

Given the presence of shared risk factors, it is essential to control for confounders when drawing causal conclusions from observational study designs. Observational studies are valuable not only for evaluating associations between an exposure and an outcome but also for making causal inferences. However, to determine whether an observed association reflects a cause-and-effect relationship, it is necessary to control for confounders, which are variables that may distort the relationship. This ensures that the independent effect of an exposure is isolated, allowing for a causal interpretation of effect estimates (Hernán, 2018). If the study's objective is solely to explore associations, adjusting for confounders may not always be necessary, as an association merely describes a relationship between two variables.

Distinguishing between associations and causal effects can be challenging, and many publications fail to clarify whether they are establishing a causal effect or merely identifying an association. The misuse of association and causal association terminology is common in research writing (M. A. Han et al., 2022). For example, researchers may state "association with" in the objectives but then present a causal claim in the conclusion, using terms such as "impact" or "attributable."

Identifying studies that explicitly or implicitly claim a causal relationship in observational research is difficult, as authors often use terminology inconsistently, frequently relating to the assessment of associations rather than causality. This lack of clarity can mislead readers and limit the practical application of findings for decision-making (Haber et al., 2022).

There is a large body of evidence on the relationship between T2DM and cancer risk, but it remains unclear whether these relationships reflect associations or causal effects. To address this challenge and support causal inference, a robust approach was adopted in this thesis.

First, a comprehensive systematic review was conducted to gather all available evidence on the association between T2DM and cancer across various cancer types. To identify the variables necessary to control for confounding, the minimum sufficient adjustment set (MSAS) was determined using Directed Acyclic Graphs (DAGs), a tool recommended for identifying confounders in epidemiology. Identifying confounders is particularly challenging in this context due to the complexity of associations between multiple risk factors. DAGs visually represent potential causal relationships between exposures, outcomes, and related variables (Sauer & VanderWeele, 2013).

To quantify the causal effect more precisely, meta-analyses were conducted on studies that accounted for the MSAS, providing a more accurate estimate of the causal relationship (Higgins et al., 2019). Finally, the results of these meta-analyses were complemented with burden of disease analyses, which offer meaningful insights to inform the design of targeted screening initiatives and public health programs aimed at reducing the cancer burden attributable to T2DM.

1.2 Research question and aim

The research question posed and answered in this thesis is:

What is the independent causal effect of T2DM on cancer incidence?

This thesis aims to determine whether T2DM has an independent causal effect on cancer risk across multiple cancer types. In this context, an *independent causal effect* refers to a relationship between T2DM and cancer that is not confounded by shared risk factors such as obesity, age, sex, alcohol consumption, and smoking. This implies that confounding has been addressed through statistical adjustments.

Several biological mechanisms, including insulin resistance, hyperinsulinemia, and hyperglycemia, have been hypothesized to contribute to the relationship between T2DM and cancer (Johnson et al., 2012; Zhao et al., 2024). However, identifying specific causal mechanisms is beyond the scope of this thesis.

1.2.1 Thesis layout

An overview of layout of the thesis is provided below which is made up of seven further chapters.

Chapter 2: Context of Evidence on T2DM and Cancer Risk

Chapter 2 provides an overview of published systematic reviews and meta-analyses examining the relationship between T2DM and cancer risk across multiple cancer types, identified through a systematic search. It outlines the objectives of these studies, distinguishing between those that explore associations and those that infer causality. The chapter also reviews their methodologies, key findings, and limitations, providing context for the research conducted in this thesis.

Chapter 3: Systematic review

Chapter 3 presents a systematic review designed to identify all relevant evidence on the association between T2DM and cancer risk across various cancer types. A comprehensive literature search was conducted across five databases from inception to December 17, 2023, without language restrictions. The risk of bias in the included studies was assessed using the Newcastle-Ottawa Scale (NOS).

The review followed the Population, Intervention, Comparison, Outcome, and Study Design (PICOS) framework. It included studies involving adults aged 18 years or older (Population), comparing individuals with T2DM (Intervention) to those without T2DM (Comparison). The outcome was cancer incidence (Outcome), measured using risk estimates such as relative risk (RR), odds ratio (OR), standardized incidence ratio (SIR), or hazard ratio (HR). Only observational studies (Study Design) were considered for inclusion.

Chapter 4: Causal Association

Estimating causal effects in observational studies requires addressing confounding bias. Therefore, Chapter 4 focuses on identifying the variables that must be controlled for to assess the causal relationship between T2DM and the cancer types examined in the systematic review (Chapter 3).

To achieve this, Directed Acyclic Graphs (DAGs) were developed through targeted literature searches to determine established risk factors for both T2DM and each cancer type. DAGs allowed the identification of the minimum sufficient adjustment set (MSAS), which represents the variables required to control for confounding.

From the studies included in the systematic review, only those that accounted for the MSAS were selected for the meta-analyses conducted in Chapter 6, which aimed to estimate the independent causal effect of T2DM on cancer risk across multiple cancer types.

Chapter 5: Assessing the Quality of Covariate Adjustment in Studies Making Causal Claims

Identifying studies that claim a causal relationship between T2DM and cancer is challenging due to inappropriate use of terminology, which can lead to misinterpretations and misunderstandings of findings.

Chapter 5 evaluates whether studies from the systematic review (Chapter 3) explicitly or implicitly claimed a causal relationship between T2DM and cancer risk by analysing terminology across different sections using a uniquely developed flowchart.

Covariate selection in causal claims should be explicitly justified. In this chapter, covariate selection was assessed based on references to clinical evidence, the use of terms such as risk factors or confounders, and explicit statements indicating reliance on prior clinical knowledge.

Additionally, to determine whether the selected covariates were appropriate given the study's explicit or implicit objectives, the minimum sufficient adjustment set (MSAS) identified in Chapter 4 using Directed Acyclic Graphs (DAGs) was compared with the covariates included in the adjusted analyses of the studies.

Chapter 6: Meta-Analyses

Chapter 6 presents meta-analyses aimed at establishing the causal effect of T2DM on cancer risk. These analyses incorporated effect estimates that accounted for the minimum sufficient adjustment set (MSAS), identified a priori using Directed Acyclic Graphs (DAGs) for each cancer type (Chapter 4).

Subgroup analyses were performed by sex, race, and study design (cohort and case-control). Additional subgroup analyses were conducted based on menopausal status for breast cancer, time since T2DM diagnosis in relation to cancer incidence, and antidiabetic drug classes, depending on data availability.

Furthermore, meta-analyses were conducted to examine the association between T2DM and cancer risk using unadjusted effect estimates and their corresponding subgroup analyses.

Random-effects meta-analyses were conducted using the DerSimonian and Laird inverse variance method. Statistical heterogeneity was quantified with the I^2 statistic, and potential publication bias was assessed through funnel plot inspection and Egger’s test.

Chapter 7: Burden of Disease

Chapter 7 presents estimates of the global burden of cancer causally associated with T2DM for 2025.

To comprehensively quantify this burden, two key metrics were estimated: the potential impact fraction (PIF) and the prevented fraction of the population (PFP). These metrics were used to calculate the number of cancer cases and cancer-related Disability-Adjusted Life Years (DALYs) attributable to or averted by T2DM. Additionally, DALYs were monetised using GDP per capita to assess the economic impact.

The study incorporated findings from the meta-analyses of cancer types causally associated with T2DM and their effect estimates (Chapter 6). Data sources included diabetes prevalence from the International Diabetes Federation (IDF), projected cancer incidence from GLOBOCAN, DALYs from the Global Burden of Disease (GBD) Compare tool (IHME), and projected 2025 GDP per capita from the World Economic Outlook Database (IMF).

Chapter 8: Discussion

Chapter 8 summarizes the main findings of the thesis, outlines its strengths and limitations, and highlights its contribution to knowledge. It provides policy recommendations, suggests directions for future research, and presents the final conclusions.

The following table provides an overview of the thesis structure. The information related to the next chapter is highlighted in yellow.

Table 1-1 Overview of the thesis structure (Next: Chapter 2)

Chapter 1: Introduction. To outline the background of the thesis, introduce the research question and aim, and provide an overview of the thesis structure.

Chapter	Objective	Methods	Key insights
2. Context of Evidence on T2DM and Cancer Risk	To provide an overview of published systematic reviews (SRs) and meta-analyses examining the relationship between T2DM and cancer risk across multiple cancer types	Summarize SR and meta-analyses examining the relationship between T2DM and cancer risk across multiple cancer types	Context for the research conducted in this thesis
3. Systematic review	To identify all relevant evidence on the association between T2DM and cancer risk across various cancer types.	Comprehensive Systematic literature review PICOS framework: men and women aged ≥ 18 years (P); individuals with T2DM (I) compared to those without T2DM (C); cancer incidence (O); observational studies (S).	Types of cancer with evidence on the association between T2DM and cancer risk. Confounders included in the adjusted analyses of the studies. Other*
4. Causal Association	To identify the variables that must be controlled for to estimate the causal relationship between T2DM and cancer across multiple cancer types.	Directed Acyclic Graphs (DAGs) were developed to identify the Minimal Sufficient Adjustment Sets (MSAS)	MSAS for each type of cancer
5. Assessing the quality of covariable adjustment in studies making causal claims	To assess whether studies making causal claims about T2DM and cancer clearly state their intent; justify their choice of covariates; and select them appropriately based on their objectives.	Studies from SR: Analysis of the terminology used across different sections to identify causal claims. Justification of variable selection using clinical evidence or terminology. Comparison of MSAS from DAGs with covariates in adjusted analyses	Emphasize the need for a clear study goal, aligned methods, and consistent terminology
6. Meta-analyses	To establish the causal effect of T2DM on cancer risk	Meta-analyses that account for the MSAS for each cancer type. Subgroup analyses.	Types of cancer causally associated with T2DM
7. Burden of disease	To quantify the future global burden of cancer causally associated with T2DM	Global burden of disease analysis, using the types of cancer causally associated with T2DM (meta-analyses)	Burden of disease metrics: PIF/PFP, the number of cancer cases, DALYs, and monetised DALYs

Cancer incidence (O), expressed as risk ratio (RR), hazard ratio (HR), standardized incidence ratio (SIR), or odds ratio (OR). Other*: Study design, follow-up, country-geographical publications, assessment method of exposure and outcome, Risk of bias assessment. Abbreviations: DALYs: Disability-adjusted life years; PICOS: Population, Intervention, Comparison, Outcome, and Study Design; PIF: population impact fraction; PFP: prevented fraction of the population

Chapter 8: Discussion. To summarize the main findings, outline strengths and limitations, provide policy recommendations, suggest directions for future research, highlight contributions to knowledge, and present final conclusions.

Chapter 2 Context of Evidence on T2DM and Cancer Risk

2.1 Summary

- Diabetes is a growing global health issue, with cases projected to reach 783 million by 2045. Type 2 diabetes mellitus (T2DM) accounts for over 90% of all cases, and cancer is increasingly recognized as one of its complications.
- Observational study designs not only help explore associations but can also support drawing causal conclusions. An association shows a relationship between two variables, while causation implies that one variable directly influences the other.
- Only four published systematic reviews and meta-analyses have examined the relationship between T2DM and multiple cancer types. Three explored associations, and one assessed the strength of causal evidence in the observed associations using bias analysis. There have been no causally interpretable meta-analyses published on the relationship between T2DM and cancer risk across a range of cancer types. This thesis addresses that gap by determining whether T2DM has an independent causal effect on cancer risk across multiple cancer types. Causal evidence can help support interventions aimed at reducing T2DM as a strategy to lower cancer risk.
- Previous reviews had limitations, such as restricting to English-language studies, searching only a few databases, focusing solely on cohort designs, or relying on outdated searches. To address these issues, the next chapter presents an updated and comprehensive systematic review of observational studies (case-control, cohort, and cross-sectional designs) conducted without language or time restrictions across five major databases. This review forms the foundation for the following chapters and supports the thesis findings.

2.2 Introduction

Objective

This thesis aims to determine whether T2DM has an independent causal effect on cancer risk across multiple cancer types. To provide context for this research, this chapter reviews published systematic reviews and meta-analyses that have examined the relationship between T2DM and cancer risk. It outlines the objectives of these studies, distinguishing between those that explore associations and those that attempt to establish causality. The chapter also reviews their methods, main findings, and limitations to highlight existing gaps and provide a basis for the research conducted in this thesis.

Background

Randomized controlled trials (RCTs) are considered the gold standard for establishing causality. This is because randomization reduces bias by evenly distributing known and unknown characteristics across comparison groups. As a result, any differences in outcomes can be attributed to the intervention itself (Foong & Vicendese, 2025). However, RCTs are often not feasible when studying exposures related to rare outcomes or conditions that take many years to develop, such as the relationship between T2DM and cancer. In these cases, observational studies are the only feasible option. These designs not only help explore associations but can also support drawing causal conclusions.

An association refers to a relationship between two variables in which changes in one (the exposure) are related to changes in the other (the outcome). In contrast, causal inference implies that the exposure has a direct effect on the outcome (Hammerton & Munafo, 2021; Schintler & McNeely, 2022).

The causal inference framework, sometimes referred to as the "What if?" theory of causation, often relies on the concept of counterfactuals, also known as potential outcomes (Bours, 2020; Hernán MA, 2020). A counterfactual comparison defines a causal effect as the difference between the outcome a person experienced when exposed or treated and the hypothetical outcome that same person would have had if they had not been exposed. One of these

outcomes is real and observable (factual), while the other is hypothetical and cannot be observed (counterfactual) (Hernan, 2004).

In theory, if all factors except the exposure were held constant in both scenarios and participants were considered exchangeable, any difference in outcomes would be attributed solely to the exposure. However, because both outcomes for the same individual cannot be observed in reality, and individuals are not identical, differences in other factors may influence the outcome. As a result, the observed effect of the exposure may be confounded by other variables (Bours, 2020; Garcia-Huidobro & Michael Oakes, 2017). Therefore, to establish a causal effect, it is necessary to identify and control for confounders. Proper control of confounding variables helps isolate the independent effect of the exposure, allowing for a valid causal interpretation of the results (Hernán, 2018).

Establishing causation provides evidence to guide healthcare decisions and support effective public health interventions and policies (Glass et al., 2013). This is particularly important for major global health challenges such as diabetes mellitus.

Over the past two decades, diabetes has emerged as a significant public health concern, with the number of affected individuals rising from 108 million in 1980 to 537 million in 2021. This number is projected to reach 783 million by 2045 (International Diabetes Federation, 2021). T2DM is the most common form, accounting for over 90% of all diabetes cases worldwide.

In addition to its well-known microvascular and macrovascular complications, T2DM has increasingly been linked to cancer risk (Tomic et al., 2022). Global health organizations, including the CDC, ADA, and ACS, have acknowledged associations between T2DM and cancers of the liver, pancreas, colon, breast, uterus, and bladder. The International Diabetes Federation (IDF) has also highlighted statistically significant associations with breast, endometrial, colorectal, and intrahepatic cholangiocarcinoma (American Cancer Society, 2021; American Diabetes Association, 2024; Centers for Disease Control and Prevention (CDC), 2023; International Diabetes Federation, 2021).

2.3 Methods

To identify **systematic reviews and meta-analyses** examining the relationship between T2DM and the risk of cancer across multiple cancer types, a systematic search was conducted in the PubMed database up to March 24, 2025.

The search strategy was guided by the **PICO framework**: men and women aged 18 years or older (Population); individuals with T2DM (Intervention) compared to those without T2DM (Comparison); and cancer incidence (Outcome).

The search was performed using Medical Subject Headings (MeSH) and title keywords. The search terms included MeSH terms “Diabetes Mellitus” and “Neoplasms,” as well as title keywords “diabetes” and “cancer”:

((diabetes[Title]) AND (cancer[Title])) OR (("Diabetes Mellitus"[Mesh]) AND "Neoplasms"[Mesh])

No restrictions were applied regarding language or publication date. Each retrieved publication was screened by title and abstract, and only those not focusing on a specific type of cancer were selected. This approach aligns with the objective of the thesis, which is to estimate the independent causal effect of T2DM on various cancer types and to assess gaps in the existing body of evidence.

It is important to highlight that Mendelian randomization studies (MRS) were excluded from this review, as they use genetic variants as instrumental variables to estimate causal relationships, which differs from the approach used in this thesis. This thesis focuses on establishing causality by identifying and controlling for confounders that are already considered in published studies (Y. H. Lee, 2020).

2.4 Results

The search retrieved 562 publications. Of these, only four examined the relationship between T2DM and multiple types of cancer (Ling et al., 2020; Noto et al., 2011; Ohkuma et al., 2018; Starup-Linde et al., 2013).

The following tables present a brief summary of each publication, including the title, objective, methods, results, limitations, and conclusion.

Table 2-1 Summary of (Noto et al., 2011)

Title	Significantly increased risk of cancer in patients with diabetes mellitus: a systematic review and meta-analysis
Objective	To conduct a systematic review and meta-analysis to examine the association between diabetes and overall cancer incidence and mortality.
Methods	A literature search was conducted using Medline and the Cochrane Library, covering studies from inception to July 5, 2010. Observational studies, including cohort, case-control, and cross-sectional designs, were included. For the meta-analyses, HRs from cohort studies were pooled using a random-effects model with inverse-variance weighting, and adjusted RRs were used to account for potential confounders. Subgroup analyses were performed by sex. Study quality was assessed using the STROBE statement. To evaluate publication bias, a funnel plot was used. Statistical analyses were performed using Review Manager (RevMan) version 5.
Results	A total of 1,314 citations were identified through the search strategy, with 26 studies included in the meta-analysis. Diabetes was associated with a statistically significantly increased risk of overall cancer incidence compared to those without diabetes (RR, 1.10; 95% CI, 1.04-1.17; $I^2 = 79\%$; $P < .00001$). In subgroup analyses by sex, both men and women with diabetes showed a statistically significant increase in cancer incidence (Men: RR, 1.14; 95% CI, 1.06-1.23; $I^2 = 81\%$; $P < .00001$; Women: RR, 1.18; 95% CI, 1.08-1.28; $I^2 = 83\%$; $P < .00001$). For cancer-related mortality, diabetes was also associated with a significantly higher risk (RR, 1.16; 95% CI, 1.03-1.30). The STROBE assessment indicated that most studies were of fair methodological quality. No publication bias was detected based on funnel plot assessment.
Limitations	The use of population-based databases limited the ability to fully control for confounding factors. Another limitation was that some studies confirmed diabetes through self-reports, which may have led to misclassification. Study populations were heterogeneous, likely due to ethnic diversity; therefore, risk by cancer type may have varied and was not assessed. Additionally, the potential impact of diabetes medications on cancer risk could not be assessed due to limited available data.
Conclusion	Diabetes is associated to a moderate increase in overall cancer incidence and mortality.

*Abbreviations.*HR: hazard ratio; RR: Risk ratio; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Table 2-2 Summary of (Starup-Linde et al., 2013)

Title	CARING (CAncer Risk and INSulin analoGues): The Association of Diabetes Mellitus and Cancer Risk with Focus on Possible Determinants - A Systematic Review and a Meta-Analysis
Objective	The main objective was to assess the association between diabetes mellitus and cancer through a systematic review and meta-analysis. Additionally, meta-regression analyses were performed as a secondary objective to determine whether extracted study characteristics were determinants of cancer risk.
Methods	A systematic review and meta-analyses were conducted using data from seven databases (PubMed, Embase, Cinahl, Bibliotek.dk, Cochrane Library, Web of Science, and SveMed+) with searches completed on November 9, 2012. No restrictions were applied to language or publication year, and all types of observational studies (case-control, cohort, and cross-sectional) were included. A random-effects model (DerSimonian and Laird) was used. Egger's regression test was performed to assess publication bias. The NOS was used to assess the risk of bias. Subgroup analyses by type of cancer, study design, and sex. Meta-regression analyses to evaluate whether extracted study characteristics (e.g., sex, study design, diabetes and cancer ascertainment, and adjustment factors such as age, BMI, smoking, and diabetes duration) influenced cancer risk. Statistical analyses were performed using Stata software (release 8).
Results	A total of 2,021 records were identified, and 193 studies were included in the systematic review (190 studies included in the meta-analyses). The meta-analysis covered 28 cancer types, including BTC and gallbladder, bladder, bone, breast, cervix, colorectal, endometrial, esophageal, kidney, leukemia, liver, lung, lymphoma, melanoma, myeloma, nervous system, NHL, ovarian, pancreatic, prostate, rectal, skin, small intestine, stomach, testes, and thyroid cancer. Statistically significant associations were found for all cancer types except for bone, larynx, lung, melanoma, myeloma, nervous system, and testicular cancer. T2DM was associated with a decreased risk of prostate (RR, 0.85; 95% CI, 0.80-0.91) and skin cancer (RR, 0.91; 95% CI, 0.83-0.99), while all other statistically significant findings indicated an associated increased cancer risk. The largest magnitude of association was observed for liver cancer (RR, 2.13; 95% CI, 1.81-2.50), pancreatic cancer (RR, 2.21; 95% CI, 1.93-2.54), endometrial cancer (RR, 1.81; 95% CI, 1.63-2.01), and biliary tract and gallbladder (RR, 1.69; 95% CI, 1.41-2.03). Subgroup analyses showed similar findings across sex (male and female) and study design (cohort and case-control). No statistical tests for subgroup differences were performed. Meta-regression analyses revealed that BMI was a negative determinant (decreased risk) for lung cancer risk but showed no significant effect on other cancers. Additionally, male gender was identified as a negative determinant for leukemia risk, suggesting that men with diabetes had a lower risk of developing leukemia. The NOS assessment found that 84% of studies were of fair quality (NOS 6-9). Publication bias showed significant bias for colorectal ($p = 0.024$), esophageal ($p = 0.022$), laryngeal ($p = 0.041$), lymphoma ($p = 0.041$), and lung cancer ($p = 0.015$), suggesting underreporting of non-significant results from small studies. Only bone and thyroid cancer meta-analyses did not show significant heterogeneity.
Limitations	Covariates such as HbA1c levels and diabetes duration were not sufficiently reported across studies, making it difficult to determine whether the elevated cancer risk results directly from diabetes or is influenced by other factors, including antidiabetic treatments. Some studies claimed to report only on T2DM and excluded T1DM diabetes by age at diagnosis (e.g., excluding cases diagnosed before age 18); however, the study populations may still include both T1DM and T2DM, potentially leading to misclassification.
Conclusion	Diabetes is associated with an increased risk of cancer.

Abbreviations. NOS: Newcastle-Ottawa Scale; RR: Risk ratio

Table 2-3 Summary of (Ohkuma et al., 2018)

Title	Sex differences in the association between diabetes and cancer: a systematic review and meta-analysis of 121 cohorts including 20 million individuals and one million events
Objective	To conduct a systematic review and meta-analyses to estimate the relative effect of diabetes on cancer risk in women compared to men.
Methods	A systematic review and meta-analyses were conducted using PubMed, with searches completed on December 23, 2016. Only observational cohort studies were included, with all studies adjusted for at least age, and the primary analyses using the maximum set of adjustments available. All-site cancer mortality was assessed as a secondary endpoint. A random-effects model with inverse-variance weighting was applied. Publication bias was evaluated using funnel plots, Egger's test, and Begg's test. The NOS was used to assess the study quality. Statistical analyses were performed using Stata software (release 13).
Results	A total of 6,371 articles were identified through the systematic search, with 108 studies included in the meta-analyses. Data on 50 cancer types were available. Diabetes was associated with a statistically significant increased risk ($p < 0.01$) in 20 cancer sites for women and 18 sites for men. The largest magnitude of association was observed for pancreatic and liver cancer in both sexes. Among men, the pooled risk ratio for pancreatic cancer was 1.94 (95% CI, 1.69-2.23), while for liver cancer, it was 2.22 (95% CI, 1.81-2.72). Among women, the corresponding pooled risk ratios were 1.96 (95% CI, 1.64-2.36) for pancreatic cancer and 2.00 (95% CI, 1.63-2.46) for liver cancer. The pooled adjusted RR for all-site cancer associated with diabetes was 1.27 (95% CI, 1.21-1.32) in women and 1.19 (95% CI, 1.13-1.25) in men. Women with diabetes had a 6% greater risk compared to men with diabetes, with a pooled RRR of 1.06 (95% CI, 1.03-1.09). Diabetes conferred a significantly greater associated risk in women than men for oral cancer (RRR, 1.13; 99% CI, 1.00-1.28), stomach cancer (RRR, 1.14; 99% CI, 1.07-1.22), kidney cancer (RRR, 1.11; 99% CI, 1.04-1.18), and leukemia (RRR, 1.15; 99% CI, 1.02-1.28). In contrast, diabetes was associated with a lower risk in women than men for liver cancer (RRR, 0.88; 99% CI, 0.79-0.99). For cancer mortality, the pooled women-to-men RRR was 1.03 (95% CI, 0.99-1.06; $p = 0.16$). The I^2 statistic for heterogeneity between studies was 66.7%, with no evidence of publication bias.
Limitations	The heterogeneity in diabetes ascertainment, study design, follow-up duration, endpoint definitions, and confounder adjustment may have influenced the results. The studies did not adjust for female-specific factors (e.g. pregnancy, menopausal status, use of hormone replacement therapy), which are also associated with both diabetes and cancer. The absence of data on diabetes duration and glycaemic control limited the ability to conduct more detailed analyses of the relationship between diabetes and cancer risk.
Conclusion	Diabetes is a risk factor for various types of cancer in both sexes. Women with diabetes had a greater risk compared to men with diabetes. The direction and magnitude of these sex differences vary depending on the type of cancer.

Abbreviations. NOS: Newcastle-Ottawa Scale; RR: Risk ratio; RRR: relative risk ratio

Table 2-4 Summary of (Ling et al., 2020)

Title	Association of Type 2 Diabetes With Cancer: A Meta-analysis With Bias Analysis for Unmeasured Confounding in 151 Cohorts Comprising 32 Million People
Objective	To conduct a systematic review and cohort-level meta-analyses, incorporating bias analysis, to evaluate the strength of causal evidence in the observed associations between T2DM and the risk of cancer incidence and mortality.
Methods	A systematic search for meta-analyses and cohort studies was conducted in PubMed, Web of Science, and the Cochrane Library, with searches completed on January 10, 2019. This study updated the umbrella review on T2DM and cancer published in 2015, which included meta-analyses up to December 2013. Only studies published in English were included. Sex stratified analyses were also conducted. Risk of bias was assessed using the NOS, with age, sex, and BMI considered as confounders. A random-effects meta-analysis was performed using the Knapp-Hartung method. Heterogeneity was quantified using the I^2 statistic. Publication bias was assessed using a funnel plot and Egger's test. The bias analysis considered the association between T2DM and cancer likely causal if at least 70% of studies reported an RR exceeding the minimum threshold, accounting for unmeasured confounding with an assumed RR of 1.5 (based on the levels of adjustments in the individual studies). The minimum threshold was set at RR 1.1 (increased risk) or 0.9 (decreased risk) to infer that the association was causal. All statistical analyses were conducted in R for Windows (version 3.6.1).
Results	From 2,138 studies identified in the systematic search, 144 articles (151 cohorts) were included in the final quantitative analysis. Meta-analyses were conducted for 19 cancer types for incidence. The pooled relative risk (RR) for all cancers combined was 1.15 (95% CI, 1.10-1.21). Statistically significant associations were found for all cancers except esophageal, lung, and melanoma. T2DM was associated with a decreased risk of prostate cancer (RR, 0.83; 95% CI, 0.79-0.88). The associations with the largest magnitude were observed for liver cancer (RR, 2.23; 95% CI, 1.99-2.49), pancreatic cancer (RR, 2.09; 95% CI, 1.88-2.33), gallbladder cancer (RR, 1.61; 95% CI, 1.34-1.93), and endometrial cancer (RR, 1.63; 95% CI, 1.41-1.88). In sex-stratified analyses, the association between T2DM and overall cancer risk was higher in women (RR 1.20, 95% CI: 1.11-1.29) than in men (RR 1.11, 95% CI: 1.04-1.19). For cancer-specific mortality, data were only available for pancreatic cancer (RR, 1.67; 95% CI, 1.30-2.14). Significant heterogeneity was observed across cohorts except for thyroid cancer. No clear evidence of publication bias was found in the analyses. Bias analyses showed that the proportion of studies with a true (causal) effect size exceeding an RR of 1.1 was nearly 100% for liver, pancreatic, and endometrial cancers; 86% for gallbladder; 67% for kidney; 64% for colon; and 62% for CRC.
Limitations	Significant statistical heterogeneity was observed across studies, likely due to methodological and clinical differences (study design, outcome assessment, ascertainment of diabetes). Furthermore, the inclusion of only English-language studies may have excluded relevant research on the topic. Some cohort studies did not differentiate between T1DM and T2DM in their inclusion criteria; however, the proportion of individuals with type 1 diabetes was likely very small.
Conclusion	This study suggests a very likely causal association between T2DM and liver, pancreatic, and endometrial cancer incidence, as well as a likely causal association with gallbladder cancer incidence. However, associations with kidney, CRC, and thyroid cancer incidence were less robust to unmeasured confounding.

Abbreviations. NOS: Newcastle-Ottawa Scale; RR: Risk ratio

Summary of the key findings from the selected publications

Four systematic reviews and meta-analyses have examined the relationship between T2DM and cancer risk across multiple cancer types. Three of them (Noto et al., 2011; Ohkuma et al., 2018; Starup-Linde et al., 2013) focused on assessing associations, while one (Ling et al., 2020) aimed to evaluate the strength of causal evidence in the observed associations using statistical methods. Although Ling et al. inferred that certain associations (e.g. with liver and pancreatic cancer) were likely causal based on assumptions about confounding, the study did not establish causal claims.

The first study was conducted by Noto and colleagues in 2011. Unlike the other reviews, it did not present findings by cancer type, as the authors noted population heterogeneity as a limitation.

Regarding the search strategies, the most recent was completed by Ling and colleagues in early 2019. This review was limited to studies published in English, which may have excluded relevant evidence, as acknowledged by the authors. In addition, this review, along with those by Ohkuma and Noto, included only cohort studies. In contrast, the study by Starup-Linde in 2013 also included case-control and cross-sectional studies.

The study by Starup-Linde was the most comprehensive, as it included searches across seven databases, compared to two in Noto's review, one in Ohkuma's, and three in Ling's. It also applied no restrictions on language or publication date and included multiple study designs. However, the search was completed in November 2012, highlighting the need for an updated review.

All meta-analyses applied a random-effects model, as it accounts for both within-study and between-study variability. Study quality was assessed using the Newcastle-Ottawa Scale (NOS) in all reviews except for Noto (2011), which used the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) scale. The majority of the included studies were considered to be of fair quality. No clear evidence of publication bias was found in most analyses, except in the review by Starup-Linde, where bias was detected in 18% (5 out of 28) of cancer types.

Significant heterogeneity was observed across nearly all cancer types in the meta-analyses, likely due to methodological and clinical differences (e.g., study design, outcome assessment, methods of diabetes ascertainment, sample size, and length of follow-up).

The types of cancer covered in these systematic reviews (SRs) and meta-analyses varied:

- Starup-Linde's meta-analysis included **28 cancer types**, with statistically significant associations found for all except seven (bone, larynx, lung, melanoma, myeloma, nervous system, and testicular cancer).
- Ling's meta-analyses were conducted for **19 cancer types**, with statistically significant associations found for all except all except three (esophageal, lung, and melanoma).
- Ohkuma's meta-analysis included data on **50 cancer types**. Diabetes was significantly associated with increased risk in 20 types among women and 18 among men ($p < 0.01$).

Starup-Linde reported a decreased risk for prostate and skin cancer, while Ling also found a decreased risk for prostate cancer; however, the association with melanoma was not statistically significant. Across the meta-analyses by Starup-Linde, Ling, and Ohkuma, the associations with the largest magnitude were observed for liver and pancreatic cancer. Additionally, both Starup-Linde and Ling reported large effect sizes for endometrial and gallbladder cancers.

No statistical tests for subgroup differences were reported by Starup-Linde or Ling; however, both studies concluded that subgroup analyses showed similar findings between males and females. In contrast, Ohkuma found that women with diabetes had a 6% higher cancer risk compared to men, with a pooled relative risk ratio (RRR) of 1.06 (95% CI: 1.03-1.09).

Several limitations were acknowledged by the authors of the SRs and meta-analyses.

- The use of population-based databases limited the ability to fully control for confounding factors. In addition, the degree of confounder adjustment may have made it difficult to determine whether the observed increased cancer risk was directly attributable to diabetes or influenced by other

variables them (Noto et al., 2011; Ohkuma et al., 2018; Starup-Linde et al., 2013).

- Significant statistical heterogeneity was observed across studies, likely due to differences in methodologies (e.g. study design) which could affect the reliability of the findings (Ling et al., 2020).
- The potential impact of diabetes medications and the duration of diabetes on cancer risk could not be assessed due to limited data availability (Noto et al., 2011; Starup-Linde et al., 2013).

2.5 Discussion

Cancer has been recognized as a complication of T2DM, with evidence derived from observational study designs. These studies are valuable not only for evaluating associations between exposures and outcomes but also for supporting causal inferences.

Only four published systematic reviews and meta-analyses have examined the relationship between T2DM and multiple types of cancer (Ling et al., 2020; Noto et al., 2011; Ohkuma et al., 2018; Starup-Linde et al., 2013). Three of these studies aimed to explore the association between T2DM and cancer, while only one specifically evaluated the strength of causal evidence in the observed associations through bias analysis.

Ling and colleagues conducted a bias analysis because, although the included studies adjusted for certain variables, they hypothesized that unmeasured confounders might still have biased the estimated associations (Ling et al., 2020). The aim of this analysis was to quantify how sensitive the results were to the presence of unmeasured confounding.

To simulate the impact of such a confounder, the researchers assigned it a specific strength of association with both T2DM and cancer, referred to as the confounder's relative risk (RR). In this study, they assumed that a hypothetical unmeasured confounder increased the risk of both T2DM and cancer by 1.5 times ($RR = 1.5$), based on the level of adjustment in the individual studies. Using this assumption, they applied a quantitative bias analysis to estimate the proportion of included studies that would still show a true effect size (point estimate)

above a pre-specified threshold ($RR > 1.1$) in the presence of such unmeasured confounding (Mathur & VanderWeele, 2020).

A high proportion (e.g., $\geq 70\%$) would suggest that the observed association is likely robust and potentially causal, whereas a low proportion would indicate that the association could plausibly be explained by residual confounding rather than a true effect of T2DM.

Therefore, the authors of the Ling et al. study did not make causal claims but used the bias analysis to assess the likelihood that the observed associations reflected causal relationships.

As a result, no causally interpretable meta-analyses have been published on the relationship between T2DM and cancer risk across a range of cancer types. This thesis identifies specific cancer types for which causality with T2DM is interpretable.

Establishing causality requires the identification and control of variables that confound the relationship between the exposure and the outcome. This is particularly important because T2DM and cancer share common risk factors, such as obesity and smoking (American Diabetes Association, 2024; Garg et al., 2014; Giovannucci et al., 2010; Maddatu et al., 2017; Malhotra et al., 2016).

Controlling for confounders is essential to isolate the independent effect of the exposure (T2DM), allowing for a valid causal interpretation of the effect estimates (Hernán, 2018).

Identifying variables that may introduce bias into effect estimates is the first step toward understanding and estimating a causal effect. In the context of complex relationships, Directed Acyclic Graphs (DAGs) have been proposed as effective tools for identifying confounders (Howards et al., 2012). This visual and structured approach was used in the thesis to determine the minimum sufficient adjustment set (MSAS) of variables that must be controlled for to remove confounding and accurately estimate the causal effect of T2DM for each cancer type (Tennant et al., 2021). The DAGs were constructed based on targeted literature reviews, which helped strengthen the validity of the findings.

Meta-analyses were then conducted to estimate the causal effect of T2DM on cancer risk, using effect estimates from studies that accounted for the MSAS derived from the DAGs. The results of these meta-analyses directly address the central research question of this thesis.

Subgroup analyses were also conducted to explore potential variations in effect estimates based on participant characteristics and study design (Higgins et al., 2019).

To further explore the impact of T2DM on cancer risk, additional analyses for breast cancer were stratified by menopausal status. Evidence suggests that postmenopausal women with T2DM have a higher risk of breast cancer compared to non-diabetic women (RR 1.25; 95% CI, 1.20 to 1.29) (Liao et al., 2011).

In addition, as noted by the authors of previous systematic reviews, key research questions remain regarding the impact of diabetes medications and the duration of diabetes on cancer risk (Noto et al., 2011; Starup-Linde et al., 2013).

Epidemiological research suggests that cancer risk varies with the duration of the disease. For example, individuals with T2DM for 4.1 to 6 years have an increased risk (HR 1.37; 95% CI, 1.25 to 1.50), whereas after 15 years, the risk appears to decline (HR 1.04; 95% CI, 0.95 to 1.14) (Y. Hu, Zhang, et al., 2021). Additionally, existing evidence suggests that some diabetes medications may influence cancer risk. For instance, pioglitazone has been associated with an increased risk (OR 1.20; 95% CI, 1.05-1.38), while metformin has shown no association (OR 1.00; 95% CI, 0.93-1.07) or even a protective effect (HR 0.50; 95% CI, 0.32-0.83) (H. J. Kim et al., 2018; Y.-C. Liu et al., 2019).

This information was also collected to support additional meta-analyses and help address these remaining research questions.

An updated and comprehensive systematic review was conducted to ensure that all relevant studies on the association between T2DM and cancer risk were identified. This review serves as the foundation for the subsequent chapters and strengthens the overall findings of the thesis.

Previous systematic reviews on this topic had several limitations. Many included only studies published in English, restricted their searches to three or fewer databases (Ling et al., 2020; Ohkuma et al., 2018), focused exclusively on cohort studies (Ling et al., 2020; Noto et al., 2011), or were based on searches conducted more than a decade ago (Starup-Linde et al., 2013).

To address these gaps, the next chapter presents a comprehensive and updated systematic review of observational studies (including case-control, cohort, and cross-sectional designs), conducted without language or time restrictions. The search was carried out across five major databases to ensure a more complete evidence base.

Conclusion

While the association between T2DM and cancer has been widely studied, no causally interpretable meta-analyses have been published on the relationship between T2DM and cancer risk across a range of cancer types. This thesis identifies specific cancer types for which causality with T2DM is interpretable.

Estimating causal effects provides robust evidence to support interventions that reduce T2DM as a strategy to lower cancer risk. This, in turn, can inform more effective public health policies and contribute to improved population health outcomes.

The following table provides an overview of the thesis structure. The information related to the next chapter is highlighted in yellow.

Table 2-5 Overview of the thesis structure (Next: Chapter 3)

Chapter 1: Introduction. To outline the background of the thesis, introduce the research question and aim, and provide an overview of the thesis structure.

Chapter	Objective	Methods	Key insights
2. Context of Evidence on T2DM and Cancer Risk	To provide an overview of published systematic reviews (SRs) and meta-analyses examining the relationship between T2DM and cancer risk across multiple cancer types	Summarize SR and meta-analyses examining the relationship between T2DM and cancer risk across multiple cancer types	Context for the research conducted in this thesis
3. Systematic review	To identify all relevant evidence on the association between T2DM and cancer risk across various cancer types.	Comprehensive Systematic literature review PICOS framework: men and women aged ≥ 18 years (P); individuals with T2DM (I) compared to those without T2DM (C); cancer incidence (O); observational studies (S).	Types of cancer with evidence on the association between T2DM and cancer risk. Confounders included in the adjusted analyses of the studies. Other*
4. Causal Association	To identify the variables that must be controlled for to estimate the causal relationship between T2DM and cancer across multiple cancer types.	Directed Acyclic Graphs (DAGs) were developed to identify the Minimal Sufficient Adjustment Sets (MSAS)	MSAS for each type of cancer
5. Assessing the quality of covariable adjustment in studies making causal claims	To assess whether studies making causal claims about T2DM and cancer clearly state their intent; justify their choice of covariates; and select them appropriately based on their objectives.	Studies from SR: Analysis of the terminology used across different sections to identify causal claims. Justification of variable selection using clinical evidence or terminology. Comparison of MSAS from DAGs with covariates in adjusted analyses	Emphasize the need for a clear study goal, aligned methods, and consistent terminology
6. Meta-analyses	To establish the causal effect of T2DM on cancer risk	Meta-analyses that account for the MSAS for each cancer type. Subgroup analyses.	Types of cancer causally associated with T2DM
7. Burden of disease	To quantify the future global burden of cancer causally associated with T2DM	Global burden of disease analysis, using the types of cancer causally associated with T2DM (meta-analyses)	Burden of disease metrics: PIF/PFP, the number of cancer cases, DALYs, and monetised DALYs

Cancer incidence (O), expressed as risk ratio (RR), hazard ratio (HR), standardized incidence ratio (SIR), or odds ratio (OR). Other*: Study design, follow-up, country-geographical publications, assessment method of exposure and outcome, Risk of bias assessment. Abbreviations: DALYs: Disability-adjusted life years; PICOS: Population, Intervention, Comparison, Outcome, and Study Design; PIF: population impact fraction; PFP: prevented fraction of the population

Chapter 8: Discussion. To summarize the main findings, outline strengths and limitations, provide policy recommendations, suggest directions for future research, highlight contributions to knowledge, and present final conclusions.

Chapter 3 Systematic review

3.1 Summary

- This systematic review (SR) presents a detailed overview of the association between type 2 diabetes mellitus (T2DM) and cancer risk, integrating observational studies without restrictions on time or language. Research in this field has expanded dramatically since 2005, with almost all publications in English.
- Over half of the studies used a cohort design, with 80% classified as low risk of bias. These studies mainly relied on medical records and cancer registries, ensuring higher diagnostic accuracy. Additionally, most cohort studies reported follow-up periods longer than three years, with the largest proportion falling between 11 and 15 years, which strengthened the temporal link between T2DM and cancer risk. In contrast, 35% of the studies were case-control designs and were generally classified as having moderate to low risk of bias due to limitations in diabetes diagnosis and control selection.
- The most commonly studied cancers in relation to T2DM were breast, pancreatic, colorectal, prostate, and lung cancers, aligning with global cancer prevalence in adults. Pancreatic cancer received particular attention due to its high mortality rate. A critical limitation in the existing body of research is its geographical focus, as most studies were conducted in American and European populations, with no studies from Africa, despite projections that Africa will experience the largest increase in diabetes cases by 2045.
- Many studies did not consider modifiable and nonmodifiable risk factors, in their adjusted analyses. Therefore, if the study aims to establish causality and not just evaluate the relationship between T2DM and cancer risk, future research should focus on identifying and controlling for confounding factors specific to each type of cancer.

3.2 Introduction

Objective

The aim of this chapter is to conduct a comprehensive systematic review of the association between T2DM and cancer risk, using five databases without any language or period restrictions.

The protocol for this review is registered with PROSPERO (ID: CRD42021257479).

Background

In the previous chapter it was presented that only four SRs and meta-analyses have examined the relationship between T2DM and multiple types of cancer (Ling et al., 2020; Noto et al., 2011; Ohkuma et al., 2018; Starup-Linde et al., 2013). Three of these studies aimed to explore the association between T2DM and cancer, and only one the one conducted by Ling and colleagues evaluated the strength of causal evidence in the observed associations using statistical methods.

All four reviews acknowledged that confounding factors may have influenced the reported associations.

Noto et al. noted that the use of population-based databases limited their ability to adjust for important confounders such as obesity and lifestyle factors.

Ohkuma et al. did not explicitly discuss causality but pointed out that the limitations of observational studies, particularly the presence of confounding, made it difficult to determine whether the increased cancer risk was directly due to T2DM.

Starup-Linde et al. acknowledged that although individuals with diabetes might have a higher risk of cancer, this association could be due to shared risk factors. As their study focused on exploring associations rather than establishing causality, they recognized the challenge of distinguishing the effect of diabetes from confounding factors.

Finally, Ling et al. provided an approach to inferring causality. They noted that while confounder adjustment is a common method for addressing confounding bias in observational studies, unmeasured confounding may still influence the findings. To assess this, they performed a sensitivity analysis, known as a bias analysis, to evaluate how robust the observed associations were to potential

unmeasured confounders. Based on their findings, they concluded that the associations were either likely robust and potentially causal or, alternatively, could be explained by residual confounding. Therefore, Ling and colleagues used statistical methods to infer causality from observational data but did not make definitive causal claims.

There have been no causally interpretable meta-analyses published on the relationship between T2DM and cancer risk across a range of cancer types. This thesis identifies specific cancer types for which causality with T2DM is interpretable.

In addition, these SRs have had limitations. The systematic reviews that investigated the relationship across various cancer types failed to include records published in languages other than English, limited their search to three or fewer databases (Ling et al., 2020; Ohkuma et al., 2018) or, if they did, the searches were conducted more than ten years ago (Starup-Linde et al., 2013), or focused only on cohort designs (Ling et al., 2020; Noto et al., 2011).

Therefore, a comprehensive and systematic review was conducted. This review formed the foundation for the subsequent chapters and supported the overall findings of the thesis. From this review, the types of cancer with published evidence on their association with T2DM were identified. In the next chapter, confounders for each identified cancer type were determined using Directed Acyclic Graphs (DAGs), as addressing confounding bias was necessary to support causal claims. Meta-analyses were carried out using studies from the systematic review that accounted for the confounders identified through the DAGs.

3.3 Methods

3.3.1 Eligibility criteria

To design the literature research strategy and establish the criteria for including or excluding studies, the Population, Intervention, Comparator, Outcome and Study design (PICOS) framework was used. The PICOS strategy organises the main concepts of the research questions into a list of terms, which allows

identifying clearly and correctly the medical literature for systematic reviewing (Cooper et al., 2018; Methley et al., 2014).

3.3.1.1 Population

The target population was **men and women aged 18 years and over**. This population was preferred since T2DM is more commonly diagnosed in adults than children. In fact, it was formerly called adult-onset diabetes (World Health Organization & International Diabetes Federation, 2006; World Health Organization, 1999).

Therefore, studies including subjects younger than 18 years only or without specifying the outcome in adults were excluded from this review.

3.3.1.2 Intervention and comparator

The focus of this systematic review was to gather all available evidence on the association between T2DM and cancer risk. Thus, **adults with T2DM are the intervention or exposure group, while adults without T2DM constitutes the comparison.**

Publications that considered subjects with other comorbidities or health conditions additional to T2DM as part of their inclusion characteristics (e.g., Non-alcoholic fatty liver disease, NAFLD) were excluded (Nouredin & Rinella, 2015; Wild et al., 2018).

It is essential to mention that the World Health Organisation (WHO) adopted the term *Type 2 diabetes* in 1980 (Chiquete et al., 2001). Consequently, before 1980, several publications did not always distinguish between type 1 or type 2 diabetes, or some authors simply used the term diabetes mellitus. To include all the available information regarding T2DM, if the type was not specified, all people aged 30 years or older at the time of the diagnosis of diabetes were considered type 2, and those <30 years were considered type 1 and excluded from the analysis (Rodbard et al., 2007; World Health Organization, 1994).

Another aspect addressed was the diagnosis of the disease. Only studies that included subjects with a confirmed diagnosis of type 2 diabetes were considered. Diagnosis was based on blood glucose concentration.

Diagnostic criteria for diabetes and prediabetes

Diabetes was defined according to international consensus criteria. A diagnosis is confirmed when any of the following thresholds are met: random plasma glucose or plasma glucose 2 hours after a 75 g oral glucose load ≥ 11.1 mmol/L (200 mg/dL), fasting plasma glucose ≥ 7.0 mmol/L (126 mg/dL), or HbA1c ≥ 48 mmol/mol. In asymptomatic individuals, two positive diagnostic tests are required to confirm diabetes.

Definitions of prediabetes vary between expert bodies:

- **Impaired fasting glucose:** fasting plasma glucose ≥ 6.1 mmol/L (110 mg/dL) and < 7.0 mmol/L (126 mg/dL) (World Health Organization), or ≥ 5.6 mmol/L (100 mg/dL) and < 7.0 mmol/L (126 mg/dL) (American Diabetes Association).
- **Impaired glucose tolerance:** fasting plasma glucose < 7.0 mmol/L (126 mg/dL) and 2-hour plasma glucose after a 75 g oral glucose drink ≥ 7.8 mmol/L (140 mg/dL) and < 11.1 mmol/L (200 mg/dL).
- **HbA1c:** 42-47 mmol/mol (National Institute for Health and Care Excellence, UK), or 39-47 mmol/mol (American Diabetes Association).

These definitions are drawn from internationally recognised diagnostic thresholds (Petrie & Boyle, 2022, p.710).

3.3.1.3 Outcome

The outcome of interest was cancer incidence, reported in subjects with T2DM and without the disease. Hence, the strength of the association of the exposure and the binary outcome between these two groups is typically expressed as risk ratio (RR), hazard ratio (HR), standardised incidence ratio (SIR) or odds ratio (OR), with their confidence intervals (Higgins et al., 2019).

The search in two databases (PubMed and Embase) without restriction in a particular type of cancer resulted in over 70,000 records. Due to the large number of records, a **scoping search** was conducted to identify the cancer types for this systematic review, making the selection process more manageable within the time constraints (the details of this review are shown in Appendix A1).

This scoping search was based on evidence derived from SRs, meta-analyses and narrative reviews related to the association between T2DM and any type of cancer. For this purpose, the PubMed database was used, with no language or time constraints enforced.

For identifying the types of cancer, only the titles and abstracts of SRs and meta-analyses were examined, unless the study sought to investigate more than one type. In that case, as in narrative reviews, the full text was screened to identify all the types of cancer explored in the publication.

From the scoping search, the following **18 types of cancer (and their subsites)** whose incidence has been related to T2DM were identified: bladder, breast, cervical, biliary tract (covering the bile duct and gallbladder), colorectal (including both colon and rectal cancers), endometrial, gastrointestinal cancers (such as those affecting the stomach and esophagus), kidney, liver (including hepatocellular carcinoma), lung, skin, oral, ovarian, pancreatic, prostate, and thyroid cancers, leukemia and lymphoma.

3.3.1.4 Study design

To evaluate the association between T2DM and cancer incidence, **observational studies**, cohort (prospective and retrospective) and case-control studies, were included. These types of studies are suitable for assessing the effects of a naturally occurring exposure, such as T2DM, on cancer risk.

Therefore, although randomized controlled trials (RCT) are considered the gold standard to evaluate the effect of an intervention, for some research questions, such as this one, RCTs are unfeasible and were excluded (Bosdriesz et al., 2020; Dekkers et al., 2019).

Compared to cross-sectional studies, which can only assess associations at a single point in time, cohort (prospective and retrospective) and case-control studies offer the advantage of establishing a temporal relationship between the exposure and the occurrence of disease. All three study types were initially included in the review; however, **cohort and case-control studies were exclusively used for statistical analyses**. Cross-sectional studies were considered as supplementary information for discussion purposes only (Song & Chung, 2010).

Also, mendelian randomisation studies (MRS) were included because they can establish causality between the exposure and the outcome (Y. H. Lee, 2020). However, they use genetic variants as instrumental variables to estimate causal relationships, which differs from the approach used in this thesis. This thesis focuses on establishing causality by identifying and controlling for confounders that are already considered in published studies. Nevertheless, MRS were included as supplemental knowledge to compare their findings against the ones from the meta-analyses with cohort and case-control studies (Chapter 6).

Grey literature (abstracts-only papers, posters, editorials) and information from websites were excluded, but their citations were reviewed. Case reports, in vitro studies, and any design involving animals were also excluded.

3.3.1.5 Geographical context

All languages were considered for this review, and no exclusion criteria were applied to geographical location.

3.3.2 Search strategy for identification of relevant studies

3.3.2.1 Electronic searches

A comprehensive search was conducted in PubMed, Embase (through OVID, since 1947), CINAHL (through EBSCO), Web of Science Core Collection-WOS-(through Clarivate Analytics) and Cochrane Library, from inception to December 17th, 2023, to identify relevant studies evaluating T2DM and the risk of the types of cancer specified in section 3.3.1.3 *Outcome*.

Medical subject headings (MeSH Terms) and/or text words of T2DM and each type of cancer, as well as the outcome of interest, were combined and specified in each strategy. The search terms and the detailed search strategy for each database are shown in Appendix A2.

3.3.2.2 Searching other resources

SRs, meta-analyses, and narrative reviews identified during the title and abstract screening were assessed, and their reference lists were screened for

eligible studies. A total of 178 additional references were screened (see Appendix A6).

3.3.2.3 Managing references

Records collected from the electronic databases were imported and managed using Mendeley Reference Manager v2.43.0.

This software was also used to identify and remove duplicates. Nonetheless, manual detection of duplicates through title screening was also carried out afterwards.

3.3.3 Process for study selection and risk of bias assessment

3.3.3.1 Selection of studies

The screening process was performed in two stages: titles and abstracts and full-text review. Three reviewers took part in the study selection.

The primary author (Fabiola Lemus) independently examined all records according to the eligibility criteria. The other two reviewers (Hanin Farhana and Watcharakorn Riabroi) validated the selection process at both screening levels by evaluating a 5% random sample blindly and independently. Disagreements between individual judgements were solved through discussion.

The **title and abstract screening** of all the references was performed in Rayyan, Intelligent Systematic Review software (Ouzzani et al., 2016). To apply the inclusion and exclusion criteria, a flowchart was developed (see Appendix A3, Figure A-1). Subsequently, the title, author and year of the references selected from the title/abstract screening were exported to Microsoft Excel (version 2013) for the **full-text screening selection process**. Studies published in a language different from English were translated with DocTranslator software which uses Google Translate. For this stage, inclusion and exclusion criteria mainly focused on the objective, target population, exposure and outcome were applied (see Appendix A3, Figure A3-2).

3.3.3.2 Risk of bias

The risk of bias for all records was assessed by one author (Fabiola Lemus) through the **Newcastle Ottawa Scale (NOS)**, which is widely used for observational studies and recommended by Cochrane Collaboration (Cochrane Scientific Committee, 2017). This activity was developed after the data extraction of each paper and once again at the end of this process to ensure consistency.

A specific version of the NOS was applied for cohort and case-control studies. Additionally, as cross-sectional studies were included, an adaptation of this scale based on previous publications was developed (Modesti et al., 2016; Ribeiro et al., 2020).

The NOS groups eight items within three domains (selection, comparability and outcome) (Luchini et al., 2017). The maximum score for cohort and case-control studies is 9 points or stars, and for cross-sectional are 10 points or stars. Studies with an overall score of **7 or more points were considered as low risk of bias**; 4 to 6 as medium risk of bias; and less than 4 as high risk of bias (Fischborn et al., 2023; Gasparini et al., 2022). The NOS used for each study design is provided in Appendix A4.

3.3.3.3 Data extraction

Data from each study were extracted by one person (Fabiola Lemus) in a standardised Microsoft Excel worksheet (version 2013) and were reviewed with additional reviewers/supervisors. Discrepancies on the categories for data extraction and unclear information reported by the authors was solved by discussion (Professor Neil Hawkins, Professor Kathleen A Boyd and Fabiola Lemus).

The data entry of all studies was verified at the end of the process. Missing data was requested twice by email to the study authors. Engauge Digitizer (version 12.1) was used to extract data from graphs not reported in text or tables.

The data collection form was designed by agreement (Professor Neil Hawkins, Professor Kathleen A Boyd and Fabiola Lemus) according to the PICOS framework: population, intervention (exposure) and comparison, outcome, study

design, and additional items of interest. The information gathered for each PICOS key element is displayed in Table 3-1.

Further data on the study's first author, publication year, title, country, language, cohort name, and project or program were also collected.

Table 3-1 Data Collected for the Main Components of the PICOS Framework

Population	• Data source • Study setting • Comparison group definition • Study period • Length of follow-up (cohort studies, mean or median in years) • Sample size (N) • Number of cancer events in each group (cohort studies) or number of T2DM subjects in cases and controls (case-control studies) • Demographics (age in years [mean or median], sex [percentage of male], BMI or obesity, and other health characteristics)
Intervention (exposure)	• T2DM Assessment method • Antidiabetic treatment (class of drug, only diet) • Duration of T2DM (mean or median in years)
Outcome	• Type of cancer (subsite) • Cancer ascertainment method • Controlled variables or confounders in the adjusted analyses • Outcome measure (RR, OR, SIR, HR) with their 95% confidence interval (95% CI). Results: overall and subgroup analyses (by sex, by the duration of T2DM, according to antidiabetic treatment, other)
Study design	• Study design (e.g. case-control, cohort)

Comparison group (C): individuals without T2DM

3.4 Results

Literature searching resulted in **18,086 publications**. Following the removal of duplicates (4,675 citations), 13,411 records were screened for eligibility. After title/abstract screening, 12 918 records were excluded (the reasons for exclusion are shown in Appendix A5). Additionally, after the full-text screening, one study was excluded because, despite two attempts to contact the author, the requested information on the odds ratio (OR) of T2DM and the risk of hilar cholangiocarcinoma was not provided and was not reported in the publication (Cai WK, et. al, 2011).

The full text of the remaining 492 studies was retrieved for detailed evaluation. Of these, 286 studies were subsequently excluded for the reasons outlined in Figure 3-1. Thus, 206 studies were eligible for this review.

Furthermore, all the 85 records labelled as "References" in the title/abstract screening exclusion criteria were downloaded to examine their reference lists. This process resulted in 497 new records. After applying the inclusion criteria to these 497 records, only 43 were ultimately selected.

To sum up, a total of 249 records (250 studies as one publication included a case-control and a cohort design) were included in this systematic review.

The search strategy results and the summary of different phases of the systematic review, according to the inclusion and exclusion criteria, are shown in the PRISMA flow diagram (Figure 3-1). An overview of the findings of the 250 studies and the risk of bias assessment is presented afterwards. Detailed information of the selected studies is presented in Appendix A7.

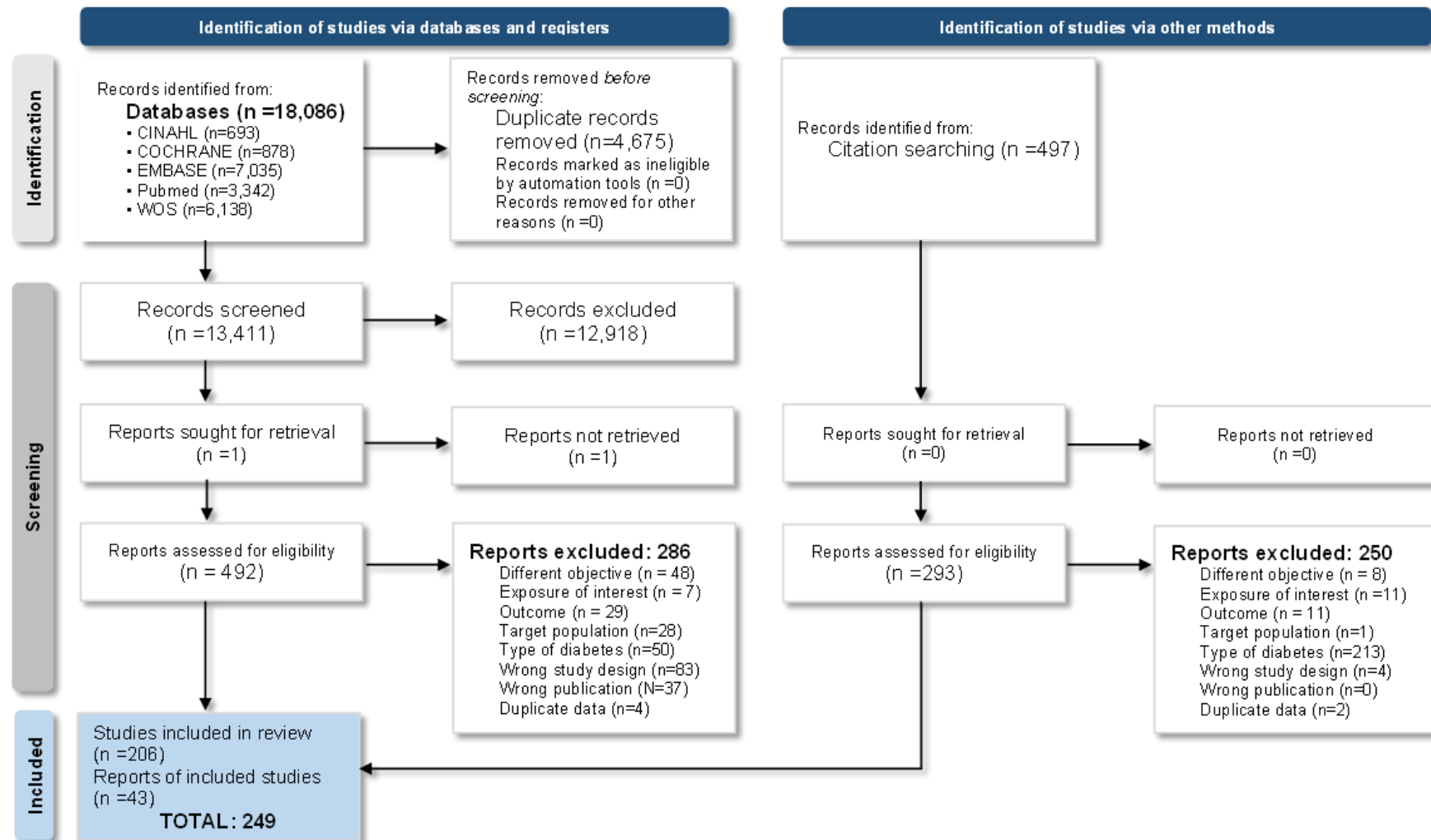


Figure 3-1 PRISMA Study flow diagram

3.4.1 Characteristics of included studies

Of the 250 studies selected, 86 (34%) had a case-control design, 141 (57%) were cohort studies (67 prospective, 74 retrospective), 7 (3%) were cross-sectional studies, and 16 (6%) used a mendelian randomization method.

The first study was published in 1992 (De Mesquita et al., 1992) however, 90% of the studies were published since 2005. Figure 3-2 presents a graph illustrating the growth in literature in this area over time by study type. The graph shows that the number of publications increased since 2009, with at least nine records published annually. In 2013, there was a peak in the number of publications, primarily due to the 23 cohort studies. The first record on the topic using MRS was released in 2015.

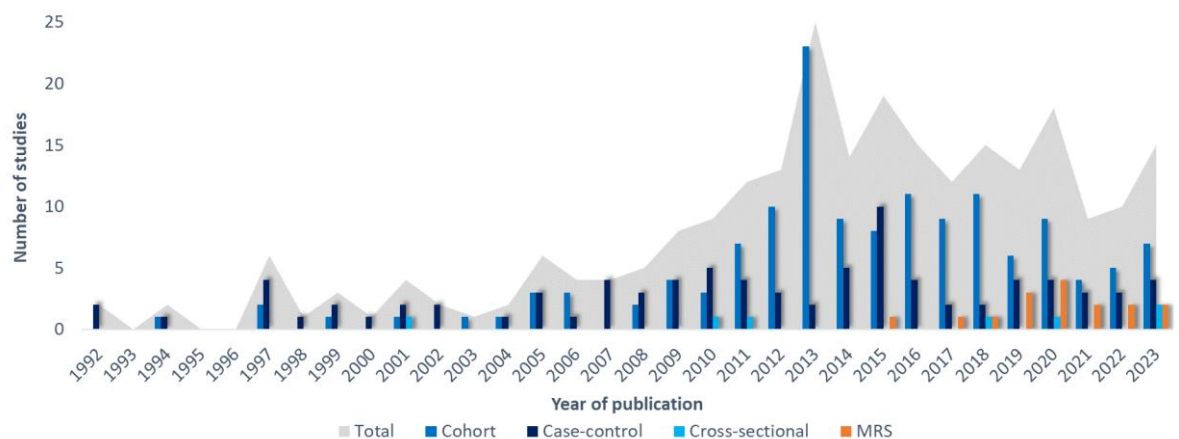


Figure 3-2 Study Period by Study Design. MRS: Mendelian Randomisation Study.

All but five studies (2 Chinese [C. et al., 2023; Liang et al., 2008], 1 German (Kajüter et al., 2014) 1 Polish [Krasnodębski et al., 2010], and 1 Spanish [Castillo-Otiniano & Yan-Quiroz, 2019]) were published in English.

The non-English were translated via Google Translate. The quality of the translated text allowed the data extraction from these papers.

From thirty different countries, at least one article on the association between T2DM subjects and cancer risk was published. America was the continent with more studies (N=86), followed by Europe (N=84), Asia (N=63) and finally Oceania (N=4). Twenty studies considered populations for more than one country within the same or different continent (Au Yeung & Schooling, 2019; Bosetti et al., 2014; Carreras-Torres et al., 2017; J. Chen et al., 2023; Disney-Hogg et al., 2018; Johansson et al., 2019; A. E. Khan, Gallo, Linseisen, Kaaks, Rohrmann, Raaschou-Nielsen, Tjonneland, et al., 2008; Lu et al., 2020; Maisonneuve et al., 2010; Molina-Montes et al., 2020; Nead et al., 2015; Nodari et al., 2023; Petrick et al., 2019; Rosato et al., 2016; Schlesinger et al., 2013; Stott-Miller et al., 2012; Tavani et al., 2002; Tsilidis, Allen, et al., 2015; Yarmolinsky et al., 2019; Yuan et al., 2020).

None of the studies retrieved evaluated this association in any population on the African continent.

In Figure 3-3 the number of publications by region and country is presented.

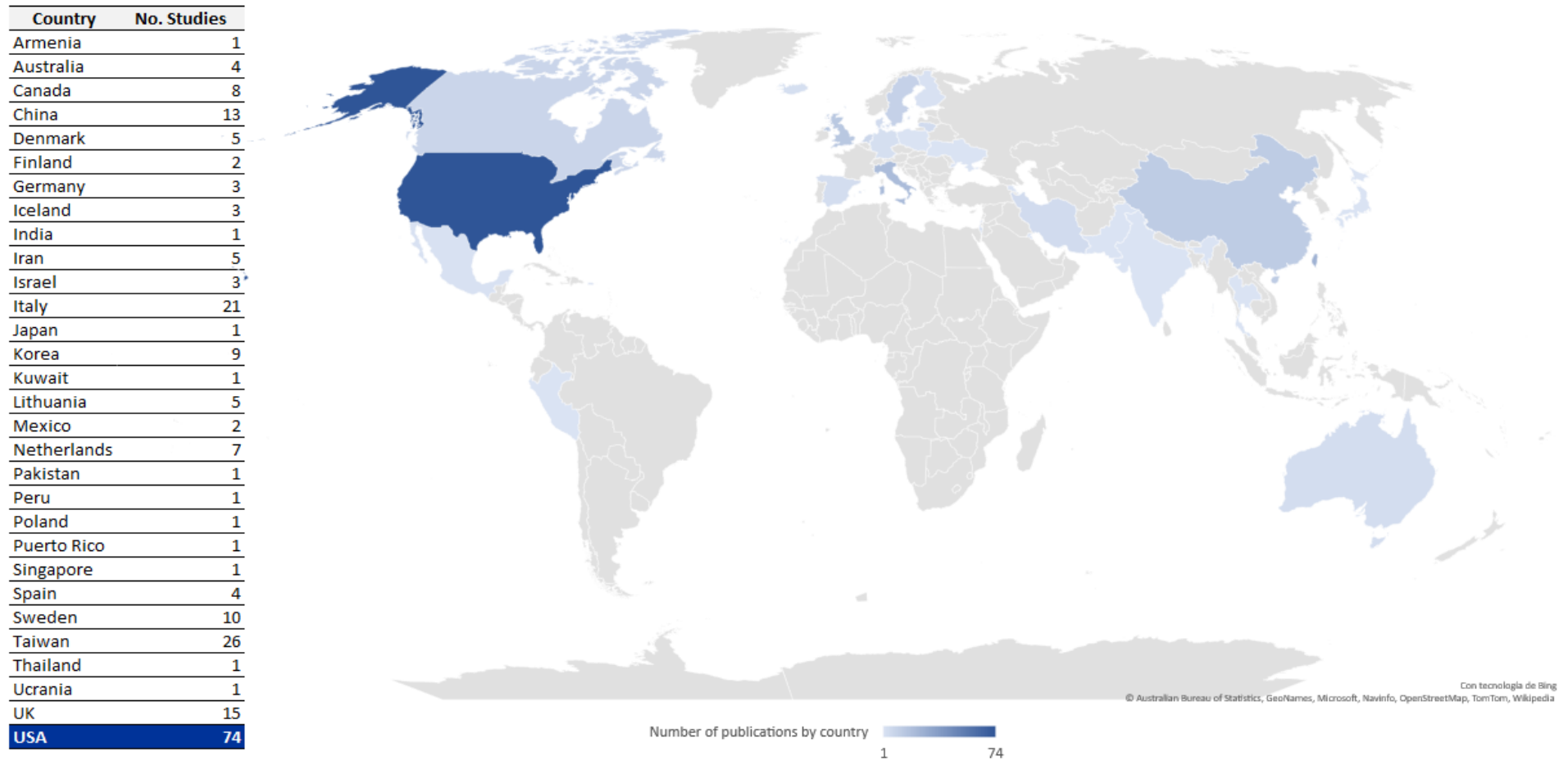


Figure 3-3 Number of Publications by Region and Country. The United States of America (USA) had the highest number of publications (N=74). Twenty studies included populations from multiple countries (not shown in the graph)

3.4.2 Findings related to the study design

Comparison group selection in case-control studies.

Sixty-three studies (73%) had a matched case-control design, usually based on age and sex. In 31 of these studies, the controls were randomly selected. The maximum number of controls per case was 10 (Yang et al., 2005).

Cohort studies follow-up

Of the 141 cohort studies, 112 reported the length of follow-up, referring to total duration rather than mean or median. Two studies had a follow-up of one year, eight reported three years, and the rest had longer follow-up periods, with some extending up to 32 and 38 years.

The mean or median follow-up duration was reported separately in only 61 studies. The reported median follow-up ranged from a minimum of 3.1 years (interquartile range [IQR], 2.1-3.9 years) (Kajüter et al., 2014) to a maximum of 27 years (IQR, 15-50 years) (He, et. al 2022). For the mean follow-up, the minimum was 2.78 years (standard deviation [SD], +0.92) (M.-S. Lai et al., 2006) and the longest follow-up was 28.4 years (SD, +11.6 years) (Maskarinec, Einarsdottir, et al., 2019).

The following graph presents the number of studies according to their follow-up duration. As shown, over 90% of the studies had a follow-up period longer than three years. The largest proportion of studies (30 studies) had follow-up durations between 11 and 15 years.

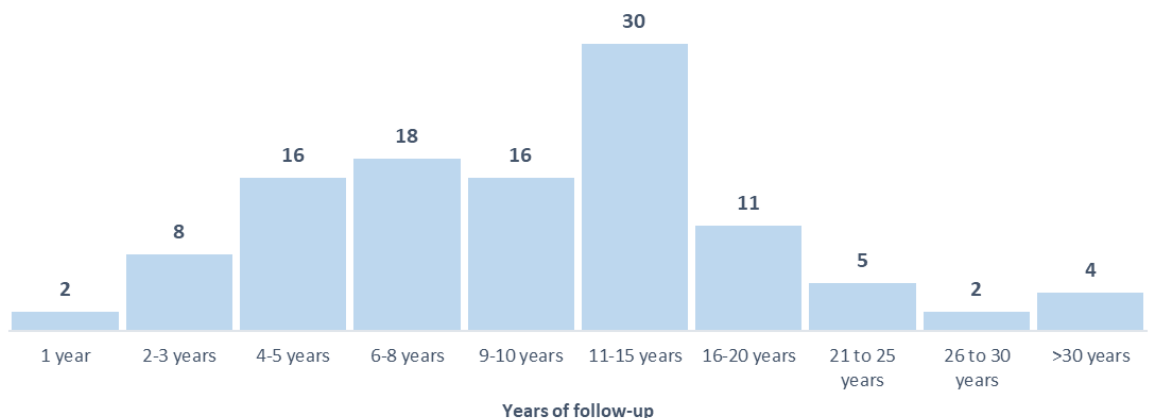


Figure 3-4 Distribution of Studies by Duration of Follow-up. This bar chart presents the number of cohort studies grouped according to their reported follow-up duration. Most studies reported more than 3 years of follow-up, with the highest proportion (30 studies) falling in the 11-15 years range.

Exposure Assessment Method

In 68% of the cohort studies and 80% of the case-control designs, the diagnosis of T2DM was based on only one method; the remaining studies used several methods to confirm the diabetes status. Among the cohort studies that used only one method, 46% were based on an available database, while 68% of the case-control designs collected this data through self-report.

The methods used to assess the diagnosis of T2DM according to the study design are shown below in Figure 3-5.

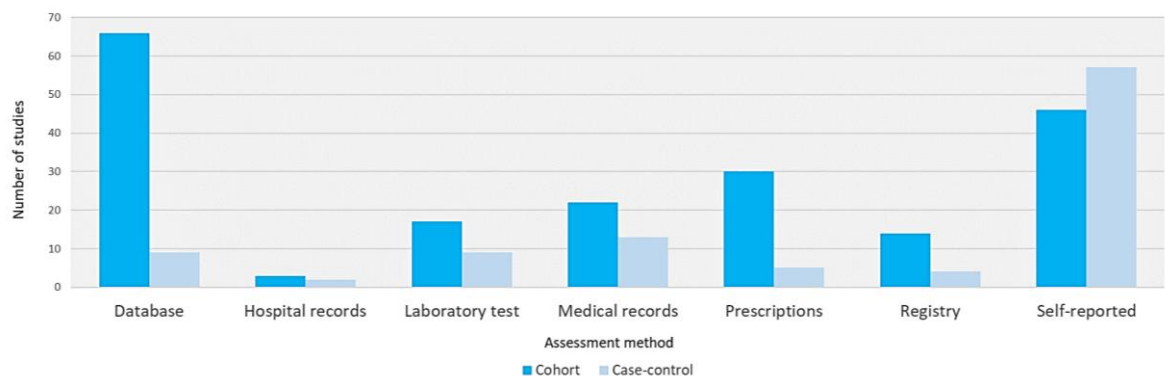


Figure 3-5 Number of Studies by Exposure Assessment Method. The number of studies exceeds the total since several studies used more than one method. Four case-control studies did not report the assessment method. Laboratory test: fasting plasma glucose level (IFG). Database (e.g. ICD-9 CM codes, UK Clinical Practice Research Datalink: General Practice Research Database). Self-reported: interview or written questionnaire. Prescriptions: antidiabetic drugs.

Outcome Assessment Method

In 77% of the cohort studies and 79% of the case-control designs, the cancer diagnosis was based on only one method and the remaining studies used several methods. Among the cohort studies that used only one method, 50% were based on registries, while 66% of case-control designs collected this data through pathology reports.

The methods used to assess the diagnosis of cancer according to the study design are shown in the figure below.

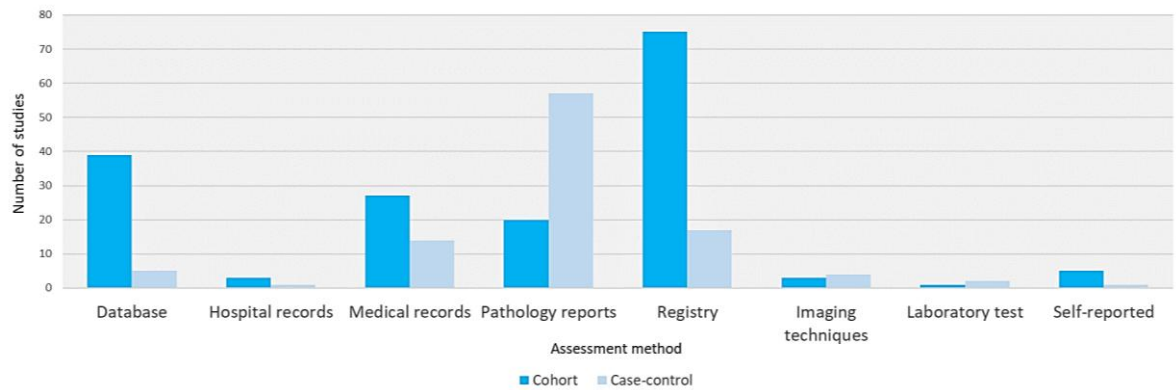


Figure 3-6 Number of Studies by Outcome Assessment Method. The number of studies exceeds the total since several studies considered more than one method. Four studies did not report the assessment method. Imaging techniques: Abdominal ultrasonography for HCC. Laboratory test: alpha-fetoprotein level greater than 400 mg/nL for HCC. None of the studies used a laboratory test or self-reported diagnosis as the only method to confirm the diagnosis of cancer.

3.4.3 Sociodemographic characteristics

Age

Most of the studies reported that the mean age of the subjects was between 50 and 60 years. Seventeen studies reported a statistical difference in the mean age between the comparison groups (Bao et al., 2018; Choi et al., 2022; Cleveland et al., 2012; de Kort et al., 2016, 2017; H. B. El-Serag et al., 2001; García-Esquinas et al., 2016; Hatia et al., 2023; X. He, Shi, et al., 2018; X. He, Wu, et al., 2018; Idilbi et al., 2013; S. K. Kim, Jang, Kim, Rhyu, Song, et al., 2020; W. P. Koh et al., 2013; H.-S. Liu et al., 2022; Rollison et al., 2007; Tseng, 2015; W. S. Yang et al., 2013). However, age was included as an adjustment variable in all but one (Idilbi et al., 2013).

Sex

One hundred forty-eight studies included both males and females. In these studies, the percentage of males ranged from 28.6% (Y. Hu, Wu, et al., 2021) to 97.5% (Jamal et al., 2009).

Of the 158,237,540 subjects in the studies that specified the sample in all comparison groups and considered both sexes, 78,152,295 participants were men (49.4%).

Ethnicity

Individuals from a specific racial or ethnic group were the study population in a limited number of publications related to the Black Women's Health Study (BWHS) and the Multiethnic Cohort (MEC). Three of the six publications evaluated the association of T2DM with breast cancer (Bosco et al., 2012; Maskarinec, Haraldsdóttir, et al., 2019; Palmer et al., 2017); one focused on endometrial cancer (Sponholtz et al., 2016); and the last two studies investigated bladder cancer (Bogumil et al., 2023), and NHL (Maskarinec et al., 2023). The findings were statistically significant in all but two studies (Bosco et al., 2012; Maskarinec et al., 2023).

Postmenopausal women

The study population of ten cohorts and two case-control studies consisted of postmenopausal women (Anderson et al., 2001; Arthur et al., 2019; Cerhan et al., 1997; Gapstur et al., 2012; García-Esquinas et al., 2016; Henry, 2014; Limburg et al., 2005; Lipscombe et al., 2006; Nicodemus et al., 2004; Prizment et al., 2013; Sellers et al., 1994; Weiderpass et al., 2000).

Six studies addressed sex-specific cancers, such as breast, ovarian and endometrial cancer. The remaining publications evaluated the relationship between T2DM and the risk of colorectal, pancreatic, NHL, kidney, and bladder cancer.

Veterans

The veteran population was another focus group in four publications (Brinton et al., 2010; H. El-Serag, 2001; Jamal et al., 2009; Makhoul et al., 2016). All of these studies were from the USA. One considered only the male population for the study of breast cancer (Brinton et al., 2010).

3.4.4 Lifestyle variables

Smoking

Smoking was one of the most commonly reported lifestyle variables in cohort studies. The proportion of smokers in the diabetic group was lower than in the non-diabetic group in 11 of the 14 studies that included this information (S. He et al., 2022; F. B. Hu et al., 1999; Idilbi et al., 2013; Jamal et al., 2009; W.-P.

Koh et al., 2013; Lawler et al., 2023; Maskarinec et al., 2017; Maskarinec, Haraldsdóttir, et al., 2019; Newton et al., 2013; Pan et al., 2018; Simon et al., 2018).

Alcohol consumption

Alcohol consumption was another baseline characteristic considered in the studies. As expected, the focus of most of these publications was the impact of T2DM on pancreatic and hepatocellular carcinoma, as it has been shown to be a risk factor for these types of cancers. In four of the six case-control studies in which this variable was reported, cases showed a higher proportion of alcohol abuse than controls (Donadon et al., 2010; La Torre et al., 2014; Sheng et al., 2023).

Hypertension

The prevalence of hypertension was also notably higher among subjects with diabetes than among nondiabetics in most studies. The difference between groups varied from 9.6% vs. 4.6% ($p < 0.0001$) (H.-C. Lai et al., 2013) to 11% vs 57.3% ($p < 0.01$) (Tseng, 2015).

Hyperlipidemia or dyslipidaemia

The proportion of hyperlipidemia or dyslipidaemia was greater in the diabetic group than in the comparison group in the ten cohort studies that included this variable (C. H. Chen et al., 2020; Choi et al., 2022; A. E. Khan, Gallo, Linseisen, Kaaks, Rohrmann, Raaschou-Nielsen, Tjønneland, et al., 2008; D. Y. Lee et al., 2018; M. Y. Lee et al., 2012; S. F. Lo et al., 2013; Y. Pan et al., 2020; Simon et al., 2018; Tseng, 2015; F. Ye et al., 2023).

The most marked difference was reported by (Tseng, 2015)(51.2 % vs. 7.4%, $p < 0.01$).

Other baseline characteristics

Body Mass index (BMI)/obesity

Among the baseline characteristics, BMI or obesity was reported in 128 studies. A statistically significant difference in the proportion of obesity between the T2DM and non-T2DM groups was observed (e.g., 25.47% vs. 4.49%, $p < 0.0001$), (Brodovicz et al., 2012) in all cohort studies.

Other characteristics

The difference in the proportion of subjects with a history of pancreatitis, cirrhosis, and hepatitis was considered in studies on pancreatic cancer or hepatocellular carcinoma, except for two publications focused on CRC risk (C. H. Chen et al., 2020), and bladder and kidney cancer (Choi et al., 2022). Most of the studies reporting this data had a case-control design, with a higher proportion of cases presenting these characteristics (Bosetti et al., 2014; H.-P. Chen et al., 2013; H. B. El-Serag et al., 2001; la Vecchia et al., 1997; Miele et al., 2015; E. J. Walker et al., 2015; Welzel et al., 2007).

All cohort studies showed that T2DM were more likely to present these comorbidities than non-diabetics (Brodovicz et al., 2012; C. H. Chen et al., 2020; Choi et al., 2022; de Jong et al., 2018; W.-Y. Lin et al., 2014; F. Ye et al., 2023).

Other variables with a statistically significant difference ($p < 0.05$) between T2DM and non-diabetics were related to cardiovascular diseases, such as coronary heart disease, cerebrovascular disease, and peripheral arterial disease (Bjornsdottir, Rawshani, Rawshani, Svensson, et al., 2020; Choi et al., 2022; Knapen et al., 2013; Sikdar et al., 2013; Tseng, 2015).

3.4.5 Variables included in the adjusted analyses

The selection of variables included in the adjusted analyses to examine the association between T2DM and the risk of cancer was diverse. The most common variables were sociodemographic covariates, followed to a lesser extent by behavioural characteristics. Although less frequent, the presence of comorbidities and the use of certain drugs were also included. Additionally, breast and endometrial cancer studies controlled for oestrogen-related variables.

A graph illustrating the variables included in the adjusted analyses and the proportion of studies that accounted for each variable is shown below.

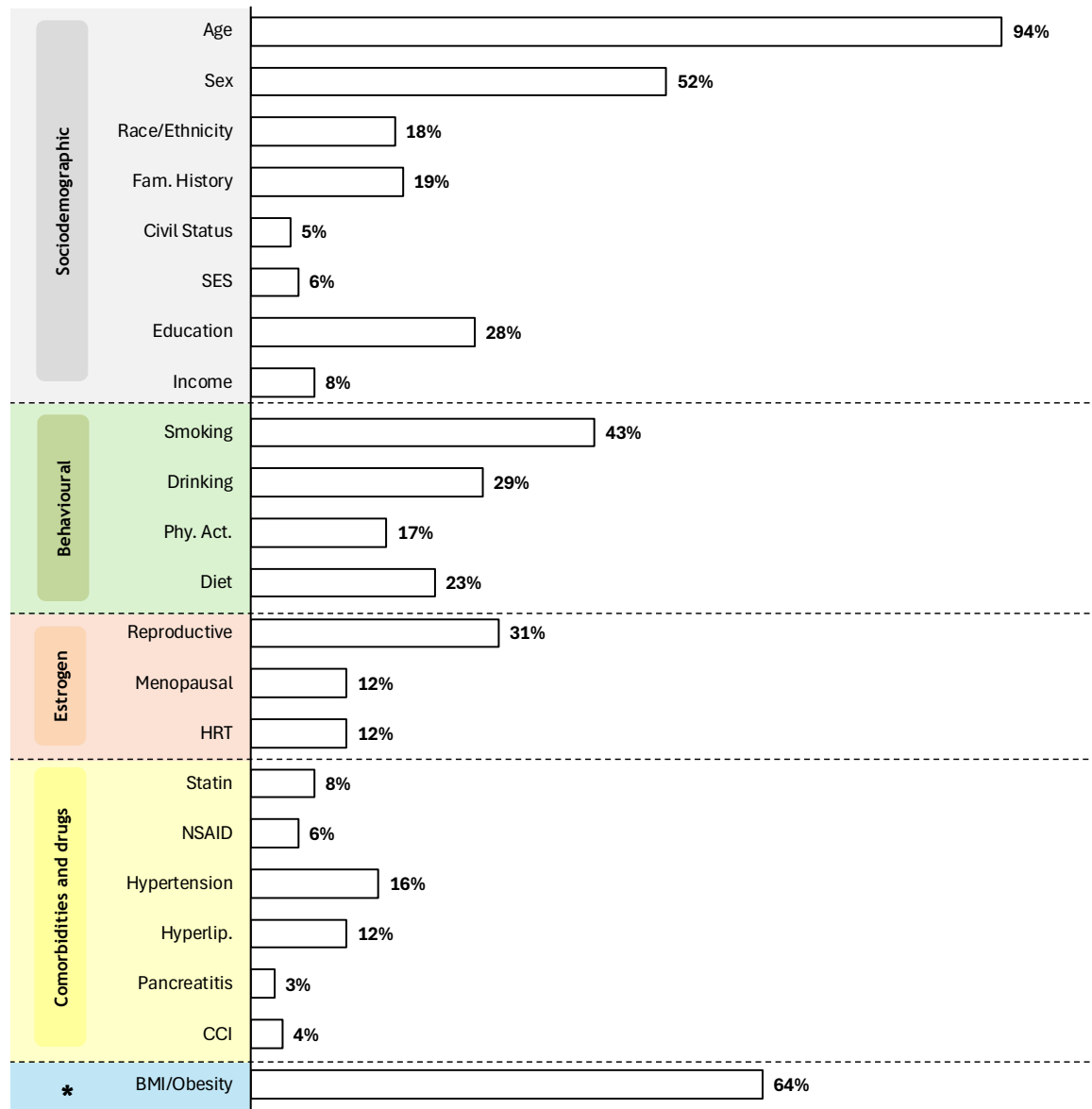


Figure 3-7 Variables included in the adjusted analyses of studies in the SR. Other variables, such as endoscopy history, colonoscopy examinations, regular checkups, and county, were rarely considered (data not shown in the graph). *BMI/obesity is a condition influenced by a variety of factors, including genetic and metabolic components. Diet: diet high in fat, total energy intake, calorie intake, fruit intake, vegetable intake, bacon, tomato sauce). Physical Activity: regular exercise, average hours per week of moderate or vigorous activity. *Abbreviations.* CCI: Charlson Comorbidity Index; HRT: hormone replacement therapy; Hyperlip.: hyperlipidemia; NSAID: of non-steroidal anti-inflammatory drugs; Phy. Act.: physical activity SES: Socioeconomic status.

Among sociodemographic variables, nearly all studies controlled for age (N=201), and approximately half adjusted for sex (N=111). The level of education was also a covariate included but in less than 30% (N=59) of the analyses. A similar proportion of studies included a family history of cancer (19% [N=41]) and race or ethnicity (18% [N=39]) in their adjustments. Fewer than 10% of the studies (N=18) controlled for income, and about 5% controlled for socioeconomic status (N=13) and civil status (N=10).

Regarding behavioral traits, 43% (N=92) of the studies controlled for smoking, alcohol consumption (drinking) was included in 29% (N=61) of the adjustments, and diet-related covariates were found in 23% (N=49). Lastly, less than 20% (N=37) of the studies considered physical activity variables.

Covariates associated with women's reproductive health, such as the number of full-term pregnancies or breastfeeding history, were included in 31% of the analyses (N=66). Menopausal status or age at menopause and the use of hormone replacement therapy (HRT) were other variables controlled for in 12% (N=25) of the studies.

Hypertension and hyperlipidemia were the comorbidities considered in the analyses of 16% (N=34) and 12% (N=26) of studies, respectively. A history of pancreatitis was included in 3% (N=6) of the studies, where the association between T2DM and the risk of pancreatic cancer was evaluated. Multimorbidity of the subjects was also measured and incorporated in 4% of the analyses (N=9) through the Charlson Comorbidity Index (CCI). Finally, 8% (N=18) of the studies controlled for the use of statins, and 6% (N=12) controlled for the use of non-steroidal anti-inflammatory drugs (NSAIDs).

Finally, 64% (N=136) of the studies controlled for BMI or obesity, a condition influenced by a variety of factors, including genetic and metabolic components (Halder & Melkani, 2025).

3.4.6 Types of cancer

The search conducted for 18 types of cancer identified from the **scoping search** resulted in **45 types of cancer (including subsites)**. The following table shows the number of studies found for each type and subsite of cancer. For example, for skin cancer, 10 studies reported the association between T2DM and overall skin cancer risk. Additionally, 10 studies specifically examined the association between T2DM and melanoma, while 2 studies focused on non-melanoma skin cancer.

It is important to note that of the 250 studies gathered from the SR, 72 studies evaluated the association between T2DM and several types of cancer in the same study, which exceeds the numbers presented in the results. In addition, four studies did not focus on a specific type of cancer (Baur et al., 2011; Berger et al., 2016; Hadaegh et al., 2020; S. He et al., 2022).

Table 3-2 Number of studies on cancer types (including subsites)

No.	Type of cancer	Subsite	No. Studies
1	Bladder cancer		29
2		Urinary tract cancer	6
3	Breast cancer		67
4	BTC		10
5		Gallbladder cancer	13
6	Cervix uteri cancer		17
7	CRC		56
8		Colon cancer	34
9		Rectal cancer	29
10		Anal cancer	3
11		Rectosigmoid cancer	1
12	Endometrial cancer		35
13	GI		6
14		Esophageal cancer	26
15		Esophagogastric junction adenocarcinoma	1
16		Stomach cancer	26
17		Large intestine cancer	1
18		Small intestine cancer	8
19	Kidney cancer		26
20	Leukemia		11
21		Acute myeloid leukaemia	1
22		Chronic Lymphocytic Leukemia	2
23		Chronic myeloid leukaemia	1
24		Lymphocytic leukaemia	1
25		Myeloid leukaemia	1
26	Liver cancer		32
27		HCC	17
28		Extrahepatic cholangiocarcinoma	3
29		Intrahepatic cholangiocarcinoma	4
30	Lung cancer		36
31	Lymphoma		5
32		Hodgkin lymphoma	6
33		Non-Hodgkin lymphoma	18
34	Skin		10
35		Melanoma	10
36		Non-melanoma	2
37	Oral cancer		7
38		Mouth cancer	2
39		Lip cancer	2
40		Salivary gland cancer	2
41		Tongue cancer	3
42	Ovarian cancer		20
43	Pancreatic cancer		59
44	Prostate cancer		51
45	Thyroid cancer		22

Abbreviations. BTC: Biliary tract cancer; CRC: Colorectal cancer; GI: Gastrointestinal.

As shown in the table, the number of studies by cancer type ranged from 1 to 67. The five types of cancer with the highest number of publications were breast cancer (N=67), pancreatic cancer (N=59), colorectal cancer (N=56), prostate cancer (N=51), and lung cancer (N=36).

According to the study design, breast cancer accounted for the largest proportion of all types of cancer in case-control (19%) and cohort (33%) studies. Among the other study designs, colorectal cancer comprised the highest proportion in cross-sectional studies (43%), and pancreatic cancer in MRS (31%). A table with the number of studies by cancer type and study design is provided in Appendix A8.

Some studies considered multiple types of cancer in the same publication, providing evidence on the association between T2DM and **other 18 types of cancer** (including subsites). The following table presents the number of studies on these additional types of cancer, along with the reference studies that assessed them.

Table 3-3 Number of studies on additional cancer types (including subsites) identified beyond the original search strategy, with respective references.

No.	Type of cancer	Subsite	No. Studies	Reference
46	Bone cancer		4	(Bjornsdottir, Rawshani, Rawshani, Franzén, et al., 2020; Chodick et al., 2010; Gini et al., 2016; C.-C. Lin, Chiang, Li, Liu, et al., 2014)
47	Brain/NS		13	(Bao et al., 2018; Barami et al., 2017; Bjornsdottir, Rawshani, Rawshani, Franzén, et al., 2020; Chodick et al., 2010; Disney-Hogg et al., 2018; Gini et al., 2016; Harding et al., 2015; X. He, Shi, et al., 2018; C.-C. Lin, Chiang, Li, Liu, et al., 2014; X. Liu et al., 2015; S.-F. Lo et al., 2013; Saarela et al., 2019; Yuan et al., 2020)
48	CUP		1	(Hemminki et al., 2016)
49	Endocrine cancer		3	(Bao et al., 2018; X. Liu et al., 2015; T. et al., 2022)
50	Eye cancer		4	(Bjornsdottir, Rawshani, Rawshani, Franzén, et al., 2020; Gini et al., 2016; X. Liu et al., 2015; Saarela et al., 2019)
51	Genital cancer		5	(Chodick et al., 2010; Gini et al., 2016; Hjalgrim et al., 1997; Idilbi et al., 2013; X. Liu et al., 2015)
52		Testicular cancer	5	(Bjornsdottir, Rawshani, Rawshani, Franzén, et al., 2020; Harding et al., 2015; C.-C. Lin, Chiang, Li, Liu, et al., 2014; Saarela et al., 2019; Yuan et al., 2020)
53		Penis cancer	3	(Bjornsdottir, Rawshani, Rawshani, Franzén, et al., 2020; C.-C. Lin, Chiang, Li, Liu, et al., 2014; Reinholdt et al., 2021)

No.	Type of cancer	Subsite	No. Studies	Reference
(Continued from previous page)				
54	Head and Neck		6	(Chodick et al., 2010; X. He, Shi, et al., 2018; S.-F. Lo et al., 2013; Stott-Miller et al., 2012; J. J. Walker et al., 2013; Yuan et al., 2020)
55		Larynx cancer	6	(Bao et al., 2018; Bjornsdottir, Rawshani, Rawshani, Franzén, et al., 2020; Gini et al., 2016; S. K. Kim, Jang, Kim, Rhyu, Lee, et al., 2020; C.-C. Lin, Chiang, Li, Liu, et al., 2014; Saarela et al., 2019)
56		Pharynx cancer	5	(Bao et al., 2018; Bjornsdottir, Rawshani, Rawshani, Franzén, et al., 2020; S.-H. Hsieh et al., 2014; C.-C. Lin, Chiang, Li, Liu, et al., 2014; Saarela et al., 2019)
57		Nose cancer	1	(Bjornsdottir, Rawshani, Rawshani, Franzén, et al., 2020)
58	Mesothelioma		2	(Gini et al., 2016; Saarela et al., 2019)
59	Myeloma		9	(Bjornsdottir, Rawshani, Rawshani, Franzén, et al., 2020; Fortuny et al., 2005; Gini et al., 2016; Harding et al., 2015; Y. Hu, Wu, et al., 2021; X. Liu et al., 2015; S.-F. Lo et al., 2013; Saarela et al., 2019; Yuan et al., 2020)
60	Peritoneum cancer		1	(C.-C. Lin, Chiang, Li, Liu, et al., 2014)
61	Placenta cancer		1	(C.-C. Lin, Chiang, Li, Liu, et al., 2014)
62	Soft tissue		4	(Bjornsdottir, Rawshani, Rawshani, Franzén, et al., 2020; Gini et al., 2016; X. He, Shi, et al., 2018; C.-C. Lin, Chiang, Li, Liu, et al., 2014)
63		Kaposi's Sarcoma	2	(Chodick et al., 2010; Gini et al., 2016)

Abbreviations. Brain/NS: brain/Nervous system; CUP: Cancer of unknown primary

Most of the additional types of cancer identified are classified as rare, such as mesothelioma and Kaposi's sarcoma. These types were found because the studies evaluated multiple types of cancer in a single study, except for two, which focused specifically on one type of cancer. One of these studies examined the association between T2DM and head and neck cancer, while the other focused on brain/nervous system cancers. Both studies were identified through the screening of additional references (3.2.2.2 Searching other sources). The study by (Stott-Miller et al., 2012) was found in the reference list of (Starup-Linde et al., 2013) for head and neck cancer, and the study by (Disney-Hogg et al., 2018) was found in the reference list of (Y. Wang et al., 2019) for brain/nervous system cancer.

The other studies, which examined over 20 types of cancer, focused primarily on those specified in the search strategy, such as breast, endometrial, gastric, liver, ovarian, and pancreatic cancers (Bao et al., 2018; Bjornsdottir, Rawshani, Rawshani, Franzén, et al., 2020; Chodick et al., 2010; Gini et al., 2016; X. He, Shi, et al., 2018; C.-C. Lin, Chiang, Li, Liu, et al., 2014).

In summary, evidence of the association between T2DM and the risk of developing cancer was found for **63 types of cancer**, including subsites.

3.4.7 Risk of bias assessment

The majority of **cohort studies (80%)**, over half of **case-control studies (52%)**, and **cross-sectional studies (57%)** were classified as **low-risk of bias** according to the NOS ($NOS \geq 7$ points).

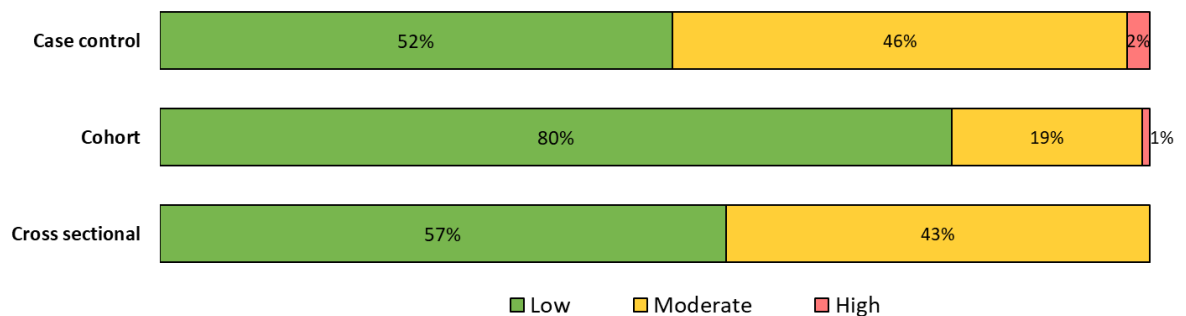


Figure 3-8 Risk of bias classification by study design. Eleven cohort studies in which the comparison group was the underlying population (with/without diabetes) were not scored. Low risk of bias ($NOS \geq 7$ points); moderate ($NOS 4-6$ points); high ($NOS < 4$ points).

The main reasons for a $NOS < 7$ in **case-control studies** were the selection of controls and the ascertainment of exposure. According to the criteria established for the NOS star allocation, controls derived from a hospitalized population are considered unacceptable, as they are not representative of the general population.

Additionally, no points were awarded to studies that based the diagnosis of T2DM on self-report, as this reduces data reliability and is prone to recall bias.

On the other hand, for **cohort studies**, one point or star is awarded for the adequacy of follow-up; that is, when complete follow-up of all or fewer than 20% of subjects is reported, or a description of those lost to follow-up is provided. However, this information was not included in the studies classified as moderate to high risk of bias.

Furthermore, in the comparability domain of the NOS, two stars were awarded if studies included age and sex (1 star) and obesity/BMI (1 star) as confounding variables in their analyses. The latter confounder was only considered in five out of the 26 cohort studies with a $NOS < 7$.

Finally, in moderate and high-risk **cross-sectional studies**, the response rate or characteristics of the included and non-included subjects were not described. Also, as in cohort studies, obesity/BMI was not considered in their analyses.

3.5 Discussion

Prevalence of T2DM

According to the latest report from the International Diabetes Federation (IDF), T2DM accounts for more than 90% of diabetes cases worldwide, although no specific global prevalence data are available (International Diabetes Federation, 2021). Notably, none of the studies included in SR were from African countries, likely due to data limitations (high-quality country data in 58% of the countries), as mentioned by the IDF. Despite this, Africa is projected to experience the largest increase in diabetes cases by 2045, with estimates suggesting 55 million people may be affected. However, while these data limitations are important, the absence of published studies on T2DM and cancer may also be due to other research questions taking precedence.

The IDF report highlights that the countries with the highest diabetes prevalence in adults are Pakistan (30.8%), French Polynesia (25.2%), and Kuwait (24.9%). However, only one study from Kuwait and one from Pakistan were identified from the SR. Similarly, while Mexico has the highest prevalence in the North American and Caribbean region (16.9%), only two studies from the Mexican population were found. In contrast, the United States of America (USA), with a lower diabetes prevalence rate than these countries, contributed the highest percentage of studies in the SR. Despite a diabetes prevalence of 11.3%, the USA leads the region with 32.2 million affected individuals living with the disease. In Europe, Turkey (14.5%) and Spain (10.3%) have the highest diabetes prevalence rates, but the SR only identified four studies from the Spanish population and none from Turkey. Conversely, a higher number of publications from countries with lower diabetes prevalence, such as Italy (6.4%) and the United Kingdom (6.3%), were identified, comprising 21 publications from Italy and 15 from the United Kingdom (UK).

Overall, although the prevalence of diabetes is high in some countries, there is little evidence related to its association with cancer risk. One possible reason for these findings could be that in these regions, there is more emphasis on other diseases or well-known complications of diabetes, such as macrovascular and microvascular diseases.

Number of studies by type of cancer

The number of studies related to specific cancer types associated with T2DM was generally consistent with the global burden of cancer. According to the Global Cancer Observatory, breast, prostate, lung, colorectal, and endometrial cancers have the highest incidence rates worldwide in populations over 20 years of age. In line with this, the largest proportion of studies included in the systematic review examined the relationship between T2DM and these cancer types. However, a considerable amount of evidence on the association with pancreatic cancer was also identified and selected. The interest in this type of cancer could be due to its mortality rate, as it is the 6th most common cause of cancer death in the world. In addition, half of the studies of this cancer were from the European population, where it is the sixth most commonly diagnosed cancer and the fourth leading cause of cancer death in this region (Ferlay et al., 2024).

Study design

The majority of the publications included in this SR utilized cohort study designs, which enable long-term data collection and evaluation of various outcomes, including different types of cancer, from a single population. This could be the reason why more studies with this design were identified in the SR, as researchers often repurposed data from existing programs or cohorts to explore a wide range of cancer associations.

In contrast, only 16 publications were MRS, with the initial study identified in 2015. The popularity of MRS has increased over the past two decades due to the greater availability of genome-wide association studies (GWAS), which provide the genetic data needed for these analyses (Sekula et al., 2016).

MRS is particularly helpful for examining causal relationships, as it reduces the risk of confounding variables and reverse causation by using genetic variants as

instrumental variables. This approach enables researchers to determine the causal effect of an exposure (such as T2DM) on an outcome (such as the risk of cancer) that traditional observational studies may not be able to establish.

Follow up

Ensuring that T2DM was diagnosed before the onset of cancer is important to support a causal interpretation. This is particularly relevant for cancers such as pancreatic cancer, which may induce diabetes (referred to as Type 3c diabetes), often within two years of tumor development (Mizuno et al., 2020). To reduce the risk of reverse causality, several authors recommend excluding cancer cases diagnosed within the first three years after diabetes onset (Bjornsdottir, Rawshani, Rawshani, Franzén, et al., 2020; Setiawan et al., 2019). This approach helps avoid misinterpreting pre-existing but undiagnosed cancers as causing T2DM.

In the systematic review conducted for this thesis, follow-up duration was incorporated into the quality assessment using the NOS. Over 90% of the included studies reported follow-up periods longer than three years, with the largest proportion falling between 11 and 15 years. This supports the conclusion that most studies provided sufficient follow-up to reduce the likelihood of reverse causality and to strengthen the temporal relationship between T2DM and cancer risk.

Exposure and outcome assessment method

One of the main reasons several case-control studies were classified as having moderate to high risk of bias was due to the method used to ascertain the diagnosis of T2DM. In the present SR, 68% of case-control studies relied on self-reported diabetes diagnoses. While self-reporting has been shown to have up to 98.5% agreement with medical records or physician validation in some studies (Jackson et al., 2014), other research indicates its accuracy may be as low as 64% (Kaye, 1991). These inconsistencies make self-reported diabetes less reliable and may introduce bias, mainly when no external validation is used (e.g. medical records, databases).

In contrast, to confirm T2DM diagnosis, cohort studies in this SR relied mainly on medical databases, such as those with ICD-coded information or the UK Clinical

Practice Research Datalink (CPRD). The use of these databases improves diagnostic accuracy, reduces misclassification risk, and provides a more valid measure of diabetes status in cohort studies.

For the diagnosis of cancer, most studies, both case-control and cohort, used as a reference pathology reports or cancer registries, because are highly reliable sources. Pathology reports are based on evidence from biopsies or surgeries, while cancer registries follow strict reporting guidelines, reducing the likelihood of diagnostic errors.

In few cohort studies, self-reported cancer diagnoses were confirmed using additional methods such as medical records or registries to verify the diagnosis.

Confounders

In the SR, the variables included in the adjusted analyses of cohort and case-control studies varied. Studies usually controlled for sociodemographic covariates, whereas behavioral characteristics were less commonly considered. The selection of variables for adjusted analyses was generally aligned with the data sources used, such as ICD-coded information or the UK CPRD, which primarily contain patient demographic data but often lack other confounders, such as lifestyle.

The choice of confounders should be consistent with the specific research objectives. When the goal is to assess causation, it is necessary to control for an appropriate set of confounders to avoid biased estimates. Evidence has shown that T2DM and multiple types of cancer share common risk factors, such as obesity (Garg et al., 2014). Nonetheless, only 64% of studies in the SR controlled for obesity. Similarly, smoking, a well-known risk factor for both T2DM and several cancers, particularly lung cancer, was considered in fewer than 50% of the studies (Maddatu et al., 2017; Malhotra et al., 2016).

Identifying potential confounding factors for each cancer type is important, even when the same data source is used. While T2DM may be associated with multiple cancer types, the confounders relevant to each cancer often differ. However, in several studies of the SR where the relationship between T2DM and various cancers was examined within the same analysis, an identical set of covariates was applied uniformly.

Therefore, to establish causal relationships in observational studies, confounders should be selected based on a thorough literature review, and statistical considerations. Failing to account for critical confounders can lead to an overestimation or underestimation of the true relationship between T2DM and cancer outcomes. To identify confounders, one recommended approach involves using causal diagrams, such as DAGs.

Strengths and limitations

This systematic review has several strengths. First, it provides a comprehensive overview of the association between T2DM and cancer risk across various cancer types using five databases combined with broad search algorithms. In order to cover the most commonly studied cancer types associated with T2DM, a scoping search was conducted. Subsequently, a search algorithm was developed based on these cancer types and applied across five databases, yielding over 18,086 publications.

Another strength is the lack of restrictions on language or publication date, which contributes to the comprehensiveness of the review. Previous SRs limited their scope to English-language publications and fewer databases (e.g., (Ling et al., 2020; Ohkuma et al., 2018)). The only published review meeting these criteria was conducted over a decade ago. Since new research on the topic has been published, this updated review is required to capture the latest evidence and address gaps in the literature.

Moreover, special attention was given to confounders in the adjusted analyses of the included studies. All reported variables from the adjusted analyses, along with their corresponding effect estimates, were extracted. This approach represents an initial step toward a more accurate method for establishing causality. In the subsequent chapter, confounders will be identified based on cancer type, and effect estimates from studies that account for these confounders will be incorporated into the meta-analyses. This approach is novel in the field, as previous research has attempted to assess only the strength of causal evidence through statistical methods (Ling et al., 2020).

However, it is important to acknowledge that this SR has certain limitations. Firstly, the search was restricted to published studies, potentially excluding relevant findings from unpublished or grey literature.

This SR is also limited by the classification of diabetes types, as some studies did not differentiate between type 1 and type 2 diabetes. To gather all relevant information on T2DM, studies were assumed to classify individuals diagnosed with diabetes after the age of 30 as having type 2 diabetes, while those diagnosed before age 30 were classified as having type 1 diabetes and excluded from the analysis (Rodbard et al., 2007; World Health Organization, 1994). Although this method helped to include studies without detailed diabetes categorization, it could have introduced potential misclassification bias.

A limitation of the included cohort studies is that individuals with type 2 diabetes were classified as cases, while the control group included people with type 1 diabetes, who also experience hyperglycaemia and insulin treatment. This may have diluted the observed effect sizes. Nevertheless, type 2 diabetes accounts for more than 90% of all cases of diabetes and is about ten times more common than type 1 diabetes. The target population in this systematic review was adults aged 18 years and older. Type 1 diabetes is the predominant form in children and adolescents, whereas type 2 diabetes is more frequently diagnosed in adults (International Diabetes Federation, 2019). For example, NHIS 2016 data in adults show that 5.8% of diagnosed cases were type 1 diabetes, 90.9% were type 2 diabetes, and 3.3% were other types (Bullard et al., 2018). Therefore, although some people in the control group may have had type 1 diabetes, their proportion was minimal and the effect sizes are unlikely to have been affected.

Finally, additional cancer types not initially identified through the scoping search were captured by reviewing the reference lists of systematic reviews and meta-analyses during the title and abstract screening stages. This approach provided new evidence on the association between T2DM and less common cancer types. However, due to time constraints, a systematic search specifically targeting these additional cancers was not conducted. Nonetheless, they were included in the systematic review and further analyses, as they may offer valuable insights for future research.

Conclusion

In conclusion, this systematic review provides evidence of the association between T2DM and cancer risk, emphasizing the need for further studies to address geographical variations and critical confounding factors.

The identification of these confounding factors is discussed in the next chapter.

The following table provides an overview of the thesis structure. The information related to the next chapter is highlighted in yellow.

Table 3-4 Overview of the thesis structure (Next: Chapter 4)

Chapter 1: Introduction. To outline the background of the thesis, introduce the research question and aim, and provide an overview of the thesis structure.

Chapter	Objective	Methods	Key insights
2. Context of Evidence on T2DM and Cancer Risk	To provide an overview of published systematic reviews (SRs) and meta-analyses examining the relationship between T2DM and cancer risk across multiple cancer types	Summarize SR and meta-analyses examining the relationship between T2DM and cancer risk across multiple cancer types	Context for the research conducted in this thesis
3. Systematic review	To identify all relevant evidence on the association between T2DM and cancer risk across various cancer types.	Comprehensive Systematic literature review PICOS framework: men and women aged ≥ 18 years (P); individuals with T2DM (I) compared to those without T2DM (C); cancer incidence (O); observational studies (S).	Types of cancer with evidence on the association between T2DM and cancer risk. Confounders included in the adjusted analyses of the studies. Other*
4. Causal Association	To identify the variables that must be controlled for to estimate the causal relationship between T2DM and cancer across multiple cancer types.	Directed Acyclic Graphs (DAGs) were developed to identify the Minimal Sufficient Adjustment Sets (MSAS)	MSAS for each type of cancer
5. Assessing the quality of covariable adjustment in studies making causal claims	To assess whether studies making causal claims about T2DM and cancer clearly state their intent; justify their choice of covariates; and select them appropriately based on their objectives.	Studies from SR: Analysis of the terminology used across different sections to identify causal claims. Justification of variable selection using clinical evidence or terminology. Comparison of MSAS from DAGs with covariates in adjusted analyses	Emphasize the need for a clear study goal, aligned methods, and consistent terminology
6. Meta-analyses	To establish the causal effect of T2DM on cancer risk	Meta-analyses that account for the MSAS for each cancer type. Subgroup analyses.	Types of cancer causally associated with T2DM
7. Burden of disease	To quantify the future global burden of cancer causally associated with T2DM	Global burden of disease analysis, using the types of cancer causally associated with T2DM (meta-analyses)	Burden of disease metrics: PIF/PFP, the number of cancer cases, DALYs, and monetised DALYs

Cancer incidence (O), expressed as risk ratio (RR), hazard ratio (HR), standardized incidence ratio (SIR), or odds ratio (OR). Other*: Study design, follow-up, country-geographical publications, assessment method of exposure and outcome, Risk of bias assessment. Abbreviations: DALYs: Disability-adjusted life years; PICOS: Population, Intervention, Comparison, Outcome, and Study Design; PIF: population impact fraction; PFP: prevented fraction of the population

Chapter 8: Discussion. To summarize the main findings, outline strengths and limitations, provide policy recommendations, suggest directions for future research, highlight contributions to knowledge, and present final conclusions.

Chapter 4 Causal association

4.1 Summary

- Estimating causal effects in observational studies requires addressing confounding bias. The use of directed acyclic graphs (DAGs) for each cancer type allowed for the identification of confounding variables that could potentially bias causal effect estimates.
- The minimum sufficient adjustment set (MSAS) of variables, derived from the hypothesized relationships between the variables in the DAGs, may have led to the omission of certain risk factors or associations. To improve the validity of the results, the causal assumptions in each DAG were informed by targeted literature searches.
- The confounders identified in the MSAS for each cancer type were mainly related to common lifestyle risk factors associated with both type 2 diabetes mellitus (T2DM) and various types of cancer. The omission of these confounders in the studies gathered from the systematic review (SR) was primarily due to unavailable data or the statistical methods employed.

4.2 Introduction

Objective

The aims of this chapter are to: (1) systematically identify risk factors for T2DM and each cancer type in order to develop DAGs, (2) define the MSAS for each cancer type based on these DAGs; and (3) compare the variables included in the identified MSAS with those used in the adjusted analyses of the studies from the SR (Chapter 3).

Background

The objective of many observational studies is to explore associations, but some also explicitly state an intention to describe either associations or causal relationships. In studies aiming to draw causal conclusions, adequate control for confounding is necessary to establish that the exposure has a causal effect on the outcome. Confounding bias occurs when another variable distorts the estimated relationship between the exposure and the outcome (Lederer et al., 2019). Chapter 3 of the SR found that a large proportion of studies did not adjust for modifiable and non-modifiable risk factors that could confound the relationship, resulting in potential bias and limiting their ability to draw valid causal conclusions. This is particularly important because T2DM and cancer share several risk factors.

A confounder is defined as a variable that is associated with both the exposure and the outcome but is not on the causal pathway between them (Cavanaugh, 2014). For simple causal structures, confounders can typically be identified based on this classic definition. However, in more complex associations, common in epidemiology due to the numerous risk factors and their interconnections, the use of other techniques, such as causal models, is critical (Howards et al., 2012). Causal models, represented graphically as DAGs, have recently become common methodologies in epidemiology. Their use is recommended for the identification of the MSAS of covariates that could distort the relationship (Tennant et al., 2021). For this research, the MSAS was identified to determine the variables that need to be controlled for in order to estimate the causal relationship between T2DM and cancer across multiple cancer types.

From the SR (outlined in Chapter 3), two studies used DAGs to inform the selection of confounding variables (Samson et al., 2016; Zabulienė et al., 2021). The first study analysed the association between diabetes and breast cancer incidence, and the second explored the impact of T2DM treatment on the risk of endometrial cancer. The authors of both publications reported that the DAGs were informed by prior knowledge of confounding factors. However, because both studies were retrospective and relied on routinely collected registry data, the DAGs were constrained by data availability. As a result, this could have led to an incomplete representation of the variables influencing the relationship

between the diseases, and some potential confounding variables may have been left unexplored.

In the other studies included in the SR in Chapter 3, the selection of variables for adjusted analyses depended on the available data and the statistical approach used. For instance, some studies included only variables with a p-value < 0.05 in the univariate analysis in their multivariable regression model (S. Agrawal et al., 2014; Alsheridah & Akhtar, 2018; Feola et al., 2022).

Key notation and terminology of DAGs

DAGs are diagrams used to represent possible causal relationships between an exposure of interest, the outcome, and related variables (Sauer & VanderWeele, 2013). DAGs are called "**directed**" because each arrow in the graph has a direction, meaning the arrows are one-way and not bidirectional. They are called "**acyclic**" because the graph does not contain any cycles; in other words, starting from any variable, it is impossible to return to the same node by following the arrows in the direction of the arrows (Hernán MA, 2020).

In the terminology of DAGs, variables are referred to as **nodes**, and the arrows between them are called **edges**. Nodes located upstream from a particular variable are known as **ancestors**, while those located downstream are referred to as **descendants** (Digitale et al., 2022; Lederer et al., 2019).

A **path** in a DAG is a sequence of arrows connecting two variables, regardless of the direction. These paths can be either **causal** or **non-causal**. A path is considered **causal** if all the arrows along the path point in the same direction. In contrast, a path is **non-causal** if the arrows do not all point in the same direction (Digitale et al., 2022). Paths in a DAG can be classified as either **open** or **blocked**. An **open path** represents a direct flow of influence between two variables, while a **blocked path** implies no influence is transmitted between the variables on that path (Tennant et al., 2021). A given exposure and outcome may be connected by multiple paths. A **mediator** is a variable on an open causal path between the exposure and outcome, while a **confounder** is a common cause of both. A **collider** is a variable with arrows pointing to it from two other variables.

A path containing a collider is automatically **closed**, blocking the flow of association between variables. Therefore, adjusting for a collider opens the path and introduces bias (Digitale et al., 2022).

A **backdoor path** is a non-causal path that connects the exposure to the outcome, where the path starts with an arrowhead pointing toward the exposure and ends with an arrowhead pointing toward the outcome. This path remains even if all arrows pointing from the exposure to other variables are eliminated.

To estimate an unbiased causal effect of the exposure on the outcome, all open backdoor paths must be identified and blocked, which is known as **backdoor criterion**. This can be done by conditioning on a minimal sufficient adjustment set of confounders (**MSAS**) (Hernán MA, 2020).

The next figure shows the key elements of a DAG, including examples of a causal path, mediator, confounder, and collider.

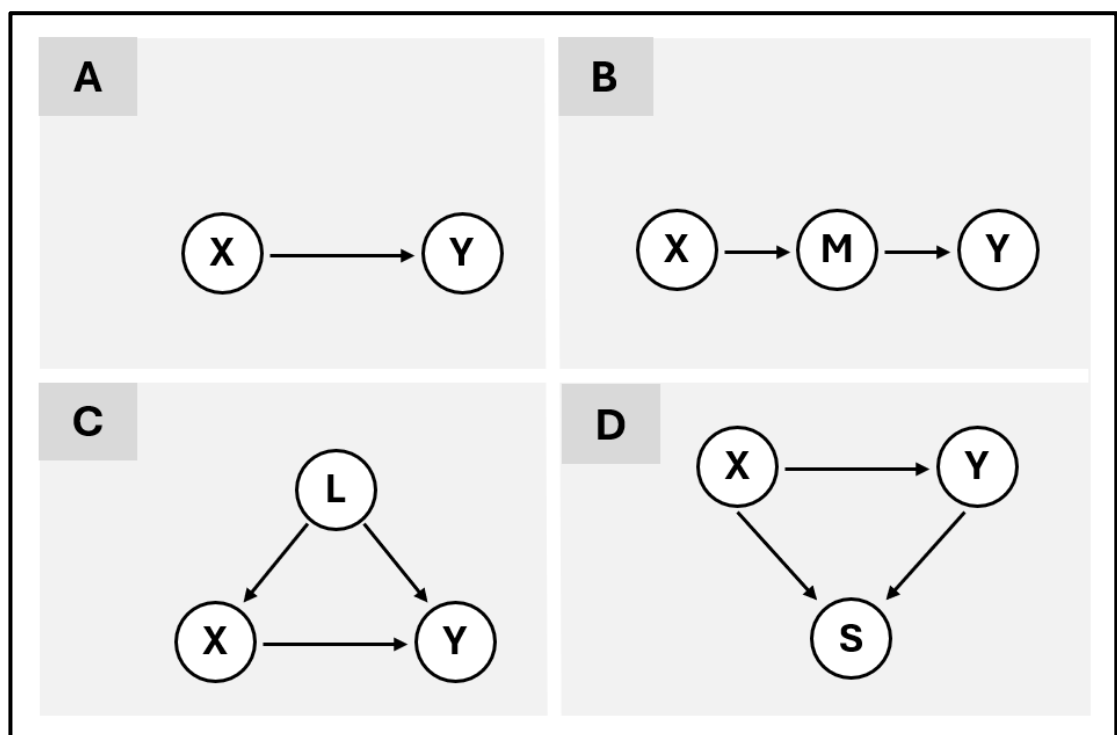


Figure 4-1 Elements of a DAG. A) A *causal path* between the exposure (X) and the outcome (Y). B) An example of a *mediator* (M), which is an intermediate variable on the causal pathway between X and Y. X influences Y through its effect on M. C) An example of a *confounder* (L), which is a common cause of X and Y. L confounds the association between X and Y and must be conditioned on to address confounding bias. D) An example of a *collider* (S), which is a common effect of X and Y. The non-causal path between X and Y is blocked by S.

4.3 Methods

To identify the MSAS of confounding variables for estimating the causal relationship between T2DM and cancer risk, DAGs were constructed. The development of each DAG followed a structured, systematic process that involved: (1) defining the exposure and outcome, (2) identifying risk factors for T2DM and each cancer type through targeted searches and expert sources, and (3) evaluating potential relationships among these variables.

A separate DAG was developed for each cancer type to identify the MSAS required to control for confounding and estimate the causal effect of T2DM. The following sections describe the steps involved in constructing the DAGs and present an example to illustrate how the MSAS was derived.

4.3.1 Directed Acyclic Graphs (DAGs) development

The DAGs for each cancer type, including the related risk factors and references, can be found in Appendix B. The DAGs were created using a systematic process described below.

Step 1 Defining the Exposure and Outcome

The first step in building DAGs is identifying the exposure and outcome of interest (Digitale et al., 2022; Poppe et al., 2024). For this research, T2DM was considered the exposure, and the risk of each cancer type was the outcome. A directed edge was drawn from the exposure to the outcome.

Step 2 Identifying risk factors

The second step involves identifying the risk factors of the exposure (T2DM) and the outcome (the risk of each type of cancer). It is recommended to include common causes of exposure and the outcome, and variables that influence only one should be included or omitted. The variables to be depicted in the DAGs should be determined based on domain knowledge from the literature and consultation with experts in the field (Digitale et al., 2022; Poppe et al., 2024; Sauer & VanderWeele, 2013).

For this step, targeted searches were conducted to identify the risk factors for T2DM and each cancer type. This information was gathered from well-known institutions, medical societies, and public health organizations, including the American Society of Clinical Oncology (ASCO), Cancer Research UK, the Centers for Disease Control and Prevention (CDC), the Canadian Cancer Society, Cancer Australia, Johns Hopkins Medicine, and the National Institute for Health and Care Excellence (NICE) (American Society of Clinical Oncology's (ASCO), 2023; Canadian Cancer Society, 2023; Cancer Australia, 2023; Cancer Research UK, 2023; Centers for Disease Control and Prevention (CDC), 2023; Johns Hopkins Medicine, 2023; National Institute for Health and Care Excellence (NICE), 2023).

There are no explicit guidelines on selecting a variable to include in the DAG other than that it should be based on domain knowledge and existing literature (Poppe et al., 2024). Therefore, for this thesis to confirm that a variable was a risk factor for the exposure or the outcome, in addition to consulting a reliable source, the variable had to be mentioned in at least two independent sources to be included in the DAG. After identifying and selecting the risk factors for T2DM and each cancer type, a directed edge was drawn from each risk factor to both the exposure and the outcome.

Additionally, it is recommended that variables that are common causes of two risk factors in the DAG should be added (Gaskell & Sleight, 2020). For this thesis, socioeconomic status (SES) was included due to its association with several risk factors, such as drinking (alcohol consumption), diet, smoking, and a sedentary lifestyle, all of which are linked to both T2DM and certain types of cancer. Research suggests that individuals with a lower SES are less likely to have healthy dietary habits, have higher rates of smoking, and are more prone to having a sedentary lifestyle compared to those with a higher SES. Conversely, individuals with a higher SES tend to consume more alcohol than those with a lower SES (Alkerwi et al., 2015; Beard et al., 2019; Hitchman et al., 2014).

Step 3 **Considering Directed Edges Between Risk Factors**

The third step involves examining potential directed edges between pairs of risk factors. It is important to justify the omission of any edges, as this implies that the variables are not related, which is a strong assumption (Sauer &

VanderWeele, 2013). Table B1-7 in the appendix B presents the omitted edges (associations) between risk factors along with their corresponding references.

To assess whether there was evidence supporting an association between any two risk factors, the same sources used to identify the risk factors for the exposure and outcome were consulted. If no information was found in these sources, a targeted search was conducted using databases, primarily PubMed. Relevant terms related to the risk factors and their potential associations were included in the search, with a focus on titles and abstracts. Observational studies, systematic reviews, and meta-analyses were considered, while grey literature (e.g., abstracts-only papers, posters, editorials), case reports, in vitro studies, and animal-based research were excluded.

4.3.2 Minimum sufficient adjustment set (MSAS) and software

The diversity of T2DM risk factors, as well as those related to each type of cancer, resulted in a complex DAG with multiple pathways. For that reason, DAGs were created with DAGitty (version 3.0), a free software available at dagitty.net (Textor et al., 2016). The MSAS was identified with this tool's automated algorithm. This software displays the different options if more than one MSAS is possible.

4.3.3 Cancer Types for Which a DAG Was Created

In the SR outlined in Chapter 3, 63 types of cancer (including subsites) were identified in relation to T2DM. However, a **DAG was created for only 56 types**, with 7 types excluded for the following reasons:

- **Chronic Lymphocytic Leukemia, Placenta Cancer, and Peritoneum Cancer:** Each of these cancers had only one study examining the association between T2DM and the respective cancer type. These studies were classified as high risk of bias according to the Newcastle-Ottawa Scale (NOS), specifically, (Benavente et al., 2018) for Chronic Lymphocytic Leukemia, and (C.-C. Lin, Chiang, Li, Liu, et al., 2014) for Placenta Cancer and Peritoneum Cancer. Only studies with a low risk of

bias were selected for the meta-analyses (Chapter 6) to ensure the most accurate effect size estimates (De Cassai et al., 2023).

- **Acute Myeloid Leukemia and Chronic Myeloid Leukemia:** A DAG for these cancer types was not created because the only study identified in the SR that evaluated the association between T2DM and these cancers used the general population, including both individuals with and without diabetes, as the comparison group, rather than only using individuals without T2DM or presenting data separately (i.e., for the general population with and without diabetes) (Harding et al., 2015). As this thesis aims to estimate the independent effect of T2DM on cancer risk, such a comparison was not appropriate.
- **Genital Cancer:** A DAG was not created to represent the association between T2DM and genital cancer because it is a broad term encompassing various cancers affecting the genital organs, such as Ovarian Cancer and Prostate Cancer, which were considered for individual DAGs (Centers for Disease Control and Prevention (CDC), 2023). Additionally, genital cancers have distinct risk factors, making it more appropriate to create separate DAGs for each type to clearly illustrate the underlying relationships between their specific risk factors. For example, in addition to sex-related risk factors such as nulliparity for Ovarian Cancer, other factors, like smoking, are recognized as risk factors for Ovarian Cancer but not for Prostate Cancer (American Cancer Society, 2021; Cancer Research UK, 2023; Centers for Disease Control and Prevention (CDC), 2023).
- **Endocrine Cancer:** Similar to genital cancer, Endocrine Cancer is a broad category that includes various cancer types that originate in glands or in cells that produce hormones (National Cancer Institute, 2023). Thyroid Cancer, for example, has already been considered individually for a DAG.

4.3.4 Example of a DAG

The following DAG illustrates the relationship between **T2DM** as the exposure and the **risk of pancreatic cancer** as the outcome.

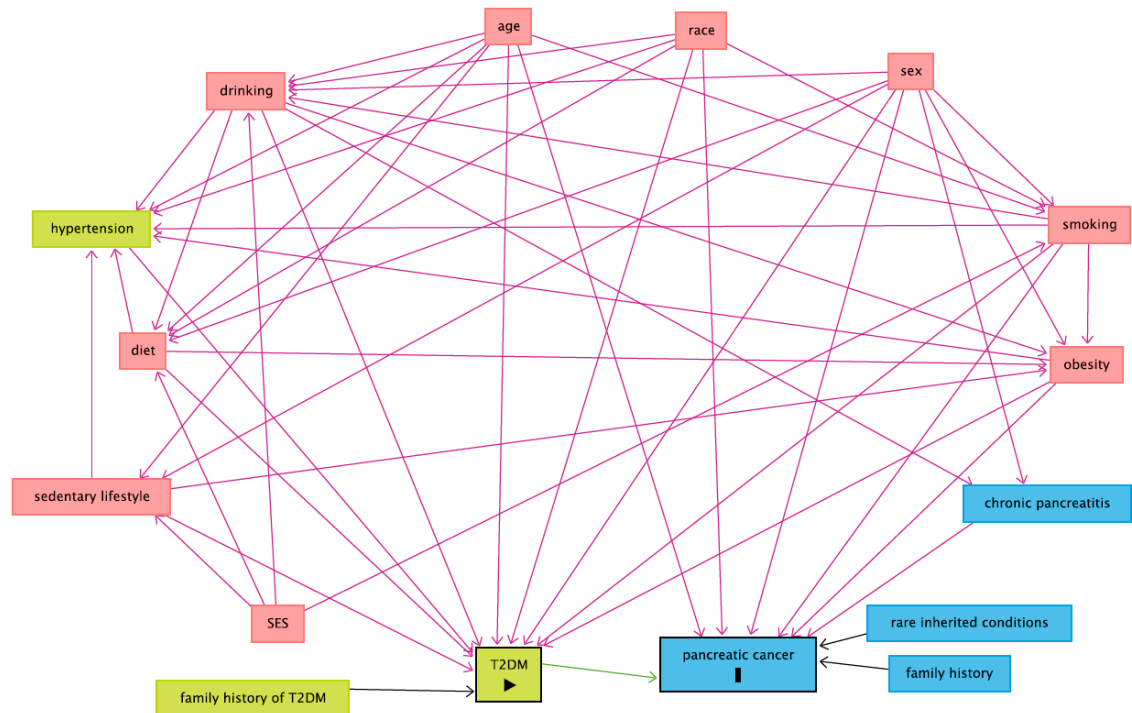


Figure 4-2 DAG of the association between T2DM and risk of pancreatic cancer, including potential confounders. Backdoor or biasing paths are marked in red.

As outlined in section 4.3.1, the DAG incorporates risk factors for both T2DM and pancreatic cancer (American Cancer Society, 2021; Canadian Cancer Society, 2023; Cancer Australia, 2023; Centers for Disease Control and Prevention (CDC), 2023; Chi et al., 2022; Han, 2020; Kao, 2001; M.-J. Kim et al., 2015; Larsson & Burgess, 2022; National Institute for Health and Care Excellence (NICE), 2023; National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), 2022; Pal et al., 2021; Sowers et al., 2001; The British Diabetic Association, 2023; Xue et al., 2022).

- **Risk factors for T2DM** include age, sex, race, sedentary lifestyle, smoking, drinking (alcohol intake), obesity, family history, hypertension, and diet.
- **Pancreatic cancer risk factors** include age, sex, race, smoking, obesity, chronic pancreatitis, family history, and rare inherited conditions, such as hereditary pancreatitis, Peutz-Jeghers syndrome, familial malignant

melanoma and pancreatic cancer, hereditary breast and ovarian cancer syndrome, Lynch syndrome, Li-Fraumeni syndrome, and familial adenomatous polyposis.

In the DAG, **T2DM** (the exposure) is represented by a **triangle in a green box**, while the outcome (**pancreatic cancer risk**) is shown in **blue with an "I."** The ancestors of the exposure (**in green**) are hypertension and family history of T2DM. The ancestors of the outcome (**in blue**) include rare inherited conditions, chronic pancreatitis, and family history. Common ancestors for both the exposure and outcome (**in pink**) include sedentary lifestyle, diet, socioeconomic status (SES), age, race, sex, smoking, and obesity.

Chronic pancreatitis is an example of a descendant of drinking, as represented by the path: **Drinking → Chronic pancreatitis.**

The graph also shows backdoor or biasing paths, marked in red. Examples of such paths include:

- **T2DM ← Age → Pancreatic cancer**
- **T2DM ← Drinking → Obesity → Pancreatic cancer**
- **T2DM ← Sedentary lifestyle ← Age → Smoking → Obesity → Pancreatic cancer**

To minimize bias in estimating the causal effect of T2DM on pancreatic cancer risk, all backdoor paths must be blocked by identifying the MSAS of covariates that could confound the relationship. There are two possible MSAS options:

- **Option 1:** Age, chronic pancreatitis, sex, obesity, race, smoking
- **Option 2:** Age, drinking, sex, obesity, race, smoking

Age, sex, race, smoking, and obesity are common confounders between T2DM and pancreatic cancer, so these covariates must be conditioned on. Additionally, drinking (alcohol consumption) or chronic pancreatitis must be conditioned on to block all remaining biased paths. Drinking is a risk factor for T2DM and is linked

to chronic pancreatitis, which is a risk factor for pancreatic cancer. Conditioning on either drinking or chronic pancreatitis is sufficient to close the remaining backdoor paths, as both appear in the same open paths.

The choice between the two MSAS options depends on the available data. In this case, chronic pancreatitis was not included as a covariate in studies on this type of cancer, while drinking (alcohol consumption) was. Therefore, **Option 2** was selected to assess the causal effect of T2DM on pancreatic cancer risk.

The following graph illustrates the open backdoor paths when neither drinking (alcohol consumption) nor chronic pancreatitis is conditioned on, even after adjusting for age, race, sex, smoking, and obesity (in grey). As shown in the graph, these open paths (in red) indicate potential bias.

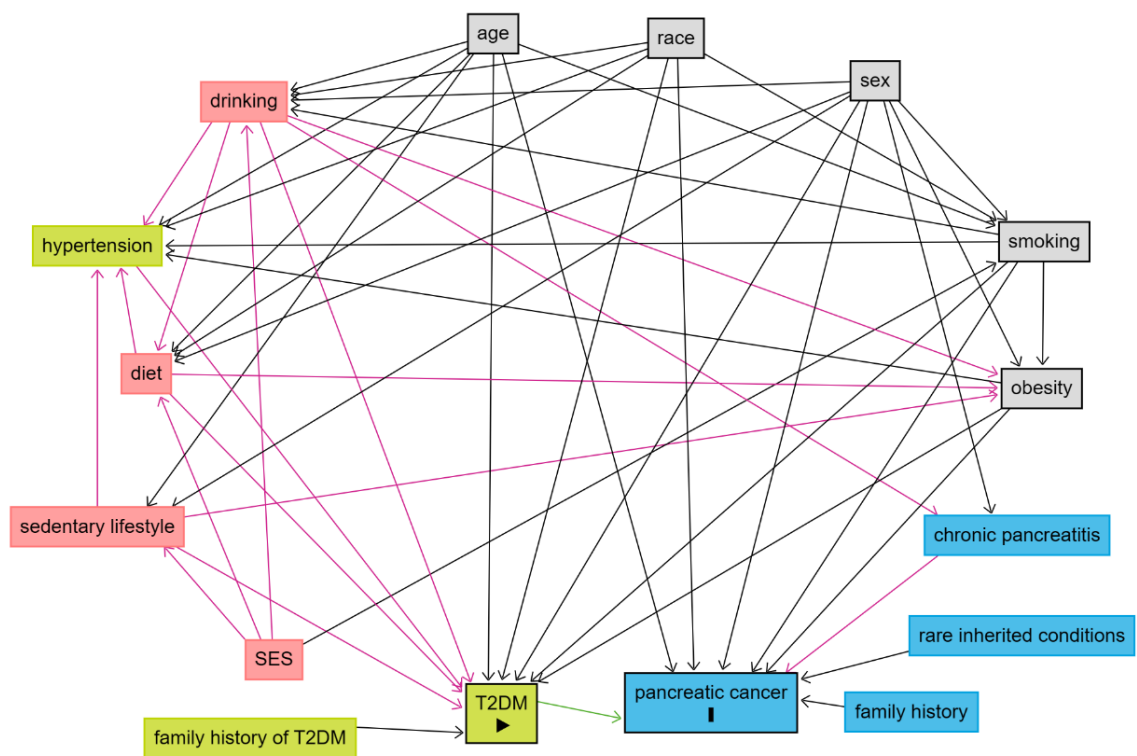


Figure 4-3 DAG of the association between T2DM and risk of pancreatic cancer, including potential confounders. Backdoor or biasing paths are marked in red. Adjusted covariates are shown in grey (age, race, sex, smoking, obesity).

4.4 Results

This section presents the results related to the MSAS identified for 56 cancer types using DAGs. It includes the MSAS for each type, the proportion of effect estimates that met MSAS criteria, cancer types with no MSAS-adjusted estimates, and the most commonly omitted confounders in non-MSAS estimates.

To formally examine the causal association between T2DM and cancer incidence, only the 149 cohort and case-control studies rated as low risk of bias in the SR were included in the analyses presented in this chapter and those that follow. These studies were selected to minimize bias in effect size estimates, as the validity and reliability of the meta-analyses in Chapter 6 depend on the quality of the included evidence (De Cassai et al., 2023).

4.4.1 MSAS for the different types of cancer

Table 4-1 presents the MSAS derived from the use of DAGs for 56 different types of cancer. Multiple MSAS options were identified for several cancer types. In these cases, conditioning on certain covariates was sufficient to block all backdoor paths, as specific variables appeared in the same open paths. This means that adjusting for either covariate effectively controlled for confounding. Among the cancer types with multiple MSAS options, the option selected for this thesis is marked with an "X." Further details are provided in 4.4.1.1 Selection of MSAS Options.

In addition, in **¡Error! No se encuentra el origen de la referencia.**, the confounders included in the MSAS that are not common risk factors between T2DM and each cancer type are **highlighted in blue**. This analysis illustrates the usefulness of DAGs in understanding the causal structure between risk factors in complex associations. In such cases, relying solely on the traditional definition of a confounder, as a variable associated with both the exposure and the outcome, but not on the causal pathway, may be insufficient (Cavanaugh, 2014). The relationship between T2DM and cancer involves numerous interconnected risk factors, making it difficult to identify the minimum set of confounders using conventional definitions alone.

Table 4-1 Minimum sufficient adjustment sets (MSAS) identified through DAGs for different cancer types.

Type of cancer	MSAS
Anal	• HIV, age, drinking, race, sex, smoking - X • SES, age, drinking, race, sex, smoking • age, diet, drinking, race, sedentary lifestyle, sex, smoking
Bladder	- X • SES, age, race, sex, smoking • age, chemical exposure, race, sex, smoking • age, diet, drinking, race, sedentary lifestyle, sex, smoking
Bone	age
Brain/NS	age, obesity
Breast	- X • SES, age, drinking, obesity, race, sedentary lifestyle • age, diet, drinking, obesity, race, sedentary lifestyle, smoking • age, drinking, first pregnancy at later age, obesity, race, sedentary lifestyle
BTC	• PSC, age, bile duct stones, liver flukes • PSC, age, bile duct stones, race • age, bile duct stones, liver flukes, sex • age, bile duct stones, race, sex - X • age, obesity, race, sex
Cervix uteri	• HIV, HPV, SES, age, other STI, smoking • HIV, HPV, age, multiparity, other STI, smoking • HIV, age, multiparity, other STI, race, smoking - X • SES, age, drinking, race, smoking • SES, age, other STI, race, smoking • age, diet, drinking, race, sedentary lifestyle, smoking
Colon	age, diet, drinking, obesity, race, sedentary lifestyle, sex, smoking
CRC	age, diet, drinking, obesity, race, sedentary lifestyle, sex, smoking
CUP	age, obesity, smoking
ECC	- X • age, drinking, obesity, race, sex, smoking • cholecystolithiasis, choledocholithiasis, cholelithiasis, drinking, hepatitis B/C, smoking • cholecystolithiasis, choledocholithiasis, cholelithiasis, drinking, sex, smoking
Endometrial	PCOS, age, obesity, race, smoking
EJA	• Helicobacter pylori infection, achalasia, age, diet, drinking, obesity, race, sex, smoking - X • SES, age, diet, drinking, obesity, race, sex, smoking • age, diet, drinking, obesity, race, sedentary lifestyle, sex, smoking
Esophageal	- X • SES, age, diet, drinking, obesity, race, sex, smoking • achalasia, age, diet, drinking, obesity, race, sex, smoking • age, diet, drinking, obesity, race, sedentary lifestyle, sex, smoking
Eye	age, race, sex
Gallbladder	age, drinking, obesity, race, sex, smoking
GI cancer	age, diet, drinking, obesity, race, sedentary lifestyle, sex, smoking
HCC	age, drinking, hypertension, obesity, race, sex, smoking
H&N	- X • SES, age, drinking, race, sex, smoking • age, diet, drinking, race, sedentary lifestyle, sex, smoking • age, drinking, environmental or occupational inhalants, race, sex, smoking
Hodgkin lymphoma	• HIV infection, age, sex - X • SES, age, race, sex, smoking • age, diet, drinking, race, sedentary lifestyle, sex, smoking
ICC	- X • age, drinking, obesity, race, sex, smoking • choledocholithiasis, cholelithiasis, drinking, hepatitis B/C, smoking • choledocholithiasis, cholelithiasis, drinking, sex, smoking
Kaposi's sarcoma	age, drinking, race, sex, smoking

Type of cancer	MSAS
(Continued from previous page)	
Kidney	age, hypertension, obesity, race, sex, smoking
Large intestine	age, diet, drinking, obesity, race, sedentary lifestyle, sex, smoking
Laryngeal	X • SES, age, drinking, race, sex, smoking • age, diet, drinking, race, sedentary lifestyle, sex, smoking • age, drinking, exposure to certain substances, race, sex, smoking
Leukemia	X • SES, age, race, sex, smoking • age, chemical exposure, race, sex, smoking • age, diet, drinking, race, sedentary lifestyle, sex, smoking
Lip	age, drinking, race, sex, smoking
Liver	age, drinking, hypertension, obesity, race, sex, smoking
Lung	X • SES, age, race, sex, smoking • age, diet, drinking, race, sedentary lifestyle, sex, smoking • age, exposure to substances, race, smoking
Lymphocytic leukemia	X • SES, age, race, sex • age, chemical exposure, race, sex • age, diet, drinking, race, sedentary lifestyle, sex, smoking
Lymphoma	X • SES, age, obesity, race, smoking • age, alcohol, diet, obesity, race, sedentary lifestyle, sex, smoking • age, chemical exposure, obesity, smoking
Melanoma	age, drinking, race, sex, smoking
Mesothelioma	X • SES, age, sex • age, asbestos exposure, sex • age, diet, drinking, race, sedentary lifestyle, sex, smoking
Mouth	age, drinking, race, sex, smoking
Myeloid leukemia	X • SES, age, sex, smoking • age, diet, drinking, race, sedentary lifestyle, sex, smoking • age, exposure to chemicals, sex, smoking
Myeloma	• HIV, SES, age, drinking, obesity, sex, smoking • HIV, age, chemical exposure, drinking, obesity, sex, smoking X • SES, age, drinking, obesity, race, sex, smoking • age, diet, drinking, hypertension, obesity, race, sedentary lifestyle, sex, smoking
NHL	age, drinking, sex, smoking
Non-melanoma	X • SES, age, drinking, race, sex, smoking • age, diet, drinking, race, sedentary lifestyle, sex, smoking • age, drinking, exposure to chemicals, race, sex, smoking
Nose	X • SES, age, race, sex, smoking • age, diet, drinking, race, sedentary lifestyle, sex, smoking • age, race, sex, smoking, workplace exposures
Oral	age, drinking, race, sex, smoking
Ovarian	X • SES, age, obesity, race, smoking • age, asbestos exposure, obesity, race, smoking • age, diet, drinking, obesity, race, sedentary lifestyle, smoking
Pancreatic	• Age, chronic pancreatitis, sex, obesity, race, smoking X • Age, drinking, sex, obesity, race, smoking
Penile	age, drinking, race, smoking
Pharyngeal	age, drinking, race, sex, smoking
Prostate	age, obesity, race
Rectal	age, diet, drinking, obesity, race, sedentary lifestyle, sex, smoking
Rectosigmoid	age, diet, drinking, obesity, race, sedentary lifestyle, sex, smoking
Salivary gland	age
Skin	X • SES, age, drinking, race, sex, smoking • age, diet, drinking, race, sedentary lifestyle, sex, smoking • age, drinking, exposure to chemicals, race, sex, smoking
Small intestine	X • Crohn's disease, age, diet, race, sex • age, diet, race, sex, smoking

Type of cancer	MSAS
(Continued from previous page)	
Soft tissue	• damaged lymph system, previous radiotherapy treatment - X • obesity
Stomach	• H. pylori infection, age, diet, drinking, obesity, race, sex, smoking - X • SES, age, diet, drinking, obesity, race, sex, smoking • age, diet, drinking, obesity, race, sedentary lifestyle, sex, smoking
Testicular	Age, race
Thyroid	age, obesity, race, sex
Tongue	• HPV, drinking, sex, smoking - X • age, drinking, race, sex, smoking
UTC	- X • SES, age, hypertension, obesity, race, sex, smoking • age, chemicals, hypertension, obesity, race, sex, smoking • age, diet, drinking, hypertension, obesity, race, sedentary lifestyle, sex, smoking

Confounding variables in the MSAS that are not shared risk factors between T2DM and the corresponding cancer type are [highlighted in blue](#). For cancer types with multiple MSAS options, the one used in this thesis for the analysis is marked with an “X.”

Drinking refers to alcohol consumption, alcohol drinks per day. Sedentary lifestyle refers to physical activity, regular exercise, average hours per week of moderate or vigorous activity.

Abbreviations. Brain/NS: Brain/Nervous System; BTC: Biliary tract cancer; CRC: Colorectal cancer; CUP: Carcinoma of unknown primary; GI: gastrointestinal; ECC Extrahepatic cholangiocarcinoma; EJA: Esophagogastric junction adenocarcinoma; H&N: Head and Neck cancer; ICC: Intrahepatic cholangiocarcinoma; MSAS: minimal sufficient adjustment sets; NHL: Non-Hodgkin lymphoma; PCOS: Polycystic ovary syndrome; SES: socioeconomic status; UTC: Urinary tract cancer.

For most cancer types, the different MSAS options included both modifiable and non-modifiable confounders. However, for bone, eye, salivary gland, and testicular cancers, only non-modifiable confounders were identified. This is because, for bone and salivary gland cancers, only age was related to other T2DM risk factors, while the remaining factors were neither modifiable nor associated with diabetes risk factors. The same applies to eye and testicular cancers, although in these cases, both age and race were related to T2DM risk factors, and the other variables were non-modifiable and unrelated to diabetes risk factors.

The analysis based on confounders included in the MSAS that are not common risk factors between T2DM and each cancer type was informative for most cancers (36 out of 56). Exceptions included bone, brain, colon, colorectal, cancer of unknown primary, eye, gallbladder, gastrointestinal, kidney, large intestine, lip, mouth, oral, pharyngeal, prostate, rectal, rectosigmoid, salivary gland, testicular, and thyroid cancers. For these, no additional confounders beyond the common T2DM and cancer risk factors were identified using DAGs.

For the types of cancer where the analysis was informative, some confounders that are not common risk factors between T2DM and each cancer type included drinking (alcohol consumption), obesity, smoking, and hypertension.

- **Drinking (alcohol consumption):** Anal, Kaposi's sarcoma, melanoma, myeloma, NHL, non-melanoma, pancreatic, penile, skin.
- **Obesity:** BTC, ECC, ICC, soft tissue.
- **Smoking:** Endometrial, Hodgkin lymphoma, lymphoma, melanoma, myeloma, NHL, non-melanoma, skin, small intestine.
- **Hypertension:** HCC, liver.

The findings emphasize the value of using DAGs to identify confounders beyond common risk factors, particularly those related to drinking (alcohol consumption) and smoking.

4.4.1.1 Selection of MSAS Options

For several cancer types, multiple MSAS options were identified. However, upon reviewing the variables included in the adjusted analyses, only one option aligned with the data available in the studies.

For example, three MSAS options were identified for ovarian cancer:

1. SES, age, obesity, race, smoking
2. Age, asbestos exposure, obesity, race, smoking
3. Age, diet, drinking (alcohol consumption), obesity, race, sedentary lifestyle, smoking

The first option was the only feasible one, as asbestos exposure was not reported in any of the included studies for this cancer type. Additionally, variables such as diet, drinking (alcohol consumption), and sedentary lifestyle were rarely included in the adjusted analyses across all cancer types. As a result, only the first option met the MSAS criteria based on the variables available in the studies.

Similarly, for lung cancer, the MSAS options were:

1. SES, age, race, sex, smoking
2. Age, diet, drinking, race, sedentary lifestyle, sex, smoking
3. Age, exposure to substances, race, smoking

Again, the first option was the only viable. Variables such as exposure to substances like asbestos, arsenic, diesel exhaust, and some forms of silica and chromium were not available in the included studies for this cancer type. Diet was also commonly omitted. Therefore, only the first MSAS option was met in the adjusted analyses for lung cancer.

Socioeconomic status (SES), whether reported as a SES or through specific indicators such as income or education, was frequently included in the adjusted analyses. For cancer types with multiple MSAS options, the option containing SES indicators often fulfilled the MSAS criteria and was therefore selected for analysis.

Therefore, the selection among different MSAS options depended on the availability of variables in the studies. In Table 4-1, the MSAS option selected for each cancer type with multiple alternatives is marked with an “X.”

Based on the selected MSAS options, whether from cancer types with multiple alternatives or those with only one option derived from DAGs, analyses were conducted to assess: (1) the proportion of effect estimates that met the MSAS criteria for each cancer type, (2) cancer types with no effect estimates meeting MSAS criteria and (3) the proportion of confounders omitted in estimates that did not meet the MSAS criteria.

4.4.2 Proportion of effect estimates with the MSAS for each type of cancer

Of the 56 cancer types, at least one effect estimate that accounted for the MSAS was identified for **31 types** (45%). The proportion of reported effect estimates that met the MSAS criteria, based on whether the variables in the adjusted analyses matched the MSAS identified through DAGs for each of these 31 cancer types is shown in Figure 4-4.

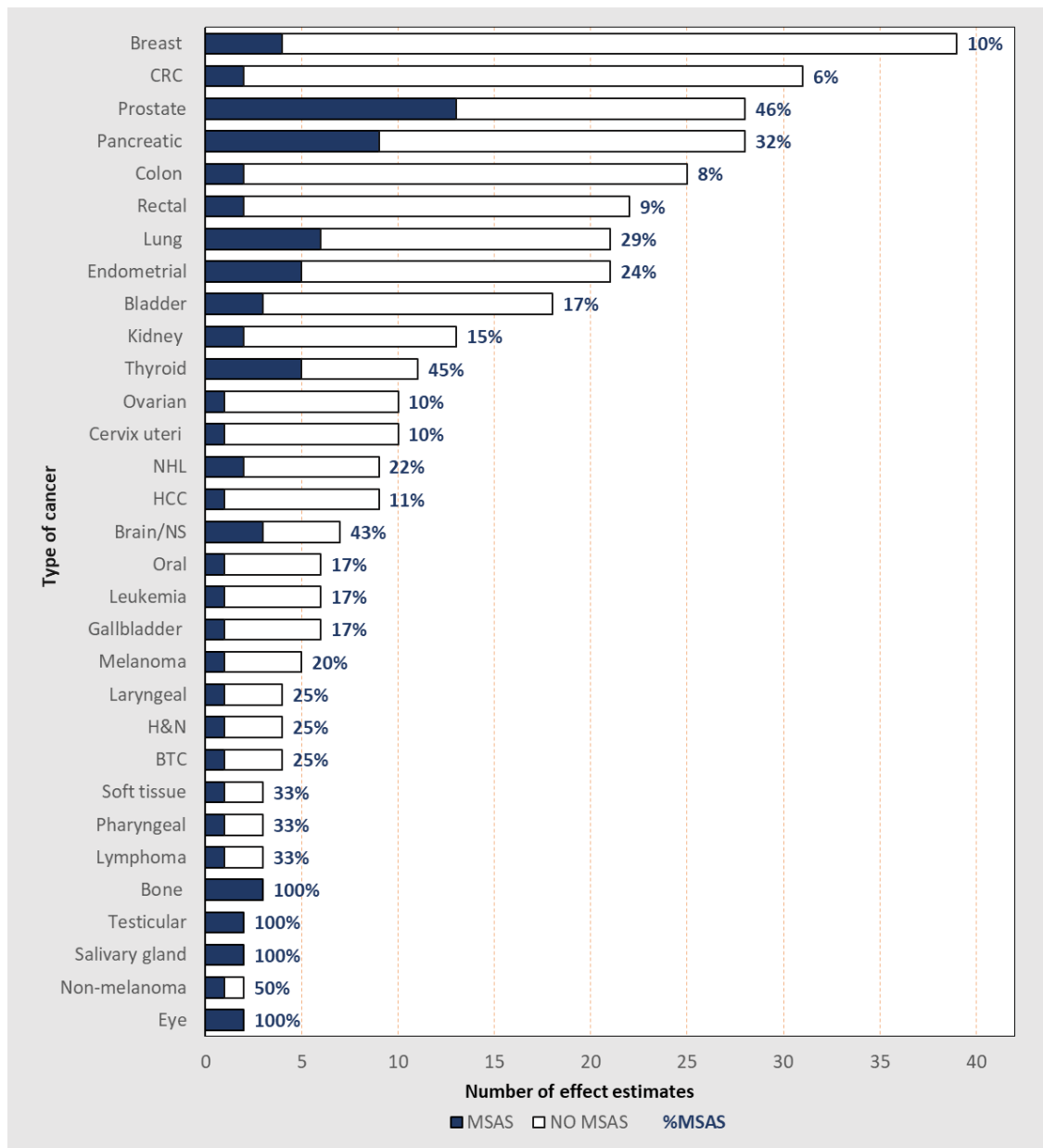


Figure 4-4 Proportion of Effect Estimates that included all variables in the MSAS. The total number of effect estimates exceeds the number of studies because some publications reported estimates for multiple cancer types. The 25 cancer types with no effect estimates meeting the MSAS criteria are not shown in the graph. MSAS: Number of effect estimates that included all variables in the minimum sufficient adjustment set. No MSAS: Number of effect estimates that did not include all MSAS variables. %MSAS: Proportion of effect estimates that met the MSAS criteria.

The two effect estimates reported for eye, salivary gland, and testicular cancer, as well as the three for bone cancer, fulfilled all the MSAS (age, sex, race).

The type of cancer with the largest number of reported effect estimates was breast cancer (N=39), but only four (10%) were adjusted with the MSAS. CRC was the second type of cancer with the highest number, but only two effect estimates (6%) were adjusted with the MSAS.

Of the 28 effect estimates related to the association between T2DM and pancreatic cancer, only 9 (32%) met the MSAS criteria.

The type of cancer with more than ten studies with an MSAS was prostate cancer. The MSAS of 13 of the 28 effect estimates (46%) was considered by the authors in the adjusted analyses.

4.4.3 Cancer types with no effect estimates meeting MSAS criteria

No effect estimates that accounted for MSAS were available for the following 25 cancer types: anal, carcinoma of unknown primary (CUP), esophageal, esophagogastric junction adenocarcinoma (EJA), extrahepatic cholangiocarcinoma (ECC), gastrointestinal (GI) cancer, Hodgkin lymphoma, intrahepatic cholangiocarcinoma (ICC), Kaposi's sarcoma, large intestine, lip, liver, lymphocytic leukemia, mesothelioma, mouth, myeloid leukemia, myeloma, nose, penile, rectosigmoid, skin, small intestine, stomach, tongue, and urinary tract cancer (UTC).

For all of these cancer types, except liver cancer, at least one confounder from the MSAS was missing in every adjusted analysis. As a result, no effect estimate met the criteria for accounting for the full MSAS.

Table 4-2 presents the confounders identified in the MSAS for each cancer type that were omitted in all adjusted effect estimates available in the included studies.

For instance, among the 13 effect estimates related to T2DM and esophageal cancer risk, none of the analyses included diet, which is a confounder identified in all three MSAS options for this cancer type. In the case of liver cancer, although no single confounder was consistently omitted across all adjusted analyses, each individual analysis failed to include the complete set of MSAS confounders. Therefore, none of the effect estimates for liver cancer met the MSAS criteria either.

Table 4-2 Confounders from the MSAS Omitted Across All Adjusted Effect Estimates, by Cancer Type

Type of cancer	Confounders Omitted Across All Effect Estimates
Anal	drinking/ smoking
CUP	obesity/smoking
Esophageal	diet
EJA	diet/drinking/SES
ECC	drinking/obesity/smoking
GI	diet
Hodgkin lymphoma	smoking
ICC	drinking/obesity/smoking
Kaposi's sarcoma	drinking/smoking
Large intestine	diet/sedentary lifestyle
Lip	drinking/smoking
Lymphocytic leukemia	SES
Mesothelioma	SES
Mouth	drinking/smoking
Myeloid leukemia	Smoking/SES
Myeloma	drinking/obesity/smoking
Nose	Smoking
Penile	drinking/smoking
Rectosigmoid	diet/sedentary lifestyle
Skin	drinking/ smoking
Small intestine	diet
Stomach	diet
Tongue	drinking/smoking
UTC	hypertension

Drinking refers to alcohol consumption, alcohol drinks per day. Sedentary lifestyle refers to physical activity, regular exercise, average hours per week of moderate or vigorous activity. *Abbreviations.* BTC: CUP: Carcinoma of unknown primary; GI: gastrointestinal; ECC Extrahepatic cholangiocarcinoma; EJA: Esophagogastric junction adenocarcinoma; ICC Intrahepatic cholangiocarcinoma; SES: socioeconomic status; UTC: Urinary tract cancer.

The next section presents the proportion of each confounder omitted from the total number of estimates that did not meet the MSAS criteria for each cancer type.

4.4.4 Proportion of confounders omitted in Non-MSAS Estimates by Cancer type

Table 4-3 shows the proportion of each confounder omitted from the effect estimates that did not meet the MSAS for each cancer type. For example, in the case of bladder cancer, the MSAS for estimating a causal effect included race, age, sex, smoking, and SES. Among the 18 effect estimates examining the association between T2DM and bladder cancer risk, only three accounted for the MSAS, meaning that 15 did not. Studies adjusted for age and sex but did not control for SES and smoking. Of the 15 effect estimates that did not meet MSAS criteria, 13 (87%) omitted smoking, and 8 (53%) omitted SES in their adjusted analyses.

Table 4-3 Proportion of Each Confounder Omitted from Effect Estimates That Did Not Meet the MSAS Criteria, by Cancer Type (Color-Coded)

Type of cancer	Number of Effect Estimates Not Meeting MSAS Criteria / Total Estimates	Confounders and Proportion of confounders not included						
		Diet	Drinking	Hypertension	Sedentary lifestyle	Obesity	Smoking	SES
Anal*	1/1		100%				100%	
Bladder	15/18						87%	53%
Brain/NS	4/7					100%		
Breast	35/39		86%		86%	40%		49%
BTC	3/4					100%		
CUP*	1/1					100%	100%	
Cervix uteri	9/10		89%				100%	33%
Colon	23/25	87%	70%		83%	43%	65%	
CRC	29/31	86%	76%		83%	52%	76%	
Endometrial	16/21					50%	88%	
Esophageal*	13/13	100%	77%			54%	69%	38%
EJA*	1/1	100%	100%					100%
ECC*	1/1		100%			100%	100%	
Gallbladder	5/6		100%			40%	100%	
GI cancer*	3/3	100%	33%		33%	33%	33%	
HCC	8/9		38%	88%		25%	50%	
H&N	3/4		100%				100%	33%
Hodgkin lymphoma*	2/2						100%	50%
ICC*	1/1		100%			100%	100%	
Kaposi's sarcoma*	2/2		100%				100%	
Kidney	11/13			73%		55%	73%	
Large intestine*	1/1	100%			100%			
Laryngeal	3/4		100%				100%	33%
Leukemia	5/6						100%	40%
Lip*	2/2		100%				100%	
Liver*	17/17		76%	88%		59%	71%	
Lung	15/21						87%	47%
Lymphocytic leukemia*	1/1							100%
Lymphoma	2/3						100%	50%
Melanoma	4/5		100%				100%	

Type of cancer	Number of Effect Estimates Not Meeting MSAS Criteria / Total Estimates	Confounders and Proportion of confounders not included						
		Diet	Drinking	Hypertension	Sedentary lifestyle	Obesity	Smoking	SES
(Continued from previous page)								
Mesothelioma*	1/1							100%
Mouth*	2/2		100%				100%	
Myeloid leukemia*	1/1						100%	100%
Myeloma*	3/3		100%			100%	100%	33%
NHL	7/9		86%				86%	
Non-melanoma	1/2		100%				100%	
Nose*	1/1						100%	
Oral	5/6		80%				100%	
Ovarian	9/10					67%	100%	22%
Pancreatic	19/28		95%			74%	84%	
Penile*	2/2		100%				100%	
Pharyngeal	2/3		100%				100%	
Prostate	15/28					100%		
Rectal	20/22	85%	65%		80%	45%	65%	
Rectosigmoid*	2/2	100%	50%		100%		50%	
Skin*	3/3		100%				100%	33%
Small intestine*	4/4	100%					75%	
Soft tissue	2/3					100%		
Stomach*	12/12	100%	83%			67%	83%	42%
Thyroid	6/11					100%		
Tongue*	2/2		100%				100%	
UTC*	2/2			100%			50%	50%

The colours represent the proportion of the total number of effect estimates without the MSAS that did not consider each variable: Green (<25%), Yellow (26-50%), Orange (>50-75%), and Red (>75%). Entries with no value in the table indicate that the variable is not included in the MSAS for that specific type of cancer.

Number of Effect Estimates Not Meeting MSAS Criteria / Total Estimates: This column shows the number of effect estimates that did not meet the MSAS criteria out of the total number of effect estimates reported for each cancer type. The total includes all reported estimates, regardless of whether they met the MSAS criteria. The cancer types marked with (*) do not have any effect estimates that account for the MSAS. Diet defined as diet high in fat, calories.

Abbreviations. BTC: Biliary tract cancer; CRC: Colorectal cancer; CUP: Carcinoma of unknown primary; GI: gastrointestinal; UTC: Urinary tract cancer; ECC Extrahepatic cholangiocarcinoma; EJA: Esophagogastric junction adenocarcinoma; H&N Head and Neck cancer; ICC Intrahepatic cholangiocarcinoma; NHL: Non-Hodgkin lymphoma; SES: socioeconomic status; UTC: Urinary tract cancer.

In most cases, the omitted variables were related to unhealthy lifestyle factors, such as diet, drinking (alcohol consumption), hypertension, sedentary lifestyle, obesity, and smoking. However, socioeconomic status (SES) and polycystic ovary syndrome (PCOS) were also among the excluded confounders.

In summary, in the adjusted analyses of 52 cancer types, the relevant MSAS criteria were not met for some cancer types.

- For these, diet was identified as a confounder in 10 types. However, in these 10 cancer types, **diet** was omitted from more than 85% of the effect estimates.
- In 31 cancer types, **drinking** (alcohol consumption) was not included in the adjusted analyses. The proportion of analyses that omitted this variable ranged from 33% to 100%, depending on the cancer type. In 28 of these types, drinking was excluded from more than 70% of the analyses.
- **Hypertension**, which according to the DAGs, should have been considered in the adjusted analyses of four types of cancer (HCC, liver, kidney, UTC), was omitted from over 70% of the effect estimates for each type.
- Physical activity (the inverse of a **sedentary lifestyle**) was identified as a confounder in the relationship between T2DM and seven types of cancer (breast, colon, CRC, GI, large intestine, rectal, and rectosigmoid). In all but the GI cancer analyses, this variable was excluded from at least 80% of the analyses.
- Among the 23 cancer types where **obesity** was part of the MSAS, the proportion of effect estimates that did not meet the MSAS and omitted obesity ranged from 25% (HCC) to 100% (brain cancer, BTC, CUP, ECC, ICC, myeloma, prostate, soft tissue, and thyroid cancers).
- The exclusion of **smoking** as a confounder varied across cancer types. It was omitted in 33% of GI cancer estimates, while in HCC, rectosigmoid, and UTC, smoking was excluded in 50% of the estimates. In colon, esophageal, kidney, liver, rectal, and small intestine cancers, smoking was omitted in

50%-75% of the effect estimates, whereas in the remaining 32 cancer types, it was excluded in more than 75% of the effect estimates.

- In contrast to smoking, the exclusion of **socioeconomic status (SES)** was less pronounced. This variable was omitted in all effect estimates without the MSAS for esophagogastric junction adenocarcinoma (EJA), lymphocytic leukemia, and mesothelioma. In the other 15 cancer types, SES was excluded in fewer than 55% of the estimates (e.g. bladder, breast).
- Polycystic ovary syndrome (**PCOS**), a confounder in the association between T2DM and endometrial cancer and was omitted in 12 analyses.

4.5 Discussion

Evidence from observational studies is valuable not only for evaluating the association between an exposure and an outcome but also for drawing causal associations by addressing confounding. Identifying variables with the potential to introduce bias in effect estimates represents the initial stage in the process of understanding and estimating a causal effect. When dealing with complex relationships, DAGs have been suggested as an effective tool for identifying confounders. Therefore, this visual tool was used in the current research to determine the MSAS that need to be controlled for in order to remove confounding and accurately estimate the causal effect of T2DM on various types of cancer.

The MSAS derived from the DAGs resulted in multiple sets of variables for nearly all cancer types, suggesting that no single set of variables was required to control for confounding. In other words, several different options could have been used, each effectively addressing confounding with the available data. However, when analysing the variables included in the adjusted analyses of the low-risk-of-bias studies from the SR (Chapter 3) across different cancer types, not all confounders from a given MSAS were incorporated. Specifically, estimates using a particular MSAS were reported for only 31 out of 56 cancer types, thus, impeding to ascertain the causal effect of T2DM on the incidence of 25 specific cancer types.

For most cancer types (excluding bone, brain, eye, salivary gland, and testicular cancers), the MSAS typically included at least one lifestyle-related confounder, as lifestyle factors are common risk factors shared by both T2DM and several types of cancer. Among these, smoking and drinking (alcohol consumption) were the primary confounders often omitted in analyses across various cancer types. Other lifestyle factors, such as hypertension, diet, socioeconomic status (SES), sedentary behaviour, and obesity, were frequently excluded from analyses.

The omission of these confounders was usually attributed to the source of the data used. Studies relying on database information often acknowledged in their limitations that they were unable to control for lifestyle-related risk factors due to the unavailability of such data.

For instance, despite smoking being widely recognized as a leading cause of lung cancer, it was omitted from several studies assessing the link between T2DM and lung cancer (American Cancer Society, 2021). It is important to note that these studies aimed to explore associations between T2DM and several cancer types, not only lung cancer, making it difficult to account for confounders specific to each cancer type (Gini et al., 2016; He et al., 2018; Hsieh et al., 2012; Johnson et al., 2011; S. K. Kim et al., 2020; Lee et al., 2012; Lega et al., 2016; Lo et al., 2013; Ogunleye et al., 2009).

In contrast, a study by Chin-Hsiao Tseng, which also used a database but focused solely on lung cancer, used chronic obstructive pulmonary disease (COPD) as a proxy for smoking due to the absence of smoking data (Tseng, 2014). However, while smoking is a primary cause of COPD, other factors, such as exposure to harmful fumes or genetic conditions, can also contribute to the disease (Centers for Disease Control and Prevention (CDC), 2023; National Health Service, 2023). This could introduce potential bias, but it offers a more accurate measure of the lung cancer association than excluding a smoking surrogate. However, the absence of smoking data remains a limitation and may lead to either underestimation or overestimation of the true effect.

Beyond lung cancer, the research by Jong et al. also highlights the lack of lifestyle data due to database reliance, as it investigated the link between T2DM and various cancer types, including pancreatic, biliary tract, gastric, esophageal, and colorectal cancers (de Jong et al., 2017). Similarly, studies focusing on

individual cancer types, such as breast, bladder, colorectal (CRC), prostate, and pancreatic cancers, have noted that limited data prevented control for key confounders (C. B. Chen et al., 2018; Häggström et al., 2017; Hossain et al., 2022; C.-C. Lin, Chiang, Li, Hsieh, et al., 2014; Peeters et al., 2015; Sikdar et al., 2013).

Interestingly, in a study on bladder cancer by Bazelier, despite the lack of detailed smoking data, the authors used smoking cessation drugs and COPD medication as a proxy for smoking (T. Bazelier et al., 2013). They adjusted for all potential confounders that changed the hazard ratio (HR) by more than 1% compared to an age/sex-adjusted analysis. Although the proxy for smoking did not meet the 1% threshold, it was included due to its clinical relevance.

In contrast to this approach, where the inclusion of variables was guided by both clinical relevance and statistical criteria, other authors based their selection of confounders solely on statistical methods. For example, Rollison and colleagues, who investigated the association between T2DM and breast cancer, had access to data on potential confounders, including alcohol consumption, collected through questionnaires and medical records rather than databases (Rollison et al., 2008). However, for the adjusted analyses, they included only variables that changed the odds ratio by at least 5%, which led to the exclusion of alcohol consumption, a confounder in the MSAS for breast cancer.

Similarly, in studies on colorectal (CRC) and esophageal cancer, despite having information on lifestyle-related factors, authors excluded covariates such as alcohol consumption and diet from the final multivariable models, based on their effect on relative risk estimates (Campbell et al., 2010; Ferreira et al., 2021; Petrick et al., 2019).

Therefore, the use of statistical methods for variable selection, like the change-in-estimate criterion (which considers a variable a confounder if removing it from a regression model alters the estimated exposure effect beyond a specified threshold) or P-value-based models, also led to the exclusion of confounders (Williamson et al., 2014).

Another reason for the omission of confounders may be the context. For instance, in populations with low alcohol consumption, such as Iranian and Armenian women, not controlling for alcohol consumption may not likely bias

the study's findings, as the prevalence of alcohol consumption is very low and infrequent in these groups (Dianatinasab et al., 2019; Khachatryan et al., 2011).

Then, from the studies gathered in the SR, the omission of confounders often resulted from unavailable data or the statistical methods used for variable selection. This omission can be appropriate, if it aligns with the author's intention (this will be discussed in the next chapter, Chapter 5). For instance, if the goal is to explore associations, statistical methods are typically employed, and controlling for confounding factors is not necessary (Diemer et al., 2021; Lederer et al., 2019). However, if the authors aim to draw causal conclusions, it is required to identify and control for the variables that confound the relationship between exposure and outcome. In such cases, these statistical methods are insufficient for unbiased causal effect estimation, as they do not consider the underlying causal structure (Hernan, 2002).

DAGs are helpful tools for visualizing and communicating the causal relationships between variables, enabling researchers to identify the MSAS for estimating causal effects (Hernan, 2002). As such, DAGs can serve as a valuable tool in statistical design or, indeed, from data collection, thus facilitating the achievement of the objectives formulated by these authors (Lipsky & Greenland, 2022; Tennant et al., 2021).

In this thesis, DAGs were used to represent the relationships between T2DM and cancer risk factors, identifying the MSAS required to estimate the least biased causal effect of T2DM on the risk of each type of cancer. From the studies in the SR (Chapter 3), those with the MSAS will be selected, and a meta-analysis will be conducted for each cancer type (Chapter 6).

Merits and Limitations of DAGs

Merits

In observational studies, drawing causal conclusions requires addressing confounding. A confounder is a variable that is associated with both the exposure and the outcome but does not lie on the causal pathway between them (Cavanaugh, 2014). For simple causal structures, confounders can often be identified with this classic definition. In complex associations with multiple risk

factors and interconnections, such as in this thesis on the causal effect of T2DM on cancer risk, this definition is insufficient. DAGs were useful in this context, as both T2DM and the different cancer types have diverse and common risk factors (Howards et al., 2012). By analysing the MSAS derived from DAGs, it was possible to identify confounders beyond the shared risk factors between T2DM and cancer, particularly drinking and smoking, because of their connections with other variables. This highlights the importance of understanding the causal structure and shows that DAGs address the limitations of the classical definition in complex epidemiological settings.

Beyond identifying confounders, DAGs also provide a transparent visual framework for stating assumptions about causal structures. By making these assumptions explicit, DAGs clarify the relationships between exposure, outcome, and other variables, helping to distinguish confounders, colliders, and mediators, which must be handled differently to avoid bias (Luijken et al., 2024). This transparency also supports communication among researchers, subject matter experts, and analysts, improving interdisciplinary understanding (Digitale et al., 2022).

DAGs also offer practical benefits. They make it possible to identify multiple alternative MSAS options, each of which can sufficiently control for confounding (Digitale et al., 2022). This flexibility allows researchers to choose adjustment sets based on data availability, as was done in this thesis. Beyond guiding statistical analyses, DAGs are also valuable for planning data collection, since identifying the necessary variables in advance helps ensure alignment with study objectives and avoids unnecessary data gathering (Lipsky & Greenland, 2022; Tennant et al., 2021).

When compared with other causal models, DAGs have the advantage of being relatively simple and broadly applicable. Potential outcome models, such as g-estimation or marginal structural models, are useful for longitudinal data with time-varying exposures but require advanced statistical programming. Structural equation models (SEM) allow detailed specification of quantitative assumptions, but they are complex and demand a strong understanding of statistical concepts. Sufficient-component cause models (SCC) can illustrate hypotheses about

biological mechanisms, although they often lack supporting data. In contrast, DAGs provide an accessible and qualitative depiction of causal assumptions that can be applied across a wide range of study designs (Greenland, Pearl, & Robins, 2002).

Finally, Mendelian randomisation studies (MRS) provide another approach to causal inference by using genetic variants as instrumental variables. They can offer strong evidence for causality, but unlike DAGs, they require access to genetic data, typically single-nucleotide polymorphisms (SNPs), which are not always available (Y. H. Lee, 2020). DAGs, in contrast, can be developed from subject-matter knowledge even in the absence of genetic information. Overall, in complex research such as the causal effect of T2DM on cancer, DAGs help identify confounders and guide statistical analyses.

Limitations

A shortcoming of DAGs is that if the variables or arrows are misspecified, meaning that relevant variables are omitted or the causal directions between them are drawn incorrectly, the resulting MSAS may also be incorrect. To reduce this risk, it has been suggested that variables should be selected using domain knowledge from the literature and through expert consultation (Digitale et al., 2022; Poppe et al., 2024; Sauer & VanderWeele, 2013). In this thesis, targeted searches and expert sources were used, and variables had to be reported in at least two independent sources to be included in the DAG. When uncertainty exists, it has been recommended to construct a series of plausible DAGs and repeat the analyses to assess whether the results differ meaningfully across them (Digitale et al., 2022).

Another limitation of DAGs is their acyclic property. This means that starting from any variable, it is not possible to return to the same node by following the arrows in their direction (Hernán, 2020). Although this ensures logical consistency, it may oversimplify biological or social processes that involve reciprocal relationships. It has been suggested that, to account for temporal order, the same factor should be represented at different time points as separate nodes, since each represents a distinct variable (Suzuki et al., 2020).

DAGs can become complex when many variables are involved. Complex graphs may be difficult to construct and interpret, which can reduce their practical usefulness. To address this challenge, software such as DAGitty is available. DAGitty is a free and user-friendly program (version 3.0) that can be used to draw variables and connections and to identify MSAS or alternative adjustment sets through its automated algorithm (Textor et al., 2016). In this thesis, DAGitty was used to support the construction of the DAGs and the identification of MSAS.

DAGs are nonparametric representations of causal processes. They do not provide information about the strength, magnitude, or functional form of relationships, and they cannot depict effect modification directly (Suzuki et al., 2020; Tennant et al., 2021). To evaluate effect modification, empirical data analysis is required. Similarly, while DAGs can highlight potential unmeasured confounders, they cannot quantify their impact on the causal effect estimate. For this reason, complementary approaches such as quantitative bias analysis are needed to assess how sensitive causal estimates are to unmeasured confounding (Mathur & VanderWeele, 2020).

Finally, while DAGs are valuable for identifying confounders, statistical methods are necessary for estimating causal effects. Relying solely on statistical approaches, such as automated covariate selection (e.g., stepwise procedures, backward elimination, or change-in-estimate criteria), can introduce bias because these methods may select colliders or mediators as covariates rather than distinguishing them from confounders (Hernan, 2002; Luijken et al., 2024; Williamson et al., 2014). This limitation arises because statistical procedures do not incorporate an understanding of causal structure, which is fundamental for confounder identification. Statistical techniques may therefore be useful for reducing the number of variables to achieve a more parsimonious model, but not for identifying confounders, where DAGs provide a more reliable option.

In summary, DAGs have limitations, but with support from literature and software they are valuable for identifying confounders, and statistical analyses alone are not sufficient for unbiased causal inference.

Conclusion

In this chapter, the use of DAGs enabled the identification of the MSAS of variables needed to estimate the causal effect of T2DM on the risk of each cancer type. Many studies from the SR did not account for all MSAS confounders, particularly those related to lifestyle factors. As a result, causal effect estimates were only feasible for 31 of the 56 cancer types.

The following table provides an overview of the thesis structure. The information related to the next chapter is highlighted in yellow.

Table 4-4 Overview of the thesis structure (Next: Chapter 5)

Chapter 1: Introduction. To outline the background of the thesis, introduce the research question and aim, and provide an overview of the thesis structure.

Chapter	Objective	Methods	Key insights
2. Context of Evidence on T2DM and Cancer Risk	To provide an overview of published systematic reviews (SRs) and meta-analyses examining the relationship between T2DM and cancer risk across multiple cancer types	Summarize SR and meta-analyses examining the relationship between T2DM and cancer risk across multiple cancer types	Context for the research conducted in this thesis
3. Systematic review	To identify all relevant evidence on the association between T2DM and cancer risk across various cancer types.	Comprehensive Systematic literature review PICOS framework: men and women aged ≥ 18 years (P); individuals with T2DM (I) compared to those without T2DM (C); cancer incidence (O); observational studies (S).	Types of cancer with evidence on the association between T2DM and cancer risk. Confounders included in the adjusted analyses of the studies. Other*
4. Causal Association	To identify the variables that must be controlled for to estimate the causal relationship between T2DM and cancer across multiple cancer types.	Directed Acyclic Graphs (DAGs) were developed to identify the Minimal Sufficient Adjustment Sets (MSAS)	MSAS for each type of cancer
5. Assessing the quality of covariable adjustment in studies making causal claims	To assess whether studies making causal claims about T2DM and cancer clearly state their intent; justify their choice of covariates; and select them appropriately based on their objectives.	Studies from SR: Analysis of the terminology used across different sections to identify causal claims. Justification of variable selection using clinical evidence or terminology. Comparison of MSAS from DAGs with covariates in adjusted analyses	Emphasize the need for a clear study goal, aligned methods, and consistent terminology
6. Meta-analyses	To establish the causal effect of T2DM on cancer risk	Meta-analyses that account for the MSAS for each cancer type. Subgroup analyses.	Types of cancer causally associated with T2DM
7. Burden of disease	To quantify the future global burden of cancer causally associated with T2DM	Global burden of disease analysis, using the types of cancer causally associated with T2DM (meta-analyses)	Burden of disease metrics: PIF/PFP, the number of cancer cases, DALYs, and monetised DALYs

Cancer incidence (O), expressed as risk ratio (RR), hazard ratio (HR), standardized incidence ratio (SIR), or odds ratio (OR). Other*: Study design, follow-up, country-geographical publications, assessment method of exposure and outcome, Risk of bias assessment. Abbreviations: DALYs: Disability-adjusted life years; PICOS: Population, Intervention, Comparison, Outcome, and Study Design; PIF: population impact fraction; PFP: prevented fraction of the population

Chapter 8: Discussion. To summarize the main findings, outline strengths and limitations, provide policy recommendations, suggest directions for future research, highlight contributions to knowledge, and present final conclusions.

Chapter 5 Assessing the quality of covariable adjustment in studies making causal claims

5.1 Summary

- Fewer studies explicitly (or implicitly) claimed a causal relationship between type 2 diabetes mellitus (T2DM) and cancer compared to those that did not (42% vs. 58%). However, some studies without explicit (or implicit) causal claims still had the potential to provide causal effect estimates, as they accounted for the minimum sufficient adjustment set (MSAS) of variables in their analyses. This shows the importance of ensuring that the study's objective is clearly defined and that the methods used, the interpretation of the findings, and the language throughout the manuscript are consistent with that objective.
- Nearly 70% of studies making causal claims justified their covariate selection based on causal pathways, compared to only one-third of No Explicit Causal Claim Studies. Non-causal studies often relied on statistical criteria for covariate selection, which aligns with their objectives.
- Although there is some evidence supporting a causal link between T2DM and most cancer types, the covariates used in the adjusted analyses for each cancer type were often inadequate. Less than 20% of the effect estimates accounted for the MSAS.
Key lifestyle-related confounders were frequently omitted due to data limitations. This highlights the need to explore additional data sources or carefully plan data collection to ensure alignment with the study's aim.

5.2 Introduction

Objective

The aims of this chapter are to: (1) review whether studies are explicit with respect to causal claims, (2) review whether the choice of covariables is explicitly justified, and (3) consider whether the choice of covariables appears appropriate given the explicit (or implicit) study objective.

Background

As described in the previous chapter, some studies retrieved in the systematic review (SR) in Chapter 3 aimed to explore the association, while others explicitly stated that their objective was to characterise an association providing evidence of a causal relationship between T2DM and cancer risk. Clear and precise language is essential in order to communicate research goals, prevent misinterpretations, and ensure a thorough understanding of the implications of conclusions.

When making causal claims, it is important to explicitly state the objective of the analysis and explicitly justify covariate selection while maintaining consistency with the language used throughout the study. Phrases and terms like "assessed a causal association" or "influence" in the objectives indicate an intent to establish causal relationships. Similarly, phrases in the methods section, such as "adjusted for potential confounders," provide valuable guidance that the authors aim to draw causal conclusions (Haber et al., 2022; Olarte Parra et al., 2021). However, the misuse of association and causal association statements without distinction is a common issue in research writing (M. A. Han et al., 2022). For example, researchers may use phrases such as "association with" in the objectives but then present a causal claim in the conclusion with words like "impact" or "attributable".

This inconsistency in language has been reported in various therapeutic areas, such as obesity and nutrition, as well as psychology (Cofield et al., 2010; Olarte Parra et al., 2021). Likewise, inconsistencies between research aims and covariate selection were evident in the studies included in the SR for this thesis. For instance, Bosco and colleagues included a section in their methods titled "Potential Confounding Variables," stating that "candidate confounding variables were identified a priori from the existing literature." However, their objectives only indicated an aim to "examine the association between T2DM and breast cancer," and their conclusions also used the term "association" without making any causal claims (Bosco et al., 2012).

An inconsistency was also identified in the study by Campbell, which focused on associations between T2DM and CRC risk (Campbell et al., 2010). Although the objectives did not indicate an intent to establish causality, the methods section

stated, “covariates were selected at the onset of the study, based on their potential to confound or modify the association....” This misalignment between the stated objectives and covariate justification was also noted in Eijgenraam's study on the association between T2DM and pancreatic cancer, as well as in the studies by Yanan and Peeters, which focused on T2DM and CRC risk (Eijgenraam et al., 2013; Ma et al., 2018; Peeters et al., 2015).

Interestingly, Suk Kyeong's research explicitly stated the objective “to determine whether the association between diabetes and cancer was causally related.”

Nonetheless, the statistical analyses lacked justification for covariate selection, and terms such as “confounders” or related phrases like “confound” or “modify” were absent from the manuscript (S. K. Kim, Jang, Kim, Rhyu, Lee, et al., 2020).

5.3 Methods

This section outlines (1) the systematic approach used to classify studies from the SR (Chapter 3) into those making explicit causal claims and those that do not, (2) the assessment of whether covariate selection in the adjusted analyses was explicitly justified, and (3) the evaluation of the appropriateness of the selected covariates by comparing them with the minimum sufficient adjustment set (MSAS) derived from Directed Acyclic Graphs (DAGs) for different cancer types (Chapter 4).

5.3.1 Identification of studies claiming causality

The first objective of this chapter was to identify studies from the SR that made either an explicit or implicit **claim of a causal relationship** between T2DM and the risk of cancer (referred to as [Explicit Causal Claim Studies](#)). However, identifying these studies was challenging, as authors often did not use clear language to explicitly denote causal relationships. In some cases, statements suggesting causality appeared only in the discussion, while the objectives or conclusions were framed in non-causal terms. As a result, sections of the same article seemed non-causal, whereas other parts could be interpreted as addressing a causal research question.

The inappropriate use of terminology made it difficult to distinguish studies that focus on causal inference. To address this, the **flowchart** shown below was uniquely developed to facilitate the selection of studies that made causal claims, either explicitly in the study objectives or implicitly through implications drawn from the conclusions, methods and discussion sections. This approach also helped to differentiate these studies from those that did not make any explicit or implicit claim of a causal relationship between T2DM and cancer (referred to as [No Explicit Causal Claim Studies](#)).

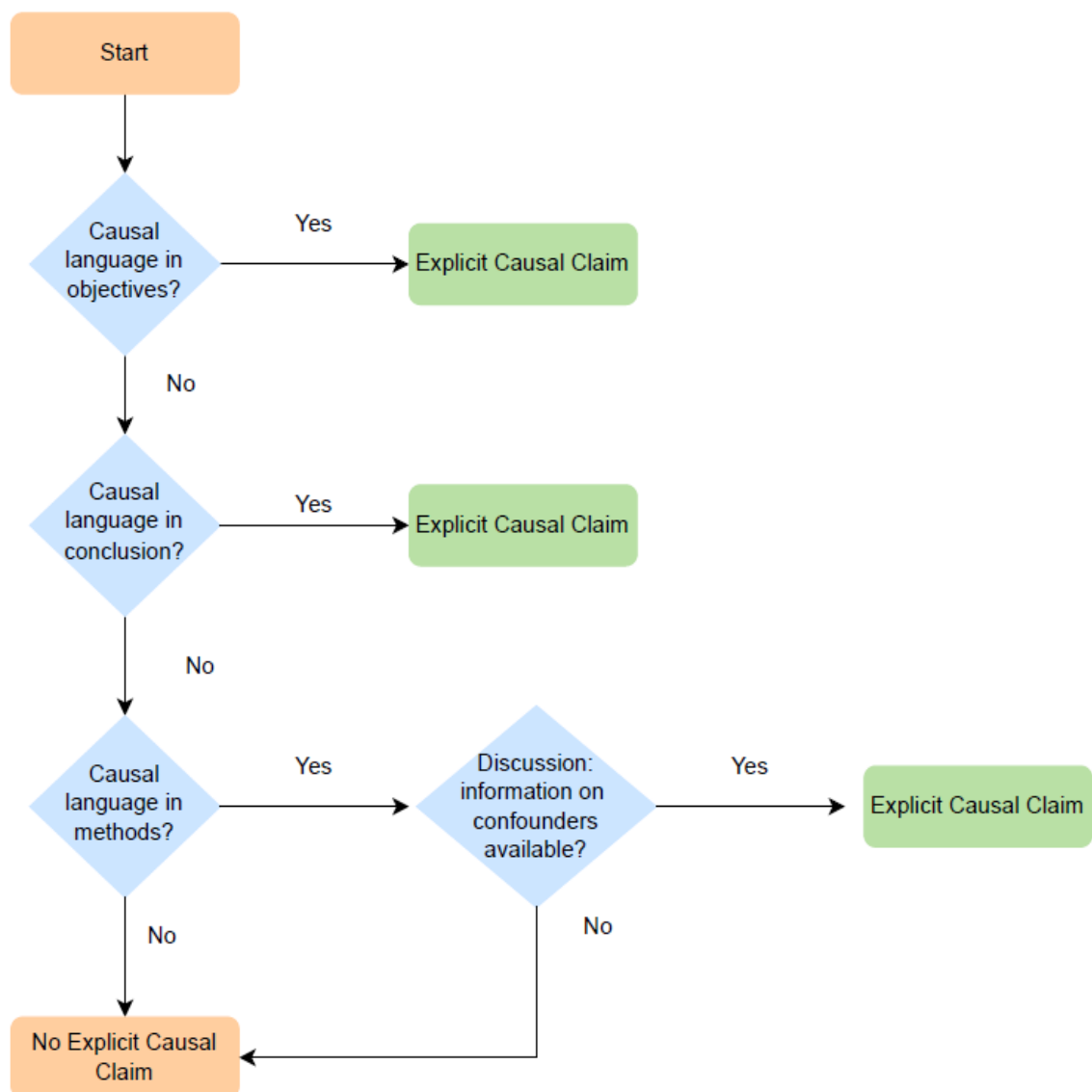


Figure 5-1 Flowchart to identify studies claiming causality. The flowchart was designed to identify the studies claiming a causal association between T2DM and the risk of cancer. *Explicit Causal Claim* refers to a study that makes an explicit or implicit claim of a causal relationship between T2DM and the risk of cancer.

The flowchart considered the language used in different parts of the paper. It was applied step by step: starting with the objectives (1), then the conclusions (2), followed by the methods (3), and finally the discussion (4). If causal language was identified at any of these stages, in this order (indicated as yes in the flowchart), the study was classified as making an “Explicit causal claim”.

Each observational study was independently reviewed for the presence of language implying causation. This systematic screening process made it possible to assess whether authors were truly addressing a causal question ([Explicit Causal Claim Studies](#)), as reflected in the statements found in these sections. In this context, causal language was defined as the use of linking words or phrases that indicate causal relationships or intent, consistent with recommendations by Olarte Parra et al. (2021), who emphasised that good causal language practice requires awareness of which terms imply causality and which do not (Olarte Parra et al., 2021).

A detailed explanation of how the flowchart was applied to each section is provided below.

1. Objectives. The flowchart started by examining the **objectives section**, where the study's interest in estimating causal effects between T2DM and cancer was identified. This was the first section assessed, as the study's goals guide the analytical approach and interpretation of the findings. In this section, authors should explicitly articulate the causal question and specify the causal effect of interest, clearly stating their objectives and using language appropriate to those aims (Hernán, 2018; Kezios, 2021).

Studies were classified as either:

- **Explicit Causal Claim Studies:** If the study made an explicit or implicit claim of a causal relationship between T2DM and cancer.
- **No Explicit Causal Claim Studies:** If the study did not make an explicit or implicit claim of a causal relationship between T2DM and cancer.

In the **Objectives section**, the use of causal wording was directly linked to the study's stated purpose. Terms such as *influence*, *effect*, *cause/causal*, and *impact* were used to identify **Explicit Causal Claim Studies**.

For example, causal language in objectives might include phrases like “investigate a causal relationship”. In contrast, terms like *association*, *relation/relationship*, *correlation*, and *risk* were associated with **No Explicit Causal Claim Studies** (M. A. Han et al., 2022; Olarte Parra et al., 2021).

However, when causal terms were not identified in the objectives, other sections of the paper could still reflect a causal interest. For this reason, the next section reviewed was the conclusions.

2. Conclusion. If the authors' intention to address a causal question was unclear after assessing the objectives, the **conclusions section** was examined. In certain cases, inconsistencies were observed between the objectives and conclusions: while the objectives were phrased in terms of assessing an association, the conclusions presented findings in causal terms. For this reason, the conclusions were also screened, as they may represent an implicit statement of the study's objectives. This step is also supported by the recommendations of Olarte, who emphasizes the importance of the objectives and conclusions sections in identifying and classifying the research objective (Olarte Parra et al., 2021).

In the **Conclusions section**, terminology such as *cause*, *contribution*, *effect*, *impact*, *recommendation for change*, *attributable*, *leads to*, *responsible for*, *influence*, *affects*, *improves*, *modifies*, *determinant*, and *etiological role* was used to identify **Explicit Causal Claim Studies**. For example, causal language in conclusions included phrases like “we did not find that X cancer incidence was influenced by diabetes” or “these results strongly suggest a possibility for an etiologic role of T2DM in the pathogenesis of X cancer.” Meanwhile, terms like *associated*, *risk*, *prediction*, and *related* were used for **No Explicit Causal Claim Studies** (M. A. Han et al., 2022; Olarte Parra et al., 2021).

Even though many authors did not use terms in the objectives or conclusions that implied causal claims, it was still evident that they were interested in causality. This was reflected in the methods section, particularly in the consideration and adjustment for confounders. Therefore, the next section screened was the methods.

3. **Methods.** If the authors' intent regarding a causal association remained unclear after reviewing the conclusions, the **methods section** was examined for terms that suggest *confounding or its derivatives (e.g., confounder, confounding, residual confounding)* (Hernán, 2018; Olarte Parra et al., 2021). Since drawing causal conclusions requires identifying and controlling for variables that confound the relationship between exposure and outcome, the use of terms related to confounding in the statistical methods section provided evidence of whether the study aimed to make causal claims (Hernán, 2002). When such terms were present, the **discussion section** was also screened to assess how the strengths and limitations related to confounders were addressed. This helped to resolve potential contradictions between the methods and discussion, and is explained further in number 4.

Conversely, when the concept of confounding or its derivatives was not mentioned, or when the selection of variables for the adjusted analyses was not explained or justified, the study was unequivocally classified as a **No Explicit Causal Claim Study**. If the aim is not to answer a causal question, adjustment for confounding is unnecessary; in such cases, variable selection may be based solely on statistical criteria. Confounding is a causal concept and does not apply to associational studies (Diemer et al., 2021; Lederer et al., 2019).

4. **Discussion.** This section was screened because it highlights potential misalignments between study goals, methods, and interpretations (Kezios, 2021).

Therefore, if the methods section specified that confounders were considered, but among the **limitations in the discussion** section, a lack of data on a variable that could potentially confound the results was mentioned, it was interpreted that the authors acknowledged their inability to fully control for confounding and were not implying a causal relationship (**No Explicit Causal Claim Study**). This approach ensured that both causal language and disclaimers or limitations were considered when interpreting causal intent. In this context, a disclaimer refers to a statement made by the authors to limit or soften their claims, in line with the recommendations of

Olarte Parra et al. (2021) and Stacey et al. (2010). Furthermore, if no mention of confounding factors or derivative terms appeared in the discussion section, the study was also classified as a **No Explicit Causal Claim Study**.

Conversely, if the methods section stated that confounders were included and among the **strengths in the discussion** section, the availability of information on potential confounders was highlighted, it was assumed that the authors were implicitly claiming causality (**Explicit Causal Claim Study**). This reflects consistency between the methods and the discussion, ensuring the methodological rigour required for causal inference (Haber et al., 2022).

5.3.1.1 Quality Checks in the Classification Process

To strengthen the reliability of the judgements made during study classification, the following quality checks were incorporated into the process.

As previously mentioned, a unique flowchart was developed to identify Explicit Causal Claim Studies. To ensure that all relevant studies were captured, minimise the risk of misclassification, and improve accuracy, it is recommended that a dual assessment be applied to a proportion of records. Any discrepancies between individual judgements should be discussed and resolved by consensus. Furthermore, if more than 20% of classifications are found to be inconsistent, it is recommended that all papers (100%) be reviewed by all reviewers (Nussbaumer-Streit et al., 2023).

For this thesis, the primary author (Fabiola Lemus) independently reviewed all records. Two additional reviewers (Professor Neil Hawkins and Professor Kathleen A. Boyd) independently (and blinded) assessed a 5% random sample of records to validate the classification process. This double-screening exercise revealed complete agreement, with no discrepancies identified. Consequently, extending the dual screening to 100% of records was not necessary.

5.3.2 Assessing Justification for Covariable Selection

Ensuring a formal justification for the considered covariates in the adjusted analyses is essential yet often not done. This reveals an understanding of the causal structure among the variables relevant to the research question and those that could potentially introduce bias in the effect estimates.

In causal inference, the selection of covariates with the potential to confound the association should be guided by clinical evidence to ensure that all relevant confounders are included. Parsimonious models, which limit variables for simplicity, can be useful, but it is important not to exclude confounders merely to reduce complexity. In causal analysis, all variables needed to control for confounding should be incorporated, even if this leads in a more complex model. Understanding these relationships is essential for identifying mediators, colliders, and confounders to avoid biased causal estimates (Luijken et al., 2024).

Given that the second objective of this chapter was to assess whether the choice of covariates was explicitly justified, with a focus on studies claiming causality, studies were categorized based on whether a **justification for variable selection** was provided in the statistical or data analysis section, using clinical evidence and terms such as **risk factors, confounders, or references to previous clinical knowledge** (Hernán, 2018; Olarte Parra et al., 2021).

Conversely, when these terms were omitted, or only a statistical technique was referenced (e.g., forward or backward selection, stepwise selection), studies were classified as not justifying variable selection due to the limitations of these techniques, which do not account for the potential position of variables in causal pathways (Hernan, 2002). Additionally, studies that provided no explanation of how variables were selected were also classified in the no-justification group.

5.3.3 Assessing the Appropriateness of Covariable Selection in Relation to Study Objectives

To assess whether the choice of covariates was appropriate given the study's explicit or implicit objective, the MSAS derived from the DAGs developed for different cancer types in Chapter 4 was compared with the covariates included in the studies' adjusted analyses.

For studies classified as Explicit Causal Claim Studies, the inclusion of the MSAS in the adjusted analyses is essential, as the MSAS represents the set of variables that must be controlled for in order to estimate a causal effect.

5.4 Results

This section presents the results of the classification of studies based on whether they made explicit causal claims, the assessment of justification for covariate selection, and the appropriateness of the selected covariates given the study's stated or implied objective.

5.4.1 Classification of Studies by Causal Claims, Covariate Selection, and Justification

Figure 5-2, shows the proportion of studies classified as Explicit Causal Claim and Non-Explicit Causal Claim. For each group, it also presents the proportion of studies that did or did not include the MSAS, and among those, whether or not a justification was provided for the choice of covariates in the adjusted analyses.

Some authors evaluated multiple cancer types within a single study, resulting in multiple effect estimates based on the number of analyses and cancer types assessed. Among the low-risk-of-bias studies, **31% (48 out of 149)** included at least one effect estimate that accounted for the **MSAS**.

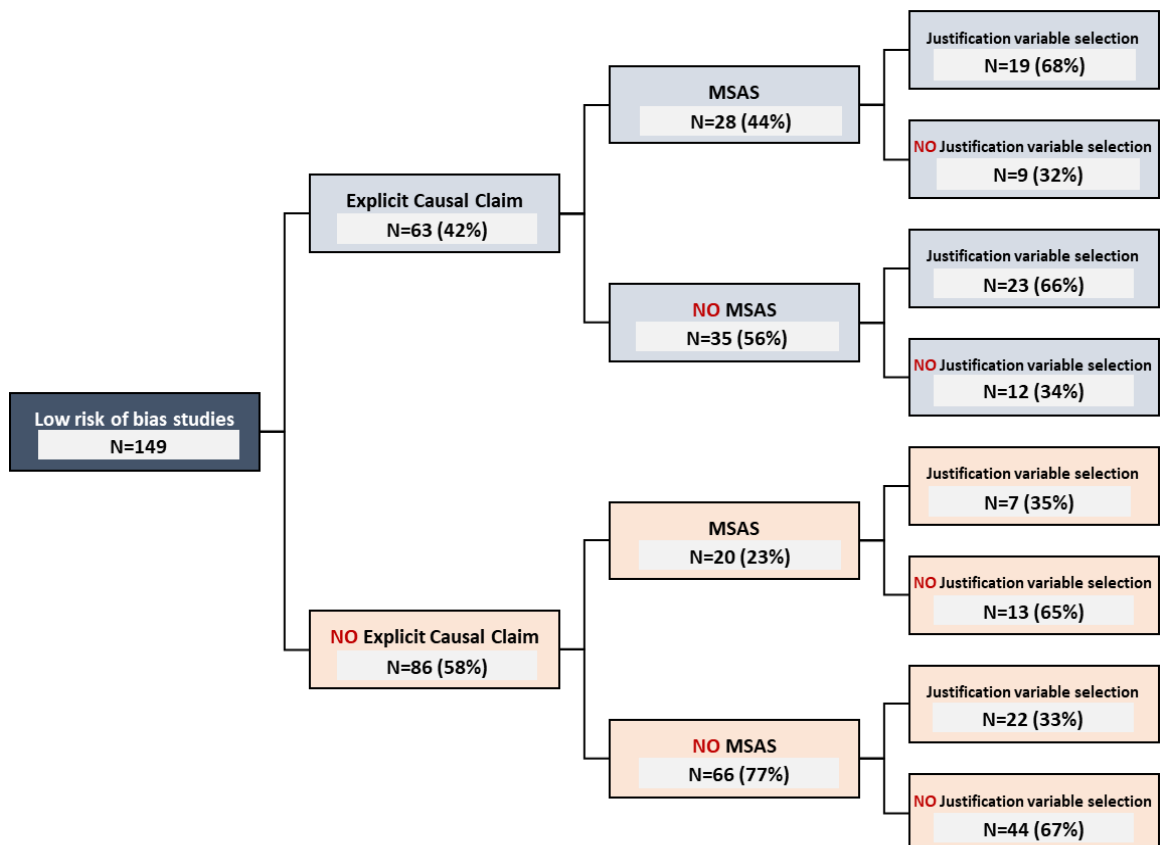


Figure 5-2 Number and proportion of studies categorized as Explicit Causal Claim and No Explicit Causal Claim, with at least one estimate accounting for the MSAS and providing a justification for variable selection.

The figure shows that, among the 149 low-risk-of-bias studies, **63 (42%) made explicit (or implicit) causal claims**, while the remaining 86 studies (58%) were classified as not making explicit (or implicit) causal claims.

For Explicit Causal Claim Studies:

- 44% (28/63) included at least one estimate that accounted for the MSAS
- 67% (42/63) provided a justification for variable selection
- Of the 28 studies that included at least one estimate with the MSAS, 19 (68%) provided a justification for variable selection

For No Explicit Causal Claim Studies:

- 23% (20/86) included at least one estimate that accounted for the MSAS
- 34% (29/86) provided a justification for variable selection
- Of the 20 studies that included at least one estimate with the MSAS, 7 (35%) provided a justification for variable selection

The proportion of studies classified as Explicit Causal Claim studies was similar across study designs, with 40% (42/104) of cohort studies and 47% (21/45) of case-control studies making explicit (or implicit) causal claims.

Of the **56 types of cancer** for which a DAG was developed (Chapter 4), **37 (66%) had at least one study** explicitly (or implicitly) **claiming a causal relationship** between T2DM and the cancer type. The proportion of MSAS by cancer type in studies making explicit causal claims is described in more detail in the next section.

On the other hand, the 19 cancer types without at least one study making explicit (or implicit) causal claims were: anal, bone, carcinoma of unknown primary (CUP), , extrahepatic cholangiocarcinoma (ECC), eye, Hodgkin lymphoma, intrahepatic cholangiocarcinoma (ICC), Kaposi's sarcoma, lip, lymphocytic leukemia, mesothelioma, mouth, myeloid leukemia, myeloma, nose, penile, salivary gland, testicular, and tongue.

To sum up, among the low-risk-of-bias studies, fewer studies made explicit causal claims compared to those that did not (42% vs. 58%). However, the proportion of studies including at least one estimate with the MSAS in the adjusted analyses was higher in the group making explicit causal claims than in the group without such claims (44% vs. 23%).

Additionally, studies classified as Explicit Causal Claim studies and accounting for at least one estimate with the MSAS were more likely to justify variable selection than those classified as No Explicit Causal Claim studies (68% vs. 35%).

5.4.2 MSAS in studies making explicit (or implicit) causal claims according to cancer type

Some authors reported findings for multiple cancer types within the same publication, leading to multiple effect estimates based on the number of analyses and cancer types studied. As a result, a total of 438 effect estimates were documented in the 149 low-risk of bias studies. Of these, **151 effect estimates (35%) were from studies making explicit (or implicit) claims of causality**, and the remaining 287 (65%) were from studies that did not make explicit claims of causality.

In studies making explicit claims of causality, **50 effect estimates (33%) incorporated the MSAS**, while 101 effect estimates (67%) did not include all covariates from the MSAS derived from the DAGs.

As shown in Figure 5-3, among the studies selected from the systematic review and **classified as making explicit (or implicit) causal claims**, CRC had the largest number of effect estimates (N = 14), followed by breast cancer (N = 12), and endometrial, colon, and rectal cancers, each with 11 effect estimates. Pancreatic cancer had 10 effect estimates.

Of these cancer types, **pancreatic cancer had the highest proportion of effect estimates considering the MSAS (70%)**, followed by endometrial cancer (45%). The remaining cancer types had proportions of less than 20%.

These findings suggest that the covariates selected to address confounding in studies making explicit (or implicit) causal claims are not adequately aligned with the study's aims, based on the MSAS derived from the DAGs.

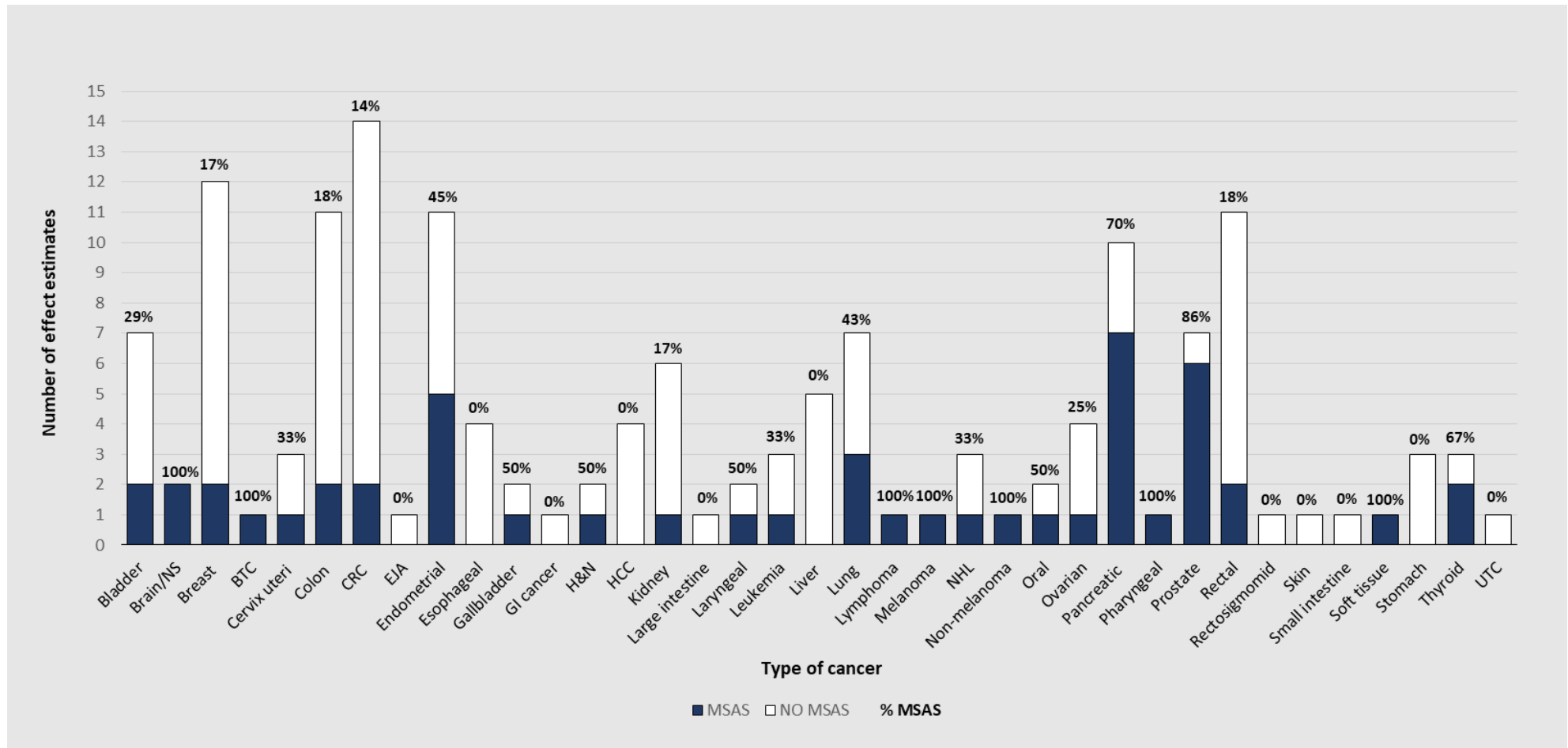


Figure 5-3 Number of effect estimates by cancer type in studies making explicit causal claims, categorized by whether the MSAS was accounted for (MSAS vs. No MSAS). Percentages represent the proportion of effect estimates that included the MSAS.

The confounders omitted from the adjusted analyses in studies making explicit (or implicit) causal claims were primarily related to an unhealthy lifestyle, including diet, drinking (alcohol consumption), hypertension, sedentary behaviour, obesity, and smoking. However, exceptions to this pattern included socioeconomic status (SES) and polycystic ovary syndrome (PCOS).

Table 5-1 shows the confounders omitted in studies making explicit causal claims that did not meet the MSAS criteria, categorized by cancer type.

Table 5-1 Confounders omitted in in Studies Making Explicit Causal Claims That Did Not Meet the MSAS Criteria, by Cancer Type

Type of cancer	NO MSAS	diet	drinking	hypertension	sedentary lifestyle	obesity	smoking	SES
Bladder	5						3	4
Breast	10		7		8	2		6
Cervix uteri	2		2				2	1
Colon	9	6	3		5	1	3	
CRC	12	9	7		8	3	7	
EJA	1	1	1					1
Endometrial	6					1	5	
Esophageal	4	4	3			1	2	2
Gallbladder	1		1				1	
GI cancer	1	1						
H&N	1		1				1	1
HCC	4			4			1	
Kidney	5			3			2	
Large intestine	1	1			1			
Laryngeal	1		1				1	
Leukemia	2						2	1
Liver	5		2	4		1	2	
Lung	4						2	3
NHL	2		1				1	
Oral	1		1				1	
Ovarian	3					1	3	1
Pancreatic	3		3			1	2	
Prostate	1					1		
Rectal	9	6	3		5	1	3	
Rectosigmoid	1	1			1			
Skin	1		1				1	1
Small intestine	1	1						
Stomach	3	3	2			1	2	1
Thyroid	1					1		
UTC	1			1				1
Total effect estimates	101	33	39	12	28	15	47	23

Diet: diet high in fat, calories (diets high in sugars, saturated and trans- fats, low fibre foods and high-sugar drinks). Drinking: alcohol consumption. SES: socioeconomic status.

Of the 101 effect estimates from studies making explicit (or implicit) causal claims that did not include all covariates from the MSAS derived from the DAGs, the following covariates were omitted: smoking in 47 effect estimates, alcohol consumption (drinking) in 39, diet in 33, sedentary lifestyle in 28, socioeconomic status (SES) in 23, obesity in 15, and hypertension in 12.

Among the omitted confounders, **smoking was the most commonly omitted** variable, as it was a covariate in the MSAS for 21 cancer types. In the context of lung cancer, it is noteworthy that despite smoking being a known risk factor for both T2DM and lung cancer, two out of the four effect estimates that did not meet the MSAS criteria failed to adjust for this variable (American Cancer Society, 2021).

Additionally, polycystic ovary syndrome (PCOS) was omitted in three effect estimates from studies on endometrial cancer. This covariate is a confounder in the association between T2DM and endometrial cancer, as it is not only a risk factor for this cancer but also frequently leads to insulin resistance in women with PCOS, increasing their risk of developing T2DM (A. Agrawal et al., 2023; American Cancer Society, 2021).

Regarding cancer types with the highest number of effect estimates that did not meet the MSAS criteria in studies making explicit causal claims:

- **Breast cancer:** SES, a confounder included in the MSAS for breast cancer, was omitted in 6 out of 10 effect estimates that did not meet the MSAS criteria.
- **Colorectal, colon, and rectal cancers:** Diet, a well-established risk factor for both T2DM and these cancers, was omitted in two-thirds of the effect estimates that did not meet the MSAS criteria.
- **Colorectal cancer (CRC):** Smoking, a common risk factor for T2DM and CRC, was omitted in 7 out of 12 effect estimates in CRC studies.

5.5 Discussion

Randomized controlled trials (RCT) are considered the gold standard for evaluating intervention effects. However, they are often not feasible for assessing exposures and their relationship with rare events with a long latency, such as cancer incidence. In such cases, observational studies are the most suitable design for investigating the association between T2DM and cancer incidence. These study designs not only allow for the exploration of associations but also enable the drawing of causal conclusions. Nevertheless, identifying studies that explicitly or implicitly claim a causal relationship between T2DM and cancer can be challenging, as authors often fail to use clear, precise, and appropriate terminology.

This chapter aimed to assess the quality of covariate adjustment in studies making causal claims. The first step involved identifying studies from the SR (Chapter 3) that made either an explicit or implicit claim of a causal relationship between T2DM and cancer risk (*Explicit Causal Claim Studies*) and distinguishing them from those that did not (*Non-Explicit Causal Claim Studies*). To support this process, a flowchart was uniquely developed to analyse the terminology used across different sections of each paper to identify causal claims. Furthermore, the appropriateness of covariate selection for confounding adjustment was evaluated by comparing the MSAS, derived using DAGs, with the covariates included in the adjusted analyses.

The findings showed that fewer studies (42%) were classified as *Explicit Causal Claim Studies* compared to *No Explicit Causal Claim Studies* (58%). However, studies making explicit causal claims were more likely to include at least one MSAS in their analyses (44% vs. 23%). This result aligns with the methodological approach defined in this thesis for establishing causality, which was based on the inclusion of the MSAS derived from DAGs to address confounding. Conversely, studies not intending to make explicit causal claims are not required to include the MSAS in their adjusted analyses, as their objectives do not focus on causal inference.

In addition, of the 56 cancer types for which a DAG was created, nearly 70% had at least one study explicitly claiming a causal relationship between T2DM and the cancer type. This indicates that the majority of cancer types have some documented evidence supporting a causal link with T2DM.

An analysis by cancer type showed that, with the exception of pancreatic and endometrial cancer, most cancer types had fewer than 20% of effect estimates that accounted for the MSAS in studies making explicit causal claims.

These findings suggest that, for most cancer types, covariate selection for confounding adjustment was not adequately aligned with the study objectives.

The omitted confounders were primarily related to unhealthy lifestyle factors, with smoking being the most frequently excluded, despite being part of the MSAS for 21 cancer types. As discussed in Chapter 4, the primary reason for the omission of these confounders was often attributed to data limitations. Studies relying on database sources frequently acknowledged their inability to control for lifestyle-related risk factors because such information was not available.

This chapter also assessed whether the selection of covariables in the adjusted analyses was explicitly justified. Justifications were evaluated based on references to clinical evidence, the use of terms such as *risk factors* or *confounders*, or *prior clinical knowledge*. Nearly 70% of studies that explicitly claimed causality provided a justification, compared to only about one-third of those classified as No Explicit Causal Claim Studies. These findings align with the criteria for defining justification and suggest that the methods used to select covariates for adjusted analyses were generally aligned with the research aims. In studies not aiming to establish causality, variables are often chosen based on statistical criteria or simply listed without considering their potential position in the causal pathway. Although this approach does not account for causal structure or confounding bias risks, it is appropriate for studies not aiming to establish causality (Gaskell & Sleight, 2020; Hernan, 2002).

Among the studies classified as No Explicit Causal Claim, 23% had the potential to establish causality because they accounted for the MSAS. This difference between the study's goal and the interpretation of the findings could be because

the authors were unaware that these variables were sufficient to control for confounding. Another possible reason is that various journals have different policies regarding using causal language when presenting data from observational studies. For example, some journals, like the *Annals of the American Thoracic Society (AnnalsATS)*, allow causal terms, while others, such as the *European Annals of Otorhinolaryngology*, the *Anatolian Journal of Emergency Medicine*, and the *Journal of the American Medical Association (JAMA Network)*, explicitly prohibit the use of terms implying causation when reporting findings from these study designs. These journals support the belief that only RCTs can establish causal relationships (Haber et al., 2022; M. A. Han et al., 2022).

This subject has been thoroughly examined in various health-related disciplines (Cofield et al., 2010; Olarte Parra et al., 2021). The authors have emphasized their concern regarding the inconsistent use of terminology across the manuscript and the need for better alignment between the study's aim and the interpretation of findings. This misalignment confuses readers and makes it harder to understand the authors' message. As a result, it complicates the practical use of the findings, which could be valuable for decision-makers (Haber et al., 2022).

In order to address this issue, authors are required to clearly define the goal of the study, ensure the research process aligns with that goal, apply appropriate analytical methods, and interpret the findings correctly. Additionally, they need to maintain consistency in language and terminology according to the research question and the study's objective across all sections of the manuscript (Kezios, 2021).

Furthermore, reviewers should recognize that observational studies can provide insights into causal relationships. Therefore, there is a growing need for the use of proper causal language to become more widespread in the research community.

Inconsistencies in study objectives, analytical methods, and terminology made it difficult to distinguish between studies making causal claims and those that did not. However, as the goal of this thesis is to assess the causal effect of T2DM on cancer risk, and since controlling for the MSAS addresses confounding bias, the

next chapter presents meta-analyses for each cancer type using studies that included the MSAS to obtain the least biased causal effect estimates.

Conclusion

This chapter assessed the quality of covariate adjustment in studies making causal claims. Among the studies included in the systematic review, fewer explicitly or implicitly claimed a causal relationship between T2DM and cancer (42%) compared to those that did not (58%). Nearly 70% of the studies making causal claims provided a justification for their covariate selection, suggesting alignment with their stated objectives.

Although evidence supporting a causal link between T2DM and most of the cancer types evaluated (31 out of 56) was identified, an analysis by cancer type showed that fewer than 20% of the effect estimates in studies making causal claims accounted for the MSAS. This suggests that, in studies making causal claims, the covariate selection for confounding adjustment was not adequately aligned with their stated objectives.

The following table provides an overview of the thesis structure. The information related to the next chapter is highlighted in yellow.

Table 5-2 Overview of the thesis structure (Next: Chapter 6)

Chapter 1: Introduction. To outline the background of the thesis, introduce the research question and aim, and provide an overview of the thesis structure.

Chapter	Objective	Methods	Key insights
2. Context of Evidence on T2DM and Cancer Risk	To provide an overview of published systematic reviews (SRs) and meta-analyses examining the relationship between T2DM and cancer risk across multiple cancer types	Summarize SR and meta-analyses examining the relationship between T2DM and cancer risk across multiple cancer types	Context for the research conducted in this thesis
3. Systematic review	To identify all relevant evidence on the association between T2DM and cancer risk across various cancer types.	Comprehensive Systematic literature review PICOS framework: men and women aged ≥ 18 years (P); individuals with T2DM (I) compared to those without T2DM (C); cancer incidence (O); observational studies (S).	Types of cancer with evidence on the association between T2DM and cancer risk. Confounders included in the adjusted analyses of the studies. Other*
4. Causal Association	To identify the variables that must be controlled for to estimate the causal relationship between T2DM and cancer across multiple cancer types.	Directed Acyclic Graphs (DAGs) were developed to identify the Minimal Sufficient Adjustment Sets (MSAS)	MSAS for each type of cancer
5. Assessing the quality of covariable adjustment in studies making causal claims	To assess whether studies making causal claims about T2DM and cancer clearly state their intent; justify their choice of covariates; and select them appropriately based on their objectives.	Studies from SR: Analysis of the terminology used across different sections to identify causal claims. Justification of variable selection using clinical evidence or terminology. Comparison of MSAS from DAGs with covariates in adjusted analyses	Emphasize the need for a clear study goal, aligned methods, and consistent terminology
6. Meta-analyses	To establish the causal effect of T2DM on cancer risk	Meta-analyses that account for the MSAS for each cancer type. Subgroup analyses.	Types of cancer causally associated with T2DM
7. Burden of disease	To quantify the future global burden of cancer causally associated with T2DM	Global burden of disease analysis, using the types of cancer causally associated with T2DM (meta-analyses)	Burden of disease metrics: PIF/PFP, the number of cancer cases, DALYs, and monetised DALYs

Cancer incidence (O), expressed as risk ratio (RR), hazard ratio (HR), standardized incidence ratio (SIR), or odds ratio (OR). Other*: Study design, follow-up, country-geographical publications, assessment method of exposure and outcome, Risk of bias assessment. Abbreviations: DALYs: Disability-adjusted life years; PICOS: Population, Intervention, Comparison, Outcome, and Study Design; PIF: population impact fraction; PFP: prevented fraction of the population

Chapter 8: Discussion. To summarize the main findings, outline strengths and limitations, provide policy recommendations, suggest directions for future research, highlight contributions to knowledge, and present final conclusions.

Chapter 6 Meta-analyses

6.1 Summary

- Type 2 diabetes mellitus (T2DM) has a causal effect on the increased risk of breast, colon, endometrial, kidney, and pancreatic cancer. Conversely, T2DM has a protective causal effect against prostate cancer.
- Adjusted (causal effect) meta-analyses provide more accurate effect estimates than unadjusted (*association*) analyses by addressing confounding bias, but this is relevant only when the goal is causal inference, as unadjusted meta-analyses are appropriate for exploring associations.
- Accounting for confounding variables reduces heterogeneity, leading to more consistent effect size across studies.
- In the adjusted (causal effect) meta-analyses, subgroup analyses showed no statistically significant differences in cancer risks between males and females and across geographic regions (as a proxy for race). Also, subgroup analyses by study design (cohort and case-control studies) revealed that while the magnitude of the estimates varied between the two designs, these differences were not statistically significant, indicating that the results from both designs are comparable. Furthermore, the subgroup analysis by menopausal status found no statistically significant difference in the risk of developing breast cancer due to T2DM between premenopausal and postmenopausal women.
- Future research focusing on the impact of diabetic treatment and the duration of T2DM will allow a better understanding of the underlying aetiology between both diseases.

6.2 Introduction

Objective

The aims of this chapter are to conduct meta-analyses to establish the causal effect of T2DM on cancer risk. Only effect estimates from adjusted analyses that include the minimum sufficient adjustment set (MSAS) of variables, identified a priori using the corresponding Directed Acyclic Graphs (DAG) for each cancer type, will be selected.

Subgroup analyses will be performed by sex and race, as well as by cohort and case-control design, to determine whether the true effect varies by study design. If sufficient data are available, additional subgroup analyses will examine menopausal status for breast cancer, time since T2DM diagnosis in relation to cancer incidence, and antidiabetic drug classes.

Finally, meta-analyses examining the association between T2DM and the risk of each type of cancer using unadjusted effect estimates will be conducted, along with the corresponding subgroup analyses.

Background

This thesis aims to assess the independent causal effect of T2DM on cancer risk across multiple cancer types. The first step involved conducting an updated systematic review (SR) to gather evidence on the association between T2DM and cancer risk across various cancer types (Chapter 3). To estimate this causal relationship, DAGs were used to identify the MSAS for each cancer type in the studies selected for the SR (Chapter 4). Although many studies in the SR did not explicitly or implicitly claim a causal relationship between T2DM and cancer risk (Chapter 5), those that accounted for the MSAS in their adjusted analyses were selected for the adjusted (causal effect) meta-analyses presented in this chapter.

Meta-analyses were conducted to statistically synthesize quantitative data from multiple studies to generate a more precise estimate. Additionally, aggregating results and pooling data increase statistical power, which is the probability of

correctly detecting a true effect when it exists. This probability improves as the sample size grows (Higgins et al., 2019).

Another reason for conducting the meta-analyses was the diversity and conflicting results of individual studies identified in the SR (Higgins et al., 2019). For example, studies assessing the relationship between T2DM and prostate cancer risk that accounted for the MSAS reported varying conclusions.

Some found a decreased risk, such as (Bao et al., 2018) with an odds ratio (OR) of 0.48 (95% confidence interval [CI]: 0.35-0.67) and (Rodriguez et al., 2005) with a relative risk (RR) of 0.67 (95% CI: 0.60-0.75). Others, such as (Dickerman et al., 2018), reported an increased risk (hazard ratio [HR]: 1.61, 95% CI: 1.07-2.43), while some, including (De Bruijn et al., 2014) and (Tavani et al., 2005), found no association with T2DM. Therefore, conducting meta-analyses helped to formally assess the diverging findings of individual studies.

Four three systematic reviews and meta-analyses have been published on the relationship between T2DM and cancer risk across various cancer types. Three of them (Noto et al., 2011; Ohkuma et al., 2018; Starup-Linde et al., 2013) focused on assessing associations, whereas the other attempted to assess the strength of causal evidence using statistical methods (Ling et al., 2020).

In the publication aiming to infer causality, Ling and colleagues conducted bias analyses for unmeasured confounding (Ling et al., 2020). The aim of this analysis was to quantify how sensitive the results were to the presence of unmeasured confounding. This type of analysis assesses the robustness of causal associations based on a prespecified minimum threshold and quantifies the proportion of studies with a true (causal) effect (Mathur & VanderWeele, 2020). However, the authors did not make causal claims but used the bias analysis to assess the likelihood that the observed associations reflected causal relationships.

In contrast to these publications, the meta-analyses conducted in this chapter aim to provide causally interpretable estimates based on studies that accounted for confounders (MSAS). These confounders were carefully identified using DAGs,

which were developed through targeted literature searches, resulting in less biased causal estimates.

6.3 Methods

To identify all relevant evidence on the association between T2DM and cancer risk, a SR was conducted using five databases without language or time restrictions (Chapter 3). This SR followed the Population, Intervention, Comparison, Outcome, and Study design (PICOS) framework: men and women aged 18 years or older (P); individuals with T2DM (I) compared to those without T2DM (C); cancer incidence (O), expressed as risk ratio (RR), hazard ratio (HR), standardized incidence ratio (SIR), or odds ratio (OR); and observational studies (S).

The SR identified 250 studies, of which 149 were cohort and case-control studies with a low risk of bias, as measured using the Newcastle-Ottawa Scale (NOS ≥ 7 points), covering 56 types of cancer. A DAG was developed for each cancer type to identify the MSAS (Chapter 4).

Two sets of meta-analyses were conducted. The first aimed to estimate the [causal effect between T2DM and cancer risk](#), relying on effect estimates that accounted for the MSAS. These analyses are referred to as *adjusted (causal effect) meta-analyses* throughout this chapter, as they use confounder adjustment to provide causally interpretable estimates. Among the 149 studies, at least one such estimate was available for 31 cancer types. However, only 17 of these had more than one effect estimate accounting for the MSAS, allowing for meta-analysis. The remaining 14 cancer types had only a single effect estimate that considered the MSAS.

The second set of meta-analysis assessed the [association between T2DM and each type of cancer](#) using unadjusted effect estimates. These are referred to as *unadjusted (association) meta-analyses* throughout this chapter, as they describe crude associations without accounting for confounding. These estimates were either reported or could be derived for 41 cancer types. However, meta-analyses were feasible for only 31 of these cancer types, as the remaining 10 had only one effect estimate.

For both sets of meta-analyses, **subgroup analyses** were conducted by sex, race, and study design, including cohort and case-control studies. If sufficient data were available, additional subgroup analyses examined the effect of menopausal status on breast cancer, antidiabetic drug classes, and time since T2DM diagnosis on cancer incidence (refer to Section 6.3.4 for further details).

The rationale behind conducting two sets of meta-analyses was to address distinct research objectives. The first set aimed to estimate the independent causal effect between T2DM and cancer risk, which is the central research question of the thesis. By properly controlling for confounding variables, this analysis isolates the independent effect of the exposure, enabling a causal interpretation of the results.

The second set of meta-analyses assessed the association between T2DM and each type of cancer using unadjusted effect estimates. This analysis was designed to explore the magnitude and direction of the relationship between T2DM and cancer risk without accounting for confounders. It also serves to reinforce the importance of controlling for confounding when estimating causal claims, as a proportion of the observed associations may be explained by shared risk factors.

The data were processed following the guidelines outlined in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins et al., 2019). Data synthesis and analyses were performed using Review Manager software (RevMan, version 5.4.1 (The Cochrane Collaboration, 2022) and R software (version 4.2.2) (Schwarzer et al., 2015). Heterogeneity and publication bias in the meta-analyses were also assessed using funnel plots and Egger's test.

6.3.1 Meta-analysis software

RevMan is an open-source software (for academic use) developed by the Cochrane Collaboration Group for conducting Cochrane Reviews (The Cochrane Collaboration, 2022). The software facilitates all stages of systematic review processes, encompassing the analysis of study test characteristics and the

creation of comparison and risk of bias tables. It also enables quantitative analysis, including meta-analysis, within subgroups of studies and across all studies. Moreover, it automatically generates forest plots accompanied by heterogeneity tests and publication-standard funnel plots.

6.3.2 Data synthesis

To combine effect measures (RR, OR, HR, SIR) from the studies, a **random-effects model** was chosen over a fixed-effects model.

A fixed-effects model assumes that all effect sizes originate from a single homogeneous population, with variation in observed effect sizes attributed only to sampling error. This model accounts for a single source of variation: within-study variance (Harrer et al., 2021; Pathak et al., 2020).

In contrast, a random-effects model accounts for variation in true effect sizes across individual studies (between-study variance). This variation may arise from clinical and methodological differences, such as disparities in participant characteristics (e.g., differences in the target populations of the studies), variations in intervention, treatment, or exposure duration, differences in control groups, outcome definitions, and study designs. Therefore, the effects observed in individual studies are assumed to deviate not only due to sampling error but also due to additional variance. This additional variance component derives from the idea that studies represent a random sample of a larger population of similar studies rather than a single homogeneous population, as assumed in a fixed-effects model.

The random-effects model assumes that there is not a single true effect size but rather a distribution of true effect sizes, and it estimates the mean of this distribution (Harrer et al., 2021; Pathak et al., 2020).

As observational data are likely to exhibit some degree of between-study heterogeneity, a random-effects model was chosen to pool the effect estimates (Dekkers et al., 2019; Ohkuma et al., 2018; Starup-Linde et al., 2013). The **DerSimonian and Laird method** was then used to estimate between-study variance (τ^2) (Harrer et al., 2021).

To calculate the pooled effect, the **inverse-variance method** was applied to determine the weight of each study. In this approach, the weight assigned to each study is based on the inverse of the variance of its effect estimate (Higgins et al., 2019).

Since the random effects model accounts for two sources of variation, between-study variance (τ^2) and within-study variance (σ^2), the weight assigned to each study is the inverse of the total variance (Harrer et al., 2021; Pathak et al., 2020):

$$w'_i = \frac{1}{\sigma^2 + \tau^2}$$

(Equation 1)

The pooled effect estimate is obtained by summing the weighted individual effect estimates and dividing by the total sum of weights, as shown in the formula below (Harrer et al., 2021; Pathak et al., 2020):

$$\hat{\mu} = \frac{\sum y_i w'_i}{\sum w'_i}$$

(Equation 2)

where $\hat{\mu}$ is the pooled effect estimate, y_i represents the individual study effect size, and w'_i is the weight assigned to each study.

Larger studies are given higher weights because they provide more precise estimates of the true effect. Smaller studies, which have larger standard errors (wider confidence intervals), contribute less weight due to their lower precision (Harrer et al., 2021).

Since risk ratios do not follow a normal distribution, the effect estimates of individual studies and their 95% confidence intervals were converted to natural logarithms. This transformation ensures asymptotic normality and centers values around zero. Once transformed, the values were entered into RevMan under the "Generic inverse variance" outcome type (Higgins et al., 2019).

As previously mentioned, **two sets of meta-analyses were conducted. The first aimed to estimate the causal effect of T2DM on cancer risk, while the second assessed the association between T2DM and each type of cancer.**

For the first set, only effect estimates from reported adjusted analyses that included the MSAS identified a priori using the corresponding DAG for each type of cancer were selected. For the second set, which explored associations between T2DM and specific cancers, unadjusted estimates were calculated when not specified, based on available summary data for each group, including the number of events and sample sizes of the exposed and unexposed groups.

Equations 3 and 4 illustrate the calculations for RR and OR, respectively, used to derive these outcome measures when not reported, in order to conduct the association (unadjusted) meta-analyses (Gordis, 2014).

$$RR = \frac{\text{cancer incidence in T2DM group}}{\text{cancer incidence in non T2DM group}}$$

(Equation 3)

$$OR = \frac{(\text{number of cancer cases in T2DM group}) * (\text{number of non T2DM group without cancer})}{(\text{number of T2DM group without cancer}) * (\text{number of cancer cases in non T2DM group})}$$

(Equation 4)

Given that the occurrence of cancer resulting from T2DM is a rare event (<1%), both the OR and HR offer a close approximation of the RR (GREENLAND, 1987; VanderWeele, 2020; Viera, 2008). Similarly, the SIR is comparable to RR, as it represents the ratio of observed to expected cases in a given population (Marti et al., 2015). Therefore, all measures of effect were interpreted as RR.

The summary effect estimates (RR) from the meta-analyses were interpreted as follows for the **adjusted (causal effect) meta-analysis** results, with the non-exposed group (individuals without T2DM) in the denominator (Gordis, 2014):

- **RR > 1** indicates an increased risk of cancer in individuals with T2DM compared to those without T2DM, suggesting a **potential causal effect** of T2DM in increasing cancer incidence.
- **RR < 1** indicates a decreased risk of cancer in individuals with T2DM compared to those without T2DM, suggesting a **potential protective causal effect** of T2DM on cancer risk.
- **RR = 1** suggests no difference in cancer risk between individuals with and without T2DM, providing **no evidence of a causal effect**.

For the **unadjusted (association) meta-analyses**, the RR were interpreted as a measure of statistical association between T2DM and cancer risk (Gordis, 2014):

- **RR > 1** indicates that T2DM is associated with an **increased** risk of cancer.
- **RR < 1** indicates that T2DM is associated with a **decreased** risk of cancer.
- **RR = 1** indicates **no observable association** between T2DM and cancer risk.

6.3.3 Heterogeneity assessment

Between-study heterogeneity refers to the variation in true effect sizes across studies within a meta-analysis. This variability is accounted for in random-effects models, as previously described (Harrer et al., 2021).

To quantify heterogeneity in the meta-analyses, the I^2 test was used:

$$I^2 = \left(\frac{Q - df}{Q} \right) \times 100\%$$

(Equation 5)

where Q is the χ^2 statistic and df its degrees of freedom.

I^2 represents the proportion of variability in effect sizes due to true differences between studies rather than sampling error. I^2 values range from 0% to 100%,

where 0% indicates no observed heterogeneity, and higher percentages reflect increasing heterogeneity (Harrer et al., 2021; Higgins et al., 2019).

For example, when $I^2 = 0\%$, none of the variation in effect sizes is due to between-study heterogeneity. Instead, all observed differences are likely due to random chance rather than actual variations between studies; in other words, the studies are highly consistent.

If I^2 is approximately 50%, about half of the variation in effect sizes is due to between-study heterogeneity.

A very high I^2 (e.g., above 75%) suggests that a large proportion of variability in effect sizes is due to real differences between studies rather than random sampling error. This can occur when studies differ in participant demographics or methodologies, such as study design (e.g., cohort vs. case-control), sample size, or follow-up duration.

For guidance, the I^2 could be interpreted as follows (Higgins et al., 2019):

- 0% to 40%: might not be important
- 30-60%: represents moderate heterogeneity
- 50-90: represent substantial heterogeneity
- 75-100% considerable heterogeneity

For this research, meta-analyses with I^2 values <30%, 30-60%, 61-75%, and >75% were classified as showing **low**, **moderate**, **substantial**, and **considerable** heterogeneity, respectively (Y. Liu et al., 2018).

6.3.4 Subgroup analyses

Subgroup analyses were conducted to assess potential variations in effect estimates based on specific participant characteristics or study design (Higgins et al., 2019).

Based on the MSAS derived from DAGs, analyses were performed by **sex**, as this covariate was identified as a confounder in estimating the causal effect of T2DM on the risk of certain types of cancer, such as CRC, HCC, and pancreatic cancer, among others.

Analyses were also conducted by **geographic region** as a proxy for **race** since this variable was identified as a confounder for certain cancer types, including breast cancer. For example, White women have a higher likelihood of developing breast cancer than Black, Hispanic, and Asian women (Yedjou et al., 2019). Geographic region can serve as an indicator of racial and ethnic differences, as regional groupings reflect some degree of genetic and environmental similarity, despite considering diverse populations (Consumer Financial Protection Bureau, 2014). Race and ethnicity are closely linked to genetic ancestry but do not always align directly with geographic regions. However, since data on racial classifications were not available in the included studies, geographic classification was used as a proxy to capture potential variations. Therefore, subgroup analyses were conducted by geographic region based on the available data from the systematic review, which included studies from multiple countries that could be grouped by **continent**, specifically **America, Asia, and Europe**.

In addition, cohort and case-control subgroup analyses (**study -design**) were accomplished to ascertain whether there was a different true effect depending on the study design.

If enough data were available, subgroup analyses based on **menopausal status** for breast cancer were conducted. Evidence suggests that postmenopausal women with T2DM have a higher risk of breast cancer compared to non-diabetic women (RR 1.25; 95% CI, 1.20-1.29) (Liao et al., 2011).

Moreover, subgroup analyses exploring the causal effect between **time since T2DM diagnosis and cancer incidence** were intended. Epidemiological research suggests that this risk varies with the duration of the disease (T2DM duration 4.1-6 years, HR 1.37 [95% CI, 1.25 to 1.50] vs T2DM after 15 years, HR 1.04 [95% CI, 0.95 to 1.14]) (Y. Hu, Zhang, et al., 2021).

Furthermore, subgroup analyses by **antidiabetic drug classes** were planned to be undertaken. Existing evidence suggests that certain medications, such as pioglitazone, may be associated with an elevated risk of cancer (OR, 1.20; 95%

CI, 1.05-1.38), while others, such as metformin, have shown no association (OR 1.00; 95% CI 0.93-1.07) or even exhibit a protective effect (HR, 0.5; 95% CI 0.32-0.83)(H. J. Kim et al., 2018; Y.-C. Liu et al., 2019).

In order to evaluate such interactions, Review Manager software (RevMan version 5.4.1) was employed, as it facilitates subgroup analysis and the execution of Cochran's Q test. This formal test for interaction compares the differences in binary outcomes between two or more subgroups. The null hypothesis of this statistical test is that the effect is the same across all the subgroups, with any observed variations being attributable solely to chance and it was assumed that a p-value lower than 0.05 indicated a significant subgroup difference (P. Liu et al., 2019).

6.3.5 Publication bias assessment

Publication bias is a form of reporting bias that arises when studies are published selectively based on the direction or strength of their findings. Those studies with positive and statistically significant results are more likely to be submitted and published than those with negative or null findings (DeVito & Goldacre, 2019; Murad et al., 2018).

Several methods can be used to assess publication bias. One subjective approach involves a **funnel plot**, which was selected for this thesis. This method involves creating a simple scatter plot of each study's effect estimate on the horizontal axis against a measure of its size or precision on the vertical axis. Larger studies are expected to cluster around the true underlying effect size (upper part of the graph), while smaller studies are anticipated to disperse widely (bottom of the graph). A symmetric (inverted) funnel plot would indicate the absence of publication bias, whereas an asymmetric one may suggest the presence of such bias. This test should be used in a meta-analysis with at least **ten studies** to avoid the potential for misleading interpretation (Higgins et al., 2019).

In addition to conducting a visual inspection of the funnel plot, the **Egger test** was employed to quantify funnel plot asymmetry. This statistical test uses a simple regression model to explore the association between the effect estimates

and their standard errors (L. Lin et al., 2018). If the intercept significantly differs from zero ($p < 0.5$), this indicates that the funnel plot is asymmetrical, suggesting evidence of publication bias. The tests were performed with R software (version 4.2.2, package “meta”) (Schwarzer et al., 2015).

6.4 Results

As the aim of this thesis is to assess the independent causal relationship between T2DM and cancer, this section presents results from **the adjusted (causal effect) meta-analyses**, which evaluated the causal effect of T2DM on cancer risk. These analyses were conducted for each cancer type, using adjusted effect estimates from individual studies that incorporated the MSAS identified through DAGs (Chapter 4).

Subgroup analyses according to study design, sex, geographic region (proxy of race), menopausal status, T2DM duration, and antidiabetic drug usage are also provided.

Additionally, publication bias was assessed using funnel plots and Egger’s test.

Results from the **unadjusted (association) meta-analyses**, along with detailed subgroup analyses and publication bias assessments, are presented in Appendix C. At the end of this section, a summary comparing the findings from both sets of meta-analyses is provided.

6.4.1 Meta-analyses of adjusted (causal effect) estimates: Overall results

Data were available to conduct adjusted (causal effect) meta-analyses of T2DM on the risk of 17 types of cancer (at least two studies per cancer type). The results are presented in Table 6-1.

Table 6-1 Pooled adjusted RR (causal effect) estimates of T2DM on cancer risk, based on studies adjusting for MSAS.

Type of cancer	No. Studies	RR (95% CI)	I ²	Forest plot
Bladder	3	1.23 (0.87, 1.74)	71%	
Bone	3	1.17 (0.85, 1.62)	0%	
Brain/NS	3	1.03 (0.64, 1.66)	0%	
Breast	4	1.11 (1.04, 1.18)	0%	
Colon	2	1.30 (1.12, 1.50)	0%	
CRC	2	1.32 (1.13, 1.53)	0%	
Endometrium	5	1.94 (1.43, 2.63)	69%	
Eye	2	1.20 (0.97, 1.49)	0%	
Kidney	2	1.67 (1.40, 1.98)	0%	
Lung	4	1.00 (0.88, 1.13)	36%	
NHL	2	1.40 (0.74, 2.64)	89%	
Pancreas	8	1.73 (1.46, 2.05)	73%	
Prostate	11	0.81 (0.72, 0.91)	78%	
Rectum	2	1.21 (0.91, 1.61)	0%	
Salivary gland	2	0.92 (0.69, 1.24)	5%	
Testicle	2	1.29 (0.23, 7.09)	84%	
Thyroid	5	0.98 (0.76, 1.27)	49%	

Statistical significance ($p < 0.05$) for effect estimates is indicated in blue. I^2 indicates heterogeneity estimate. *Abbreviations.* Brain/NS: Brain/Nervous system; CI: confidence interval; CRC: colorectal cancer; NHL: non-Hodgkin lymphoma; RR: pooled relative risk.

As shown in Table 6-1, the point estimate was above one for all cancer types except prostate, salivary gland, and thyroid, based on the adjusted (causal effect) meta-analyses of T2DM and cancer risk. However, **statistically significant causal results** ($p < 0.05$) were found only for **breast** (RR, 1.11; 95% CI, 1.04-1.18), **colon** (RR, 1.30; 95% CI, 1.12-1.50), **CRC** (RR, 1.32; 95% CI, 1.13-1.53), **endometrial** (RR, 1.94; 95% CI, 1.43-2.63), **kidney** (RR, 1.67; 95% CI, 1.40-1.98), **pancreatic** (RR, 1.73; 95% CI, 1.46-2.05), and **prostate cancer** (RR, 0.81; 95% CI, 0.72-0.91).

Pooled adjusted estimates for more than half of the cancer types exhibited low to moderate heterogeneity, while six of the 17 cancer types (bladder, endometrial, NHL, pancreatic, prostate, and testicular) showed substantial ($I^2 = 50$ -90%) or considerable heterogeneity ($I^2 = 75$ -100%).

Among cancers with a statistically significant causal effect ($p < 0.05$) of T2DM, endometrial ($I^2 = 69\%$), pancreatic ($I^2 = 73\%$), and prostate cancer ($I^2 = 78\%$) exhibited substantial heterogeneity. In contrast, breast, colon, and kidney cancer, which also showed a statistically significant causal effect of T2DM, had no detectable heterogeneity ($I^2 = 0\%$), indicating consistency across studies.

Meta-analyses could not be conducted for 14 cancer types, as only a single effect estimate accounting for the MSAS was available for each (Table 6-2).

Table 6-2 Adjusted RR estimates of the causal effect of T2DM on cancer risk, based on a single study per cancer type, that accounted for MSAS variables

Type of cancer	Effect estimate (95% CI)	Forest plot
BTC	1.79 (0.88-3.62)	
Cervix uteri	0.62(0.08,4.72)	
Gallbladder	0.47(0.11,1.96)	
H&N	0.87 (0.73,1.04)	
HCC	5.68 (4.25,7.59)	
Larynx	1.15(0.25,5.20)	
Leukemia	0.31(0.07,1.26)	
Lymphoma	0.34(0.11,1.10)	
Melanoma	0.67(0.34,1.33)	
Non-melanoma	0.87(0.69,1.11)	
Oral	0.28(0.04,2.08)	
Ovary	1.30(0.46,3.71)	
Pharynx	10.55(2.95,37.67)	
Soft tissue	1.64 (0.62, 4.37)	

The study for H&N used a case-control design, while all others were cohort studies. The outcome measure was RR for BTC, HR for HCC and soft tissue, and OR for the remaining cancers. Statistical significance ($p < 0.05$) for effect estimates is indicated in blue. *Abbreviations.* BTC: Biliary tract cancer; CI: confidence interval; HCC: hepatocellular cancer; H&N: head and neck; HR: Hazard ratio; OR: Odds ratio; RR: relative risk.

The findings suggest that T2DM increases the risk of pharyngeal cancer (OR, 10.55; 95% CI, 2.95- 37.67) and may be causally associated with a reduced risk of H&N (HR, 0.87; 95%CI,0.73- 1.04). However, as these estimates are based on a single study, further research is needed to confirm these findings.

6.4.2 Meta-analyses of adjusted (causal effect) estimates: Subgroup analyses

6.4.2.1 Study design

Subgroup analyses by study design (cohort and case-control studies) were conducted for five types of cancer (Table 6-3).

Table 6-3 Subgroup analyses by study design (cohort and case control) from the adjusted (causal effect) meta-analyses of T2DM on cancer risk

Type of cancer	Effect estimate (95% CI)	p	I ²	p subgroup
Endometrium				
Cohort	1.52 (1.05, 2.21)	0.03	N/A	0.24
Case-control	2.07 (1.45, 2.95)	<0.0001	71%	
Pancreas				
Cohort	1.65 (1.39, 1.96)	<0.00001	67%	0.07
Case-control	2.16 (1.72, 2.72)	<0.00001	N/A	
Prostate				
Cohort	0.80 (0.70, 0.92)	0.001	82%	0.67
Case-control	0.85 (0.69, 1.04)	0.12	0%	
Thyroid				
Cohort	1.07 (0.80, 1.44)	0.64	50%	0.10
Case-control	0.70 (0.46, 1.07)	0.10	N/A	

Adjusted RR estimates in each subgroup. P-values in the third column indicate if the subgroup-specific effects are significant ($p < 0.05$). I², indicate heterogeneity in each subgroup; values are N/A for subgroups in which only 1 study was found, and therefore no subgroup I² value could be calculated. P subgroup represents the p-value for the test of subgroup differences (statistical significance [$p < 0.05$] is indicated in blue). Effect estimate: cohort studies (RR/OR) and case-control studies (OR). *Abbreviations.* CI: confidence interval; N/A: not applicable; OR: odds ratio; RR: relative risk.

The test of subgroup differences was not statistically significant ($p \geq 0.05$) for all types of cancer. Therefore, although the magnitude of the estimates varies between the two study designs, these differences are not statistically significant, suggesting that the results from both designs are comparable.

6.4.2.2 Sex

Subgroup analyses by sex were conducted for 16 types of cancer (Table 6-4). The test of subgroup differences was not statistically significant ($p \geq 0.05$) for all types of cancer. These results suggest that male and female appear to have comparable cancer risks in relation to T2DM.

Table 6-4 Subgroup analyses by sex from the adjusted (causal effect) meta-analyses of T2DM on cancer risk

Type of cancer	RR (95% CI)	p	I ²	p subgroup
Bladder				
Male	1.11 (0.61, 2.01)	0.73	N/A	0.26
Female	1.64 (1.17, 2.30)	0.004	0%	
Bone				
Male	1.12 (0.49, 2.57)	0.79	N/A	0.93
Female	1.19 (0.38, 3.75)	0.77	N/A	
Brain/NS				
Male	0.93 (0.45, 1.94)	0.85	0%	0.94
Female	0.97 (0.43, 2.19)	0.95	2%	
Colon				
Male	1.40 (1.07, 1.84)	0.02	N/A	0.61
Female	1.28 (1.03, 1.59)	0.02	31%	
CRC				
Male	1.42 (1.11, 1.81)	0.005	N/A	0.44
Female	1.26 (1.04, 1.52)	0.02	34%	
Gallbladder				
Male	0.53 (0.07, 4.14)	0.54	N/A	0.85
Female	0.40 (0.05, 3.01)	0.37	N/A	
H&N				
Male	0.81 (0.66, 0.99)	0.04	N/A	0.12
Female	1.13 (0.78, 1.63)	0.51	N/A	
Lung				
Male	0.83 (0.61, 1.12)	0.23	0%	0.52
Female	0.95 (0.72, 1.25)	0.70	0%	
Melanoma				
Male	1.06 (0.50, 2.23)	0.88	N/A	0.10
Female	0.18 (0.02, 1.29)	0.09	N/A	
NHL				
Male	1.39 (0.68, 2.86)	0.36	77%	0.91
Female	1.32 (0.75, 2.35)	0.34	79%	
Non-melanoma				
Male	0.90 (0.66, 1.23)	0.51	N/A	0.85
Female	0.86 (0.60, 1.24)	0.42	N/A	
Pancreas				
Male	1.14 (0.58, 2.21)	0.71	43%	0.41
Female	1.51 (1.33, 1.70)	<0.00001	0%	
Pharynx				
Male	15.64 (3.18, 76.99)	0.0007	N/A	0.41
Female	4.51 (0.38, 52.84)	0.23	N/A	
Rectum				
Male	1.48 (0.89, 2.47)	0.13	N/A	0.36
Female	1.11 (0.79, 1.57)	0.55	N/A	
Soft tissue				
Male	1.40 (0.26, 7.66)	0.7	N/A	0.93
Female	1.54 (0.43, 5.48)	0.5	N/A	
Thyroid				
Male	0.96 (0.59, 1.58)	0.88	0%	0.90
Female	0.91 (0.45, 1.84)	0.79	77%	

Adjusted RR estimates in each subgroup. P-values in the third column indicate if the subgroup-specific effects are significant ($p < 0.05$). I^2 , indicate heterogeneity in each subgroup; values are N/A for subgroups in which only 1 study was found, and therefore no subgroup I^2 value could be calculated. P subgroup represents the p-value for the test of subgroup differences (statistical significance [$p < 0.05$] is indicated in blue). *Abbreviations.* Brain/NS: Brain/Nervous system; CI: confidence interval; CRC: colorectal cancer; H&N: head and neck; NHL: non-Hodgkin lymphoma; N/A: not applicable; RR: relative risk.

6.4.2.3 Geographic region (as a proxy for race)

Subgroup analyses by geographic region (as a proxy for race), were conducted for 11 types of cancer (Table 6-5).

Table 6-5 Subgroup analyses by geographic region (America, Asia, Europe) from the adjusted (causal effect) meta-analyses of T2DM on cancer risk

Type of cancer	RR (95% CI)	p	I ²	p subgroup
Bladder				
America	1.27 (0.77, 2.10)	0.35	85%	0.78
Asia	NR	N/A	N/A	
Europe	1.15 (0.69, 1.92)	0.59	N/A	
Bone				
America	NR	N/A	N/A	0.93
Asia	1.14 (0.58, 2.24)	0.70	N/A	
Europe	1.18 (0.82, 1.70)	0.37	0%	
Brain/NS				
America	NR	N/A	N/A	0.93
Asia	1.04 (0.61, 1.78)	0.88	0%	
Europe	0.99 (0.35, 2.79)	0.98	N/A	
Breast				
America	1.10 (1.03, 1.18)	0.006	0%	0.61
Asia	1.14 (1.00, 1.31)	0.05	0%	
Europe	NR	N/A	N/A	
Endometrium				
America	1.78 (1.41, 2.23)	<0.00001	0%	0.60
Asia	NR	NR	N/A	
Europe	2.16 (1.06, 4.41)	0.03	90%	
Kidney				
America	1.60 (1.18, 2.17)	0.003	N/A	0.75
Asia	1.70 (1.38, 2.10)	<0.00001	N/A	
Europe	NR	N/A	N/A	
Lung				
America	NR	N/A	N/A	0.37
Asia	1.01 (0.88, 1.15)	0.92	46%	
Europe	0.77 (0.44, 1.36)	0.37	N/A	
NHL				
America	1.04 (0.96, 1.13)	0.35	N/A	0.002
Asia	2.00 (1.32, 3.03)	0.001	N/A	
Europe	NR	N/A	N/A	
Pancreas				
America	1.33 (1.04, 1.70)	0.02	N/A	0.08
Asia	1.59 (1.24, 2.03)	0.0002	70%	
Europe	1.97 (1.55, 2.49)	<0.00001	54%	
Prostate				
America	0.77 (0.66, 0.90)	0.0008	77%	0.18
Asia	0.89 (0.84, 0.94)	<0.0001	0%	
Europe	0.81 (0.62, 1.05)	0.11	77%	
Thyroid				
America	1.17 (0.70, 1.96)	0.55	85%	0.51
Asia	0.96 (0.71, 1.30)	0.78	57	
Europe	NR	N/A	N/A	

Adjusted RR estimates in each subgroup. P-values in the third column indicate if the subgroup-specific effects are significant ($p < 0.05$). I², indicate heterogeneity in each subgroup; values are N/A for subgroups in which only 1 study or no study (NR) was found, and therefore no subgroup I² value could be calculated. P subgroup represents the p-value for the test of subgroup differences (statistical significance [$p < 0.05$] is indicated in blue). *Abbreviations.* Brain/NS: Brain/Nervous system; CI: confidence interval; NHL: non-Hodgkin lymphoma; N/A: not applicable; NR: not reported; RR: relative risk.

The test of subgroup differences was statistically significant only for NHL ($p=0.002$), although these findings are based on data from a single study per region.

These results suggest that populations from different geographic regions appear to have comparable cancer risks in relation to T2DM.

6.4.2.4 Menopausal status

The subgroup analysis by menopausal status showed no statistically significant difference in the risk of developing breast cancer due to T2DM between premenopausal and postmenopausal women ($p = 0.35$). These results suggest that menopausal status does not significantly modify the causal effect of T2DM on breast cancer risk (Table 6-6).

Table 6-6 Subgroup analyses by menopausal status (premenopausal and postmenopausal) from the adjusted (causal effect) meta-analyses of T2DM on cancer risk

Menopausal status	RR (95% CI)	p	I ²	p subgroup
Premenopausal	0.95 (0.66, 1.35)	0.760	0%	0.35
Postmenopausal	1.12 (1.04, 1.21)	0.002	0%	

Adjusted RR estimates in each subgroup. P-values in the third column indicate if the subgroup-specific effects are significant ($p<0.05$). I², indicate heterogeneity in each subgroup. P subgroup represents the p-value for the test of subgroup differences (statistical significance [$p < 0.05$] is indicated in blue).

Abbreviations. CI: confidence interval; RR: relative risk.

6.4.2.5 T2DM duration

Evidence supporting a causal effect of T2DM duration on cancer incidence was limited and inconsistent due to discrepancies in the reported time periods across studies. Meta-analysis with two studies were only possible to conduct for bladder and thyroid cancer (Table 6-7).

Table 6-7 Subgroup analyses by T2DM duration (time [years] since diagnosis) from the adjusted (causal effect) meta-analyses of T2DM on cancer risk

Type of cancer	RR (95% CI)	p	I ²	p subgroup
Bladder				
≤ 5 years T2DM diagnosis	1.17 (0.60, 2.29)	0.65	74%	0.92
5-10 years T2DM diagnosis	1.38 (0.74, 2.59)	0.31	74%	
> 10 years T2DM diagnosis	1.20 (0.80, 1.78)	0.38	53%	
Thyroid				
< 5 years T2DM diagnosis	0.46 (0.29, 0.75)	0.002	32%	0.001
> 5 years T2DM diagnosis	1.11 (0.89, 1.39)	0.34	0%	

Adjusted RR estimates in each subgroup. P-values in the third column indicate if the subgroup-specific effects are significant ($p<0.05$). I², indicate heterogeneity in each subgroup. P subgroup represents the p-value for the test of subgroup differences (statistical significance [$p < 0.05$] is indicated in blue).

Abbreviations. CI: confidence interval; RR: relative risk.

The test of subgroup differences was statistically significant for thyroid cancer ($p=0.001$). Diabetics have a significantly decreased risk of developing thyroid cancer within 5 years after the diagnosis of T2DM (RR 0.46; 95% CI, 0.29, 0.75); beyond this period, the risk increases, although without statistical significance (RR 1.11; 95% CI, 0.89, 1.39). The results suggest that the risk of developing thyroid cancer is significantly influenced by the duration of T2DM.

6.4.2.6 Antidiabetic drug usage

The adjusted effect estimates (accounting for the MSAS) related to antidiabetic drug use on cancer incidence were rarely reported. Meta-analyses were only feasible for breast and pancreatic cancer, each based on two studies (Table 6-8).

Table 6-8 Subgroup analyses of antidiabetic drug usage from the adjusted (causal effect) meta-analyses of T2DM on cancer risk

Type of cancer	RR (95% CI)	p	I ²	p subgroup
Breast				
Use of antidiabetic drug				
Yes	1.23 (1.13, 1.35)	<0.00001	22%	0.16
No	0.99 (0.73, 1.33)	0.92	65%	
Pancreas				
Type of antidiabetic drug				
Metformin	1.02 (0.15, 6.77)	0.99	97%	0.52
Insulin	4.45 (0.81, 24.43)	0.09	90%	
Glitazones	2.20 (0.68, 7.08)	0.19	76%	

Adjusted RR estimates in each subgroup. P-values in the third column indicate if the subgroup-specific effects are significant ($p<0.05$). I², indicate heterogeneity in each subgroup. P subgroup represents the p-value for the test of subgroup differences (statistical significance [$p < 0.05$] is indicated in blue).

Abbreviations. CI: confidence interval; RR: relative risk.

For breast cancer, the studies reported estimated RRs comparing women with T2DM using an antidiabetic drug and women with T2DM not using an antidiabetic drug, with non-diabetic women as the reference group. The test of subgroup differences was not statistically significant ($p=0.16$), suggesting that the risk of developing breast cancer is not influenced by the use of antidiabetic drugs.

For pancreatic cancer, the studies reported estimated RRs comparing different types of antidiabetic drug use (metformin, insulin, glitazones) among individuals with T2DM, using non-diabetic individuals as the reference group. The test of subgroup differences was not statistically significant ($p=0.52$), suggesting that the risk of developing pancreatic cancer is not influenced by the type of antidiabetic drug used.

6.4.3 Publication bias assessment

Egger's test was used to assess potential publication bias in meta-analyses with at least ten studies via funnel plot asymmetry (Higgins et al., 2019). Among the different cancer types, only prostate cancer had a sufficient number of studies to conduct this test. The corresponding funnel plot, shown below, suggested no evidence of publication bias, and this finding was further supported by Egger's test ($p = 0.656$).

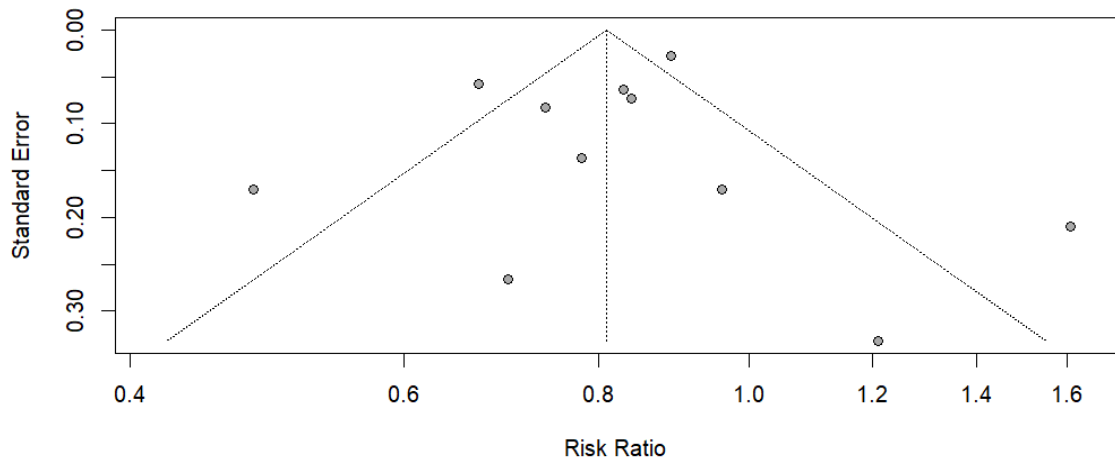


Figure 6-1 Funnel plot from the adjusted (causal effect) meta-analysis of T2DM and prostate cancer risk. Each point represents a separate study for the indicated causal effect. The funnel plot shows no evidence of publication bias. Egger's test for a regression intercept gave a p-value of 0.656, indicating no evidence of publication bias.

6.4.4 Adjusted (causal effect) versus unadjusted (association) meta-analyses

Table 6-9 presents a summary comparing findings from meta-analyses aimed at estimating the causal effect of T2DM on cancer risk (adjusted for MSAS) with findings from meta-analyses designed only to examine associations (using unadjusted estimates), for the 13 cancer types for which both sets of meta-analyses could be conducted.

For the cancer types for which both sets of meta-analyses could be conducted, the magnitude of associations was larger than the causal effects (except for rectal and kidney), suggesting that a proportion of the association was explained by common factors and highlighting the importance of adjustment for making causal inferences or estimation.

Table 6-9 Summary comparing findings from adjusted (causal effect) versus unadjusted (association) meta-analyses of T2DM and cancer risk.

Type of cancer	Causal effect		Association	
	RR (95% CI)	I ²	RR (95% CI)	I ²
Bladder	1.23 (0.87, 1.74)	71%	1.51 (1.15, 1.98)	99%
Brain/NS	1.03 (0.64, 1.66)	0%	1.07 (0.78, 1.45)	68%
Breast	1.11 (1.04, 1.18)	0%	1.20 (1.00, 1.44)	99%
Colon	1.30 (1.12, 1.50)	0%	1.51 (1.24, 1.85)	97%
CRC	1.32 (1.13, 1.53)	0%	1.58 (1.13, 2.20)	100%
Endometrium	1.94 (1.43, 2.63)	69%	2.37 (1.83, 3.08)	94%
Kidney	1.67 (1.40, 1.98)	0%	1.48 (1.08, 2.02)	97%
Lung	1.00 (0.88, 1.13)	36%	1.46 (1.02, 2.09)	99%
NHL	1.40 (0.74, 2.64)	89%	1.49 (1.18, 1.87)	89%
Pancreas	1.73 (1.46, 2.05)	73%	2.45 (1.97, 3.05)	97%
Prostate	0.81 (0.72, 0.91)	78%	1.09 (0.91, 1.29)	99%
Rectum	1.21 (0.91, 1.61)	0%	1.20 (1.04, 1.39)	76%
Thyroid	0.98 (0.76, 1.27)	49%	1.17 (0.96, 1.42)	70%

Statistical significance ($p < 0.05$) for effect estimates is indicated in blue. I², indicates heterogeneity estimate. *Abbreviations.* Brain/NS: Brain/Nervous system; CI: confidence interval; CRC: colorectal cancer; NHL: non-Hodgkin lymphoma; RR: pooled relative risk.

A higher heterogeneity (I²) in association meta-analyses compared to causal effect meta-analyses may also be due to confounding bias. This is because association meta-analyses were conducted using unadjusted effect estimates, resulting in greater variability. In contrast, causal effect meta-analyses accounted for confounders by using the MSAS, leading to more consistent estimates.

6.5 Discussion

The aim of this chapter was to provide causally interpretable estimates of the effect of T2DM on cancer risk. These analyses incorporated effect estimates exclusively from studies that controlled for MSAS, identified through DAGs in Chapter 4. By accounting for these confounders, the adjusted (causal effect) meta-analyses provided more accurate effect size estimates of the impact of T2DM on cancer risk.

Compared to the unadjusted (association) meta-analyses of T2DM on cancer risk, the magnitude of associations was consistently larger than the causal effects (except for rectal and kidney). This suggests that a proportion of the observed associations was explained by shared risk factors, reinforcing the importance of adjusting for confounding when the goal is to infer causality. However, if the

objective is solely to explore associations, confounder adjustment may not always be necessary. Therefore, the interpretation of research findings depends on the research objective, which should be clearly defined to ensure accurate conclusions, as discussed in Chapter 5.

The increased heterogeneity (I^2 statistic) observed in unadjusted (association) meta-analyses may reflect the influence of confounding bias. Because these analyses did not account for potential confounding variables, the effect estimates from individual studies varied more widely, leading to inflated heterogeneity. When at least two common confounders, such as age and sex, were accounted for, as in case-control studies, where controls were matched to cases based on these characteristics, heterogeneity levels decreased in meta-analyses of this study design compared to cohort studies. Appendix C (Table C1-3) presents the results of subgroup analyses by study design (cohort and case-control) on the association between T2DM and cancer risk. Moreover, Appendix A, related to systematic review findings (Table A7-3, column "Control group selection method"), provides details on the variables used to match control groups to cases.

Although heterogeneity was less pronounced in the adjusted (causal effect) meta-analyses, it remained considerable for prostate, testicular, and NHL cancers. Subgroup analyses were conducted to explore potential sources of heterogeneity based on available information, including study design, sex (when applicable), and geographic region (as a proxy for race). Among these analyses, a statistically significant difference was observed only for NHL; however, this finding was based on data from a single study per region, so further research and additional data are needed to confirm these findings.

Findings from Adjusted (Causal Effect) Meta-Analyses Compared with Published Systematic Reviews and Meta-Analyses

Among the four published SRs and meta-analyses on the relationship between T2DM and cancer risk, three focused on associations (Noto et al., 2011; Ohkuma et al., 2018; Starup-Linde et al., 2013), and one evaluated the strength of causal evidence in the observed associations using bias analysis (Ling et al., 2020).

Although the study by Starup-Linde et al. aimed to explore associations, its findings closely align with the adjusted (causal effect) meta-analyses presented in this chapter, except for pancreatic cancer. Specifically, they reported a RR of 2.21 (95% CI, 1.93-2.54), compared to a RR of 1.73 (95% CI, 1.46-2.05) in the adjusted (causal effect) meta-analyses. However, since their study did not aim to establish causality, the findings combined studies with varying levels of confounder adjustment, which may have contributed to the higher estimated risk for pancreatic cancer.

Unlike the objective of Starup and colleagues' publication, the study by Ling and colleagues aimed to evaluate the strength of causal evidence in the observed associations between T2DM and cancer incidence risk (Ling et al., 2020). To assess the likelihood that the observed associations reflected causal relationships, they used bias analyses for unmeasured confounding. Using this methodology, they identified a very likely causal association between T2DM and pancreatic and endometrial cancer, and a likely causal association for kidney, colon, and CRC. These findings were consistent with the summary effect estimates derived from the adjusted (causal effect) meta-analyses presented in this chapter.

For breast and prostate cancer, Ling et al. concluded that the association was more likely due to residual confounding rather than a direct causal effect of T2DM, as the proportion of studies supporting a causal association was below 70% under an unmeasured confounding assumption of a relative risk of 1.5. However, it is important to note that the pooled effect estimates from their meta-analyses for these cancer types closely aligned with the findings presented in this chapter. Specifically, for breast cancer, Ling et al. reported a RR of 1.10 (95% CI: 1.05-1.15), while this chapter reported an RR of 1.11 (95% CI: 1.04-1.18). For prostate cancer, Ling et al. reported an RR of 0.83 (95% CI: 0.79-0.88), while this chapter reported an RR of 0.81 (95% CI: 0.72-0.91).

They also performed sex-stratified analyses, and although tests for subgroup differences were not reported, cancer risk in relation to T2DM appeared comparable between males and females, as confidence intervals overlapped for all cancer types except kidney cancer. These findings align with the subgroup

analyses by sex conducted in this chapter, where no statistically significant subgroup differences were observed, suggesting similar cancer risks for males and females. However, due to limited evidence, a sex-specific subgroup analysis for kidney cancer was not feasible in this thesis.

The bias analysis method used by Ling and colleagues to infer that certain associations were likely causal does not require the identification of the MSAS in the relationship between T2DM and each type of cancer. For this reason, they considered only age, sex, and BMI as primary confounders, incorporating additional variables based on the study cohorts. As a result, the differences in findings between their publication and the adjusted (causal effect) meta-analyses presented in this chapter arise from differing objectives, which led to the use of distinct methodologies.

The aim of the bias analysis is to quantify the sensitivity of results to the presence of unmeasured confounding. These analyses depend on assumptions about the strength of unmeasured confounding (Mathur & VanderWeele, 2020). In the case of Ling and colleagues, they assumed a threshold of 1.5, meaning that an unmeasured confounder would be associated with both T2DM and cancer risk with a relative risk of 1.5, reflecting a moderate level of confounding. In contrast, the objective of this thesis is to establish causality, which involves identifying the MSAS in the association between T2DM and each type of cancer to address confounding. Therefore, both methodologies are useful, depending on the aim of the research.

Findings from Adjusted (Causal Effect) Meta-Analyses Compared to Published Mendelian Randomization Studies (MRS)

Mendelian Randomization Studies (MRS) are valuable for establishing causality from an observed association between an exposure and an outcome (Sekula et al., 2016). In the SR (Chapter 3), MRS were included to compare their findings with those from the adjusted (causal effect) meta-analyses conducted in this thesis. Consistent with the findings of this thesis, MRS further supported the causal effect of T2DM on the risk of kidney and pancreatic cancer.

For kidney cancer, the study by Yuan and colleagues reported a borderline increased risk of developing this cancer (OR 1.08; 95% CI 1.00, 1.17) (Yuan et al., 2020).

In the MRS examining the potential causal effect of T2DM on pancreatic cancer, statistically significant causal effect estimates were reported by (Lu et al., 2020) (OR 1.09; 95% CI, 1.01-1.17) and (Yuan et al., 2020) (OR 1.13; 95% CI, 1.04-1.22).

Conversely, other MRS that assessed the causal effect of T2DM on pancreatic cancer did not find strong evidence to support causality. These studies suggested that the statistically significant associations observed in traditional epidemiological studies may be explained by confounding and reverse causation (Carreras-Torres et al., 2017; Goto et al., 2020; Molina-Montes et al., 2021). They argued that unmeasured confounding from shared risk factors, such as lifestyle-related variables, may have distorted the effect estimates. Additionally, reverse causation, where cancer development precedes the exposure rather than the exposure causing cancer, was considered a plausible explanation (Mizuno et al., 2020). This was particularly relevant for pancreatic cancer, where T2DM may develop as an early manifestation of undiagnosed cancer rather than being a risk factor (Carreras-Torres et al., 2017).

However, it is important to emphasize that all the studies included in the current meta-analyses had a follow-up duration longer than three years. This criterion was established to assess the risk of bias using the Newcastle-Ottawa Scale (NOS) for all studies identified in the SR, not just for pancreatic cancer. A follow-up period of more than three years helps reduce the risk of reverse causation by limiting the influence of undiagnosed cancer on diabetes diagnosis.

Moreover, Goto and colleagues acknowledged that the null findings in their MRS regarding the causal association for pancreatic cancer, as well as for other types of cancer such as breast, CRC, and prostate cancer, were more likely due to insufficient statistical power to detect a significant association (Goto et al., 2020).

The present research also identified a statistically significant causal effect of T2DM on CRC. The point estimates were above 1 for both colon and rectal

cancer, yet statistical significance was observed only for colon cancer. One possible explanation for this difference is the larger surface area of the colon, which allows for increased absorption and prolonged exposure to cancer-causing molecules, compared to the rectum, which is shorter and may have less exposure to these substances (Canadian Cancer Society, 2023; Hua, 2019). Another explanation may be the larger number of cases in the CRC and colon cancer groups. Since statistical power increases with sample size, the lower number of rectal cancer cases could have led to wider confidence intervals, making it more difficult to detect a statistically significant causal association. This aligns with the conclusions reported by Murphy and colleagues, who acknowledged that the null findings in their MRS for CRC or its anatomical subregions (colon and rectum) may have been influenced by sample size limitations (Murphy et al., 2022).

Future Research

Subgroup analyses exploring the causal effect of time since T2DM diagnosis on cancer incidence were planned, as epidemiological research suggests that cancer risk varies with disease duration (Y. Hu, Zhang, et al., 2021). Nonetheless, due to limited evidence, these analyses were only feasible for bladder and thyroid cancer. For both cancer types, only two studies that included the MSAS reported findings for the same duration from T2DM diagnosis to cancer incidence. For bladder cancer, findings were available for three time periods based on T2DM duration (<5 years, 5-10 years, ≥ 10 years), and the test for subgroup differences showed no statistically significant variation in cancer risk across these time periods ($p > 0.05$).

In contrast, for thyroid cancer, findings were available for two time periods (<5 years, ≥ 5 years), and the test for subgroup differences showed a statistically significant variation in cancer risk ($p < 0.05$). Individuals with T2DM had a decreased risk of developing thyroid cancer within five years after diagnosis, but the risk increased beyond this period.

This protective effect of T2DM on thyroid cancer risk has been reported in other studies (Seo et al., 2017; Tseng, 2012). It has been proposed that cancer risk may be influenced by insulin levels. During the initial stages of T2DM, pancreatic

beta cells produce excess insulin, but as the disease progresses, insulin production decreases due to beta-cell dysfunction (Y. Hu, Wu, et al., 2021). However, this argument would suggest a higher risk in the early years following a T2DM diagnosis, which contrasts with the findings for thyroid cancer in this thesis. Therefore, additional evidence is needed to validate these findings.

Evidence suggest that diabetic treatment may influence the risk of developing certain types of cancer, such as CRC and pancreatic cancer (Bodmer et al., 2012; H. J. Kim et al., 2018; Y.-C. Liu et al., 2019). Studies have shown that metformin, the primary treatment for T2DM, can reduce insulin levels, which may, in turn, lower cancer risk (Y. Chen et al., 2022; Haq et al., 2024; Jo et al., 2023; H. Li et al., 2023; Luo et al., 2020; Najafi et al., 2023; Rennert et al., 2020; Tseng, 2012). However, subgroup analyses on the impact of antidiabetic drug use on cancer risk were not feasible due to limited evidence, except for the causal effect of T2DM on breast and pancreatic cancer risk. For these two cancer types, only two studies that included the MSAS reported findings related to antidiabetic drug use and cancer risk.

For breast cancer, the studies reported adjusted RRs comparing non-diabetic women (reference group) with women with T2DM using or not using antidiabetic drugs. Similarly, for pancreatic cancer, RRs were estimated for different antidiabetic drugs (metformin, insulin, glitazones), using non-diabetic individuals as the reference group. The test of subgroup differences was not statistically significant for either breast cancer ($p = 0.16$) or pancreatic cancer ($p = 0.52$), suggesting no influence of antidiabetic drug use or drug type on the risk of these cancers. Further research is needed to confirm these results.

Potential Biological Mechanisms between T2DM and Risk of Cancer

Several mechanisms have been proposed to explain the contribution of T2DM in cancer development, including insulin resistance, hyperinsulinemia, and hyperglycemia (Cosmin Stan & Paul, 2024; Pearson-Stuttard et al., 2021). However, further research is needed to fully understand the complex biological mechanisms involved. Below are some of the proposed mechanisms that may help explain the results of the meta-analyses.

Insulin resistance and hyperinsulinemia

- High levels of insulin in individuals with T2DM may increase cancer risk by modulating various growth factors. For example, hyperinsulinemia can suppress the production of a protein called insulin-like growth factor binding protein 1 (IGFBP-1), which normally binds to insulin-like growth factor 1 (IGF-1). As a result, more free IGF-1 becomes available in the body. IGF-1 is a key regulator of tumor cell proliferation (Gligorijević et al., 2022; Janssen, 2021).
- Insulin resistance and hyperinsulinemia can also modulate hormone levels. When the body becomes resistant to insulin, it compensates by producing more, which can reduce levels of sex hormone-binding globulin (SHBG). SHBG regulates the amount of oestrogen in circulation. Lower SHBG levels result in increased free oestrogen, which may promote the development of oestrogen-dependent cancers, such as breast and endometrial cancer (Arcidiacono et al., 2012; Johnson et al., 2012; Nead et al., 2015; Zhao et al., 2024).

Hyperglycemia

- Hyperglycemia is hypothesised to influence cancer risk by increasing the rate of glucose uptake by cancer cells, known as the Warburg effect. This high glucose requirement supports the survival and growth of tumours (Asiri et al., 2024).
- One proposed mechanism for the inverse association between T2DM and prostate cancer is that hyperglycemia may reduce testosterone levels. Lower testosterone levels have been associated with a reduced risk of prostate cancer, as testosterone can promote the growth of prostate cancer cells (Z. Wang et al., 2023).

Numerous studies have found that the risk of breast, endometrial, and colorectal cancers is higher shortly after a T2DM diagnosis and even during prediabetes, but the risk decreases over time. This suggests that hyperinsulinemia, rather than hyperglycemia, may be more strongly associated with cancer risk (Shlomai et al., 2016).

Conclusion

The aim of this chapter was to provide causally interpretable estimates of the effect of T2DM on cancer risk. To achieve this, adjusted (causal effect) meta-analyses were conducted, as these account for confounders and therefore provide estimates of the independent causal effect. In addition, unadjusted (association) meta-analyses were performed to assess the overall association between T2DM and each cancer type. These unadjusted analyses were not intended to provide causal estimates but rather to illustrate the importance of confounder adjustment by showing how effect sizes and heterogeneity can be inflated when confounding is not addressed. They also served to explore the magnitude and direction of the associations.

When comparing the two sets of results, the effect estimates derived from the unadjusted (association) meta-analyses were consistently larger than those from the adjusted (causal effect) meta-analyses, except for rectal and kidney cancer. This suggests that part of the observed associations was explained by shared risk factors. In terms of direction, both sets of meta-analyses generally indicated that T2DM increased the risk of cancer. However, the adjusted (causal effect) meta-analyses showed a statistically significant protective effect for prostate cancer, whereas the unadjusted (association) meta-analyses suggested an increased, though non-statistically significant ($p > 0.05$), risk. Statistically significant increased risks ($p < 0.05$) were also observed in the unadjusted (association) meta-analyses for bladder, lung, NHL, and rectal cancer, but these associations did not remain statistically significant after adjustment.

Differences were also observed in heterogeneity. The unadjusted (association) meta-analyses showed higher between-study heterogeneity (I^2), likely reflecting the influence of confounding bias. In contrast, the adjusted (causal effect) meta-analyses reduced variability and produced more consistent estimates, although heterogeneity remained considerable for prostate, testicular, and NHL cancers.

In conclusion, the unadjusted (association) meta-analyses were useful for exploring the direction and magnitude of the association between T2DM and cancer risk, but these estimates were affected by confounding. The adjusted (causal effect) meta-analyses, by accounting for shared risk factors, provided more reliable estimates of the independent causal effect of T2DM on cancer risk.

For the purpose of this thesis, the adjusted (causal effect) meta-analyses showed that T2DM has a causal effect on the increased risk of breast, colon, endometrial, kidney, and pancreatic cancers, while appearing protective against prostate cancer. Subgroup analyses found no statistically significant differences by sex, geographic region, study design, or menopausal status, indicating that the observed causal effects are consistent across these factors. These causal estimates will be used to quantify the future global burden of cancer attributable to T2DM, which is the focus of Chapter 7.

The following table provides an overview of the thesis structure. The information related to the next chapter is highlighted in yellow.

Table 6-10 Overview of the thesis structure (Next: Chapter 7)

Chapter 1: Introduction. To outline the background of the thesis, introduce the research question and aim, and provide an overview of the thesis structure.

Chapter	Objective	Methods	Key insights
2. Context of Evidence on T2DM and Cancer Risk	To provide an overview of published systematic reviews (SRs) and meta-analyses examining the relationship between T2DM and cancer risk across multiple cancer types	Summarize SR and meta-analyses examining the relationship between T2DM and cancer risk across multiple cancer types	Context for the research conducted in this thesis
3. Systematic review	To identify all relevant evidence on the association between T2DM and cancer risk across various cancer types.	Comprehensive Systematic literature review PICOS framework: men and women aged ≥ 18 years (P); individuals with T2DM (I) compared to those without T2DM (C); cancer incidence (O); observational studies (S).	Types of cancer with evidence on the association between T2DM and cancer risk. Confounders included in the adjusted analyses of the studies. Other*
4. Causal Association	To identify the variables that must be controlled for to estimate the causal relationship between T2DM and cancer across multiple cancer types.	Directed Acyclic Graphs (DAGs) were developed to identify the Minimal Sufficient Adjustment Sets (MSAS)	MSAS for each type of cancer
5. Assessing the quality of covariable adjustment in studies making causal claims	To assess whether studies making causal claims about T2DM and cancer clearly state their intent; justify their choice of covariates; and select them appropriately based on their objectives.	Studies from SR: Analysis of the terminology used across different sections to identify causal claims. Justification of variable selection using clinical evidence or terminology. Comparison of MSAS from DAGs with covariates in adjusted analyses	Emphasize the need for a clear study goal, aligned methods, and consistent terminology
6. Meta-analyses	To establish the causal effect of T2DM on cancer risk	Meta-analyses that account for the MSAS for each cancer type. Subgroup analyses.	Types of cancer causally associated with T2DM
7. Burden of disease	To quantify the future global burden of cancer causally associated with T2DM	Global burden of disease analysis, using the types of cancer causally associated with T2DM (meta-analyses)	Burden of disease metrics: PIF/PFP, the number of cancer cases, DALYs, and monetised DALYs

Cancer incidence (O), expressed as risk ratio (RR), hazard ratio (HR), standardized incidence ratio (SIR), or odds ratio (OR). Other*: Study design, follow-up, country-geographical publications, assessment method of exposure and outcome, Risk of bias assessment. Abbreviations: DALYs: Disability-adjusted life years; PICOS: Population, Intervention, Comparison, Outcome, and Study Design; PIF: population impact fraction; PFP: prevented fraction of the population

Chapter 8: Discussion. To summarize the main findings, outline strengths and limitations, provide policy recommendations, suggest directions for future research, highlight contributions to knowledge, and present final conclusions.

Chapter 7 Burden of disease

7.1 Summary

- By 2025, 2% of global cases of breast, colon, endometrial, kidney, and pancreatic cancers will be due to the effect of type 2 diabetes mellitus (T2DM), with the Middle East and North Africa (MENA) region projected to have the highest share of cancer cases due to T2DM.
- The Western Pacific is expected to experience the highest absolute number of cancer cases and disability-adjusted life years (DALYs) due to the effect of T2DM, primarily driven by the large population of countries such as China.
- Pancreatic cancer is projected to contribute the largest proportion of DALYs and the highest economic burden among cancers attributable to the effect of T2DM globally.
- Endometrial cancer is anticipated to represent the highest proportion of cancer cases attributable to the effect of T2DM in most regions, particularly in the MENA region, where over one in ten cases may be due to T2DM.
- Africa will experience the lowest burden of cancers attributable to T2DM in 2025 due to its low prevalence of diabetes among all regions. However, this trend will likely change as T2DM prevalence and related cancers are expected to rise sharply in the region by 2045.
- Addressing undiagnosed T2DM remains critical, particularly in regions like Africa, the Western Pacific, and Southeast Asia, where more than half of T2DM cases are undiagnosed. This problem affects early cancer detection efforts, with only half of the cancer cases causally associated with T2DM, being identified under current conditions.
- To maximize health and economic outcomes, interventions must be adapted to the specific priorities, needs, and contexts of each country, as T2DM prevalence, cancer burdens, and economic conditions vary both between and within regions.

7.2 Introduction

Objectives

The objective of this chapter is to quantify the future global burden of cancer causally associated with T2DM. Using the adjusted (causal effect) estimates for cancers causally associated with T2DM (Chapter 6), up-to-date prevalence data on diabetes, and considering the lowest level of diabetes prevalence achievable through a healthier lifestyle (i.e., individuals in the lowest risk category for modifiable risk factors), the findings will be reported as the proportion and number of incident cancer cases and DALYs attributable to or averted due to T2DM. Additionally, monetised DALYs will be calculated to assess the economic impact.

Background

Research undertaken for this thesis (Chapter 3 systematic review and Chapter 6 meta-analyses) has shown that T2DM is causally associated with the risk of breast, endometrial, colon, kidney, pancreatic, and prostate cancer.

Specifically, individuals with T2DM have a 30% higher risk of colon cancer, 67% for kidney cancer, and 73% for pancreatic cancer, beyond the risk associated with shared risk factors of T2DM and cancer. In women with T2DM, the risks of breast and endometrial cancer increase by 11% and 94%, respectively, while men with T2DM have a 19% lower risk of prostate cancer. Subgroup analyses confirmed that differences in cancer risks were not statistically significant across sexes or geographic regions.

T2DM is considered a public health problem. If current trends continue, it is estimated that the number of adults will reach 783 million by 2045 globally (International Diabetes Federation, 2021). Nevertheless, evidence suggests that adopting a healthier lifestyle can prevent over 80% of T2DM cases (Steinbrecher et al., 2011). Hence, if the incidence of T2DM is reduced, it is likely that the number of breast, colon, endometrial, and pancreatic cancers will also decrease in the future. For prostate cancer, however, the meta-analyses in Chapter 6 suggest that T2DM may lower risk. This means that if the incidence of T2DM is reduced, it is likely that the number of avoided prostate cancer cases will be less pronounced. Nevertheless, the overall health benefits of reducing diabetes

would lead to larger reductions in serious complications, such as cardiovascular diseases (e.g., coronary heart disease, myocardial infarction), which would ultimately lower mortality rates (De Rosa et al., 2018).

The Global Burden of Diseases, Injuries and Risk Factors (GBD) Study, recognized as the most comprehensive epidemiological effort worldwide since 1990, has been providing estimates of deaths, prevalence, years of life lost (YLL), years lived with disability (YLD), and DALYs attributable to a wide range of diseases, including diabetes (Mathers, 2020). In their latest publication, based on estimates from 2021, the authors aimed to report all non-fatal outcomes of diabetes (e.g. diabetic neuropathy, foot ulcers) but noted that their findings underestimated the disease's burden as they did not include its impact on chronic kidney disease or cancer (Ong et al., 2023).

The research conducted by Pearson and colleagues is the only study to date that has attempted to quantify the global cancer burden attributed to T2DM and high body mass index (BMI), both individually and in combination (Pearson-Stuttard et al., 2018). Their primary analysis focused on estimating the number of cancer cases attributable to these risk factors in 2012, assuming that 1% is the lowest achievable prevalence of T2DM. To calculate their estimates, they relied on the findings reported in an umbrella review and meta-analysis (Tsilidis, Kasimis, et al., 2015). However, the literature search for this umbrella review covered studies up to 2013, and the authors acknowledged that they did not adequately address confounding bias. Although Pearson's study provides valuable insights into the impact of T2DM on cancer burden, its findings may be influenced by confounding bias.

Therefore, in this chapter, the global burden of cancer causally associated with T2DM is quantified using relative risks (RR) derived from the adjusted (causal effect) meta-analyses (Chapter 6). The studies included in these meta-analyses were gathered from a comprehensive and updated systematic review (Chapter 3). Only those studies that accounted for the minimum sufficient adjustment set (MSAS), identified through Directed Acyclic Graphs (DAGs) (Chapter 4), were selected for inclusion in the meta-analyses.

In contrast, Pearson's analysis relied on a publication based on a literature search conducted over a decade ago. As discussed in Chapter 3, new evidence on this topic has been published annually, emphasizing the need for an updated search. This publication also failed to address confounding bias, meaning the effect of T2DM was not isolated, and it cannot be concluded that the findings were solely due to the exposure.

Additionally, to quantify the proportional reduction in cancer cases if T2DM exposure were reduced to a counterfactual scenario, Pearson assumed a minimum achievable T2DM prevalence of 1%. However, a more realistic approach was used in this thesis by incorporating the proportion of T2DM cases that cannot be prevented due to non-modifiable risk factors. Unlike Pearson's assumption, this data was supported by published evidence.

Finally, while Pearson and colleagues measured the burden of disease by estimating the proportion and number of cancer cases attributable to T2DM, the present research adopts a more comprehensive approach. In addition to these measures, this thesis evaluates the impact of T2DM-related cancer on quality of life by reporting DALYs. It also incorporates the economic impact by monetising DALYs, using gross domestic product (GDP) per capita to assign a monetary value to the DALYs attributable to T2DM-related cancer.

7.3 Methods

Since cancer develops gradually and risk factors take time to have an effect, there is usually a delay of several years between exposure to a risk factor and the onset of the disease. Research suggests that carcinogenesis may begin in the early stages of pre-diabetes through direct or indirect mechanisms involving hyperinsulinemia and insulin resistance (Huang et al., 2014). Evidence indicates that the risk of cancer is highest 4 to 8 years after a T2DM diagnosis (Y. Hu, Zhang, et al., 2021). Based on this, the present analysis assumes a **4-year lag**, estimating the cancer incidence in 2025 attributable to the prevalence of diabetes in 2021.

To measure the future global burden of cancer causally associated with T2DM, two key metrics were used: the potential impact fraction (PIF) and the prevented fraction of the population (PFP). These metrics were used to calculate the number of cancer cases and the number of cancer-related DALYs attributable or averted to T2DM. Additionally, DALYs were monetised based on GDP per capita to assess the economic impact.

The following sections briefly describe each metric, the formulas used, and the sources of reference. The analysis was conducted using Excel.

7.3.1 Measures of potential impact

To estimate the proportional reduction of a disease within a population than can be attributed to an exposure, the **potential impact fraction (PIF)** was used (Zapata-Diomedes et al., 2018).

The formula for the PIF is:

$$PIF = \frac{(P - P')(RR - 1)}{P(RR - 1) + 1}$$

(Equation 7)

where P is the prevalence of diabetes; P' is the counterfactual prevalence of diabetes (minimum exposure); RR is the relative risk of cancer causally associated with T2DM.

The PIF is an epidemiological measure that calculates the proportional change in disease incidence (cancer cases) following a change in exposure (T2DM). Calculating this change in the exposure requires both the current prevalence of T2DM and a counterfactual prevalence of the disease, which represents a hypothetical scenario where the prevalence is reduced to its minimum possible level, often referred to as the “theoretical minimum risk” (Zapata-Diomedes et al., 2018).

The findings presented in Chapter 6 from the adjusted (causal effect) meta-analyses showed that T2DM increases the risk of developing **breast, endometrial, colon, kidney, and pancreatic cancer**, with statistical significance. Therefore,

the PIF was used to estimate the potential reduction in the incidence of these cancers resulting from a decrease in T2DM prevalence.

However, when exposure acts as a protective factor, the **prevented fraction of the population (PFP)** is recommended instead of the PIF (Strain et al., 2020). The PFP is defined as the proportion of a disease outcome that has been averted or prevented due to the exposure of a protective factor (Strain 2020). For this reason, the PFP was used to calculate the proportion of **prostate cancer** cases that would be averted if T2DM prevalence were reduced to the theoretical minimum level (counterfactual scenario). The PFP is calculated as:

$$PFP = \frac{(P - P')(1 - RR)}{1 - (1 - RR)(1 - P)}$$

(Equation 8)

where P is the prevalence of diabetes; P' is the counterfactual prevalence of diabetes (minimum exposure); RR is the relative risk of prostate cancer associated with T2DM.

7.3.2 Number of cancer cases due to T2DM

To estimate the number of breast, endometrial, colon, kidney, and pancreatic cancer cases in 2025 attributable to T2DM prevalence in 2021, the following formula was used (Volesky 2019):

$$\text{number of cancer cases attributable} = PIF * Nx$$

(Equation 9)

where Nx represents the projected number of breast, endometrial, colon, kidney and pancreatic cancer incident cases in adults over 20 years of age for 2025, for each country and sex (except for breast and endometrial cancer, which are female-specific).

It is important to highlight that the PIF for each country was calculated by summing all cancer cases attributable to T2DM, across the different cancer types and dividing this sum by the total projected number of cases for these cancer types (Nx), as projected by GLOBOCAN for 2025.

To estimate the number of prostate cases that could have been averted if the prevalence of T2DM had been reduced to the counterfactual scenario, the following formula was used (Strain 2021):

$$\text{number of cancer cases averted} = \frac{Nx}{1 - PFP} - Nx$$

(Equation 10)

where Nx represents the projected number of prostate cancer incident cases in adults over 20 years of age for 2025, for each country.

7.3.3 Number Disability-Adjusted Life Years (DALYs) Due to T2DM

DALY is a measure used to assess the overall burden of a disease. It combines the years of life lost due to premature mortality (YLLs) and years of healthy life lost due to disability (YLDs). One DALY is equivalent to the loss of one year of full health (Bloom et al., 2011).

To estimate the number cancer-DALYs (breast, endometrial, colon, kidney and pancreatic) in 2025 attributable to T2DM prevalence in 2021, the following formula was used (Cui et al., 2016) :

$$\text{number of cancer DALYs attributable} = PIF * DALYx$$

(Equation 11)

where $DALYx$ represents the projected number of DALYs for breast, colon, endometrial, kidney, and pancreatic cancer in adults over 20 years of age for 2025, for each country and sex (except for breast and endometrial cancer, which are female-specific).

To estimate the number of prostate cancer DALYs that could have been averted if the prevalence of T2DM had been reduced to the counterfactual scenario, the following formula was used (Strain et al., 2020):

$$\text{number of cancer DALYs averted} = \frac{DALYx}{1 - PFP} - DALYx$$

(Equation 12)

where $DALYx$ represents the projected number of DALYs for prostate cancer in adults over 20 years of age for 2025, for each country.

7.3.4 Monetised DALYs (economic burden)

The gross domestic product (GDP) per capita was used to assign a monetary value to the DALYs of cancer-related to T2DM. GDP measures the total monetary value of goods and services produced within a country in one year. When divided by the population, it is referred to as GDP per capita, providing an estimate of the average economic performance per person (OECD, 2013).

In 2001, the WHO's Commission on Macroeconomics and Health recommended valuing DALYs between one- and three-times GDP per capita (Bloom et al., 2011). However, this threshold has since been debated. Daroudi and colleagues suggested that the threshold should vary according to a country's Human Development Index (HDI) (Daroudi et al., 2021).

The HDI, developed by the United Nations Development Programme (UNDP), is a composite measure that captures social and developmental aspects through three dimensions: health (life expectancy at birth), education (average and expected years of schooling) and standard of living (adjusted income through Gross national income [GNI] per capita)(United Nations Development Programme, 2024).

Based on HDI classification, Daroudi and colleagues proposed the following thresholds for the cost per DALY averted:

- **Low HDI countries: 0.34 times GDP per capita**
- **Medium HDI countries: 0.67 times GDP per capita**
- **High HDI countries: 1.22 times GDP per capita**
- **Very high HDI countries: 1.46 times GDP per capita**

While GDP per capita focuses on economic output, HDI incorporates health and education, providing a broader view of well-being. Adjusting GDP-based thresholds with HDI allows for more equitable comparisons across countries. Therefore, in this analysis, **monetised DALYs were calculated based on GDP per capita and adjusted for each country's HDI tier.**

7.3.5 Data sources

7.3.5.1 Prevalence of diabetes

The data source used for diabetes prevalence was the age-adjusted comparative **prevalence of diabetes (%) for individuals aged 20-79 in 2021**, as reported by the International Diabetes Federation (International Diabetes Federation, 2021). Estimates were available for 215 countries and territories. Since the data covers all types of diabetes, and type 2 accounts for 90% of cases, the prevalence estimates were multiplied by 0.9 for this analysis (International Diabetes Federation, 2021). As the estimates are not provided by sex, it was assumed that the prevalence rates are comparable between males and females, using the same data for both.

The IDF is a global non-profit organization founded in 1950, consisting of 240 national diabetes associations from 160 countries and territories (International Diabetes Federation, 2021). Since 2000, the IDF has provided reliable estimates based on strict methods for selecting and analysing high-quality data. These estimates primarily rely on population-based studies from peer-reviewed publications and national health surveys. To a lesser extent, national registries and reports from health regulatory bodies are also used. Moreover, for the latest edition (10th), data sources published between January 2019 and December 2020 were also screened. Sources rated as having the best quality are those with a nationally representative sample, conducted within the last five years, published in peer-reviewed journals, and using objective diagnostic methods rather than self-reports.

Counterfactual prevalence of diabetes

For many exposures or risk factors, such as smoking, researchers define the counterfactual scenario as zero or no exposure. However, for T2DM, achieving a zero prevalence is not feasible due to non-modifiable risk factors, such as age, sex, genetics, and family history, which cannot be eliminated. As a result, some level of T2DM prevalence will always remain.

Therefore, the counterfactual scenario for the present analysis was based on the **proportion of T2DM cases that could not be prevented due to T2DM non-modifiable risk factors**. To calculate this, data from the cohort study by (Steinbrecher et al., 2011) were used. The aim of this study was to estimate the proportion of T2DM cases that could be prevented if all the individuals were in the lowest-risk category for modifiable risk factors. This category was defined as: BMI <25 kg/m² (normal weight), physical activity > 4 hours per week, diet (lowest quintile of processed red meat intake and highest quintile of dietary fibre intake [expressed as g/4184 kJ per day], alcohol intake (non-drinker [<0.4 g/day, equivalent to one drink per month]), smoking status (non-smoker).

The authors found that 78% (95% CI: 63-87%) of T2DM cases in men and 83% (95% CI: 69-91%) in women could be prevented if individuals maintained the lowest-risk category for modifiable lifestyle factors. These results suggest that the proportion of preventable T2DM cases through healthy lifestyle changes is comparable between men and women.

Steinbrecher et al.'s (2011) findings indicate that 22% of T2DM cases in men and 17% in women could not be avoided, even if individuals adopted these optimal lifestyle behaviours. In other words, in the worst-case scenario, **one in five T2DM cases (22 out of 100) would remain unavoidable due to non-modifiable risk factors**.

To calculate the minimum T2DM prevalence for the counterfactual scenario, the 2007 prevalence in the United States of America (USA) was used, as it aligns with the population and follow-up period of the Steinbrecher study. According to the IDF, 9.2% of the USA population had diabetes that year, with 8.3% diagnosed with T2DM, as this type accounts for 90% of all diabetes cases (International Diabetes Federation, 2021).

For the counterfactual scenario, the 22% of T2DM cases that cannot be avoided was multiplied by the 8.3% T2DM prevalence, resulting in a **1.8% minimum prevalence for T2DM**.

7.3.5.2 Cancer cases

The projected number of cancer **incident cases for adults over 20 years in 2025** was consulted in Globocan (Ferlay J, Laversanne M, Ervik M, n.d.). For colon, kidney and pancreatic cancer, the number of cases were extracted by sex and country. In contrast, data for breast, endometrial, and prostate cancers were retrieved only by country, as these are sex-specific.

Globocan is an online database launched in 2013 and managed by the International Agency for Research on Cancer (IARC). It provides global cancer statistics (incidence, mortality, prevalence) for 36 types of cancers across 185 countries. Projections through 2050 are based on estimates from 2022, which rely on data from population-based cancer registries (via the International Association of Cancer Registries), the World Health Organization, and peer-reviewed studies (Ferlay J, Laversanne M, Ervik M, n.d.).

7.3.5.3 DALYs

The DALYs used in this analysis were obtained from the Global Burden of Disease (GBD) Compare tool, a visualization platform developed by the Institute for Health Metrics and Evaluation (IHME). This tool is based on data from the GBD study, a large-scale collaborative project involving researchers from over 160 countries. The data are gathered from various sources, such as hospitals, government agencies, surveys, and databases (Institute for Health Metrics and Evaluation (IHME), 2019).

The GBD Compare tool provides multiple metrics to visualize the burden of diseases, such as YLLs (years of life lost), YLDs (years lived with disability), prevalence, incidence, mortality rates, the maternal mortality ratio, probability of death, life expectancy, and HALE (healthy life expectancy).

For this thesis, the DALYs of cancers causally associated with T2DM were extracted from this database. DALYs are calculated by combining YLLs due to premature mortality and YLDs due to disability caused by these cancers. **The DALYs were expressed as rates per 100,000 adults aged 20 years and older** and were available for 204 countries in 2019. Data were disaggregated by sex, except for sex-specific cancers (breast, endometrial, and prostate cancer).

The present analysis used five-year age group rates (e.g., 20-24, 25-29, up to 80+) to account for population structure. These rates were then applied to the 2025 world population projections, which were also grouped into five-year age categories and stratified by age group, sex, and country, to improve the accuracy of the estimated DALYs (United Nations, Department of Economic and Social Affairs, 2022).

7.3.5.4 Monetised DALYs

The projected GDP per capita for 2025 for each country was sourced from the World Economic Outlook Database, published by the International Monetary Fund (IMF). Founded in 1944, the IMF is a global organization with 190 member countries that provide (*World Economic Outlook, October 2023, 2023*) *World Economic Outlook, October 2023, 2023*).

For the present analysis, the purchasing power parity (PPP)-adjusted GDP per capita, measured in constant 2017 international dollars (INT \$), was used from this database. PPPs adjust for price differences between countries, allowing for a fairer comparison of purchasing power by showing how much money is needed in one country to buy the same goods and services as in another (OECD, 2013).

The classification of each country HDI tier was obtained from the United Nations Development Programme website (United Nations Development Programme, 2024).

7.3.5.5 Relative risks

The relative risks (RRs) derived from the **adjusted (causal effect) meta-analyses** (Chapter 6), which accounted for the MSAS identified through DAGs (Chapter 4), showed that T2DM has an independent causal effect on cancer risk.

The RRs for the cancers found to be causally associated with T2DM were: **breast** (RR, 1.11; 95% CI, 1.04-1.18), **colon** (RR, 1.30; 95% CI, 1.12-1.50), **endometrial** (RR, 1.94; 95% CI, 1.43-2.63), **kidney** (RR, 1.67; 95% CI, 1.40-1.98), **pancreatic** (RR, 1.73; 95% CI, 1.46-2.05), and **prostate** cancer (RR, 0.81; 95% CI, 0.72-0.91).

7.3.6 Further analysis: Number of Cancer Cases Attributable to Confirmed T2DM

Early diagnosis of T2DM can help prevent future complications, but nearly 50% of people with diabetes are unaware of their condition (Ogurtsova et al., 2022). Undiagnosed diabetes refers to individuals with blood glucose levels consistent with diabetes (i.e., fasting plasma glucose ≥ 126 mg/dL [7.0 mmol/L], or 2-hour plasma glucose ≥ 200 mg/dL (11.1 mmol/L) after a 75 g oral glucose tolerance test, or glycated haemoglobin [HbA1c] $\geq 6.5\%$ [48 mmol/mol]) but without a formal medical diagnosis (American Diabetes Association, 2024; International Diabetes Federation, 2021).

To provide a more realistic estimate of the burden of cancer due to T2DM, the number of cancer cases due to confirmed T2DM was calculated by subtracting the proportion of undiagnosed diabetes from the total diabetes prevalence used in the main analysis (International Diabetes Federation, 2021).

The IDF reports that estimates of undiagnosed diabetes prevalence are derived from population studies. In these studies, participants who self-report as not having diabetes are tested for the condition. Those who test positive are classified as undiagnosed cases. This method provides an estimate of the overall prevalence of diabetes and the proportion of undiagnosed cases.

7.4 Results

The IDF provided country-level diabetes prevalence data, while the GBD Compare tool provided DALY data. By merging these sources, diabetes prevalence and DALY data from 179 countries were obtained and used for the present analysis.

However, the reported T2DM prevalence for Benin and Gambia fell below the minimum counterfactual prevalence considered, resulting in negative values for the burden of disease, an outcome that is not feasible. According to the IDF, this may be due to limitations in the quality of T2DM prevalence data for these countries. Consequently, these countries were excluded from the analysis, as also recommended by (Strain et al., 2020), reducing the dataset to **177 countries**.

Countries were then grouped into seven regions based on the IDF classification: Africa, Europe, the Middle East and North Africa (MENA), North America and the Caribbean, South and Central America, South-East Asia, and the Western Pacific.

Additionally, **countries were categorized into four groups according to their HDI to estimate the monetary value of DALYs**. Monetised DALYs could not be calculated for five countries (Cuba, the Democratic People's Republic of Korea, Guam, Somalia, and the Syrian Arab Republic) due to unavailable data on GDP projections and HDI classification.

The following sections present global results and detailed findings stratified by sex, region, and cancer type for each metric: the number of cancer cases, DALYs, and monetised DALYs. Section 7.8 summarizes key findings and provides recommendations for each region.

Appendix D presents the burden of disease analysis findings by country.

7.4.1 Cancers attributable to T2DM

This section focuses on cancers for which T2DM is causally associated with an increased risk, namely **breast, endometrial, colon, kidney, and pancreatic cancers**.

7.4.1.1 Overall findings

Table 7-1 presents the PIF, the estimated number of cancer cases attributable to T2DM, the estimated DALYs lost due to T2DM, and the corresponding monetised value of these DALYs for the year 2025.

Results indicate that **105,084 incident cancer cases in 2025 are expected to be attributable to T2DM**. These cases will represent **2% of the total number of breast, endometrial, colon, kidney, and pancreatic cancer cases** (0.5% of all 21,572,577 types of cancer cases projected by GLOBOCAN).

Table 7-1 PIF, projected number of cancer cases, DALYs loss and economic value of DALYs loss (monetised DALYs) attributable to T2DM in 2025

Type of cancer	Total number of cancer cases	PIF	Estimated number of attributable cancer cases	Total number of DALYs loss	Estimated number of DALYs loss	Estimated monetised DALYs loss in millions (INT. \$)
Breast	2,436,025	0.7%	17,091	23,149,056	187,941	\$4,600.05
Colon	1,278,694	1.8%	22,948	27,761,919	531,007	\$19,295.36
Endometrial	447,447	5.4%	24,072	2,624,460	152,645	\$5,068.35
Kidney	453,450	3.8%	17,444	4,147,803	178,867	\$6,600.41
Pancreatic	557,317	4.2%	23,530	12,890,749	566,259	\$21,844.70
Total	5,172,933	2.0%	105,084	9,693,896	1,616,719	\$57,408.87

The "Total number of DALYs loss" represents the overall burden of disease in terms of DALYs lost due to cancer, without considering the contribution of T2DM. The "Estimated number of DALYs loss" refers to the portion of DALYs lost that can be specifically attributed to T2DM, based on the estimated population attributable fraction (PIF). The total number of cancer cases are based on Globocan projections. The total number of DALYs lost are derived from the GBD Compare tool and World Population Prospects, as reported by the United Nations. INT \$: international dollars at constant 2017 prices.

Among the different cancers, **endometrial cancer will be the largest contributor** to the burden (PIF 5.4%), with 24,072 attributable cases expected. This is driven by the high relative risk (RR) of endometrial cancer associated with T2DM (RR: 1.94, 95% CI: 1.43-1.63) and the fact that it ranks as the 8th most common cancer worldwide, with 447,447 new cases projected in 2025.

The distribution of T2DM-attributable cases will be relatively even across cancer types. Endometrial cancer will account for 22.9% (24,072 cases), followed by pancreatic cancer (22.4%, 23,530 cases), colon cancer (21.8%, 22,948 cases), kidney cancer (16.6%, 17,444 cases), and breast cancer (16.3%, 17,091 cases).

The 105,084 cancer cases attributable to T2DM, are expected to result in the **loss of 1.6 million DALYs**, with an estimated **economic burden valued at INT \$57,408.87 million**. While the distribution of cases is relatively even across cancer types, pancreatic and colon cancers will together account for nearly 70% of the total DALYs lost, 566,259 DALYs (INT \$21,844.70 million) and 531,007 DALYs (INT \$19,295.36 million), respectively. However, the burden from colon

cancer may be overestimated since GBD Compare tool reports DALYs for colorectal cancer (CRC) as a whole, including rectal cancer.

The differences between the number of cancer cases and DALYs attributed to T2DM arise from variations in how each outcome is distributed. For instance, breast cancer is expected to have a much higher incidence (2,436,025 cases) than pancreatic cancer (557,317 cases) by 2025, with a ratio of roughly 4:1. However, due to pancreatic cancer's poorer prognosis, this relationship does not hold when considering DALYs, with 23,149,056 DALYs attributed to breast cancer compared to 12,890,749 for pancreatic cancer. As a result, even though the same PIF is applied, the number of cases and DALYs derived for each cancer type differ.

7.4.1.2 By Sex

Figure 7-1 illustrates the projected number of cancer cases and cancer-related DALYs by sex and cancer type.

Of the 105,084 cases attributable to T2DM, 35,917 (24%) will be expected to occur in men and 69,167 (76%) in women. Despite the larger number of cases in women, the overall burden (PIF) will be higher in men (2.9%) compared to women (1.8%). This is because the high number of breast cancer cases projected for 2025 (nearly 2.5 million) greatly increases the total cancer cases in women. However, since T2DM slightly increases the risk of breast cancer, the number of cases causally associated with T2DM is small. As a result, T2DM has a smaller overall burden of cancer due to T2DM (PIF) in women compared to men.

In men, pancreatic cancer will constitute the main burden among the different types of cancer (PIF: 4.3%, 12,700 cases), followed by kidney cancer (PIF: 3.9%, 11,092 cases) and colon cancer (PIF: 1.8%, 12,126 cases).

In women, endometrial cancer will represent the highest burden (PIF: 5.4%, 24,072 cases), followed by pancreatic cancer (PIF: 4.1%, 10,830 cases). Breast cancer will contribute the smallest burden (PIF: 0.7%, 17,091 cases).

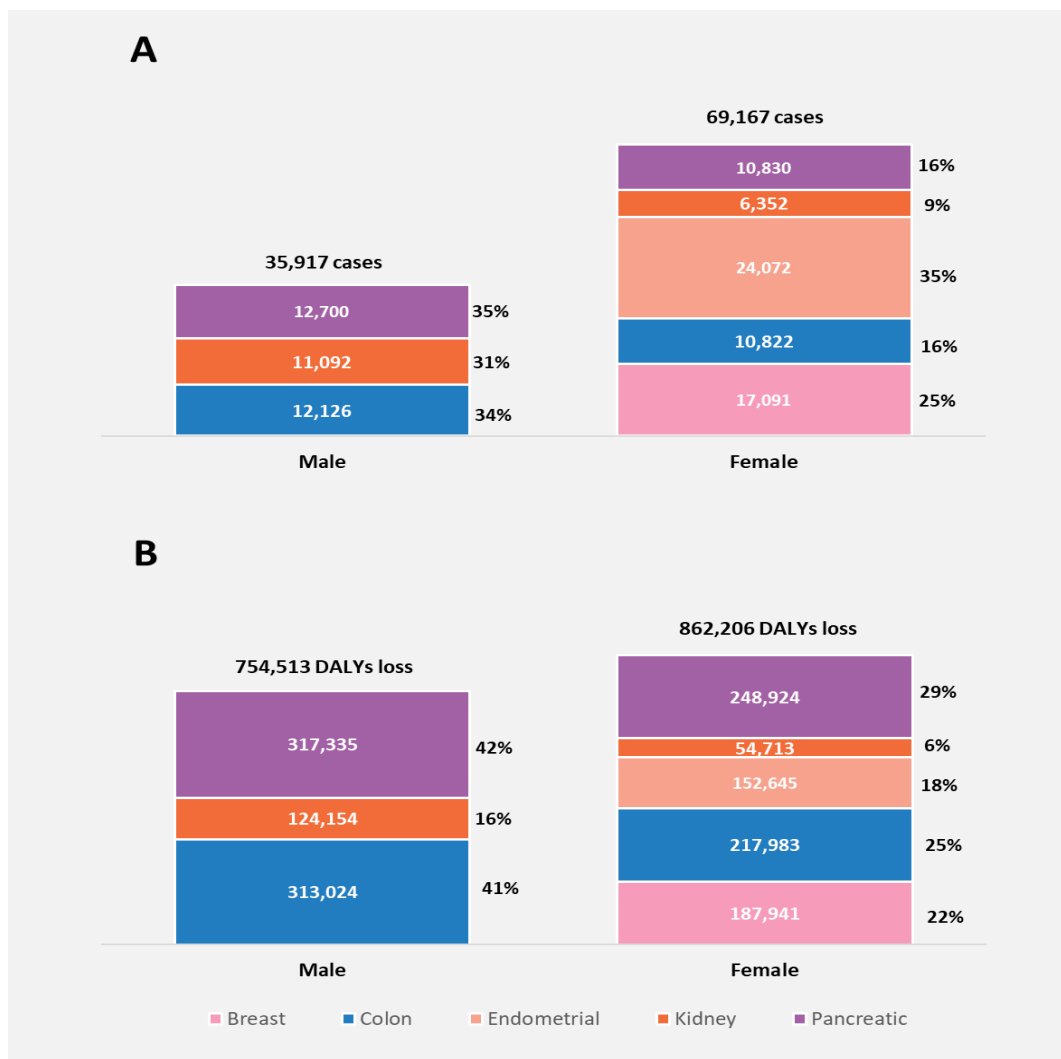


Figure 7-1 A) Projected number of cancer cases by sex and cancer type, with percentages indicating the share of each cancer type relative to the total cases within each sex. **B) Projected cancer-related DALYs by sex and cancer type**, with percentages showing the contribution of each cancer type to the total DALYs within each sex.

In women, endometrial cancer will represent 35% of T2DM-attributable cancer cases (24,072 cases), followed by breast cancer at 25% (17,091 cases). In both men and women, kidney cancer will account for the lowest proportion (31% [11,092 cases] in men and 9% [6,352 cases] in women).

Pancreatic cancer will represent the largest DALY burden for both sexes, with 317,335 DALYs (42% of total male cancer DALYs) in men and 248,924 DALYs (29% of total female cancer DALYs) in women. In economic terms, pancreatic cancer will have the highest economic impact for both sexes, amounting to INT \$12,275.29 million for men and INT \$9,569.40 million for women. These findings indicate that, although pancreatic and colon cancers account for a smaller share of total cases, they contribute to a greater extent to DALY loss and economic burden.

7.4.1.3 By Region and Country level

Globally, the countries with the highest PIF (5.7%) will be Kuwait and Pakistan, with 106 and 2,469 expected cancer cases attributable to T2DM, respectively. In contrast, Guinea-Bissau and Sierra Leone will have the lowest PIF (< 0.01%), translating to fewer than one attributable case in each. Figure 7-3 presents the PIF of all cancer incidence by country for 2025.

Table 7-2 provides an overview of the cancer burden attributed to T2DM by region. Across the seven global regions, the burden of cancer attributable to T2DM will vary, ranging from 0.7% in Africa to 3.1% in the MENA, consistent with the International Diabetes Federation (IDF) data showing MENA as the region with the highest global diabetes prevalence (16.2%).

Table 7-2 PIF, expected number of attributable cancer cases, DALYs loss and economic value of DALYS loss (monetised DALYs) attributable to T2DM in 2025 by region

Region	Total number of cancer cases	PIF	Estimated number of attributable cancer cases	Total number of DALYs loss	Estimated number of DALYs loss	Estimated monetised DALYs loss in millions (INT. \$)
Africa	208,354	0.7%	1,554	4,251,031	36,969	\$198.58
Europe	1,434,766	1.5%	21,796	15,735,784	256,949	\$14,572.51
Middle East and North Africa	236,066	3.1%	7,325	5,184,461	217,217	\$2,749.80
North America and Caribbean	728,039	2.8%	20,442	7,973,847	243,875	\$19,893.50
South and Central America	383,686	1.7%	6,555	4,955,073	93,441	\$1,932.84
South-East Asia	322,644	1.8%	5,663	7,854,952	164,762	\$964.47
Western Pacific	1,859,378	2.2%	41,750	24,618,839	603,505	\$17,097.16

The "Total number of DALYs loss" represents the overall burden of disease in terms of DALYs lost due to cancer, without considering the contribution of T2DM. The "Estimated number of DALYs loss" refers to the portion of DALYs lost that can be specifically attributed to T2DM, based on the estimated population attributable fraction (PIF). INT \$: international dollars at constant 2017 prices.

In Africa, Tanzania will have the highest PIF (2.73%), followed by South Africa (2.18%) and Zambia (2.08%). Turkey will lead Europe with a PIF of 4.03%, while Kuwait (5.74%) and Pakistan (5.69%) will top the MENA region. In North America

and the Caribbean, Mexico's PIF (4.17%) will surpass that of the USA (2.78%) due to Mexico's higher diabetes prevalence (16.9% vs. 10.7%). Puerto Rico will report the highest PIF in South and Central America (3.13%), Mauritius will lead Southeast Asia (4.87%), and Guam will have the highest PIF in the Western Pacific (5.27%).

Countries with the highest PIF are likely to differ from those with the most cancer cases attributable to T2DM, as the number of attributable cases depends on the total cancer incidence projected by GLOBOCAN. Figure 7-2 illustrates the projected number of cancer cases and cancer-related DALYs by region and cancer type, while Figure 7-4 and Figure 7-5 display these outcomes at the country level.

Of the total number of cancer cases expected to be attributable to T2DM, Africa will account for the lowest proportion, with 1,554 cases (1.5%). In contrast, nearly 39.7% (41,750 cases) of the global cancer burden is expected to occur in the Western Pacific region, with China contributing 25,517 cases (71% of the region's total). Europe is expected to report 21,796 T2DM-related cancer cases, roughly half of the Western Pacific's burden, while North America and the Caribbean will have a similar count (20,442 cases). However, the disease burden will be higher in North America and the Caribbean, with a PIF of 2.8%, compared to Europe's 1.5%. The Russian Federation will account for 12% of Europe's T2DM-related cancer burden, with 2,619 cases. Meanwhile, 79% of North America and the Caribbean's cancer cases (16,230) will occur in the USA.

In terms of DALYs, Africa will report the lowest loss, with 36,969 DALYs equivalent to an economic value of INT \$198.58 million. In contrast, the Western Pacific region will have the largest DALY burden, primarily driven by China, with 412,923 DALYs lost, representing INT \$10.47 billion. Figure 7-6 presents the projected monetised DALY Loss by country for 2025.

China's DALY loss will account for 68% of the region's total and 26% of the global DALYs attributable to T2DM-related cancers. Furthermore, China's economic loss will contribute 61% of the regional and 26% of the global economic burden associated with these DALYs.

The USA will rank second in DALY burden, with 184,816 DALYs (75.8% of the region's total and 11.4% globally), equivalent to INT \$17.88 billion. Although China will report more DALYs than the USA the higher GDP per capita in the USA (INT \$66,257.75 vs. INT \$20,782.64) results in the USA will account for a larger share of the global economic burden, at 31%.

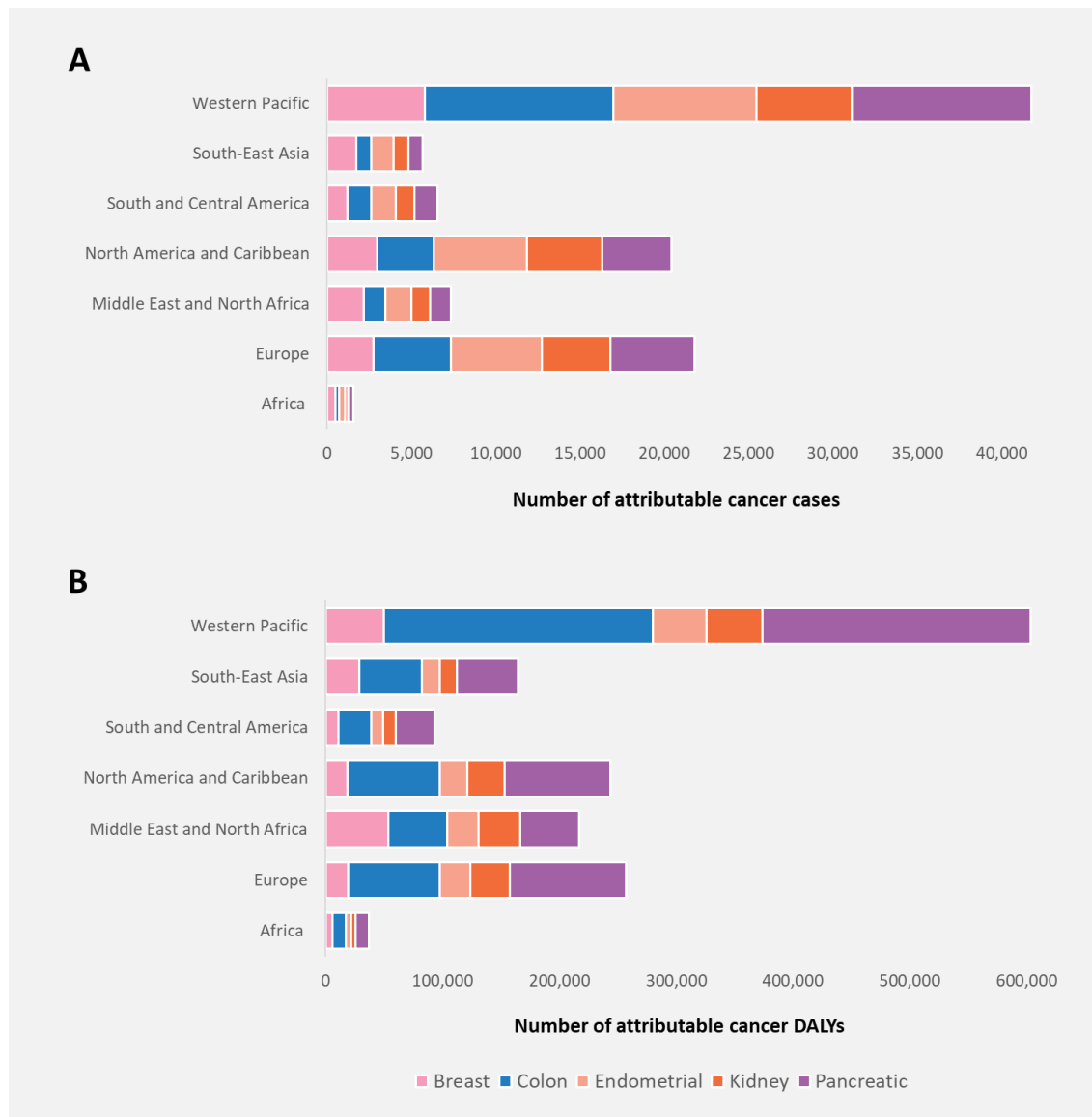


Figure 7-2 A) Projected number of cancer cases by region and cancer type. B) Projected cancer-related DALYs by region and cancer type.

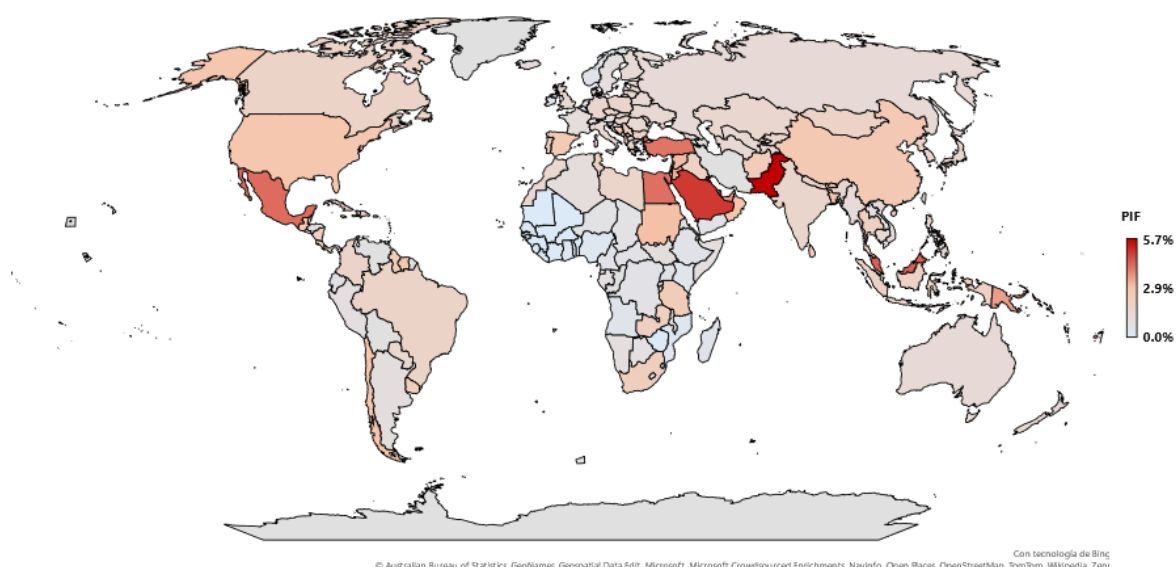


Figure 7-3 Population Impact Fraction (PIF) of All Cancer Incidence in 2025. In red are the countries with the highest PIF (5.7%), including Kuwait and Pakistan. In orange are countries with a PIF around 2.9%, such as Chile, Singapore, and Sudan. In light blue are countries with the lowest PIF, such as Sierra Leone and Guinea-Bissau, with a PIF of 0.01%.

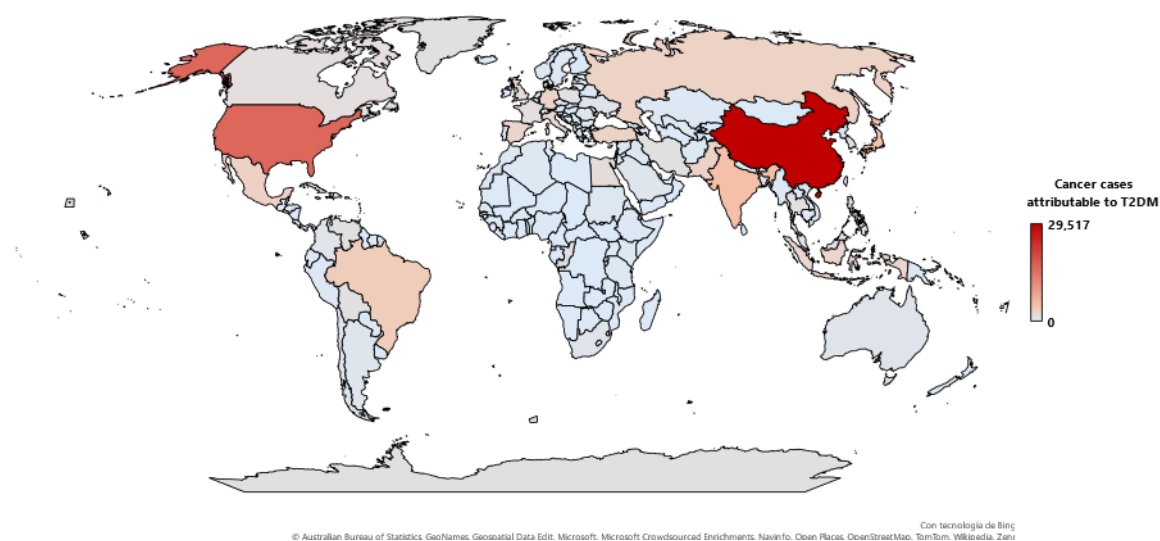


Figure 7-4 Projected Number of Cancer Cases Attributable to T2DM in 2025. In red are the countries with the highest number of cases, led by China (29,517 cases) and followed by the USA (16,230 cases). In orange are countries with approximately 4,000 cases, such as India (4,750 cases) and Japan (4,369 cases). In light blue are countries with fewer cases, primarily from the African region.

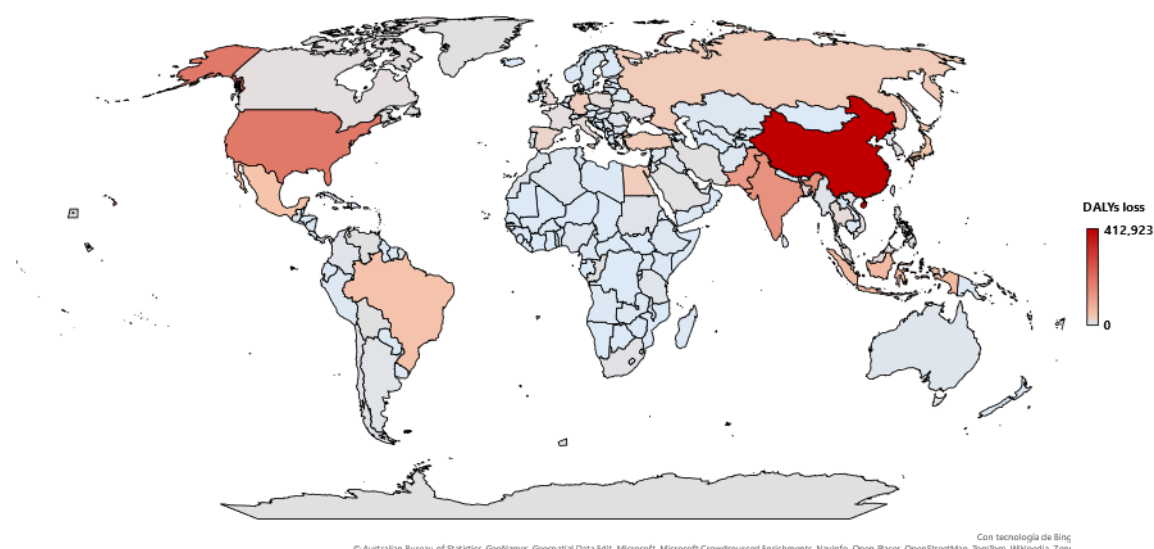


Figure 7-5 Projected Cancer-Related DALY Loss Due to T2DM in 2025. In red is China, with the highest DALY loss (412,923 DALYs). In orange are countries with DALY losses of around 40,000, including Japan (39,000 DALYs) and Mexico (42,642 DALYs). In light blue are countries with the lowest cancer-related DALY losses due to T2DM, mainly from the African region.

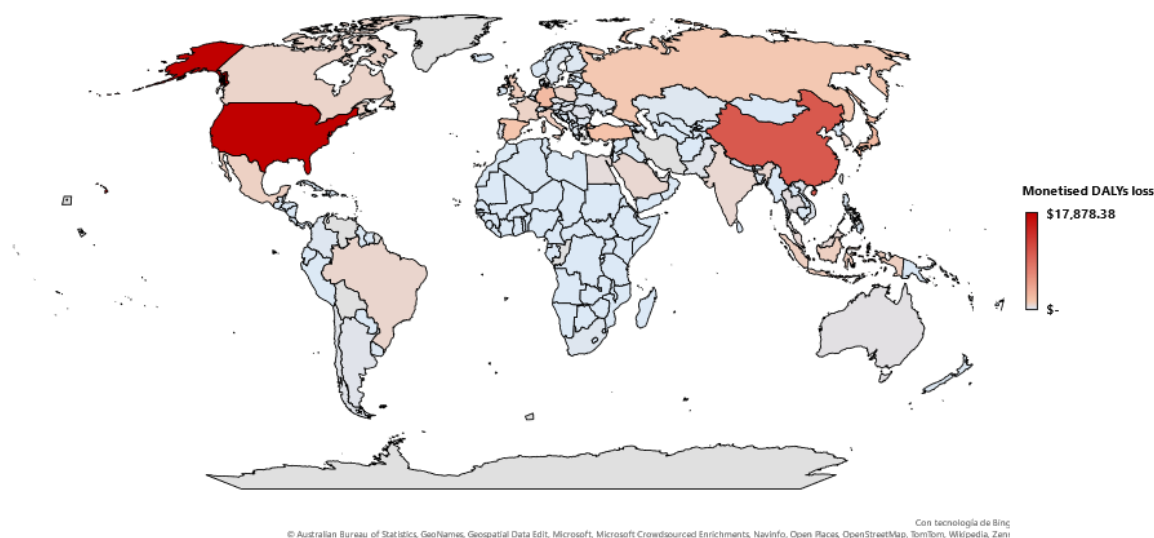


Figure 7-6 Projected Monetised DALY Loss (in Millions) in 2025. In red are the countries with the highest economic burden, including the USA (INT \$17,878.39 million) and China (INT \$10,469.59 million). In orange are countries with monetised DALY losses of around INT \$1,500 million, such as Spain (INT \$1,475.16 million). In light blue are countries with the least economic impact, primarily from the African region. INT \$ refers to international dollars at constant 2017 prices.

7.4.1.4 Type of cancer

PIF

As shown in Figure 7-7, endometrial cancer will have the highest PIF across all regions, ranging from 3.0% in Africa to 11.2% in the MENA region. Pancreatic cancer will consistently have the second-highest PIF, except in the MENA region, with values ranging from 2.4% in Africa to 5.4% in North America and the Caribbean.

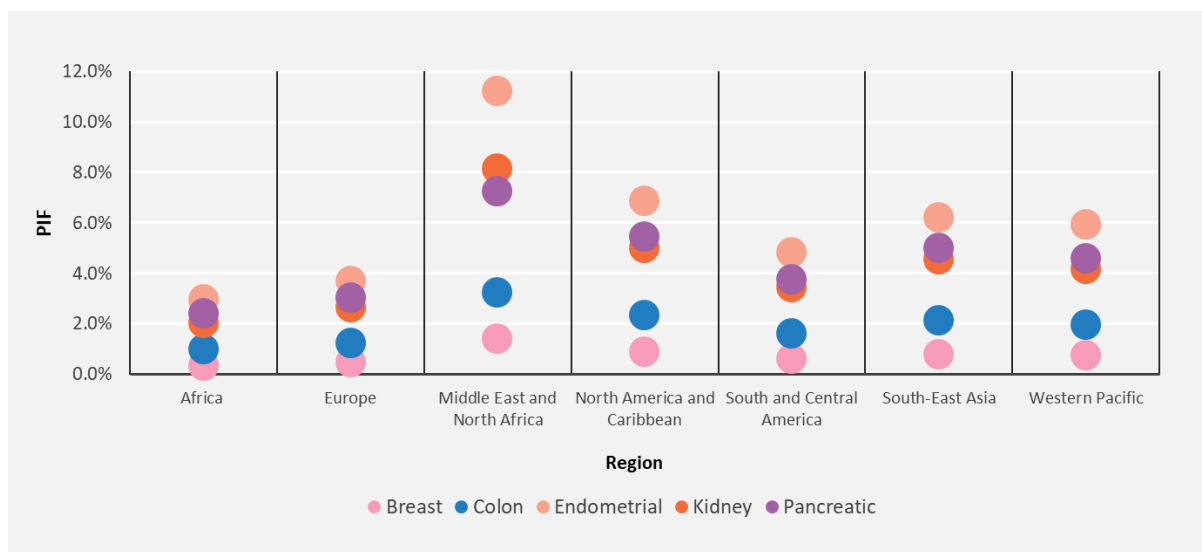


Figure 7-7 Expected PIF by region and type of cancer.

Cancer cases attributable to T2DM

The proportions of specific cancer types among T2DM-related cancer cases will vary by region, except in South and Central America, where colon (1,402 cases, 21.4%), endometrial (1,458 cases, 22.2%), and pancreatic cancers (1,380 cases, 21%) will have similar proportions. In this region, Brazil will report the highest number of cases: colon (635), endometrial (725), and pancreatic (666).

Breast cancer will account for the largest proportion of T2DM-related cancer cases in Africa (485 cases, 31.2%), MENA (2,152 cases, 29.4%), and South-East Asia (1,741 cases, 30.7%).

- Africa: South Africa will report the most cases (151).
- MENA: Pakistan will report the highest number (845).
- South-East Asia: India will account for the largest number, with 1,493 cases.

Endometrial cancer will represent the highest proportion of T2DM-related cancer cases in Europe (5,393 cases, 24.7%) and in North America and the Caribbean (5,492 cases, 26.9%).

- Europe: The Russian Federation will have the most cases (845).
- North America and the Caribbean: The USA will report the majority, with 4,404 cases.

In the Western Pacific region, colon cancer will account for 26.8% of T2DM-related cancer cases (11,178 cases). China will contribute the most, with 8,099 cases.

DALYs attributable to T2DM

Colon and pancreatic cancer DALYs due to T2DM will account for more than 60% of the total DALYs and economic burden (monetised DALYs) across all regions except in MENA. In MENA, these cancers will constitute only 47% of DALYs but will contribute 65% to the region's economic burden.

The proportion of **colon and pancreatic** DALYs will be similar across Africa, South-East Asia, and the Western Pacific region:

- In Africa, colon and pancreatic cancer will contribute to 62% of the region's DALYs. Colon cancer will account for 11,419 DALYs (INT \$59.22 million), and pancreatic cancer for 11,476 DALYs (INT \$74.07 million). South Africa will report the highest numbers for both cancers, with 2,811 colon DALYs (INT \$45.18 million) and 3,744 pancreatic DALYs (INT \$60.17 million).
- In South-East Asia, these two cancers will represent 64% of DALYs. Colon cancer will contribute 53,682 DALYs (INT \$311.99 million), while pancreatic cancer will account for 52,480 DALYs (INT \$305.83 million). India will report the most DALYs, with 46,677 for colon cancer (INT \$256.80 million) and 45,077 for pancreatic cancer (INT \$248.00 million).
- In the Western Pacific region, colon and pancreatic cancers will constitute 76% of DALYs. Colon cancer will account for 229,755 DALYs (INT \$6,475.84 million) and pancreatic cancer for 229,364 DALYs (INT \$6,689.28 million). China will report the highest numbers, with 160,735 colon DALYs (INT \$4,075.42 million) and 166,453 pancreatic DALYs (INT \$4,220.38 million).

In Europe, North America and the Caribbean, and South and Central America, **pancreatic cancer** will account for the largest proportion of T2DM-related DALYs:

- In Europe, pancreatic cancer will represent 39% of DALYs (INT \$5,720.84 million), with Turkey reporting the highest number (13,911 DALYs; INT \$716.17 million).
- In North America and the Caribbean, pancreatic cancer will account for 37% of DALYs (INT \$7,342.74 million), with the USA reporting the highest number (67,948 DALYs; INT \$6,573.07 million).
- In South and Central America, pancreatic cancer will represent 35% of DALYs (INT \$687.24 million), with Brazil reporting the highest number (15,657 DALYs; INT \$297.94 million).

In the MENA region, breast cancer will account for the highest proportion, with 24.7% of T2DM-related DALYs (INT \$387.50 million), with Pakistan contributing the highest number (39,188 DALYs; INT \$77.20 million).

7.4.2 Cancers averted to T2DM

This section focuses on cancer for which T2DM is causally associated with a decreased risk (protective effect), namely **prostate cancer**.

7.4.2.1 Overall findings

The **global PFP was 1.3%, equating to 21,351 prostate cancer cases** that could have been averted in 2025 if the prevalence of T2DM had been reduced to the minimum level (counterfactual scenario). However, based on 2021 T2DM prevalence estimates, 28,213 prostate cancer cases (1.7%) are expected to be averted because T2DM acts as a protective factor. Therefore, a reduction in T2DM prevalence could lead to an increase in prostate cancer cases, meaning fewer cases could be prevented.

The **21,351 prostate cancer cases** that could have been averted under the minimum T2DM scenario would correspond to **138,305 DALYs saved**, with an associated economic burden reduction valued at **INT \$5,272.93 million**. In contrast, based on 2021 T2DM prevalence, 180,293 DALYs (INT \$6,929.29 million) are expected to be averted due to T2DM's protective effect against prostate cancer.

7.4.2.2 By Region and Country level

Globally, Pakistan will have the highest PFP at 5.7%, with 320 prostate cancer cases expected to be averted due to the protective effect of T2DM. Pakistan's high PFP could be because it has the highest diabetes rate in the world (30.8%). The burden of cancer cases prevented will vary across regions, ranging from 0.8% in Africa to 2.5% in the MENA region. The PFP for the remaining regions will fall between 1% and 2%.

Table 7-3 provides an overview of the cancer burden averted to T2DM by region.

Table 7-3 PFP, expected number of averted cancer cases, DALYs averted and economic value of DALYS averted (monetised DALYs) to T2DM in 2025 by region.

Region	Total number of cancer cases	PFP	Estimated number of averted cancer cases	Total number of DALYs	Estimated number of DALYs averted	Estimated monetised DALYs averted in millions (INT. \$)
Africa	90,961	0.8%	730	1,117,971	7,572	\$39.16
Europe	547,402	1.0%	5,360	2,453,378	23,488	\$1,388.52
Middle East and North Africa	42,800	2.5%	1,112	435,192	14,579	\$225.47
North America and Caribbean	298,973	1.9%	5,655	1,469,592	28,420	\$2,201.33
South and Central America	212,655	1.3%	2,870	1,116,447	14,634	\$291.63
South-East Asia	45,918	1.6%	770	786,236	13,288	\$77.84
Western Pacific	346,344	1.4%	4,855	2,315,079	36,325	\$1,048.97

The "Total number of DALYs" refers to the overall burden of disease in terms of DALYs lost due to cancer, without considering the impact of reducing T2DM prevalence. The "Estimated number of DALYs averted" represents the portion of DALYs that would be avoided if T2DM prevalence were reduced to the theoretical minimum level (counterfactual scenario). INT \$: international dollars at constant 2017 prices.

Cases averted

As shown in Table 7-3, Africa will have the lowest proportion of prostate cancer cases prevented due to T2DM, with only 730 cases (3% of the global total [21,351 cases]), followed closely by the South-East Asia region, with 770 cases (4%). In contrast, nearly 75% of all averted prostate cancer cases will occur in three regions: Europe (5,360 cases), the Western Pacific (4,855 cases), and North America and the Caribbean (5,655 cases). Among these, North America and the Caribbean will exhibit the highest PFP, driven primarily by the USA, which will lead globally with 4,144 averted cases, accounting for 73% of the region's total. The large number of cases prevented in the USA reflects both the high prevalence of T2DM and the considerable number of prostate cancer cases projected for 2025. In the USA, 227,207 prostate cancer cases are projected by that year, nearly twice the number projected for China (141,012), despite China's larger population.

In Europe, the 5,360 averted cases will come from several countries, with none preventing more than 800 cases. In the Western Pacific, the majority of averted cases will come from China (2,542 cases) and Japan (1,069 cases), together accounting for most of the region's total.

In South and Central America, 2,870 prostate cancer cases are expected to be averted due to T2DM, with Brazil contributing 57% (1,654 cases). In the MENA region, 1,112 cases will be prevented, with Pakistan accounting for 29% (320 cases) of the region's total.

DALYs averted

In terms of DALYs, China will report the highest number of DALYs averted globally (21,436), followed by the USA with 19,895 DALYs. However, the monetised value of averted DALYs will be higher in the USA than in China (USA: INT \$1,924.53 million vs. China: INT \$543.52 million), reflecting differences in GDP per capita. Notably, the USA contribution will represent 36% of the total global economic burden reduction.

By region, Africa will report the fewest DALYs averted (7,572 DALYs; INT \$39.16 million), with South Africa contributing the highest share (2,700 DALYs; INT \$30.54 million). In contrast, the Western Pacific region will account for the largest proportion of DALYs averted (36,325 DALYs; INT \$1,048.97 million),

driven primarily by China. North America and the Caribbean will report 28,420 DALYs averted (INT \$2,201.33 million), largely due to the USA.

Pakistan and Brazil will account for most DALYs averted in the MENA and South and Central America regions, respectively. Brazil will contribute 6,648 DALYs (INT \$126.50 million), while Pakistan will contribute 7,376 DALYs (INT \$14.53 million).

In the South-East Asia region, the number of DALYs averted will be slightly lower than in in the MENA and South and Central America regions, with India accounting for 82% of the region's DALYs (10,912 DALYs) and 77% of its economic burden reduction (INT \$60.04 million).

In Europe, more than 2,000 DALYs will be averted in four countries, with the United Kingdom contributing the most (2,613 DALYs; INT \$179.69 million). Other key contributors include the Russian Federation (2,200 DALYs; INT \$93.76 million), Spain (2,521 DALYs; INT \$153.35 million), and Turkey (2,429 DALYs; INT \$125.06 million). However, due to differences in GDP per capita, the economic value of averted DALYs in the Russian Federation will be lower than the other countries despite its substantial DALY reduction.

7.4.3 Further analyses: Number of Cancer Cases Attributable to Confirmed T2DM

7.4.3.1 Cancers attributable to T2DM

To provide a more accurate estimate of the cancer burden attributable to T2DM, Table 7-4 compares the number of cancer cases associated with a confirmed T2DM diagnosis against those derived in the base case, which considers T2DM prevalence (undiagnosed + diagnosed).

The findings show that due to the prevalence of undiagnosed T2DM, only 50% of the cancer cases attributable to T2DM could be identified. As a result, the **projected number of cancer cases** attributed to T2DM decreases from 105,084 to **52,642**.

In this scenario, the 52,642 cases will represent 1% of the total breast, endometrial, colon, kidney, and pancreatic cancer cases, compared to 2% if all individuals with T2DM were aware of their condition.

Table 7-4 PIF and projected number of cancer cases attributable to T2DM in 2025 among individuals with a confirmed T2DM diagnosis against those derived in the base case

Type of cancer	Base case		Confirmed diagnosis of T2DM		Difference in cases
	PIF	Estimated number of attributable cancer cases	PIF	Estimated number of attributable cancer cases	
Breast	0.7%	17,091	0.3%	8,232	8,859
Colon	1.8%	22,948	0.9%	10,872	12,076
Endometrial	5.4%	24,072	2.8%	12,365	11,707
Kidney	3.8%	17,444	2.1%	9,363	8,081
Pancreatic	4.2%	23,530	2.1%	11,611	11,919
Total	2.0%	105,084	1.0%	52,442	52,642

Confirmed T2DM diagnosis was calculated by subtracting the proportion of undiagnosed diabetes from the total diabetes prevalence used in the main analysis, both measures reported by the International Diabetes Federation (IDF). Difference in cases: Estimated number of attributable cancer cases in the base case minus those in the confirmed T2DM diagnosis scenario.

Among the different cancers, kidney cancer shows the smallest underestimation, with the PIF dropping from 3.8% to 2.1%, a reduction of less than 50%. The decrease is more pronounced for other cancer types, with the burden estimated at almost half of the cases projected in the main analysis.

Despite the differences in undiagnosed prevalence, the distribution of T2DM-attributable cancer cases remains consistent across cancer types, similar to the trend observed in the base case. Endometrial cancer will account for 23.6% (12,365 cases), followed by pancreatic cancer (22.1%, 11,611 cases), colon cancer (20.7%, 10,872 cases), kidney cancer (17.9%, 9,363 cases), and breast cancer (15.7%, 8,232 cases).

The findings, stratified at the regional level, are illustrated in Figure 7-8. As shown, Africa will experience the greatest impact from undiagnosed T2DM, with the PIF decreasing from 0.7% to 0.2%, resulting in 1,146 fewer cancer cases that could be prevented or diagnosed early with confirmed diabetes.

The Southeast Asia region will also show an important difference, with the PIF dropping from 1.8% to 0.6%, equating to 3,599 fewer cancer cases. Similarly, in the Western Pacific region, the PIF will decrease from 2.2% to 0.8%, a reduction

equivalent to 27,788 cancer cases. In contrast, the North American and Caribbean region will show the smallest gap between the base case and the confirmed-diagnosis scenario, with the PIF dropping only from 2.8% to 2.2% (2,541 fewer cases).

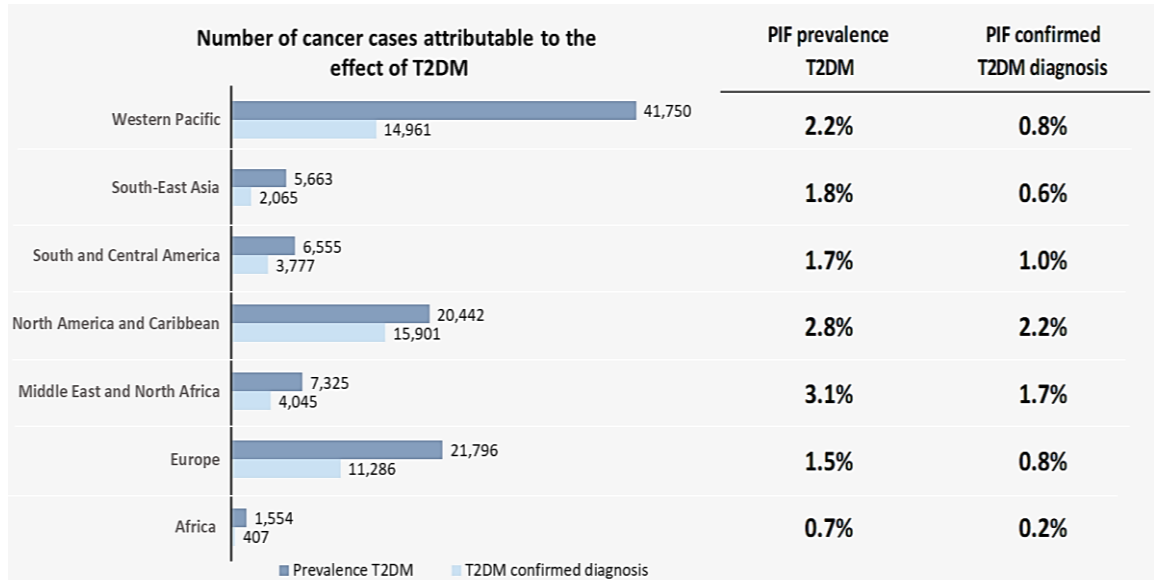


Figure 7-8 PIF and projected number of cancer cases attributable to T2DM in 2025 among individuals with a confirmed T2DM diagnosis compared to those in the base case, by region.

These findings emphasize the need for improved health programs focused on identifying undiagnosed T2DM, as such efforts could also contribute to earlier cancer detection and prevention. Africa would benefit the most from improved diabetes screening, as many cancer cases could be avoided with earlier diagnosis. Other regions with large differences, such as Southeast Asia and the Western Pacific, would also gain from improved diabetes screening and management efforts.

7.4.3.2 Cancer cases averted to T2DM

In the scenario with a confirmed T2DM diagnosis, the global PFP for prostate cancer cases averted would be 0.7%, equating to 11,525 cases prevented, compared to 1.3% (21,351 cases) in the base case. However, since T2DM acts as a protective factor against prostate cancer, the number of cases prevented remains unaffected, regardless of diagnosis confirmation. Nonetheless, a confirmed diagnosis is important for managing the disease effectively and for preventing other complications.

7.4.4 Summary of findings and recommendations

To summarize the findings, Table 7-5 presents the results of different measures of disease burden by region and for those cancer types attributable to T2DM. These are the types of cancer for which T2DM is causally associated with an increased risk. Therefore, a reduction in T2DM will prevent cases of these types of cancer.

Table 7-5 Results of the different measures of disease burden by region and type of cancer: PIF, expected number of attributable cancer cases, DALYs loss and economic value of DALYs loss (monetised DALYs) attributable to T2DM in 2025

Measure of burden of disease	Region	Type of cancer				
		Breast	Colon	Endometrial	Kidney	Pancreatic
PIF	Africa	0.3%	1.0%	3.0%	2.0%	2.4%
	Europe	0.5%	1.2%	3.7%	2.6%	3.0%
	MENA	1.4%	3.2%	11.2%	8.1%	7.3%
	NAC	0.9%	2.3%	6.9%	5.0%	5.4%
	SACA	0.6%	1.6%	4.8%	3.4%	3.8%
	SEA	0.8%	2.1%	6.2%	4.6%	5.0%
	WP	0.7%	2.0%	5.9%	4.2%	4.6%
Estimated number of attributable cancer cases	Africa	485	239	355	174	300
	Europe	2,740	4,602	5,393	4,087	4,975
	MENA	2,152	1,290	1,566	1,089	1,227
	NAC	2,988	3,357	5,492	4,461	4,145
	SACA	1,191	1,402	1,458	1,125	1,380
	SEA	1,741	881	1,305	884	851
	WP	5,794	11,178	8,503	5,624	10,651
Estimated number of DALYs loss	Africa	6,316	11,419	4,277	3,482	11,476
	Europe	19,436	78,435	26,300	33,657	99,121
	MENA	53,751	50,490	26,659	35,832	50,486
	NAC	18,558	79,081	23,809	32,079	90,349
	SACA	10,847	28,146	10,321	11,144	32,983
	SEA	28,944	53,682	15,256	14,400	52,480
	WP	50,089	229,755	46,023	48,274	229,364
Estimated monetised DALYs loss in millions (INT. \$)	Africa	\$27.03	\$59.22	\$21.28	\$16.98	\$74.07
	Europe	\$1,093.65	\$4,481.73	\$1,405.08	\$1,871.21	\$5,720.84
	MENA	\$387.50	\$754.46	\$187.63	\$395.51	\$1,024.70
	NAC	\$1,410.05	\$6,624.76	\$1,958.50	\$2,557.45	\$7,342.74
	SACA	\$228.30	\$587.36	\$194.94	\$235.00	\$687.24
	SEA	\$165.79	\$311.99	\$91.35	\$89.50	\$305.83
	WP	\$1,287.72	\$6,475.84	\$1,209.56	\$1,434.76	\$6,689.28
cost/DALY*	Africa	\$0.0043	\$0.0052	\$0.0050	\$0.0049	\$0.0065
	Europe	\$0.0563	\$0.0571	\$0.0534	\$0.0556	\$0.0577
	MENA	\$0.0072	\$0.0149	\$0.0070	\$0.0110	\$0.0203
	NAC	\$0.0760	\$0.0838	\$0.0823	\$0.0797	\$0.0813
	SACA	\$0.0210	\$0.0209	\$0.0189	\$0.0211	\$0.0208
	SEA	\$0.0057	\$0.0058	\$0.0060	\$0.0062	\$0.0058
	WP	\$0.0257	\$0.0282	\$0.0263	\$0.0297	\$0.0292

*The cost per DALY was calculated to compare the impact of the different types of cancer burden due to T2DM in a standardized unit. The cost per DALY was calculated by dividing the estimated monetised DALY loss (in millions, INT \$) by the estimated number of DALYs lost. *Abbreviations.* INT \$: international dollars at constant 2017 prices; MENA: Middle East and North Africa; NAC: North America and Caribbean; SACA: South and Central America; SEA: South-East Asia; WP: Western Pacific.

Based on these findings, the following recommendations are proposed for each region to most effectively address the effect of T2DM on the cancer burden.

- Africa:** Pancreatic cancer will constitute a high burden (PIF of 2.4%), along with the largest DALY loss (11,476), economic burden (INT \$74.07 million), and cost per DALY (\$0.0065) in the region. These findings support prioritizing **pancreatic** cancer prevention and early screening programs for T2DM patients to reduce health and economic burdens effectively. However, if the goal is to address the largest number of cancer cases, interventions targeting **breast** cancer should take precedence, as this type accounts for the highest proportion of T2DM-related cancer cases in the region (485).
- Europe:** Endometrial cancer will represent the highest burden (PIF 3.7%) and the highest number of attributable cancer cases (5,393). However, **pancreatic** cancer, will constitute the second-highest PIF (3.05%) and case count (4,975), will surpass all other cancers in DALY loss (99,121), economic burden (INT \$5,720.84 million), and cost per DALY (\$0.0577). Interventions targeting both pancreatic and endometrial cancers could significantly reduce the health and economic burden of T2DM-associated cancers in Europe.
- Middle East and North Africa (MENA):** **Pancreatic** cancer will constitute the highest economic impact (INT \$1,024.70 million), cost per DALY (\$0.0203), and DALY loss in the region (50,486), making it a critical target for interventions. However, **endometrial** cancer also warrants attention, with one in ten cases attributable to T2DM (PIF 11.2%). Screening initiatives for endometrial cancer should be prioritized to address its substantial burden of disease.
- North America and the Caribbean:** **Pancreatic cancer**, with a PIF of 5.4%, the highest DALY loss (90,349), economic burden (INT \$7,342.74 million), and cost per DALY (\$0.0813), will make it the most impactful target for interventions in this region. If the focus will be on reducing the largest proportional change in disease incidence and case counts, **endometrial** cancer would be the priority, with a PIF of 6.9% and 5,492 attributable cases.

- **South and Central America:** **Pancreatic** cancer will constitute the second-highest burden (PIF 4.8%) and will account for the highest DALY loss (32,983) and economic burden (INT \$687.24 million). Targeted interventions for pancreatic cancer in T2DM patients could mitigate its substantial health and economic effects. Alternatively, **endometrial** cancer, with the highest PIF (4.8%) and case count (1,458), could be prioritized if reducing incidence and case numbers is the primary goal.
- **South-East Asia:** **Pancreatic** cancer will rank as the second-highest in PIF (5.0%) and will exhibit a significant DALY loss (52,480) and economic burden (INT \$305.83 million). Interventions focusing on pancreatic cancer prevention and early detection would help alleviate the region's health and financial challenges.
- **Western Pacific:** **Pancreatic cancer** will constitute the second-highest burden (PIF 4.6%), contribute with a high case count (10,651) and DALY loss (229,364), and represent the largest economic impact (INT \$6,689.28 million) and cost per DALY (\$0.0292). This suggests that managing pancreatic cancer in T2DM patients would yield the greatest economic and health benefits for this region.

Overall, the findings suggest that implementing strategies for **pancreatic cancer** prevention and early detection in all regions, and **endometrial cancer** in some regions, could help reduce the cancer burden attributable to T2DM.

These targeted interventions focus on the specific cancer types with the highest burden in each region, considering PIF, cancer cases, DALY losses, and economic impact. They would likely represent the most efficient use of resources for reducing the effect of T2DM on the global cancer burden.

It is also important to note that the research question of this thesis is to estimate the independent causal effect of T2DM on cancer risk. Therefore, the recommendations focus solely on T2DM's effect on cancer risk. However, in practice, T2DM prevention strategies also target shared risk factors for both T2DM and cancer. Addressing these factors could lead to broader health benefits, potentially exceeding the estimates presented.

7.5 Discussion

This global burden of disease analysis in this thesis projects that by 2025, 2% of breast, colon, endometrial, kidney, and pancreatic cancer cases will be attributable to T2DM across 177 countries. Regionally, the MENA is expected to contribute with the highest burden, with a PIF of 2.8%. Within this region, Kuwait and Pakistan will lead both regionally and globally with a PIF of 5.7%, mainly due to the high T2DM prevalence of approximately 25%. In contrast, Africa will have the lowest PIF at 0.7%; however, certain countries, such as Tanzania, will reach a PIF of 2.73%, showing that there is variability within regions.

These findings align with IDF data on T2DM prevalence, as the calculated PIF directly reflects diabetes prevalence, which was the only country-specific data applied, while the RR of each type of cancer causally associated with T2DM was applied uniformly across all regions. Consequently, countries with higher diabetes prevalence report higher PIFs, while those with lower prevalence show reduced PIFs. For instance, as noted by the IDF, MENA has the highest global diabetes prevalence at 16.2%, while African countries report the lowest at 4.5%.

The Western Pacific region, however, is projected to experience the highest number of cancer cases and DALYs attributable to T2DM, with an estimated 41,750 cases (35% of all cancer cases across regions) and 603,505 DALYs lost (37% of total DALYs across regions). In contrast, Africa is expected to account for the lowest burden, with 1,554 cases (1% of all cases) and 36,969 DALYs lost (2% of total DALYs). The difference between regions with the highest PIF and those with the highest absolute number of cases and DALYs arises because cancer cases attributable to T2DM are based on projected incidence from GLOBOCAN, while DALYs are calculated using Global Burden of Disease (GBD) study data.

Thus, although the proportional reduction of cancer due to T2DM is lower in the Western Pacific compared to the MENA region, the large population of the Western Pacific results in a greater number of T2DM-attributable cancer cases and a higher disease burden measured in DALYs.

Within the Western Pacific and globally, China is projected to contribute to the largest cancer burden due to the effect of T2DM, with an estimated 25,517 cases and 412,923 DALYs lost.

Despite this meaningful health impact, the financial burden in China remains lower than in the USA, with a total economic impact of INT \$10,469.59 million compared to INT \$17,878.38 million in the USA. This difference is due to the economic assessment methodology used, which considers GDP per capita and HDI classification. Then, the USA, with a higher GDP per capita and a "very high" HDI, incurs greater economic costs per DALY lost than China, which has a lower GDP per capita and is classified as "high" HDI. This disparity reflects that, while both nations will cope with substantial health burdens, the economic impact per DALY is proportionally higher in wealthier countries, where each DALY represents a greater financial loss due to higher average income, indirectly measured through the GDP per capita.

Among the cancers attributable to the effect of T2DM, endometrial cancer is projected to have the highest global burden, with 1 in 20 cases attributable to T2DM, equating to 24,072 cases. Pancreatic cancer is expected to follow closely, with 1 in 25 cases attributable to T2DM, totalling 23,530 cases. Among men, pancreatic cancer is expected to contribute the greatest burden, with 12,700 cases compared to 10,830 among women. Pancreatic cancer will also contribute to the largest proportion of global DALYs, highlighting its severe impact on quality of life and survival.

Of the 1.6 million DALYs projected to be lost to all cancers due to T2DM, pancreatic cancer is expected to account for 35% (566,259 DALYs). It will also represent the highest financial burden worldwide, with an economic impact of INT \$21,844.70 million and the highest cost per DALY across all types, at INT \$0.038 million per DALY. These findings emphasize the need for interventions designed to prevent T2DM and improve early cancer detection, especially for pancreatic and endometrial cancers, in order to alleviate future health and economic burdens.

Regionally, these two cancer types are also projected to represent the highest burdens when measured by PIF. Endometrial cancer is expected to have the highest burden across all regions, particularly in the MENA region, where it could

reach 11.2%. This suggests that over one in ten cases will be due to T2DM. Consequently, reducing the prevalence of T2DM could significantly decrease the incidence of cancer in this region. Pancreatic cancer is anticipated to rank as the second-highest burden in most regions; however, in MENA, kidney cancer is projected to surpass pancreatic cancer with a PIF of 8.1%.

The distribution of cancer burden by attributable case counts is expected to vary across regions. Endometrial cancer is projected to dominate in Europe and North America, while breast cancer will account for the largest proportion of cases in Africa, MENA, and South-East Asia. Colon cancer is expected to lead in the Western Pacific, whereas in South and Central America, colon, pancreatic, and endometrial cancers are forecasted to contribute similarly to the burden.

When measured by DALYs and economic costs, pancreatic cancer is anticipated to remain the primary contributor in Africa, Europe, North America, and South and Central America. In South-East Asia and the Western Pacific, both pancreatic and colon cancers are expected to account for the largest shares of DALYs and economic impact. In MENA, while breast cancer is projected to lead in DALYs, colon and pancreatic cancers will follow, with pancreatic cancer contributing to the highest economic burden in the region. These findings highlight the importance of tailoring intervention strategies to regional and specific cancer burdens to better address the health and financial impacts of cancers attributable to the effect of T2DM.

These results of the present burden of disease study suggest that the choice of cancer type for targeted interventions will depend on the region and the specific measure of disease burden being addressed. Globally, pancreatic cancer is projected to have the highest economic impact across most regions, except in Southeast Asia, where it is expected to rank second after colon cancer. In terms of DALYs lost, pancreatic cancer is anticipated to remain among the leading contributors across all regions except in the MENA region. Meanwhile, endometrial cancer, which is expected to have the highest PIF in all regions, will likely represent the greatest burden when measured by this metric. This emphasises the need for region-specific strategies to address both the economic and health impacts of cancer due to T2DM in the future.

It is noteworthy that although Africa is projected to have the lowest cancer burden attributable to the effect of T2DM among all regions due to its currently low T2DM prevalence rates, this situation is expected to change dramatically. By 2045, Africa is forecasted to exhibit a 129% increase in the number of people with diabetes, the highest growth rate among all regions, reaching an estimated 55 million individuals (International Diabetes Federation, 2021). Furthermore, by 2045, the number of new cancer cases causally associated with T2DM is anticipated to show the highest percentage increase across all regions, with more than a 100% rise in all types of such cancers (Ferlay J, Laversanne M, Ervik M, n.d.). This sharp rise in T2DM cases and related cancers is expected to intensify the overall cancer burden in the region. Therefore, implementing interventions to prevent T2DM will remain a critical and urgent strategy, particularly in Africa, to address this growing health challenge.

In addition to preventing T2DM, raising awareness about the condition is crucial to ensure early detection of the types of cancers attributable to T2DM. According to the IDF, one in two adults is unaware they have T2DM. This lack of awareness presents a challenge in monitoring people with T2DM to detect cancer in early stages, making such efforts ineffective or inefficient. For instance, based on scenario analysis results, only 50% of cancer cases attributable to T2DM could currently be identified due to the high prevalence of undiagnosed diabetes. Therefore, interventions to improve T2DM diagnosis should be prioritized across all regions. This is particularly urgent in Africa, the Western Pacific, and Southeast Asia, where more than 50% of T2DM cases remain undiagnosed, compared to 24.1% in the North America and Caribbean region, which has the lowest prevalence of undiagnosed diabetes. Strengthening diagnostic efforts globally will be essential for effectively addressing the dual burden of T2DM and cancer.

The proportion of cancer cases attributable to the effect of T2DM remained consistent between the burden of disease estimates published by Pearson and colleagues and the findings presented in this chapter combination (Pearson-Stuttard et al., 2018). Pearson et al. reported that in 2012, 2.1% (280,100) of cancer cases were attributable to diabetes prevalence in 2002, while in the present analysis, 2% (105,084 cases) of cancer cases by 2025 are expected to be attributable to T2DM in 2021. However, differences in the number of reported

cases can be attributed to variations in the types of cancers included and the methodological approaches used.

Pearson and colleagues based their findings on an umbrella review and meta-analysis, identifying T2DM as a risk factor for colorectal cancer (CRC), gallbladder, pancreatic, liver, endometrial, and breast cancers (Tsilidis, Kasimis, et al., 2015). In contrast, the present study, using results from the adjusted (causal effect) meta-analysis (Chapter 6), found that T2DM increased the risk of breast, colon, endometrial, kidney, and pancreatic cancers. Pearson's study included CRC rather than colon cancer alone and considered liver cancer instead of kidney cancer. However, no effect estimates accounting for MSAS were available for liver cancer in the studies identified in the systematic review (Chapter 3), which restricted the ability to conduct a causal effect meta-analysis for this cancer type.

The relative risks (RRs) used by Pearson were slightly higher for breast and pancreatic cancer, and similar for CRC (colon) and endometrial cancers. The RRs were: breast (RR: 1.20 [95% CI: 1.12-1.28] vs. 1.11 [95% CI: 1.04-1.18]), pancreatic (RR: 1.95 [95% CI: 1.66-2.28] vs. 1.73 [95% CI: 1.46-2.05]), CRC (RR: 1.27 [95% CI: 1.21-1.34] vs. 1.30 [95% CI: 1.12-1.50]), and endometrial (RR: 1.97 [95% CI: 1.71-2.27] vs. 1.94 [95% CI: 1.43-2.63]).

Pearson's study also included liver cancer (RR: 2.31, 95% CI: 1.87-2.84), which led to a higher case count (115,500) in their analysis.

However, the umbrella review underlying Pearson's analysis did not account for confounding bias, which may affect the reliability of the estimates, especially when quantifying the cancer burden attributable to T2DM, as its effect could not be isolated.

Another methodological difference was the lower counterfactual diabetes prevalence used in Pearson's study, which contributed to their higher case count. Moreover, Pearson's analysis did not include broader measures of disease burden, such as DALYs or monetised DALYs, which limited its ability to comprehensively capture the overall impact of cancers attributable to T2DM.

Additionally, the umbrella review that Pearson relied on did not identify any cancers with a protective effect of T2DM, such as prostate cancer, which was considered in the present analysis. If T2DM prevalence decreases, the number of prostate cancer cases attributable to the effect of T2DM would rise.

Nevertheless, early detection interventions for T2DM could improve prognosis in

these cases. More importantly, reducing T2DM prevalence would prevent far more cases of breast, endometrial, colon, pancreatic, and kidney cancers than the additional prostate cancer cases. Furthermore, addressing T2DM would not only decrease the burden of associated cancers but also help prevent complications such as heart attacks, strokes, and kidney failure, which are leading causes of morbidity and mortality. Efforts to lower the prevalence of T2DM could yield health benefits beyond cancer prevention.

Strengths and limitations

This burden of disease analysis has several strengths. Unlike the GBD study's most recent publication on T2DM, it includes cancer as a complication of diabetes to provide a more complete understanding of the disease's impact. Compared to the work of Pearson and colleagues, the only published research that also examined the cancer burden attributable to T2DM, this analysis explicitly identifies cancers causally associated with T2DM, using a robust methodology.

These cancer types were determined based on the adjusted (causal effect) meta-analyses presented in Chapter 6. To ensure the least biased estimates of the causal effect of T2DM on cancer risk, only studies that accounted for the MSAS, which represents the set of variables required to control for confounding and was identified using DAGs (Chapter 4) were selected. In contrast, Pearson's work relied on an umbrella review and meta-analyses that did not account for confounding bias, limiting the reliability of their estimates.

Moreover, the umbrella review used by Pearson and colleagues for their burden of disease estimates only included publications up to 2013, while the studies used in the meta-analyses of this thesis were drawn from a comprehensive and updated systematic review of the T2DM-cancer relationship across various cancer types. This systematic review incorporated five databases, employed broad search algorithms, and imposed no language or time restrictions, covering literature up to December 2023 (Chapter 3). Therefore, by using updated evidence and a rigorous approach to address confounding bias, this study provides a more accurate estimate of the cancer burden from the effect of T2DM.

To provide a realistic picture and improve the applicability of the findings, the present analysis considered a counterfactual prevalence of diabetes based on the lowest achievable rate, assuming individuals were in the lowest risk category for modifiable factors. Additionally, a scenario analysis was conducted to assess the impact of undiagnosed diabetes prevalence on cancer prevention efforts and to raise awareness of this issue.

Furthermore, the study adopts a holistic approach by incorporating multiple burden of disease metrics, including PIF/PFP, the number of cancer cases, DALYs, and monetised DALYs. The use of reliable and up-to-date T2DM prevalence data from the IDF further strengthens the analysis, ensuring the findings are based on accurate epidemiological data.

Despite these strengths, the study has some limitations. The projected DALYs focus solely on cancer, potentially underestimating the total burden, as evidence suggests that patients with both T2DM and cancer experience worse prognoses, including higher risks of death and disease progression (Cortellini et al., 2023). Additionally, while monetised DALYs provide valuable insights into the economic burden of T2DM-related cancer, the analysis could be further enriched by considering the specific needs and contexts of individual countries. Including the costs of available treatments, interventions, and indirect costs, such as productivity loss, would offer policymakers a more comprehensive framework for resource allocation based on national needs and priorities.

Conclusions

In conclusion, by 2025, 2% of global cases of breast, colon, endometrial, kidney, and pancreatic cancers will be attributable to the effect of T2DM, with the MENA region projected to have the highest share of cancer cases due to T2DM. The findings of this study underscore the need for interventions to prevent or reduce T2DM prevalence in order to decrease the associated cancer burden. However, the results may underestimate the full impact of such interventions, as this analysis focused specifically on the effect of T2DM on cancer risk. In practice, T2DM prevention strategies also address shared risk factors, such as obesity, through lifestyle changes. These risk factors not only contribute to T2DM but also increase the risk of several cancers. By addressing these interrelated risk factors, broader health benefits could be achieved, potentially exceeding the estimates presented.

Chapter 8 Discussion

8.1 Introduction

Diabetes mellitus (DM) is a major public health concern, with global projections estimating 783 million cases by 2045. Type 2 diabetes mellitus (T2DM) is the most prevalent form, accounting for over 90% of all cases worldwide (International Diabetes Federation, 2021). In addition to its well-known complications, growing evidence suggests a link between T2DM and increased cancer risk (American Cancer Society, 2021; American Diabetes Association, 2024).

While the association between T2DM and cancer has been widely studied, no causally interpretable meta-analyses have been published on the relationship between T2DM and cancer risk across a range of cancer types. This thesis addresses that gap by determining whether T2DM has an independent causal effect on cancer risk across multiple cancer types. In this context, independent causal effect refers to a relationship that is not influenced by confounding factors.

Recognizing these causally interpretable associations can inform the development of more targeted public health interventions, particularly in cancer screening and prevention strategies for individuals with T2DM.

This chapter summarizes the key findings, outlines its strengths and limitations, provides policy recommendations, suggests directions for future research, highlights its contribution to knowledge, and presents the final conclusions.

8.2 Main findings

8.2.1 Context (Chapter 2) and Systematic Literature Review (Chapter 3)

As presented in Chapter 2, only three systematic reviews (SRs) have explored the relationship between T2DM and multiple cancer types (Ling et al., 2020; Ohkuma et al., 2018; Starup-Linde et al., 2013). These studies reported findings for each

cancer type examined, but their results varied in terms of statistical significance, magnitude of association, and the specific cancers included. The discrepancies in their findings can be attributed to several methodological differences. Some SRs excluded non-English studies, restricted their search to only a few databases (Ling et al., 2020; Ohkuma et al., 2018), or conducted their searches more than a decade ago, even when including a larger number of databases (Starup-Linde et al., 2013). Additionally, some studies included only cohort studies (Ling et al., 2020; Ohkuma et al., 2018). Differences in how confounders were adjusted for may have also contributed to the varying results, as acknowledged by the authors.

Confounding adjustment is necessary for causal inference but is not always required when exploring only associations. Of the three publications, two explored associations, while one assessed the strength of causal evidence in the observed associations (Ling et al., 2020). Therefore, no causally interpretable meta-analyses have been published on the relationship between T2DM and cancer risk across a range of cancer types.

The research question that remains unanswered and serves as the basis of this thesis is:

What is the independent causal effect of T2DM on cancer incidence?

To address this research question, the first step was to identify all relevant evidence on the relationship between T2DM and cancer risk across multiple cancer types. Compared to previous SRs, which often searched a limited number of databases, applied language restrictions, or included only cohort studies, there was a clear need for a more comprehensive and up-to-date review. This SR followed the Population, Intervention, Comparison, Outcome, and Study design (PICOS) framework: men and women aged 18 years or older (P); individuals with T2DM (I) compared to those without T2DM (C); cancer incidence (O), expressed as risk ratio (RR), hazard ratio (HR), standardized incidence ratio (SIR), or odds ratio (OR); and observational studies (S).

A total of 18,086 publications were identified through the literature search. After title and abstract screening, followed by full-text screening, 250 studies were included in this SR.

The majority of studies were cohort studies (57%), followed by case-control studies (34%). Among these, 80% of cohort studies and 52% of case-control studies were classified as having a low risk of bias (NOS ≥ 7), providing more reliable effect size estimates for the meta-analyses presented in Chapter 6.

Follow-up duration was a key criterion in assessing whether T2DM preceded cancer onset rather than being a consequence of cancer. Most cohort studies had follow-up periods longer than three years, with the largest proportion ranging between 11 and 15 years. This supports the conclusion that most studies provided adequate follow-up to minimize the likelihood of reverse causality and to strengthen the temporal relationship between T2DM and cancer risk.

Although the studies included in the SR covered 30 countries, none were from African nations, likely due to data limitations, as noted by the International Diabetes Federation, which reports high-quality country data for only 58% of African countries. This is relevant as Africa is projected to experience the highest rise in diabetes cases by 2045 (International Diabetes Federation, 2021). However, while these data limitations are important, the absence of published studies on T2DM and cancer may also be due to other research questions taking precedence.

Similarly, limited evidence on the association between T2DM and cancer risk was found for countries with high rates of diabetes prevalence in adults, such as Kuwait, Pakistan, and Mexico. A possible explanation is that research in these regions has prioritized other diseases or well-known diabetes complications, such as macrovascular and microvascular diseases.

The association between T2DM and cancer risk was examined for 63 cancer types. Breast, pancreatic, colorectal, prostate, and lung cancers were the most frequently studied, reflecting their high prevalence among adults worldwide. Pancreatic cancer was of particular interest due to its high mortality rate (Ferlay J, Laversanne M, Ervik M, n.d.).

The variables included in the adjusted analyses of cohort and case-control studies varied widely. Sociodemographic factors, such as age and sex, were commonly accounted for, whereas behavioural factors, such as smoking and alcohol consumption, were less frequently included. The choice of adjustment variables generally reflected the data sources utilized, such as ICD-coded records or the UK Clinical Practice Research Datalink (CPRD), which predominantly provide demographic information but often lack behavioural covariates.

The selection of variables for adjustment should reflect the study's objective. If the aim is to explore associations, adjusting for confounders may not always be necessary, but accounting for them helps ensure that the observed relationship is not influenced by other factors. However, to assess causality, which is the aim of this thesis, it is essential to control for confounders to isolate the independent effect of the exposure and support a causal interpretation.

One recommended method for identifying confounders is through the use of causal diagrams, such as Directed Acyclic Graphs (DAGs). This approach was the focus of the next chapter (Chapter 4: Causal Association).

8.2.2 Causal Association (Chapter 4)

Estimating causal effects in observational studies requires accounting for confounding bias. To determine which variables need to be controlled for when estimating the causal relationship between T2DM and cancer across multiple cancer types, DAGs were created for each cancer type identified in the SR (Chapter 3). DAGs facilitated the identification of the minimum sufficient adjustment set (MSAS), representing the variables necessary to control for confounding (Tennant et al., 2021). After determining the MSAS, these variables were compared with those included in the adjusted analyses of low-risk-of-bias studies from SR.

Of the 56 cancer types for which a DAG was developed, at least one effect estimate meeting the MSAS criteria was reported for 31 types (45%). This indicates that there is evidence supporting a causal effect for many types of

cancer. Prostate cancer had the most studies meeting the MSAS criteria, with 13 of 28 effect estimates (46%) accounting for MSAS.

For most cancer types, the MSAS typically included at least one lifestyle-related confounder, as these variables are common risk factors for both T2DM and several types of cancer. Smoking and alcohol consumption were the most frequently omitted confounders in the adjusted analyses of the studies, despite being part of the MSAS.

DAGs helped identify confounders, including behavioural factors such as alcohol consumption, obesity, smoking, and hypertension, which were not direct risk factors shared by T2DM and the cancer types but were included in the MSAS due to their interrelationships with other risk factors. This illustrates the usefulness of DAGs in understanding the causal structure between risk factors in complex associations and identifying confounders beyond the traditional definition. A variable associated with both the exposure and the outcome, but not on the causal pathway, may be insufficient as a confounder (Cavanaugh, 2014).

Therefore, not all confounders from a given MSAS were incorporated in the adjusted analyses of all the types of cancer identified in the SR. This may be appropriate for studies exploring associations, but not for those making causal claims, where addressing confounding is necessary. The next chapter (Chapter 5: Assessing the quality of covariate adjustment in studies making causal claims) examined whether such studies clearly state their intent, justify their choice of covariates, and select them appropriately based on their objectives.

8.2.3 Assessing the quality of covariable adjustment in studies making causal claims (Chapter 5)

When making causal claims, it is important to explicitly state the objective of the analysis and explicitly justify covariate selection while ensuring consistency in terminology throughout so that the findings can be properly understood and used. However, identifying studies that claim a causal relationship between

T2DM and cancer is challenging due to inconsistencies in objectives, analytical methods, and terminology, which can lead to misinterpretation of findings. This chapter assessed whether studies making causal claims about T2DM and cancer explicitly stated their intent, justified covariate selection, and appropriately selected variables based on their objectives.

From the SR, fewer studies (42%) were classified as studies making causal claims (Explicit Causal Claim Studies) compared to those not making causal claims (No Explicit Causal Claim Studies) (58%).

Some studies assessed multiple cancer types, resulting in several effect estimates per study. Studies making causal claims were more likely to include at least one MSAS in their analyses (44% vs. 23%) compared to those not making such claims. This suggests alignment with their respective objectives: studies aiming for causal inference are expected to address confounding by including MSAS, whereas studies not making causal claims are not required to adjust for confounders.

Nearly 70% of examined cancers had at least one study explicitly claiming causality. However, the extent to which these studies accounted for confounding varied by cancer type. Apart from pancreatic and endometrial cancer, where a higher proportion of studies considered the MSAS, fewer than 20% of effect estimates in Explicit Causal Claim studies for other cancer types included the necessary covariates for proper adjustment. This indicates that, despite claiming causality, many studies did not sufficiently control for confounding variables.

Almost 70% of Explicit Causal Claim studies justified covariate selection, compared to one-third of No Explicit Causal Claim studies. This suggests that studies aiming to establish causality were more likely to align covariate selection with their research objectives. In contrast, studies that did not make causal claims often relied on statistical criteria or listed variables without considering their role in the causal pathway, which was consistent with their aim.

Some studies without explicit causal claims still provided causal effect estimates by accounting for the MSAS, possibly because authors were unaware that these

variables were sufficient to control for confounding. Another reason could be that journal policies restrict causal language in observational studies (Haber et al., 2022; M. A. Han et al., 2022).

Therefore, authors should clearly define study objectives, align methods accordingly, and use consistent terminology throughout the paper to ensure readers understand the practical implications of the findings. Reviewers should also recognize that observational studies can provide causal insights, supporting the acceptance of causal language in research.

8.2.4 Meta-analyses (Chapter 6)

This chapter presented adjusted (causal effect) meta-analyses of T2DM and cancer risk, using effect estimates from studies that accounted for the MSAS, derived from DAGs, in Chapter 4.

Meta-analyses were conducted for 17 cancer types. However, analyses could not be performed for 14 cancer types, as only a single effect estimate accounting for the MSAS was available for each.

T2DM was found to increase the risk of colon (30%), kidney (67%), and pancreatic (73%) cancer. In women, it increased the risk of breast and endometrial cancer by 11% and 94%, respectively. Conversely, T2DM reduced the risk of prostate cancer in men by 19%.

Subgroup analyses showed no statistically significant differences in cancer risk by sex, geographic region (proxy for race), or study design (case-control and cohort studies). Breast cancer risk due to T2DM was comparable in premenopausal and postmenopausal women, suggesting consistent effects across these groups.

More evidence is needed to confirm if cancer risk varies by time since T2DM diagnosis and by antidiabetic drug use. Analyses by duration of T2DM were only possible for bladder and thyroid cancers, as only two MSAS-adjusted studies were available for each. A significant difference between subgroups was found only

for thyroid cancer. Similarly, for drug impact, only breast and pancreatic cancer had enough data, with two MSAS-adjusted studies each, but no significant subgroup differences were observed.

Unadjusted (association) meta-analyses were also conducted to explore associations between T2DM and each cancer type. Compared to adjusted (causal effect) meta-analyses, the magnitude of associations was consistently larger than the causal effects (except for rectal and kidney cancers), suggesting that shared risk factors contributed to the observed relationships. These findings highlight the importance of confounder adjustment when estimating causality.

This chapter provided causally interpretable estimates of T2DM on various cancers. However, it also highlights the limited number of studies that accounted for the MSAS, which restricted analyses for certain cancer types, on the impact of diabetic treatments on cancer risk and the influence of T2DM duration on cancer incidence.

8.2.5 Burden of disease (Chapter 7)

This chapter quantified the global burden of cancer (across 177 countries grouped into seven regions) causally associated with T2DM using findings from the adjusted (causal effect) meta-analyses of breast, endometrial, colon, kidney, pancreatic, and prostate cancers.

Results were reported as proportions (population impact fraction [PIF] and prevented fraction of the population [PFP]) and the number of incident cancer cases and disability-adjusted life years (DALYs) attributable to or averted due to T2DM. Additionally, monetised DALYs were used to assess the economic impact.

In 2025, 105,084 cancer cases (2% of total cases for these cancer types) are expected to be attributable to T2DM prevalence in 2021. Endometrial cancer will have the highest burden, with 1 in 20 cases (24,072) attributable to T2DM, followed by pancreatic cancer, with 1 in 25 cases (23,530). Pancreatic cancer is also projected to account for the highest DALY and economic burden (INT \$21,844.70 million). These findings highlight the need for targeted interventions

to prevent T2DM and enhance early cancer detection, particularly for pancreatic and endometrial cancers, to reduce future health and economic burdens.

Regionally, the Middle East and North Africa (MENA) is projected to experience the highest burden (PIF 2.8%), while Africa will have the lowest (PIF 0.7%). Although Africa currently has the lowest cancer burden attributable to T2DM due to low diabetes prevalence, this is expected to change. By 2045, the number of diabetes cases in Africa is projected to increase by 129%, the highest growth rate globally (International Diabetes Federation, 2021). Additionally, the incidence of cancers causally associated with T2DM in Africa is expected to rise by 100% (Ferlay J, Laversanne M, Ervik M, n.d.). Consequently, the number of new cancer cases causally associated with T2DM is anticipated to increase sharply in this region, highlighting the urgent need for strategies to address this growing health challenge.

The Western Pacific region is projected to have the highest number of cancer cases and DALYs attributable to T2DM, with 41,750 cases (35% of all cases) and 603,505 DALYs lost (37% of total DALYs). This is mainly because the region includes China, which is a country with a large population.

The findings of this analysis also showed that if T2DM prevalence declines, prostate cancer cases due to T2DM may increase, as T2DM is a protective factor. However, reducing T2DM would prevent far more cases of breast, endometrial, colon, pancreatic, and kidney cancers, as well as T2DM-related microvascular and macrovascular complications.

Detecting T2DM early can help prevent complications and improve outcomes. However, since nearly 50% of T2DM cases remain undiagnosed, many T2DM attributable cancer cases cannot be recognized (Ogurtsova et al., 2022). As a result, as shown in the analysis considering only the number of cancer cases due to a confirmed T2DM diagnosis, half of the estimated cancer cases can be identified, reducing the projection from 105,084 to 52,642.

The findings of this global burden of disease analysis highlight the need for interventions to reduce T2DM prevalence, particularly in countries with high

rates of the disease. Raising awareness of T2DM diagnosis can help manage the condition, prevent complications, and support early cancer detection.

This analysis emphasizes the importance of reducing T2DM to lower the cancer burden attributable to it. However, the impact may be underestimated, as T2DM prevention strategies also target shared risk factors like obesity, which contribute to both T2DM and cancer. Addressing these factors could provide broader health and financial benefits beyond the results presented.

8.3 Strengths and Limitations

8.3.1 Strengths

The systematic review in Chapter 3 comprehensively identified evidence on the association between T2DM and cancer risk across multiple cancer types. A broad search across five databases was conducted without language or time restrictions, ensuring an exhaustive review that included various observational study designs. This approach resulted in 18,086 publications, with breast, pancreatic, colorectal, prostate, and lung cancers being the most frequently studied, aligning with the global cancer burden. Identifying all available evidence on this topic was essential, as it served as the basis for subsequent chapters and strengthened the thesis findings.

A strength of Chapter 4 is the systematic approach used to develop the DAGs. Targeted literature searches ensured that only risk factors reported in at least two authoritative sources (well-known institutions, medical societies, and public health organizations) were included, and all relationships between risk factors were supported by published clinical evidence. This approach strengthened the validity of the MSAS derived from the DAGs.

This was essential because the MSAS determined which studies were included in the adjusted (causal effect) meta-analyses in Chapter 6. By selecting only studies that accounted for the MSAS, the analyses provided less biased estimates and addressed the core research question of this thesis.

Additionally, the MSAS identified for each cancer type were used in Chapter 5 to assess whether studies making causal claims appropriately selected confounders in their adjusted analyses.

In Chapter 5, a unique flowchart was developed to identify studies making causal claims. This flowchart was based on terminology used in different manuscript sections and aligned with recommendations from methodological publications on causal language (M. A. Han et al., 2022; Olarte Parra et al., 2021). This approach ensured transparency, was evidence-based, and provided a reproducible framework for future research.

In Chapter 7, the burden of disease analysis included only cancers causally associated with T2DM allowing the effect of T2DM to be isolated from other common risk factors. The analysis used the most recent T2DM prevalence data from the IDF, which enhanced the accuracy of the results (International Diabetes Federation, 2021). Another strength was the holistic approach, which incorporated multiple burden of disease metrics (PIF/PFP, cancer cases, DALYs, and monetised DALYs) to quantify the health and economic impact of T2DM on cancer. This approach provided a better understanding of the cancer burden attributable to T2DM and supported the development of recommendations at the country, regional, and global levels.

8.3.2 Limitations

Although a comprehensive and systematic search was conducted, no studies from African countries examining the association between T2DM and cancer were identified in the systematic review (Chapter 3). As a result, the findings may not be generalizable to these populations. However, for the burden of disease analysis, it was assumed that the adjusted (causal effect) estimates from the meta-analyses could be applied to African countries, given that Africa is projected to experience the highest increase in diabetes cases by 2045 (International Diabetes Federation, 2021). This assumption was made to reflect the potential impact of reducing diabetes prevalence on cancer burden in the region.

Another limitation of the systematic review is the classification of diabetes types, as some studies did not differentiate between type 1 and type 2 diabetes. To ensure relevant data on T2DM was included, individuals diagnosed with diabetes after age 30 were assumed to have type 2 diabetes, while those diagnosed before age 30 were classified as type 1 and excluded (Rodbard et al., 2007; World Health Organization, 1994). Although this approach allowed for the inclusion of studies without explicit diabetes classification, it may have introduced misclassification bias. Nonetheless, as type 2 accounts for approximately 90% of diabetes cases, the proportion of individuals with type 1 is likely to be minimal.

The systematic review was also restricted to published studies, potentially excluding relevant findings from unpublished or grey literature.

Additionally, for the systematic review, the selection of cancer types was based on a scoping search. The systematic search strategy included only those cancer types identified during this scoping search. This approach was chosen because an initial search in just two databases (PubMed and Embase), without limiting the cancer types, returned over 70,000 records, making the review unmanageable. However, additional cancer types were later identified during the screening stages of the systematic review by reviewing the reference lists of other systematic reviews and meta-analyses. Due to time constraints, a systematic search was not conducted for these additional cancers, meaning some evidence on their association with T2DM may have been missed. Most of these additional cancer types are considered rare. Nevertheless, their data were included in the systematic review findings and subsequent analyses, as they may offer valuable insights for future research.

For example, in Chapter 4 (Causal Association), after identifying the MSAS for each cancer type, the studies were assessed to determine whether these confounders were included in their adjusted analyses. It was found that for 25 cancer types, the MSAS was not accounted for. Of these, only six types (Kaposi's sarcoma, mesothelioma, myeloma, nasal, penile, and cancer of unknown primary [CUP]) were among those for which a systematic search had not been conducted.

Similarly, in Chapter 6 (Meta-Analyses), no statistically significant causal effect was found for bone, brain, eye, and testicular cancers, which were also among the types not systematically searched, indicating a need for further investigation. Additionally, for pharyngeal cancer (type of cancer without systematic search), only one study accounted for MSAS. Although it showed a causal association (OR 10.55, 95% CI: 2.95-37.67), the wide confidence interval suggested uncertainty.

Therefore, although a systematic search was not conducted for these additional cancer types, which are mostly rare, the systematic review in this thesis included a wide range of cancers, covering those with the highest incidence and mortality worldwide. As a result, the findings are valid and support the main research objective: to estimate the causal effect of T2DM on cancer risk.

One limitation of the burden of disease analysis is that DALYs were reported as rates per 100,000 adults aged 20 years and older in 2019 using the GBD Compare tool. Since the study focuses on 2025, applying these rates to 2025 population projections may have led to either overestimation or underestimation of DALYs for cancers due to T2DM. Additionally, the GBD Compare tool reports DALYs for colorectal cancer (CRC) as a whole, including rectal cancer, whereas this thesis only considered colon cancer, which may have further contributed to overestimation. Lastly, sex-specific diabetes prevalence data were unavailable, so men and women were assumed to have comparable rates, potentially introducing bias into the global burden of disease estimates.

8.4 Future Research and Policy Recommendations

Causal Association (Chapter 4)

Establishing causal claims in observational studies requires addressing confounding bias. DAGs have been recommended for identifying the MSAS of covariates that could distort the relationship (Tennant et al., 2021). However, DAGs in published studies are often based on the variables available in the dataset, which can limit their ability to accurately identify the true MSAS.

To improve the validity of DAGs, it is recommended that the variables included and the causal assumptions depicted be based on clinical evidence from reliable sources, such as well-known institutions, medical societies, and public health organizations.

For exposures and outcomes with risk factors (both non-modifiable and modifiable), such as T2DM and cancer, the variables needed to control for confounding may include both non-modifiable and modifiable risk factors, particularly lifestyle-related factors. Therefore, it is essential to identify the MSAS a priori to ensure that all relevant confounders are collected for estimating the independent effect of the exposure being assessed.

In addition to guiding statistical analysis, DAGs are also valuable for planning data collection. Identifying in advance which variables need to be collected helps ensure the study aligns with its objectives and avoids unnecessary data collection.

Assessing the quality of covariable adjustment in studies making causal claims (Chapter 5)

In studies making causal claims, many lifestyle-related confounders (e.g., diet, smoking, and alcohol consumption) were often omitted from the MSAS in adjusted analyses across various cancer types. This limited their ability to properly address confounding. In most cases, authors acknowledged that this was due to data limitations.

To ensure alignment with study objectives, researchers should carefully plan data collection to include relevant confounding variables. Alternatively, if such data are not available, researchers should explore additional data sources.

Clear and consistent language is necessary for correct interpretation of findings. However, identifying studies that make causal claims using observational designs can be difficult, as authors often avoid using language that explicitly denote causal relationships. This may be due to journal policies which discourage causal language in non-randomized studies (Haber et al., 2022; M. A. Han et al., 2022).

Reviewers should recognize that observational studies can provide causal insights and encourage the appropriate use of causal terminology. Authors should maintain consistent language that reflects the study's objectives throughout the manuscript.

Meta-analyses (Chapter 6)

Epidemiological research suggests that cancer risk may vary with the duration of T2DM and antidiabetic drug classes (Y. Hu, Wu, et al., 2021; H. J. Kim et al., 2018; Y.-C. Liu et al., 2019).

Subgroup analyses exploring time since T2DM diagnosis and the impact of antidiabetic drug use were planned but not feasible due to limited data. Future studies that adjust for confounders and aim to estimate the causal effect of diabetes duration and treatment on cancer risk are needed to better understand the causal relationship between T2DM and cancer.

Burden of disease analysis (Chapter 7)

Detecting T2DM early can help prevent complications and improve outcomes. However, since nearly 50% of T2DM cases remain undiagnosed, many T2DM attributable cancer cases may not be recognized (Ogurtsova et al., 2022). This emphasizes the need for improved health programs for early T2DM detection, which can lead to earlier cancer diagnoses and better management of related T2DM related comorbidities. This is particularly urgent in Africa, the Western Pacific, and Southeast Asia, where over 50% of T2DM cases remain undiagnosed.

In this global burden of disease analysis, DALYs were monetised using GDP per capita to assess the economic impact of cancer attributable to T2DM. However, to support more informed policy decisions at the country level, future analyses should incorporate country-specific data. Including the costs of treatments, interventions, and indirect costs (e.g., productivity loss) would offer policymakers a more complete framework for resource allocation based on national needs.

Additional recommendations

Findings from the adjusted (causal effect) meta-analyses revealed that T2DM is causally associated with the risk of breast, endometrial, colon, kidney, pancreatic, and prostate cancer, independently of confounding factors. Therefore, prediction models for T2DM should consider these cancers as potential complications, given the limited evidence available on this topic.

A scoping review (Ndjaboue et al., 2022), which included studies published between 2000 and 2020, identified 260 risk prediction models for complications in people with diabetes. Of these, only one model included cancer risk in individuals with T2DM, based on a Chinese prospective cohort (X. Yang et al., 2008). Additionally, only a few prediction models based on Chinese and Taiwanese populations have been developed to assess the risk of specific cancers, such as bladder, pancreatic, colorectal, liver, and breast cancer, in patients with T2DM (S. Chen et al., 2023; M. H. Hsieh et al., 2018; M.-H. Hsieh et al., 2018, 2019; Qiu et al., 2024; Rau et al., 2016; Yau et al., 2024).

From a health economics perspective, another scoping review focusing on economic models for T2DM identified 34 models in studies published between 1997 and 2022 (X. Li et al., 2023). Aside from the Archimedes model, which simulates pathways for multiple diseases, including diabetes and cancer (colon, breast, and lung), none explicitly listed cancer as a complication of T2DM.

To capture the full impact of T2DM, it is recommended that T2DM models include the risk of these cancers.

8.5 Contribution to knowledge

The main contribution of this thesis is the findings of causally interpretable meta-analyses examining the relationship between T2DM and cancer risk. This thesis identifies specific cancer types (breast, endometrial, colon, kidney, pancreatic, and prostate cancers) that are causally associated with T2DM.

This thesis provided several additional contributions:

First, a comprehensive and updated systematic review was conducted (Chapter 3) to identify all available evidence on the association between T2DM and cancer risk across multiple cancer types. Unlike previous reviews, this search used five databases without language or time restrictions and included various observational study designs. The included studies differed in design, geographical setting, confounder adjustment, and whether they sought to establish causality.

Second, this thesis introduced the use of DAGs as a novel and systematic tool for identifying confounders between T2DM and 56 cancer types (Chapter 4). The DAGs were developed using risk factors supported by clinical evidence, thereby strengthening the validity of the MSAS. The identified confounders were both non-modifiable and modifiable.

Third, the thesis showed that although some evidence supports a causal link between T2DM and many cancer types, many studies did not address confounding. Fewer than 20% of effect estimates in studies making causal claims accounted for the MSAS, suggesting that many studies were not aligned with their stated objectives and did not adequately adjust for confounding.

Finally, Chapter 7 presented a global burden of disease analysis quantifying the health and economic impact of cancers causally associated with T2DM. It estimated that in 2025, 105,084 cancer cases (2% of total cases for these cancer types) would be attributable to T2DM prevalence in 2021.

8.6 Conclusion

The research question addressed in this thesis was: *What is the independent causal effect of T2DM on cancer incidence?*

This thesis provides evidence that T2DM has an independent causal effect on increasing the risk of breast, colon, endometrial, kidney, and pancreatic cancers. In contrast, T2DM appears to have a protective causal effect against prostate cancer.

Identifying the types of cancer causally associated to T2DM can support the development of more targeted public health interventions, particularly in cancer screening and prevention strategies for people with T2DM. At a global level, the burden of disease analysis further emphasized the importance of preventing T2DM and improving early cancer detection, especially for pancreatic and endometrial cancers, to reduce the future health and economic burden of cancer due to T2DM.

A novel aspect of this thesis was the use of DAGs to identify the MSAS required to control for confounding in the relationship between T2DM and various cancer types.

The research was grounded in a comprehensive, updated, and systematic review of the literature to ensure that all relevant evidence was considered.

This thesis provides a framework for causal inference in T2DM and cancer research and contributes with new evidence to inform research and public health strategies.

Appendices

Appendix A. Systematic Review

This appendix provides supplementary information for Chapter 3: Systematic Review.

Appendix B. Causal association

This appendix provides supplementary information for Chapter 4: Causal association

Appendix C. Meta-analyses

This appendix provides supplementary information for Chapter 6: Meta-analyses.

Appendix D. Burden of disease

This appendix provides supplementary information for Chapter 7: Burden of Disease.

Appendix A. Systematic Review

This appendix provides supplementary information for Chapter 3: Systematic Review.

A1. Scoping search to identify the types of cancer associated with T2DM

A1.1 Objective

This search aimed to identify the types of cancer associated with T2DM.

A1.2 Materials and methods

A search was conducted in PubMed from inception to February 17th, 2021, to find the types of cancer whose incidence is related to T2DM. Text words for diabetes and cancer were combined and specified in each strategy. It is important to note that the type of diabetes was not defined in order to compile all articles on the subject. However, during the screening process, additional emphasis was placed on type 2 diabetes, and only cancers associated with this type of diabetes were collected.

Only meta-analysis and systematic reviews were considered at the first stage, but to ensure all the possible types of cancer-related to diabetes, a further search on narrative reviews was performed.

All languages were accepted, and no exclusion criteria were applied to geographical location. Only studies involving human subjects were included.

According to the rationale of the review, which aimed to identify cancers associated with T2DM, the screening process was carried out in one stage (titles and abstracts) for all retrieved records. A second stage (full-text screening) was conducted only for meta-analyses, systematic reviews, and narrative reviews that reported multiple cancer types.

Two search strategies were developed, one for meta-analysis and systematic reviews, and another for narrative reviews.

Table A1-1 Search strategies to identify the types of cancer-related to T2DM

Search strategies
("cancer"[Title] AND "diabetes"[Title]) AND ((meta-analysis[Filter] OR systematicreview[Filter]) AND (humans[Filter]))
("cancer"[Title] AND "diabetes"[Title]) AND ((review[Filter]) AND (humans[Filter]))

A1.3 Results

The search strategy for meta-analysis and systematic reviews retrieved 419 records, from which 78 were selected after reading the titles/abstracts. The following graph shows the percentage of studies categorized by cancer type.

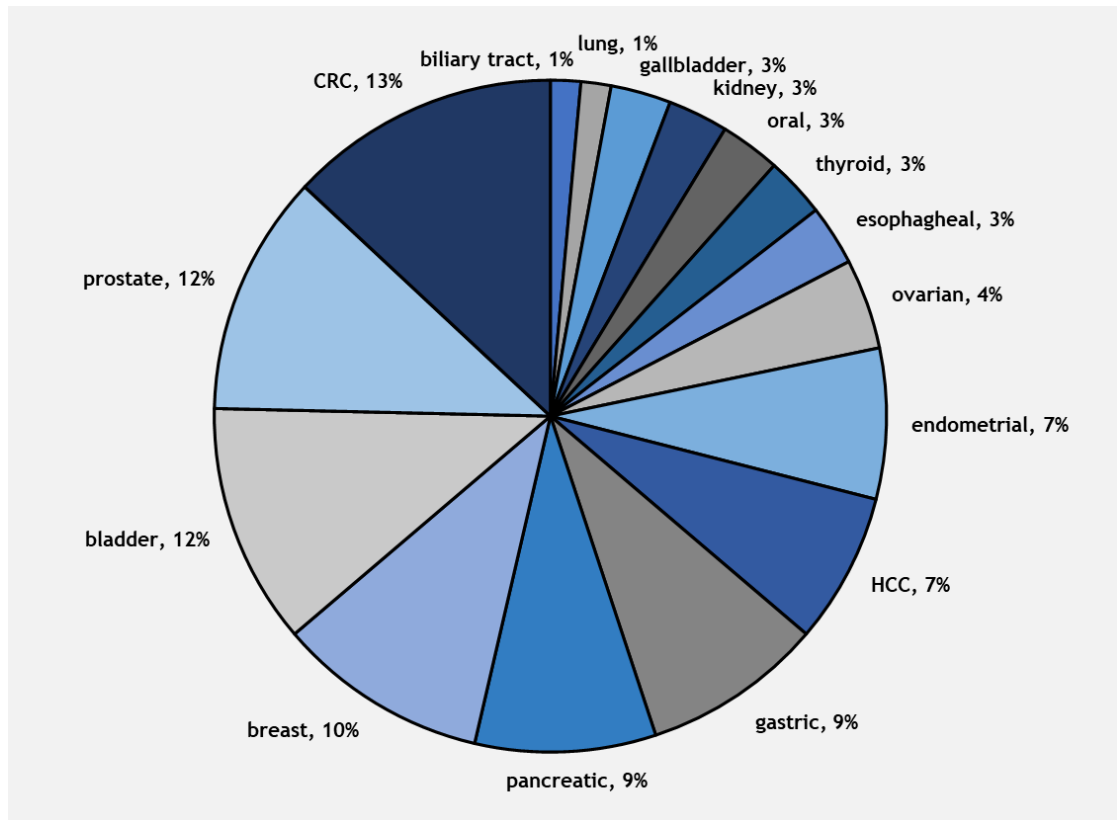


Figure A1-1 Percentage of studies by cancer type from the meta-analyses and systematic reviews on T2DM and cancer risk. *Abbreviations.* HCC: hepatocellular carcinoma. Biliary tract cancer includes bile ducts.

The graph illustrates the 16 types of cancer (and their subsites) with evidence from meta-analyses and systematic reviews regarding their association with T2DM: biliary tract, bladder, breast, colorectal, endometrial, gallbladder, HCC, kidney, lung, oesophagus, oral, ovarian, pancreas, prostate, stomach (gastric), and thyroid. Breast, bladder, colorectal and prostate cancers have more publications on the topic.

In addition, 9 out of the 78 meta-analyses and systematic reviews examined multiple types of cancer within the same publication. Among these, leukemia, lymphoma, cervical cancer (cervix uteri), NHL, intrahepatic cholangiocarcinoma, and bile duct cancer were also identified.

The search strategy for narrative reviews retrieved 419 papers. After the title/abstract screening, 37 papers were selected, and a sample of 20 papers was considered for full-text screening. The graph below shows the types of cancer related to T2DM and their respective percentages.

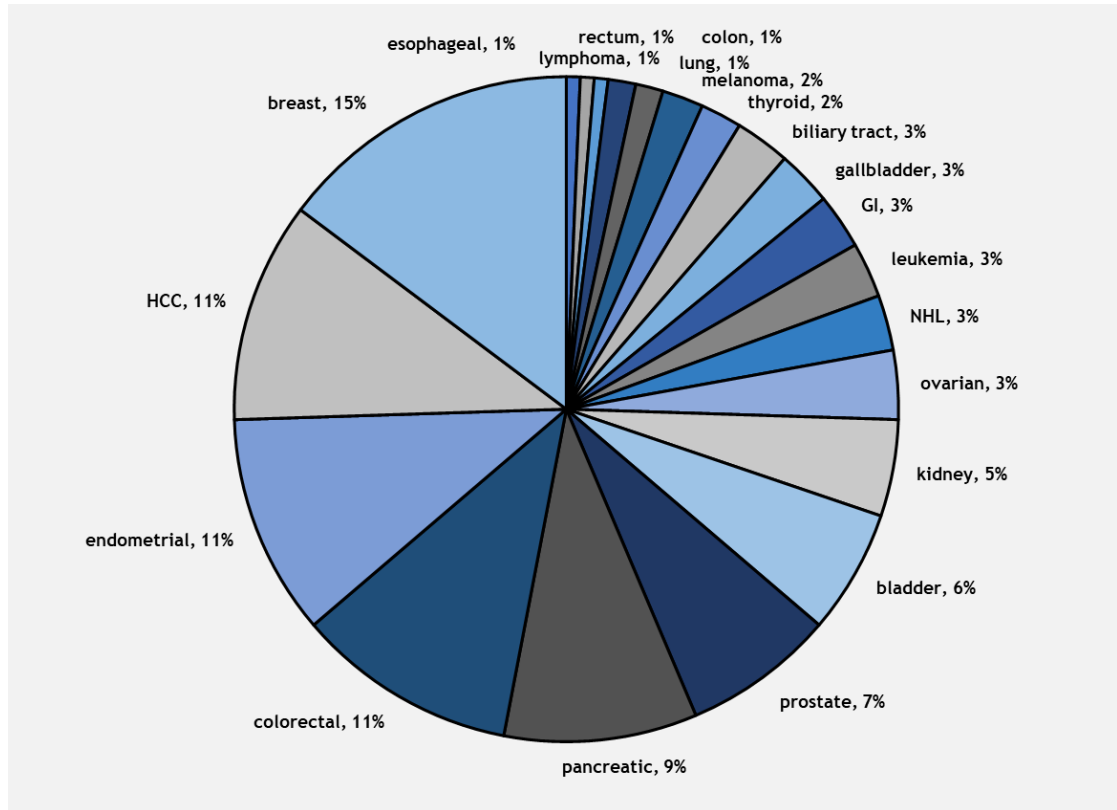


Figure A1- 2 Percentage of studies by cancer type from narrative reviews on T2DM and cancer risk. *Abbreviations.* GI: gastrointestinal; HCC: hepatocellular carcinoma; NHL: Non-Hodgkin lymphoma. Biliary tract cancer includes bile ducts.

The 21 types of cancer (and their subsites) associated with T2DM, identified from narrative reviews, were biliary tract, bladder, breast, colon, colorectal, endometrial, esophageal, gallbladder, GI, HCC, kidney, lymphoma, leukemia, lung, melanoma, NHL, ovarian, pancreatic, prostate, rectum, and thyroid. Colorectal, endometrial, HCC, and breast cancers had the most publications on the topic.

As a result, narrative reviews also identified melanoma as an additional type of cancer associated with T2DM.

A1.4 Conclusions

From the scoping search, 18 types of cancer (and their subsites) whose incidence has been related to T2DM were identified.

1. Bladder
2. Breast
3. Biliary tract cancer (BTC), including (bile duct, gallbladder)
4. Cervical cancer
5. Colorectal cancer (CRC), including colon and rectum
6. Endometrial
7. Gastrointestinal cancer (GI), including gastric (stomach), esophageal
8. Kidney
9. Leukemia
10. Liver, including HCC (hepatocellular cancer)
11. Lung
12. Lymphoma
13. Skin
14. Oral
15. Ovarian
16. Pancreatic
17. Prostate
18. Thyroid

From the list, the **most common** types of cancer include those related to the hepatobiliary system: liver cancer (LVC), comprising hepatocellular carcinoma (HCC), and biliary tract cancer (BTC), which includes bile duct (BD) and gallbladder (GB). Other common cancers include bladder cancer (BlaC), breast cancer (BC), colorectal cancer (CRC), endometrial cancer (EndC), gastrointestinal cancer (GI), kidney cancer (KC), ovarian cancer (OC), pancreatic cancer (PC), and prostate cancer (PrC).

Less common cancers studied in relation to this topic include cervical cancer (CeuC), leukemia (Le), lymphoma (Ly), lung cancer (LC), melanoma (M), oral cancer (Oral), and thyroid cancer (TC).

A2. Systematic Review. Terms and search strategies

The terms and search strategies for the SR are described in this section. The following text words were shared between the five databases with MeSH terms introduced for each database except WOS, which does not use this terminology.

Type 2 diabetes: adult-onset diabetes, diabetes mellitus type 2, diabetes type 2, DM2, ketoacidosis, ketoacidosis-resistant diabetes, ketosis-resistant diabetes, maturity-onset diabetes, NIDDM, non-insulin dependen*, noninsulin dependen*, non-insulin-dependen*, T2D, T2DM, type 2 diabet*, type II diabet*

Cancer: adenocarcinoma, cancer, carcino*, malignan*, metastas*, neoplas*, tumor*, tumour*

Risk/Incidence: associat*, diagnos*, factor*, incidence, predict*, predispos*, relationship*, risk*, risk factor*

RCT: randomised controlled trial*

Types of cancer

Most common cited types

- Bladder: bladder, transitional cell, transitional urothelial, urinary bladder, urothel*, vesica urinaria
- Breast: breast, mammary
- Colorectal: colon, colonic, colorectal, rectal, rectum
- Endometrial: corpus uteri, decidu*, endometri*, myometri*, pelvi*, uteri*, uterus
- Gastrointestinal: stomach, gastric, gastroint*, intestin*, esophag*
- HCC, BTC (Hepatobiliary cancers): biliary duct, gallbladder, liver, ampulla of Vater, biliary, bile, gallbladder, hepatic, hepatocellular, intrahepatic, extrahepatic, liver, intrahepatic cholangiocarcinoma, extrahepatic cholangiocarcinoma
- Kidney: kidney, renal cell
- Ovarian: ovarian, ovary
- Pancreatic: pancrea*
- Prostate: prostat*, Pca

Less Common cited types

- Cervix: cervix, cervic*
- Leukemia: leukem*
- Lymphoma: lymph*, lymphoma*, NHL
- Lung: lung
- Melanoma: skin, melanoma
- Oral: mouth, oral squamous cell, oral
- Thyroid: thyroid

A2.1 CINAHL search strategy

A2.1.1 MeSH Terms

Type 2 diabetes: MH "Diabetes Mellitus, Type 2"

Cancer: MH "Neoplasms+"

Risk/Incidence: MH "Risk Factors+", MH "Incidence"

RCT: MH "Randomized Controlled Trials+"

Types of cancer:

Most common cited types

- Bladder: MH "Bladder"
- Breast: MH "Breast+"
- Colorectal: MH "Colon+", MH "Rectum+"
- Endometrial: MH "Endometrium"
- Gastrointestinal: MH "Gastrointestinal System+"
- HCC, BTC (Hepatobiliary cancers): (MH "Biliary Tract+") OR (MH "Bile Ducts") OR (MH "Gallbladder") OR (MH "Liver")
- Kidney: MH "Kidney+"
- Ovarian: MH "Ovary+"
- Pancreatic: MH "Pancreas"
- Prostate: MH "Prostate"

Less common cited types

- Cervix: MH "Cervix"
- Lung: MH "Lung+"
- Melanoma: MH "Skin+"
- Oral: MH "Mouth+"
- Thyroid: MH "Thyroid Gland"

Types of cancer (full term)

Most common cited types

- Bladder: MH "Bladder Neoplasms"
- Breast: MH "Breast Neoplasms+"
- Colorectal: (MH "Colorectal Neoplasms+")
- Endometrial: MH "Endometrial Neoplasms"
- Gastrointestinal: MH "Gastrointestinal Neoplasms+"
- HCC, BTC (Hepatobiliary cancers): (MH "Biliary Tract Neoplasms") OR (MH "Bile Duct Neoplasms") OR (MH "Gallbladder Neoplasms") OR (MH "Liver Neoplasms")
- Kidney: (MH "Kidney Neoplasms+")
- Ovarian: (MH "Ovarian Neoplasms+")
- Pancreatic: (MH "Pancreatic Neoplasms+")
- Prostate: (MH "Prostatic Neoplasms+")

Less common cited types

- Cervix: (MH "Cervix Neoplasms+")
- Leukemia: (MH "Leukemia+")
- Lymphoma: (MH "Lymphoma+")
- Lung: (MH "Lung Neoplasms+")
- Melanoma: (MH "Melanoma+")
- Oral: (MH "Mouth Neoplasms+")
- Thyroid: (MH "Thyroid Neoplasms+")

A2.1.2 Search Strategy

Table A2-1 Search strategy for CINAHL

No.	Keyword searches
1	TI ("adult-onset diabetes" OR "diabetes mellitus type 2" OR "diabetes type 2" OR DM2 OR ketoacidosis OR "ketoacidosis-resistant diabetes" OR "ketosis-resistant diabetes" OR "maturity-onset diabetes" OR NIDDM OR "non-insulin dependen*" OR "noninsulin dependen*" OR "non-insulin-dependen*" OR T2D OR T2DM OR "type 2 diabet*" OR "type II diabet*") OR AB ("adult-onset diabetes" OR "diabetes mellitus type 2" OR "diabetes type 2" OR DM2 OR ketoacidosis OR "ketoacidosis-resistant diabetes" OR "ketosis-resistant diabetes" OR "maturity-onset diabetes" OR NIDDM OR "non-insulin dependen*" OR "noninsulin dependen*" OR "non-insulin-dependen*" OR T2D OR T2DM OR "type 2 diabet*" OR "type II diabet*")
2	MH "Diabetes Mellitus, Type 2"
3	S1 OR S2
4	TI (associat* OR diagnos*OR factor* OR incidence OR predict*OR predispos* OR relationship* OR risk* OR "risk factor*") OR AB (associat* OR diagnos*OR factor* OR incidence OR predict*OR predispos* OR relationship* OR risk* OR "risk factor*")
5	MH "Risk Factors+" OR MH "Incidence"
6	S4 OR S5
7	TI (Bladder OR "transitional cell" OR "transitional urothelial" OR "urinary bladder" OR urothel* OR "vesica urinaria" OR breast OR mammary OR colon OR colonic OR colorectal OR rectal OR rectum OR "corpus endometri" OR decidu* OR endometri* OR myometri* OR pelvi* OR uteri* OR uterus) OR AB (Bladder OR "transitional cell" OR "transitional urothelial" OR "urinary bladder" OR urothel* OR "vesica urinaria" OR breast OR mammary OR colon OR colonic OR colorectal OR rectal OR rectum OR "corpus endometri" OR decidu* OR endometri* OR myometri* OR pelvi* OR uteri* OR uterus)
8	TI (stomach OR gastric OR gastroint* OR intestin* OR esophag* OR "ampulla of Vater" OR biliary OR bile OR gallbladder OR hepatic OR hepatocellular OR intrahepatic OR extrahepatic OR liver OR kidney OR renal cell OR ovarian OR ovary OR pancrea* OR prostat* OR Pca) OR AB (stomach OR gastric OR gastroint* OR intestin* OR esophag* OR "ampulla of Vater" OR biliary OR bile OR gallbladder OR hepatic OR hepatocellular OR intrahepatic OR extrahepatic OR liver OR kidney OR renal cell OR ovarian OR ovary OR pancrea* OR prostat* OR Pca)
9	TI (cervix OR cervic* OR lymph* OR lung OR skin OR mouth OR "oral squamous cell" OR oral OR thyroid) OR AB (cervix OR cervic* OR lymph* OR lung OR skin OR mouth OR "oral squamous cell" OR oral OR thyroid)
10	7 OR 8 OR 9
11	TI (adenocarcinoma OR cancer OR carcino* OR malignan* OR metastas* OR neoplas* OR tumor* OR tumour*) OR AB (adenocarcinoma OR cancer OR carcino* OR malignan* OR metastas* OR neoplas* OR tumor* OR tumour*)
12	S10 AND S11
13	TI ("intrahepatic cholangiocarcinoma" OR "extrahepatic cholangiocarcinoma" OR leukem* OR lymphoma* OR NHL OR melanoma) OR AB ("intrahepatic cholangiocarcinoma" OR "extrahepatic cholangiocarcinoma" OR leukem* OR lymphoma* OR NHL OR melanoma)
14	S12 OR S13
15	MH "Bladder" OR MH "Breast+" OR MH "Colon+" OR MH "Rectum+" OR MH "Endometrium" OR (MH "Gastrointestinal System+") OR (MH "Biliary Tract+") OR (MH "Bile Ducts") OR (MH

	"Gallbladder") OR (MH "Liver") OR MH "Kidney+" OR MH "Ovary+" OR MH "Pancreas" OR MH "Prostate" OR MH "Cervix" OR MH "Lung+" OR MH "Skin+" OR MH "Mouth+" OR MH "Thyroid Gland"
16	MH "Neoplasms+"
17	S15 AND S16
18	(MH "Bladder Neoplasms") OR (MH "Breast Neoplasms+") OR (MH "Colorectal Neoplasms+") OR (MH "Endometrial Neoplasms") OR (MH "Gastrointestinal Neoplasms+") OR (MH "Biliary Tract Neoplasms") OR (MH "Bile Duct Neoplasms") OR (MH "Gallbladder Neoplasms") OR (MH "Liver Neoplasms") OR (MH "Kidney Neoplasms+") OR (MH "Ovarian Neoplasms+") OR (MH "Pancreatic Neoplasms+") OR (MH "Prostatic Neoplasms+") OR (MH "Cervix Neoplasms+") OR (MH "Leukemia+") OR (MH "Lymphoma+") OR (MH "Lung Neoplasms+") OR (MH "Melanoma+") OR (MH "Mouth Neoplasms+") OR (MH "Thyroid Neoplasms+")
19	S14 OR S17 OR S18
20	S3 AND S6 AND S19
21	TI ("randomized controlled trial*") OR AB ("randomized controlled trial*")
22	MH "Randomized Controlled Trials+"
23	S21 OR S22
24	S20 NOT S23

"AB" indicates term in the abstract field; "TI" indicates term in the title field; "*" indicates truncation to find multiple endings of the word; quotes (") around words indicate exact search for the entire phrase; "MH" indicates the exact subject heading, searching for major and minor headings; "+" indicates to explode the heading, so it retrieves references of the selected subject heading and more specific headings below; "OR" indicates similar keywords; "AND" links the concepts; "NOT" excludes specific words.

A2.2 COCHRANE search strategy

A2.2.1 MeSH terms

Type 2 diabetes: MeSH descriptor: [Diabetes Mellitus, Type 2] explode all trees

Cancer: MeSH descriptor: [Neoplasms] explode all trees

Risk/Incidence: MeSH descriptor: [Risk Factors] explode all trees, MeSH descriptor: [Incidence] explode all trees

RCT: MeSH descriptor: [Randomized Controlled Trial] explode all trees

Types of cancer:

Most common cited types

- Bladder: MeSH descriptor: [Urinary Bladder] explode all trees
- Breast: MeSH descriptor: [Breast] explode all trees
- Colorectal: MeSH descriptor: [Colon] explode all trees, MeSH descriptor: [Rectum] explode all trees
- Endometrial: MeSH descriptor: [Endometrium] explode all trees
- Gastrointestinal: MeSH descriptor: [Upper Gastrointestinal Tract] explode all trees
- HCC, BTC (Hepatobiliary cancers): MeSH descriptor: [Biliary Tract] explode all trees, MeSH descriptor: [Liver] explode all trees
- Kidney: MeSH descriptor: [Kidney] explode all trees
- Ovarian: MeSH descriptor: [Ovary] explode all trees
- Pancreatic: MeSH descriptor: [Pancreas] explode all trees
- Prostate: MeSH descriptor: [Prostate] explode all trees

Less common cited types

- Cervix: MeSH descriptor: [Cervix Uteri] explode all trees
- Lung: MeSH descriptor: [Lung] explode all trees

- Melanoma: MeSH descriptor: [Skin] explode all trees
- Oral: MeSH descriptor: [Mouth] explode all trees
- Thyroid: MeSH descriptor: [Thyroid Gland] explode all trees

Types of cancer (full term)

Most common cited types

- Bladder: MeSH descriptor: [Urinary Bladder Neoplasms] explode all trees
- Breast: MeSH descriptor: [Breast Neoplasms] explode all trees
- Colorectal: MeSH descriptor: [Colorectal Neoplasms] explode all trees
- Endometrial: MeSH descriptor: [Endometrial Neoplasms] explode all trees
- Gastrointestinal: MeSH descriptor: [Gastrointestinal Neoplasms] explode all trees
- HCC, BTC (Hepatobiliary cancers): MeSH descriptor: [Biliary Tract Neoplasms] explode all trees, MeSH descriptor: [Liver Neoplasms] explode all trees
- Kidney: MeSH descriptor: [Kidney Neoplasms] explode all trees
- Ovarian: MeSH descriptor: [Ovarian Neoplasms] explode all trees
- Pancreatic: MeSH descriptor: [Pancreatic Neoplasms] explode all trees
- Prostate: MeSH descriptor: [Prostatic Neoplasms] explode all trees

Less common cited types

- Cervix: MeSH descriptor: [Uterine Cervical Neoplasms] explode all trees
- Leukemia: MeSH descriptor: [Leukemia] explode all trees
- Lymphoma: MeSH descriptor: [Lymphoma] explode all trees
- Lung: MeSH descriptor: [Lung Neoplasms] explode all trees
- Melanoma: MeSH descriptor: [Melanoma] explode all trees
- Oral: MeSH descriptor: [Mouth Neoplasms] explode all trees
- Thyroid: MeSH descriptor: [Thyroid Neoplasms] explode all trees

A2.2.2 Search strategy

Table A2-2 Search strategy for COCHRANE

No.	Keyword searches
1	((bladder OR "transitional cell" OR "transitional urothelial" OR "urinary bladder" OR urothel* OR "vesica urinaria" OR breast OR mammary OR colon OR colonic OR colorectal OR rectal OR rectum OR "corpus uteri" OR decidu* OR endometri* OR myometri* OR pelvi* OR uteri* OR uterus OR stomach OR gastric OR gastroint* OR intestin* OR esophag* OR "ampulla of Vater" OR biliary OR bile OR gallbladder OR hepatic OR hepatocellular OR intrahepatic OR extrahepatic OR liver OR kidney OR "renal cell" OR ovarian OR ovary OR pancrea* OR prostat* OR Pca OR cervix OR cervic* OR lymph* OR lung OR skin OR mouth OR "oral squamous cell" OR oral OR thyroid)):ti,ab,kw
2	(adenocarcinoma OR cancer OR carcino* OR malignan* OR metastas* OR neoplas* OR tumor* OR tumour*):ti,ab,kw
3	#1 AND #2
4	("intrahepatic cholangiocarcinoma" OR "extrahepatic cholangiocarcinoma" OR leukem* OR lymphoma* OR NHL OR melanoma):ti,ab,kw
5	#3 OR #4
6	MeSH descriptor: [Urinary Bladder] explode all trees

7	MeSH descriptor: [Breast] explode all trees
8	MeSH descriptor: [Colon] explode all trees
9	MeSH descriptor: [Rectum] explode all trees
10	MeSH descriptor: [Endometrium] explode all trees
11	MeSH descriptor: [Upper Gastrointestinal Tract] explode all trees
12	MeSH descriptor: [Biliary Tract] explode all trees
13	MeSH descriptor: [Liver] explode all trees
14	MeSH descriptor: [Kidney] explode all trees
15	MeSH descriptor: [Ovary] explode all trees
16	MeSH descriptor: [Pancreas] explode all trees
17	MeSH descriptor: [Prostate] explode all trees
18	MeSH descriptor: [Cervix Uteri] explode all trees
19	MeSH descriptor: [Lung] explode all trees
20	MeSH descriptor: [Skin] explode all trees
21	MeSH descriptor: [Mouth] explode all trees
22	MeSH descriptor: [Thyroid Gland] explode all trees
23	#6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22
24	MeSH descriptor: [Neoplasms] explode all trees
25	#23 AND # 24
26	MeSH descriptor: [Urinary Bladder Neoplasms] explode all trees
27	MeSH descriptor: [Breast Neoplasms] explode all trees
28	MeSH descriptor: [Colorectal Neoplasms] explode all trees
29	MeSH descriptor: [Endometrial Neoplasms] explode all trees
30	MeSH descriptor: [Gastrointestinal Neoplasms] explode all trees
31	MeSH descriptor: [Biliary Tract Neoplasms] explode all trees
32	MeSH descriptor: [Liver Neoplasms] explode all trees
33	MeSH descriptor: [Kidney Neoplasms] explode all trees
34	MeSH descriptor: [Ovarian Neoplasms] explode all trees
35	MeSH descriptor: [Pancreatic Neoplasms] explode all trees
36	MeSH descriptor: [Prostatic Neoplasms] explode all trees
37	MeSH descriptor: [Uterine Cervical Neoplasms] explode all trees
38	MeSH descriptor: [Leukemia] explode all trees
39	MeSH descriptor: [Lymphoma] explode all trees
40	MeSH descriptor: [Lung Neoplasms] explode all trees
41	MeSH descriptor: [Melanoma] explode all trees
42	MeSH descriptor: [Mouth Neoplasms] explode all trees
43	MeSH descriptor: [Thyroid Neoplasms] explode all trees
44	#26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40 OR #41 OR #42 OR #43
45	#5 OR #25 OR #44 ((“adult-onset diabetes” OR “diabetes mellitus type 2” OR “diabetes type 2” OR DM2 OR ketoacidosis OR “ketoacidosis-resistant diabetes” OR “ketosis-resistant diabetes” OR “maturity-onset diabetes” OR “NIDDM” OR (non-insulin NEXT dependen*) OR (noninsulin NEXT dependen*) OR (non NEXT insulin NEXT dependen*) OR T2D OR T2DM OR (type NEXT 2 NEXT diabet*) OR (type NEXT II NEXTdiabet*))) :ti,ab,kw
46	MeSH descriptor: [Diabetes Mellitus, Type 2] explode all trees
47	#46 OR # 47
48	((associat* OR diagnos*OR factor* OR incidence OR predict*OR predispos* OR relationship* OR risk* OR (risk NEXT factor*))) :ti,ab,kw
49	MeSH descriptor: [Risk Factors] explode all trees
50	MeSH descriptor: [Incidence] explode all trees
51	#50 OR #51
52	#49 OR #52
53	#45 AND #48 AND #53
54	(randomized NEXT controlled NEXT trial*) in Trials
55	MeSH descriptor: [Randomized Controlled Trial] explode all trees
56	#55 OR #56 in Cochrane Reviews, Cochrane Protocols, Trials
57	#54 NOT #57 in Cochrane Reviews, Cochrane Protocols, Trials
58	child*
59	MeSH descriptor: [Child] explode all trees
60	#59 OR #60

62	MeSH descriptor: [Adult] this term only
63	adult*
64	#62 OR #63
65	(#58 AND #64) NOT #61 in Cochrane Reviews, Special Collections

" ab, ti, kw" indicate terms in title or abstract or keyword fields; "*" indicates truncation to find multiple endings of the word; quotes ("") around words indicate exact search for the entire phrase; MeSH descriptor indicates the indexing term [MeSH]; "explode all trees" retrieves references of the selected subject heading and more specific headings below; "OR" indicates similar keywords; "AND" links the concepts; "NOT" excludes specific words.

A2.3 EMBASE search strategy

A2.3.1 MeSH terms

Type 2 diabetes: exp non-insulin dependent diabetes mellitus/

Cancer: exp malignant neoplasm/

Risk/Incidence: exp risk factor/, exp incidence/

RCT: exp randomized controlled trial/

Types of cancer:

Most common cited types

- Bladder: exp bladder/
- Breast: exp breast/
- Colorectal: exp colon/, exp rectum/
- Endometrial: exp endometrium/
- Gastrointestinal: exp esophagus/, exp intestine/, exp stomach/
- HCC, BTC (Hepatobiliary cancers): exp hepatobiliary system/
- Kidney: exp kidney/
- Ovarian: exp ovary/
- Pancreatic: exp pancreas/
- Prostate: exp prostate/

Less common cited types

- Cervix: exp uterine cervix/
- Lung: exp lung/
- Melanoma: exp skin/
- Oral: exp mouth/
- Thyroid: exp thyroid gland/

Types of cancer (full term)

Most common cited types

- Bladder: exp bladder cancer/
- Breast: exp breast cancer/
- Colorectal: exp colon cancer/, exp rectum cancer/
- Endometrial: exp endometrium cancer/
- Gastrointestinal: exp esophagus cancer/, exp intestine cancer/, exp stomach cancer/
- HCC, BTC (Hepatobiliary cancers): exp hepatobiliary system cancer/

- Kidney: exp kidney cancer/
- Ovarian: exp ovary cancer/
- Pancreatic: exp pancreas cancer/
- Prostate: exp prostate cancer/

Less common cited types

- Cervix: exp uterine cervix carcinoma/
- Leukemia: exp leukemia/
- Lymphoma: exp lymphoma/
- Lung: exp lung cancer/
- Melanoma: exp melanoma/
- Oral: exp mouth cancer/
- Thyroid: exp thyroid cancer/

A2.3.2 Search Strategy

Table A2-3 Search strategy for Embase (OVID)

No.	Keyword searches
1	(bladder or transitional cell or transitional urothelial or urinary bladder or urothel* or vesica urinaria or breast or mammary or colon or colonic or colorectal or rectal or rectum or corpus endometri or decidu* or endometri* or myometri* or pelvi* or uteri* or uterus or stomach or gastric or gastroint* or intestin* or esophag* or ampulla of Vater or biliary or bile or gallbladder or hepatic or hepatocellular or intrahepatic or extrahepatic or liver or kidney or renal cell or ovarian or ovary or pancrea* or prostat* or Pca or cervix or cervic* or lymph* or lung or skin or mouth or oral squamous cell or oral or thyroid).tw.
2	(adenocarcinoma or cancer or carcino* or malignan* or metastas* or neoplas* or tumor* or tumour*).tw.
3	1 and 2
4	(intrahepatic cholangiocarcinoma or extrahepatic cholangiocarcinoma or leukem* or lymphoma* or NHL or melanoma).tw.
5	3 or 4
6	exp bladder/ or exp breast/ or exp colon/ or exp rectum/ or exp endometrium/ or exp esophagus/ or exp intestine/ or exp stomach/ or exp hepatobiliary system/ or exp kidney/ or exp ovary/ or exp pancreas/ or exp prostate/ or exp uterine cervix/ or exp lung/ or exp skin/ or exp mouth/ or exp thyroid gland/
7	exp malignant neoplasm/
8	6 and 7
9	exp bladder cancer/ or exp breast cancer/ or exp colon cancer/ or exp rectum cancer/ or exp endometrium cancer/ or exp esophagus cancer/ or exp intestine cancer/ or exp stomach cancer/ or exp hepatobiliary system cancer/ or exp kidney cancer/ or exp ovary cancer/ or exp pancreas cancer/ or exp prostate cancer/ or exp uterine cervix carcinoma/ or exp leukemia/ or exp lymphoma/ or exp lung cancer/ or exp melanoma/ or exp mouth cancer/ or exp thyroid cancer/
10	5 or 8 or 9
11	(adult-onset diabetes or diabetes mellitus type 2 or diabetes type 2 or DM2 or ketoacidosis or ketoacidosis-resistant diabetes or ketosis-resistant diabetes or maturity-onset diabetes or NIDDM or non-insulin dependen* or noninsulin dependen* or non-insulin-dependen* or T2D or T2DM or type 2 diabet*OR type II diabet*).tw.
12	exp non insulin dependent diabetes mellitus/
13	11 or 12
14	(associat* or diagnos*OR factor* or incidence or predict*OR predispos* or relationship* or risk* or risk factor*).tw.
15	exp risk factor/ or exp incidence/
16	14 or 15
17	10 and 13 and 16
18	randomised controlled trial*.tw.

19	exp randomised controlled trial/
20	18 or 19
21	17 not 20
22	limit 21 to human
23	limit 22 to (adult <18 to 64 years> or aged <65+ years>)

“tw” indicate terms in title or abstract; “*” indicates truncation to find multiple endings of the word; “exp/” indicates to explode the heading, so it retrieves references of the selected subject heading and more specific headings below; “OR” indicates similar keywords; “AND” links the concepts; “NOT” excludes specific words.

A2.4 PUBMED Search strategy

A2.4.1 MeSH Terms

Type 2 diabetes: "Diabetes Mellitus, Type 2"[Mesh]

Cancer: "Neoplasms"[Mesh]

Risk/Incidence: "Diagnosis"[Mesh], "Epidemiologic Factors"[Mesh], "Epidemiologic Measurements"[Mesh], "Epidemiologic Methods"[Mesh], "Incidence"[Mesh], "Risk Assessment"[Mesh]

RCT: "Randomized Controlled Trial"[Publication Type]

Types of cancer:

Most common cited types

- Bladder: "Urinary Bladder"[Mesh]
- Breast: "Breast"[Mesh]
- Colorectal: "Colon"[Mesh], "Rectum"[Mesh]
- Endometrial: "Endometrium"[Mesh]
- Gastrointestinal: "Gastrointestinal Tract"[Mesh], "Stomach"[Mesh]
- HCC, BTC (Hepatobiliary cancers): "Ampulla of Vater"[Mesh], "Biliary Tract" [Mesh], "Liver"[Mesh]
- Kidney: "Kidney"[Mesh]
- Ovarian: "Ovary"[Mesh]
- Pancreatic: "Pancreas"[Mesh]
- Prostate: "Prostate"[Mesh]

Less common cited types

- Cervix: "Cervix Uteri"[Mesh]
- Lung: "Lung"[Mesh]
- Melanoma: "Skin"[Mesh]
- Oral: "Mouth"[Mesh]
- Thyroid: "Thyroid Gland"[Mesh]

Types of cancer (full term)

Most common cited types

- Bladder: "Urinary Bladder Neoplasms"[Mesh]
- Breast: "Breast Neoplasms"[Mesh]
- Colorectal: "Colonic Neoplasms"[Mesh], "Rectal Neoplasms"[Mesh], "Colorectal Neoplasms"[Mesh]
- Endometrial: "Endometrial Neoplasms"[Mesh]
- Gastrointestinal: "Gastrointestinal Neoplasms"[Mesh]

- HCC, BTC (Hepatobiliary cancers): "Biliary Tract Neoplasms"[Mesh], "Liver Neoplasms"[Mesh], "Cholangiocarcinoma"[Mesh]
- Kidney: "Kidney Neoplasms"[Mesh]
- Ovarian: "Ovarian Neoplasms"[Mesh]
- Pancreatic: "Pancreatic Neoplasms"[Mesh]
- Prostate: "Prostatic Neoplasms"[Mesh]

Less common cited types

- Cervix: "Uterine Cervical Neoplasms"[Mesh]
- Leukemia: "Leukemia"[Mesh]
- Lymphoma: "Lymphoma"[Mesh]
- Lung: "Lung Neoplasms"[Mesh]
- Melanoma: "Melanoma"[Mesh]
- Oral: "Mouth Neoplasms"[Mesh]
- Thyroid: "Thyroid Neoplasms"[Mesh]

A2.4.2 Search Strategy

Table A2-4 Search strategy for PUBMED

Search strategy by concept

Diabetes search strategy

```
((((((((((((((adult-onset diabetes[Title/Abstract]) OR (diabetes mellitus type
2[Title/Abstract])) OR (diabetes type 2[Title/Abstract])) OR (dm2[Title/Abstract])) OR
(ketoacidosis[Title/Abstract])) OR (ketoacidosis-resistant diabetes[Title/Abstract])) OR
(DM2[Title/Abstract])) OR (ketosis-resistant diabetes[Title/Abstract])) OR (maturity-onset
diabetes[Title/Abstract])) OR (NIDDM[Title/Abstract])) OR (non-insulin
dependen*[Title/Abstract])) OR (noninsulin dependen*[Title/Abstract])) OR (non-insulin-
dependen*[Title/Abstract])) OR (T2D[Title/Abstract])) OR (T2DM[Title/Abstract])) OR (type 2
diabet*[Title/Abstract])) OR (Type II diabet*[Title/Abstract])) OR ("Diabetes Mellitus, Type
2"[Mesh])
```

Cancer Search strategy

```
((((((((((((((((((((((bladder[Title/Abstract]) OR (transitional cell[Title/Abstract])) OR
(transitional urothelial[Title/Abstract])) OR (urinary bladder[Title/Abstract])) OR
(urothel*[Title/Abstract])) OR (vesica urinaria[Title/Abstract])) OR ((breast[Title/Abstract])
OR (mammary[Title/Abstract])) OR (((colon[Title/Abstract]) OR (colonic[Title/Abstract]))
OR (colorectal[Title/Abstract])) OR (rectal[Title/Abstract])) OR (rectum[Title/Abstract])) OR
((((((((corpus uteri[Title/Abstract]) OR (decidu*[Title/Abstract])) OR
(endometri*[Title/Abstract])) OR (myometri*[Title/Abstract])) OR (pelvi*[Title/Abstract])) OR
(uteri*[Title/Abstract])) OR (uterus[Title/Abstract])) OR (((((((stomach[Title/Abstract]) OR
(gastric[Title/Abstract])) OR (gastroint*[Title/Abstract])) OR (intestin*[Title/Abstract])) OR
(esophag*[Title/Abstract])) OR (((((((ampulla of Vater[Title/Abstract]) OR (ampulla of
vater[Title/Abstract])) OR (biliary[Title/Abstract])) OR (bile[Title/Abstract])) OR
(gallbladder[Title/Abstract])) OR (hepatic[Title/Abstract])) OR
(hepatocellular[Title/Abstract])) OR (intrahepatic[Title/Abstract])) OR
(extrahepatic[Title/Abstract])) OR (liver[Title/Abstract])) OR ((kidney[Title/Abstract]) OR
(renal cell[Title/Abstract])) OR ((ovarian[Title/Abstract]) OR (ovary[Title/Abstract])) OR
(pancrea*[Title/Abstract])) OR ((prostat*[Title/Abstract]) OR (Pca[Title/Abstract])) OR
((cervix[Title/Abstract]) OR (cervic*[Title/Abstract])) OR (lymph*[Title/Abstract])) OR
(lung[Title/Abstract])) OR (skin[Title/Abstract])) OR (((mouth[Title/Abstract]) OR (oral
squamous cell[Title/Abstract])) OR (oral[Title/Abstract])) OR (thyroid[Title/Abstract])) AND
((((((((adenocarcinoma[Title/Abstract]) OR (cancer[Title/Abstract])) OR
(carcino*[Title/Abstract])) OR (malignan*[Title/Abstract])) OR (metastas*[Title/Abstract])) OR
(neoplas*[Title/Abstract])) OR (tumor*[Title/Abstract])) OR (tumour*[Title/Abstract])) OR
((((((((intrahepatic cholangiocarcinoma[Title/Abstract]) OR (extrahepatic
cholangiocarcinoma[Title/Abstract])) OR (leukem*[Title/Abstract])) OR
```

((lymphoma*[Title/Abstract])) OR (NHL[Title/Abstract])) OR (melanoma[Title/Abstract])) OR
 (((((((((((((((("Urinary Bladder"[Mesh]) OR ("Breast"[Mesh]) OR ("Colon"[Mesh]) OR
 ("Rectum"[Mesh]) OR ("Endometrium"[Mesh]) OR ("Gastrointestinal Tract"[Mesh]) OR
 ("Stomach"[Mesh]) OR ("Ampulla of Vater"[Mesh]) OR ("Biliary Tract"[Mesh]) OR
 ("Liver"[Mesh]) OR ("Kidney"[Mesh]) OR ("Ovary"[Mesh]) OR ("Pancreas"[Mesh]) OR
 ("Prostate"[Mesh]) OR ("Cervix Uteri"[Mesh]) OR ("Lung"[Mesh]) OR ("Skin"[Mesh]) OR
 ("Mouth"[Mesh]) OR ("Thyroid Gland"[Mesh]) AND ("Neoplasms"[Mesh])) OR
 (((((((((((((((("Urinary Bladder Neoplasms"[Mesh]) OR ("Breast Neoplasms"[Mesh]) OR
 ("Colonic Neoplasms"[Mesh]) OR ("Rectal Neoplasms"[Mesh]) OR ("Colorectal
 Neoplasms"[Mesh]) OR ("Endometrial Neoplasms"[Mesh]) OR ("Gastrointestinal
 Neoplasms"[Mesh]) OR ("Biliary Tract Neoplasms"[Mesh]) OR ("Liver Neoplasms"[Mesh]) OR
 ("Cholangiocarcinoma"[Mesh]) OR ("Kidney Neoplasms"[Mesh]) OR ("Ovarian
 Neoplasms"[Mesh]) OR ("Pancreatic Neoplasms"[Mesh]) OR ("Prostatic Neoplasms"[Mesh]) OR
 ("Uterine Cervical Neoplasms"[Mesh]) OR ("Leukemia"[Mesh]) OR ("Lymphoma"[Mesh]) OR
 ("Lung Neoplasms"[Mesh]) OR ("Melanoma"[Mesh]) OR ("Mouth Neoplasms"[Mesh]) OR
 ("Thyroid Neoplasms"[Mesh]))

Risk search strategy

((((((((((associat*[Title/Abstract]) OR (diagnos*[Title/Abstract])) OR (factor*[Title/Abstract]))
 OR (incidence[Title/Abstract])) OR (predict*[Title/Abstract])) OR (predispos*[Title/Abstract]))
 OR (relationship*[Title/Abstract])) OR (risk*[Title/Abstract])) OR (risk factor*[Title/Abstract]))
 OR (((((((("Diagnosis"[Mesh]) OR ("Epidemiologic Factors"[Mesh]) OR ("Epidemiologic
 Measurements"[Mesh]) OR ("Epidemiologic Methods"[Mesh]) OR ("Incidence"[Mesh]) OR ("Risk
 Assessment"[Mesh]))

RCT search strategy

((("randomized controlled trial*[Title/Abstract] AND "humans"[MeSH Terms]) OR ("Randomized
 Controlled Trial"[Publication Type] AND "humans"[MeSH Terms])) AND (humans [Filter])

Complete Strategy

(((((("bladder"[Title/Abstract] OR "transitional cell"[Title/Abstract] OR (("transit"[All Fields] OR
 "transited"[All Fields] OR "transiting"[All Fields] OR "transition"[All Fields] OR "transitional"[All
 Fields] OR "transitionals"[All Fields] OR "transitioned"[All Fields] OR "transitioning"[All Fields]
 OR "transitions"[All Fields] OR "transits"[All Fields]) AND "urothelial"[Title/Abstract]) OR
 "Urinary Bladder"[Title/Abstract] OR "urothel*" [Title/Abstract] OR "vesica
 urinaria"[Title/Abstract] OR ("Breast"[Title/Abstract] OR "mammary"[Title/Abstract]) OR
 ("Colon"[Title/Abstract] OR "colonic"[Title/Abstract] OR "colorectal"[Title/Abstract] OR
 "rectal"[Title/Abstract] OR "Rectum"[Title/Abstract]) OR ("corpus uteri"[Title/Abstract] OR
 "decidu*" [Title/Abstract] OR "endometri*" [Title/Abstract] OR "myometri*" [Title/Abstract] OR
 "pelvi*" [Title/Abstract] OR "uteri*" [Title/Abstract] OR "uterus"[Title/Abstract]) OR
 ("Stomach"[Title/Abstract] OR "gastric"[Title/Abstract] OR "gastroint*" [Title/Abstract] OR
 "intestin*" [Title/Abstract] OR "esophag*" [Title/Abstract]) OR ("Ampulla of
 Vater"[Title/Abstract] OR "Ampulla of Vater"[Title/Abstract] OR "biliary"[Title/Abstract] OR
 "bile"[Title/Abstract] OR "gallbladder"[Title/Abstract] OR "hepatic"[Title/Abstract] OR
 "hepatocellular"[Title/Abstract] OR "intrahepatic"[Title/Abstract] OR
 "extrahepatic"[Title/Abstract] OR "Liver"[Title/Abstract]) OR ("Kidney"[Title/Abstract] OR
 "renal cell"[Title/Abstract]) OR ("ovarian"[Title/Abstract] OR "Ovary"[Title/Abstract]) OR
 "pancrea*" [Title/Abstract] OR ("prostat*" [Title/Abstract] OR "Pca"[Title/Abstract]) OR
 ("cervix"[Title/Abstract] OR "cervic*" [Title/Abstract]) OR "lymph*" [Title/Abstract] OR
 "Lung"[Title/Abstract] OR "Skin"[Title/Abstract] OR ("Mouth"[Title/Abstract] OR "oral squamous
 cell"[Title/Abstract] OR "oral"[Title/Abstract]) OR "thyroid"[Title/Abstract]) AND
 ("adenocarcinoma"[Title/Abstract] OR "cancer"[Title/Abstract] OR "carcino*" [Title/Abstract]
 OR "malignan*" [Title/Abstract] OR "metastas*" [Title/Abstract] OR "neoplas*" [Title/Abstract] OR
 "tumor*" [Title/Abstract] OR "tumour*" [Title/Abstract])) OR ("intrahepatic
 cholangiocarcinoma"[Title/Abstract] OR "extrahepatic cholangiocarcinoma"[Title/Abstract] OR
 "leukem*" [Title/Abstract] OR "lymphoma*" [Title/Abstract] OR "NHL"[Title/Abstract] OR
 "Melanoma"[Title/Abstract]) OR ((("Urinary Bladder"[MeSH Terms] OR "Breast"[MeSH Terms] OR
 "Colon"[MeSH Terms] OR "Rectum"[MeSH Terms] OR "Endometrium"[MeSH Terms] OR
 "Gastrointestinal Tract"[MeSH Terms] OR "Stomach"[MeSH Terms] OR "Ampulla of Vater"[MeSH
 Terms] OR "biliary tract"[MeSH Terms] OR "Liver"[MeSH Terms] OR "Kidney"[MeSH Terms] OR
 "Ovary"[MeSH Terms] OR "Pancreas"[MeSH Terms] OR "Prostate"[MeSH Terms] OR "Cervix
 Uteri"[MeSH Terms] OR "Lung"[MeSH Terms] OR "Skin"[MeSH Terms] OR "Mouth"[MeSH Terms]
 OR "Thyroid Gland"[MeSH Terms]) AND "Neoplasms"[MeSH Terms]) OR ("Urinary Bladder

Neoplasms"[MeSH Terms] OR "Breast Neoplasms"[MeSH Terms] OR "Colonic Neoplasms"[MeSH Terms] OR "Rectal Neoplasms"[MeSH Terms] OR "Colorectal Neoplasms"[MeSH Terms] OR "Endometrial Neoplasms"[MeSH Terms] OR "Gastrointestinal Neoplasms"[MeSH Terms] OR "Biliary Tract Neoplasms"[MeSH Terms] OR "Liver Neoplasms"[MeSH Terms] OR "Cholangiocarcinoma"[MeSH Terms] OR "Kidney Neoplasms"[MeSH Terms] OR "Ovarian Neoplasms"[MeSH Terms] OR "Pancreatic Neoplasms"[MeSH Terms] OR "Prostatic Neoplasms"[MeSH Terms] OR "Uterine Cervical Neoplasms"[MeSH Terms] OR "Leukemia"[MeSH Terms] OR "Lymphoma"[MeSH Terms] OR "Lung Neoplasms"[MeSH Terms] OR "Melanoma"[MeSH Terms] OR "Mouth Neoplasms"[MeSH Terms] OR "Thyroid Neoplasms"[MeSH Terms])) AND "humans"[MeSH Terms] AND (("adult onset diabetes"[Title/Abstract] OR "diabetes mellitus type 2"[Title/Abstract] OR "diabetes type 2"[Title/Abstract] OR "DM2"[Title/Abstract] OR "ketoacidosis"[Title/Abstract] OR ("ketoacidosis-resistant"[All Fields] AND "diabetes"[Title/Abstract]) OR "DM2"[Title/Abstract] OR "ketosis resistant diabetes"[Title/Abstract] OR "maturity onset diabetes"[Title/Abstract] OR "NIDDM"[Title/Abstract] OR "non insulin dependen*"[Title/Abstract] OR "noninsulin dependen*"[Title/Abstract] OR "non insulin dependen*"[Title/Abstract] OR "T2D"[Title/Abstract] OR "T2DM"[Title/Abstract] OR "type 2 diabet*"[Title/Abstract] OR "type ii diabet*"[Title/Abstract] OR "diabetes mellitus, type 2"[MeSH Terms]) AND "humans"[MeSH Terms]) AND (("associat*"[Title/Abstract] OR "diagnos*"[Title/Abstract] OR "factor*"[Title/Abstract] OR "Incidence"[Title/Abstract] OR "predict*"[Title/Abstract] OR "predispos*"[Title/Abstract] OR "relationship*"[Title/Abstract] OR "risk*"[Title/Abstract] OR "risk factor*"[Title/Abstract] OR ("Diagnosis"[MeSH Terms] OR "Epidemiologic Factors"[MeSH Terms] OR "Epidemiologic Measurements"[MeSH Terms] OR "Epidemiologic Methods"[MeSH Terms] OR "Incidence"[MeSH Terms] OR "Risk Assessment"[MeSH Terms])) AND "humans"[MeSH Terms]) AND "humans"[MeSH Terms]) NOT (((("randomized controlled trial*"[Title/Abstract] AND "humans"[MeSH Terms]) OR ("Randomized Controlled Trial"[Publication Type] AND "humans"[MeSH Terms])) AND "humans"[MeSH Terms])) AND ((humans[Filter]) AND (alladult[Filter]))

"[Title/Abstract]" indicate terms in title or abstract; "*" indicates truncation to find multiple endings of the word; quotes ("") around words indicate exact search for the entire phrase; "[MeSH]" indicates the Medical Subject Heading term; "OR" indicates similar keywords; "AND" links the concepts; "NOT" excludes specific words.

A2.5 Web of Science search strategy

A2.5.1 Search Strategy

Table A2-5 Search strategy for WOS

No.	Keyword searches
1	T1= ("adult-onset diabetes" OR "diabetes mellitus type 2" OR "diabetes type 2" OR DM2 OR ketoacidosis OR "ketoacidosis-resistant diabetes" OR "ketosis-resistant diabetes" OR "maturity-onset diabetes" OR NIDDM OR "non-insulin dependen*" OR "noninsulin dependen*" OR "non-insulin-dependen*" OR T2D OR T2DM OR "type 2 diabet*" OR "type II diabet*")
2	T1 = (associat* OR diagnos*OR factor* OR incidence OR predict*OR predispos* OR relationship* OR risk* OR "risk factor*")
3	T1 =(bladder OR "transitional cell" OR "transitional urothelial" OR "urinary bladder" OR urothel* OR "vesica urinaria" OR breast OR mammary OR colon OR colonic OR colorectal OR rectal OR rectum OR "corpus endometri" OR decidu* OR endometri* OR myometri* OR pelvi* OR uteri* OR uterus OR stomach OR gastric OR gastroint* OR intestin* OR esophag* OR "ampulla of Vater" OR biliary OR bile OR gallbladder OR hepatic OR hepatocellular OR intrahepatic OR extrahepatic OR liver OR kidney OR renal cell OR ovarian OR ovary OR pancrea* OR prostat* OR Pca OR cervix OR cervic* OR lymph* OR lung OR skin OR mouth OR "oral squamous cell" OR oral OR thyroid)
4	T1= (adenocarcinoma OR cancer OR carcino* OR malignan* OR metastas* OR neoplas* OR tumor* OR tumour*)
5	3 AND 4
6	T1=("intrahepatic cholangiocarcinoma" OR "extrahepatic cholangiocarcinoma" OR leukem* OR lymphoma* OR NHL OR melanoma)
7	5 OR 6

8	1 AND 2 AND 7
9	TI= ("randomized controlled trial*")
10	#8 NOT #9
11	AB= ("adult-onset diabetes" OR "diabetes mellitus type 2" OR "diabetes type 2" OR DM2 OR ketoacidosis OR "ketoacidosis-resistant diabetes" OR "ketosis-resistant diabetes" OR "maturity-onset diabetes" OR NIDDM OR "non-insulin dependen*" OR "noninsulin dependen*" OR "non-insulin-dependen*" OR T2D OR T2DM OR "type 2 diabet*" OR "type II diabet*")
12	AB= (associat* OR diagnos*OR factor* OR incidence OR predict*OR predispos* OR relationship* OR risk* OR "risk factor*")
13	AB=(bladder OR "transitional cell" OR "transitional urothelial" OR "urinary bladder" OR urothel* OR "vesica urinaria" OR breast OR mammary OR colon OR colonic OR colorectal OR rectal OR rectum OR "corpus endometri" OR decidu* OR endometri* OR myometri* OR pelvi* OR uteri* OR uterus OR stomach OR gastric OR gastroint* OR intestin* OR esophag* OR "ampulla of Vater" OR biliary OR bile OR gallbladder OR hepatic OR hepatocellular OR intrahepatic OR extrahepatic OR liver OR kidney OR renal cell OR ovarian OR ovary OR pancrea* OR prostat* OR Pca OR cervix OR cervic* OR lymph* OR lung OR skin OR mouth OR "oral squamous cell" OR oral OR thyroid)
14	AB= (adenocarcinoma OR cancer OR carcino* OR malignan* OR metastas* OR neoplas* OR tumor* OR tumour*)
15	13 AND 14
16	AB= ("intrahepatic cholangiocarcinoma" OR "extrahepatic cholangiocarcinoma" OR leukem* OR lymphoma* OR NHL OR melanoma)
17	15 OR 16
18	11 AND 12 AND 17
19	AB= ("randomized controlled trial*")
20	#18 NOT #19
21	10 OR 20

"AB" indicates term in the abstract field; "TI" indicates term in the title field; "*" indicates truncation to find multiple endings of the word; quotes ("") around words indicate exact search for the entire phrase; "OR" indicates similar keywords; "AND" links the concepts; "NOT" excludes specific words.

A3. Criteria for screening selection

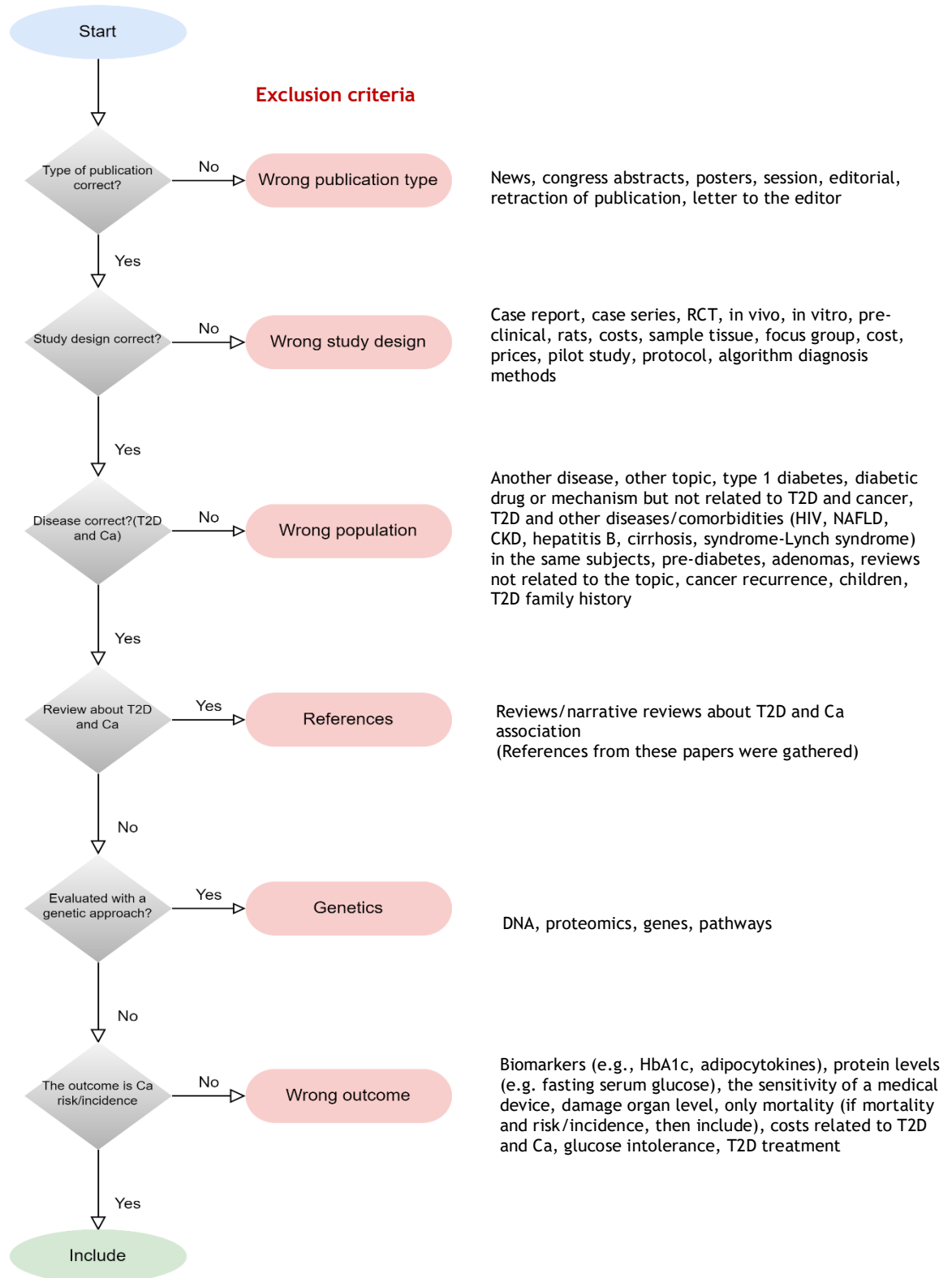


Figure A3-1 Flowchart for title/abstract screening stage. The flowchart was designed to identify the exclusion criteria for title/abstract screening. "Wrong population" criterion was also applied when the same paper considered: a case report and a literature review; costs and health outcomes; T2DM treatment on cancer survival; health review but not specifically on T2DM and cancer. "Wrong study design" criterion was also applied when the same paper considered: cost and resource use; paper retracted.

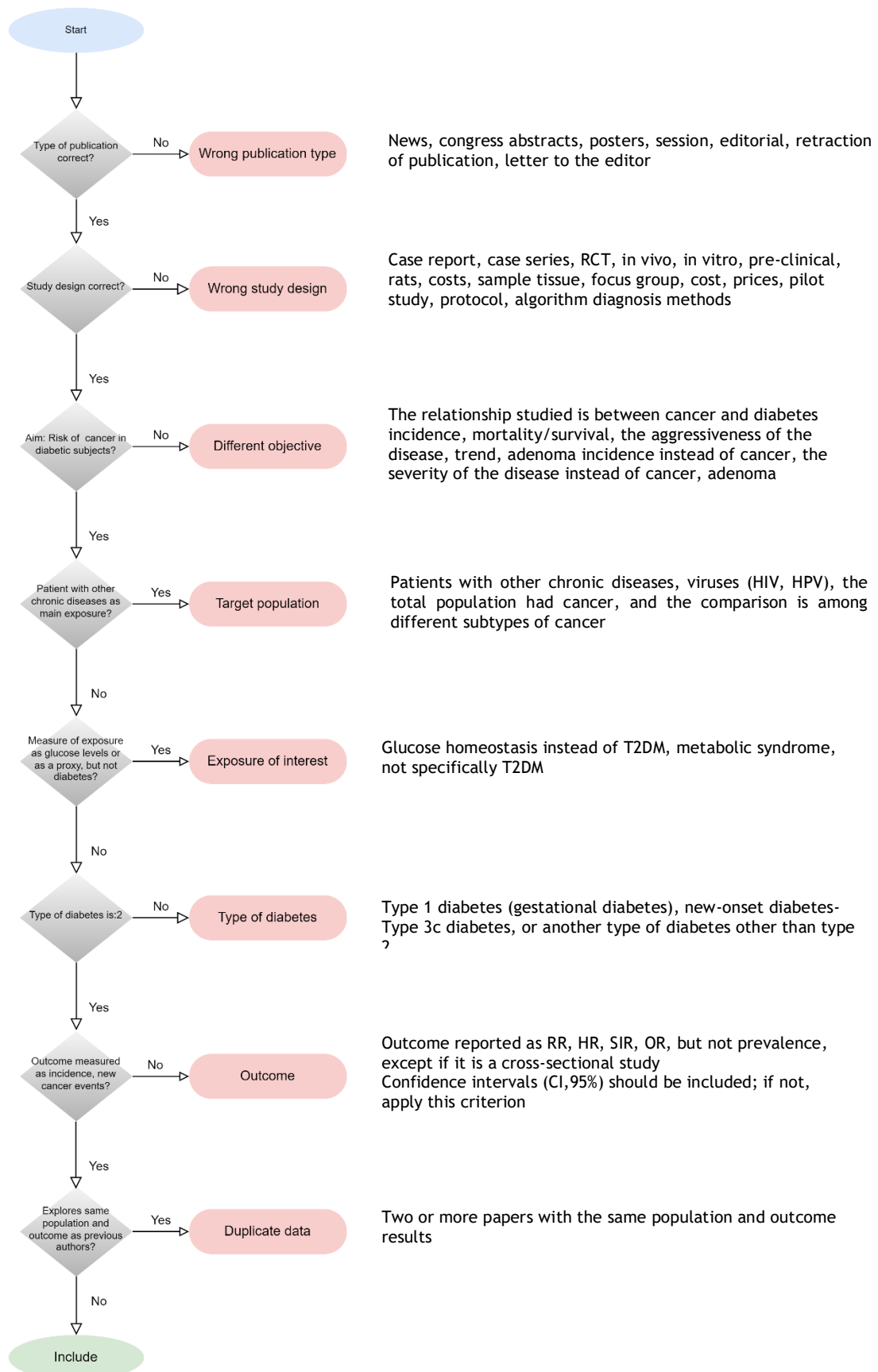


Figure A3-2 Flowchart for full-text screening stage. The flowchart was designed to identify the exclusion criteria for full-text screening. "Wrong publication type" and "Wrong study design" were included again in this stage as some records could only be excluded once the entire paper was screened. If the type of diabetes was not specified, a diagnosis of the disease after 30 years old was considered type 2.

A4. Newcastle-Ottawa Scale (NOS)

Table A4-1 Criteria for the Newcastle-Ottawa scale (NOS) star allocation to assess the risk of bias of cohort studies

Domain	Item	Acceptable (Stars awarded)	Unacceptable (No stars awarded)
Selection	1. Representativeness of the exposed cohort	– Truly representative of the average, veterans, ethnic group, community, women/men – Somewhat representative of the veterans, ethnic group, community, women/men	– A selected group of patients (nurses, physicians) – No description of the derivation of the cohort
	2. Selection of the non-exposed cohort	– Drawn from the same community as the exposed cohort	– Drawn from a different source – No description of the derivation of the non-intervention cohort
	3. Ascertainment of exposure	– Secure records (e.g health care records, surgical records) – Structured interview	– Written self-report – Other / no description
	4. Demonstration that outcome of interest was not present at the start of the study	– Yes	– No
Comparability	5. Comparability of cohorts on the basis of the design or analysis	– Study controls for at least age and sex – An additional star if the study controls for BMI or obesity, or overweight or WHR.	– No confounders included
Outcome	6. Assessment of outcome	– Independent blind assessment – Record linkage	– Self-report – Other/no description
	7. Was follow-up long enough for outcomes to occur	– Media/Median duration of follow up ≥ 3 or more years	– Media/median duration of follow up < 3 years – Not reported media/median duration of follow up
	8. Adequacy of follow up of cohorts	– Complete follow up (all subjects accounted for) – Subjects lost to follow up <20%, or description of those lost – Follow up rate >80% and no description of those lost – In the case of a database, the number of missing values reported and $\leq 5\%$	– No statement of follow-up or missing values

Each item can be awarded one star, except for “Comparability”, which could be given two stars. The follow-up period was established to confirm the type of diabetes and, at the same time, exclude reverse causality (type 3c diabetes, a manifestation of PC, appears within 2-3 years before cancer diagnosis (Mizuno et al., 2020)) and detection bias (Bjornsdottir et al., 2020).

Table A4-2 Criteria for the Newcastle-Ottawa scale (NOS) star allocation to assess the risk of bias of case-control studies

Domain	Item		Acceptable (Stars awarded)	Unacceptable (No stars awarded)
Selection	1. Is the case definition adequate?	–	With independent validation (<1person/record/time/process to extract information)	– Record linkage (e.g., ICD codes in a database) or based on self-reports with no reference to the primary record
		–	Reference to primary record source (e.g., x-rays, medical/hospital records)	– No description
	2. Representativeness of the cases	–	All eligible cases with the outcome of interest over a defined period of time, all cases in a defined catchment area, hospital or clinic, group of hospitals, health maintenance organization, or an appropriate sample of those cases	– Not satisfying requirements to meet acceptable criteria or not stated
	3. Selection of Controls	–	Community controls	– Hospital controls, within the same community as cases but derived from a hospitalised population
				– No description
	4. Definition of Controls	–	No history of the disease (outcome)	– No description of the source or no mention of the history of the outcome
Comparability	5. Comparability of cases and controls on the basis of the design or analysis	–	Study controls for at least age and sex	– No confounders included
		–	An additional star if the study controls for BMI or obesity, or overweight or WHR	
Exposure	6. Ascertainment of exposure	–	Secure records (e.g., surgical records)	– Interview not blinded to case/control status
		–	A structured interview where blind to case/control status	– Written self-report or medical record only
				– No description
	7. Same method of ascertainment for cases and controls	–	Yes	– No
	8. Non-Response rate	–	Same rate for both groups	– No respondents described
				– Rate different ($\geq 20\%$) between groups and no designation

Each item can be awarded one star, except for “Comparability”, which could be given two stars.

Table A4-3 Criteria for the Newcastle-Ottawa scale (NOS) star allocation to assess the risk of bias of Cross-sectional studies

Domain	Item	Acceptable (Stars awarded)	Unacceptable (No stars awarded)
Selection	1. Representativeness of the sample	– Truly representative of the average in the target population (all subjects or random sampling) – Somewhat representative of the average in the target population (non-random sampling)	– A selected group of users – No description
	2. Sample size	– Justified and satisfactory	– Not justified
	3. Non-included subjects	– Comparability between the characteristics of included and non-included subjects was established; and if the response rate was satisfactory	– No description of the response rate or the characteristics of the included and non-included subjects
	4. Ascertainment of exposure	– Validated measurement tool (2 stars) – Non-validated measurement tool, but the tool is available or described (1 star)	– No description of the measurement tool
Comparability	5. Subjects in different outcome groups are comparable	– Study controls for at least age and sex – An additional star if the study controls for BMI or obesity, or overweight or WHR.	– No confounders included
Outcome	6. Assessment of outcome	– Independent blind assessment (2 stars) – Record linkage (2 stars) – Self-report (1 star)	– No description
	7. Statistical test	– The statistical test is clearly described and appropriate	– The statistical test is not described, appropriate or incomplete

Each item can be awarded one star, except for “Ascertainment of exposure”, “Comparability”, “Assessment of outcome”, which could be given two stars.

A5. Reasons for Title and Abstract Screening Exclusions

Table A5-1 Title and abstract screening exclusion reasons

Reason for exclusion	Number of studies
Wrong population	7,304
Wrong publication type	1,242
Wrong study design	3,848
Wrong outcome	291
Genetics	148
References	85
Total	12,918

The main reason for excluding studies during the title and abstract screening stage was related to the target population. Specifically, the study subjects had another type of diabetes, or the primary focus of the papers was on another disease, with T2DM considered only as a comorbidity.

A6. Searching other sources

Systematic reviews, meta-analyses and narrative reviews were selected through the title/abstract screening process, and their reference lists were screened for eligible studies. In total, 126 records were gathered for this purpose:

(Abudawood, 2019; Agache et al., 2018; Anastasi et al., 2018; Anwar et al., 2014; Baffy, 2012; Bansal et al., 2013; Bartosch-Härlid & Andersson, 2010; Ben et al., 2011; Bernard et al., 2016; Bonora & DeFronzo, 2020; Bosetti et al., 2012; Boyle et al., 2012; Bugianesi, 2005; Cantiello et al., 2015; Castillo et al., 2012; Chao & Page, 2008; Chen et al., 2015; Collier et al., 2012; Coman et al., 2021; Craig et al., 2016; Crawley et al., 2018; Cui & Andersen, 2012; Dąbrowski, 2021; de Souza et al., 2016; Deng et al., 2012; Dong et al., 2022; Eibl et al., 2018; Feng et al., 2020; Ferrís-i-Tortajada et al., 2011; Franzen et al., 2018; Fröhlich & Wahl, 2012; Gallagher & LeRoith, 2013; Giovannucci, 2001, 2007; Gong et al., 2015; Goodwin, 2021; Gouveri et al., 2011; Gu et al., 2016; J. Guo et al., 2022; J.-Z. Guo et al., 2022; Guraya, 2015; Hara, 2012; Hippisley-Cox & Coupland, 2015; Hsieh et al., 2019; D. Q. Huang et al., 2023; Hussain & Chowdhury, 2019; Huxley et al., 2005; J. et al., 2023; Jian Gang et al., 2015; Jiang et al., 2011; Joost, 2014; Joung et al., 2015; Kallergi et al., 2015; Katzke et al., 2020; Kilvert & Fox, 2020; Kramer et al., 2012; Lann & LeRoith, 2008; Larsson et al., 2007;

Lega & Lipscombe, 2020; D. Li, 2012; Y. Li et al., 2019; Y. R. Li et al., 2018; Ling et al., 2020; López-Suárez, 2019; Lu et al., 2023; Magruder et al., 2011; Malaguarnera et al., 2017; Man Chu, 2018; Mantovani & Targher, 2017; Marengo et al., 2016; Maskarinec et al., 2018; McFarland & Cripps, 2010; Muller, 1998; Nagel & Göke, 2006; Nicolucci, 2010; Ohkuma et al., 2018b; Onitilo et al., 2012; Patrascu Mihaita MD PhD, Diculescu Mircea MD PhD et al., 2018; Pearson-Stuttard et al., 2021; Pereira et al., 2020; Pierce, 2012; Pizzato et al., 2019; Qi et al., 2014; Rau et al., 2016; Raychaudhuri et al., 2022; Reeves et al., 2016; Rose et al., 2007; Rysz et al., 2020; Sacerdote & Ricceri, 2018; Saka Herrán et al., 2018; Samuel et al., 2018; Sayyid & Fleshner, 2016; Schott et al., 2010; Siddiqui, 2011; D. Simon & Balkau, 2010; Soltani et al., 2021; Suchanek et al., 2016a, 2016b; Suh & Kim, 2019; Sun & Yu, 2012; Szablewski, 2014; Tan et al., 2017; Tang et al., 2017; Tarantino et al., 2021; “The Relationship between Diabetes Mellitus and Gastric Cancer and the Potential Benefits of Metformin: An Extensive Review of the Literature,” 2021; Tolman et al., 2004; Tomic et al., 2022; Tsilidis, Kasimis, et al., 2015; Vigneri et al., 2009; Wainwright et al., 2017; C. Wang et al., 2012; Y. Wang et al., 2016, 2020; Y. Wang, O’Neil, et al., 2019; Wojciechowska et al., 2016; Wolf et al., 2005; Wu et al., 2013; B. Xu et al., 2017; Xue & Michels, 2007; Yan, Wang, Fu, et al., 2021; Yan, Wang, Yu, et al., 2021; Yee et al., 2020; Yeo, 2015; Yu et al., 2016; Zanders et al., 2014; Z.-H. Zhang et al., 2013)

Since some of the publications identified in the bibliographies of the records mentioned above were also meta-analyses, systematic reviews, or narrative reviews, the same process was repeated, resulting in 52 new references:

(Anothaisintawee et al., 2013; Bao et al., 2013; Batabyal et al., 2014; Bonovas et al., 2004; de Bruijn et al., 2013; Fang et al., 2013, 2018; Friberg et al., 2007; Ge et al., 2011; Hardefeldt et al., 2012; W. Huang et al., 2012; James Everhart, MD, MPH, David Wright, 1995; Jiang et al., 2011; Jing et al., 2012; Kasper & Giovannucci, 2006; Larsson et al., 2005, 2006; Larsson & Wolk, 2011; J. Y. Lee et al., 2013; J.-Y. Lee et al., 2013; H. Li & Qian, 2017; C. Liao et al., 2014; S. Liao et al., 2011; Long et al., 2012; S. Luo et al., 2016; W. Luo et al., 2012; Miao et al., 2017; Mitri et al., 2008; Noto et al., 2010, 2011; Pang et al., 2017; Ren et al., 2011; S. et al., 2020; Saed et al., 2019b; Schmid et al., 2013; Shimoyama, 2013; S. Song et al., 2015; Starup-Linde, Karlstad, Eriksen, Vestergaard,

Bronsveld, Vries, Andersen, Auvinen, Haukka, Hjellvik, Bazelier, Boer, Furu, & de Bruin, 2013; Tian et al., 2012a, 2012b; L. Wang et al., 2017, 2020; P. Wang et al., 2012; Y. Wang, Sun, et al., 2019; H. Xu et al., 2013; X. Xu et al., 2013; X.-Q. Yang et al., 2013; Yoon, 2013; D. Zhang et al., 2017; F. Zhang et al., 2012; Zhao et al., 2016; Zhu et al., 2013)

In total, 178 additional references were screened.

A7. Selected studies

Table A 7-1 lists the abbreviations and acronyms for the sources of the studies selected in the systematic review, while Table A 7-2 provides the abbreviations for the types of cancer included in the studies.

A description of the selected studies according to study design is provided in Tables A7-3 (case-control), A7-4 (cohort), A7-5 (cross-sectional), and A7-6 (Mendelian randomization).

Table A 7-1 Abbreviations and acronyms for the sources of the studies selected in the systematic review.

Abbreviations/ Acronym	Expanded Form
AAWC	Armenian-American Wellness Center
ANECs	Australian National Endometrial Cancer Study
BCAC	Breast Cancer Association Consortium
BCLHD	British Columbia Linked Health Databases
BEACON	Barrett's and Esophageal Adenocarcinoma Consortium
BIP	Bezafibrate Infarction Prevention
BWHS	Black Women's Health Study
CCDRD	Cancer and Chronic Disease Research Database
CKB Study	China Kadoorie Biobank Study
CLL-MCC	CLL-Multi Case-Control
COSM	Cohort of Swedish Men
CPRD	Clinical Practice Research Datalink
CPS-II Nutrition Cohort	Cancer Prevention Study (CPS)-II Nutrition Cohort
D2C cohort	Diabetes II-to-Cancer' (D2C) cohort
Database of HIC	Health Informatics Centre
DETECT study	Diabetes Cardiovascular Risk and Evaluation: Targets and Essential Data for Commitment of Treatment (DETECT) study
DIAGRAM consortium	DIAbetes Genetics Replication And Metaanalysis consortium
EDGE Study	Estrogen, Diet, Genetics, and Endometrial Cancer (EDGE) Study

Abbreviations/ Acronym	Expanded Form
eMERGE network	Electronic Medical Records and Genomics network
EPIC	European Prospective Investigation into Cancer and Nutrition
FDS	Fremantle Diabetes Study
GLIOGENE Consortium	Glioma Gene Consortium
GPRD	General Practice Research Database
HIM	Health Interview Study of Men
HIWATE Project	Health Impacts of long-term exposure to disinfection by-products in drinking WATER project
HPFS	Health Professionals Follow-up Study
IARC	International Agency for Research on Cancer
IWHS	Iowa Women's Health Study
JPHC	Japan Public Health Center-based Prospective
KCIS	Keelung Community-Based Integrated Screening
KNHI	Korean National Health Insurance
KNHIS	Korean National Health Insurance Services
KORA	Cooperative Health Research in the Augsburg Region ("Kooperative Gesundheitsforschung in der Region Augsburg")
KPSC	Kaiser Permanente Southern California
LCPP	Liver Cancer Pooling Project
LHID 2005	Longitudinal Health Insurance Database 2005
LIBCSP	Long Island Breast Cancer Study Project
LTR	Louisiana Tumor Registry
MAGIC	Meta-Analyses of Glucose and Insulin-related traits Consortium
MCC	Multi Case-Control
MDA	MD Anderson Cancer Center
MEC	Multiethnic Cohort
NDSS	National Diabetes Services Scheme
NHANES	National Health and Nutrition Examination Survey
NHI	National Health Insurance
NHIC	National Health Insurance Corporation
NHIF	National Health Insurance Fund
NHIRD	National Health Insurance Research Database
NHIS-HEALS	Korean National Health Insurance Service (NHIS)-National Health Screening Cohort (HEALS)
NHRI database	National Health Research Institute database
NHS	Nurses' Health Study
NJMU GSA Project	Global Screening Array (GSA) Project of Nanjing Medical University (NJMU)
NJMU GWAS	Genome-Wide Association Study (GWAS) of Nanjing Medical University (NJMU)
NLCS	Netherlands Cohort Study on diet and cancer
NOC	National Oncology Center
NORI	Nuclear Medicine, Oncology and Radiotherapy Institute
NPCR	National Prostate Cancer Register
OCAC	Ovarian Cancer Association Consortium
OCS	Olmsted County Study
PanC4	Pancreatic Cancer Case-Control Consortium
PANDoRA consortium	PANcreatic Disease ReseArch consortium
PanGenEU	European Study into Digestive Illnesses and Genetics
PanScan	Pancreatic Cancer Cohort Consortium
PCBaSe 3.0	Prostate Cancer data Base Sweden 3.0

Abbreviations/ Acronym	Expanded Form
PMBB	Penn Medicine BioBank
POPGEN	Population Genetic Cohort
PRACTICAL	Prostate Cancer Association Group to Investigate Cancer-Associated Alterations in the Genome
PROS	Prostate Research Study
PROtEuS	Prostate Cancer & Environment Study
RCGLD	Research Center for Gastroenterology and Liver Diseases
REP	Rochester Epidemiology Project
SAYO	South Australian Young Onset Colorectal Polyp and Cancer Study
SCI-DC dataset	Scottish Care Information -Diabetes Collaboration dataset
SCSS	The Southern Community Cohort Study
SEARCH	Studies of Epidemiology and Risk factors in Cancer Heredity study
SEER	Surveillance, Epidemiology, and End Results
SMHS	Shanghai Men's Health Study
STR	Nationwide Swedish Twin Registry
Swedish NDR	Swedish National Diabetes Register
SWHS	Shanghai Women's Health Study
T-CALOS study	Thyroid Cancer Longitudinal Study
TCS	Thai Cohort Study
TLGS	Tehran Lipid and Glucose Study
TNHIRD	Taiwan's National Health Insurance Research Database
UCSF	University of California, San Francisco
UEBMI database	Urban Employee Basic Medical Insurance database
United States NCI	United States National Cancer Institute
VISN16	Veterans Integrated Services Network 16 database
WHI cohort	Women's Health Initiative (WHI) cohort

Table A 7-2 Abbreviations for the types of cancer included in the studies of the systematic review.

Abbreviation	Type of cancer
AML	Acute myeloid leukaemia
BC	Breast cancer
BlaC	Bladder cancer
BraNSC	Brain/nervous system cancer
BTC	Biliary tract cancer
C. in situ	Carcinoma in situ
CC	Colon cancer
CeuC	Cervix uteri cancer
CLL/SLL	Chronic Lymphocytic Leukemia
CML	Chronic myeloid leukaemia
CRC	Colorectal cancer
CUP	Cancer of unknown primary
EC	Esophagus cancer
ECC	Extrahepatic cholangiocarcinoma
EJA	Esophagogastric junction adenocarcinoma
EnC	Endocrine cancer
EndC	Endometrial cancer
GaC	Gastric cancer

Abbreviation	Type of cancer
GB	Gallbladder cancer
GC	Genital cancer
GI	Gastrointestinal cancer
H&N	Head and Neck
HCC	Hepatocellular cancer
Hematologic.	Haematological neoplasms
HLy	Hodgkin lymphoma
ICC	Intrahepatic cholangiocarcinoma
KC	Kidney cancer
LaC	Larynx cancer
LC	Lung cancer
Le	Leukemia
LiC	Large intestine cancer
LL	Lymphocytic leukemia
LVC	Liver cancer
Ly	Lymphoma
M	Melanoma
MesC	Mesothelioma
MLe	Myeloid leukemia
My	Myeloma
NHL	Non-Hodgkin lymphoma
NM	Nonmelanoma
OC	Ovary cancer
PC	Pancreatic cancer
PhC	Pharynx cancer
PrC	Prostate cancer
RC	Rectal cancer
RsC	Rectosigmoid cancer
SaG	Salivary gland cancer
SiC	Small intestine cancer
ST	Soft tissue
TC	Thyroid cancer
TeC	Testicular cancer
UADT	Upper aerodigestive tract
UTC	Urinary tract cancer

Table A 7-3 Main characteristics of the case-control studies included in the systematic review

No.	Author, year	Country	Type of cancer	Data source	Study setting	Control group selection method	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	T2DM in Cases/ Controls	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	N O S
1	(Agrawal et al., 2014)	USA	EC	Veterans Affairs Medical Center	Hospitals	Patients with Barrett's Esophagus	Mean 62.3y (11.2)	568 (97.4%)	Cases <25, 42.5%; 25-<30, 34.5%; ≥30, 23% Controls <25, 22.0%; 25-<30, 39.4%; ≥30, 38.6%	583	Cases 39/115 Controls 119/468	Medical records	Medical records	Age, alcohol intake, BMI, PPIs, smoking, statin	7
2	(Alsheridah & Akhtar, 2018)	Kuwait	CRC	Kuwait Cancer Control Centre	Hospital	Ratio 2:1 Matched by age, sex, and nationality	25-80 y	168 (54.4%)	NR	309	Cases 19/103 Controls 82/206	Interview (self-report)	Pathology records	Age, BMI, constipation, education level, egg intake, fruits and vegetables intake, hypercholesterolaemia, nationality, physical activity, red meat-intake, sex	6
3	(Antwi et al., 2016)	USA	PC	Biospecimen Resource for Pancreas Research	Hospitals	Matched by age, race, sex	Mean Cases 66.7y (10.3) Controls 65.4y (10.3)	1,416 (55%)	Mean Cases 27.7 (5.1) Controls 26.9 (4.6)	2,573	Cases 211/817 Controls 151/1,756	Medical records	Pathology records	Age, BMI, education level, race, sex, smoking	6
4	(B. S. Lee et al., 2015)	Korea	BTC, ECC, ICC	Cheju Halla General Hospital	Hospital	Ratio 2:1 Matched by age, sex, and date of diagnosis. Randomly selected	Mean Cases 67.8y (12.5) Controls 67.5y (12.5)	418 (50.4%)	Obesity Cases 23.2% Controls 24.3%	828	Cases 65/276 Controls 69/552	Interview (self-report), medical records	Pathology reports	Age, alcohol intake, cholecystolithiasis, choledocholithiasis, HBV, hepatolithiasis, liver fluke infestation, sex	5
5	(Balbi et al., 2010)	Italy	HCC	Liver Unit and Diabetic Clinic of 3rd Internal Medicine	Hospital	Matched by age, sex, BMI, history of diabetes, time of admission	Mean Cases 68.5y (8.9) Controls 69.4y (13.8)	749 (78.6%)	Mean Cases 25.2(3.1) Controls 25.1(2.9)	955	Cases 145/465 Controls 62/490	Medical records	Pathology records, imaging techniques	Age, alcohol intake, BMI, history of diabetes, sex	5
6	(Barami et al., 2017)	USA	BraNSC	KPSC	Population	Ratio 1:1 Matched by age, sex, length of membership Randomly selected	Median 65y (57-74)	1,074 (55.4%)	<30 Cases 27.7% Controls 32.1%	1,938	Cases 152/969 Controls 163/969	Registry	Registry	Age, race, sex	7
7	(Baron et al., 2001)	USA	BC	State tumor registries (Wisconsin, Massachusetts, and New Hampshire)	Population	Randomly selected	50-75 y Mean Cases 65.3y (8.0) Controls 64.1y (8.0)	N/A	NR	11,587	Cases NR/5,659 Controls NR/5,928	Interview-blinded (self-report)	Registry	Age, alcohol intake, BMI, oestrogen-related variables (age at first birth, age at menopause, HRT use, menopausal status, oral contraceptive, parity), family history of BC, state	8
8	(Barrios-Rodriguez et al., 2022)	Spain	PrC	MCC-Spain	Population	Matched by age, sex, study region	Mean Cases 65.9y (7.4) Controls 66.4y (8.6)	N/A	Cases Non- T2DM <25, 26.8%; 25-30, 51.2%; ≥30, 22%	2,426	Cases 136/1,047 Controls 217/	Interview (self-report)	Medical records and pathology records	Age, BMI, education level, family history of PrC, sex	9

No.	Author, year	Country	Type of cancer	Data source	Study setting	Control group selection method	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	T2DM in Cases/ Controls	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	N O S	
									T2DM <25, 16.2%; 25-30, 47.8%; >30, 36% Controls Non-T2DM <25, 27%; 25-30, 49.9%; >30, 23.1% T2DM <25, 16.6%; 25-30, 50.2%; ≥30, 33.2%		1,379					
9	(Benavente et al., 2018)	Spain	CLL/ SLL	CLL-MCC study	Population	Matched by sex, age, province of recruitment. Randomly selected	Mean Non-DM 64.2y (10.9) DM 66.5y (10.4)	1,394 (58%)	NR	2,405	Cases 69/560 Controls 265/1,845	Interview (self-report)	Pathology records	Age, education level, region of recruitment, sex	6	
10	(Blanc-Lapierre et al., 2015)	Canada	PrC	PROtEuS	Population	Matched by age, district. Randomly selected	≥20 y	N/A	Mean Cases 26.8(4.0) Controls 27.2(4.4)	3,932	Cases 205/1,937 Controls 348/1,995	Interview (self-report)	Pathology records	Age, ancestry, family history of PrC, family income, PrC screening	6	
11	(Bosetti et al., 2014)	Australia, Canada, Czech Republic, Slovakia, Greece, Italy, Netherlands, Poland, USA	PC	PanC4 consortium	Hospital/ Population	NR	Cases 60-68y Controls 56-70y	12,741 (57.1%)	Cases <20,4.8%; 20-<25,33.3%; 25-<30,37.4%; ≥30,19.5% Controls <20,4.8%; 20-<25,38.5%; 25-<30,36.8%; ≥30,13.8%	22,292	Cases 1,767/8,305 Controls 1,284/13,987	Self-reported	NR	Age, alcohol intake, BMI, education level, history of pancreatitis, race/ethnicity, sex, smoking, study, center	6	
12	(Braga et al., 1997)	Italy	HCC	Hospitals in the Greater Milan area	Hospitals	No	Median Cases 60 y (24-74) Controls 56 y (22-74) 20-74 y	1,266 (73.3%)	NR	1,728	Cases NR/320 Controls NR/1,408	Interview (self-report)	Pathology reports	Age, area of residence, history of liver cirrhosis, smoking status, total carrot/green vegetable/fresh fruit consumption	6	
13	(Brinton et al., 1992)	USA	EndC	7 hospitals of 5 areas (Chicago, Hershey, Irvine, Long Beach, Minneapolis)	Population	Ratio 1:1 Matched by age, race, location of residence. Randomly selected	Mean Cases 59.2 y Controls 58.0 y 40-60 y	N/A	NR	702	Cases NR/405 Controls NR/297	Interview (self-report)	Pathology reports	Age, education level, oestrogen-related variables (HRT, number of births, oral contraceptive use), race, weight	7	
14	(Castillo-Otiniano & Yan-Quiroz, 2019)	Peru	BC	Regional Institute of Neoplastic Diseases North	Hospital	No	Mean Cases 52.3y (9.1) Controls	N/A	NR	140	Cases 22/70 Controls 10/70	Medical records	Pathology reports	NR	4	

No.	Author, year	Country	Type of cancer	Data source	Study setting	Control group selection method	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	T2DM in Cases/ Controls	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	N O S
15	(Cleveland et al., 2012)	USA	BC	LIBCSP	Population	Matched by age. Randomly selected.	54.3y (10.3) Mean T2DM 63.6y Non-T2DM 57.4y	N/A	Mean Non-T2DM 26.1 T2DM 30.9	2,900	Cases 121/1,447 Controls 95/1,453	Interview (self-report)	Medical records	Age, menopausal status, obesity, race	8
16	(D. Li et al., 2009)	USA	PC	University of Texas M. D. Anderson Cancer Center	Hospital	Matched by sex, age, and race	Approx. 85% older than 50 y	1,125 (61.3%)	Cases <25, 55.4%; 25-<30, 35.3%; ≥30: 9.3% Controls <25, 75.1%; 25-<30, 21.2%; ≥30: 3.7%	1,836	Cases 259/973 Controls 109/863	Interview (self-report) and medical records	Pathology reports	Age, alcohol intake, BMI, diabetes duration, family history of cancer, insulin use, race, sex, smoking	7
17	(Davtalab Esmaeili et al., 2019)	Iran	CRC	General and referral hospitals (Imam Reza, Shahid Madani, and Sina)	Hospitals	Matched by sex and age	30-75 y Mean Cases 60.64y (13.23) Controls 59.51y (13.73)	220 (53%)	NR	414	Cases 48/207 Controls 21/207	Hospital records, laboratory test, self-report	Pathology reports	Age, sex	6
18	(De Mesquita et al., 1992)	Netherlands	PC	SEARCH	Population	Ratio 2:1 Matched by age, sex. Randomly selected	35-79 y	329 (49.6%)	NR	663	Cases 24/176 Controls 50/487	Interview (self-report)	Medical records, pathology reports, registry	Age, response status, sex, smoking	6
19	(Dianatinasab et al., 2019)	Iran	BC	Namazi Hospital	Hospital	Matched by age	Approx. 80% younger than 60y	N/A	Cases <5, 30.7%; 25-<30, 45.1%; ≥30, 24.2% Controls <25, 33.2%; 25-<30, 48.5%; ≥30, 16.0%	2,018	Cases 85/1,009 Controls 67/1,009	Interview (self-report)	Pathology reports	Age, BMI, chest X-ray history, deliberate weight loss, education level, ethnicity, family history of BC, history of benign breast disease, oestrogen-related variables (age at first delivery, breastfeeding, history of miscarriage, menarche age, menopause status, oral contraceptives, regular menstrual), occupation, physical activity, smoking	7
20	(Diaz-Algorri et al., 2015)	Puerto Rico	CRC	VA Caribbean Healthcare System and the Puerto Rico Medical Centre	Hospitals	No	Mean 61.1y (11.9)	269 (59.6%)	<30 Cases, 29% Controls, 70% ≥30 Cases, 29% Controls, 70%	374	Cases 51/218 Controls 76/156	Self-report	Medical records, pathology reports	Age, family history of diabetes, study center	6
21	(Donadon et al., 2010)	Italy	HCC	Pordenone General Hospital	Hospital	Ratio 3:1 Matched by age, sex, BMI, time of hospital admission. Randomly selected	Mean Cases 68.5y (8.9) Controls 69.4y (13.8)	1,845 (80%)	Mean Cases 25.2(3.1) Controls 25.1(2.9)	2,306	Cases 190/610 Controls 215/1,696	Laboratory test	Imaging techniques, pathology reports	Age, alcohol intake, BMI, sex	7

No.	Author, year	Country	Type of cancer	Data source	Study setting	Control group selection method	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	T2DM in Cases/ Controls	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	N O S
22	(Donadon, 2010)	Italy	HCC	Pordenone General Hospital	Hospital	Matched by age, sex, BMI, time of hospital admission	60-75 y	749 (78.4%)	Median Cases 25.6(2.9) Controls 25.1(2.4)	955	Cases 145/465 Controls 62/490	Laboratory test	Imaging techniques, pathology reports	Age, alcohol intake, BMI, sex	7
23	(Draho et al., 2016)	UK	EC	UK CPRD	Population	Ratio 5:1 Matched by date of diagnosis, birth year, sex, year of entry in CPRD, and general practice	Mean Cases 69.2y (11.3) Controls 68.9y (11.2)	2,793 (80%)	Obesity Cases 24.7% Controls 20.2%	3,493	Cases 44/592 Controls 212/2,901	Database	Database	Age, sex, years of CPRD data	8
24	(E. J. Walker et al., 2015)	USA	PC	University of California San Francisco medical clinics	Hospitals	Matched by age and sex	21-85 y Median Cases 63y Controls 60y	704 (50.1%)	Cases ≤25, 49.8%; 25-<30, 38.3%; ≥30, 11.9% Controls ≤25, 54.3%; 25-<30, 30.5%; ≥30, 15.2%	1,405	Cases 81/536 Controls 89/869	Interview (self-report)	Medical records, pathology reports, registry	Age, alcohol intake, BMI, family history of PC, diabetes drugs, history of pancreatitis, race, sex, smoking	6
25	(Lin et al., 2020)	Sweden	PrC	NPCR	Population	Ratio 5:1 Matched by year of birth and county of residence. Randomly selected	>85% of cases and controls older than 60 y	N/A	NR	186,227	Cases 4,186/31,415 Controls 24,965/154,812	Prescriptions, registry	Registry	Age, age at year of PrC onset, CCI, civil status, education level	7
26	(El-Serag, 2001)	USA	HCC	Department of Veterans Affairs	Population	Ratio 4:1 Randomly selected	Mean Cases 62y (10) Controls 60y (11)	4,177 (97.5%)	NR	4,282	Cases 274/823 Controls 1,035/3,459	Hospital records	Medical records	Age, comorbidities (AIDS, CVD, COPD, CHF, dementia, hemiplegia or paraplegia, MI, peptic ulcer disease, PVD, renal disease, rheumatologic disease), ethnicity	6
27	(F. Wang et al., 2006)	USA	PC	San Francisco Bay Area individuals	Population	Matched by sex and age. Randomly selected	21-85 y	1,434 (64.2%)	NR	2,233	Cases 58/532 Controls 110/1,701	Interview (self-report)	Pathology reports	Age, sex	6
28	(Fall et al., 2013)	Sweden	PrC	PCBase (based on NPCR-Swedish National Prostate Cancer Register)	Population	Ratio 5:1 Matched on calendar time, age, and county of residence. Randomly selected	Mean 71y (64-78)	N/A	NR	265,847	Cases 1,481/44,352 Controls 9,240/221,495	Prescriptions, registry	Laboratory test, registry	Age, age at prostate cancer diagnosis CCI, marital status, SES, diabetes prevalence in county of residence	7
29	(Feola et al., 2022)	Italy	EnC	Three Italian centres	Hospital	Matched by age and sex	Mean 58.3y (15.8)	Cases 74 (50%) Controls NR	Obesity Cases 44.1% Controls 28.1%	358	Cases 34/148 Controls 25/210	NR	Pathology reports	Age, family history of cancer, hypertriglyceridemia, obesity, sex	4
30	(Ferreira et al., 2021)	India	CRC	Hospital in Goa	Hospital	Ratio 1:1 Matched by age and sex	NR	132 (60%)	Cases Overweight 18.2% Obese 27.3%	220	Cases 33/110 Controls 16/110	Interview (self-report)	Pathology reports	Age, BMI, fruit/vegetable consumption, sex	7

No.	Author, year	Country	Type of cancer	Data source	Study setting	Control group selection method	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	T2DM in Cases/ Controls	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
									Controls Overweight 15.5% Obese 12.7%						
31	(Fortuny et al., 2005)	Spain	HLy, Ly, My	Centres: Barcelona, Tortosa, Reus, Madrid	Hospitals	Matched by age, sex, and study centre	Mean 59y	630 (54%)	NR	1,166	Cases 96/565 Controls 72/601	Interview for prescriptions (blinded), self-reported	Laboratory test, pathology records	Age, BMI, centre, SES, sex	7
32	(Fortuny et al., 2009)	USA	EndC	EDGE Study	Population	Matched by age and area. Randomly selected	Median Cases 61y (33-88) Controls NR	N/A	Cases <25, 25%; 25<30, 27%; 30<35, 17%; ≥35, 30% Controls <25, 45%; 25<30, 33%; 30<35, 15%; ≥35, 7%	936	Cases 76/469 Controls 38/467	Interview (self-report)	Pathology records, registry	Age, BMI, education level, endometrial fibrosis, oestrogen-related variables (menarche, HRT, OC, age at menopause, parity), family history of EndC, hypertension, osteoporosis, smoking, use of drugs (bisphosphonates, calcitonin, statins, sulfonylureas, insulin, biguanides, fibrates, antihypertensive, diuretics), race	8
33	(Garcia-Esquinas et al., 2016)	Spain	BC	MCC-Spain	Population	Matched by age and region. Randomly selected	Mean Cases Non-T2DM, 62.0y (8.7); T2DM, 68.9y (8.0) Controls Non-T2DM, 64.1y (9.14); T2DM, 71.7y (8.4)	N/A	Cases Non-T2DM <25, 37.8%; 25-30, 40.5%; ≥30, 21.7% T2DM <25, 17.3%; 25-30, 35.8%; ≥30, 46.9% Controls Non-T2DM <25, 46.4%; 25-30, 34.1%; ≥30, 19.5% T2DM <25, 19.2%; 25-30, 38.4%; ≥30, 42.4%	2,010	Cases 81/916 Controls 99/1,094	Interview (self-report)	Pathology records	Age, age at first birth, age of menarche, BMI, educational level, family history, previous biopsies	8
34	(Giraldi et al., 2021)	Italy	LC, PC	Three Italian centres	Hospitals	No	>50% older than 50 y	191 (44.2%)	Cases <18.5, 3.3%; 18.5<25, 20.7%; 25<30, 54.9%; ≥30, 21.2% Controls <18.5, 4.8%; 18.5<25, 27.8%; 25<30, 54.4%; ≥30, 12.1%	432	Cases 19/184 Controls 8/248	Interview (self-report)	NS	LC: age, family history of cancer, sex, smoking PC: age, family history of cancer, sex	5
35	(Gong et al., 2012)	China	EC, GaC, PC, BTC	Chinese PLA General Hospital	Hospital	No	≥30 y	Approx, 50% male	NR	4,065	Cases NR/3,314 Controls NR/751	Laboratory test	Pathology reports	NR	5

No.	Author, year	Country	Type of cancer	Data source	Study setting	Control group selection method	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	T2DM in Cases/ Controls	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	N O S
36	(H.-P. Chen et al., 2013)	Taiwan (Baur et al., 2011)	HCC	TNHIRD	Population	Ratio 2:1 Matched by age, sex, date of physician visit.	Mean Cases 62.49y (0.04) Controls 62.48y (0.03)	212,875 (72.8%)	NR	292,290	Cases 22,047/ 97,430 Controls 25,773/ 194,860	Database, prescriptions	Pathology records, database	Age, date of physician visit, sex	7
37	(Hassan et al., 2002)	USA	HCC	University of Texas M. D. Anderson Cancer Center	Hospital	Ratio 2:1 Matched by sex, age, and year of diagnosis	Mean 59.2y (10.9)	261 (75.7%)	NR	345	Cases 15/115 Controls 9/230	Medical records	Pathology reports	Age, alcohol intake, hepatitis virus infection, sex, smoking	5
38	(Hatia et al., 2023)	USA	ECC, ICC	MD Anderson Cancer Center, Texas	Hospital	Ratio 2.5:1 Matched by age and sex	Mean Cases 60.97y (12.34) Controls 59.66y (10.85)	Cases 232 (54%) Controls 670 (63.1%)	NR	1,490	Cases 82/ 429 Controls 109/ 1,061	Interview (self-report)	Pathology reports	Age, alcohol intake, race/ethnicity, family history of cancer, gallstones, sex, smoking, state of residency, viral hepatitis infection	6
39	(Henry et al., 2013)	USA	PC	Hospitals of the Twin Cities and the Mayo Clinic	Population	Matched by age and sex. Randomly selected	Mean Cases 65.4y (12.5) Controls 66.1y (11.4)	503 (57.6%)	NR	873	Cases 28/200 Controls 42/673	Interview (self-report)	Pathology reports	Age, alcohol intake, education level, race, sex, smoking	5
40	(Hossain et al., 2022)	USA	BC	LTR	Population	Ratio 5:1 Matched by age and race	≥ 20 y	N/A	Mild and high obesity approx. >40%	21,305	Cases 549/ 2,370 Controls 1,151/ 18,935	Database	Registry	Age, obesity, race	9
41	(Jordan et al., 2009)	Thailand	BC	TCS	Population-students	Matched by age. Randomly selected	28-51y	N/A	Cases <18.5,12%; 18.5-<23,43%; 23-<25,17%; ≥25,28% Controls <18.5,6%; 18.5-<23, 56% 23-<25, 19% ≥25,19%	903	Cases 3/43 Controls 10/860	Self-report	Medical records, self-report	Age, ever had children, height, income	4
42	(Khachatryan et al., 2011)	Armenia	BC	NOC, AAWC	Population	Randomly selected	35-70 y Mean Cases 55.79y (7.89) Controls 51.11y (9.94)	N/A	Cases Mean 29.03 (4.28) 19-<25,18.4%; 25-<30,39.7%; ≥30,41.9% Controls Mean 27.67 (4.57) 19-<25,27.6%; 25-<30,46.3%; ≥30,26.0%	302	Cases 22/150 Controls 5/152	Interview (self-report)	Hospital records	Age, BMI, oestrogen-related variables (abortion, age at first pregnancy, age at menarche, age at menopause, breastfeeding duration, HRT, live births)	8
43	(Khoshirsat S, Peyvandi AA, Oroei M,	Iran	TC	Loghman Hakim Hospital	Hospital	Matched by period	Mean Cases 44.17y (17.05)	673 (66.2%)	NR	1,016	Cases 12/87	Medical records	Medical records	Age, sex	5

No.	Author, year	Country	Type of cancer	Data source	Study setting	Control group selection method	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	T2DM in Cases/ Controls	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	N O S
	Peyvandi H, 2020)						Controls 44.14y (8.09)				Controls 36/929				
44	(La Torre et al., 2014)	Italy	PC	Teaching Hospital "Agostino Gemelli"	Hospital	Ratio 5:1 Matched by age, sex. Randomly selected	Mean Cases 63.59y (11.05) Controls 64.92y (10.95)	249 (52.7%)	NR	472	Cases 19/80 Controls 54/392	Interview (self-report), laboratory test, prescriptions	Pathology reports	Age, alcohol intake, sex, smoking, work activity	6
45	(La Vecchia et al., 1994)	Italy	EndC, LVC, PC	Teaching and general hospitals in the Greater Milan area	Hospitals	No	Median Controls 53 y (NR for cases)	6,326 (35.5%)	NR	17,825	Cases 233/9,991 Controls NR/7,834	Interview (self-report)	Pathology reports	Age, BMI, education, smoking	7
46	(La Vecchia et al., 1997a)	Italy	CC, CRC, RC	Hospitals of six Italian areas	Hospitals	No	Median Cases 62y (23-74) Controls 58y (20-74)	3,198 (52.3%)	CC cases <24, 32%; 24-27.1, 33%; ≥27.1, 34% RC cases <24, 33%; 24-27.1, 33%; ≥27.1, 34% Controls <24, 32%; 24-27.1, 34%; ≥27.1, 34%	6,107	Cases 109/1,953 Controls 158/4,154	Interview (self-report)	Pathology reports	Age, alcohol intake, BMI, center, education level, family history of CRC, fat, fibre intake, physical activity, sex, total energy	7
47	(La Vecchia et al., 1997b)	Italy	HCC	Teaching and general hospitals in the Greater Milan area	Hospitals	No	Median Cases 60y (23-74) Controls 56y (20-74)	1,418 (73.5%)	NR	1,930	Cases 52/428 Controls 64/1,502	Interview (self-report)	Pathology reports	Age, alcohol intake, area of residence, BMI, calendar year at interview, education level, history of hepatitis and liver cirrhosis, family history of LVC, sex, smoking	7
48	(Li L, Huang HL, Zhang XY, Zhang XY, 2011)	China	CRC	Three Hospitals in Jingdezhen	Hospitals	No	30-75 y	2,702 (55.6%)	NR	4,860	Cases 105/1,506 Controls 174/3,354	Medical records	Pathology reports	NR	4
49	(Liang et al., 2008)	China	EndC	Hospital of Guangxi Medical University	Hospital	No	Mean Cases 51.4y (10.3) Controls 52.1y (12.1)	N/A	NR	1,648	Cases 45/462 Controls 51/1,186	NR	Pathology reports	NR	3
50	(Lipworth et al., 2011)	Italy	PC	Major teaching and general hospitals	Hospitals	No	Median Study 1 Cases, 59 y Controls, 55 y Study 2 Cases and controls, 63 y	1,892 (65.4%)	NR	2,892	Cases 68/688 Controls 85/2,204	Interview (self-report)	Imaging techniques, pathology reports	Age, BMI, education level, sex, study centre, smoking, year of interview	7

No.	Author, year	Country	Type of cancer	Data source	Study setting	Control group selection method	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	T2DM in Cases/ Controls	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	N O S
51	(Lu et al., 2015)	UK	PC	Health Improvement Network	Population	Matched by sex, age, and calendar year. Randomly selected	≥75% older than 65 y	3,166 (57.3%)	Cases <20, 3.2%; 20-25, 29.7%; 25-30, 31.9%; ≥30, 22.2%; Missing, 13.0% Controls <20, 3.2%; 20-25, 26.0%; 25-30, 35.1%; ≥30, 20.4%; Missing, 15.3%	5,529	Cases 175/529 Controls 856/5,000	Database	Medical records	Age, alcohol intake, BMI, calendar year, GP visit, sex, smoking, TDI, time since diagnosis of diabetes	9
52	(Macías-Gómez et al., 2015)	Mexico	BC	Alfredo Pumarejo Hospital	Hospital	No	Mean Cases 47.2 y Controls 50.7 y	N/A	NR	770	Cases 79/389 Controls 57/381	NR	NR	NR	1
53	(Maisonneuve et al., 2010)	Australia, Canada, Netherlands, Poland	PC	SEARCH	Population	Matched by age and sex. Randomly selected	Approx. ≥ 40y	1,314 (52.5%)	NR	2,502	Cases NR/823 Controls NR/1,679	Interview (self-report)	Registry	Age, center, education level, response status, sex, smoking	6
54	(Miele et al., 2015)	Italy	HCC	Teaching Hospital “Agostino Gemelli”	Hospital	No	Approx. ≥ 60 y	402 (65.6%)	NR	613	Cases 69/224 Controls 52/389	Interview (self-report)	Pathology reports	Age, alcohol intake, sex, smoking	5
55	(Mikael et al., 2021)	Australia	CRC	SAYO	Population	No	Median Cases 44y (18-54) Controls 45y (18-54)	155 (47%)	NR	330	Cases 17/90 Controls 12/240	Interview (self-report) and medical records	Medical records	NR	6
56	(Mkuu et al., 2023)	USA	CeuC	Electronic health records (EHR) data from the OneFlorida+ Data Trust	Population	Ratio 4:1 Matched by year of encounter / diagnosis and age	≥ 25 y	N/A	NR	503,695	Cases 2,833/100,739 Controls 9,271/402,956	Database	Database	Age, insurance type, race/ethnicity, rurality, social vulnerability	6
57	(Nodari et al., 2023)	11 European countries (Greece, Italy, Germany, Netherlands, Denmark, Czech Republic, Hungary, Poland, Ukraine, Lithuania, UK, Japan, Brazil)	PC	PANDoRA consortium	Population	No	Median 45y (30-81)	3,584 (55%)	NR	2,009	Cases 550/1,498 Controls 72/511	Database	Database	Age, sex	4
58	(Oberman et al., 2015)	USA	TC	Penn State Hershey Medical Center	Hospital	Matched by calendar date. Randomly selected	Mean Cases 49.13y (14.73) Controls 53.81y (15.22)	70 (27.7%)	Mean Cases 37.83 (6.95) Controls 29.53 (8.26)	220	Cases 14/48 Controls 33/172	NR	Pathology reports	Age, BMI, race, sex	4
59	(Parazzini et al., 1999)	Italy	EndC	Hospitals in the Greater Milan area	Hospitals	No	Median Cases	N/A	Cases <25, 39.6%; 25-30, 33.9%;	3,339	Cases 128/748	Interview (self-report) and medical records	Pathology reports	Age, calendar year at interview, education level, oestrogen-related	8

No.	Author, year	Country	Type of cancer	Data source	Study setting	Control group selection method	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	T2DM in Cases/ Controls	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
							60y (28-74) Controls 54y (25-74)		≥30, 26.5% Controls <25, 60.6%; 25-30, 29.8%; ≥30, 9.6%		Controls 101/ 2,591			variables (age at menopause, oral contraceptive and HRT, parity), hypertension, Quetelet's index, smoking	
60	(Petrick et al., 2019)	Australia, Canada, Netherlands, UK, USA	EC, EJA	BEACON studies	Population	Ratio 2:1 in NLCS; 4:1 in NIHAARP; and 8:1 in MHC.	Mean Cases EC 62.3y (8.8); EJA 62.0y (8.7) Controls 60.1y (9.4)	11,852 (73.7%)	Mean EC 28.0 (4.9) EJA, 27.4(4.7) Controls 26.4 (4.4)	16,088	Cases 270/ 4,247 Controls 1,003/ 11,841	Self-reported	Medical records, pathology reports, registry	Age, BMI, sex, smoking	7
61	(Pierce et al., 2008)	USA	PrC	HIM, PROS, SEER	Population	Matched by age and period. Randomly selected	≥ 35y	N/A	Cases <25, 31.1%; 25-<30, 49.0%; ≥30, 20.0% Controls <25, 30.1%; 25-<30, 47.5%; ≥30, 22.5%	3,396	Cases 147/ 1,752 Controls 136/ 1,644	Interview (self-report)	Pathology reports and registry	Age, BMI, family history, PrC screening history, race, study	8
62	(Rollison et al., 2008)	USA	BC	4 Corners Breast Cancer Study	Population	Matched by age. Randomly selected	Mean Non-T2DM 55.5y (12.1) T2DM 57y (12.3)	N/A	<20 Non-T2DM 61.1% T2DM 53.9% 20-25 Non-T2DM 29.3% T2DM 34.2% ≥25 Non-T2DM 3.4% T2DM 7.3%	4,847	Cases 204/ 2,324 Controls 247/ 2,523	Interview (self-report)	Registry	Age, BMI, family history of BC, oestrogen-related variables (age at menarche, age at first pregnancy, breastfeeding history, menopausal status, number of full-term pregnancies), physical activity, study site	7
63	(Rosato et al., 2011)	Italy	EndC	Hospitals of three Italian areas	Hospitals	Matched by age, study center, and period of enrolment	Median Cases 60y (18-79) Controls 61y (19-79)	N/A	Cases ≤30, 68.7%; >30, 31.3% Controls ≤30, 88.2%; >30, 11.8%	1,252	Cases 53/454 Controls 54/798	Interview (self-report)	Pathology reports	Age, education level, oestrogen-related variables (age at menarche, menopausal status, oral contraceptive, HRT, parity), study center, year of interview	6
64	(Rosato et al., 2016)	Italy, Spain	CC, CRC, RC	HIWATE Project	Hospital/ Population	No	Median Italy Cases 67y (35-80); Controls 66y (31-80) Spain Cases	1,748 (63.7%)	Cases <25, 34.8%; 25-30, 44.0%; ≥30, 21.1%; Missing, 0.1% Controls <25, 34.1%; 25-30, 44.3%;	2,741	Cases 159/ 1,147 Controls 188/ 1,594	Interview (self-report)	Pathology reports	Age, alcohol intake, BMI, education level, physical activity, sex, smoking, study center, use of aspirin or statin	7

No.	Author, year	Country	Type of cancer	Data source	Study setting	Control group selection method	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	T2DM in Cases/ Controls	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	N O S
							68y (22-85); Controls 66y (28-85)		≥30, 21.3%; Missing, 0.3%						
65	(Safaei et al., 2009)	Iran	CRC	RCGLD	Hospital	Randomly selected	≥ 40y	877 (50.8%)	Cases <18.5, 5.8%; 18.5-25, 41.1%; 25-30, 46.2%; ≥30, 7.0% Controls <18.5, 2.3%; 18.5-25, 34.2%; 25-30, 46.8%; ≥30, 16.7%	1,724	Cases 86/862 Controls 15/862	Registry	Pathology reports	Age, BMI, sex, smoking	6
66	(Salinas-Martinez et al., 2014)	Mexico	BC	Center of the public health	Hospital	No	Mean Cases 57.4y (12.0) Controls 55.2y (9.1)	N/A	Cases Central obesity, 85.5% 18-<25, 16.0%; 25-<30, 35.4%; ≥30, 48.5% Controls Central obesity, 68.8% 18-<25, 17.1%; 25-<30, 39.7%; ≥30, 43.2%	646	Cases 117/240 Controls 138/406	Interview (self-report) and laboratory test	Pathology reports	Age, oestrogen-related variables (breastfeeding, HRT, maternal age, parity), obesity	6
67	(Saltzman et al., 2007)	USA	EndC	Based case-control studies: King, Pierce, or Snohomish County	Population	Matched by diagnosis year, age, and county	45-74 y	N/A	Cases <25, 42.8%; 25-<30, 29.1%; 30-<35, 12.9%; ≥35, 15.2% Controls <25, 59.3%; 25-<30, 25.7%; 30-<35, 10.2%; ≥35, 4.8%	3,055	Cases 118/1,289 Controls 72/1,766	Interview (self-report)	Pathology reports	Age, BMI, county, HTR, reference year	8
68	(Sanderson et al., 2010)	USA	BC	Clinic (Lower Rio Grande Valley, Texas)	Hospital	Ratio 4:1. Matched by age. Randomly selected.	30-79 y Mean Cases 54.3y (11.1) Controls high risk 49.2y (10.1); low-risk 55.1y (8.4)	N/A	Cases <25, 11.3%; 25-<30, 29.6%; 30-<35, 30.7%; ≥35, 28.5% Controls High risk <25, 16.1%; 25-<30, 33.5%; 30-<35, 30.1%; ≥35, 20.3% Low risk <25, 13.4%; 25-<30, 29.7%; 30-<35, 31.4%; ≥35, 25.5%	1,169	Cases 47/190 Controls 241/979	Laboratory test, medical records, interview (self-report)	Pathology reports	Age, mammography screening, menopausal status	7

No.	Author, year	Country	Type of cancer	Data source	Study setting	Control group selection method	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	T2DM in Cases/ Controls	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
69	(Schumacher et al., 2021)	USA	CC, CRC, RC	KPSC	Population	Ratio 5:1 Matched by age, sex, length of membership	Mean Cases 43.0 y (5.70) Controls 43.5 y (5.73)	3,078 (50%)	Obesity Cases 48.3% Controls 39.3%	6,160	Cases 104/1,032 Controls 477/5,128	Laboratory test, medical records, prescriptions	Medical records, registry	Age, census, dyslipidaemia, education level, ER visit/hospitalization, family history, hypertension, income level, number of outpatient visits, obesity, race/ethnicity, sex, smoking	9
70	(Seo et al., 2017)	Korea	TC	T-CALOS study	Hospitals	Matched by age and sex	Mean 54.9y (10.2)	398 (48%)	Obesity Cases 33.7% Controls 28.0%	830	Cases 50/415 Controls 67/415	Interview (self-report) and laboratory test	Imaging techniques	Age, alcohol intake, comorbidities (benign thyroid disease, CHD, dyslipidaemia, hypertension), education level, income, obesity, regular exercise, sex, smoking, other cancers	8
71	(Sheng et al., 2023)	China	GB	Hospital of Medical School, Xi'an Jiaotong University	Hospital	Matched by age, ethnicity, occupation, alcohol consumption	Mean 61.4y (11.2)	1,965 (45%)	Mean Cases 24.45 (6.0) Controls 24.12 (3.0)	4,420	Cases 272/2,210 Controls 223/2,210	Interview (self-report)	Pathology reports	Age, alcohol intake, BMI, ethnicity, fasting blood glucose, fasting blood insulin, gallbladder stone disease, HDL, hypertension, NAFLD, occupation, retinol binding- protein, triglyceride	6
72	(Shoff & Newcomb, 1998)	USA	EndC	Residents of Wisconsin	Population	Matched by age. Selected randomly	Mean Cases Non-T2DM 62.6y (10.1); T2DM 64.8y (8.7) Controls Non-T2DM 63.1 y (8.4); T2DM 65.7y (7.3)	N/A	Mean Cases Non-T2DM 29.1 (7.4) T2DM 34.5 (7.0) Controls Non-T2DM 26.1 (5.0) T2DM 29.1 (5.5)	3,014	Cases NR/723 Controls NR/2,291	Interview-blinded (self-report)	Pathology reports, registry	Age, BMI, education level, oestrogen- related variables (HRT, parity), smoking	9
73	(Stott-Miller et al., 2012)	Germany, Italy, Japan, Switzerland, USA	H&N	Pooled analysis of 12 studies	Hospitals/ Population	Matched by age, sex, neighbourhood, centre, race/ ethnicity	17-93 y	1,462 (74%)	Cases <18.5, 4.6%; 18.5-<25,50.1%; 25-<30,32.7%; ≥30, 12.5% Controls <18.5, 2.1%; 18.5-<25,46.5%; 25-<30,37.9%; ≥30,13.5%	14,413	Cases 258/4,944 Controls 463/9,469	Interview (self-report)	Registry	Age, alcohol intake, BMI, education level, race/ethnicity, sex, smoking, study center	7
74	(Tabassum et al., 2016)	Pakistan	BC	NORI	Hospital	No	30-70 y	N/A	NR	400	Cases 27/200 Controls 10/200	Interview (self-report)	NR	Age at diagnosis, BMI, dietary fat, oestrogen-related variables (birth control pills, HRT, menopausal status), physical activity, smoking	4
75	(Talamini et al., 1997)	Italy	BC	Teaching and general hospitals	Hospitals	No	Median Cases 55y (23-74)	N/A	NR	5,157	Cases 107/2,569	Interview (self-report)	Pathology reports	Age, BMI, education level, menopausal	7

No.	Author, year	Country	Type of cancer	Data source	Study setting	Control group selection method	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	T2DM in Cases/ Controls	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
							Controls 56y (20-74)				Controls 86/2,588			status, parity, study area	
76	(Tavani et al., 2002)	Greece, Italy	PrC	Major teaching and general hospitals	Hospitals	No	Median Greek study Cases 70y Controls 71y Italian study Cases: 67y Controls: 60y	N/A	NR	1,616	Cases 40/608 Controls 36/1,008	Interview (self-report)	Pathology reports	Age, BMI, calendar year at interview, centre, education level	6
77	(Tavani et al., 2005)	Italy	PrC	Major teaching and general hospitals (four Italian areas)	Hospitals	No	Median Cases 66y (46-74)	N/A	NR	2,765	Cases 77/1,294 Controls 89/1,471	Interview (self-report)	Pathology reports	Age, BMI, calorie intake, centre, education level, family, history of PrC, smoking	7
78	(Turati et al., 2015)	Italy	BlaC	Centres: Aviano, Milan, Naples, Catania	Hospitals	Matched by study centre, sex, and age	Median Cases 67y (25-84) Controls 66y (27-84)	1,156 (85.3%)	NR	1,355	Cases 106/690 Controls 55/665	Interview (self-report)	Pathology reports	Age, education level, sex, smoking, study centre, year of interview	6
79	(Weiderpass et al., 2000)	Sweden	EndC	Six regional cancer registries	Population	Matched by age	50-74 y Mean Cases 65.4y (6.0) Controls 64.0y (6.6)	N/A	Cases <22.5, 16.5%; 22.5-<25, 21.7%; 25-<28, 20.3%; 28-<30, 12.6%; 30-<34, 16.9%; ≥34, 12.0% Controls <22.5, 22.9%; 22.5-<25, 27.6%; 25-<28, 26.5%; 28-<30, 10.3%; 30-<34, 8.0%; ≥34, 3.3%	3,581	Cases 76/702 Controls 159/2,879	Self-reported	Pathology reports	Age, BMI, oestrogen-related variables (age at last birth, age at menarche, age at menopause, HRT, oral contraceptives, parity), smoking	7
80	(Weiss et al., 1999)	USA	BC	US Cancer registries	Population	Matched by geographical area and age. Randomly selected	20-54 y	N/A	NR	4,138	Cases 33/2,158 Controls 31/1,980	Interview (self-report)	Registry	Age, alcohol intake, BMI, oestrogen-related variables (age at first birth, menopausal status, number of births), family history, number of mammograms, previous breast biopsy, race, site	7
81	(Welzel et al., 2007)	USA	ECC, ICC	SEER registries	Population	Matched by geographical area. Randomly selected	Mean Cases 79 y Controls 77.1 y	37,680 (36.3%)	Obesity Cases ECC, 2.9% ICC, 4.3% Controls 3.1%	103,866	Cases 342/1,084 Controls 22,764/102,782	Database	Registry	Age, geographic location, race, sex, state buy-in status	8
82	(Wu et al., 2007)	USA	BC	Los Angeles County Cancer Surveillance Program	Population	Matched by age, Asian ethnicity and neighbourhood of residence	25-74 y	N/A	Cases <20.4, 24.8%; >20.4-22.7, 26.2%; >22.7-24.6 21.0%; ≥24.6, 27.5%	2,346	Cases 90/1,224 Controls 48/1,122	Interview (self-report)	Pathology reports	Age, Asian ethnicity, birthplace, BMI, education level, oestrogen-related variables (age at	7

No.	Author, year	Country	Type of cancer	Data source	Study setting	Control group selection method	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	T2DM in Cases/ Controls	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
									Controls <20.4, 24.9%; >20.4-22.7, 30.8%; >22.7-24.6, 19.3%; ≥24.6, 24.7%					menarche, age at menopause, parity, type of menopause), marital status, physical activity, soy and tea intake, years of residence in the USA, WHR	
83	(Y. X. Yang et al., 2005)	UK	CC, CRC, RC	GPRD	Population	Ratio 10:1 Matched by year of birth, enrolment date, and duration of GPRD follow-up	NR	49,406 (43%)	Cases <20: 4.6%; ≥20-<25, 25.0%; ≥25-<30, 23.4%; ≥30, 8.7%; Missing, 38.2% Controls <20: 4.1%; ≥20-<25, 23.1%; ≥25-<30, 22.9%; ≥30, 9.2%; Missing, 40.6%	114,876	Cases 289/ 10,447 Controls NR/ 104,429	Database	Database	Age, BMI, duration of follow-up, enrolment year	8
84	(Z. Li et al., 2022)	USA	CC, CRC, RC	IBM® MarketScan® Commercial Database	Population	Ratio 8:1 Matched by age, sex, geographical region, duration of insurance enrolment before index diagnosis, prescription drug coverage	Mean Cases 43.0y (5.8) Controls 42.8y (5.8)	29,479 (50.7%)	Obesity Cases 7.1% Controls 5.7%	58,105	Cases 302/ 6,001 Controls 1,946/ 52,104	Database	Pathology reports	Age, IBD, CCI, duration of insurance enrolment, employment status, family history of GI, faecal occult blood test, health plan, obesity, prescription drug coverage, region, residence, screening colonoscopy, sex	8
85	(Zhu, 2004)	USA	PrC	Physicians' Health Study	Population	Matched by age and randomization date. Randomly selected	40-84 y	N/A	Mean 24.9 (2.87)	2,220	Cases 32/1,110 Controls 44/1,110	Interview (self-report)	Medical records	Age, aspirin/beta-carotene use, race	5
86	(Zucchetto et al., 2007)	Italy	KC	Teaching and general hospitals of Pordenone, Gorizia, Milan, Latina, Naples	Hospitals	Ratio 2:1 Matched by age, sex and study centre	Median 62y (22-79)	1,482 (64.4%)	Cases <25, 64:12%; 25-<30, 25.3%; ≥30, 5.0% Controls <25, 65:7%; 25-<30, 22.4%; ≥30, 3.7%	2,301	Cases 37/767 Controls 61/1,534	Interview (self-report)	Pathology reports	Age, BMI, calendar period of interview, center, education level, family history of KC, sex, smoking	7

Ratio: controls matched to cases. Age: values are presented as range, mean (SD), median (IQR), depending on what was reported in the publication. BMI: Values are presented as percentages, mean (SD), median (IQR), categories (overweight [BMI > 25-29.9 kg/m²], obesity [BMI > 30 kg/m²]), depending on what was reported in the publication. Sample size: total number of subjects (cases and controls). T2DM in Cases (No. T2DM/No. cases)/Controls (No. T2DM/No. controls). NOS: range, 1-9 points. Adjustment variables: the maximum adjustment reported, but for quantitative analyses, the effect estimates based on the selection of the different adjustment variables were extracted separately. BMI: body mass index; IQR: interquartile range; N/A: not applicable; NOS: New-castle Ottawa Scale; NR: not reported; SD: standard deviation; y: years. Abbreviations for cancer types, data sources, and adjustment variables are shown in the "Abbreviations" section.

Table A 7-4 Main characteristics of the cohort studies included in the systematic review

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
1	(Anderson et al., 2001)	USA	EndC	IWHS	Population	12y (P)	55-69 y Mean Non-T2DM 61.7y T2DM 62.5y	N/A	Mean Non-T2DM 26.6 T2DM Insulin users, 31.3 Non-insulin users, 30.5	24,664	T2DM 34/1,325 Non-T2DM 307/23,150	Self-reported	Registry	Age, BMI, oestrogen-related variables (gravidity, HRT, menstrual irregularities, ovulatory span), hypertension, WHR	7
2	(Arthur et al., 2019)	USA	PC	WHI cohort	Population	Median 17.9y (9.0-19.4) (P)	50-79y	N/A	Mean Cases 28.1(5.9) Controls 28.0 (5.9)	155,075	T2DM 71/9,130 Non-T2DM 973/145,945	Self-reported	Medical records and pathology records	Age, alcohol intake, allocation, educational level, metabolic equivalent task hrs/week, race, smoking, WHR	8
3	(Ballotari et al., 2017)	Italy	BC, BlaC, BTC, CRC, EndC, GaC, KC, LC, LVC, NHL, OC, PC, PrC, TC	Registry of Inhabitants of the Reggio Emilia province	Population	4y (P)	20-84 y Median Non-T2DM 45y (35-60) T2DM 67y (58-75)	200,958 (49.3%)	NR	407,157	T2DM 1,464/23,358 Non-T2DM 9,558/383,799	Registry	Registry	Age, foreign status, sex	7
4	(Bao et al., 2018)	Sweden	BC, BlaC, BraNSC, LC, CC, CeuC, EC, EnC, GaC, GB, KC, LaC, Le, LiC, LVC, Ly, M, NM, OC, Oral, PhC, PrC, PC, RC, SiC	STR	Population	16y (P)	Mean Non-T2DM 62.7y (9.3) T2DM 60.9y (8.0)	11,900 (47.3%)	Mean Non-T2DM 24.8(3.3) T2DM 27.5 (4.1)	25,154	T2DM 301/1,766 Non-T2DM NR/23,388	Registry, medical records	Registry	Age, alcohol intake, BMI, education level, marital status, sex, smoking	7
5	(Berger et al., 2016)	Denmark	Overall	Nationwide registers	Population	Mean 12.6y (5.7) (R)	20-89 y Mean 41.4y (18.9)	2,403,419 (49.8%)	NR	4,826,142	T2DM NR/65,690 Non-T2DM NR/4,760,452	Prescriptions	Registry	Age, calendar year	7
6	(Bjornsdottir et al., 2020)	Sweden	BlaC, Bone, BC, BraNSC, CeuC, CRC, EC, Eye, GaC, GB, HLy, KC, LaC, Le, Lip, LVC, LC, M, Mouth, My, NHL, Nose, OC, Oral, PC, Penis, PhC, Pleura, PrC, RC, SiC, ST, TeC, TC, Tongue, Uterine	Swedish NDR	Population	Median 6.6y (R)	Mean 65.2y (12.6)	12,724,487 (54.53%)	Mean Non-T2DM NR T2DM 29.7 (5.4)	23,334,838	T2DM NR/457,473 Non-T2DM NR/22,877,365	Registry	Registry	Age, education, geographical region of birth, income, marital status, sex	8

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
7	(Bogumil et al., 2023)	USA	BlaC	MEC	Population	Mean 19.7y (P)	45-75 y	83,766 (45.2%)	NR	185,059	T2DM NR/65,278 Non-T2DM NR/119,781	Self-reported, database	Registry	Age, alcohol intake, BMI, oestrogen/progesterone, menopause status, parity, race/ethnicity, sex, smoking	9
8	(Bosco et al., 2012)	USA	BC	BWHS	Population	Mean 10.5y (2.9) (P)	21-69 y	N/A	Non-T2DM <25, 56.5%; 25-29, 32.3%; ≥30, 10.4% T2DM <25, 8.5%; 25-29, 24.3%; ≥30, 66.2%	49,172	T2DM 110/1,900 Non-T2DM 1118/47,272	Self-reported	Medical records or cancer registry	Age, BMI, education level, individual cardiometabolic factors, physical activity	8
9	(Boustany et al., 2023)	USA	CRC	The Explorys database	Population	20y (R)	18-65 y	21,390,330 (45%)	NR	47,714,750	T2DM 14,050/2,193,180 Non-T2DM 68,370/45,521,570	Database	Database	Age, alcohol intake, IBS, history of H pylori infection, obesity, sex, smoking	9
10	(Bowker et al., 2011)	Canada	BC	BCLHD	Population	10y Mean T2D 4.4y (2.9) Non-T2DM 4.5y (2.9) (R)	Mean 61.8y (14.2)	N/A	NR	169,012	T2DM 1,178/84,506 Non-T2DM 1,203/84,506	Database	Database	Age, number of physician visits, SES, year of diagnosis	7
11	(Brinton et al., 2010)	USA	BC	U.S. Veterans Affairs computerized medical care system database	Population	Mean Cancer 7.8y Non-cancer 11.7y (R)	Mean With cancer 58.1 y Without cancer 52.5 y	N/A	NR	4,501,578	NR	Database	Database	Age, calendar year, latency between study entry and exit, number of hospital visits, race	8
12	(Brodovicz et al., 2012)	UK	PC	UK GPRD	Population	6 y (R)	Mean Non-T2DM 46.35y (17.81) T2DM 63.88 y (13.92)	1,453,895 (49%)	Obesity Non-T2DM, 4.49% T2DM 24.57%	2,965,233	T2DM 436/146,692 Non-T2DM 1,467/2,818,541	Database, medical records, prescriptions	Database, medical records, prescriptions	Age, alcohol intake, CCI, history of chronic pancreatitis, gallbladder disease, obesity, sex, smoking	8
13	(Bronsveld et al., 2017)	UK	BC	CPRD	Population	1,582,751 py (R)	Median 64y (51-74)	N/A	Non-T2DM <20, 7.9%; 20-<25, 32.3%; 25-<30, 32.2%; 30-35, 17.0%; ≥35: 10.6%; Missing, 76.0% T2DM <20, 1.9%; 20-<25, 13.1%; 25-<30, 27.5%;	295,996	T2DM 2,371/147,998 Non-T2DM 2,252/147,998	Medical records, prescriptions	Medical records	Age	6

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
									30-35,26.3%; ≥35: 31.2%; Missing,45.1%						
14	(C. B. Chen et al., 2018)	Canada	PrC	Population Data British Columbia database	Population	Median T2DM 9.2y Non-T2DM 9.4y (R)	Mean 64.7 y (9.4)	N/A	NR	160,566	T2DM 3,557/ 80,001 Non-T2DM 4,370/ 80,565	Database, prescriptions	Database	Age, number of physician visits, SES	8
15	(C. Tseng, 2013)	Taiwan	HCC	NHI reimbursement databases	Population	3y (R)	≥ 18y	494,080 (49.6%)	NR	996,921	T2DM 909/ 114,979 Non-T2DM 1,414/ 881,942	Database	Database	Age, alcohol-related diagnoses, antidiabetic drugs, comorbidities (CHD, COPD, dyslipidaemia, eye disease, hypertension, nephropathy, nonalcohol-related chronic liver disease, PAD, stroke), living region, obesity, occupation, potential detection bias	7
16	(C. Tseng, 2014)	Taiwan	BC	NHI reimbursement databases	Population	3y (R)	Approx. ≥50 y	N/A	NR	501,747	T2DM 326/ 62,752 Non-T2DM 871/ 438,995	Database	Database	Age, comorbidities (alcohol-related diagnosis, benign breast conditions CHD, COPD, dyslipidaemia, eye disease, hypertension, nonalcohol-related chronic liver disease, nephropathy, PAD, postmenopausal syndrome, stroke), drug use (antidiabetic, antihypertensives, antiplatelet, fibrates, oestrogen, statins), living region, obesity, occupation, potential detection examinations	7
17	(C.-C. Lin, Chiang, Li, Hsieh, et al., 2014)	Taiwan	PC	NHI program	Population	11y (R)	Mean Non-DM 50.47y (12.07) T2DM 50.71y (11.80)	185,404 (55.56%)	Morbid obesity Non-T2DM 0.07% T2DM 0.36%	333,700	T2DM NR/166,850 Non-T2DM NR/166,850	Database	Database	Age, insurance premium, residential area, sex, urbanisation indicator	7
18	(C.-C. Lin, Chiang, Li, Liu, et al., 2014)	Taiwan	BC, Blac, Bone, BransC, C. in situ, CeuC, CRC, EC, EndC, GaC, GB, KC,	NHI program	Population	10y (R)	≥20y	5,419,238 (54.8%)	NR	9,884,228	T2DM 83,694/ 472,979 Non-T2DM	Database	Database	Age, area registered for NIH program, sex	6

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
			HLy, LC, LaC, Le, LVC, OC, Oral, PC, Penis, Peritoneum PhC, Placenta, Respiratory, intrathoracic organs, SiC, Skin, ST, TeC								667,889/ 9,411,249				
19	(C.-H. Chen et al., 2019)	Taiwan	CC, CRC, RC	NHI database	Population	12y Mean Non-T2DM 6.97y (4.00) T2DM 6.75y (3.91) (R)	≥20y Mean Non-T2DM 56.8y (13.6) T2DM 57.6y (13.2)	70,967 (53.5%)	Obesity Non-T2DM 4.59% T2DM 1.65%	132,648	T2DM 819/66,324 Non-T2DM 627/66,324	Database	Database	Age, alcohol-related illness, CAD, CKD, cirrhosis, colorectal adenomas, hyperlipidemia, hypertension, IBD, IBS, examinations (colonoscopy, sigmoidoscopy, stool guaiac, test and faecal immunochemical test), obesity, urbanization level, sex, stroke	9
20	(C.-H. Tseng, 2011a)	Taiwan	BlaC	NHI reimbursement databases	Population	3y (R)	≥40 y	495,199 (45.6%)	NR	998,947	T2DM 227/ 115,731 Non-T2DM 362/ 883,216	Database	Database	Age, comorbidities (COPD, CHD, eye disease, dyslipidaemia, hypertension, PAD, stroke), drug use (antihypertensive, fibrate, statin), living region, occupation, sex	6
21	(C.-H. Tseng, 2011b)	Taiwan	PrC	NHI reimbursement databases	Population	3y (R)	≥40 y	N/A	NR	494,630	T2DM 362/52,131 Non-T2DM 527/ 442,499	Database	Database	Age, comorbidities (CHD, COPD, dyslipidaemia, eye disease, hypertension, nephropathy, PAD, stroke), drug use (antidiabetic, antihypertensive, fibrate, statin), living region, obesity, occupation	7
22	(C.-H. Tseng, 2012a)	Taiwan	NHL	NHI reimbursement databases	Population	1y (R)	Mean 59.7y (11.4)	163,092 (49.7%)	NR	328,267	T2DM 64/90,222 Non-T2DM 84/238,045	Database	Database	Age, living region, occupation, sex	6
23	(C.-H. Tseng, 2012b)	Taiwan	TC	LHID 2005	Population	10y (R)	NR	495,673 (49.5%)	NR	999,730	T2DM 149/ 115,928 Non-T2DM 794/ 883,802	Database	database	Age, comorbidities (benign thyroid disease, CHD, COPD, dyslipidaemia, eye disease, hypertension, nephropathy, PAD, stroke, other cancer), drug use	8

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
24	(C.-H. Tseng, 2013)	Taiwan	Oral	NHI reimbursement databases	Population	3y (R)	>25y	494,817 (49.5%)	NR	998,540	T2DM 214 / 115,691 Non-T2DM 552 / 882,849	Database	Registry, database	(antidiabetic, antihypertensive, antiplatelet drugs, fibrate, NSAID, statin), living region, obesity, occupation, potential detection examinations, sex Age, alcohol-related diagnoses, comorbidities (CHD, COPD, dyslipidaemia, eye disease, hypertension, nephropathy, PAD, stroke), drugs use (antidiabetic, antihypertensive, fibrates, statins), living region, obesity, occupation, potential oral cancer detection	7
25	(C.-H. Tseng, 2014)	Taiwan	LC	NHI reimbursement databases	Population	1y (R)	Approx. ≥40 y	494,014 (49.5%)	NR	996,975	T2DM 321 / 114,920 Non-T2DM 450 / 882,055	Database	Database	Age, chest X-ray examination (CXR), comorbidities (CHD, COPD, dyslipidaemia, eye disease, hypertension, nephropathy, PAD), diabetes duration, drug use (antidiabetic, antihypertensive, fibrate, statin), living region, obesity, occupation	7
26	(C.-H. Tseng, 2015)	Taiwan	KC	NHI reimbursement databases	Population	3y (R)	Mean Non-T2DM 33y (19.4) T2DM 57.3y (16.7)	495,114 (49.5%)	Obesity Non-T2DM 0.6% T2DM 2.5%	998,728	T2DM 193 / 115,655 Non-T2DM 292 / 883,073	Database	Database	Age, comorbidities (alcohol-related diagnoses, CHD, COPD, dyslipidaemia, end-stage renal disease, eye disease, hypertension, nephropathy, stroke, PAD), diabetes duration, drug use (antidiabetic, antihypertensive, fibrate, statin), living region, obesity, occupation, sex, smoking	7
27	(C.-S. Wang et al., 2009)	Taiwan	HCC	A-Lein Township, Taiwan's Kaohsiung County	Population	8y (50,889 py) (P)	≥35 y	2,579 (43.5%)	<30, 86.3% ≥30, 13.6%	5,929	T2DM NR / 5383 Non-T2DM NR / 546	Laboratory test, prescriptions	Imaging techniques, registry	Age, alcohol intake, BMI, sex, smoking	7

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
28	(Campbell et al., 2010)	USA	CC, CRC, RC	CPS-II Nutrition Cohort	Population	15y (P)	50-74 y	73,312 (47.3%)	Male Non-T2DM <18.5, 0.5%; 18.5-<25, 35.9%; 25-<30, 48.9%; 30-<35, 11.1%; ≥35, 2.1% T2DM <18.5, 0.5%; 18.5-<25, 25.9%; 25-<30, 47.6%; 30-<35, 18.5%; ≥35, 5.8% Female Non-T2DM <18.5, 1.9%; 18.5-<25, 51.5%; 25-<30, 30.8%; 30-<35, 10.3%; ≥35, 3.9% T2DM <18.5, 0.9%; 18.5-<25, 29.7%; 25-<30, 33.8%; 30-<35, 20.7%; ≥35, 13.0%	154,975	T2DM 335/11,335 Non-T2DM 2,474/143,640	Self-reported and medical records	Self-reported confirmed with medical records or cancer registry	Age, alcohol intake, BMI, education level, endoscopy history, family history of CRC, NSAID use, physical activity	9
29	(Campbell et al., 2016)	USA	LVC	LCPP	Population	12-30y (P)	Mean 58.2y	640,367 (40.7%)	Mean 26.6	1,570,023	T2DM 449/102,051 Non-T2DM 1,697/1,467,972	Self-reported	Registry and medical records	Age, alcohol intake, BMI, race, sex, smoking, study	7
30	(Cavichia et al., 2013)	USA	CC, CRC, RC	South Carolina Medicaid	Population	18y (R)	≥30 y Mean With cancer 58.7y (0.55) Without cancer 59.1y (0.06)	27,294 (29.7%)	NR	91,846	T2DM 59/6,019 Non-T2DM 618/85,827	Medical records	Registry	Age, hypertension, marital status, number of months Medicaid eligibility, race, sex	7
31	(Cerhan et al., 1997)	USA	NHL	IWHS	Population	7y 251,489 py (P)	55-69 y	N/A	NR	37,934	T2DM 13/NR Non-T2DM 99/NR	Self-reported	Registry	Age, marital status, history of cancer, residence, transfusion history	6
32	(Chodick et al., 2010)	Israel	Anal, BC, Bone, BraNSC, CC, CeuC, EC, GaC, GB, GC, H&N, Kaposi's sarcoma, OC, PC, RC, SiC, Skin,	MHS	Population	Mean 8y (2.5) (R)	Mean Non-T2DM 61.6y (13.4) T2DM 61.6y (13.1)	52,913 (52.6%)	Non-T2DM <25, 12.4%; 25-30, 25.5%; >30, 27.8%; Missing, 34.2% T2DM <25, 16.1%; 25-30, 23.6%;	100,595	T2DM 1,639/16,721 Non-T2DM 7,945/83,874	Registry, prescriptions	Registry	Age, BMI, history of CVD, region, SES, use of healthcare services	9

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
			Respiratory, TC, UTC, Uterine						>30,15.2%; Missing,45.1%						
33	(Choi et al., 2022)	Korea	BlaC, KC	KNHI database	Population	≥5 y (P)	Mean Non-T2DM 46.0 y (13.9) T2DM 57.2y (12.0)	5,358,995 (54%)	Mean Non-T2DM 23.6 (3.2) T2DM 25.0 (3.3)	977,3462	T2DM 4,504/ 842,554 Non-T2DM 20,484/ 8,930,908	Database, laboratory test, prescriptions	Database	Age, alcohol intake, BMI, CKD, CVD, regular activity, sex, smoking	9
34	(Colmers et al., 2013)	Canada	BlaC	BCLHD	Population	10 y Median 4y (R)	Mean 60.7y (13.5)	201,188 (54.34%)	NR	370,200	T2DM 603/ 185,100 Non-T2DM 568/ 185,100	Database	Database	Age, SES, sex	8
35	(D. Y. Lee et al., 2018)	Korea	PC	NHIC	Population	Median 4.9y (3.9-6.0) (R)	Mean 57.2y (12.5)	2,796,125 (55.6%)	Non-T2DM 23.9 (3.0) T2DM NR	5,027,045	T2DM 2,916/ 1,005,409 Non-T2DM 5,673/ 4,021,636	Database, prescriptions	Registry	Age, acute pancreatitis, alcoholism, biliary disease, chronic pancreatitis, HBV, HCV, income, NAFLD, place of residence, sex	9
36	(Davis & Davis, 2022)	Australia	PC	FDS	Population	Median T2DM 15.3y (7.5-22.4) Non-T2DM 20.5y (11.4-23.5) (R)	Mean Non-T2DM NR T2DM 64.0y (11.2)	3,141 (48.70%)	Mean Non-T2DM NR T2DM/no cancer 29.6 (5.4) T2DM/cancer 28.0 (4.8)	6,449	T2DM 22/1,291 Non-T2DM 45/5,158	Database	Database	Age, CCI, postcode, sex	7
37	(De Bruijn et al., 2014)	Netherlands	BC, CC, LC, PC, PrC, RC, RsC	Rotterdam Study (I and II)	Population	Mean 11y (5.8) (P)	>55 y	4,087 (40%)	Mean Non-T2DM 26(3.8) T2DM 29(4.0)	10,181	T2DM 152/906 Non-T2DM 2,086/ 9,275	Laboratory test, prescriptions	Pathology records	Age, alcohol intake, BMI, sex, smoking status, year of inclusion in the study	9
38	(de Jong et al., 2017)	Netherlands	BTC, CC, EC, GaC, GI, PC, RC	PHARMO Database Network	Population	13y (R)	Mean 63.9y (12.6)	34,686 (51%)	NR	68,076	T2DM 583/34,038 Non-T2DM 351/34,038	Database, prescriptions	Registry	Age, use of antihypertensives, PPIs, statins, sex	7
39	(de Jong et al., 2018)	UK	BTC, CC, EC, GaC, GI, LVC, PC, RC, SiC	UK CPRD	Population	14y (R)	Median 61.8y (52-73)	366,782 (55%)	Non-T2DM <20,4.0%; 20-24,26.2%; 25-29,28.7%; 30-34,10.9%; ≥35,4.4%; Missing,25.8% T2DM <20,1.5%; 20-24,13.6%; 25-29,28.8%;	666,876	T2DM 6,165/ 333,438 Non-T2DM 4,812/ 333,438	Database, prescriptions	Database	Sex, year of birth	6

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
									30-34,22.1% ≥35,17.6%, Missing, 16.4%						
40	(de Kort et al., 2016)	Netherlands	CRC	Subcohort of the NLCS	Population	17.3y (P)	50-69 y Mean Male Non-T2DM 61.2y (4.2) T2DM 62.9y (4.4) Female Women Non-T2DM 61.3y (4.3) T2DM 62.7y (4.2)	1,890 (48.2%)	Mean Male Non-T2DM 24.9 (2.5) T2DM 25.1 (2.8) Female Non-T2DM 25.0 (3.5) T2DM 26.4 (3.9)	3,919	T2DM 92/122 Non-T2DM 2,964/3,797	Self-reported	Registry and pathology records	Age, alcohol intake, BMI, dietary habits, family history of CRC, nonoccupational physical activity, pants/skirt size, smoking	6
41	(de Kort et al., 2017)	Netherlands	CC, CRC, RC	PHARMO Database Network	Population	Mean T2DM 4.5y (3.2) Non-T2DM 7.4y (4.1) (R)	Mean Non-T2DM 51.2y (15.0) T2DM 64y (12.8)	152,775 (47%)	NR	325,054	T2DM 394/41,932 Non-T2DM 1,939/283,122	Database, prescriptions	Registry	Age, sex, statin use	8
42	(Dickerman et al., 2018)	Iceland	PrC	The Reykjavik Study	Population	Mean 23.7y (11.8) (P)	Mean 50.6 y (7.9)	N/A	NR	9,061	T2DM 24/234 Non-T2DM 1,060/8,827	Self-reported	Registry	Age, BMI, family history of PrC, education level, fasting blood glucose, height, regular checkups, smoking	8
43	(Dulskas, Patasius, Kaceniene, et al., 2020)	Lithuania	GaC	NHIF	Population	12y (R)	≥40	NR	NR	NR	T2DM 337/99,992 Non-T2DM NR	Database, prescriptions	Registry	NR	N/A
44	(Dulskas, Patasius, Linkeviciute-Ulinskiene, et al., 2020)	Lithuania	CC, CRC, RC	NHIF	Population	13y (R)	≥40	NR	NR	NR	T2DM 1,213/111,109 Non-T2DM NR	Database, prescriptions	Registry	NR	N/A
45	(Eijgenraam et al., 2013)	Netherlands	PC	NLCS	Population	16.3y (P)	55-59 y Mean Male 61.2y (4.2) Female 61.4y (4.3)	1,944 (49%)	Mean Male 24.9 (2.6) Female 25.1 (3.5)	5,000	T2DM 23/127 Non-T2DM 421/4,873	Self-reported	Registry and pathology records	Age, alcohol intake, BMI, family history of PC, SES, sex, smoking, and cigarettes smoked	7
46	(F. B. Hu et al., 1999)	USA	CC, CRC, RC	NHS	Health professionals	18y (P)	30-55 y Mean Non-T2DM 42y T2DM	N/A	Mean Non-T2DM, 24 T2DM, 28	118,403	T2DM 62/7069 Non-T2DM 830/111,003	Self-reported and medical records	Medical records, pathology reports	Age, alcohol intake, BMI, aspirin use, menopausal status, multivitamin supplement use, physical activity, parental history of	7

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
							45y							CRC, red meat quintile categories, smoking	
47	(Gapstur et al., 2012)	USA	OC	CPS-II Nutrition Cohort	Population	15y (P)	Mean Non-T2DM 62.2 y (6.45) T2DM 63.16y (6.21)	N/A	Non-T2DM <18.5,2%; 18.5-<25,52.4%; 25-<30,33.3%; ≥30,13.8% T2DM <18.5,1%; 18.5-<25,31%; 25-<30,33.8%; ≥30,32.3%	63,440	T2DM 39/3,577 Non-T2DM 485/59,863	Self-reported	Registry and medical records	Age, BMI, education level, HRT, race	8
48	(Gini et al., 2016)	Italy	BC, Blac, Bone BraNSC, BTC/GB, CC, CRC, EC, EndC, Eye, GaC, GC, Kaposi's sarcoma, KC/UTC, LaC, LC, Le, LL, LVC, M, MesC, MLe, Mouth, My, NHL, OC, PC, PrC, RC, ST, TC, Tongue, UADT	Health databases of the Friuli Venezia Giulia (FVG) region	Population	Median 3.65y (1.66-5.60) (R)	40-84 y Median 65y	17,827 (55.2%)	NR	1,050,765	T2DM 2,069/32,247 Non-T2DM NR/1,018,518	Database, hospital discharge, prescriptions	Registry	Age, sex, year of cancer diagnosis	7
49	(Goossens et al., 2015)	UK	Blac	UK CPRD	Population	Mean T2DM 5.91y Non-T2DM 5.66y (R)	≥ 18 y	335,009 (52.6%)	Non-T2DM <20,5.5%; 20-<25,31.4%; 25-<30,33.6%; ≥30,18.2%; Missing,11.3% T2DM <20,2%; 20-<25,16.5%; 25-<30,32.1%; ≥30,45.6%; Missing,3.8%	605,660	T2DM 1,251/298,345 Non-T2DM 1,071/307,315	Database, prescriptions	Database	Age, BMI, sex, smoking	9
50	(Graff et al., 2018)	USA	KC	NHS, HPFS	Health professionals	38y (P)	Mean Male Non-T2DM 61.6y (9.6) T2DM 66.6y (8.5) Female Non-T2DM 60.6y (7.2) T2DM 63.4y (6.5)	46,380 (29.4%)	Mean HPFS Non-T2DM 25.9 y (3.6) T2DM 28.5 y (5.0) NHS Non-T2DM 26.1 y (5.1) T2DM 31.4 (6.4)	157,624	T2DM 80/6,859 Non-T2DM 640/150,765	Self-reported	Pathology records	Age, alcohol intake, BMI, duration of non aspirin NSAID drug use, hypertension, parity, physical activity, smoking and pack-years	6
51	(H.-C. Lai et al., 2013)	Taiwan	GB	NHRI database	Population	8y (R)	≥20 y	212,840 (50.5%)	Obesity Non-T2DM 0.01% T2DM 0.07%	421,039	T2DM 206/214,179 Non-T2DM 147/206,860	Database, registry	Registry	Age, gallstones, hyperlipidemia, hypertension, obesity, sex	8

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
52	(H.-S. Liu et al., 2022)	Taiwan	Uterine	NHI program	Population	11y (P)	Mean Non-T2DM 45.13y (10.2) T2DM 54.55 y (9.2)	N/A	Obesity Non-T2DM 0.69% T2DM 3.09%	5,104,242	T2DM 678/ 147,772 Non-T2DM 13,720/ 4,956,470	Database	Registry	Age, HRT, hyperlipidemia, obesity, PCOS, statin	9
53	(Hadaegh et al., 2020)	Iran	Overall	TLGS	Population	Median 13.9y (9.9-14.5) (P)	Mean 47.8y (12.4)	3,935 (45.8%)	Mean With cancer 27.2 (4.8) Without cancer 27.5 (4.6)	8,599	T2DM 50/1,193 Non-T2DM 197/7,406	Self-reported, laboratory test	Pathology records	NR	7
54	(Häggström et al., 2017)	Sweden	PrC	PCBaSe 3.0	Population	Mean 5y (3.0) (P)	Mean 72y (9.0)	N/A	NR	612,846	T2DM 3,951/ 119,571 Non-T2DM 21,931/ 493,275	Register	Registry	Age, CCI, county, educational level	8
55	(Harding et al., 2015)	Australia	AML, Anal, BC, BlCa, BraNSC, CeuC, CML, CRC, EC, GaC, GB, HLY, KC, LC, LVC, M, My, NHL, OC, PC, PrC, TC, TeC, Uterine	NDSS	Population	Median 5.8y (R)	Median Non-T2DM NR T2DM 60.4y (51.1-69.7)	NR	NR	NR	T2DM 70,406/ 872,706 Non-T2DM NR	Registry	Registry	NR	N/A
56	(Hemminki et al., 2016)	Sweden	CUP	Patients' registries (nationwide)	Population	Median 4y (R)	≥30y	NR	NR	NR	T2DM 391/ 178,444 Non-T2DM NR	Registry	Registry and pathology records	Age, period of diagnosis, region, SES, sex	7
57	(Henry et al., 2013)	USA	PC	IWHS	Population	22y (P)	55-69 y Mean With cancer 62.4 y (4.1) Without cancer 61.7y (4.2)	N/A	Mean With cancer 27.2 (5.2) Without cancer 27.0 (5.1)	36,084	T2DM 25/2,205 Non-T2DM 267/33,879	Self-reported	Registry	Age, BMI, education level, smoking	8
58	(Hense et al., 2011)	Germany	BC, CRC, LC, LVC, PC, PrC	DZC cohort	Population	Median Male 3.54y (0.3-5.8) Female 3.79y (0.2-5.8) (P)	40-79 y Median Male 62.6y (40-79) Female 65.2y (40-49)	NR	NR	NR	T2DM 1,364/ 26,742 Non-T2DM NR	Database	Registry	Age	N/A
59	(Hjalgrim et al., 1997)	Denmark	BC, Oral, EC/GaC/Peritoneum, GC, Ly/	Denmark residents and National	Population	NR (P)	≥30	NR	NR	NR	T2DM 101/772 Non-T2DM	Medical records	Registry	NR	N/A

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM NR	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
			Haemato-poietic tissues, Respiratory system, Skin, UTC	Conscript Registry											
60	(Idilbi et al., 2013)	Israel	BC, CRC, EC/GaC/SiC, GC, PC	Clalit Health Care Services (Health Insurance Registry)	Population	10y 817,506 py (R)	Mean Non-T2DM 49.5y (13.4) T2DM 62.5y (12.3)	43,632 (49.6%)	Non-T2DM <25,26.4%; 25-30,34.6%; ≥30,27.5% T2DM <25,13.3%; 25-30,34%; ≥30,50.5%	87,934	T2DM 752/13,450 Non-T2DM 2,045/74,484	Registry, medical records, laboratory test, prescriptions	Registry	NR	6
61	(J. J. Walker et al., 2013)	UK	BC, BlaC, CeuC, CRC, EC, EndC, GaC, H&N, KC, LC, LVC, M, NHL, OC, PC, PrC	SCI-DC dataset	Population	7y (R)	55-79 y	NR	NR	NR	T2DM 4,285/80,838 Non-T2DM NR	Database	Registry	Age, SES	6
62	(J. Kim et al., 2020)	Korea	LC	NHIS-HEALS cohort	Population	Median 12.86y (R)	40-69 y Mean Male Non T2DM: 50.5 y (7.8) T2DM, on metformin 55 y (7.7); no metformin 52.4 y (8.0) Female Non T2DM, 51.3 y (8.0) T2DM, on metformin 57.7y (7.4); no metformin 55.4 y (8.3)	175,737 (52.3%)	Mean Male Non-T2DM 23.8 (2.8) T2DM Metformin, 24.6 (2.8); No metformin, 24.1 (3.1) Female Non-T2DM 23.7 (2.9) T2DM Metformin, 25.2 (3.2); No metformin, 24.5 (3.3)	336,168	T2DM NR/21,877 Non-T2DM NR/314,291	Prescription, laboratory test	Medical records	Age, alcohol intake, ALT, blood glucose levels, BMI, household income, hypertension history, physical activity, residential area, smoking status, systolic blood pressure, total cholesterol	9
63	(J.-Y. Wang et al., 2013)	Taiwan	CRC	LHID2005	Population	3y (R)	>18 y	91,207 (49.3%)	NR	185,005	T2DM 314/37,001 Non-T2DM 614/148,004	Database	Database	Age, autoimmune disease, CLD, CVD, nephropathy, sex, urbanization level	6
64	(Jamal et al., 2009)	USA	BTC, GB, PC	Patients' treatment file of Veterans Health Administration (VHA) hospitals	Population	10y (R)	Mean Non-T2DM 64.8 (12.6) y T2DM 65.8y (11.3)	1,087,168 (97.5%)	Obesity Non-T2DM 7.5% T2DM 15.7%,	1,115,044	T2DM 2,871/278,761 Non-T2DM 2,676/836,283	Medical records	Medical records	Age, ethnicity, obesity, pancreatitis and other pancreatic disorders, sex, smoking	6
65	(Joh et al., 2011)	USA	KC	NHS	Health professionals	32y (P)	30-55 y Mean Non-T2DM	N/A	Non-T2DM 25.5 T2DM	118,177	T2DM 60/16,819	Self-reported and medical records	Medical records, pathology reports	Age, BMI, hypertension, smoking status, pack-years of smoking, parity	7

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
							56.8 T2DM 57y		30.5		Non-T2DM 270/ 101,358				
66	(Johnson et al., 2011)	Canada	BC, CeuC, CRC, EndC, LC, OC, PrC, TC	BCLHD	Population	Mean Non-T2DM 4.4y (2.8) T2DM 4.3y (2.8)	Mean 60.7y (13.5)	201,188 (54.3%)	NR	370,000	T2DM 12,438/ 185,000 Non-T2DM 11,473/ 185,000	Database	Database	Age, index date, number of physician visits, sex, SES	8
67	(Kajüter et al., 2014)	Germany	BC, CRC, LC, LVC, NHL, PC, PrC, uterine	Statutory health insurance company (AOK NordWest) and Disease Management Program for diabetes mellitus type 2	Population	(R) Median Male 3.1y (2.1-3.9) Female 3.2y (2.2-4.6) (R)	40-79 y	NR	NR	NR	T2DM 3,664/ 67,546 Non-T2DM NR	Database	Registry	NR	N/A
68	(Kang et al., 2021)	Korea	LC	KNHIS	Population	10y (P)	40-80 y	388,760 (53.1%)	Mean Non-T2DM 24.2 (4.7) T2DM 21.9 (6.9)	732,199	T2DM NR/46,205 Non-T2DM NR/685,994	Database	Database	Age, alcohol intake, BMI, CCI, drug use (aspirin, statin), income, sex, smoking,	8
69	(Kao et al., 2013)	Taiwan	BC, BlaC, CeuC, CRC, GaC, GB, KC, LC, Ly, LVC, Oral, PC, PrC	NHIRD	Population	9y (R)	Mean 56.5y (13.2)	63,114 (55.1%)	NR	114,546	T2DM 900/22,910 Non-T2DM 2,952/ 91,636	Database	Database	Age, sex	7
70	(Kasmari et al., 2017)	USA	HCC	MarketScan Database	Population	5y (R)	19-64 y Mean 57.7y (9.08)	17,934 (60.6%)	NR	29,583	T2DM 2,541/ 7,955 Non-T2DM 5,414/ 21,628	Database	Database	Age, HCV, hyperlipidemia, medications (for diabetes and hyperlipidemia), sex	7
71	(Kasper et al., 2009)	USA	PrC	HPFS	Health professionals	T2DM 42,444 py Non-T2DM 702,580 py (P)	Mean Non-T2DM 53.7 y T2DM 60.1 y	N/A	Mean Non-T2DM 24.9 T2DM 25.7	47,781	T2DM 313/1,613 Non-T2DM 4,198/ 46,168	Self-reported and medical records	Medical records, pathology reports	Age, ancestry, BMI, family history, height, intakes of calories (bacon, tomato sauce, alpha-linolenic acid, calcium, fish, vitamin E supplementation), physical activity, smoking	7
72	(Khan et al., 2008)	Denmark, France, Germany, Greece, Italy, Netherlands, Norway, Spain, Sweden, UK	NHL	EPIC	Population	Mean 8.5y (P)	≥ 20 y	139,431 (35.4%)	Non-T2DM <25, 48.0%; 25-<30, 37.6%; ≥30, 14.3% T2DM <25, 22.3%; 25-<30, 41.2%; ≥30, 36.5%	392,591	T2DM 45/10,253 Non-T2DM 1,168/ 382,338	Self-reported	Registry and pathology records	Age, BMI, smoking	7

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
73	(Kincius et al., 2020)	Lithuania	PrC	NHIF	Population	490,187.88 py (R)	>40 y	N/A	NR	NR	T2DM 2,751/ 64,000 Non-T2DM NR	Database	Registry	No	N/A
74	(Knapen et al., 2013)	Denmark	CRC	Danish National databases	Population	Mean T2DM 5.3y Non-T2DM 6.2y (R)	Mean Non-T2DM 62.3y T2DM 62.4y	345,044 (52.7%)	NR	654,928	T2DM 2,823/ 177,281 Non-T2DM 7,116/ 477,647	Database, prescriptions	Registry	Age, drug use (antidiabetic, aspirin, statins), history of IBD and other malignancies, sex	8
75	(Koh et al., 2013)	Singapore	HCC	Singapore Chinese Health Study	Population	45-74 y Mean 14y (3.7) (P)	Mean Non-T2DM 56.0y (7.9) T2DM 60.0y (7.7)	27,349 (44.6%)	Mean Non-T2DM, 23.0 (3.2) T2DM 24.1 (3.3)	61,321	T2DM NR Non-T2DM 412/55,852	Self-reported	Registry	Age, alcohol intake, BMI, consumption of coffee and tea, dialect group, education level, sex, smoking, year of recruitment	8
76	(Larsson et al., 2005)	Sweden	CC, CRC, RC	COSM	Population	Mean 6.2y (P)	45-79 y	N/A	NR	45,550	T2DM 69/NR Non-T2DM 342/NR	Self-reported	Registry	Age, aspirin use, BMI, consumption of dairy foods / fruits / red/ meat / vegetables, education level, family history of CRC, multivitamin supplement use, physical activity, smoking	8
77	(Lawler et al., 2023)	USA	CC, CRC, RC	SCSS	Population	NR (P)	Median 51y (46-58)	19,811 (36%)	Obesity Non-T2DM, 37% T2DM, 62%	54,597	T2DM 578/25,992 Non-T2DM 394/28,605	Self-reported	Registry	Alcohol intake, BMI, CRC screening, education level, enrolment source, family history of CRC, health insurance, income, race/ethnicity, sex, smoking	8
78	(Lawrence et al., 2013)	Israel	PrC	BIP study	Population	Median 12.7y (0-15.7) (P)	36-74 y Median 59 y (7.3)	N/A	Obesity No metabolic syndrome Non-T2DM, 4% T2DM, 3% Metabolic syndrome Non-T2DM, 30% T2DM, 26%	11,516	T2DM 48/2,046 Non-T2DM 411/9,470	Laboratory test	Registry	Age	6
79	(Lega et al., 2016)	Canada	BC, BlaC, CeuC, CRC, EndC, LC, LVC, OC, PC, PrC, TC	Health care databases in Ontario	Population	Median 5.38y (R)	Mean 58.0y (13.8)	554,782 (53.7%)	NR	1,032,438	T2DM 36,140/ 516,219 Non-T2DM 33,156/ 516,219	Database	Registry	Age, income, number of physician visits before diabetes diagnosis, sex	8
80	(Liaw et al., 2015)	Taiwan	BC	TNHIRD	Population	7y	Mean Non-T2DM	N/A	NR	10,822,341	T2DM 9,105/	Database	Database	Age, insurance cost, region	7

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
						(P)	33.06y (19.05)				830,546				
							T2DM 60.16y (13.91)				Non-T2DM 48,164/ 9,991,795				
81	(Limburg et al., 2005)	USA	CC, CRC, RC	IWHS	Population	14y (P)	55-69 Mean Non-T2DM 61.5y (0.02) T2DM 62.3y (0.10)	N/A	Mean Non-T2DM 26.8 (0.03) T2DM 30.6 (0.1)	34,972	T2DM 131/1,900 Non-T2DM 1,596/ 33,072	Self-reported	Registry	Age, BMI, calcium intake and vitamin E intake, total energy intake	8
82	(Limburg et al., 2006)	USA	CRC	REP	Population	19,158 py (R)	Mean 61.0 y (14.1)	NR	<25.85: 25% 25.85-29.34:25% 29.35-33.94: 25% ≥33.95: 25%	NR	T2DM 51/1,975 Non-T2DM NR	Medical records and prescriptions		Age, calendar period, sex	N/A
83	(Lind et al., 2012)	Sweden	BC, PrC	Diab-Base system	Hospitals	Mean 6.9y (P)	Mean Male 52.4y (17.2) Female 52.2y (19.2)	NR	Mean Male 27.4 (4.7) Female 27.5 (5.8)	NR	T2DM NR/8,952 Non-T2DM NR	Database	Registry	Age, chronological time	5
84	(Lipscombe et al., 2006)	Canada	BC	Ontario health plan databases	Population	Median 4.5y (R)	55-79 y Mean Non-DM 64.9y (6.9) T2DM 66.2y (6.7)	N/A	NR	465,510	T2DM 1,011/ 73,796 Non-T2DM 5,096/ 391,714	Database	Registry, medical records, pathology reports	Age, income, physician visits	8
85	(Lo et al., 2013)	Taiwan	BC, BlaC, BraNSC, BTC, CeuC, CRC, EC, GaC, H&N, KC, LC, Le, LVC, M, My, NHL, OC, PC, PrC, Skin, TC, Uterine	NHRI database	Population	7y (R)	Mean Non-T2DM 60.4 y (13.7) T2DM 60.5y (13.6)	879,316 (49.1%)	NR	1,790,868	T2DM 58,289/ 895,434 Non-T2DM 52,093/ 895,434	Database	Registry, pathology reports	Age, hyperlipidemia, hypertension, sex, urbanization	7
86	(M.-C. Hsieh et al., 2012)	Taiwan	BC, CC, LC, LVC, PC, PrC, GaC	NHI database	Population	9y (R)	Mean Non-T2DM NR T2DM 61.44y (13.23)	T2DM 31,568 (51.1%)	NR	739,155	T2DM 6,554/ 61,777 Non-T2DM 26,754/ 677,378	Database	Database	Age, sex	7
87	(M.-S. Lai et al., 2006)	Taiwan	HCC	KCIS	Community	Mean 2.78y (0.92) (P)	>30 y	20,848 (38%)	With cancer <25, 55.6%; ≥25,44.4% With cancer <25, 57.7%; ≥25,42.3%	54,916	T2DM 18/5,732 Non-T2DM 85/49,184	Laboratory test	Registry, laboratory test, imaging techniques	Age, alcohol intake, HBV status/HCV status, sex, smoking	6
88	(M.-Y. Lee et al., 2012)	Taiwan	BC, CC, LC, LVC, PrC	NHI database	Population	11y (P)	≥30 y	488,778 (49.6%)	NR	985,815	T2DM 632/ 104,343 Non-T2DM 34,309/	Database, prescriptions	Database	Age, dyslipidaemia, gout, hypertension, sex	7

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM 881,472	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
89	(Ma et al., 2018)	USA	CC, CRC, RC	NHS, HPFS	Health professionals	32y (3,547,837 py) (P)	Mean NHS Non-T2DM 60.2y (11.3) T2DM 67.7y (9.4) HPFS Non-T2DM 63.8y (11.2) T2DM 70.7y (9.1)	47,240 (35%)	Mean NHS Non-T2DM 24.7 (4.2) T2DM 30.2 (5.7) HPFS Non-T2DM 25.7 (3.3) T2DM 29.1 (4.6)	134,763	T2DM 240/NR Non-T2DM 2,760/NR	Self-reported and medical records	Medical records, pathology reports	Age, alcohol intake, BMI, aspirin use, family history of diabetes, history of CRC, history of endoscopy/ sigmoidoscopy, HRT, physical activity, race, smoking, total calorie/energy-adjusted intake (folate, calcium, vitamin D, processed meats, beef, pork, or lamb)	7
90	(Magliano et al., 2012)	Australia	BC, CRC, PC, PrC	FDS	Population	71,015.8 py (R)	Mean Non-T2DM NR T2DM 64.1y (11.3)	NR	Non-T2DM NR T2DM/ no cancer <25,17.4%; 25-<30, 40.5%; ≥30,42.1% T2DM/cancer <25,17.5%; 25-<30, 37.9%; ≥30,44.7%	6,445	T2DM 307/1,289 Non-T2DM 1,127/5,156	Database	Database	Age, postcode, sex	7
91	(Makhoul et al., 2016)	USA	PC	VISN16	Hospitals	T2DM, 558,142.91 py Non-T2DM 1,298,683.39 py (R)	Median Non-T2DM 62y (19-94) T2DM 64y (21-96)	315,196 (97.7%)	NR	322,614	T2DM 124/ 110,919 Non-T2DM 140/ 211,695	Database	Database	Age, facility, sex	7
92	(Mariusdottir et al., 2016)	Iceland	KC	The Reykjavik Study	Population	Median 25y (1month - 38 y) (P)	Mean 53y (9.0)	9,100 (48.3%)	Mean With cancer 26 (4) No cancer 25(4)	18,840	T2DM 3/748 Non-T2DM 222/18,092	Self-reported	Registry	Age, sex	7
93	(Maskarinec et al., 2017)	USA	BC	MEC	Population	Mean T2DM 14.1y (4.4) Non-DM 14.9y (4.1) (P)	45-75 y Mean Non-T2DM 59.7 y (8.9) T2DM 61.6y (8.1)	N/A	Non-T2DM <20, 8.9%; ≥20-<25, 40.8%; 25-<30, 31.2%; ≥30,19.2% T2DM <20, 2.3%; ≥20-<25, 19.0%; 25-<30, 34.8%; ≥30,44.0%	103,721	T2DM 896/14,558 Non-T2DM 5,703/ 89,163	Self-reported, database	Registry	Age, alcohol intake, BMI, education level, ethnicity, oestrogen-related variables (age at first live birth, age at menarche, HRT, parity), family history of BC, physical activity	9
94	(Maskarinec et al., 2019)	Iceland	BC	The Reykjavik Study	Population	Mean 28.4y (11.6) (P)	Mean Non-T2DM 53.0y (8.9) T2DM	N/A	Non-T2DM <18.5, 1.7%; 18.5-<25,23.5%; 25-<30,38.4%; ≥30,36.4%	9,606	T2DM 16/294 Non-T2DM 712/9,312	Self-reported, laboratory test	Registry	Age, birth cohort, BMI, height, education level, oestrogen-related variables (age at first live birth, age	9

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
						60.4y (9.2)			T2DM <18.5, 2.4%; 18.5-<25, 53.6%; 25-<30, 33.1%; ≥30, 10.9%					at menarche, parity), place of residence, physical activity, smoking	
95	(Maskarinec et al., 2023)	USA	NHL	MEC	Population	Mean 19.2y (6.6) (P)	Mean 59.9y (8.8)	87,256 (45.3%)	NR	192,424	T2DM 1,176/68,850 Non-T2DM 2,296/123,574	Self-reported, database	Registry	Age, alcohol intake, BMI, education level, physical activity, race/ethnicity, sex, smoking	9
96	(Michels et al., 2003)	USA	BC	NHS	Health professionals	22y (P)	30-55 y Mean Non-T2DM 52.1y T2DM 59.1y	N/A	Mean Non-T2DM, 25.0 T2DM, 30.7	116,488	T2DM 202/6,220 Non-T2DM 5,403/110,268	Self-reported and medical records	Medical records, pathology reports	Age, alcohol intake, BMI, oestrogen-related variables (age at first child's birth, age at menarche, age at menopause, menopausal status, parity, HRT), family history of BC, history of benign breast disease, height, physical activity	8
97	(Newton et al., 2013)	USA	BlaC	CPS-II Nutrition Cohort	Population	Mean 11.9y (P)	Approx. ≥55 y	80,805 (46.7%)	Non T2DM <18.5, 1.4%; 18.5-<25, 44.6%; 25-<30, 39.0%; 30-<35, 10.6%; ≥35, 3.1% Missing, 1.4% T2DM <18.5, 0.7%; 18.5-<25, 27.7%; 25-<30, 41.1%; 30-<35, 19.5%; ≥35, 9.1% Missing, 1.7%	172,791	T2DM 216/12,863 Non-T2DM 1,636/159,928	Self-reported	Self-reported confirmed with medical records or cancer registry	Age, alcohol intake, BMI, education level, race, sex, smoking status	8
98	(Nicodemus et al., 2004)	USA	KC	IWHS	Population	15y (P)	55-69 y	N/A	NR	34,637	T2DM 10/NR Non-T2DM 114/NR	Self-reported	Registry	Age	5
99	(Ogunleye et al., 2009)	UK	BC, BlaC, CC, EC, EndC, GaC, KC, LC, LVC, NHL, PC, PrC, RC, Skin	Database of HIC	Population	Mean T2DM 3.88y Non-T2DM 4.04y (R)	Mean 62 y	15,314 (53.3%)	NR	28,731	T2DM 661/9,577 Non-T2DM 1,364/19,154	Database	Registry	Age, deprivation decile, GP practice, index date, sex	7
100	(Onitilo et al., 2014)	USA	BC	Marshfield Clinics	Hospitals	Mean T2DM 7.5y (5-10)	Mean Non-T2DM 64.0y (54-74) T2DM	N/A	Mean Non-T2DM 28.6 (26-30) T2DM 33.8 (29-38)	31,769	T2DM 181/5,423 Non-T2DM 671/26,346	Medical records	Registry	Age, BMI, comorbidities (CHD, COPD, CVD, MI, PVD, RHD, renal disease), insurance status,	9

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
						Non-T2DM 6.4y (3-9) (R)	63.6y (54-74)							reference date, residence, smoking	
101	(Onitilo, Berg, Engel, Glurich, et al., 2013)	USA	CC	Marshfield Clinics	Hospitals	Mean T2DM 7.4y (4.4-10.0) Non-T2DM 6.1y (2.8-8.8) (R)	Mean Non-T2DM 63.2y (53-72) T2DM 62.9y (53-73)	33,832 (51.57%)	Mean Non-T2DM 28.9 (26-39) T2DM 33.4 (28-37)	65,601	T2DM 106/11,236 Non-T2DM 341/54,365	Medical records	Registry	Age, BMI, comorbidities (CHD, COPD, CVD, MI, PVD, RHD, renal disease), diagnosis reference date, insurance status residence, sex, smoking	9
102	(Onitilo, Berg, Engel, Stankowski, et al., 2013)	USA	PrC	Marshfield Clinics	Hospitals	Mean T2DM 7.2y (4.2-9.9) Non-T2DM, 5.9y (2.6-8.5) (R)	Mean Non-T2DM 62.5y (53-72) T2DM 62.3y (53-71)	N/A	Mean Non-T2DM 29.2 (27-30) T2DM 33.1 (29-39)	33,832	T2DM 237/5,813 Non-T2DM 1,073/28,019	Medical records	Registry	Age, BMI, comorbidities (CHD, COPD, CVD, MI, PVD, RHD, renal disease), diagnosis reference date, insurance status residence, sex, smoking	9
103	(Overbeek et al., 2019)	Netherlands	CC, CRC, RC	PHARMO Database Network	Population	Mean Male T2DM 3.7 y (2.2) Non-T2DM 3.5y (2.2) Female T2DM 3.8y (2.2) Non-T2DM 3.7y (2.2) (R)	Mean Male Non-T2DM 62.0 y (10.9) T2DM 62.1y (10.9) Female Non-T2DM 64.8 y (12.2) T2DM 64.9y (12.2)	146,266 (54%)	NR	271,052	T2DM 301/55,045 Non-T2DM 886/216,007	Database, medical records, prescriptions	Registry	Age, drugs (aspirin, non-aspirin NSAIDs, statins, antihypertensives, HRT), SES	8
104	(Palmer et al., 2017)	USA	BC	BWHS	Population	870,358 py (P)	Mean Non-T2DM 47.0y (11.2) T2DM 56.1y (10.6)	N/A	Mean Non-T2DM 29.5 (6.8) T2DM 34.9 (8.4)	54,337	T2DM 163/6,694 Non-T2DM 1,688/47,643	Laboratory test	Medical records or cancer registry	Age, BMI, oestrogen-related variables (age at first birth, age at menarche, duration of oral contraceptive/ HRT, parity), family history of BC	8
105	(Park et al., 2019)	Korea	LC	KNHIS	Population	9y (P)	>40 y	103,288 (34.4%)	Non-T2DM <18, 1.3%; 18-<23, 37.0%; 23-<25, 27.3%; ≥25, 34.4% T2DM <18, 0.7%; 18-<23, 23.7%; 23-<25, 26.0%; ≥25, 49.6%	300,518	T2DM 66/7,499 Non-T2DM 1,906/293,019	Database, prescriptions	Hospital discharge records	Age, alcohol intake, BMI, CCI, income, physical activity, sex, systolic blood pressure, total serum cholesterol	8

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
106	(Pasqual et al., 2023)	USA	TC	Sister Study	Population	Median 12.5y (11.3-14.1) (P)	Median 55.4y (48.9-62.1)	N/A	NR	47,739	T2DM 17/NR Non-T2DM 220/NR	Self-reported	Self-reported and confirmed with medical records/ pathology report	Age, benign thyroid disease, BMI, education level, healthcare utilization, race/ethnicity, smoking	8
107	(Peeters et al., 2015)	UK	CC, CRC, RC	CPRD	Population	Mean T2DM 5.6y (4.5) Non-T2DM 6.0y (4.7) (R)	Median 61y (50-71)	316,618 (52.8%)	Mean Non-T2DM 25.5 (23.3-29.1) T2DM 30.0 (26.4-34.5)	600,078	T2DM 2,759/300,039 Non-T2DM 2,359/300,039	Medical records, prescriptions	Medical records	Age, alcohol intake, sex, smoking, statin	8
108	(Prizment et al., 2013)	USA	BlaC	IWHS	Population	25y (P) T2DM 62.6y (4.2)	55-69 y Mean Non-T2DM 61.7y (4.2) T2DM 62.6y (4.2)	N/A	Mean Non-T2DM 26.8(4.9) T2DM 30.5 (6.3)	37,327	T2DM 45/2,274 Non-T2DM 232/35,053	Self-reported	Registry	Age, alcohol intake, BMI, education level, occupation, marital status, physical activity, smoking, WHR	8
109	(Redaniel et al., 2012)	UK	BC	UK GPRD	Population	T2DM 594,694.9 py Non-T2DM 255,747.5 py (R)	≥35 y	N/A	Non-T2DM <18.5, 2.61%; 18.5-<25, 43.52%; 25-<30, 34.39%; ≥30, 19.48% T2DM <18.5, 0.72%; 18.5-<25, 18.23%; 25-<30, 34.46%; ≥30, 47.59%	82,867	T2DM 873/52,657 Non-T2DM 714/30,210	Medical records	Medical records	Age, BMI, period, region	7
110	(Reinholdt et al., 2021)	Denmark	Anal, Penis	Nationwide health registries and clinical databases	Population	39,652,441 py (P)	>15 y	N/A	NR	2,500,566	NR	Database, registry	Database, registry	Age, calendar period, educational level	6
111	(Rodriguez et al., 2005)	USA	PrC	CPS-II Nutrition Cohort	Population	T2DM 48,369 py Non-T2DM 487,193 py (P)	Approx. ≥60 y	N/A	Non-T2DM <25, 37.4%; 25-<30, 49.0%; ≥30, 12.3% T2DM <25, 24.1%; 25-<30, 47.7%; ≥30, 26.6%	72,670	T2DM 343/10,053 Non-T2DM 4,975/62,617	Self-reported	Self-reported confirmed with medical records or cancer registry	Age, BMI, education level, family history of PrC, energy-adjusted intakes (lycopene, calcium), prostate-specific antigen, race, total fat	8
112	(S. He et al., 2022)	China	Overall	Da Qing IGT and Diabetes Study	Population	Median 27y (15-50) (P) T2DM 48.4y (8.8)	Mean Non-T2DM 44y (8.9) T2DM 47.7y (13.2)	Approx 50%	Obesity Non-T2DM, 31.5% T2DM 54.4%	1,125	T2DM 89/611 Non-T2DM 77/514	Laboratory test	Database, medical records	Age, BMI, hypertension, smoking	9
113	(S. K. Kim et al., 2020)	Korea	BC, BlaC, CeuC, CRC, EC, EndC, GaC, KC, LaC, LC, Le, LVC, OC, Oral/PhC, PC, PrC, TC	KNHIS	Population	Mean 8.6y (1.5) (R) T2DM 58.0y (12.6)	Non-T2DM 47.3y (13.1) T2DM 58.0y (12.6)	12,540,261 (48.78%)	NR	25,709,497	T2DM NR/697,935 Non-T2DM NR/25,011,562	Database	Database	Age, CHD, CKD, CLD, hyperlipidemia, hypertension, income, place, sex	8

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
114	(S.-H. Hsieh et al., 2014)	Taiwan	BC, CC, Cervical, EC, GaC, hematologic, LC, LVC, OC, PC, PhC, PrC, Skin, TC, UTC, Uterine	Chang Gung Memorial Hospital	hospital	NR (R)	≥20y	23,666 (56%)	18.5-~25, Non-T2DM 20.6 (2.4) T2DM 20.9 (2.3) 29-~30 Non-T2DM 25.3 (0.9) T2DM 25.4 (0.9) ≥30 Non-T2DM 29.7 (2.8) T2DM 30.0 (2.9)	42,229	T2DM 4,619/ 7,876 Non-T2DM 19,873/ 34,353	Medical records	Imaging techniques, pathology reports	NR	4
115	(Saarela et al., 2019)	Finland	BC, BTC, BlaC, BraNSC, CC, CeuC, EC, EndC, Eye, GaC, HLy, KC, LaC, LC, Le, Lip, LVC, Ly, MesC, NH, NM, My, OC, PC, PhC, PrC, RC/Anal, SaG, SiC, Skin/M, TC, TeC, Tongue, UTC	Finnish National Diabetes Register	Population	10.5y (R)	>30y	NR	NR	NR	T2DM 74,063/ 428,326 Non-T2DM NR	Registry, hospital discharge, prescriptions	Registry	NR	N/A
116	(Samson et al., 2016)	USA	BC	South Carolina Medicaid	Population	6y (R)	NR	N/A	NR	7,219	T2DM NR/1,293 Non-T2DM NR/5,926	Database	Registry	Diabetes medication, menopausal age	6
117	(Schlesinger et al., 2013)	Denmark, France, Germany, Greece, Italy, Netherlands, Norway, Spain, Sweden, UK	BTC, HCC	EPIC	Population	Mean 8.5y (P)	25-70 y	128,273 (35.3%)	Mean Non-T2DM 25.8 (4.3) T2DM, 29.2 (4.9)	363,426	T2DM 21/8,588 Non-T2DM 343/ 354,838	Self-reported	Registry and pathology records	Age, alcohol intake, BMI, education level, smoking, WHR	7
118	(Sellers et al., 1994)	USA	BC	IWHS	Population	5y (167,644 py) (P)	55-69 y	N/A	NR	36,603	T2DM 31/NR Non-T2DM 553/NR	Self-reported	Registry	Age	5
119	(Sikdar et al., 2013)	Canada	CRC	CCDRD	Population	10.5y (751,562 py) (R)	>30y Mean Males Non-T2DM 55.9y (12.8) T2DM 56.3y (12.8); Females Non-T2DM	61,156 (50.03%)	NR	122,228	T2DM 414/25,304 Non-T2DM 1,249/ 96,924	Database	Database	Age, CCI	7

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
							58.3y (14.3) T2DM 58.7y (14.3)								
120	(Simon et al., 2018)	USA	HCC, ICC	NHS, HPFS	Health professionals	32y (4,488,410 py) (P)	Mean Non-T2DM 62.5y (7.9) T2DM 65.7y (7.5)	44,710 (29%)	Mean Non-T2DM, 26.2 (4.7) T2DM 29.6 (5.9)	153,895	T2DM 37/10,110 Non-T2DM 75/143,785	Self-reported	Medical records	Age, alcohol intake, BMI, drug use (aspirin, hypoglycaemic medications), family history of diabetes, physical activity, race, sex, smoking	6
121	(Sponholtz et al., 2016)	USA	EndC	BWHS	Population	665,749 py (P)	21-69 y	N/A	Mean <25, 22.3(21.5); 25-<30, 27.3 (18.2); 30-<35, 32.3 (20.0); 35-40, 37.3 (21.5); ≥40, 45.1 (75.5)	47,557	T2DM 57/1,452 Non-T2DM 218/44,441	Self-reported	Medical records	Age, BMI, oestrogen-related variables (age at menarche, menopausal status, oral contraceptive use/HRT, parity), physical activity, smoking, use of drugs (statin, metformin)	7
122	(T. Bazelier et al., 2013)	Denmark	BlaC	Danish Medicines Agency (National Pharmacological Database)	Population	Mean T2DM 5.3y NonT2D 6.2y (R)	Mean Non-T2DM 62.4y T2DM 62.5y	348,691 (52.48%)	NR	664,375	T2DM 838/179,056 Non-T2DM 2,228/485,319	Database, prescriptions	Registry	Age, chronic UTI, COPD medication, history of malignancies (other than bladder), sex, smoking cessation drugs	8
123	(Tsilidis et al., 2015)	Denmark, France, Germany, Greece, Italy, Netherlands, Norway, Spain, Sweden, and the UK	PrC	EPIC	Population	Median 12y (0-18) (P)	35-70 y	N/A	Mean Non-DM 26.6 (26.5-26.6) T2DM 28.1 (28.0-28.2)	139,131	T2DM NR/4,761 Non-T2DM 4,374/134,370	Self-reported, laboratory test, registry, prescriptions	Registry and pathology records	Age, BMI, education level, physical activity, smoking status, waist circumference	8
124	(Ulcickas Yood et al., 2009)	USA	BlaC, CRC, LVC, M, PC	Privately insured Population database	Population	1,544,613 py (R)	Mean Non-T2DM, 53.1y (53.1-53.2) T2DM 56.2y (56.1-56.3)	229,615 (51.86%)	NR	442,712	T2DM 813/191,223 Non-T2DM 352/251,489	Database	Pathology records	Age, selected cancer risk factors (CRC: Crohn's disease, history of polyps, ulcerative colitis; BlaC: pelvic radiation, schistosomiasis; LVC: alcoholism, cirrhosis, HBV/HCC; PC: alcoholism, chronic pancreatitis, deep venous thrombosis, dermatomyositis/poly myositis, HBV/HCC, history of polyps, partial gastrectomy), sex	7
125	(Undzyte et al., 2020)	Lithuania	KC	NHIF	Population	17y (R)	NR	NR	NR	NR	T2DM 588/115,923 Non-T2DM NR	Database	Registry	NR	N/A

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
126	(Valent, 2015)	Italy	CC, EC, GaC, GB, LVC, PC, RC, SiC	Regional health information system	Population	11,921,662 py (R)	20-94 y	NR	NR	1,189,515	NR/109,255 Non-T2DM 19,733/1,080,260	Registry	Hospital discharge records	Age, sex	6
127	(Vatseba & Sokolova, 2019)	Ukraine	BC, CC, GaC, LC, PC, PrC, Skin, uterine	Precarpathian clinical oncology center and in the Ivano-Frankivsk regional clinical hospital	Hospitals	5y (P)	NR	NR	NR	700,654	272/42,532 Non-T2DM 2,413/658,122	Hospital records	Hospital records	NR	1
128	(W. S. Yang et al., 2013)	China	LVC	SMHS, SWHS	Population	Mean SWHS 11.2y SMHS 5.5y (P)	Mean Male Non-T2DM 54.9y (9.6) T2DM 60.5y (9.7) Female Non-T2DM 51.9y (8.9) T2DM 58.3y (8.3)	60,183 (45.2%)	Mean Male Non-T2DM 23.6 (3.1) T2DM 24.6 (3.0) Female Non-T2DM 23.8 (3.3) T2DM 26.0 (3.8)	133,288	T2DM 51/10,916 Non-T2DM 293/122,372	Self-reported	Registry and medical records	Age, alcohol intake, birth cohort, BMI, education level, family history of cancer, history of chronic hepatitis or CLD, HRT, income, menopausal status, occupation, physical activity, smoking, total energy intake/fruit intake/vegetable intake	8
129	(W.-S. Yang et al., 2014)	China	LC	SMHS, SWHS	Population	Median SWHS 12.2y SMHS 6.3y (P)	Mean Male Non-T2DM 54.89y (9.63) T2DM 60.48y (9.61) Female Non-T2DM 51.94y (8.91) T2DM 58.51y (8.34)	59,910 (45%)	Mean Male Non-T2DM 23.64 (3.07) T2DM 24.61 (3.04) Female Non-T2DM 23.82 (3.33) T2DM 26.00 (3.76)	133,024	T2DM 98/10,890 Non-T2DM 919/122,134	Self-reported	Registry and medical records	Age, alcohol intake, birth cohort, BMI, education level, family history of LC, HRT, income, menopausal status, occupation, physical activity, smoking, total energy intake/fruit intake/vegetable intake	8
130	(W.-S. Yang et al., 2016)	China	NHL	SMHS, SWHS	Population	Mean SWHS 12.9y SMHS 7.4y (P)	Mean Male Non-T2DM 54.9y (9.6) T2DM 60.5y (9.5) Female Non-T2DM 51.9y (8.9) T2DM 58.5y (8.3)	59,971 (45%)	Mean Male Non-T2DM 23.6 (3.1) T2DM 24.6 (3.0) Female Non-T2DM 23.8 (3.3) T2DM 26.1 (3.8)	133,097	T2DM 32/10,902 Non-T2DM 140/122,195	Self-reported	Registry and medical records	Age, alcohol intake, birth cohorts, BMI, education level, oestrogen-related variables (HRT, menopausal status), family history of cancer, income, occupation, physical activity, smoking, total energy intake/fruit intake/vegetable intake	8
131	(W.-S. Yang et al., 2017)	Taiwan	GI, LC, UTC	Chin-Shan township habitants	Community	Median 10.5y (P)	>35 y Mean 54.9y (12.3)	1,703 (47.3%)	Mean Non-T2DM 23.3 (3.35) T2DM	3,561	T2DM 232/473 Non-T2DM 43/3,088	Self-reported, laboratory test	Self-reported confirmed with medical records	Age, alcohol intake, BMI, exercise, sex, smoking, total cholesterol/creatinine/	8

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
									24.7 (3.75)					lipoprotein/ triglycerides	
132	(Wallner et al., 2011)	USA	PrC	OCS	Population	15y (P)	40-79 y	N/A	NR	2,445	T2DM 9/NR Non-T2DM 197/NR	Self-reported	Pathology records	Age	5
133	(X. He et al., 2018)	China	BC, Blac, BraNSC, Cervix, CRC, EC, GaC, GB, KC, H&N, LC, Le/My, LVC, NHL, OC, PC, PrC, Skin, ST, TC, Uterine	UEBMI database	Population	5y (R)	Mean Non-T2DM 43.2 y (15.1) T2DM 58.8y (12.0)	414,274 (52.19%)	Obesity Non-T2DM 0.001% T2DM 0.09%	793,795	NR	Database	Database	Age, comorbidities (alcohol-related disease, hypertension, CHD, COPD, dyslipidaemia, genital-urinary, tract-diseased, nephropathy), concomitant medications, obesity, sex	8
134	(X. Liu et al., 2015)	Sweden	BC, Blac, BraNSC, CeuC, CC, EC, EnC, EndC, Eye, GaC, HLY, KC, LC, Le, LVC, M, My, NHL, OC, Other genital (GC), PC, PrC, RC, SiC, Skin, TC, Upper aerodigestive tract, Skin	Nationwide database of hospitals and primary health care	Population	Median 4y (2-8) (R)	>39 y	NR	NR	NR	T2DM 32,084/380,196 Non-T2DM NR	Registry	Registry	Age, region, SES, sex	6
135	(X. F. Pan et al., 2018)	China	BC, CRC, EC, LC, LVC, PC, GaC	CKB Study	Population	3,612,769 py Mean 7.1y (P)	Mean 51.5y (10.7)	208,832 (41%)	No T2DM <18.5, 4.5%; 18.5-<24,52.8%; 24-<28,32.7%; ≥28,10% T2DM <18.5, 2.4%; 18.5-<24,36.3%; 24-<28, 41.7%; ≥28,19.6%	508,892	T2DM 1,457/29,835 Non-T2DM 16,006/479,057	Self-reported and laboratory test	Registry	Age, alcohol intake, BMI, education level, menopausal status (for BC), parental history of cancer, physical activity, sex, smoking, study area	9
136	(Y. Hu et al., 2021)	USA	BC, Blac, CRC, KC, EC, EndC, GB, LC, Le, LVC, M, My, NHL, OC, PC, PrC, TC	NHS, HPFS	Health professionals	4,294,078 py (P)	Mean NHS Non-T2DM 58.7y (11.9) T2DM 68.0 y (9.6) HPFS Non-T2DM 64.5y (11.0) T2DM 71.3y (9.4)	45,604 (28.6%)	Mean NHS Non-T2DM 25.6 (4.9) T2DM 30.6 (6.5) HPFS Non-T2DM 25.9 (3.6) T2DM 28.7 (5.0)	159,033	T2DM 3,947/NR Non-T2DM 39,902/NR	Self-reported	Medical records	Age, alcohol intake, alternative healthy eating index, endoscopy screening, ethnicity, family history of cancer, family history of diabetes, fasting glucose screening, multivitamin use, physical activity, smoking, total energy intake. For women: BMI, HRT, insulin use, mammography screening, oral	6

No.	Author, year	Country	Type of cancer	Data source	Study setting	Follow-up, y (Design)	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	Cancer in T2DM/ Non-T2DM	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	NOS
137	(Y. Pan et al., 2020)	Taiwan	BlaC	TNHIRD	Population	16y (R)	20-100 y	54,453 (56.84%)	NR	95,796	T2DM 67/31,932 Non-T2DM 152/63,864	Database	Database	contraceptive, oral hypoglycaemic Age, comorbidities (cataract, CHD, CHF, CVD, coagulation defects, COPD, dementia, glaucoma, hemiplegia or paraplegia, hyperlipidemia, hypertension, kidney disease, liver disease, peptic ulcer disease, PVD, RHD), income, length of hospital stay, sex, urbanization	7
138	(Ye et al., 2023)	UK	HCC	UK Biobank cohort	Population	Median 12y (11.4-12.7) (P)	Mean 56.7y (8.0)	204,556 (45.5%)	Mean Non-T2DM 27.04 (4.50) T2DM 31.75 (5.89)	436,202	T2DM 99/22,934 Non-T2DM 157/ 413,268	Laboratory test, medical records, prescriptions, self-report	Medical records	Age, alcohol intake, BMI, center, dyslipidaemia, education level, hypertension, income, physical activity, PRS, sex, smoking, Townsend deprivation index	9
139	(Yeo et al., 2021)	Korea	PrC	KNHI database	Population	NR (R)	>40 y	N/A	<18.5, 2%; 18.5-<23, 32%; 23-<25, 28.2%; 25-<30, 35%; ≥30, 2.8%	1,339,820	T2DM 2,655/ 184,769 Non-T2DM 11,219/ 1,155,051	Database, laboratory test, prescriptions	Database	Age, alcohol intake, BMI, hypertension, physical activity, smoking	8
140	(Z. Yang et al., 2023)	China	BC, GI, LC	Yinzhou Health Information System database	Population	Median 9.2y (7.87-10.49) (P)	Mean 57.89y (12.29)	107,364 (52%)	Mean Non-T2DM 23.08 (2.75) T2DM 24.08 (17.62)	208,289	T2DM 3,359/ 42,279 Non-T2DM 12,370/ 166,010	Laboratory test, prescriptions	Medical records	Age, alcohol intake, BMI, educational level, marital status, physical activity, sex, smoking	9
141	(Zabulienė et al., 2021)	Lithuania	EndC	NHIF database	Population	Mean 9.6y (R)	>40 y	N/A	NR	NR	T2DM 995/77,708 Non-T2DM NR	Database	Registry	NR	N/A

Follow-up: values are presented as total years, patient-years, range, mean (SD), median (IQR), depending on what was reported in the publication. Design: prospective (P) or retrospective (R). Age: values are presented as range, mean (SD), median (IQR), depending on what was reported in the publication. BMI: Values are presented as percentages, mean (SD), median (IQR), categories (overweight [BMI > 25-29.9 kg/m²], obesity [BMI > 30 kg/m²], morbid obesity [BMI > 40 kg/ kg/m²]), depending on what was reported in the publication. Sample size: total number of subjects (T2DM and non-T2DM). Cancer in T2DM (No. cancer cases/No. T2DM) /Non-T2DM (No. cancer cases/No. non-T2DM). NOS: range, 1-9 points; N/A because the comparison group was the general population (no distinction between T2DM and non T2DM reported). Adjustment variables: the maximum adjustment reported, but for quantitative analyses, the effect estimates based on the selection of the different adjustment variables were extracted separately. BMI: body mass index; IQR: interquartile range; N/A: not applicable; NOS: New-castle Ottawa Scale; NR: not reported; py: person-years; SD: standard deviation; y: years. Abbreviations for *cancer types*, *data sources* and *adjustment variables* are shown in the "Abbreviations" section. "Overall" stands for Non-specific neoplasms.

Table A 7-5 Main characteristics of the cross-sectional included in the systematic review

No.	Author, year	Country	Type of cancer	Data source	Study setting	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	T2DM in Cases	T2DM in Controls	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	Effect size (OR, 95% CI)	NOS
1	(Baur et al., 2011)	Germany	Overall	DETECT study	Population	Mean Non-T2DM/ no cancer 55.5y (14.4) Non-T2DM/ no cancer 65.5y (11.6) T2DM/ no cancer 66.4 y (10) T2DM/cancer 70.4 y (8.0)	NR	Mean Non-T2DM/ no cancer 26.6 (4.7) Non-T2DM/ no cancer 26.7(4.2) T2DM/ no cancer 29.8 (5) T2DM/cancer 28.3(4.2)	7,519	66/1,308	185/6,211	Laboratory test, medical records, prescriptions	Medical records	Age, BMI, HbA1c, sex, smoking	1.64(1.12-2.41)	10
2	(Elangovan et al., 2021)	USA	CRC	Explorys, IBM Watson Health database	Population	≥20- 74 y	11,012,010 (43%)	Obesity 31.5%	37,483,140	17,490/ 2,330,460	144,660/ 35,152,680	Database	Database	NR	Male 20-39 years, 3.42(2.85-5.37) 40-49 years, 2.00(1.75-2.28) 50-74 years, 1.35(1.30-1.39) Female 20-39 years, 0.95(0.84-1.07) 40-49 years 0.76(0.68-1.05) 50-74 years, 1.97(1.92-2.02)	6
3	(Krasnodebski et al., 2010)	Poland	CRC	Medical University of Warsaw	Hospital	Non-T2DM 61.8 y (15.5) T2DM 66.4 y (12.2)	105 (37.2%)	Non-T2DM 25.17 (4.46) T2DM 27.89 (5.30)	282	6/43	11/239	Self-report	Pathology reports	NR	3.0 (1.9, 9.3)	4
4	(M. Chen et al., 2018)	China	PrC	Shanghai Tenth People's Hospital	Hospital	50-90 y Mean 71.66 y (8.74)	N/A	Mean T2DM 23.75 (3.07) T2DM 24.34 (2.95)	2,032	NR	499/1,565	Laboratory test	Pathology reports	Age, BMI, digital rectal examination, prostate volume, PSA	Diagnosis of T2DM 0-5y, 1.56 (0.99-2.46) 6-10y 1.53 (0.94-2.49) 2.91 (1.45, 5.83)	8
5	(Sandhu, 2001)	UK	CRC	EPIC-Norfolk study	Population	45-74 y	13,193 (45%)	NR	29,343	9/561	138/ 28,782	Self-report	Registry	Age, alcohol intake, BMI, education level, family history of CRC, sex, smoking	Male 3.11 (1.32, 7.33) Female, 2.53 (0.77, 8.33) BlaC 1.81(1.75-1.86)	8
6	(Suarez Arbelaez et al., 2023)	USA	BlaC, KC, PrC, UTC	TriNetX Research Network database	Population	>37 -65 y	Approx. 16,610,321 (45%)	NR	36,911,824	54,912/ 4,353,929	145,185/ 32,557,895	Database	Database	Age, hyperlipidaemia, hypertension	KC 2.76(2.68-2.84) PrC 0.83 (0.82-0.84)	6

No.	Author, year	Country	Type of cancer	Data source	Study setting	Age, y	Sex (Male) N (%)	BMI kg/m ²	Sample size	T2DM in Cases	T2DM in Controls	T2DM ascertainment method	Cancer ascertainment method	Adjustment variables	Effect size (OR, 95% CI)	NOS
															UTC 2.32 (2.07-2.60)	
7	(Sun et al., 2023)	USA	BC	NHANES	Population	>45 y Non-T2DM, 38% T2DM, 81%	N/A	Mean Non-T2DM 27.7 (0.150) T2DM 34 (0.29)	6,144	82/ 1,452	64/4,692	Laboratory test	Database	Age, BMI, clinical parameters (cholesterol, oestradiol, glycohemoglobin, serum cotinine, serum creatinine, serum glucose, triglycerides), educational level, marital status, race, reproductive health	1.51(1.00-2.28)	9

Age: values are presented as range, mean (SD), depending on what was reported in the publication. BMI: Values are presented as mean (SD), category (obesity [BMI > 30 kg/m²]), depending on what was reported in the publication. Sample size: total number of subjects. Cancer in T2DM (No. cancer cases/No. T2DM). Cancer in non-T2DM (No. cancer cases/No. non-T2DM). NOS: range, 1-10 points. Adjustment variables: the maximum adjustment reported. BMI: body mass index; CI: confidence interval; N/A: not applicable; NOS: New-castle Ottawa Scale; NR: not reported; OR: odds ratio; SD: standard deviation; y: years. Abbreviations for *cancer types*, *data sources* and *adjustment variables* are shown in the "Abbreviations" section. "Overall" stands for Non-specific neoplasms.

Table A 7-6 Main characteristics of the mendelian randomization studies included in the systematic review

No.	Author, year	Country	Type of cancer	Data source	Sample size (total)	Effect size
1	(Au Yeung & Schooling, 2019)	Europe	BC, PrC	DIAbetes Genetics Replication And Meta-analysis consortium; European ancestry from BCAC; PRACTICAL	898,130	BC IVW: 0.98 (0.95 - 1.01) MR-Egger regression: 1.01 (0.95 - 1.08) Weighted median: 1.02 (0.99 - 1.06) PrC IVW: 1.02 (0.97 - 1.07) MR-Egger regression: 0.98 (0.86 - 1.10) Weighted median: 0.97 (0.93 - 1.02)
2	(Carreras-Torres et al., 2017)	Canada, China, Finland, France, USA	PC	PanScan, PanC4	14,374	1.03(0.95-1.11)
3	(Disney-Hogg et al., 2018)	France, Germany, UK, USA	BraNSC	Groupe Hospitalier Pitié-Salpêtrière Paris, KORA; POPGEN, and the Heinz Nixdorf Recall study, GLIOGENE Consortium, The University of Texas M.D. Anderson Cancer Center, GiomaScan (NIH), UCSF cases, and Mayo Clinic Biobank, UCSF (SFAGS), the INTERPHONE study	30,657	1.04 (0.97, 1.11)
4	(Fussey et al., 2020)	UK	TC	UK Biobank	310,601	IVW: 0.99 (0.87-1.13) Egger: 0.92 (0.71-1.81) WM: 0.94 (0.76-1.17) PWM: 0.94 (0.76-1.71)
5	(Goto et al., 2020)	Japan	BC, CRC, GaC, LC, LVC, PC, PrC	JPHC	10,536	Overall: HR 1.20(1.04-1.39); IVW 1.03(0.92-1.15); Egger 0.93(0.77-1.13) GaC: HR 1.40(1.07-1.83); IVW 0.98(0.78-1.23); Egger 0.81(0.55-1.21) CRC: HR 1.05(0.77-1.44); IVW 0.90(0.74-1.10); Egger 0.93(0.66-1.33) LVC: HR 1.40(0.84-2.32); IVW 0.80(0.57-1.14); Egger 0.79(0.42-1.47) PC: HR 1.43(0.77-2.64); IVW 1.08(0.73-1.59); Egger 0.92(0.46-1.82) LC: HR 1.40(1.01-1.95); IVW 1.00(0.80-1.25); Egger 1.02(0.68-1.52) BC: HR 0.73(0.32-1.66); IVW 1.09(0.82-1.44); Egger 1.17(0.69-1.97) PrC: HR 0.95(0.62-1.46); IVW 1.13(0.91-1.40); Egger 1.28(0.90-1.82)
6	(Hong et al., 2021)	China	LC	NJMU GSA Project; the NJMU GWAS; the NJMU OncoArray GWAS	26,655	0.96 (0.91, 1.01)
7	(Huang et al., 2022)	Finland	TC	FinnGen Study	218,792	IVW: 1.13 (0.99, 1.30) Weighted median: 1.18 (0.92, 1.51) MR-Egger: 1.10 (0.80, 1.52)
8	(J. Chen et al., 2023)	Canada, China, Finland, France, UK, USA	CRC, GaC, LCV	FinnGen consortium, UK Biobank, Celiac UK, Inflammatory Bowel Disease Genetics Consortium, Pancreatic Cancer Cohort Consortium (PanScan)	1,139,296	CRC: 1.00 (0.95, 1.05) GaC: 1.02 (0.92, 1.13) LVC: 1.12 (0.97, 1.30)
9	(Johansson et al., 2019)	UK, USA	KC	IARC, the United States NCI, MDA, and United Kingdom cancer research centers	31,190	0.99 (0.93-1.05)
10	(Lu et al., 2020)	Canada, China, Finland, France, USA	PC	PanScan, PanC4	15,824	IVW:1.09 (1.01 to 1.17) Egger: 0.95 (0.82 to 1.12) WM: 1.02 (0.90 to 1.15)
11	(Molina-Montes et al., 2020)	Germany, Ireland, Italy, Spain, Sweeden, UK	PC	PanGenEU (European Study into Digestive Illnesses and Genetics) study	3,558	2.5 (2.05-3.05)
12	(Murphy et al., 2022)	UK	CC, CRC, RC	UK Biobank	898,130	CRC IVW: 1.04 (1.01, 1.07) Weighted median: 1.00 (0.96, 1.04) MR-Egger: 0.97 (0.90, 1.04) MR-PRESSO: 1.04 (1.01, 1.07) CC IVW: 1.03 (1.00, 1.07) Weighted median: 0.98 (0.94, 1.03) MR-Egger: 0.97 (0.90, 1.05) MR-PRESSO: 1.04 (1.01, 1.08) RC IVW: 1.04 (1.00, 1.08) Weighted median: 1.00 (0.93, 1.07)

No.	Author, year	Country	Type of cancer	Data source	Sample size (total)	Effect size
						MR-Egger: 0.97 (0.88, 1.07) MR-PRESSO: 1.04 (1.00, 1.08) IVW: 0.91 (0.79 to 1.04)
13	(Nead et al., 2015)	Australia, UK	EndC	ANECs, SEARCH	9,560	IVW (p value, 0.98) MR-Egger (p value, 0.85) Weighted Median (p value, 0.40)
14	(Xiao et al., 2023)	USA	BC	PMBB and eMERGE network data sets	71,008	IVW: 1.00 (0.97-1.02) MR-Egger regression: 0.94 (0.81-1.10) MR-Egger intercept: 1.01 (0.99-1.03) Weighted median: 1.01 (0.97-1.04) Weighted mode: 1.02 (0.96-1.08)
15	(Yarmolinsky et al., 2019)	Australia, Belarus, Belgium, Canada, Denmark, Finland, Germany, Netherlands, Norway, Poland, Spain, Sweden, UK, USA	OC	OCAC	66,450	Overall cancer IVW: 1.00 (0.98-1.01) MR-Egger: 1.01 (0.99-1.04) Weighted median: 1.01 (0.98-1.03) MR-PRESSO: 1.00 (0.98-1.01)
16	(Yuan et al., 2020)	Australia, Belarus, Belgium, Canada, China, Denmark, Finland, France, Germany, Greece, Israel, Italy, Japan, Korea, Malaysia, Netherlands, Norway, Poland, Russia, Singapore, Spain, Sweden, UK, USA, Taiwan, Thailand	BC, Blac, BraNSC, BTC, CeuC, CRC, EC, H&N, KC, LC, Le, LVC, M, MM, NHL, OC, PC, PrC, GaC, TC, TeC, Uterine	DIAGRAM consortium; MAGIC; PanScan, PanC4; BCAC; UK Biobank	N/A	IVW LVC: 1.16(0.99-1.36) PC:1.13 (1.04-1.22) KC:1.08(1.00-1.17) TC:1.08(0.94-1.24) Uterine:1.08(1.01-1.15) CeuC: 1.07 (1.01-1.15) BTC: 1.07 (0.93-1.23) OC: 1.05 (0.97-1.12) Le: 1.05 (0.97-1.12) H&N: 1.03 (0.96-1.11) Blac: 1.02 (0.97-1.08) LC: 1.02 (0.96-1.08) GaC: 1.01 (0.91-1.13) NHL: 1.01 (0.95-1.07) CRC: 1.01 (0.97-1.05) BC (ER+):1.00(0.97-1.03) BC (ER-):0.99(0.96-1.01) TeC: 0.98(0.87-1.09) PrC:0.97(0.93-1.01) MM:0.96(0.86-1.07) M:0.93(0.89-0.97) BraNSC:0.92(0.84-1.02) EC: 0.89(0.81-0.98)

Sample size: total number of subjects. Effect size. Values are presented as OR (95%, CI) if nothing else is specified. CI: confidence interval; IVW: inverse variance weighted; MR-PRESSO: Mendelian Randomization Pleiotropy RESidual Sum and Outlier; N/A: not applicable; OR: odds ratio. Abbreviations for *cancer types* and *data sources* are shown in the "Abbreviations" section. "Overall" stands for Non-specific neoplasms.

A8. Number of Studies by Cancer Type and Study Design

Table A8-1 Number of studies by cancer type and study design

Type of cancer	Case-control		Cohort		Cross-sectional		MRS	
	No. Studies	%	No. Studies	%	No. Studies	%	No. Studies	%
AML			1	1%				
Anal			3	2%				
Breast	16	19%	46	33%	1	14%	4	25%
Bladder	1	1%	26	18%	1	14%	1	6%
BraNSC	1	1%	1	7%			2	13%
BTC	2	2%	7	5%			1	6%
Bone			4	3%				
Colon	5	6%	28	2			1	6%
Cervix uteri	1	1%	15	11%			1	6%
CLL/SLL	2	2%						
CML			1	1%				
CUP			1	1%				
CRC	12	14%	37	26%	3	43%	4	25%
Esophagus	4	5%	21	15%			1	6%
EJA	1	1%						
ECC	3	3%						
Endometrial	9	1	24	17%			2	13%
Endocrine	1	1%	2	1%				
Eye			3	2%				
Gastric	1	1%	23	16%			2	13%
Gallbladder	1	1%	12	9%				
Genital			5	4%			1	6%
GI			6	4%				
HCC	9	1	8	6%				
H&N	1	1%	4	3%			1	6%
HL	1	1%	5	4%				
ICC	3	3%	1	1%				
Kidney			19	13%	1	14%	1	6%
Larynx			6	4%				
Lung	1	1%	32	23%			3	19%
Leukemia			1	7%			1	6%
Large intestine			1	1%				
Lip			2	1%				
Liver	1	1%	28	2			3	19%
Lymphoma	1	1%	2	1%				
Lymphatic			2	1%				
Lymphocytic leukemia			1	1%				
Melanoma			9	6%			1	6%
Mesothelioma			2	1%				
Mouth			2	1%				
Myeloma	1	1%	7	5%			1	6%
Myeloid leukemia			1	1%				

Type of cancer	Case-control		Cohort		Cross-sectional		MRS	
	No. Studies	%	No. Studies	%	No. Studies	%	No. Studies	%
NHL			17	12%			1	6%
Non-melanoma			2	1%				
Nose			1	1%				
Ovarian			18	13%			2	13%
Oral			7	5%				
Penis			3	2%				
Pharynx			5	4%				
Placenta			1	1%				
Prostate	8	9%	38	27%	2	29%	3	19%
Pancreas	15	17%	39	28%			5	31%
Peritoneum			1	1%				
Rectum	5	6%	23	16%			1	6%
RCC	1	1%	3	2%			1	6%
Rectosigmoid			1	1%				
Salivary gland			2	1%				
Skin			1	7%				
Small intestine			8	6%				
Soft tissue			4	3%				
Thyroid	3	3%	16	11%			3	19%
Testicles			4	3%			1	6%
Tongue			3	2%				
UTC			5	4%	1	14%		
Other			6	4%				
Overall			3	2%	1	14%		

The total number of studies exceeds the overall count because some publications included multiple cancer types. "Overall" refers to non-specific neoplasms. "Other" includes cancers related to the respiratory and haematological systems, intrathoracic organs, and other female and male genital organs. MRS: Mendelian Randomization Studies.

Appendix B. Causal association

This appendix provides supplementary information for Chapter 4: Causal association

B1. Risk factors

B1.1 T2DM

The following risk factors for T2DM were considered in the DAGs: age, sex, race, sedentary lifestyle, smoking, drinking, obesity, family history, hypertension, and diet and family history (Centers for Disease Control and Prevention (CDC), 2023; Chi et al., 2022; Han, 2020; Kao, 2001; Kim et al., 2015; Larsson & Burgess, 2022; National Institute for Health and Care Excellence (NICE), 2023; National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), 2022; Pal et al., 2021; Sowers et al., 2001; The British Diabetic Association, 2023; Xue et al., 2022).

B1.2 Types of cancer

The risk factors for the 56 types of cancer for which a DAG was developed, are presented in the tables below. The cancers were classified by histological type (lymphoma, leukemia, myeloma, sarcoma, and carcinoma) (National Cancer Institute, 2023).

Lymphoma

The sources consulted for these types of cancer were the American Society of Clinical Oncology's (ASCO), Cancer Australia, American Cancer Society, Lymphoma Action, and Lymphoma Australia (American Cancer Society, 2021; American Society of Clinical Oncology's (ASCO), 2023; Cancer Australia, 2023; Lymphoma Action, 2023; Lymphoma Australia, 2023).

Table B1-1 Risk factors for types of lymphoma

Type of cancer	Risk factors
Hodgkin lymphoma	1. Age 2. Sex 3. Epstein-Barr virus (EBV) 4. HIV 5. Family history of cancer
Lymphoma	1. Age 2. Family history of cancer (lymphoma/blood cancer) 3. Infections (Viral: Epstein-Barr virus [EBV], Hepatitis C virus [HCV], Human T-lymphotropic virus type 1 [HTLV-1], Human herpesvirus 8 [HHV-8]; Bacterial: Helicobacter pylori, Chlamydia psittaci, Campylobacter jejuni, Borrelia burgdorferi, Moraxella catarrhalis)

Type of cancer	Risk factors
	- Suppressed immune system - Autoimmune treatment - Chemical exposure (industrial chemicals, solvents, weed killers and insecticides) - Previous cancer treatment - Obesity
NHL	- Age - Sex - Suppressed immune system - Infections (Helicobacter pylori, Epstein-Barr virus [EBV], HIV, Epstein-Barr virus, human T-cell leukaemia/lymphoma virus [HTLV-1] or human herpesvirus 8) - Exposure to ionizing radiation (atomic bombs, nuclear reactor accidents, and medical radiation therapy)

Each risk factor was reported in at least two references. NHL stands for non-Hodgkin lymphoma.

Leukemia

The most consulted sources for these types of cancer were the American Society of Clinical Oncology's (ASCO), Cancer Research UK, the American Cancer Society, the National Cancer Institute, the Canadian Cancer Society, and the University of Rochester Medical Center (URMC) (American Cancer Society, 2021; American Society of Clinical Oncology's (ASCO), 2023; Canadian Cancer Society, 2023; Cancer Research UK, 2023; National Cancer Institute, 2023; University of Rochester Medical Center (URMC), 2023).

Table B1-2 Risk factors for types of leukemia

Type of cancer	Risk factors
Leukemia	- Age - Sex - Smoking - Rare congenital Diseases (Down syndrome, ataxia-telangiectasia, and Bloom syndrome) - MSD Myelodysplastic Syndromes (blood disorders) - Previous cancer treatment (chemotherapy) - Exposing to ionizing radiation - Family history of cancer - Chemical exposure (benzene, formaldehyde) - Race
Lymphocytic leukemia	- Age - Sex - Radiation exposure (bomb, treatment) - Chemical exposure (benzene, petrol industry, chemical industry, pharmaceutical industry, rubber industry, shoe production, printing industry) - Genetic syndromes (Down's syndrome, Fanconi anaemia, ataxia telangiectasia, Bloom syndrome) - Race - Family history
Myeloid leukemia	- Age - Sex - Smoking - Genetic disorders (Down syndrome, Ataxia telangiectasia, Li-Fraumeni syndrome, Klinefelter syndrome, Fanconi anaemia, Wiskott-Aldrich syndrome, Bloom syndrome, Familial Platelet Disorder syndrome, Germline mutations which are present at birth, the most common of which are changes to the GATA2, ETV6, CEBPA, and RUNX1 genes) - Radiation - Chemotherapy - Chemicals (benzene, oil refineries and chemical and petrochemical plants) - Bone marrow disorders (myelodysplastic syndrome, myeloproliferative disorders)

Each risk factor was reported in at least two references.

Myeloma

The sources consulted for this type of cancer were the Cancer Research UK, and the Rare Cancers Australia Ltd (Cancer Research UK, 2023; Rare Cancers Australia Ltd, 2023).

Table B1-3 Risk factors for myeloma

Type of cancer	Risk factors
Myeloma	1. Age 2. Sex 3. Chemical exposure 4. Family history of myeloma 5. MGUS (monoclonal gammopathy of undetermined significance, a rare medical condition) 6. Autoimmune treatment 7. Suppressed immune system 8. Exposing to ionizing radiation 9. HIV 10. Obesity

Each risk factor was reported in at least two references.

Sarcoma

The most consulted sources for these types of cancer were the American Society of Clinical Oncology's (ASCO), Cancer Research UK, Cancer Council NSW, American Cancer Society, University of Rochester Medical Center (URMC), and Stanford Medicine (American Cancer Society, 2021; American Society of Clinical Oncology's (ASCO), 2023; Cancer Council NSW, 2023; Cancer Research UK, 2023; Stanford Medicine, 2023; University of Rochester Medical Center (URMC), 2023).

Table B1-4 Risk factors for types of sarcoma

Type of cancer	Risk factors
Bone cancer	1. Age 2. Previous cancer treatment 3. Bone diseases (Enchondroma or osteochondroma, Ollier's disease, Paget disease) 4. Genetic factors/syndromes (Li-Fraumeni syndrome, Hereditary retinoblastoma, Werner syndrome)
Brain/nervous system cancer	1. Age 2. Obesity 3. Family history and genetic conditions (neurofibromatosis [NF] type 1 and type 2, tuberous sclerosis [TSC], Li-Fraumeni syndrome, Von Hippel-Lindau syndrome [VHL], Turner syndrome, Turcot syndrome, Gorlin syndrome) 4. Exposure to ionizing radiation
Kaposi's sarcoma	1. Sex 2. Race 3. Suppressed immune system 4. Infection (human herpes virus-8 [HHV-8]) 5. Sexual activity (men who have sex with men)
Mesothelioma	1. Age 2. Sex 3. Asbestos 2. Radiotherapy 5. Genetics
Soft tissue	1. Damaged lymph system 2. Previous radiotherapy treatment 3. Exposure to ionizing radiation 4. Genetic factors

Each risk factor was reported in at least two references.

Carcinoma

The most consulted sources for these types of cancer were the American Society of Clinical Oncology's (ASCO), Cancer Research UK, Cancer Council NSW, Centers for Disease Control and Prevention, Johns Hopkins Medicine, National Health Service, American College of Obstetricians and Gynecologists, Cancer Australia, Australian Cancer Research Foundation, Canadian Cancer Society, National Institute for Health and Care Excellence (NICE), American Cancer Society, National Cancer Institute, American Society of Colon and Rectal Surgeons (American Cancer Society, 2021; American College of Obstetricians and Gynecologists, 2024; American Society of Clinical Oncology's (ASCO), 2023; American Society of Colon and Rectal Surgeons, 2020; Australian Cancer Research Foundation, 2017; Canadian Cancer Society, 2023; Cancer Australia, 2023; Cancer Council NSW, 2023; Cancer Research UK, 2023; Centers for Disease Control and Prevention (CDC), 2023; Johns Hopkins Medicine, 2023; National Cancer Institute, 2023; National Health Service, 2023; National Institute for Health and Care Excellence (NICE), 2023).

Table B1-5 Risk factors for types of carcinoma

Type of cancer	Risk factors
Anal cancer	1. Age 2. Sex 3. HPV 4. Other cancers (cervix, vagina, or vulva) 5. HIV infection 6. Smoking 7. Suppressed immune system 8. Race
Bladder cancer	1. Age 2. Sex 3. Smoking 4. Race 5. Chemicals (amylamines, Polycyclic aromatic hydrocarbons [PAH]) 6. Radiation therapy 7. Cyclophosphamide 8. Frequent urinary tract infections (UTIs) 9. Family history 10. Long term urinary catheter-use
Breast cancer	1. Age 2. Obesity 3. Alcohol 4. Sedentary lifestyle 5. Genetic mutations 6. Family history (breast or ovarian cancer) 7. Hormone replacement therapy after menopause (HRT) 8. Breast density 9. Non-cancerous breast diseases 10. Early menstruation 11. Late menopause 12. Ethnicity 13. First pregnancy at a later age 14. Nulliparity
BTC	1. Age 2. Primary Sclerosing Cholangitis (PSC) 3. Congenital conditions (Choledochal cyst and Caroli syndrome) 4. Liver flukes

Type of cancer	Risk factors
	5. Bile duct stones (hepatolithiasis) 6. Thorotrast
Carcinoma of unknown primary	1. Age 2. Smoking 3. Sun exposure 4. Obesity
Cervix uteri cancer	1. Age 2. HIV 3. Smoking 4. Contraceptive pill (using birth control pills for a long time) 5. Multiparity 6. Human papillomavirus (HPV) 7. Other sexually transmitted infections (STI) 8. Family history 9. Previous radiotherapy treatment
Colon cancer	1. Age 2. Sex 3. Obesity 4. Drinking 5. Smoking 6. Race 7. Family history 8. Polyps 9. IBD 10. Inherited Genetic Syndromes (Lynch syndrome and familial adenomatous polyposis [FAP]) 11. Diet (low-fibre, high-fat diet) 12. Sedentary lifestyle
CRC	1. Age 2. Sex 3. Obesity 4. Drinking 5. Smoking 6. Race 7. Family history 8. Polyps 9. IBD 10. Inherited Genetic Syndromes (Lynch syndrome and familial adenomatous polyposis [FAP]) 11. Diet 12. Sedentary lifestyle
Endometrial cancer	1. Age 2. Family history 3. Race 4. Genetics (Lynch syndrome) 5. Obesity 6. Endometrial hyperplasia 7. PCOS 8. HRT (postmenopausal oestrogen-only therapy) 9. Early menstruation 10. Late menopause 11. Nulliparity 12. Tamoxifen
Esophageal cancer	1. Age 2. Sex 3. Obesity 4. Smoking 5. Drinking 6. Diet (low in fruit and vegetables; high in processed meat, smoked, salted or pickled food) 7. Barrett's oesophagus 8. Medical conditions (achalasia) 9. Gastroesophageal reflux disease (GORD/GERD) 10. Race
Esophagogastric junction adenocarcinoma	1. Age 2. Sex 3. Helicobacter pylori infection 4. Family history 5. Race 6. Diet (foods high in salt, processed meat) 7. Pernicious anaemia 8. Smoking 9. Drinking 10. Obesity 11. GORD/GERD (long-term- severe acid reflux/ gastritis) 12. Barrett's oesophagus 13. Medical conditions (achalasia)
Extrahepatic cholangiocarcinoma	1. Choledochal cyst 2. Choledocholithiasis 3. Cirrhosis

Type of cancer	Risk factors
	4. Cholelithiasis 5. Hepatitis B/C 6. Alcohol 7. Cholecystolithiasis 8. IBD 9. Smoking
Eye cancer	1. Age 2. Sex 3. Race 4. Medical conditions (nevus of Ota, Nevi) 5. Inherited cancer syndromes (mutation of BAP1, Dysplastic nevus syndrome)
Gallbladder cancer	1. Age 2. Sex 3. Family history 4. Race 5. Smoking 6. Drinking 7. Obesity 8. Inflammation of the gallbladder (cholecystitis) 9. Gallstones 10. Porcelain gallbladder 11. Primary sclerosing cholangitis (PSC) 12. Pancreas and bile duct abnormalities (born with choledochal cysts, abnormality of the join between the bile duct and the pancreas) 13. Gallbladder Polyps
GI cancer	1. Age 2. Sex 3. Diet 4. Race 5. Inherited syndromes (familial adenomatous polyposis [FAP]- MUTYH-associated polyposis, Lynch syndrome, Peutz-Jeghers syndrome, Cystic fibrosis [CF]) 6. Celiac disease 7. Sedentary lifestyle 8. Obesity 9. Drinking 10. Smoking 11. Family history 12. Polyps 13. IBD (Crohn's disease) 14. Helicobacter pylori infection 15. Pernicious anaemia 16. GORD/GERD (long-term severe acid reflux/ gastritis)
Head and Neck cancer	1. Age 2. Sex 3. Smoking 4. Alcohol 5. Prolonged sun exposure (UV-light) 6. Exposure to radiation tx (X-rays) 7. Human papillomavirus (HPV) 8. Epstein-Barr virus (EBV) 9. Environmental or occupational inhalants (asbestos, wood dust, paint fumes, and certain chemicals-nickel dust, formaldehyde) 10. Race
Hepatocellular carcinoma	1. Age 2. Sex 3. Race 4. Cirrhosis 5. Drinking 6. Non-alcoholic fatty liver disease (NAFLD) 7. Hepatitis B/C infection 8. Smoking 9. Suppressed immune system 10. Family history 11. Gallbladder removal 12. Obesity 13. Aflatoxin
Intrahepatic cholangiocarcinoma	1. Choledochal cyst 2. Choledocholithiasis 3. Cirrhosis 4. Cholelithiasis 5. Hepatitis B/C 6. Alcohol 7. IBD 8. Smoking
Kidney cancer	1. Age 2. Sex

Type of cancer	Risk factors
	3. Obesity 4. Smoking 5. Hypertension 6. Family history 7. Inherited genetic conditions (von Hippel-Lindau disease, Hereditary papillary renal cell carcinoma, Hereditary leiomyoma-renal cell carcinoma, Birt-Hogg-Dube [BHD] syndrome, Familial renal cancer, Cowden syndrome, Tuberous sclerosis) 8. Advanced kidney disease 9. Race 10. Hepatitis C
Large intestine cancer	1. Age 2. Sex 3. Sedentary lifestyle 4. Obesity 5. Drinking 6. Smoking 7. Race 8. Family history 9. Polyps 9. IBD 10. Inherited Genetic Syndromes (Lynch syndrome and familial adenomatous polyposis [FAP]) 11. Diet 13. HPV 14. Other cancers (cervix, vagina, or vulva) 15. HIV infection 16. Suppressed immune system
Larynx cancer	1. Age 2. Sex 3. Smoking 4. Drinking 5. Exposure to certain substances (asbestos, sulphuric acid) 6. Family history 7. Race
Lip cancer	1. Age 2. Sex 3. Race 4. Alcohol 5. Smoking 6. UV light (sun or sunbeds) 7. Suppressed immune system 8. Human papillomavirus (HPV) 9. Chewing Betel quid 10. Family history 11. Genetic syndromes (Fanconi anaemia, Dyskeratosis congenita)
Liver cancer	1. Age 2. Sex 3. Hepatitis B /C virus infection 4. Cirrhosis 5. Drinking 6. Aflatoxin 7. Non-alcoholic fatty liver disease (NAFLD) 8. Smoking 9. Genetic conditions (untreated hereditary hemochromatosis, alpha-1 antitrypsin deficiency, glycogen storage disease, porphyria cutanea tarda, Wilson disease) 10. Race 11. Obesity
Lung cancer	1. Age 2. Race 3. Smoking 4. Radon 5. Exposure to substances (asbestos, silica, diesel engine) 6. Air pollution (diesel engine exhaust, benzene, particulate matter and some PAHs) 7. Family History 8. Passive smoking
Melanoma	1. Age 2. Sex 3. Sun exposure (UV) 4. Indoor tanning 5. Moles (many moles or unusual moles called dysplastic nevi or atypical moles) 6. Fair skin, blonde or red hair and blue, green or grey eyes 7. Genetics (familial melanoma, xeroderma pigmentosum, retinoblastoma, Li-Fraumeni syndrome, Werner syndrome, and certain hereditary breast and ovarian cancer syndromes) 8. Race 9. Suppressed immune system 10. Family history
Mouth cancer	1. Age 2. Sex

Type of cancer	Risk factors
	3. Race 4. Alcohol 5. Smoking 6. Human papillomavirus (HPV) 7. Chewing Betel quid 8. Family history 9. Genetic syndromes (Fanconi anaemia, Dyskeratosis congenita)
Non-melanoma	1. Age 2. Sex 3. History of sunburns 4. Sun exposure (UV) 5. Fair skin, blonde or red hair and blue, green or grey eyes 6. Indoor tanning 7. Family history of skin cancer 8. Skin conditions (Actinic keratosis) 9. Genetics (Gorlin syndrome, Xeroderma pigmentosum) 10. Exposure to chemicals (arsenic compounds) 11. Race 12. Suppressed immune system
Nose cancer	1. Age 2. Sex 3. Workplace exposures (wood dust-carpentry, leather dust-shoemaking, nickel stainless steel and radium) 4. Smoking 5. HPV 6. Race 7. Radiation treatment for retinoblastoma
Oral cancer	1. Age 2. Sex 3. Race 4. Alcohol 5. Smoking 6. UV light (sun or sunbeds) 7. Suppressed immune system 8. Human papillomavirus (HPV) 9. Chewing Betel quid 10. Family history 11. Genetic syndromes (Fanconi anaemia, Dyskeratosis congenita)
Ovarian cancer	1. Age 2. Family history 3. Genetic mutation (BRCA1 or BRCA2, or one associated with Lynch syndrome) 4. Race 5. Endometriosis 6. Nulliparity 7. Early menstruation 8. Late menopause 9. Hormone replacement therapy (HRT) 10. Smoking 11. Asbestos exposure 12. Obesity
Pancreatic cancer	1. Age 2. Sex 3. Race 4. Smoking 5. Obesity 6. Family history 7. Rare inherited conditions (Hereditary pancreatitis [HP], Peutz-Jeghers syndrome [PJS], Familial malignant melanoma and pancreatic cancer [FAMM-PC], Hereditary breast and ovarian cancer [HBOC] syndrome, Lynch syndrome, Li-Fraumeni syndrome [LFS], Familial adenomatous polyposis [FAP]) 8. Chronic Pancreatitis
Penile cancer	1. Age 2. Human papillomavirus (HPV) 3. Race 4. Smoking 5. Suppressed immune system 6. Uncircumcised men with phimosis 7. Psoriasis treatment (psoralen and light therapy [phototherapy])
Pharynx cancer	1. Age 2. Sex 3. Epstein-Barr virus (EBV) 4. Drinking 5. Smoking 6. Family history 7. Human papillomavirus (HPV) 8. Race

Type of cancer	Risk factors
Prostate cancer	1. Age 2. Race 3. Family history 4. Obesity 5. Inherited genes and syndromes (BRCA2, Lynch syndrome)
Rectal cancer	1. Age 2. Sex 3. Diet 4. Smoking 5. Drinking 6. Obesity 7. Sedentary lifestyle 8. Family history of colorectal cancer 9. IBD 10. Polyps 11. Race
Rectosigmoid cancer	1. Age 2. Sex 3. Obesity 4. Drinking 5. Smoking 6. Race 7. Family history 8. Polyps 9. IBD 10. Inherited Genetic Syndromes (Lynch syndrome and familial adenomatous polyposis [FAP]) 11. Diet 12. Sedentary lifestyle
Salivary gland cancer	1. Age 2. Radiation therapy to the head or neck (medical treatment or due to their workplace) 3. Family history
Skin cancer	1. Age 2. Sex 3. Sun exposure (UV) 4. Indoor tanning 5. Moles (many moles or unusual moles called dysplastic nevi or atypical moles) 6. Fair skin, blonde or red hair and blue, green or grey eyes 7. Genetics (familial melanoma, xeroderma pigmentosum, retinoblastoma, Gorlin syndrome, Li-Fraumeni syndrome, Werner syndrome, and certain hereditary breast and ovarian cancer syndromes) 8. Race 9. Suppressed immune system 10. Family history 11. History of sunburns 12. Skin conditions (Actinic keratosis) 13. Exposure to chemicals (arsenic compounds)
Small intestine cancer	1. Age 2. Sex 3. Diet (high in red meat and salted or smoked foods, high fat diet) 4. Race 5. Inherited syndromes (familial adenomatous polyposis [FAP]- MUTYH-associated polyposis, Lynch syndrome, Peutz-Jeghers syndrome, Cystic fibrosis [CF]) 6. Celiac disease 7. Crohn's disease
Stomach cancer	1. Age 2. Sex 3. Helicobacter pylori infection 4. Family history 5. Race 6. Diet (foods high in salt, processed meat) 7. Pernicious anaemia 8. Smoking 9. Drinking 10. Obesity 11. GORD/GERD (long-term- severe acid reflux/ gastritis)
Testicular cancer	1. Age 2. Undescended testicles (cryptorchidism) 3. Family history 4. Race
Thyroid cancer	1. Age 2. Sex 3. Race 4. Obesity 5. Non-cancerous (benign) thyroid conditions (enlarged thyroid [goitre], Hashimoto's disease, adenomas) 6. Family history

Type of cancer	Risk factors
	7. Inherited faulty genes (MEN2A and MEN2B, altered RET oncogene) 8. Radiation (radiotherapy, nuclear accidents or weapons)
Tongue cancer	1. Sex 2. Smoking 3. Drinking 4. Human papillomavirus (HPV)
UTC	1. Age 2. Sex 3. Smoking 4. Obesity 5. Race 6. Chemicals (amylamines, Polycyclic aromatic hydrocarbons [PAH]) 7. Radiation therapy 8. Cyclophosphamide 9. Frequent urinary tract infections (UTIs) 10. Family history 11. Long term urinary catheter-use 12. Hypertension 13. Inherited genetic conditions (von Hippel-Lindau disease, Hereditary papillary renal cell carcinoma, Hereditary leiomyoma-renal cell carcinoma, Birt-Hogg-Dube (BHD) syndrome, Familial renal cancer, Cowden syndrome, Tuberous sclerosis) 14. Advanced kidney disease 15. Hepatitis C

Each risk factor was reported in at least two references. The risk factors for Rectosigmoid cancer were assumed to be the same as CRC; the ones for Skin cancer were the result of non-melanoma and melanoma risk factors; UTC factors were assumed to be bladder and kidney cancer risk factors; large intestine risk factors were assumed to be CRC plus anal ones. *Abbreviations.* BTC: Biliary tract carcinoma; CRC: colorectal cancer; GI: gastrointestinal; UTC: urinary tract cancer.

B1.3 Associations between risk factors

The following sources were consulted to support an association between risk factors.

Worldwide organizations, societies, and universities (American Academy of Orthopaedic Surgeons, 2023; American Cancer Society, 2021; American College of Gastroenterology, 2024; American College of Obstetricians and Gynecologists, 2024; Canadian Cancer Society, 2023; Cancer Research UK, 2023; Centers for Disease Control and Prevention (CDC), 2023; ERN Board of Member States, 2020; Johns Hopkins Medicine, 2023; Lymphoma Action, 2023; Lymphoma Australia, 2023; National Cancer Institute, 2023; National Health Service, 2023; National Institute for Health and Care Excellence (NICE), 2023; National Organization for Rare Disorders, 2023; New York State Department of Health, 2023; Public Health England, 2019; The American Society of Retina Specialists, 2024; University of Florida Health, 2022; University of Oxford. Medical Sciences Division, 2024; University of Rochester Medical Center (URMC), 2023)

Publications (Alkerwi et al., 2015; Andrén-Sandberg, 2012; Baden et al., 2014; Berbudi et al., 2019; Britt & Bonci, 2013; Ceylan & Özerdoğan, 2015; Chakraborty, 2014; Chen-Sankey et al., 2021; Dare et al., 2015; De Socio et al., 2020; Garawi et al., 2014; Griffith et al., 2007; Husain et al., 2014; Inelmen et

al., 2007; Isayama et al., 2018; Kusunoki & Hayashi, 2008; Lin et al., 2015; Liu & Zhang, 2017; Martins et al., 2021; McCullough et al., 2022; Molina et al., 2010; Morimoto et al., 2021; Newby et al., 2003; Peters et al., 2009; Reis et al., 2018; Shackley et al., 2017; Silveira et al., 2022; Sopori, 2002; Sripa et al., 2007; Standerwick et al., 2007; Ullrich et al., 2007; Weyand & Goronzy, 2016)

Table B1-6 Associations between risk factors

Edge originates from	Edge terminates at
Age	Smoking, diet, hypertension, sedentary lifestyle, drinking (alcohol), bile duct stones/gallstones, myelodysplastic syndromes, PSC, suppressed immune system, urinary catheter use (long term), UTI, breast density, non-cancerous breast diseases, HPV, STI, polyps-colon/bowel, endometriosis, endometrial hyperplasia, Barret's esophagus, achalasia, porcelain gallbladder, gallbladder polyps, H. pylori infection, pernicious anaemia, celiac disease, NAFLD, advanced kidney disease (dialysis), contraceptive pill, not full-term pregnancy, HRT (only oestrogen/combined)
Autoimmune treatment	Suppressed immune system
Cancer treatment	Suppressed immune system
Celiac disease	Pernicious anaemia
Chemical exposure	Myelodysplastic syndromes
Chemotherapy	Myelodysplastic syndromes
Crohn's disease	Pernicious anaemia
Diet	Obesity, hypertension, GERD/GORD
Drinking alcohol	GERD/GORD, obesity, hypertension, diet, suppressed immune system, STI, cirrhosis, chronic pancreatitis
Early menstruation	Endometriosis, endometrial hyperplasia
Endometriosis	Pregnancy/infertility/nulliparity
Ethnicity	Aflatoxin
First pregnancy at a later age	Late menopause
Gallbladder polyps	Cholecystitis
Gallstones	Polyps-colon/bowel, cholecystitis, porcelain gallbladder, gallbladder polyps
GERD/GORD	Barret's esophagus
Genetics	Bone diseases (Enchondroma or osteochondroma, Ollier's disease, Paget disease)
Hepatitis B/C infection	Cirrhosis
HIV	Smoking, suppressed immune system
HRT	Breast cancer, breast density
HRT (only oestrogen)	Endometrial hyperplasia
Hypertension	NAFLD, advanced kidney disease (dialysis)
IBD	PSC, polyps-colon/bowel
Late menopause	Endometriosis, endometrial hyperplasia
Life-long irregular periods	Late menopause
Multiparity	Late menopause
NAFLD	Cirrhosis
Nulliparity	Endometrial hyperplasia
Obesity	Hypertension, bile duct stones/gallstones, damaged lymph system or lymphedema, breast density, polyps-colon/bowel, endometriosis, endometrial hyperplasia, GERD/GORD, Barret's esophagus, NAFLD, advanced kidney disease (dialysis), not full-term pregnancy
PCOS	T2DM, multiparity/nulliparity/infertility, endometrial hyperplasia, NAFLD
Pregnancy/breastfeeding	Breast density
PSC	Gallbladder polyps
Psoriasis treatment (PUVA)	Suppressed immune system
Race	Smoking, diet, hypertension, HIV, liver flukes, chemicals (pesticide), bile duct stones/gallstones, drinking (alcohol), HPV, STI, nevus of ota, nevi, Barret's esophagus, H. pylori infection, pernicious anaemia, Crohn's disease, NAFLD, advanced kidney disease (dialysis), chewing betel quid, fair skin
Radiation	Suppressed immune system, myelodysplastic syndromes
Radiotherapy	Damaged lymph system or lymphedema
Sedentary lifestyle	Obesity
Sex	Smoking, diet, obesity, sedentary lifestyle, HIV, myelodysplastic syndromes, bile duct stones/gallstones, PSC, drinking (alcohol), UTI, polyps-colon/bowel, Barret's esophagus, achalasia, porcelain gallbladder, sulphuric acid fumes exposure, nickel-dust exposure, asbestos exposure, Hepatitis B/C infection, chronic pancreatitis, silica exposure, urinary catheter use (Long term), HPV, arsenic exposure, celiac disease
Sexual activity	Infections
Smoking	Obesity, hypertension, drinking (alcohol), suppressed immune system, IBD, polyps-colon/bowel, endometrial hyperplasia, Barret's esophagus, GERD/GORD, Crohn's disease, NAFLD, not full-term pregnancy
Suppressed immune system	Infections, HPV, Hepatitis B/C infection

Edge originates from	Edge terminates at
T2DM	Cholecystitis, suppressed immune system, endometrial hyperplasia, NAFLD, advanced kidney disease (dialysis)
Tamoxifen	Endometrial hyperplasia
Urinary catheter use	UTI

The association was reported in at least one reference source (worldwide organizations, societies, universities, publications). The risk factors for cholecystolithiasis were the same as those for gallstones. *Abbreviations.* GERD/GORD: Gastroesophageal reflux; HIV: human immunodeficiency virus; HPV: human papillomavirus; HRT: hormone replacement therapy; IBD: inflammatory bowel disease; NAFLD: non-alcoholic fatty liver disease; PCOS: polycystic ovarian syndrome; PSC: Primary Sclerosing Cholangitis; PUVA: psoralen and long-wave ultraviolet radiation (UVA); STI: sexually transmitted infections; UTI: urinary tract infections.

Table B1-7 Omitted edges (associations) between risk factors

Edge originates from	Edge terminates at	Reference
Age	Obesity	(Leahy et al., 2014; National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), 2022)
Age	Infections: HCV, H. pylori	(Mhaskar et al., 2013; Mohd Hanafiah et al., 2013)
Age	HIV	(Centers for Disease Control and Prevention (CDC), 2023)
Age	Epstein-Barr virus	(Glaser et al., 1997)
HIV	Drinking (alcohol)	(Ikeda et al., 2016)
HIV	Obesity	(Tshikuka et al., 2024)
Race	Breast density	(del Carmen et al., 2007)
Sedentary lifestyle	Drinking (alcohol)	(Rhodes et al., 2012)
SES	Suppressed immune system	(Hoes et al., 2018)
SES	HIV	(Ekholuenetale et al., 2021; Ogunmola et al., 2014)
Sex	Hypertension	(Centers for Disease Control and Prevention (CDC), 2023)
Sex	Infections: H. pylori, Chlamydia	(Crichton et al., 2015; Zamani et al., 2018)
Sex	Chemical exposure (dioxin)	(Apostoli et al., 2005; Harden et al., 2007; Schettgen et al., 2011)

The association was reported in at least one reference source (worldwide organizations, societies, universities, publications). *Abbreviations.* H. pylori: Helicobacter pylori; HCV: hepatitis C virus; HIV: human immunodeficiency virus; SES: Socioeconomic Status.

B2. Directed Acyclic graphs

The colours in the DAGs represent the following (Textor et al., 2016):

	Lines in green represent a front door (causal path), arrowheads pointing mainly from the exposure to the outcome of interest.
	Lines in purple represent part of an open back door (noncausal or biased path). Generally, indicate common causes (confounders) of the exposure and the outcome. An open backdoor means that, in addition to the direct causal path between the exposure and the outcome, there is also an indirect path through the confounder variables, which bias the association.
	Lines in black represent neither a front or back door path.
	Green variables are ancestors (“upstream” variables) of the exposure
	Blue variables are ancestors (“upstream” variables) of the outcome
	Red variables are ancestors of the exposure and the outcome
	The exposure is represented in green with a triangle
	The outcome is represented in blue with a “I”

The following graphs show the DAGs for each type of cancer, classified by histological type.

Lymphoma

Hodgkin Lymphoma

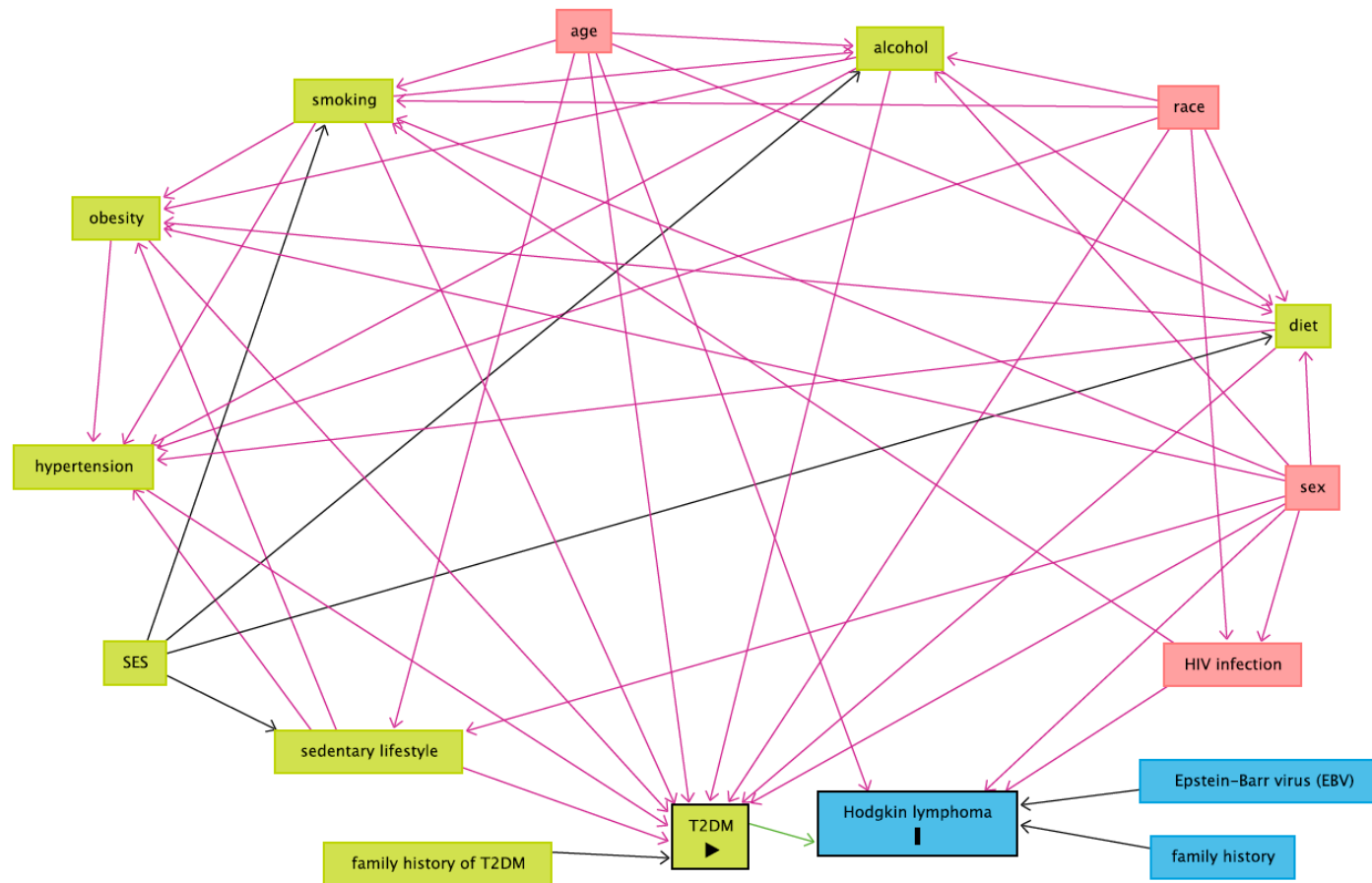


Figure B2-1 DAG of the association between T2DM and risk of Hodgkin lymphoma, including potential confounders. *Abbreviations.* HIV: human immunodeficiency virus; SES: socioeconomic status. Family history of cancer: having a first-degree relative with Hodgkin lymphoma.

Lymphoma

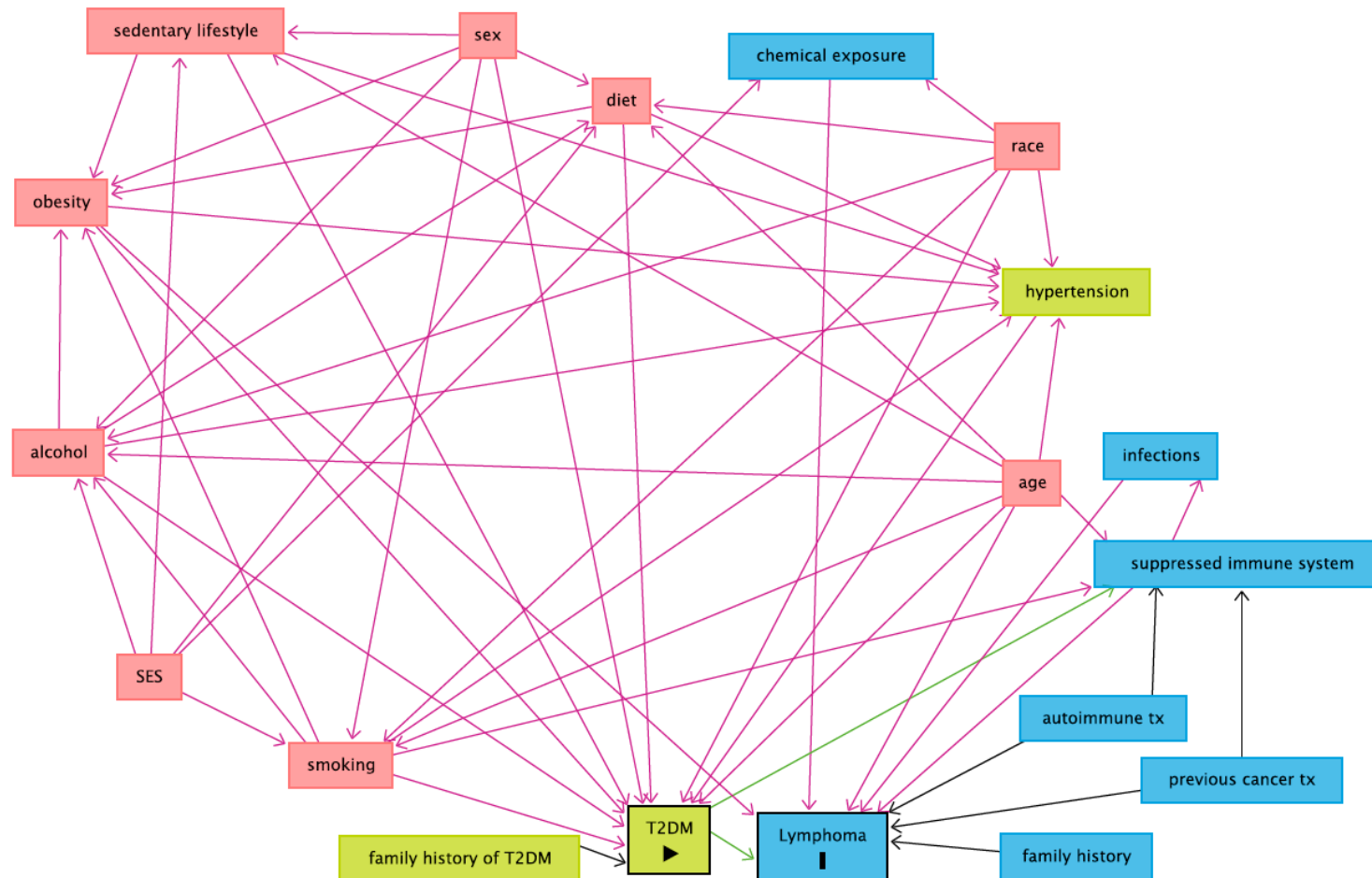


Figure B2-2 DAG of the association between T2DM and risk of lymphoma, including potential confounders. Family history of cancer: lymphoma/blood cancer. Infections: Viral. Epstein-Barr virus (EBV); Hepatitis C virus (HCV); Human T-lymphotropic virus type 1 (HTLV-1); Human herpesvirus 8 (HHV-8). Bacterial: *Helicobacter pylori*; *Chlamydia psittaci*; *Campylobacter jejuni*; *Borrelia burgdorferi*; *Moraxella catarrhalis*. Chemical exposure: industrial chemicals, solvents, weed killers and insecticides. *Abbreviations.* SES: socioeconomic status; tx: treatment.

NHL

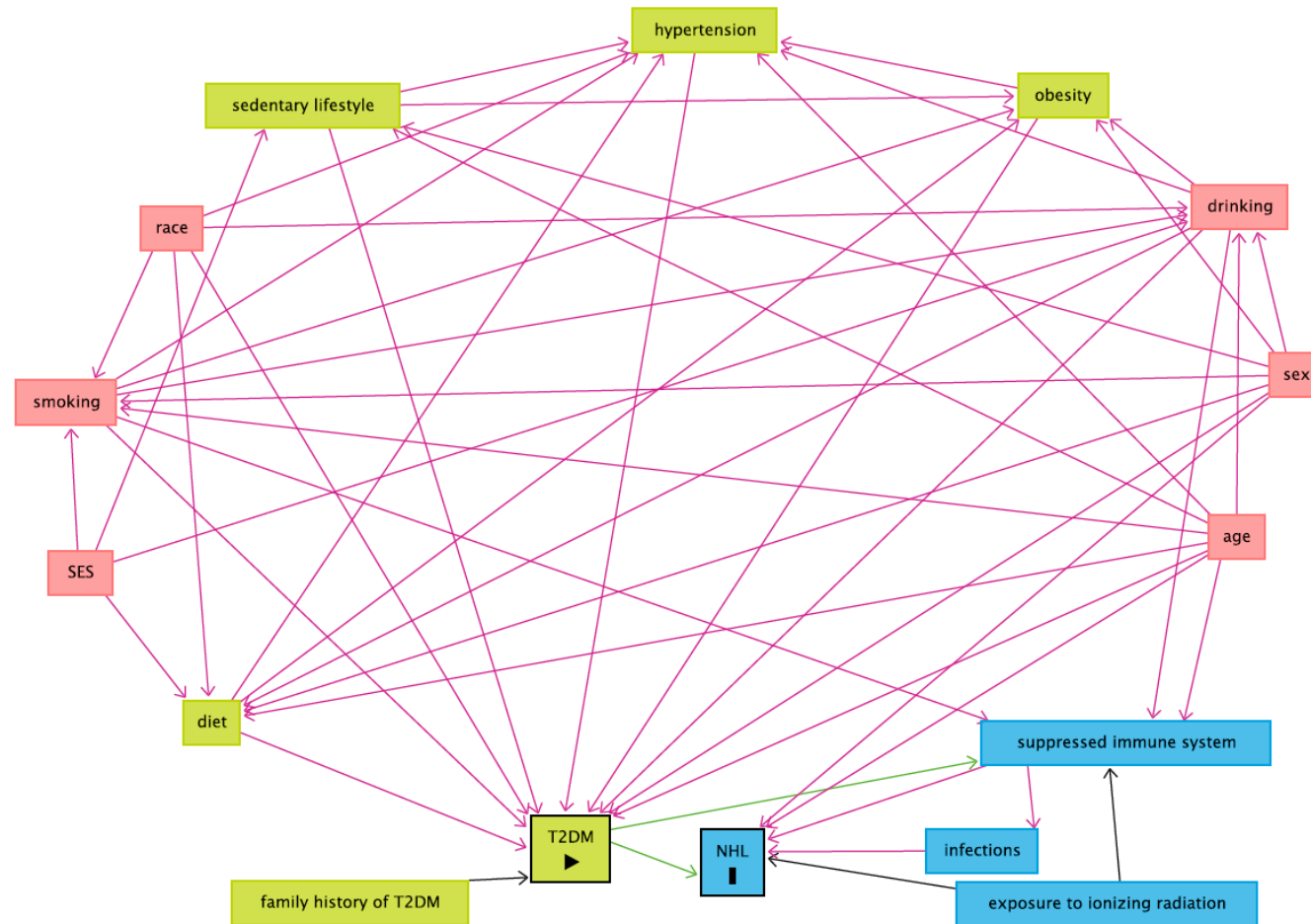


Figure B2-3 DAG of the association between T2DM and risk of NHL, including potential confounders. Infections: *Helicobacter pylori*, Epstein-Barr virus (EBV), HIV, human T-cell leukaemia/lymphoma virus (HTLV-1) or human herpesvirus 8. Ionizing radiation: atomic bombs, nuclear reactor accidents, and medical radiation therapy. Drinking defined as alcohol intake. SES: socioeconomic status.

Leukemia

Leukemia

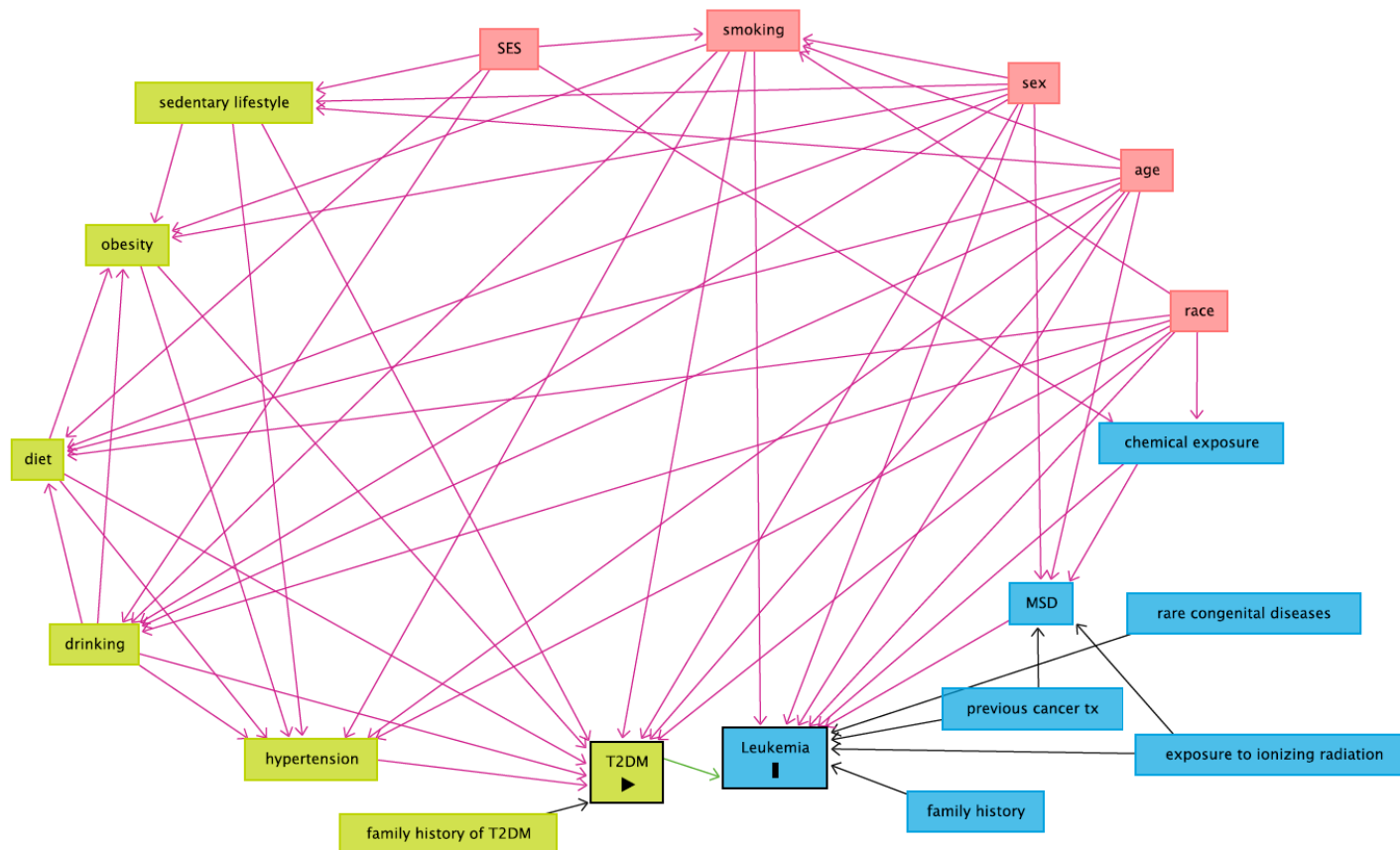


Figure B2-4 DAG of the association between T2DM and risk of leukemia, including potential confounders. Chemical exposure: benzene, formaldehyde. Family history: relatives who have AML or CLL. Rare congenital diseases: Down syndrome, ataxia-telangiectasia, and Bloom syndrome. Drinking defined as alcohol intake. MSD: Myelodysplastic Syndromes. *Abbreviations.* SES: socioeconomic status; tx: treatment.

Lymphocytic leukemia

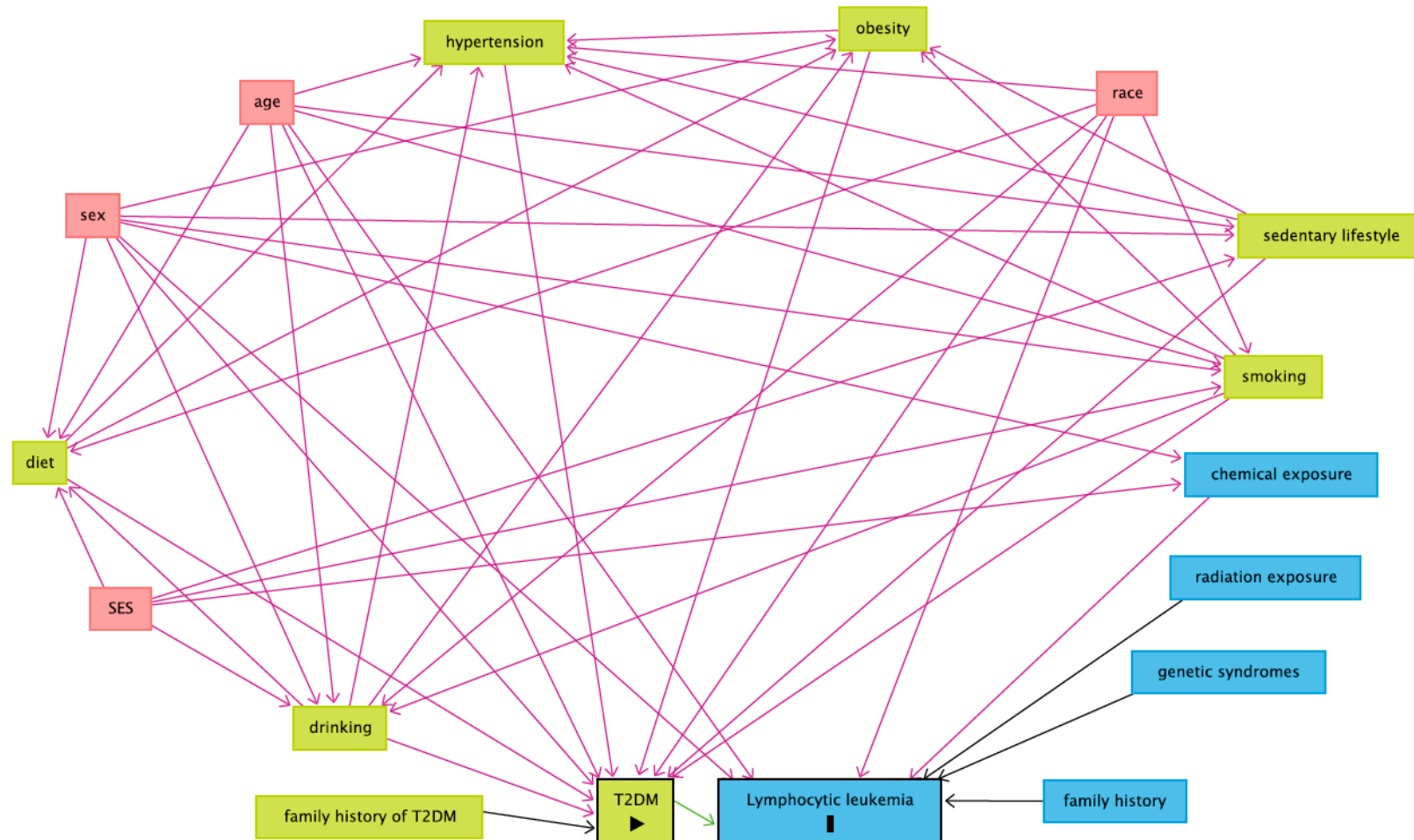


Figure B2-5 DAG of the association between T2DM and risk of Lymphocytic leukemia, including potential confounders. Chemical exposure: benzene- petrol industry, chemical industry, pharmaceutical industry, rubber industry, shoe production, printing industry. Genetic syndromes: Down's syndrome, Fanconi anaemia, ataxia telangiectasia, Bloom syndrome. Drinking defined as alcohol intake. SES: socioeconomic status.

Myeloid leukemia

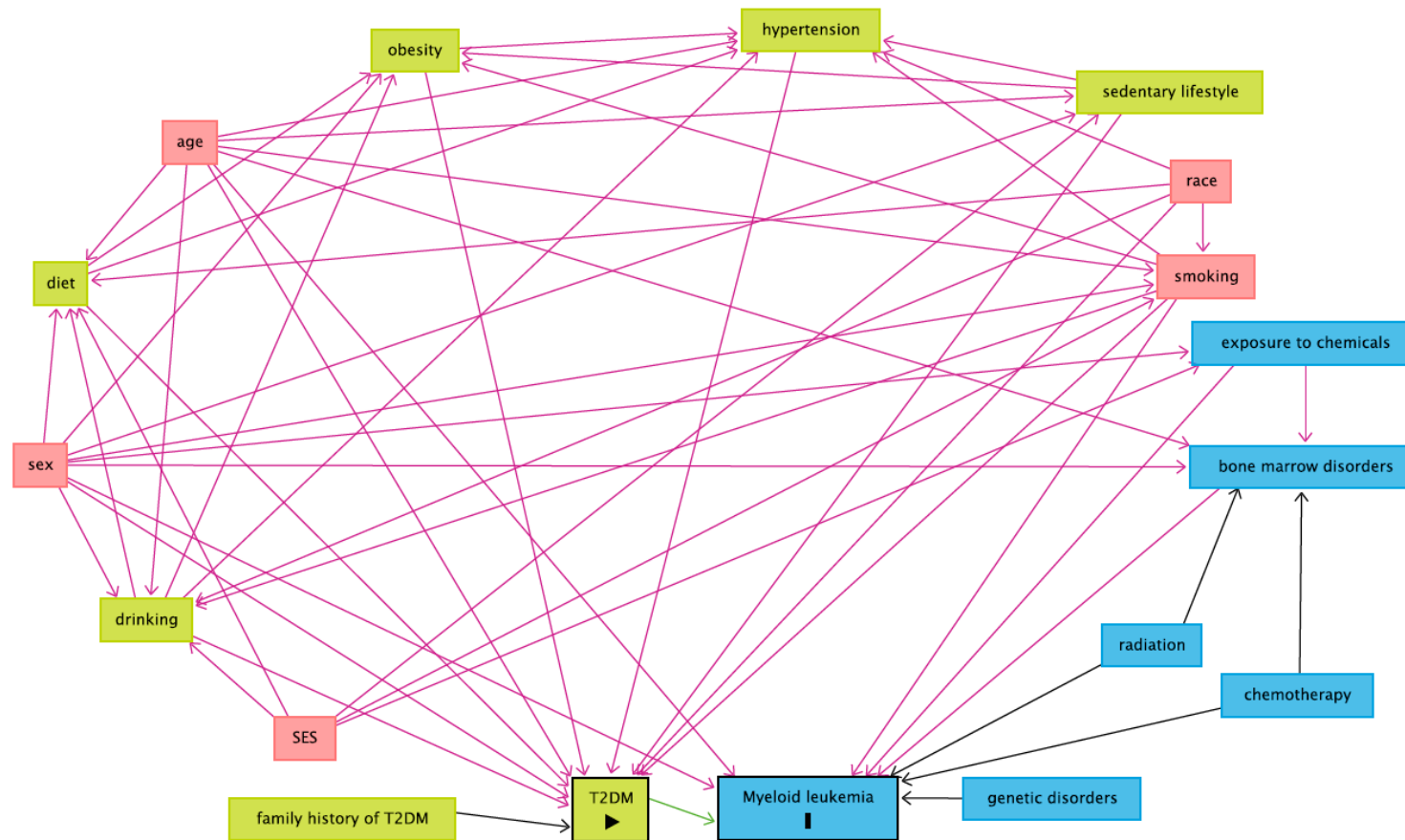


Figure B2-6 DAG of the association between T2DM and risk of Myeloid leukemia, including potential confounders. Bone marrow disorders: myelodysplastic syndrome, myeloproliferative disorders. Exposure to chemicals: benzene, oil refineries and chemical and petrochemical plants. Genetic disorders: Down syndrome, Ataxia telangiectasia, Li-Fraumeni syndrome, Klinefelter syndrome, Fanconi anaemia, Wiskott-Aldrich syndrome, Bloom syndrome, Familial Platelet Disorder syndrome, Germline mutations (GATA2, ETV6, CEBPA, and RUNX1 genes). Drinking defined as alcohol intake. SES: socioeconomic status.

Myeloma

Myeloma

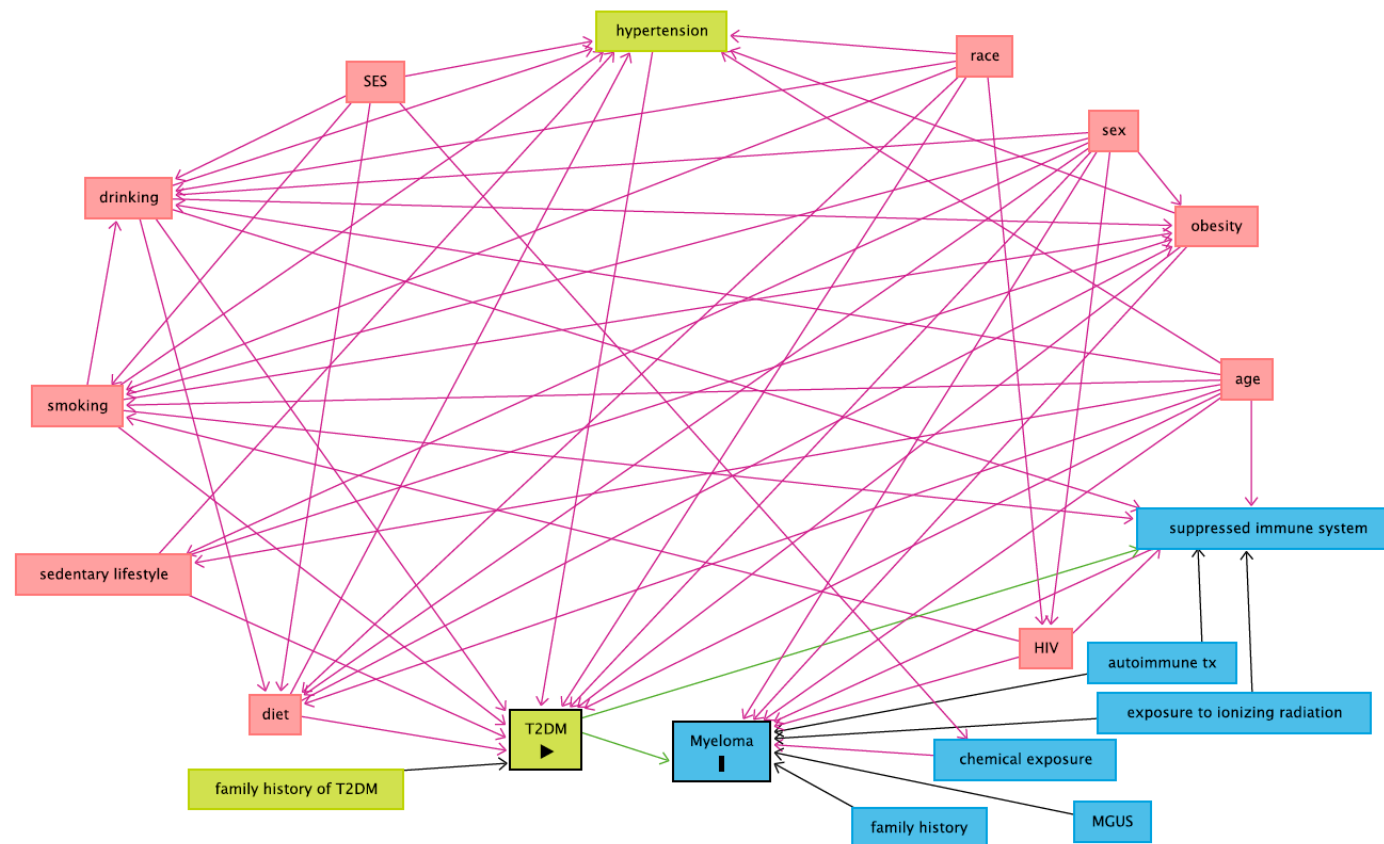


Figure B2-7 DAG of the association between T2DM and risk of Myeloma, including potential confounders. Suppressed immune system: take medicines to lower immunity after an organ transplant, HIV, exposure to high-dose radiation, exposure to certain chemicals (dioxin). Drinking defined as alcohol intake. *Abbreviations.* HIV: human immunodeficiency virus; MGUS: monoclonal gammopathy of undetermined significance; SES: socioeconomic status; tx: treatment.

Sarcoma

Bone cancer

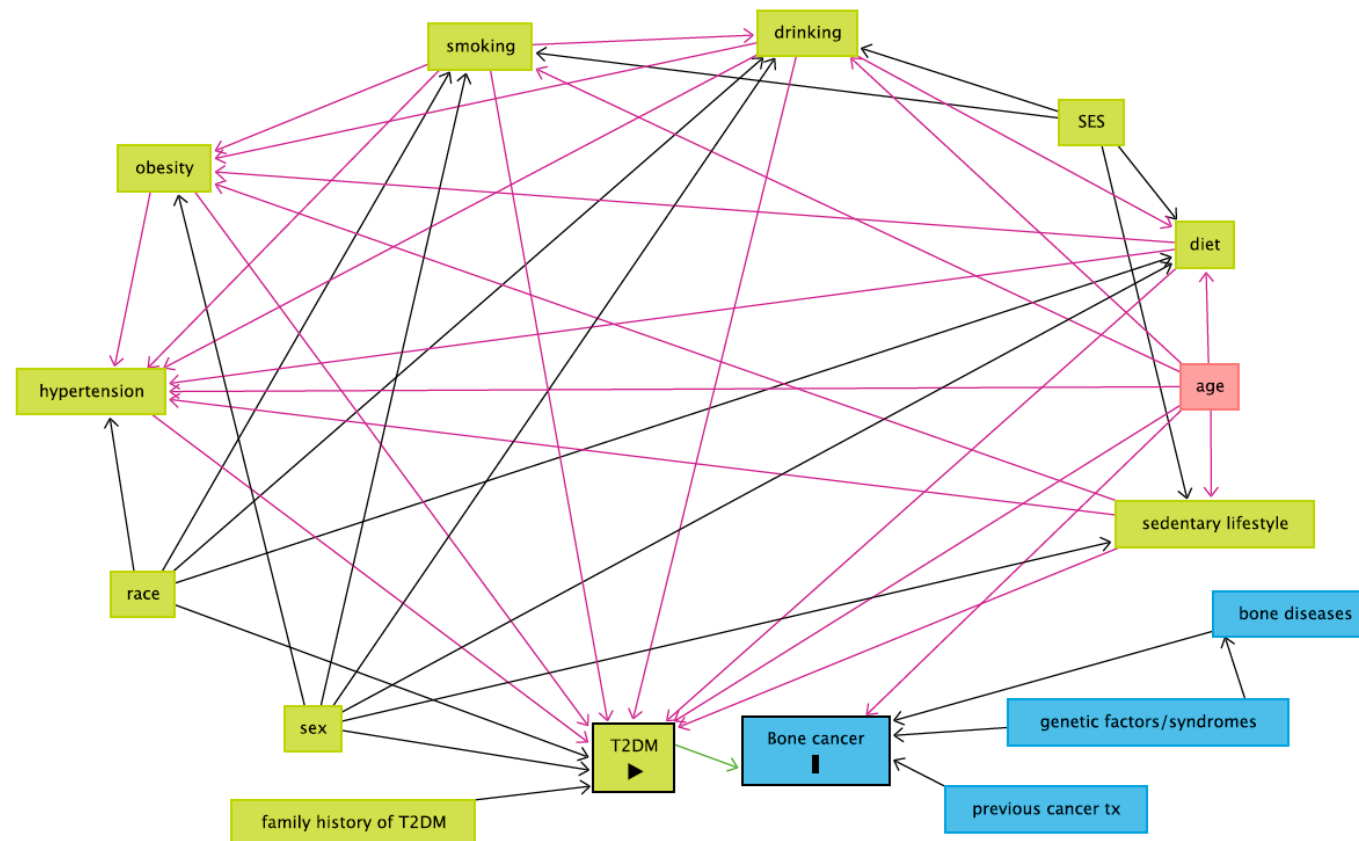


Figure B2-8 DAG of the association between T2DM and risk of bone cancer, including potential confounders. Bone diseases: Enchondroma or osteochondroma, Ollier's disease, Paget disease. Genetic factors/syndromes: Li-Fraumeni syndrome, Hereditary retinoblastoma, Werner syndrome. Drinking defined as alcohol intake. *Abbreviations.* SES: socioeconomic status; tx: treatment.

Brain/nervous system cancer

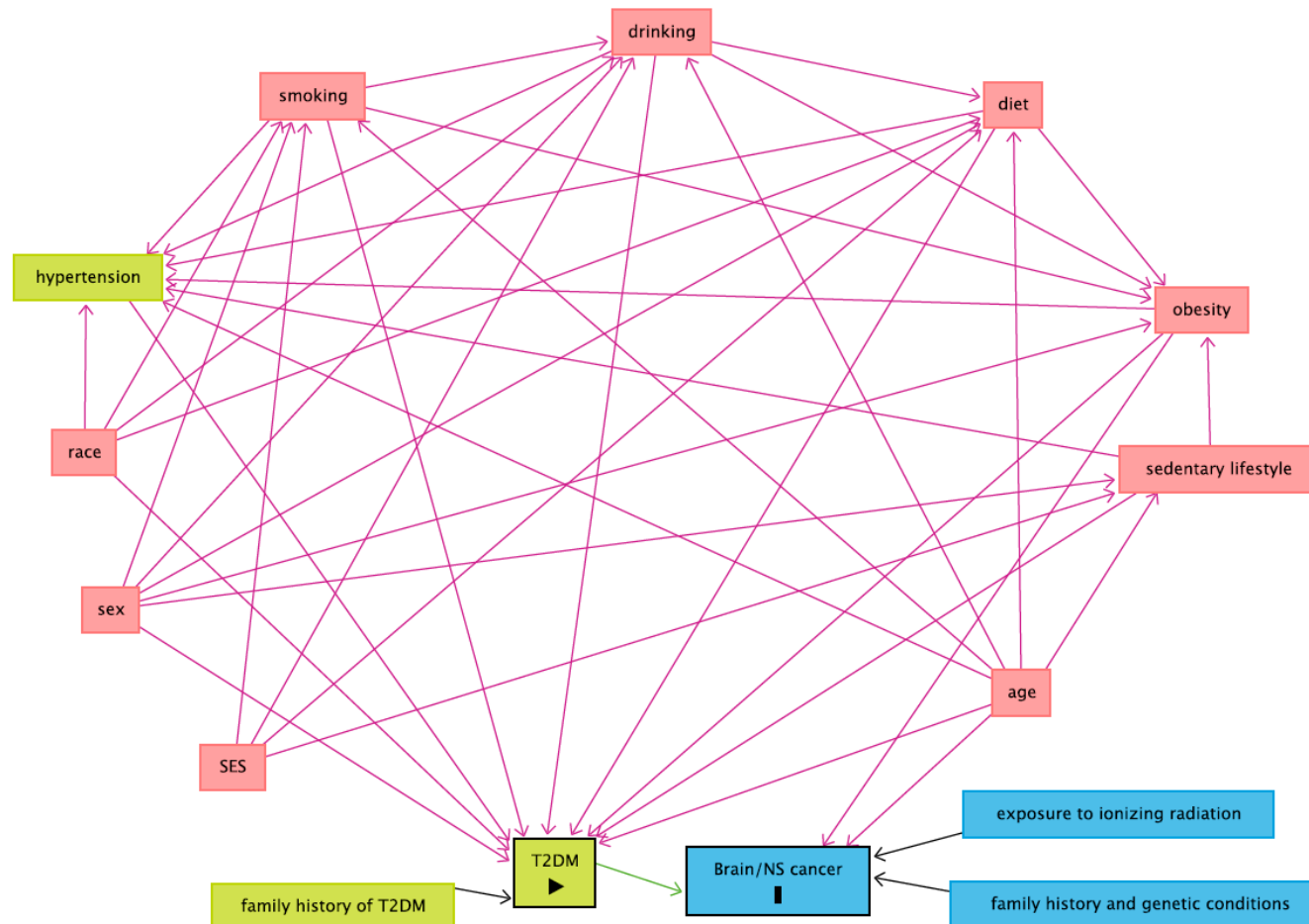


Figure B2-9 DAG of the association between T2DM and risk of brain/nervous system cancer, including potential confounders. Family history and genetic conditions: neurofibromatosis (NF) type 1 and type 2, tuberous sclerosis (TSC), Li-Fraumeni syndrome, Von Hippel-Lindau syndrome (VHL), Turner syndrome, Turcot syndrome, Gorlin syndrome. Drinking defined as alcohol intake. *Abbreviations.* NS: nervous system; SES: socioeconomic status.

Kaposi's sarcoma

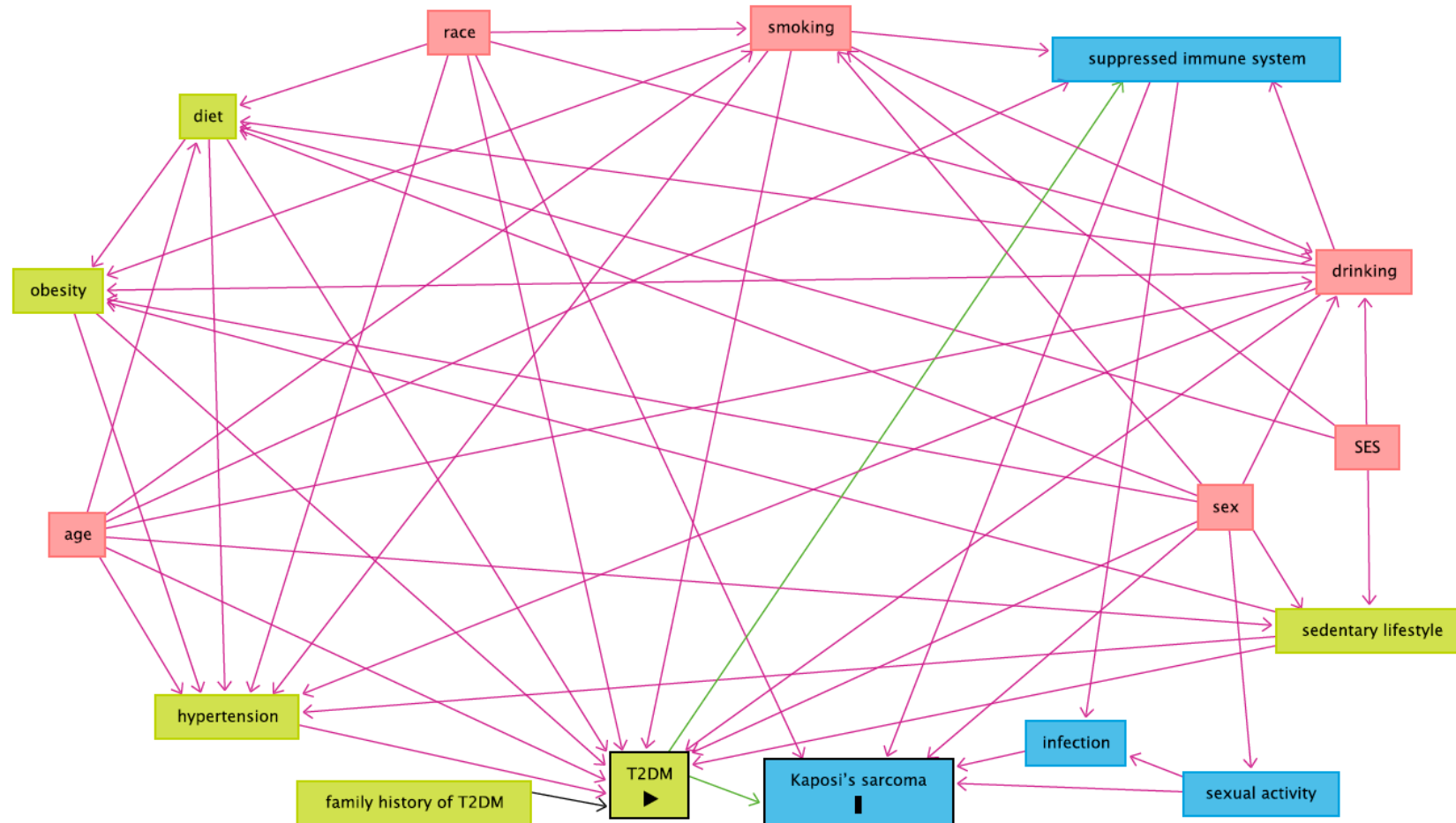


Figure B2-10 DAG of the association between T2DM and risk of Kaposi's sarcoma, including potential confounders. Infection: human herpes virus-8 (HHV-8). Sexual activity: men who have sex with men are at higher risk for Kaposi's sarcoma. Drinking defined as alcohol intake. SES: socioeconomic status.

Mesothelioma

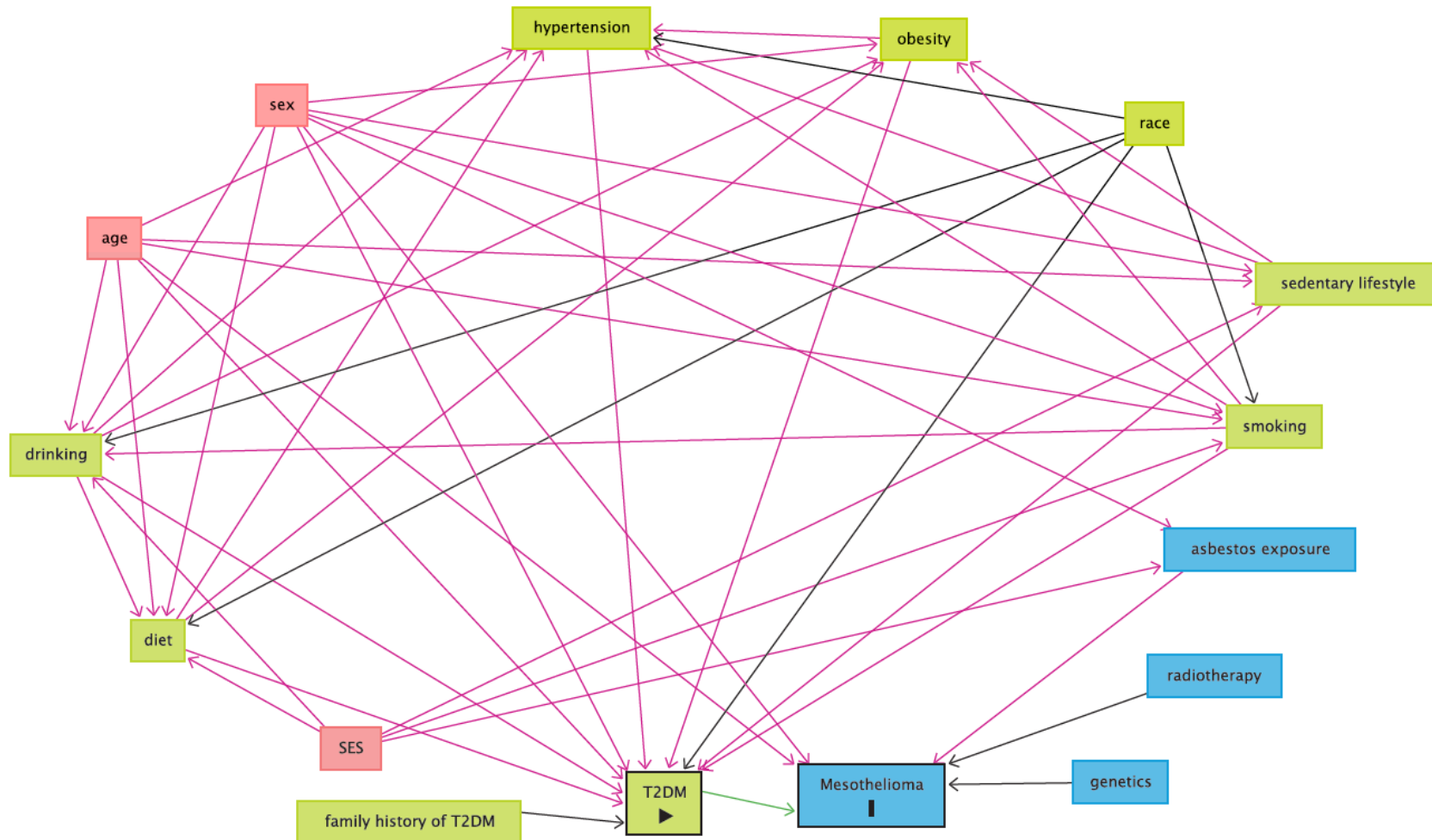


Figure B2-11 DAG of the association between T2DM and risk of Mesothelioma, including potential confounders. Drinking defined as alcohol intake. SES: socioeconomic status.

Soft tissue sarcoma

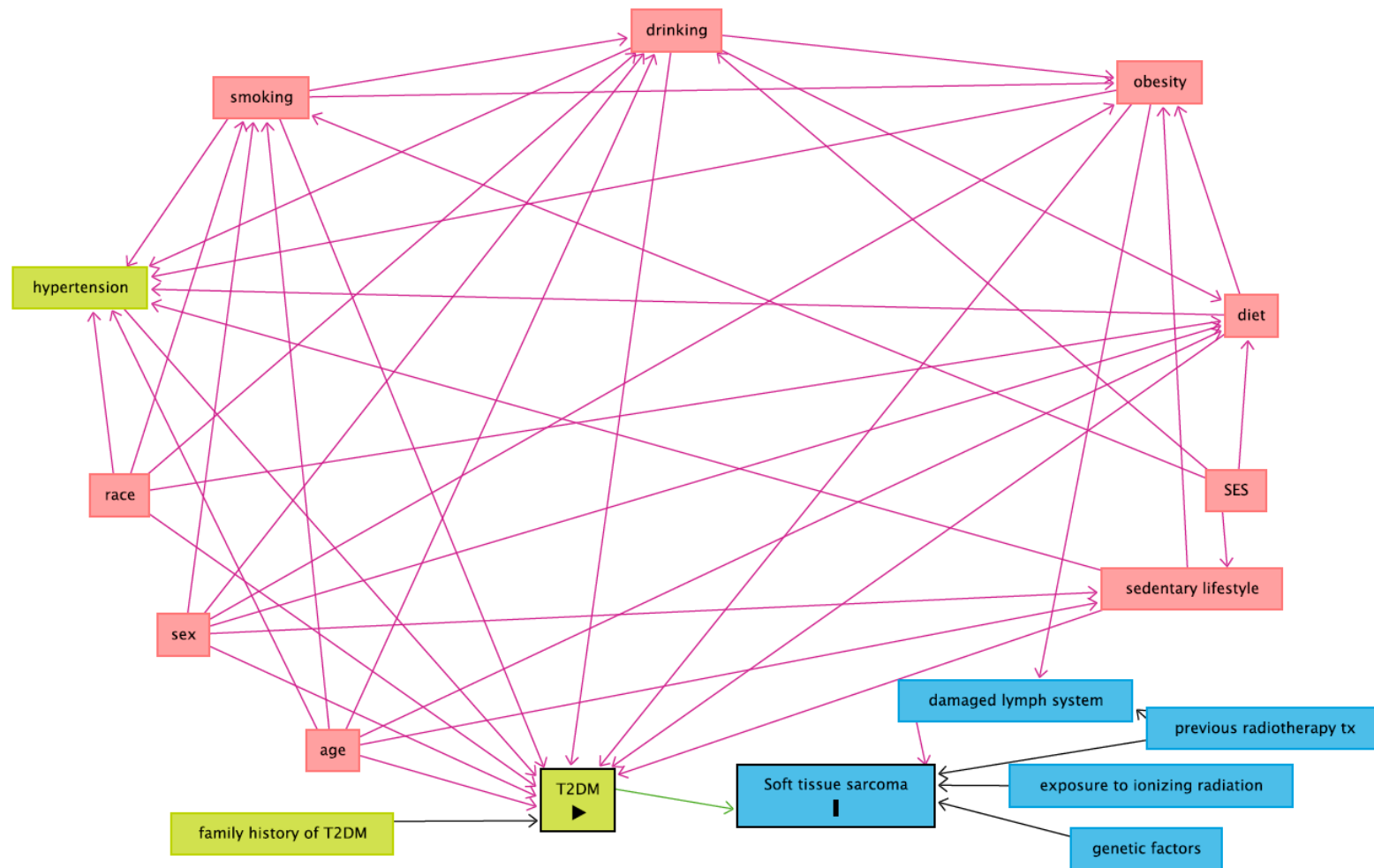


Figure B2-12 DAG of the association between T2DM and risk of soft tissue sarcoma, including potential confounders. Genetic factors: Li-Fraumeni syndrome, Retinoblastoma, Neurofibromatosis (NF), Werner syndrome, Gorlin syndrome, Tuberous sclerosis. Drinking defined as alcohol intake. *Abbreviations.* SES: socioeconomic status; tx: treatment.

Carcinoma

Anal cancer

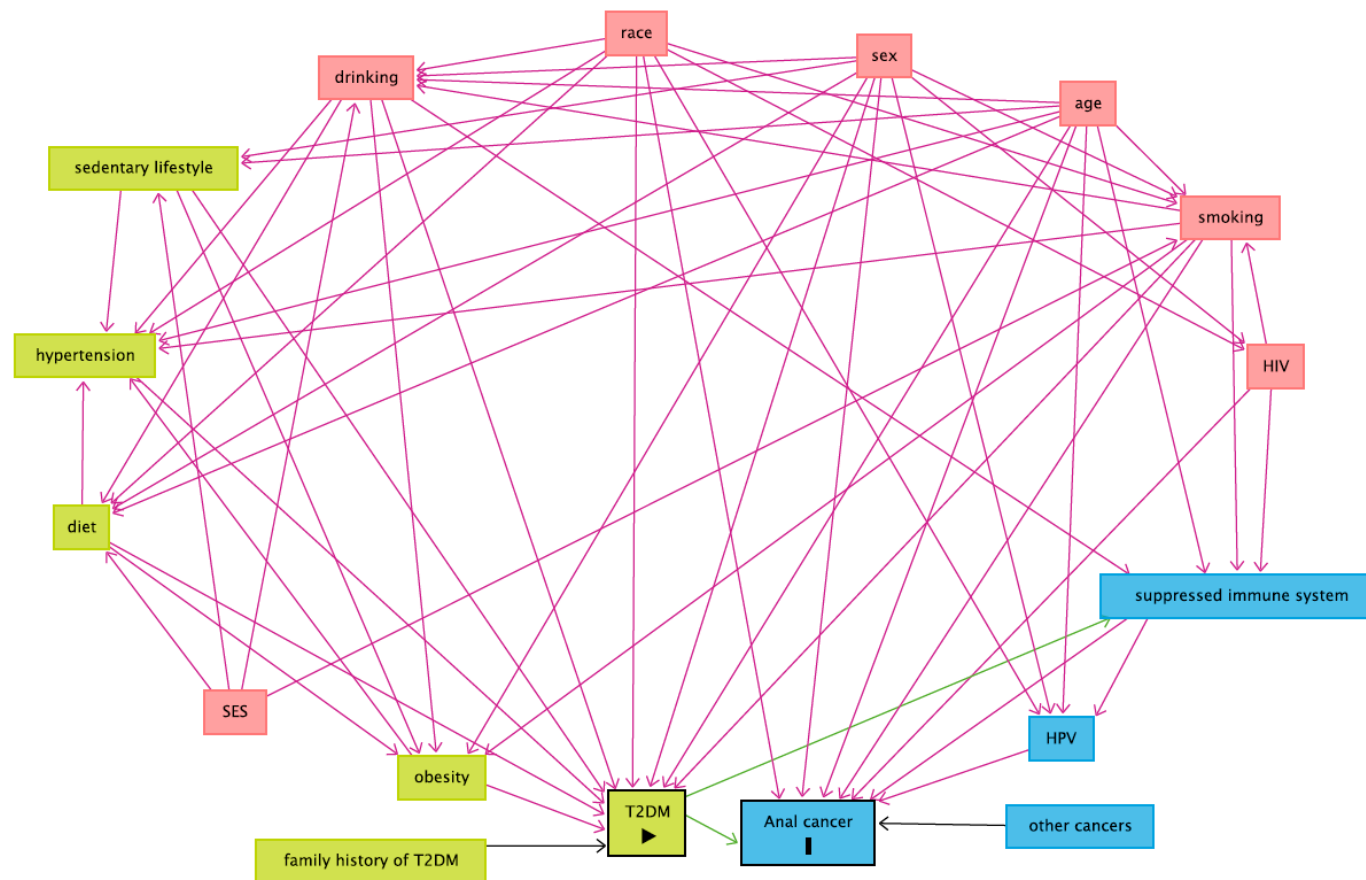


Figure B2-13 DAG of the association between T2DM and risk of anal cancer, including potential confounders. Other cancers: cervix, vagina, or vulva. Drinking defined as alcohol intake. *Abbreviations.* HIV: human immunodeficiency virus; HPV: human papillomavirus; SES: socioeconomic status.

Bladder cancer

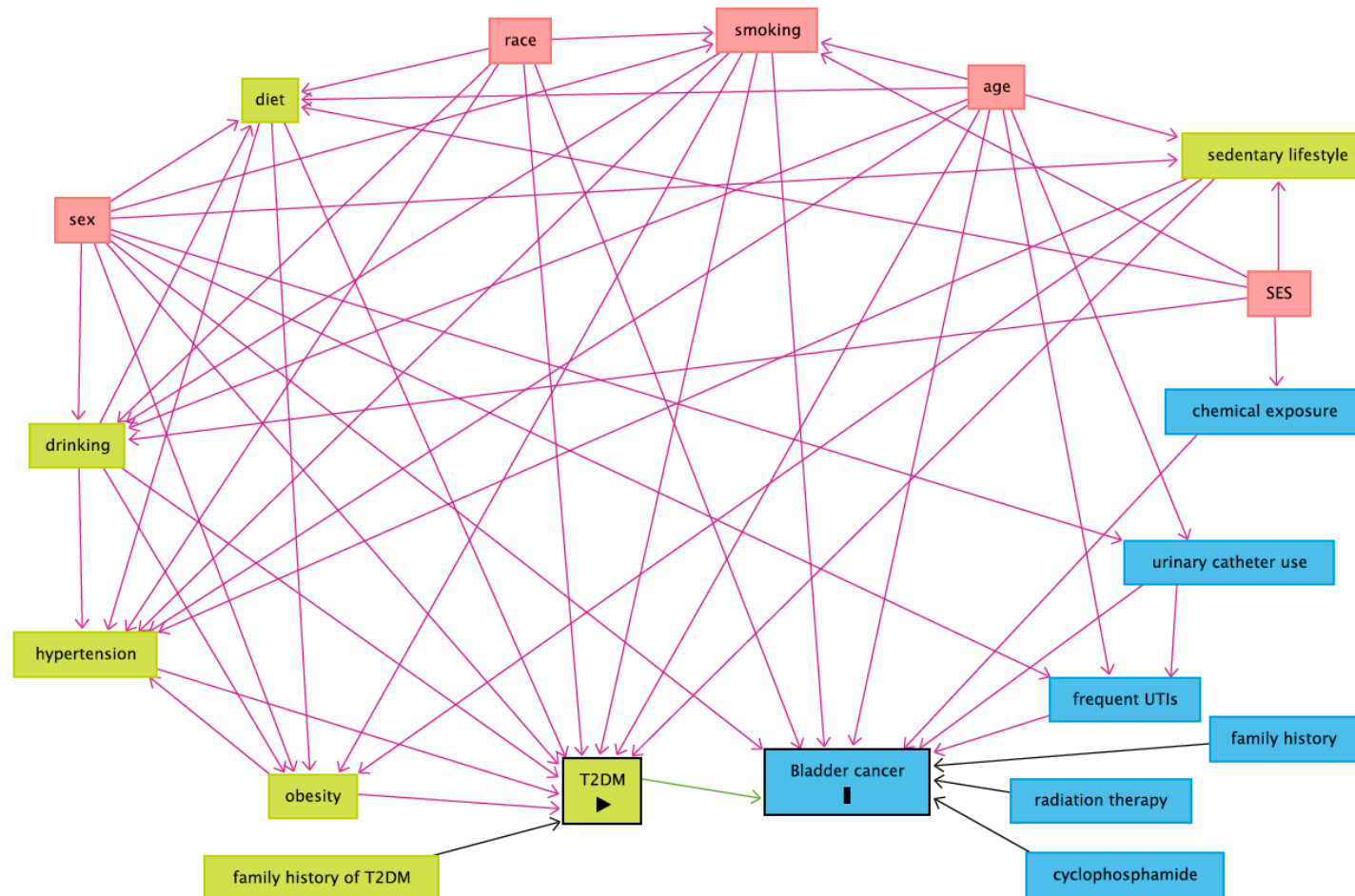


Figure B2-14 DAG of the association between T2DM and risk of bladder cancer, including potential confounders. Chemical exposure: alylamines (industries that produce dyes, rubber, textiles, chemicals or plastics), aromatic amines Polycyclic aromatic hydrocarbons (PAH). Long-term urinary catheter use. Drinking defined as alcohol intake. *Abbreviations.* SES: socioeconomic status; UTIs: urinary tract infections.

Breast cancer

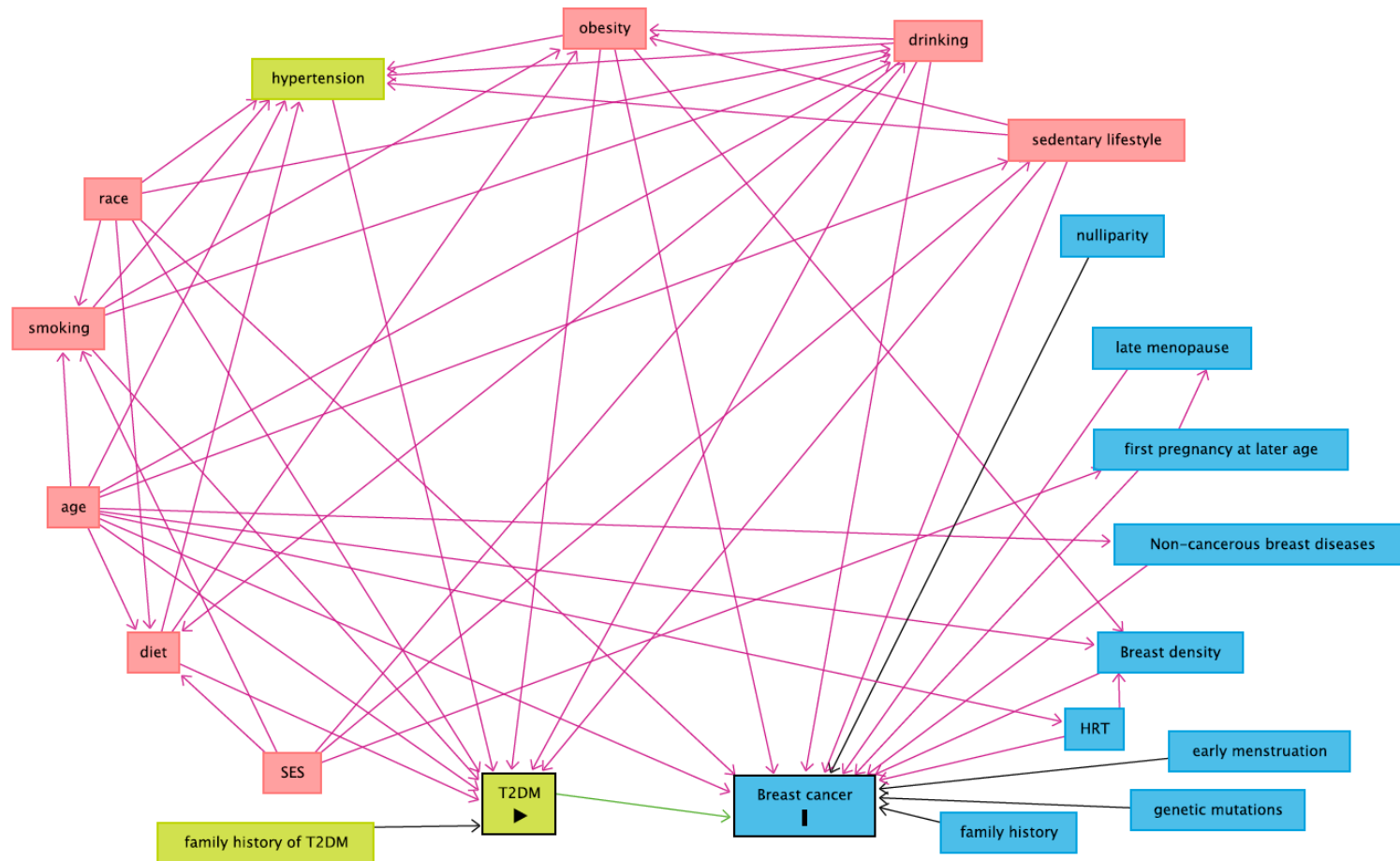


Figure B2-15 DAG of the association between T2DM and risk of breast cancer, including potential confounders. Benign breast disease: non proliferative, proliferative without atypia, proliferative with atypia (atypical hyperplasia). Breast density: dense breast tissue. Family history: breast or ovarian cancer. Non-cancerous breast diseases: DCIS (ductal carcinoma in situ) and LCIS (lobular carcinoma in situ). Drinking defined as alcohol intake. *Abbreviations.* HRT: hormone replacement therapy; SES: socioeconomic status.

Biliary tract cancer (BTC)

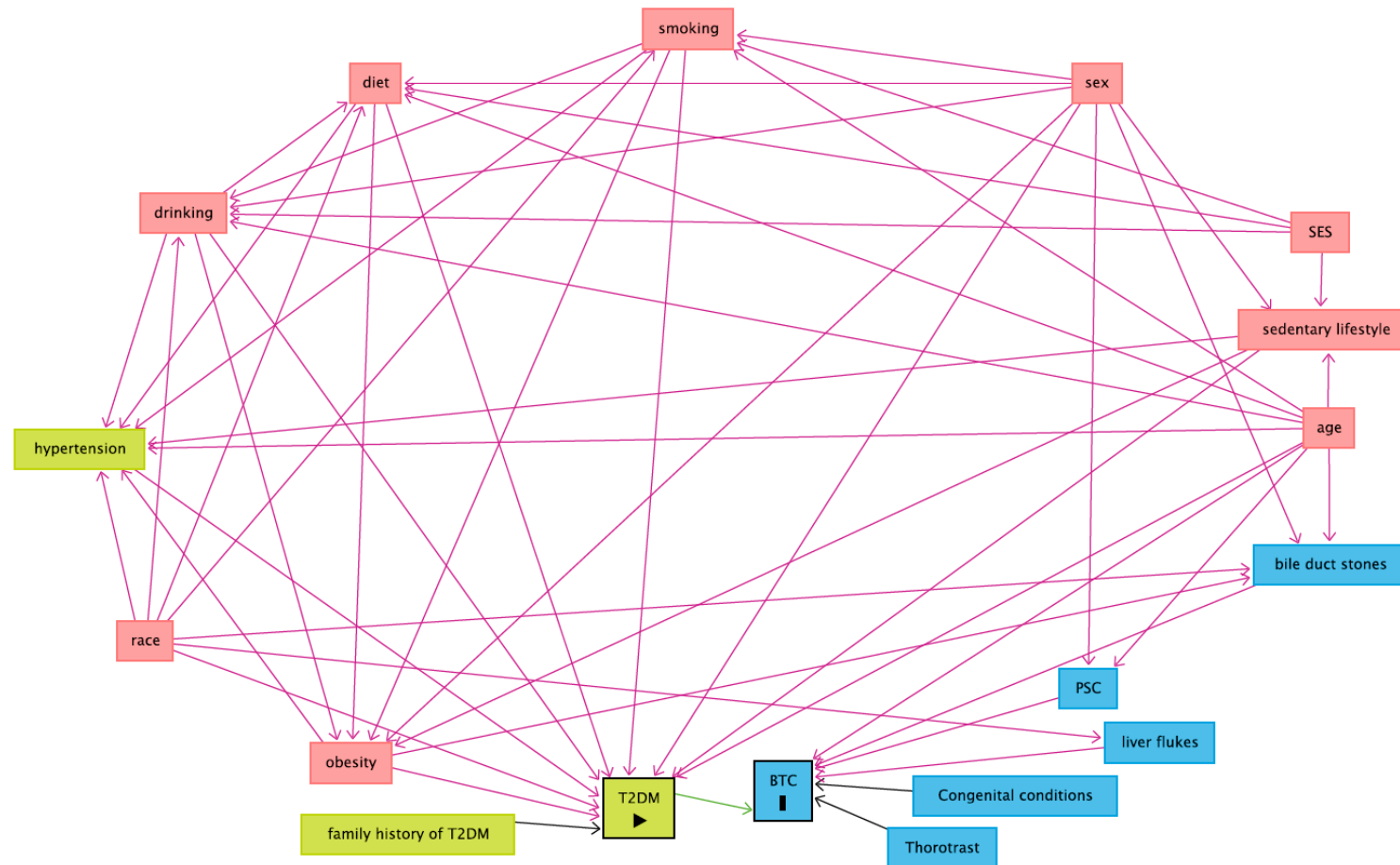


Figure B2-16 DAG of the association between T2DM and risk of BTC, including potential confounders. Congenital conditions: Choledochal cyst and Caroli syndrome. Drinking defined as alcohol intake. *Abbreviations.* IBD: Inflammatory bowel disease. BTC: biliary tract cancer; SES: socioeconomic status; PSC: Primary Sclerosing Cholangitis.

Carcinoma of unknown primary

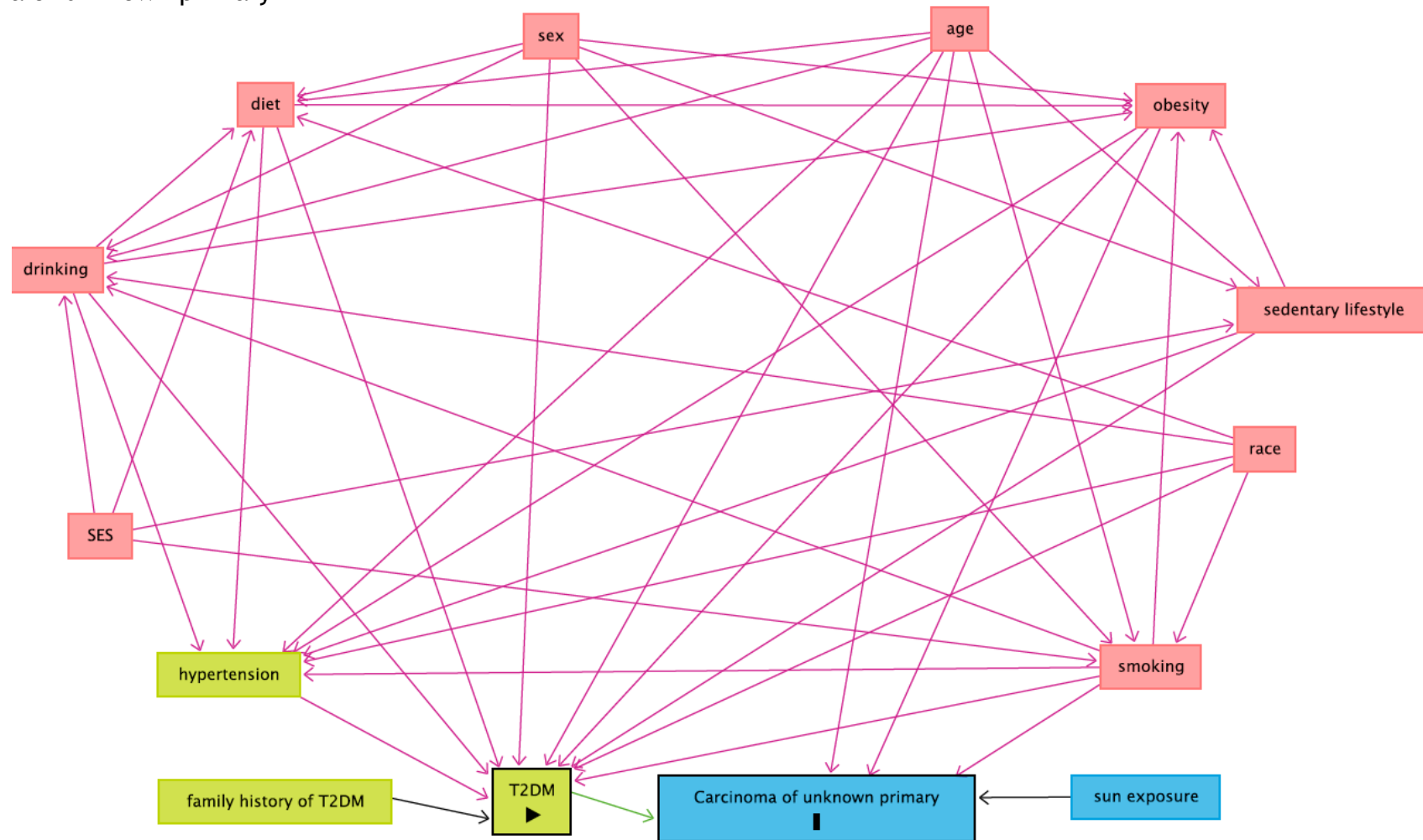


Figure B2-17 DAG of the association between T2DM and risk of carcinoma of unknown primary, including potential confounders. Drinking defined as alcohol intake. SES: socioeconomic status.

Cervical cancer (cervix-uteri)

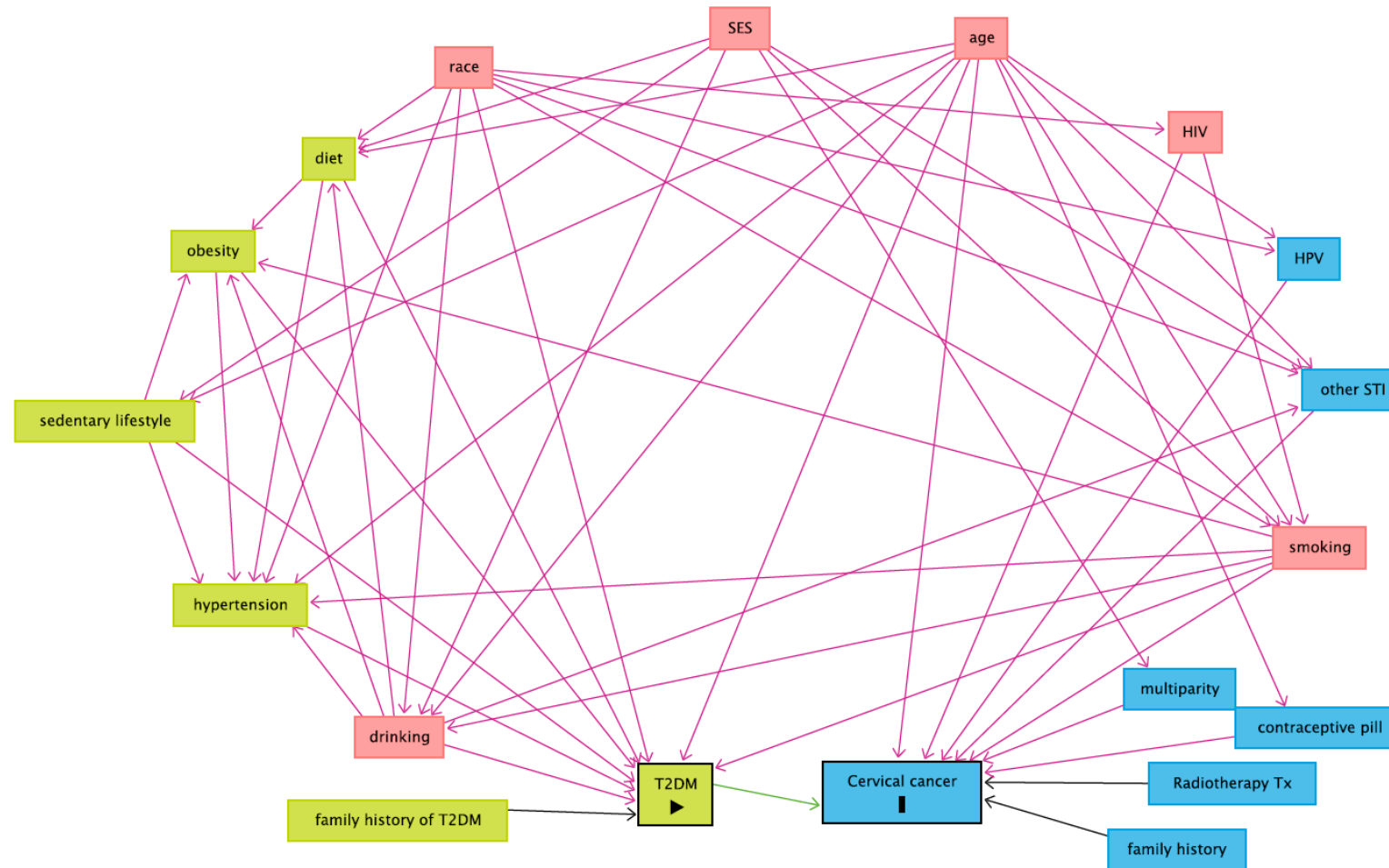


Figure B2-18 DAG of the association between T2DM and risk of cervical cancer, including potential confounders. Contraceptive pill: using birth control pills for a long time (five or more years). Multiparity: three or more children. Drinking defined as alcohol intake. *Abbreviations.* HIV: human immunodeficiency virus; HPV: human papillomavirus; SES: socioeconomic status; STI: sexually transmitted infections; tx: treatment.

Colon cancer

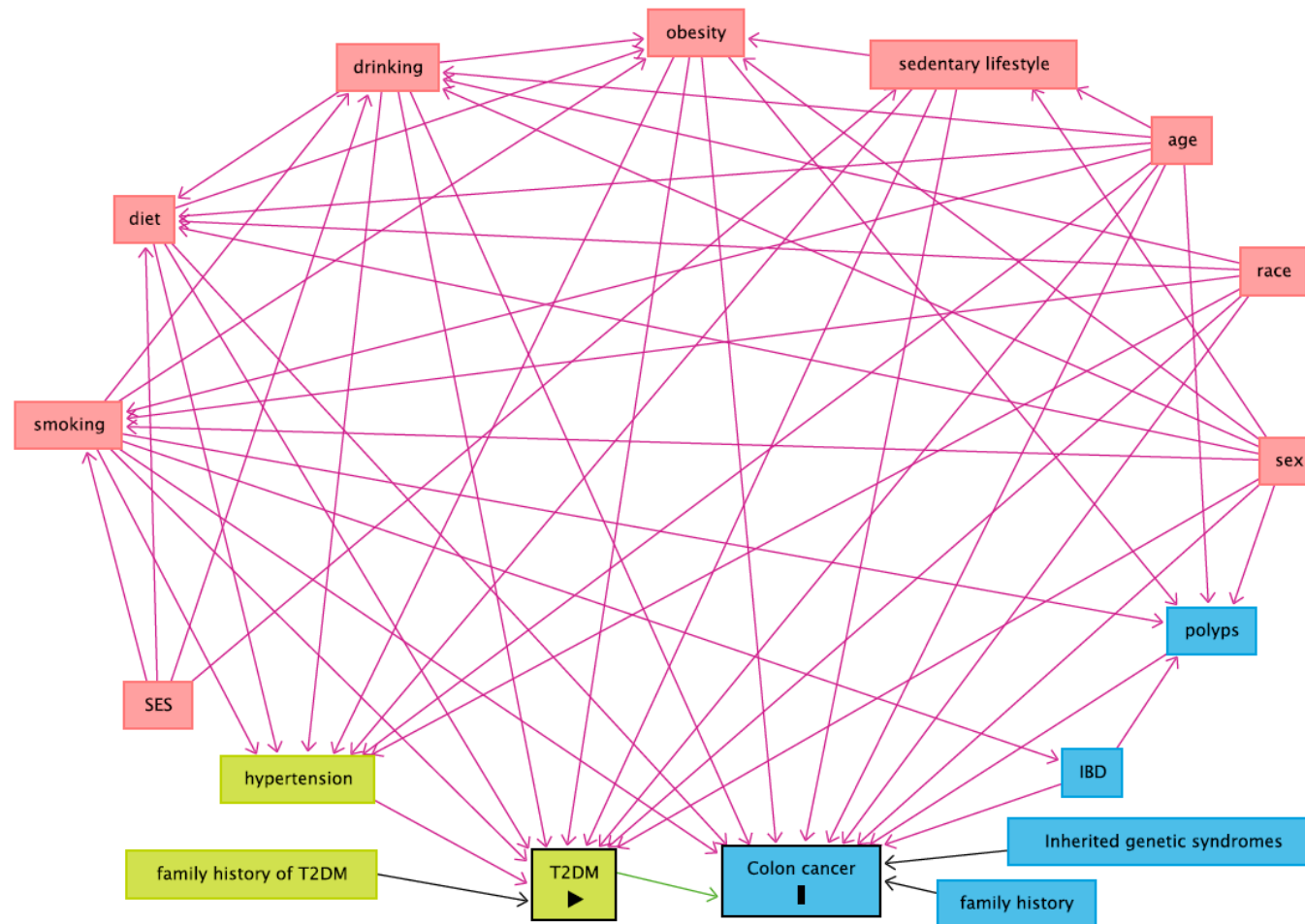


Figure B2-19 DAG of the association between T2DM and risk of colon cancer, including potential confounders. Inherited Genetic Syndromes: Lynch syndrome and familial adenomatous polyposis. Drinking defined as alcohol intake. *Abbreviations.* IBD: inflammatory bowel disease; SES: socioeconomic status.

Colorectal cancer (CRC)

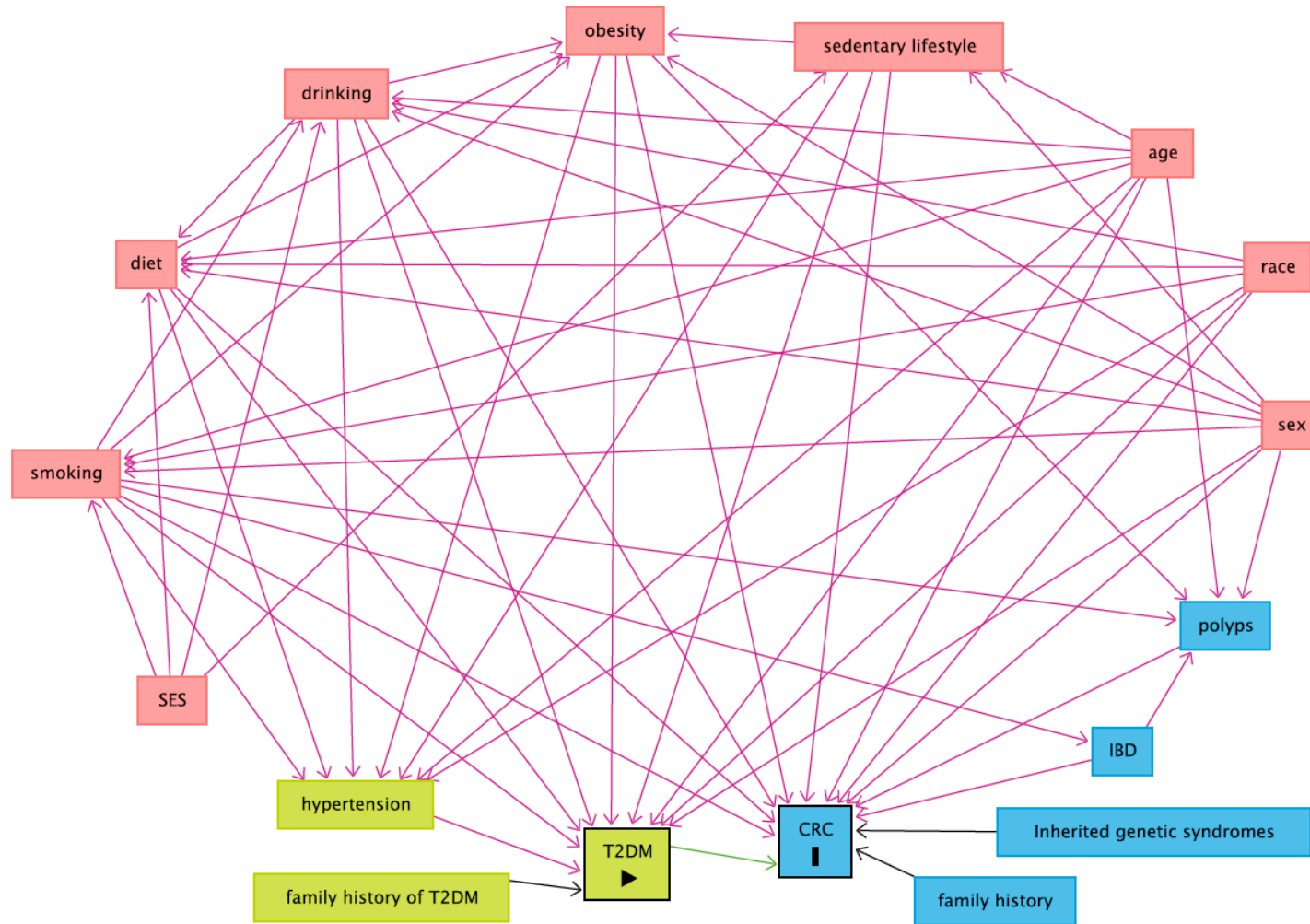


Figure B2-20 DAG of the association between T2DM and risk of CRC, including potential confounders. Inherited Genetic Syndromes: Lynch syndrome and familial adenomatous polyposis. Drinking defined as alcohol intake. *Abbreviations.* CRC: colorectal cancer; IBD: inflammatory bowel disease; SES: socioeconomic status.

Endometrial cancer

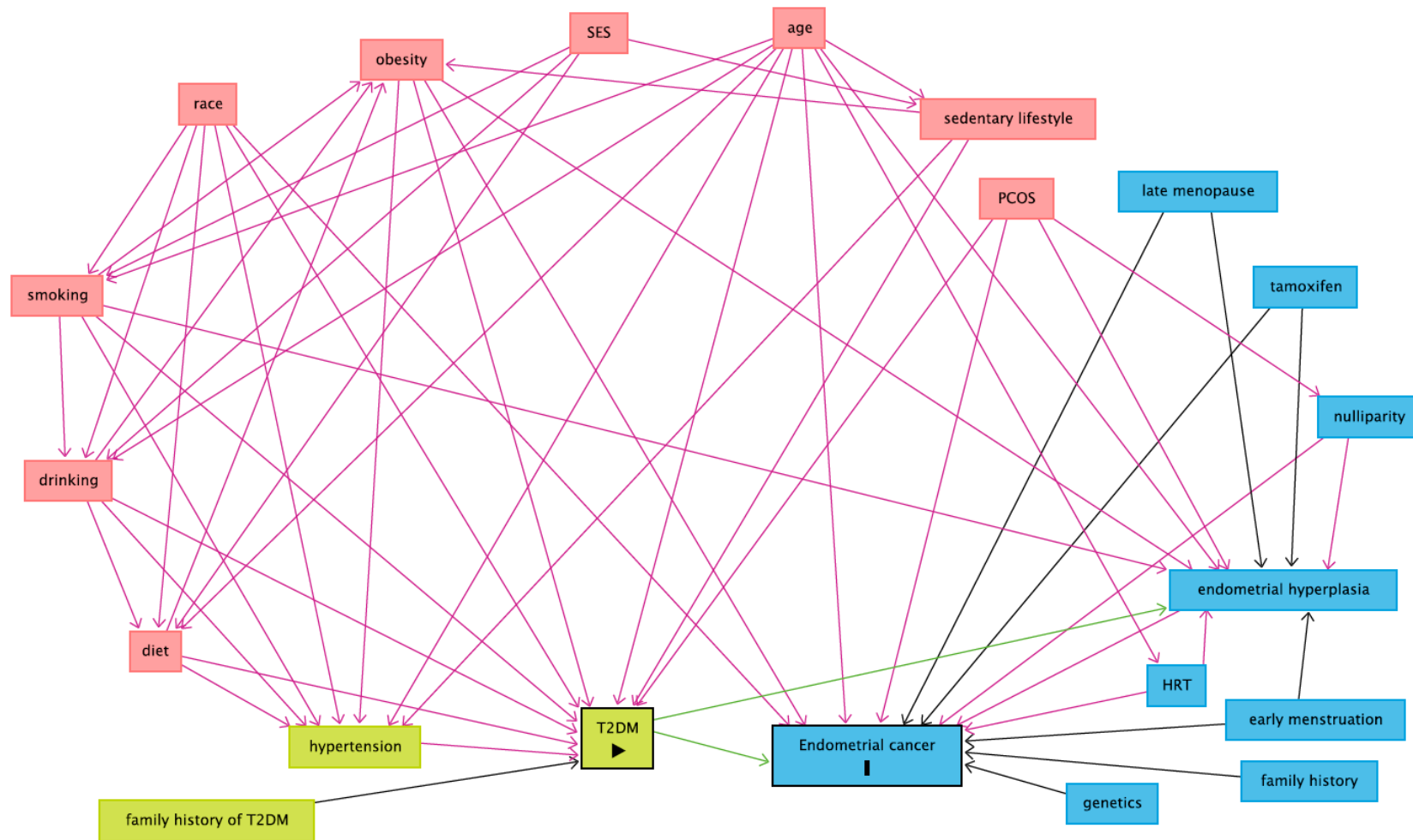


Figure B2-21 DAG of the association between T2DM and risk of endometrial cancer, including potential confounders. Genetics: Lynch syndrome. Drinking defined as alcohol intake. *Abbreviations.* HRT: hormone replacement therapy (oestrogen-only therapy); PCOS: polycystic ovarian syndrome; SES: socioeconomic status.

Esophageal cancer

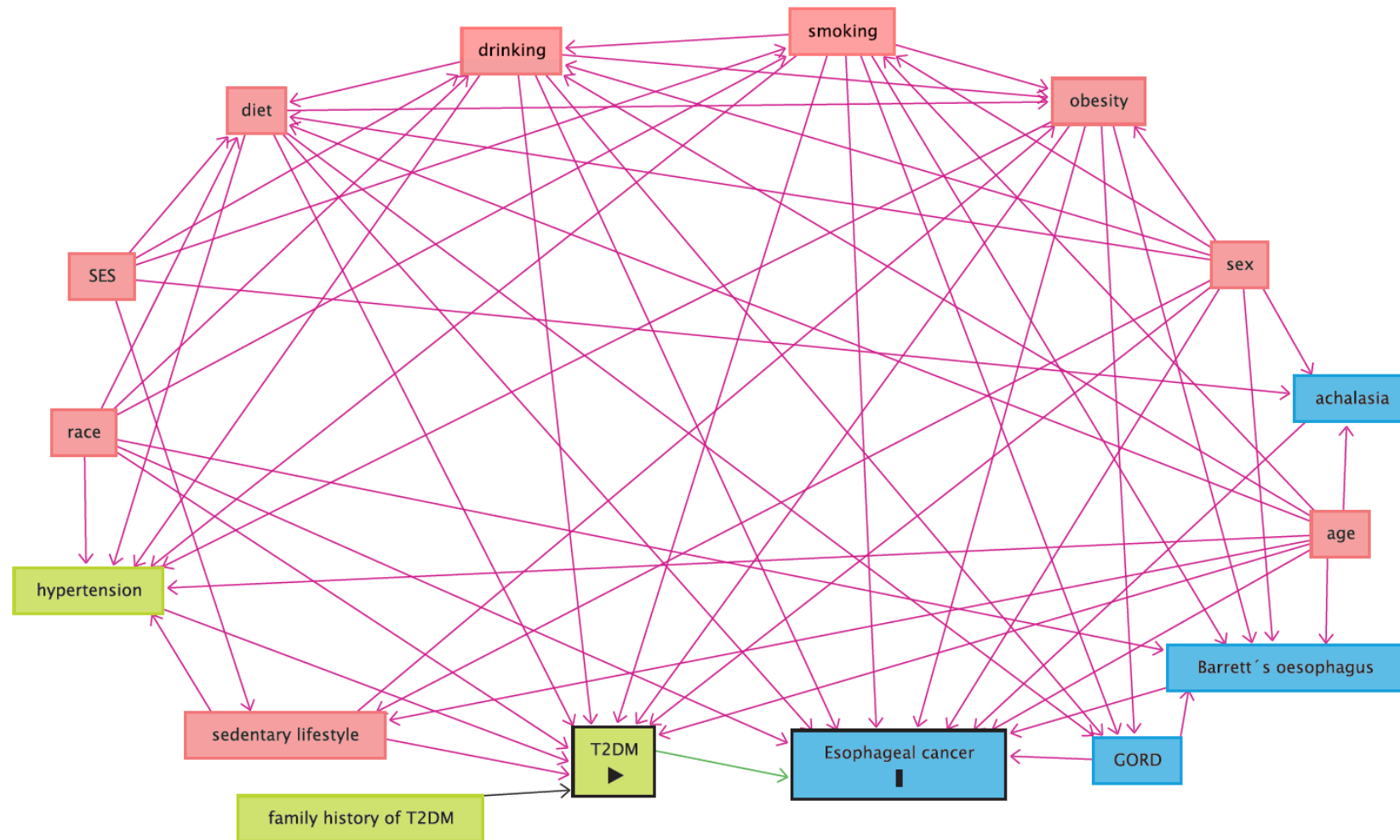


Figure B2-22 DAG of the association between T2DM and risk of esophageal cancer, including potential confounders. Drinking defined as alcohol intake.
Abbreviations. GORD/GERD: Gastroesophageal reflux disease; SES: socioeconomic status.

Esophagogastric junction adenocarcinoma

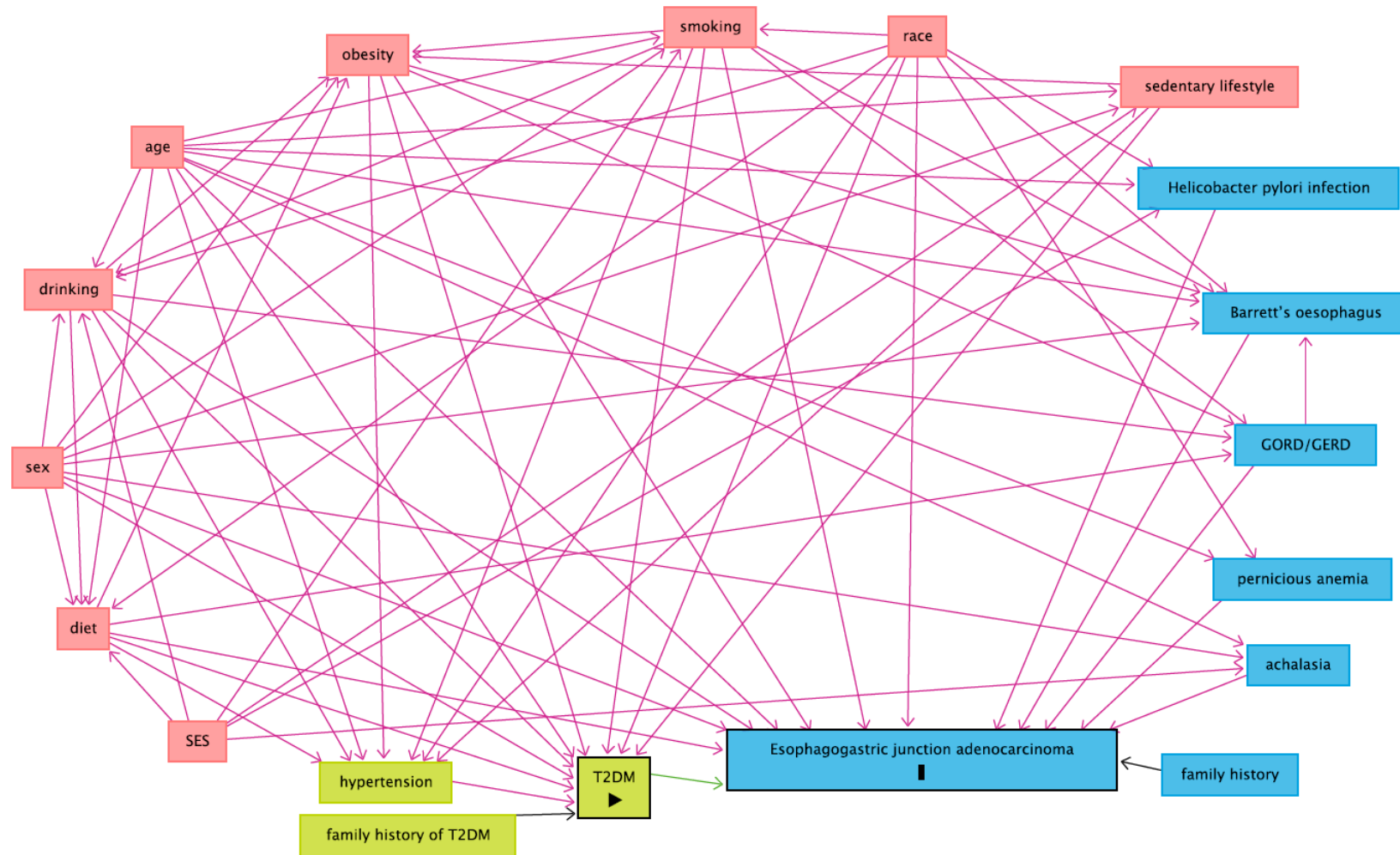


Figure B2-23 DAG of the association between T2DM and risk of esophagogastric junction adenocarcinoma, including potential confounders. Drinking defined as alcohol intake. *Abbreviations.* GORD/GERD: Gastroesophageal reflux disease; SES: socioeconomic status.

Extrahepatic cholangiocarcinoma

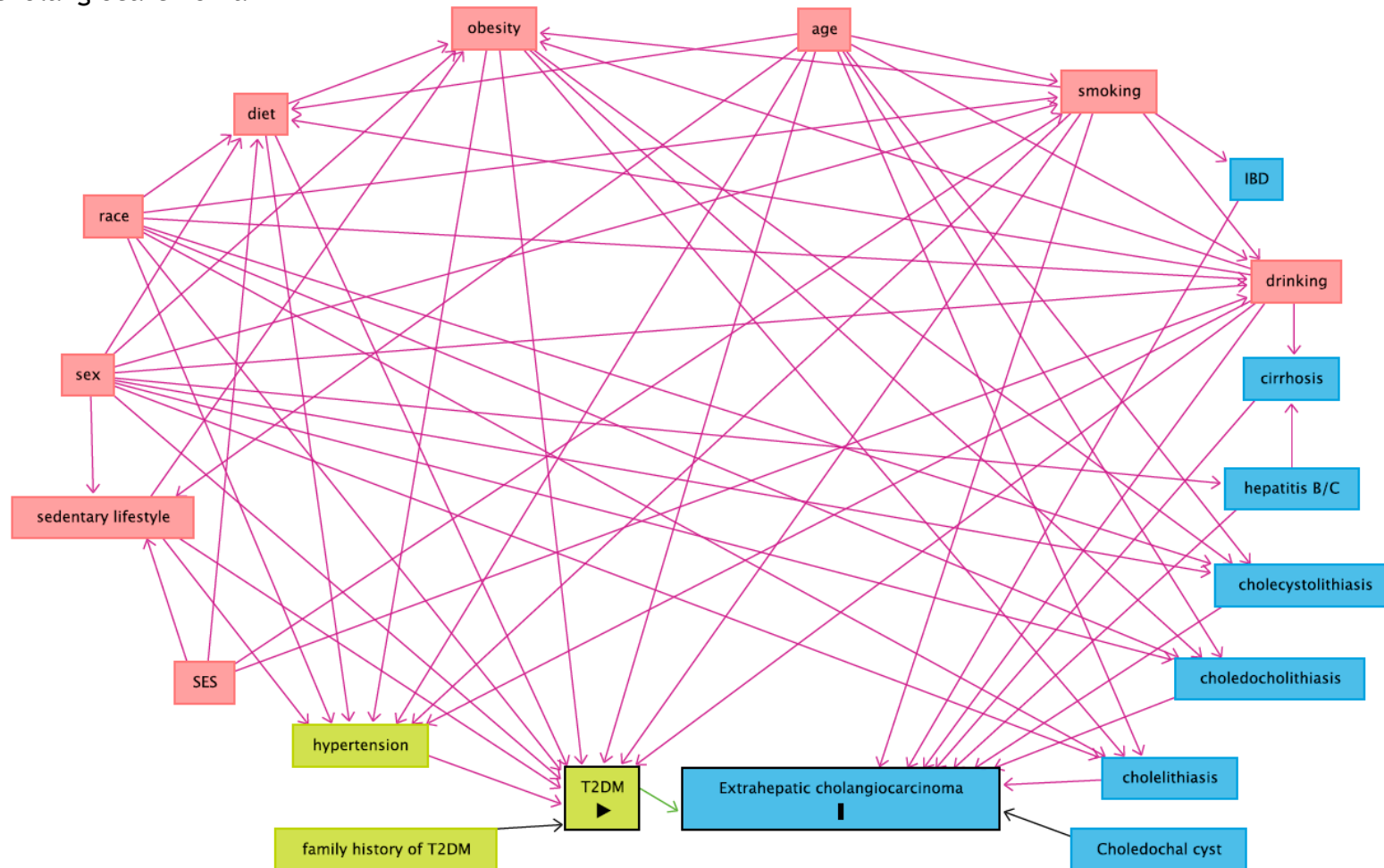


Figure B2-24 DAG of the association between T2DM and risk of extrahepatic cholangiocarcinoma, including potential confounders. Drinking defined as alcohol intake. *Abbreviations.* IBD: inflammatory bowel disease; SES: socioeconomic status.

Eye cancer

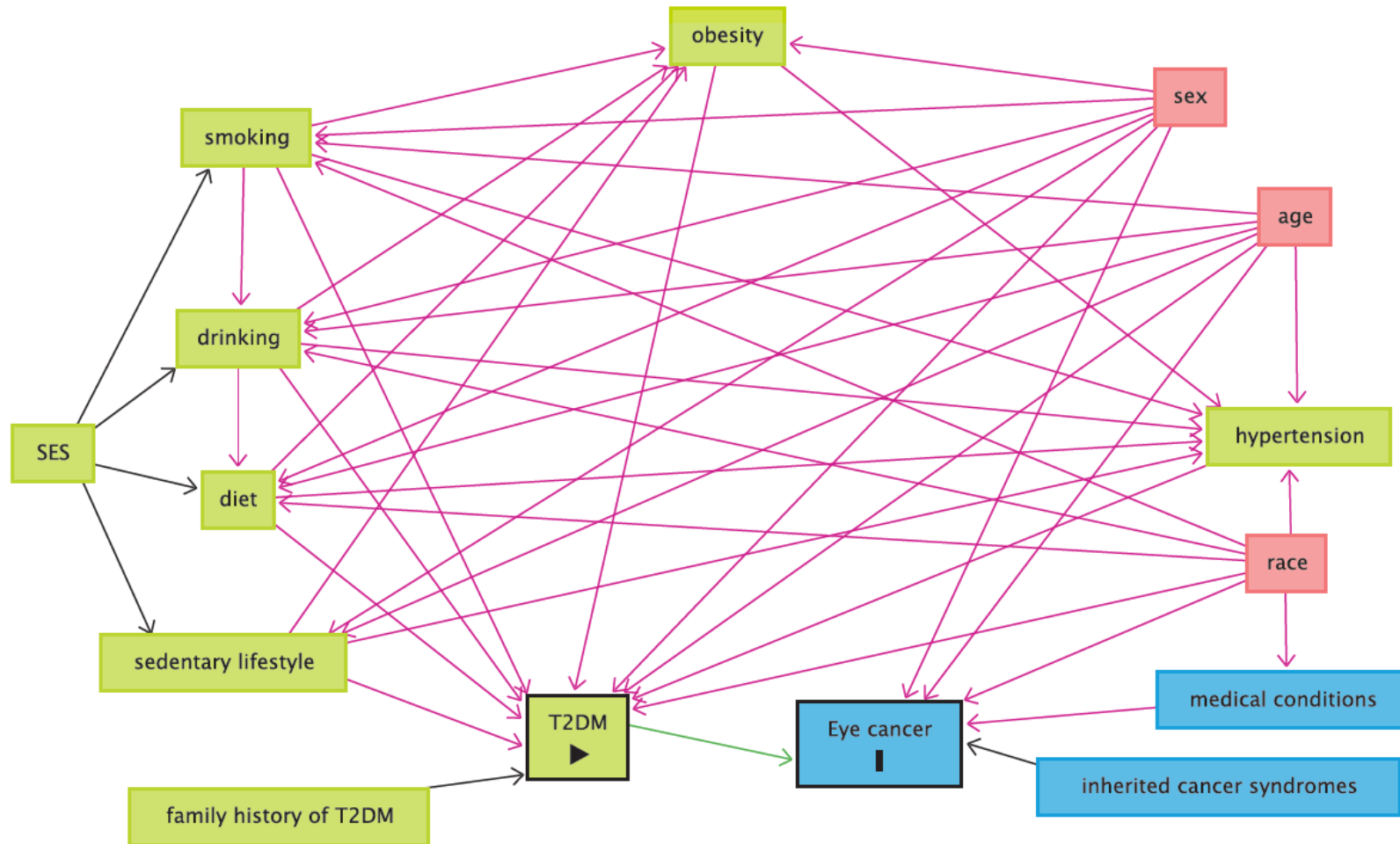


Figure B2-25 DAG of the association between T2DM and risk of eye cancer, including potential confounders. Inherited cancer syndromes: mutation of BAP1, Dysplastic nevus syndrome. Medical conditions: nevus of Ota, Nevus. Drinking defined as alcohol intake. SES: socioeconomic status.

Gallbladder cancer

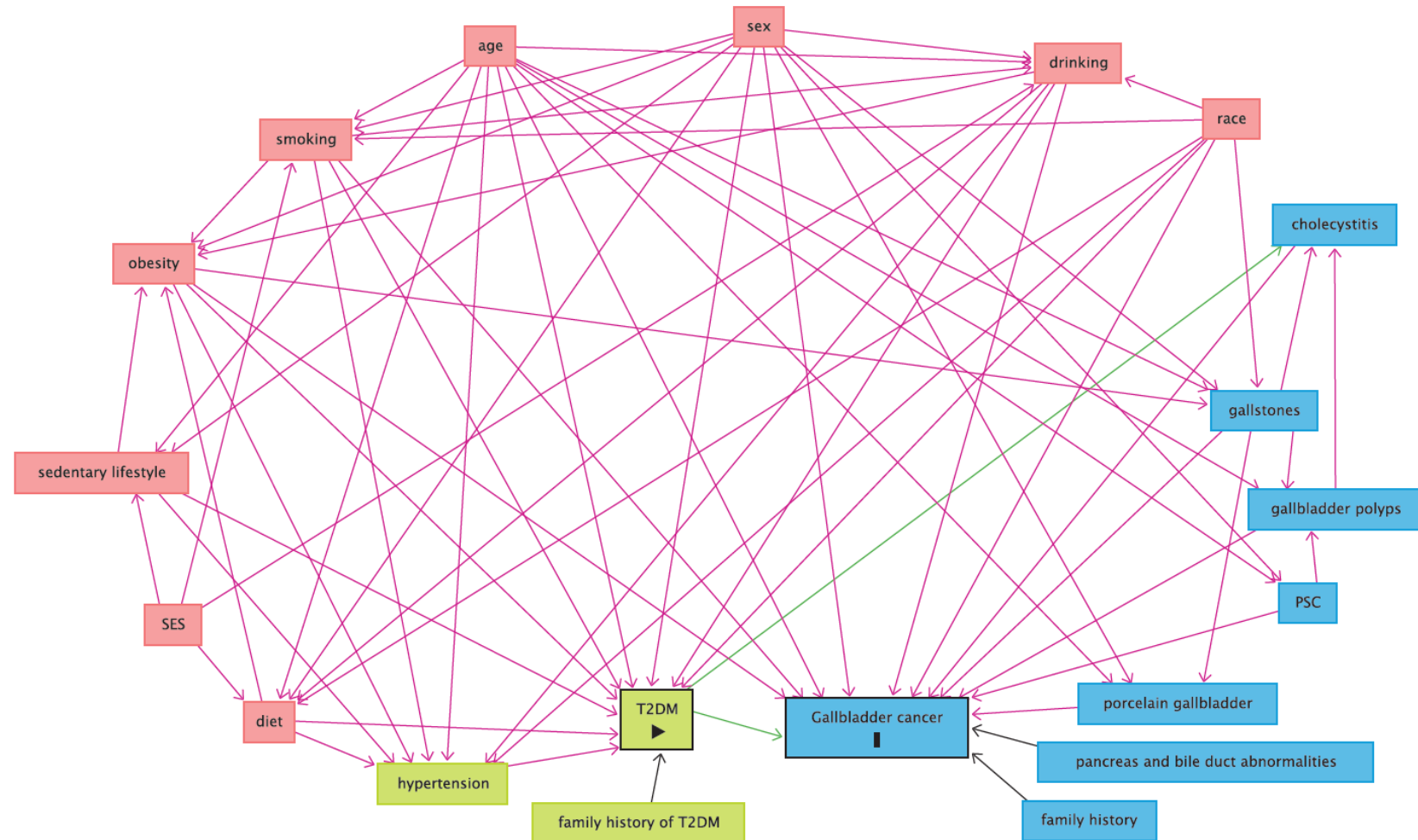


Figure B2-26 DAG of the association between T2DM and risk of gallbladder cancer, including potential confounders. Pancreas and bile duct abnormalities born with choledochal cysts, abnormality of the join between the bile duct and the pancreas. Drinking defined as alcohol intake. *Abbreviations.* PSC: Primary sclerosing cholangitis; SES: socioeconomic status.

Gastrointestinal (GI) cancer

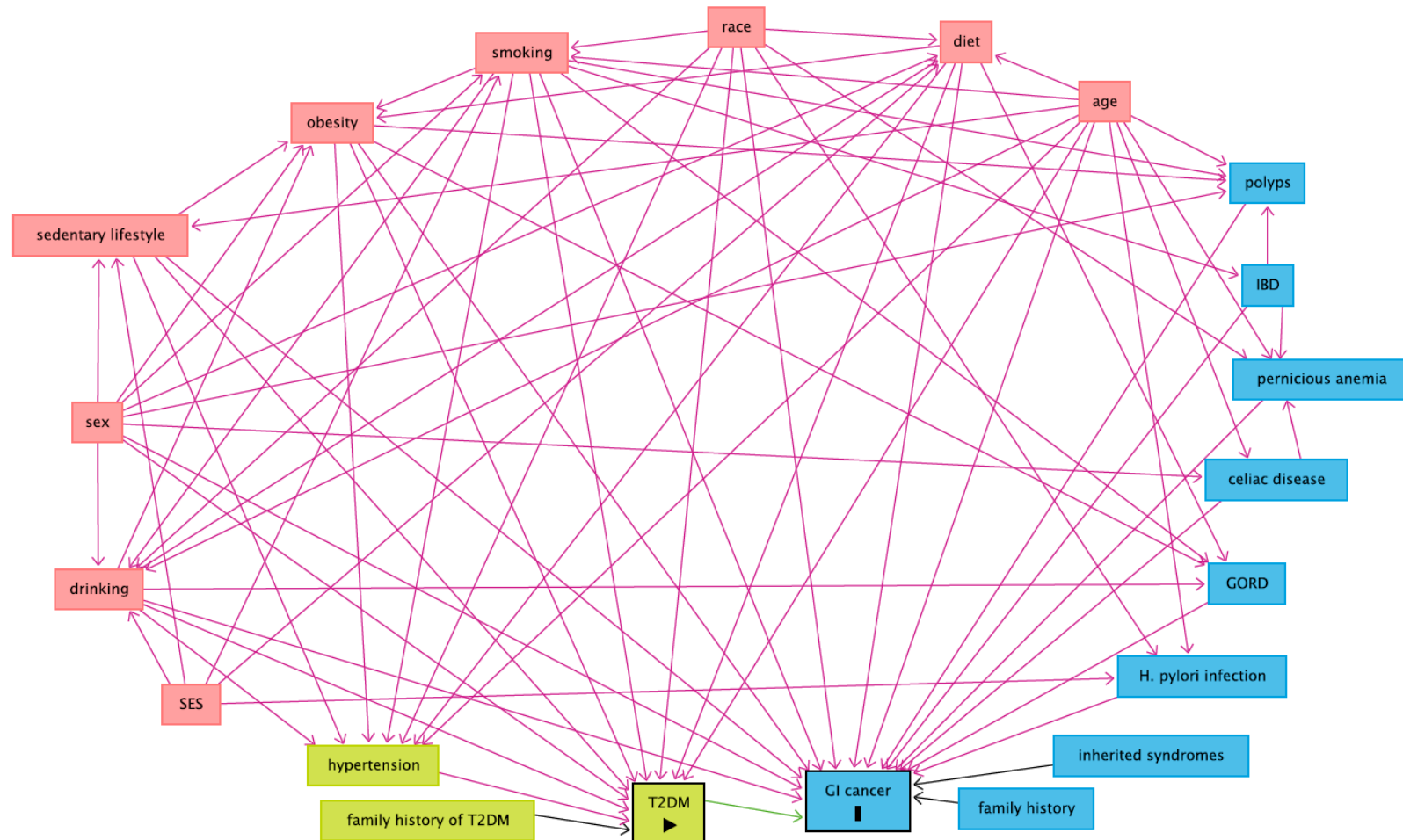


Figure B2-27 DAG of the association between T2DM and risk of gastrointestinal (GI) cancer, including potential confounders. Inherited syndromes: Familial adenomatous polyposis (FAP)- MUTYH-associated polyposis, Lynch syndrome, Peutz-Jeghers syndrome, Cystic fibrosis. Drinking defined as alcohol intake. *Abbreviations.* GORD/GERD: Gastroesophageal reflux disease. IBD: inflammatory bowel disease (Crohn's disease); SES: socioeconomic status.

Head and neck cancer

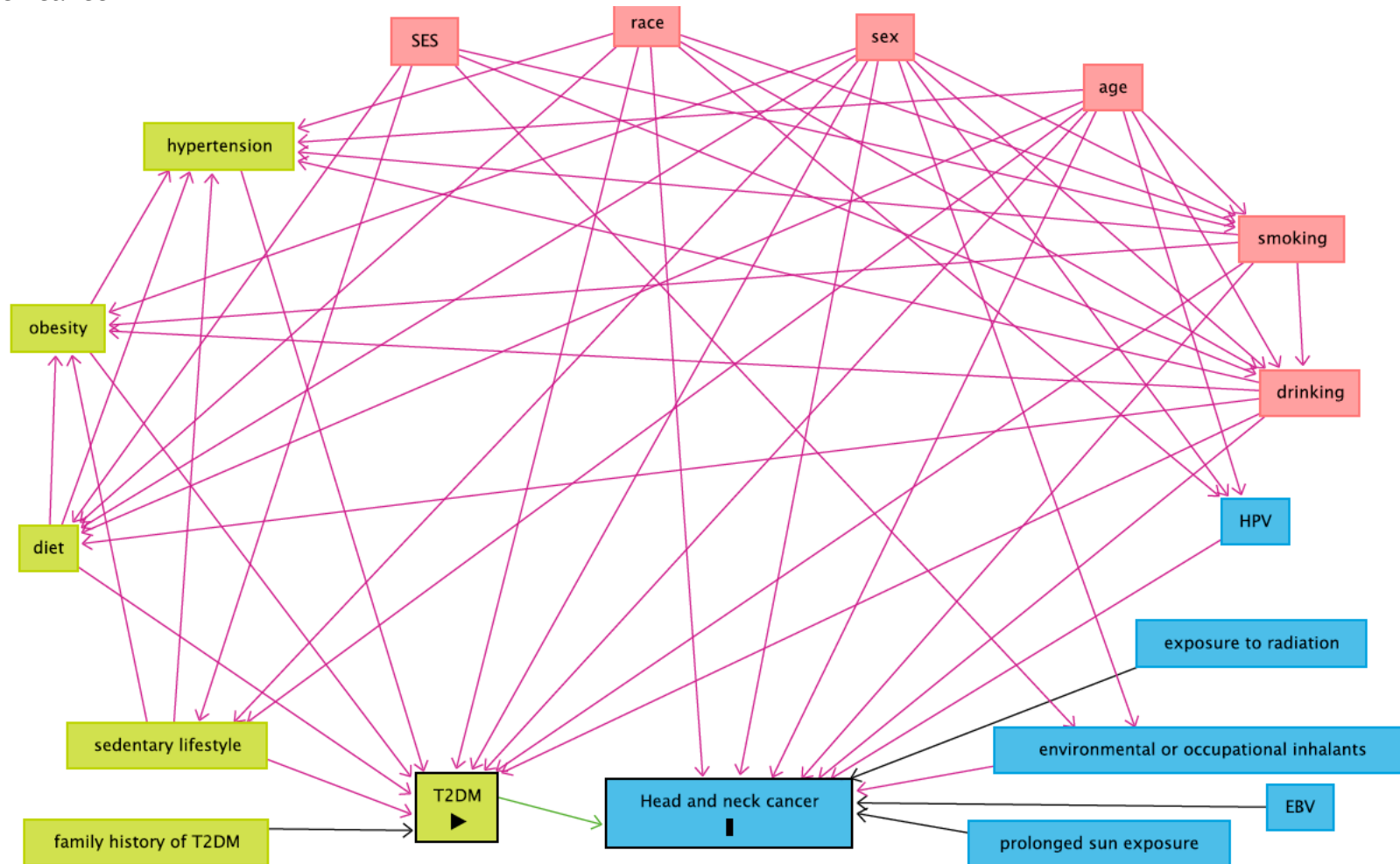


Figure B2-28 DAG of the association between T2DM and risk of head and neck cancer, including potential confounders. Environmental or occupational inhalants: asbestos, wood dust, paint fumes, and certain chemicals-nickel dust, formaldehyde. Exposure to radiation: X-rays, medical radiation. Drinking defined as alcohol intake. *Abbreviations.* EBV: Epstein-Barr virus; HPV: human papillomavirus; SES: socioeconomic status.

Hepatocellular carcinoma (HCC)

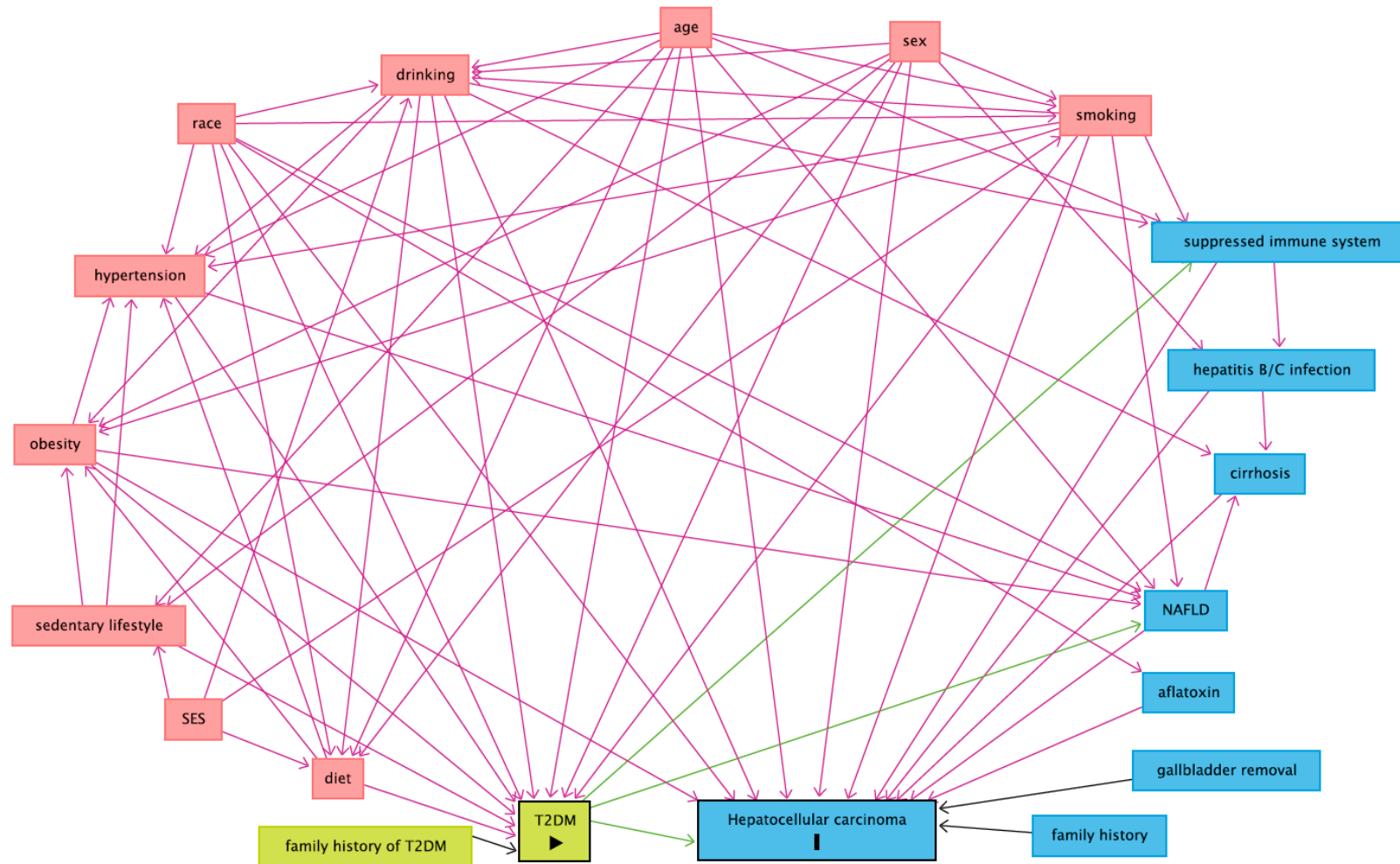


Figure B2-29 DAG of the association between T2DM and risk of HCC, including potential confounders. Drinking defined as alcohol intake. *Abbreviations.* HCC: hepatocellular carcinoma; NAFLD: non-alcoholic fatty liver disease; SES: socioeconomic status.

Kidney cancer

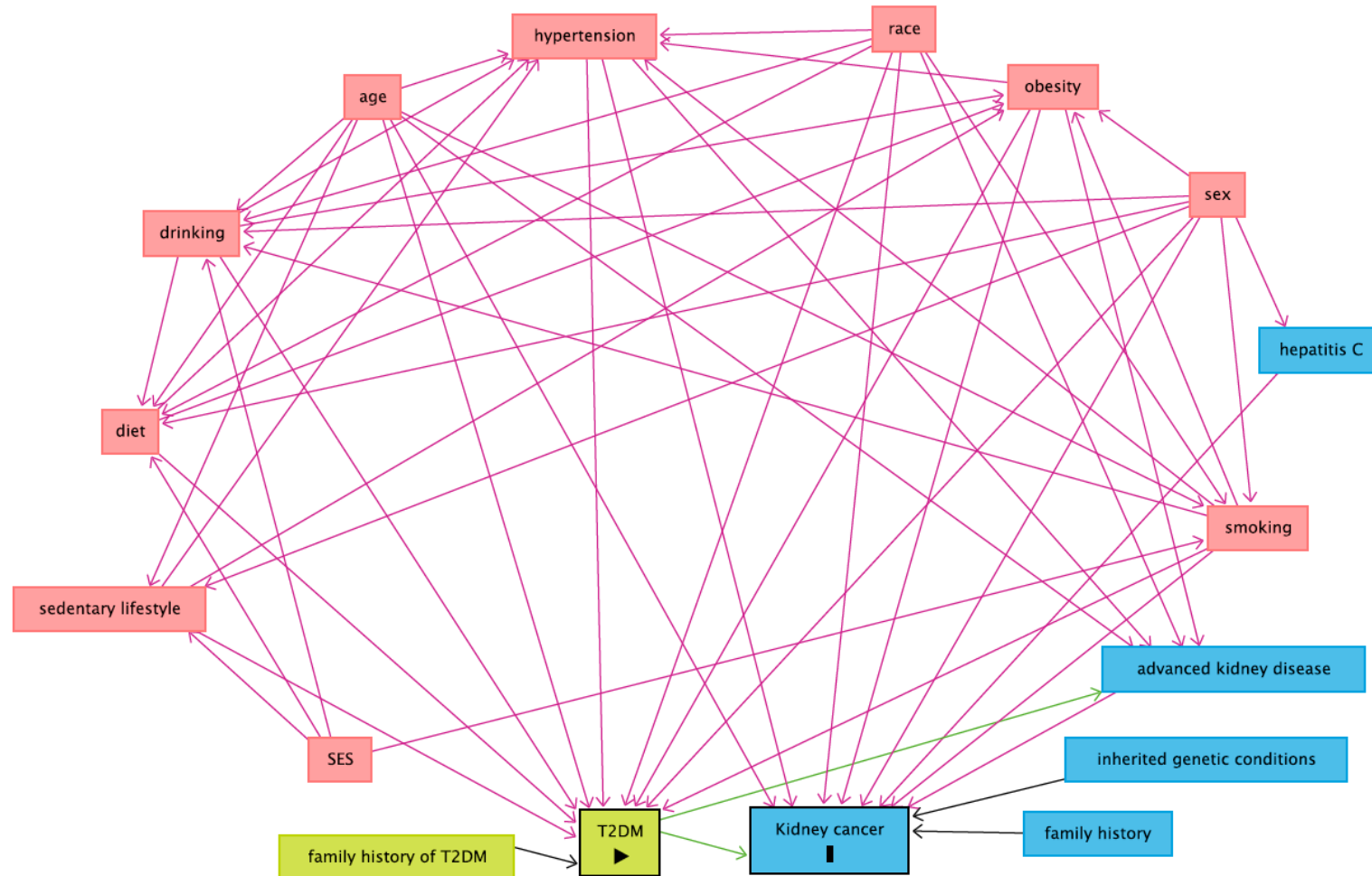


Figure B2-31 DAG of the association between T2DM and risk of kidney cancer, including potential confounders. Advanced kidney disease: especially those in dialysis. Drinking defined as alcohol intake. SES: socioeconomic status.

Large intestine cancer

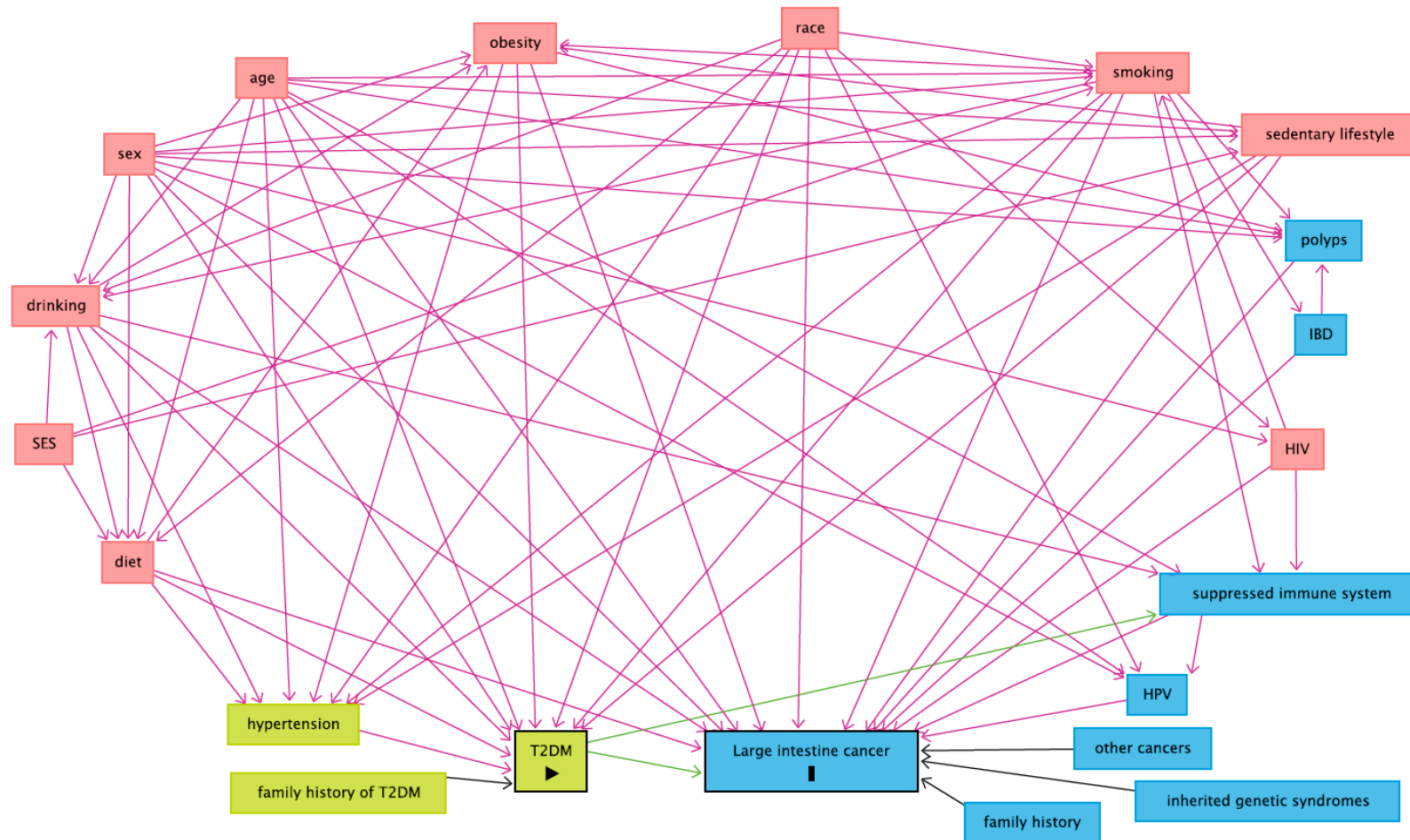


Figure B2-32 DAG of the association between T2DM and risk of large intestine cancer, including potential confounders. Inherited Genetic Syndromes: Lynch syndrome and familial adenomatous polyposis. Other cancers: cervix, vagina, or vulva. Drinking defined as alcohol intake. *Abbreviations.* HIV: human immunodeficiency virus; HPV: human papillomavirus; IBD: inflammatory bowel disease; SES: socioeconomic status.

Larynx cancer

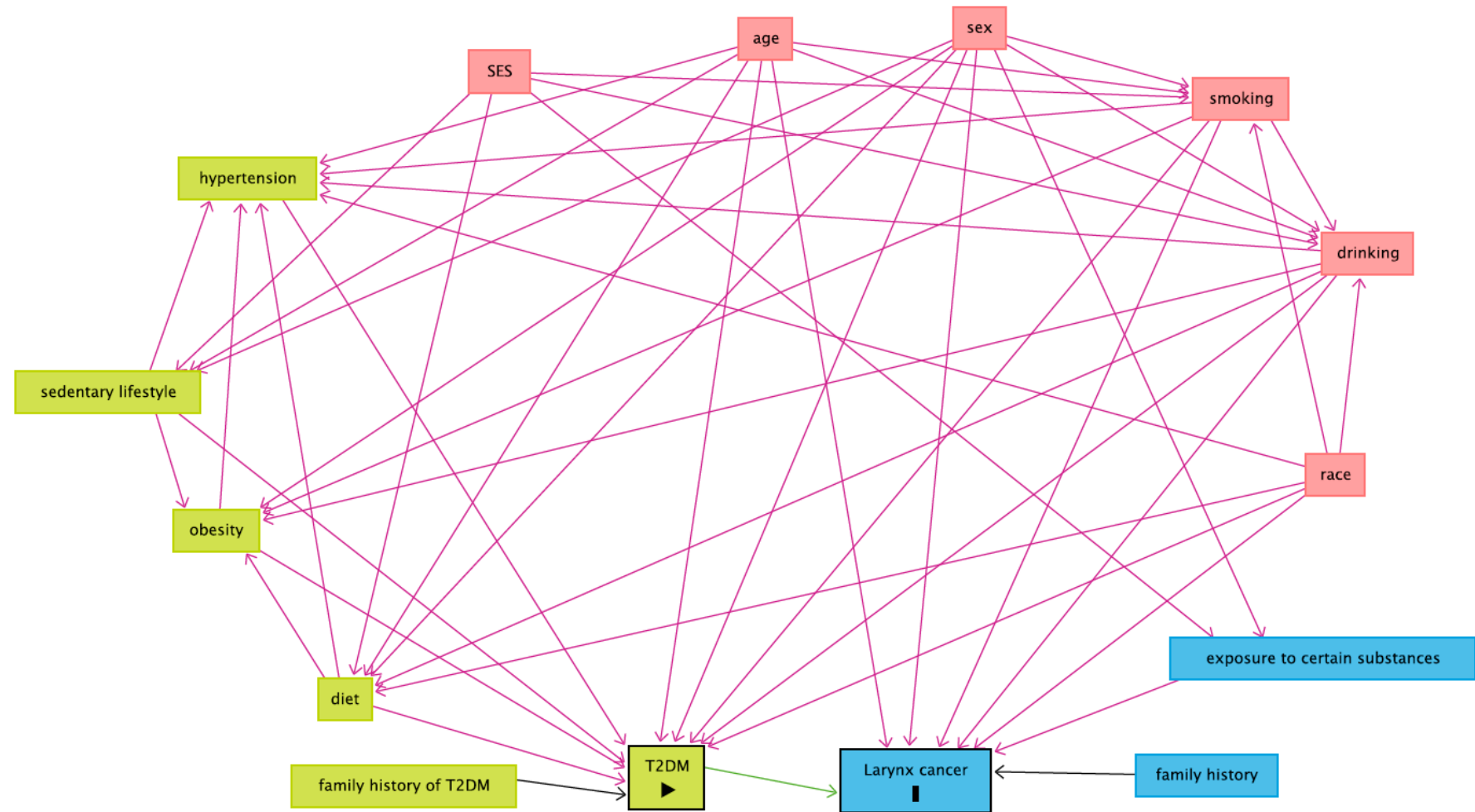


Figure B2-33 DAG of the association between T2DM and risk of larynx cancer, including potential confounders. Exposure to certain substances: asbestos, sulphuric acid. Drinking defined as alcohol intake. SES: socioeconomic status.

Lip cancer

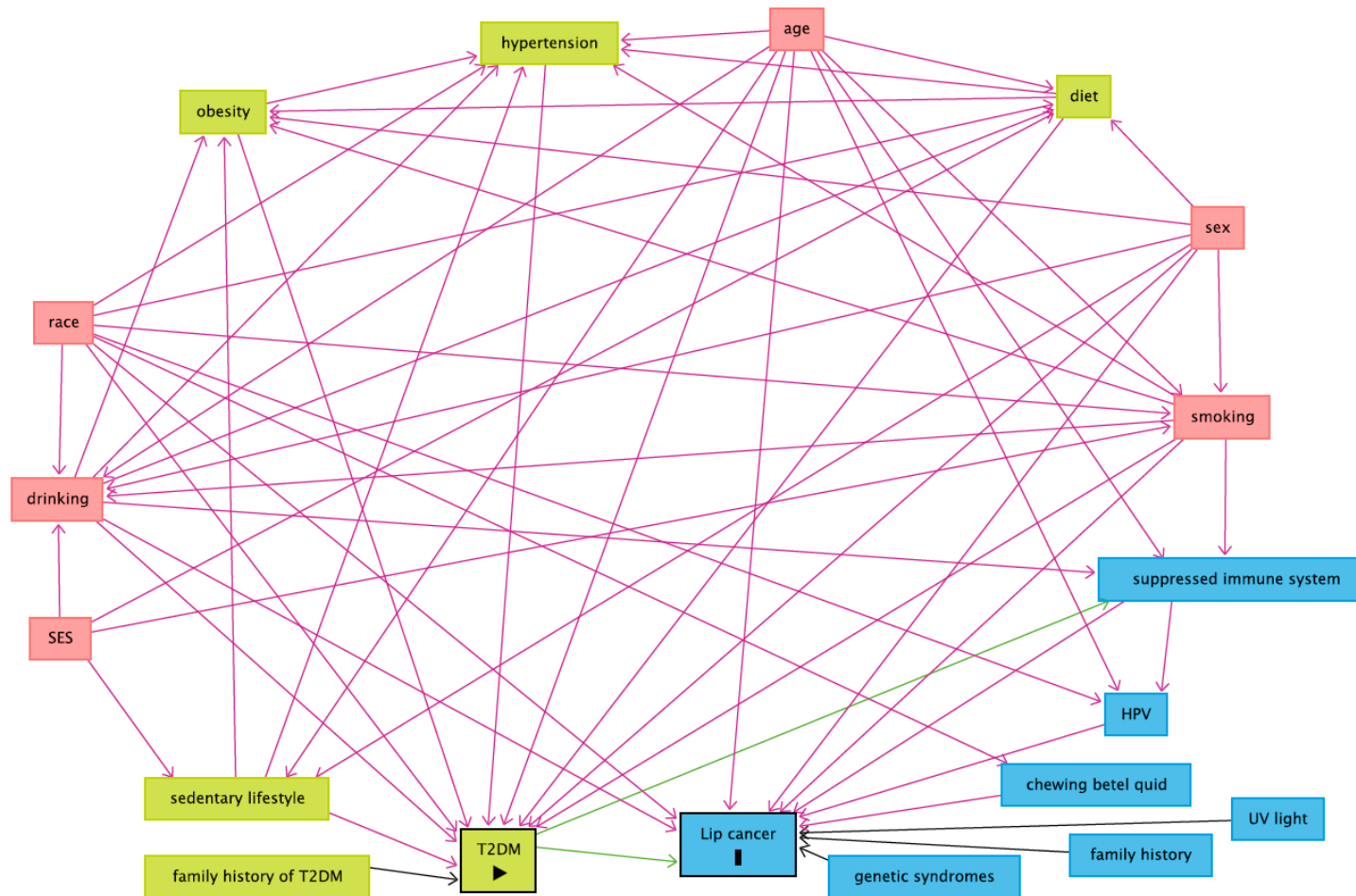


Figure B2-34 DAG of the association between T2DM and risk of lip cancer, including potential confounders. Genetic syndromes: Fanconi anaemia, Dyskeratosis congenita. Drinking defined as alcohol intake. *Abbreviations.* HPV: human papillomavirus; SES: socioeconomic status.

Liver cancer

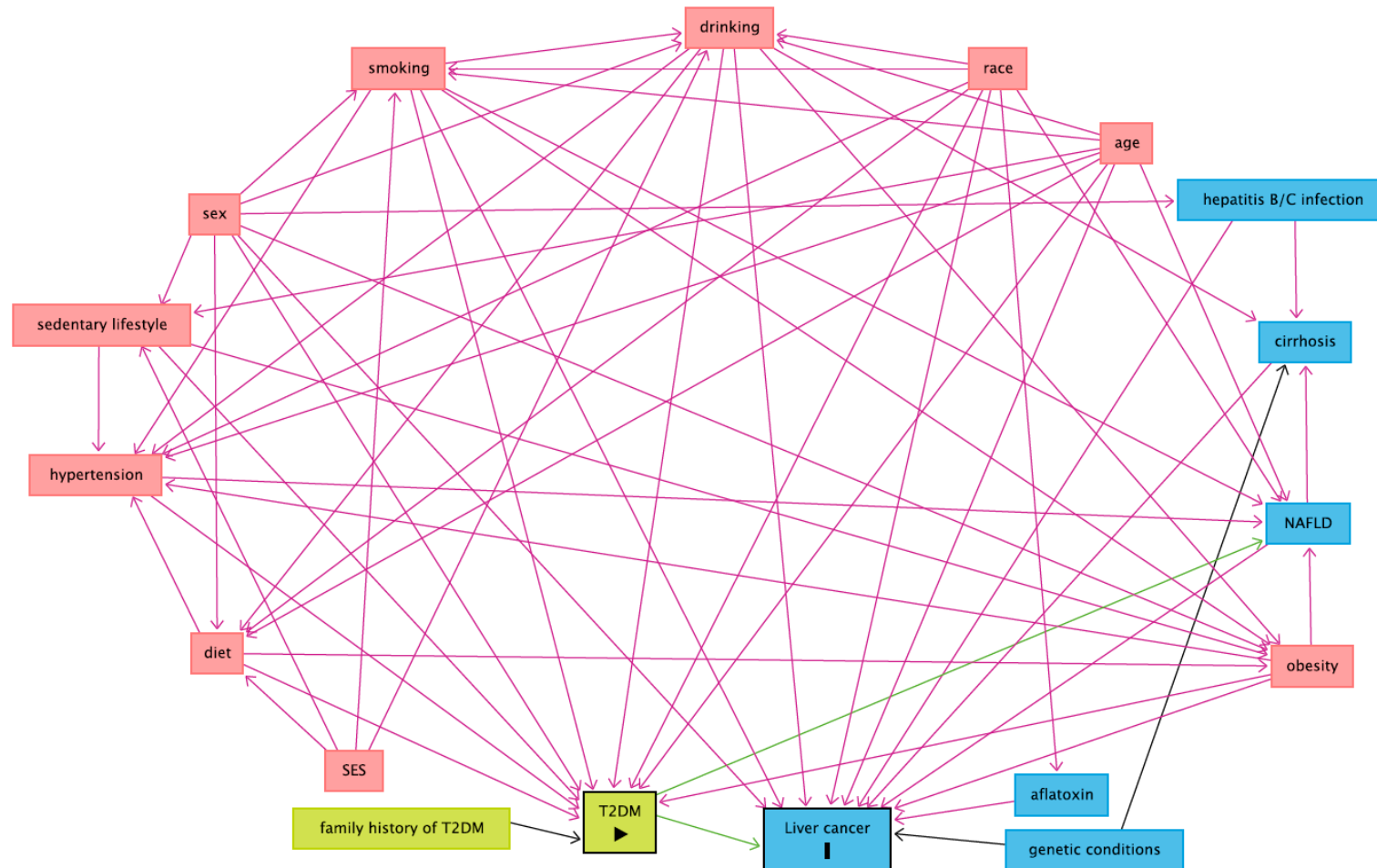


Figure B2-35 DAG of the association between T2DM and risk of liver cancer, including potential confounders. Genetic conditions: haemochromatosis, alpha 1 antitrypsin deficiency. Drinking defined as alcohol intake. *Abbreviations.* NAFLD: non-alcoholic fatty liver disease; SES: socioeconomic status.

Lung cancer

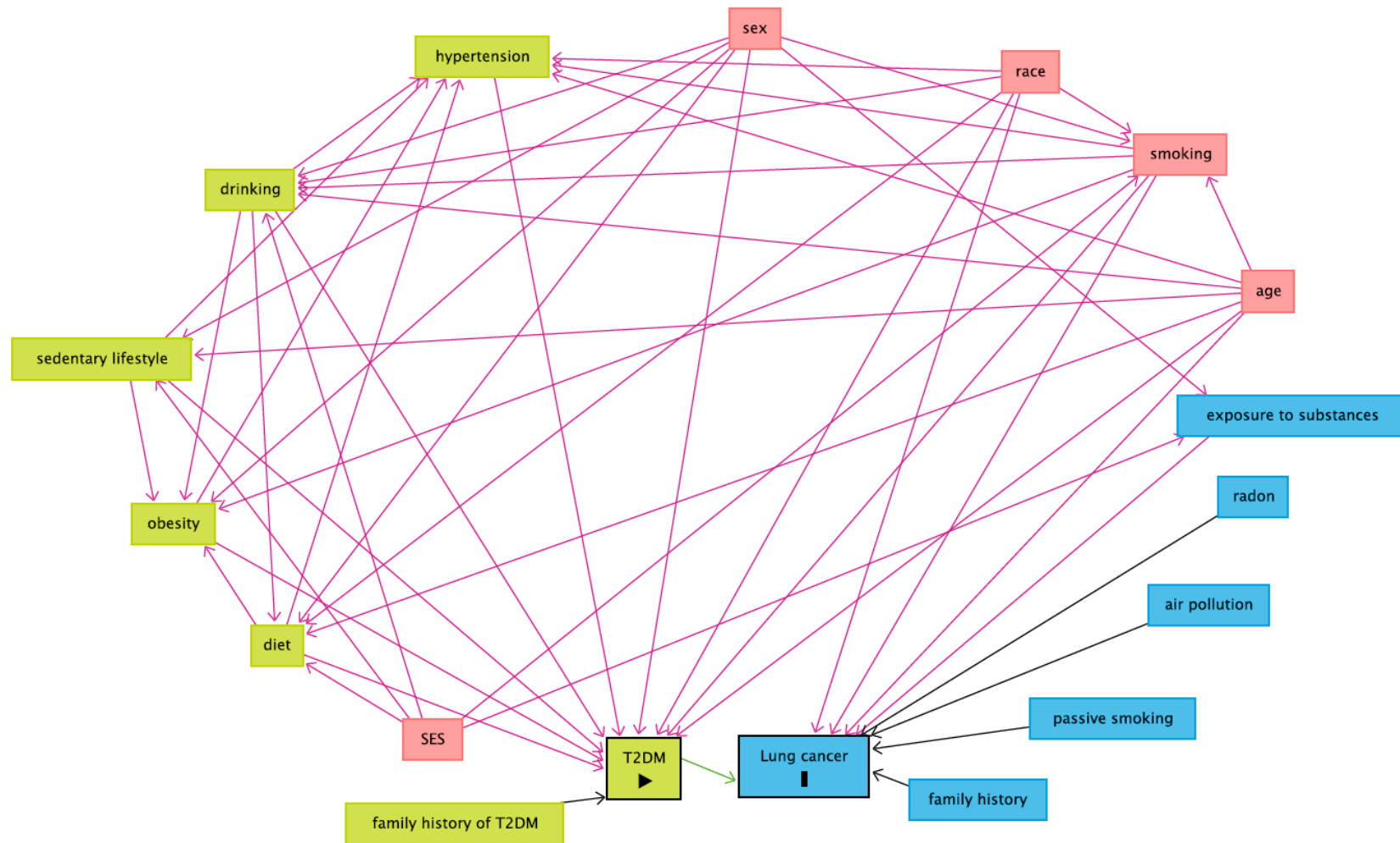


Figure B2-36 DAG of the association between T2DM and risk of lung cancer, including potential confounders. Drinking defined as alcohol intake. SES: socioeconomic status.

Melanoma

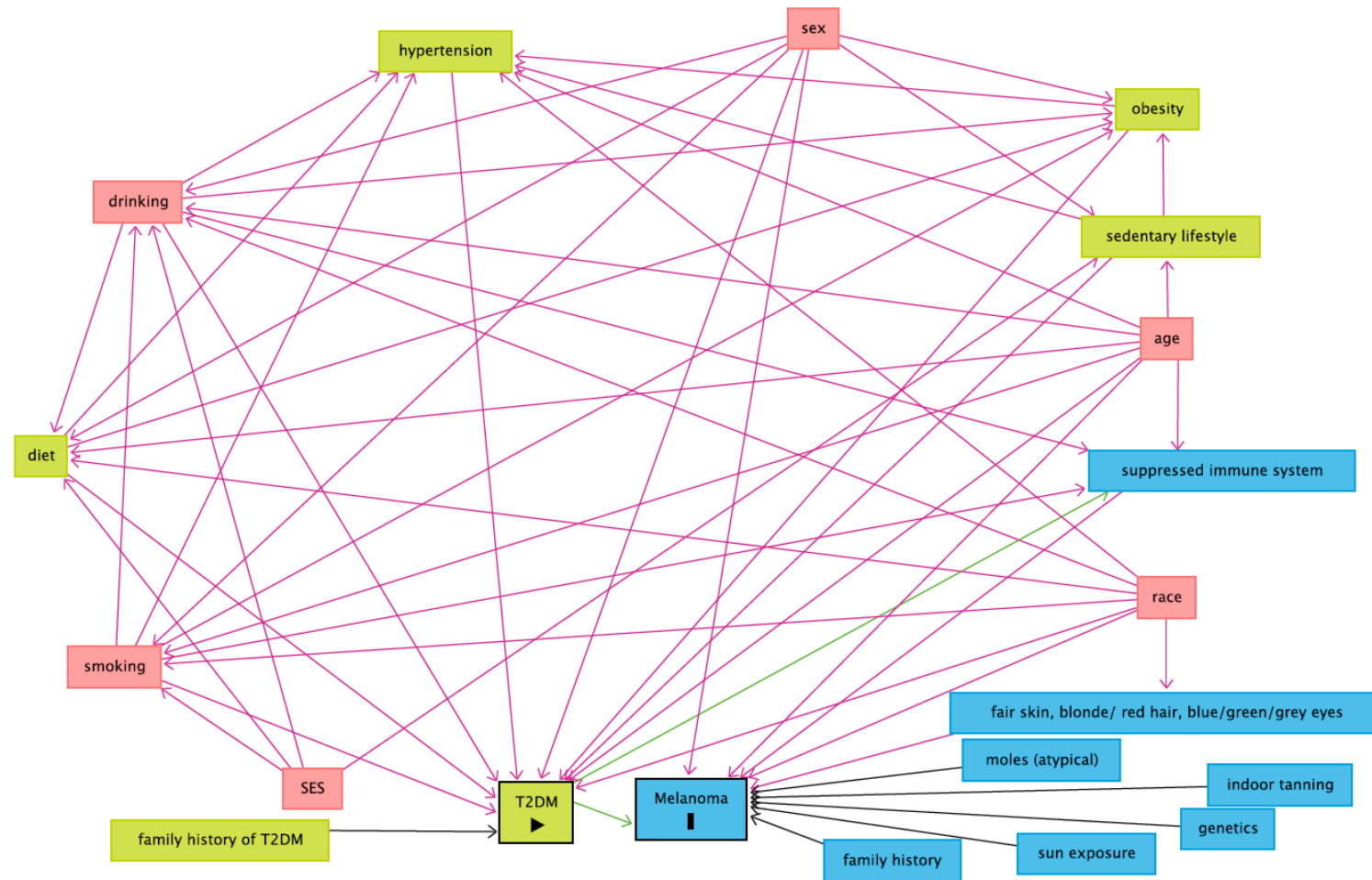


Figure B2-37 DAG of the association between T2DM and risk of melanoma, including potential confounders. Drinking defined as alcohol intake. SES: socioeconomic status.

Mouth cancer

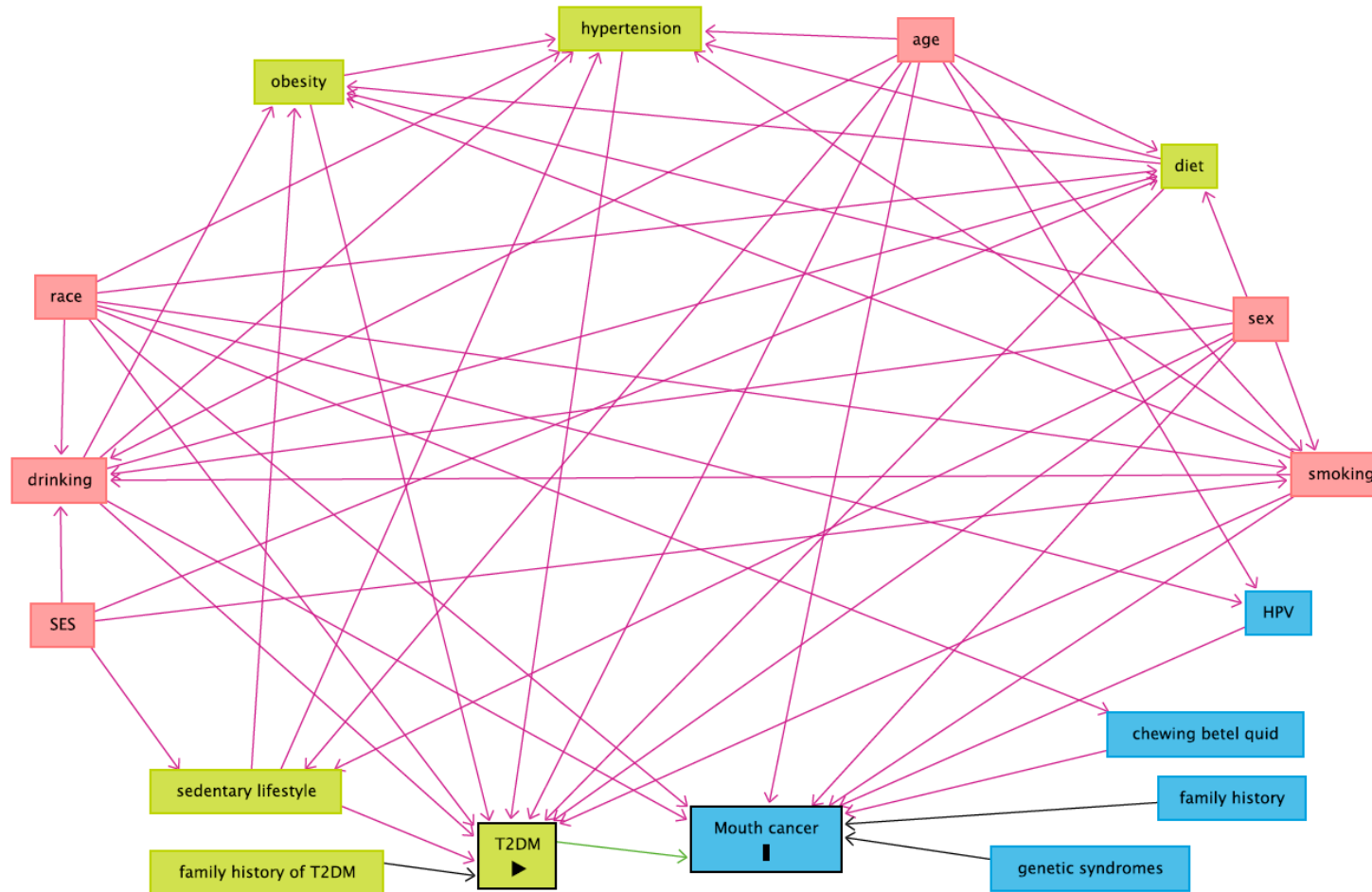


Figure B2-38 DAG of the association between T2DM and risk of mouth cancer, including potential confounders. Genetic syndromes: Fanconi anaemia, Dyskeratosis congenita. Drinking defined as alcohol intake. *Abbreviations.* HPV: human papillomavirus; SES: socioeconomic status.

Non-melanoma

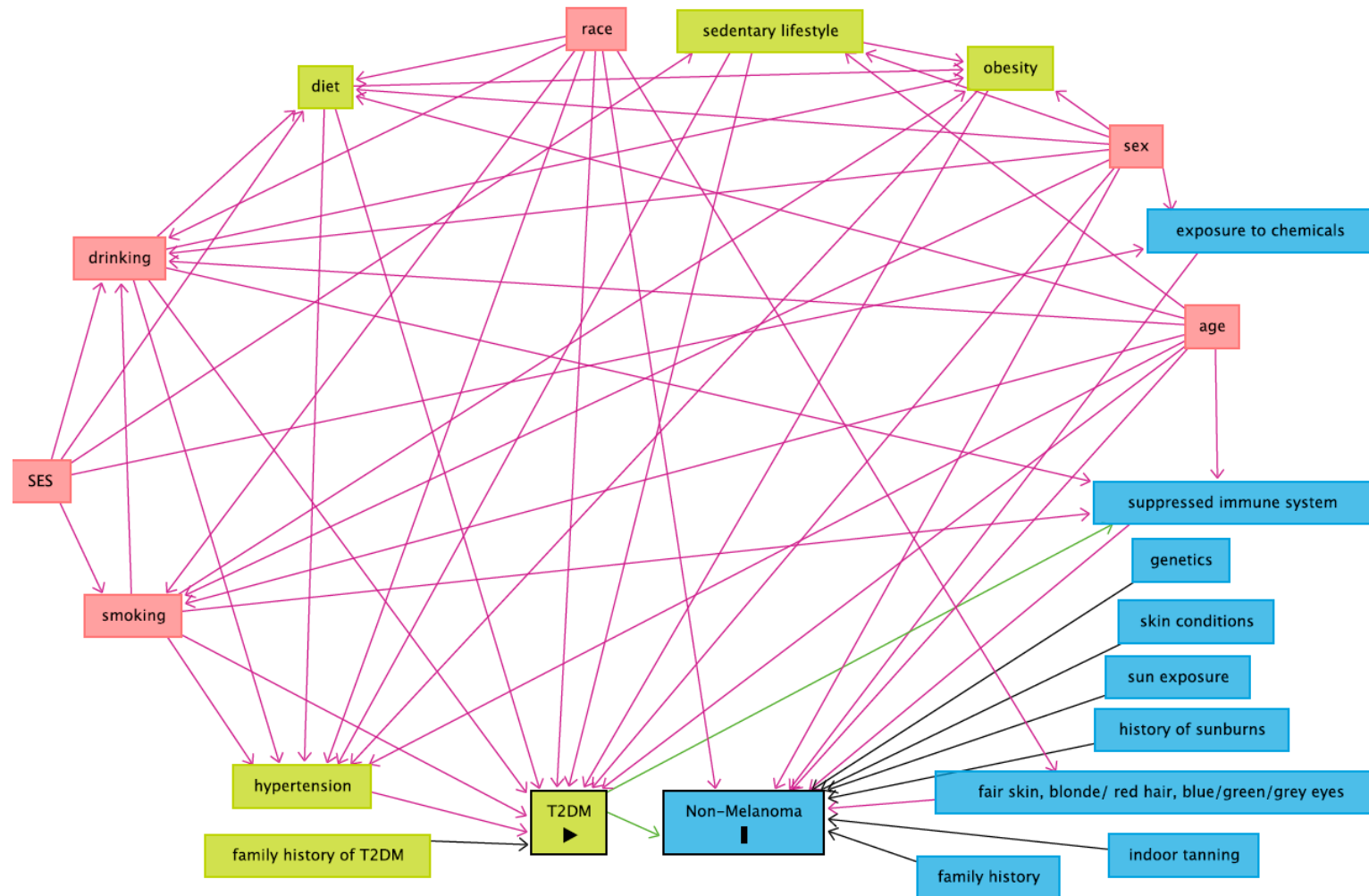


Figure B2-39 DAG of the association between T2DM and risk of non-melanoma, including potential confounders. Exposure to chemicals: arsenic compounds. Genetics: Gorlin syndrome, Xeroderma pigmentosum. Skin conditions: Actinic keratosis. Drinking defined as alcohol intake. SES: socioeconomic status.

Nose cancer

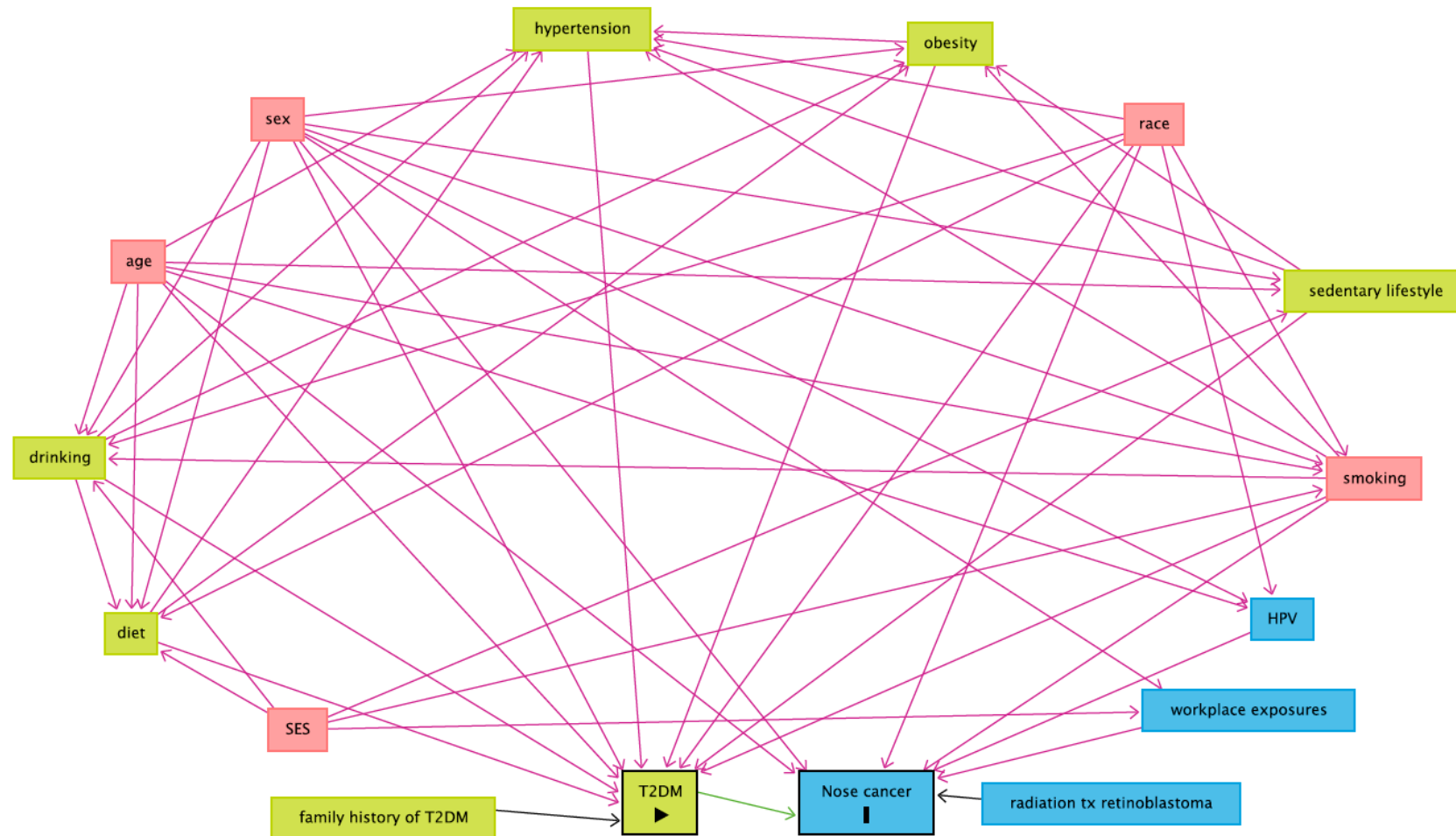


Figure B2-40 DAG of the association between T2DM and risk of nose cancer, including potential confounders. Workplace exposures: wood dust-carpentry, leather dust-shoemaking, nickel- stainless steel and radium. Drinking defined as alcohol intake. *Abbreviations.* HPV: human papillomavirus; SES: socioeconomic status; tx: treatment.

Oral cancer

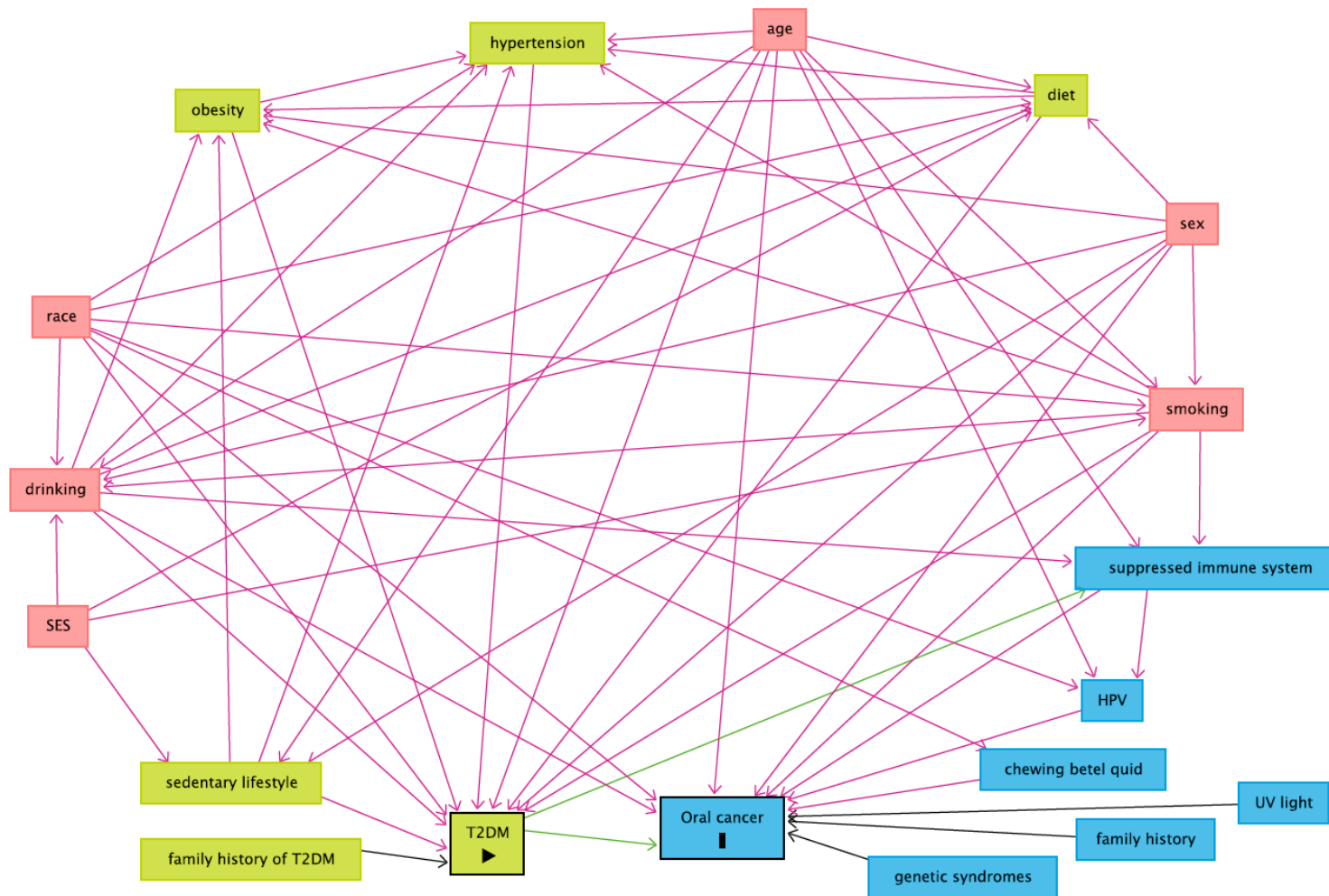


Figure B2-41 DAG of the association between T2DM and risk of oral cancer, including potential confounders. Genetic syndromes: Fanconi anaemia, Dyskeratosis congenita. Drinking defined as alcohol intake. *Abbreviations.* HPV: human papillomavirus; SES: socioeconomic status.

Ovarian cancer

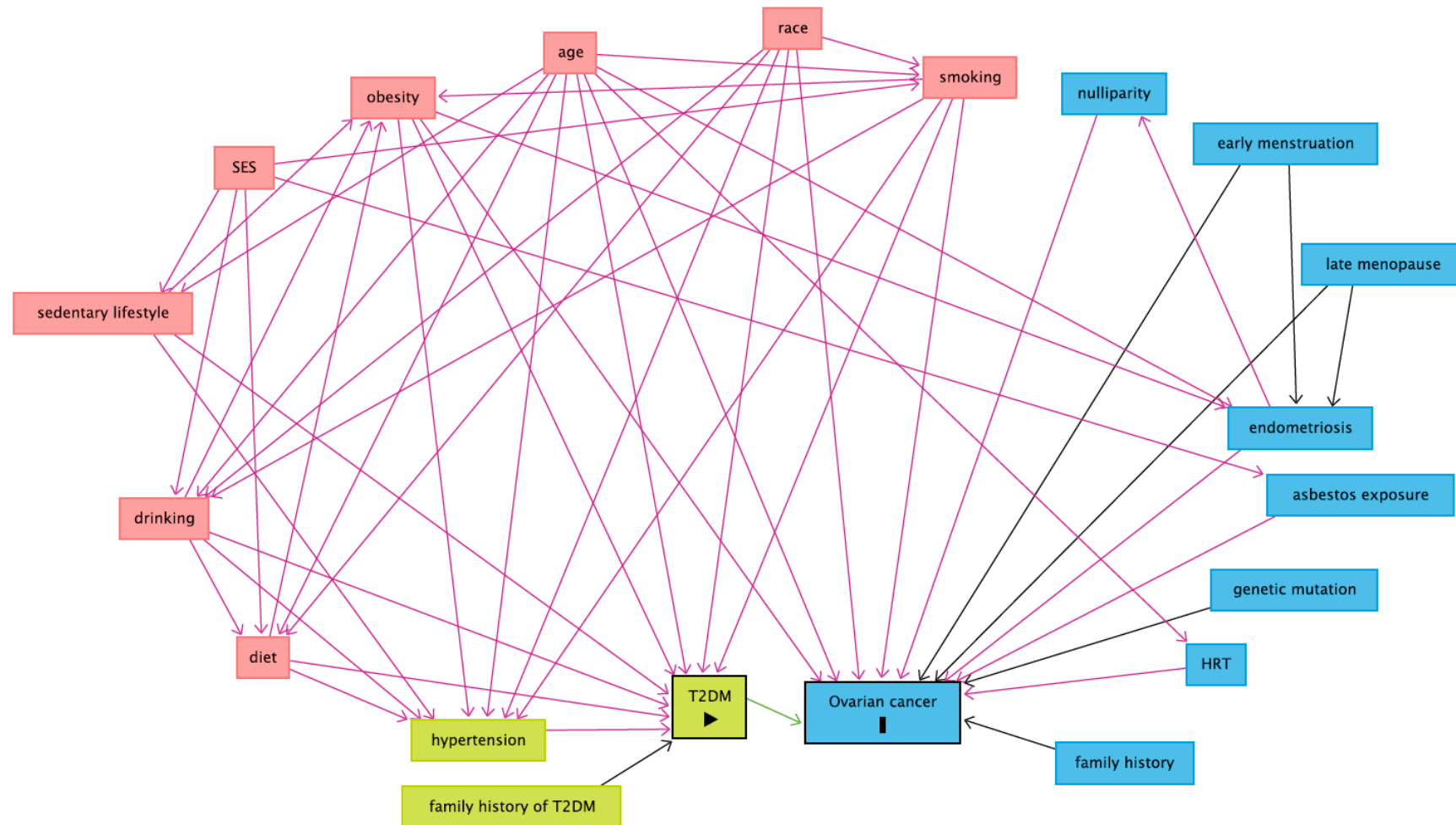


Figure B2-42 DAG of the association between T2DM and risk of ovarian cancer, including potential confounders. Genetic mutation: BRCA1 or BRCA2, or one associated with Lynch syndrome. Drinking defined as alcohol intake. *Abbreviations.* HRT: hormone replacement therapy; SES: socioeconomic status.

Pancreatic cancer

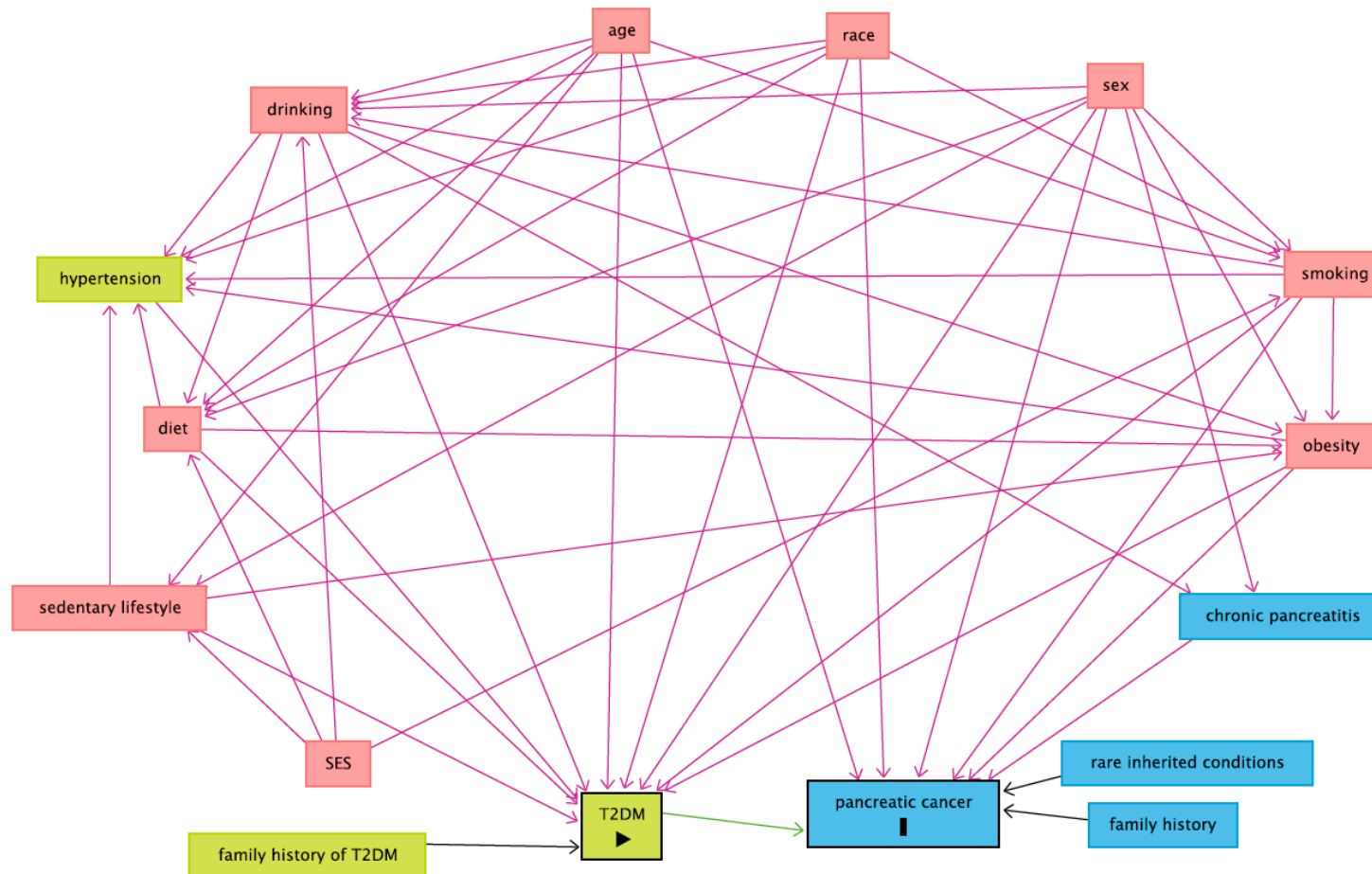


Figure B2-43 DAG of the association between T2DM and risk of pancreatic cancer, including potential confounders. Rare inherited conditions: hereditary pancreatitis (HP), Peutz-Jeghers syndrome (PJS), Familial malignant melanoma and pancreatic cancer (FAMM-PC), hereditary breast and ovarian cancer (HBOC) syndrome, Lynch syndrome, Li-Fraumeni syndrome (LFS), Familial adenomatous polyposis (FAP). Drinking defined as alcohol intake. SES: socioeconomic status.

Penile cancer

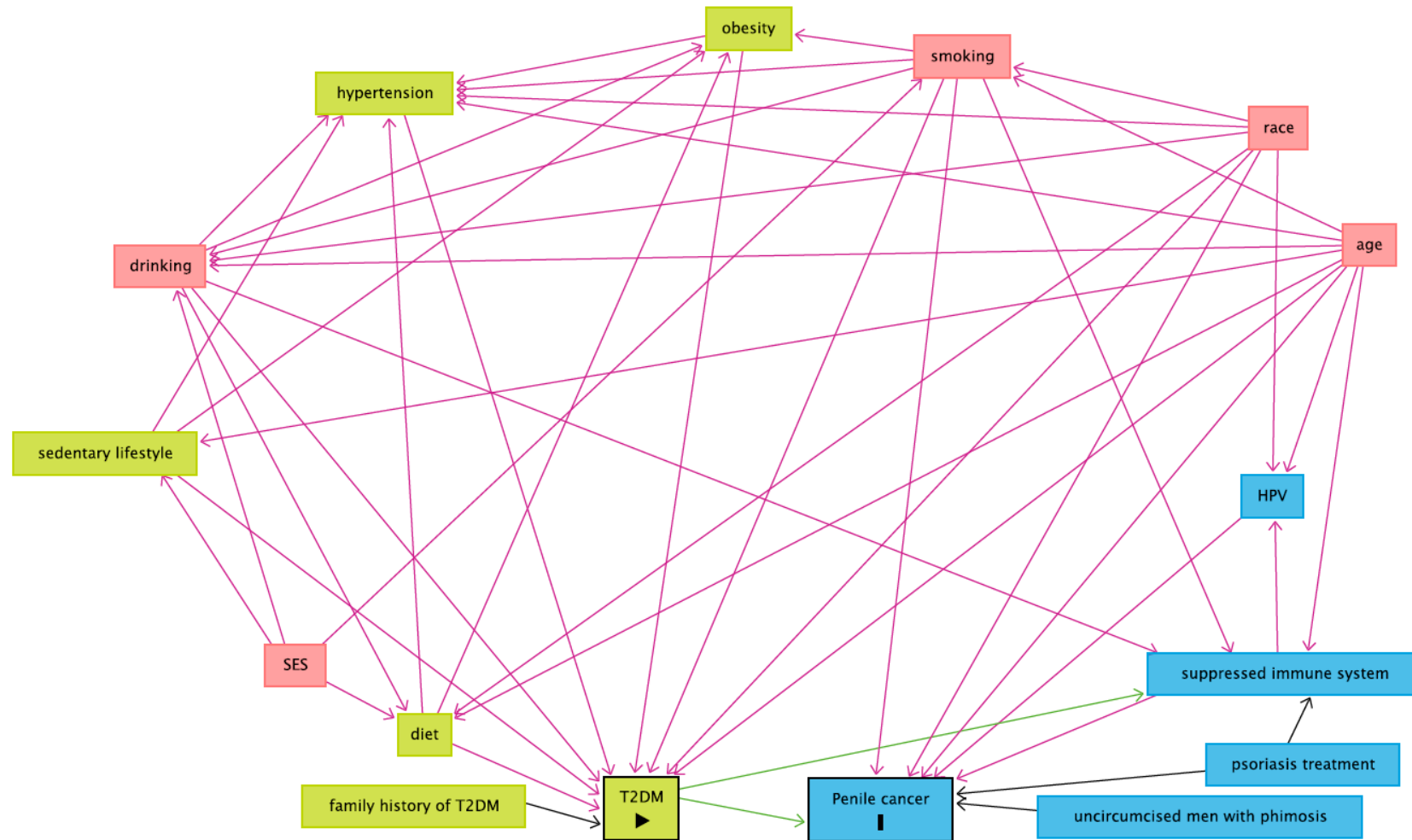


Figure B2-44 DAG of the association between T2DM and risk of penile cancer, including potential confounders. Psoriasis treatment: psoralen and light therapy (phototherapy). Drinking defined as alcohol intake. *Abbreviations.* HPV: human papillomavirus; SES: socioeconomic status.

Pharyngeal cancer

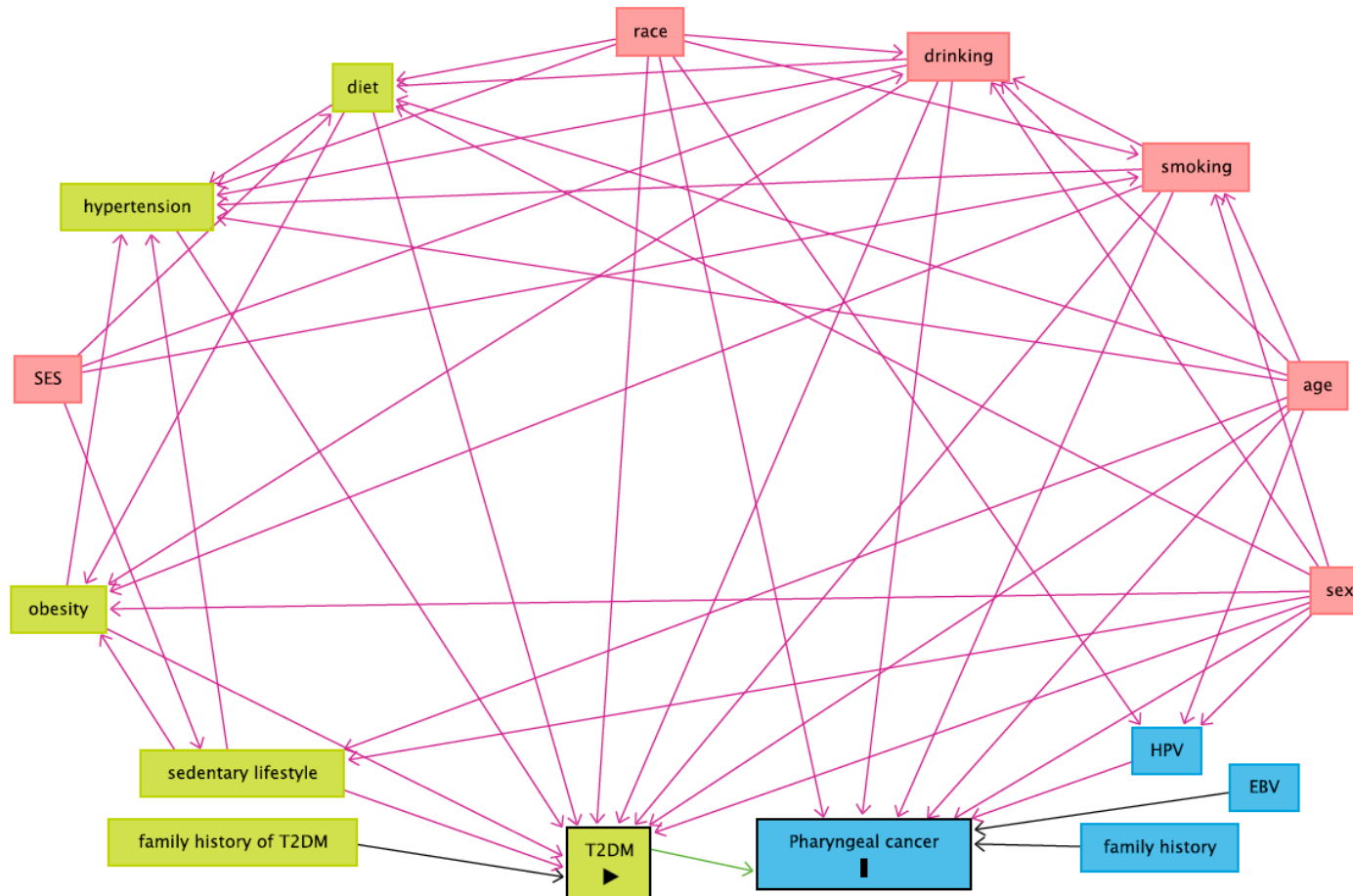


Figure B2-45 DAG of the association between T2DM and risk of pharyngeal cancer, including potential confounders. Drinking defined as alcohol intake.

Abbreviations. EBV: Epstein-Barr virus; HPV: human papillomavirus; SES: socioeconomic status.

Prostate cancer

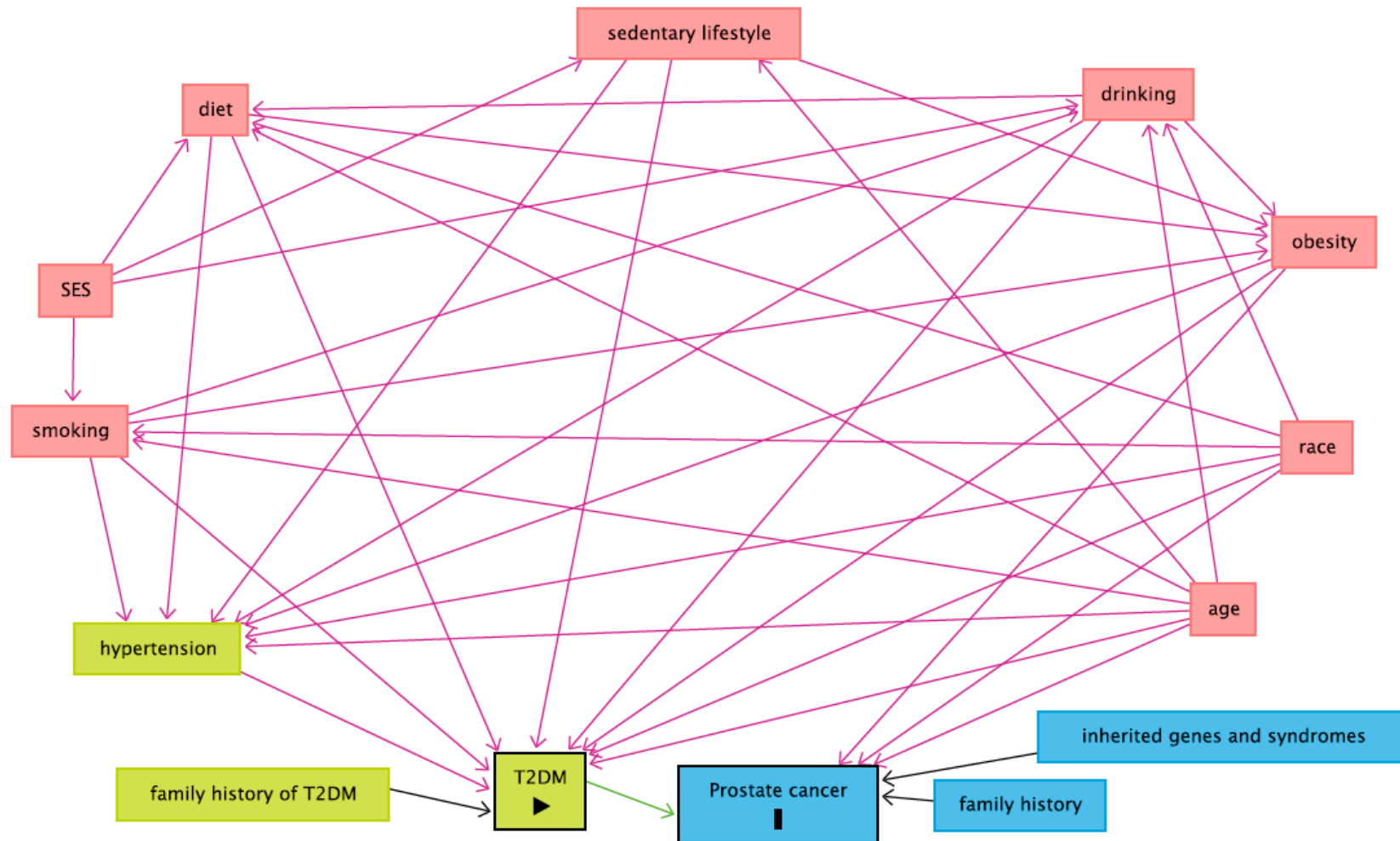


Figure B2-46 DAG of the association between T2DM and risk of prostate cancer, including potential confounders. Inherited genes and syndromes: genetic changes (BRCA2), Lynch syndrome. Drinking defined as alcohol intake. SES: socioeconomic status.

Rectal cancer

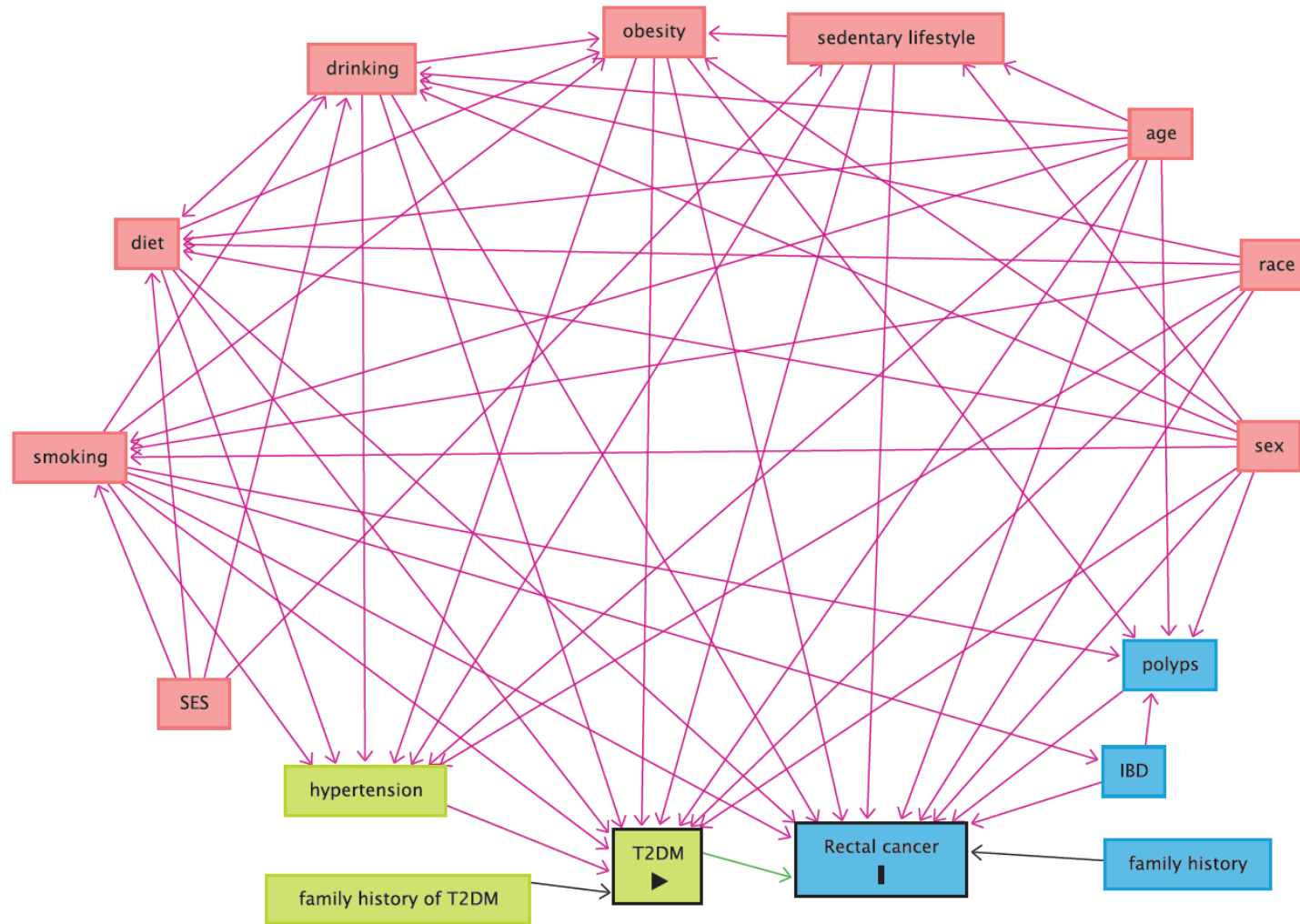


Figure B2-47 DAG of the association between T2DM and risk of rectal cancer, including potential confounders. Drinking defined as alcohol intake. *Abbreviations.* IBD: inflammatory bowel disease; SES: socioeconomic status.

Rectosigmoid carcinoma

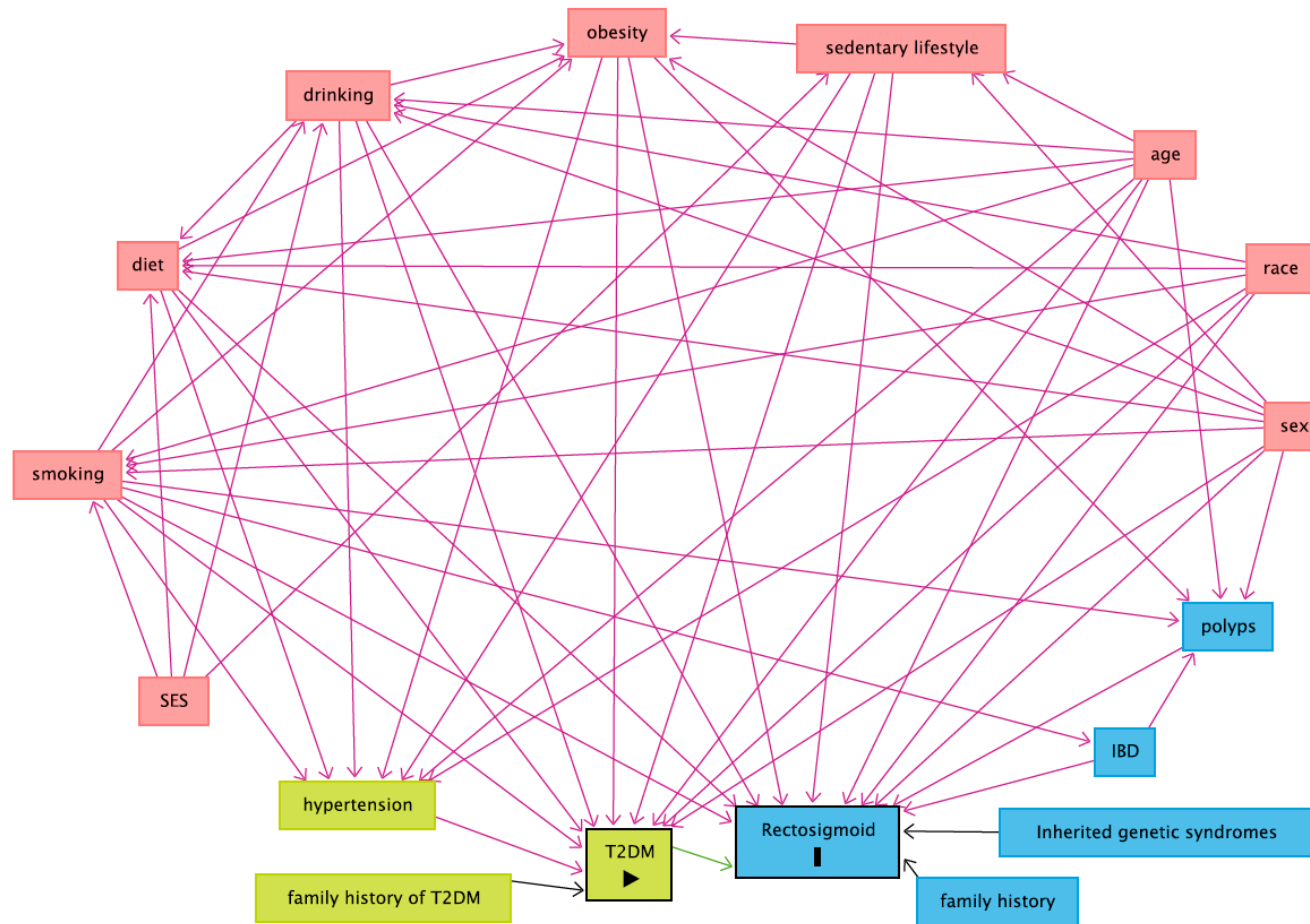


Figure B2-48 DAG of the association between T2DM and risk of rectosigmoid cancer, including potential confounders. Inherited Genetic Syndromes: Lynch syndrome and familial adenomatous polyposis. Drinking defined as alcohol intake. *Abbreviations.* IBD: inflammatory bowel disease; SES: socioeconomic status.

Salivary gland cancer

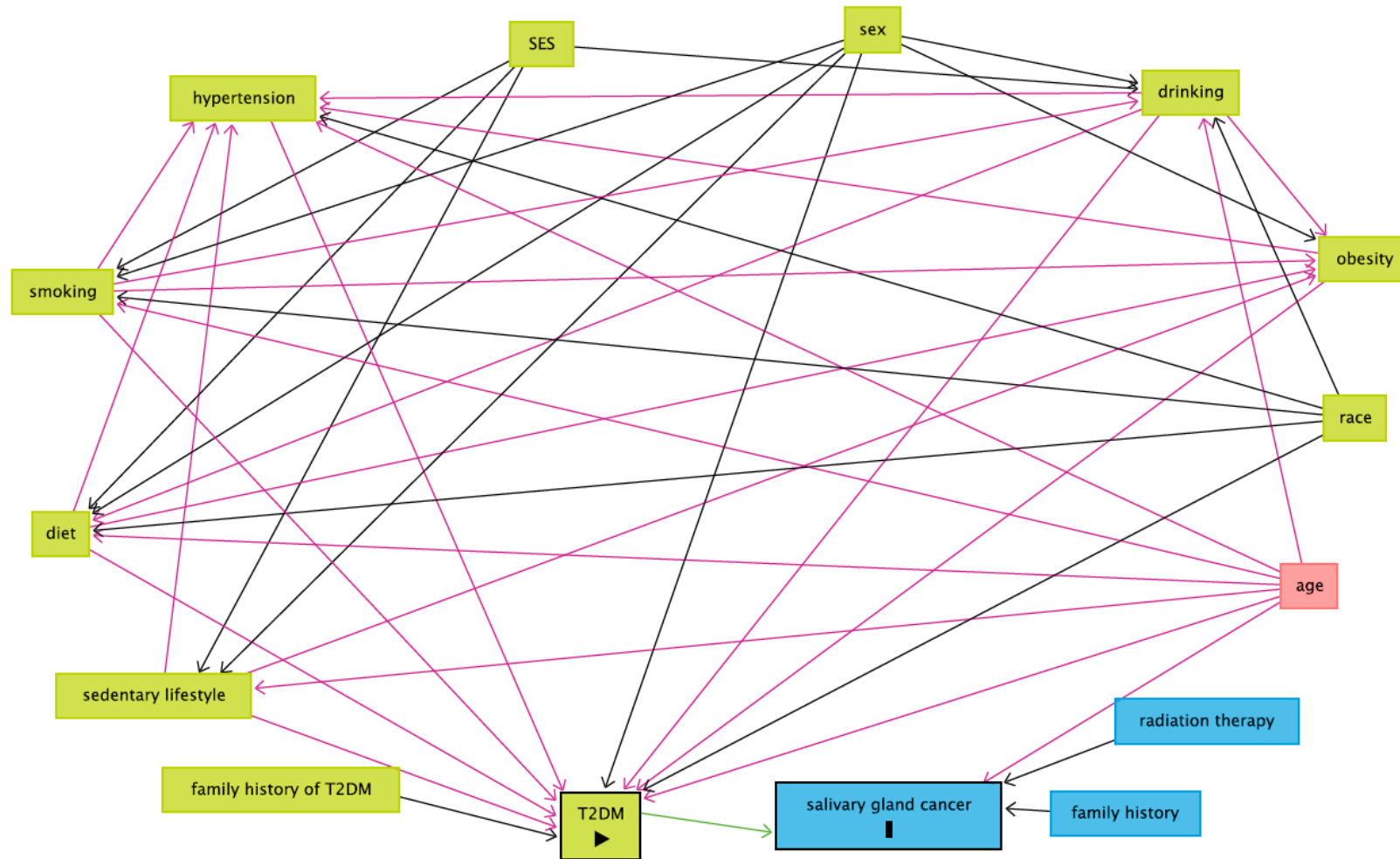


Figure B2-49 DAG of the association between T2DM and risk of salivary gland cancer, including potential confounders. Radiation therapy to the head or neck: medical treatment or due to workplace. Drinking defined as alcohol intake. SES: socioeconomic status.

Skin cancer

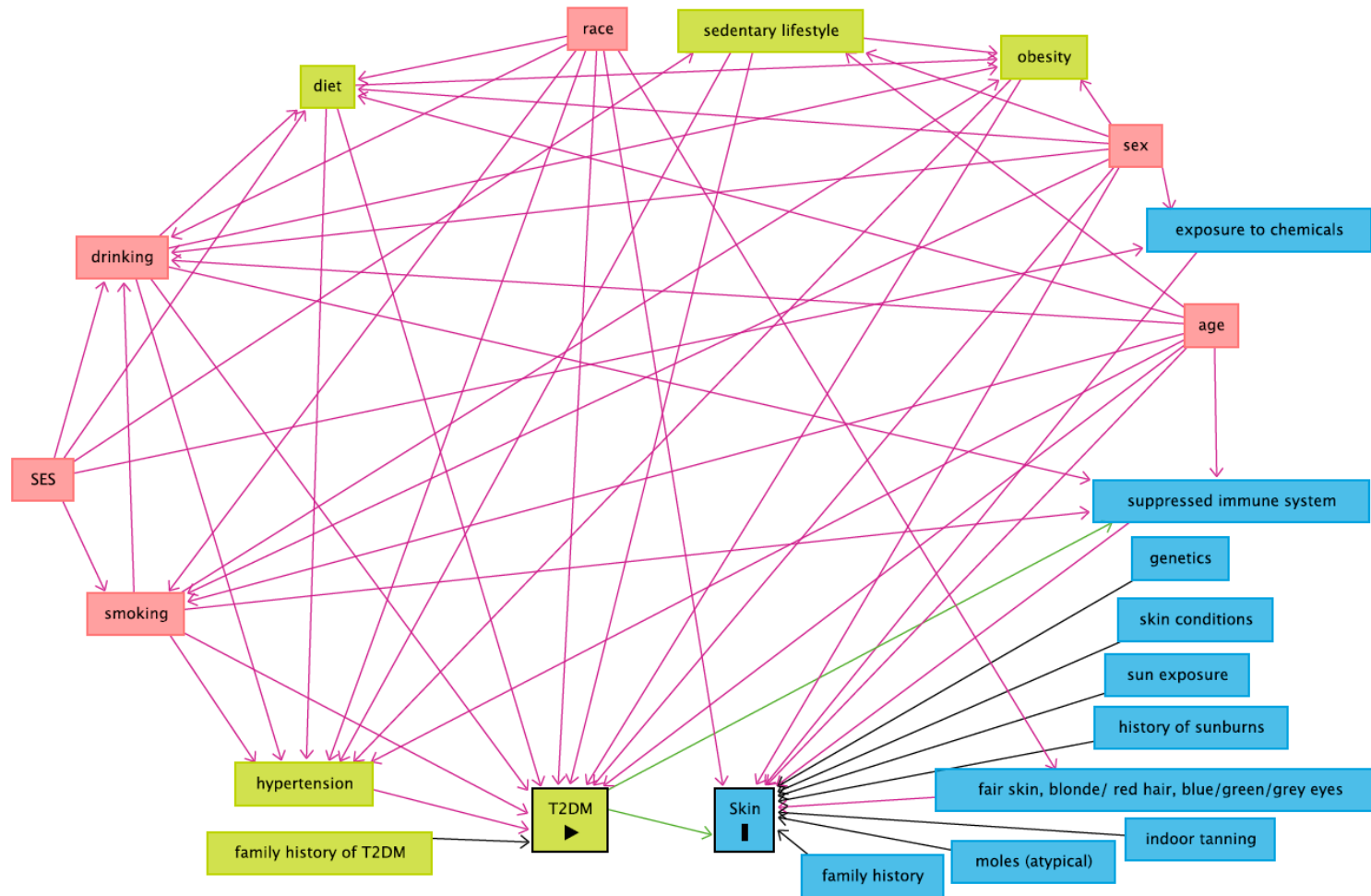


Figure B2-50 DAG of the association between T2DM and risk of skin cancer, including potential confounders. Exposure to chemicals: arsenic compounds. Genetics: familial melanoma, xeroderma pigmentosum, retinoblastoma, Gorlin syndrome, Li-Fraumeni syndrome, Werner syndrome, and certain hereditary breast and ovarian cancer syndromes. Skin conditions: Actinic keratosis. Drinking defined as alcohol intake. SES: socioeconomic status.

Small intestine cancer

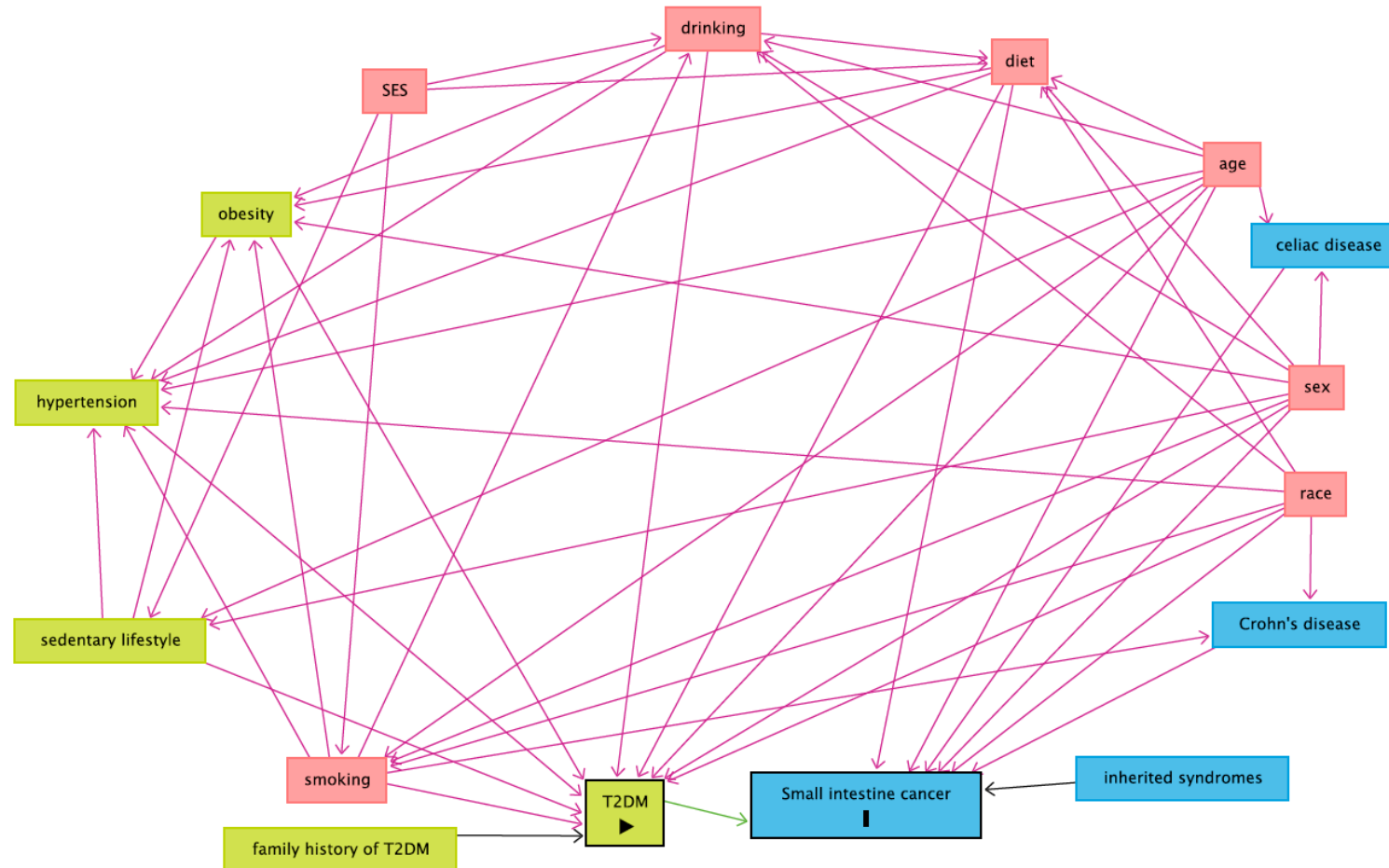


Figure B2-51 DAG of the association between T2DM and risk of small intestine cancer, including potential confounders. Inherited syndromes: Familial adenomatous polyposis (FAP)- MUTYH-associated polyposis, Lynch syndrome, Peutz-Jeghers syndrome, Cystic fibrosis (CF). Drinking defined as alcohol intake. SES: socioeconomic status.

Stomach (gastric) cancer

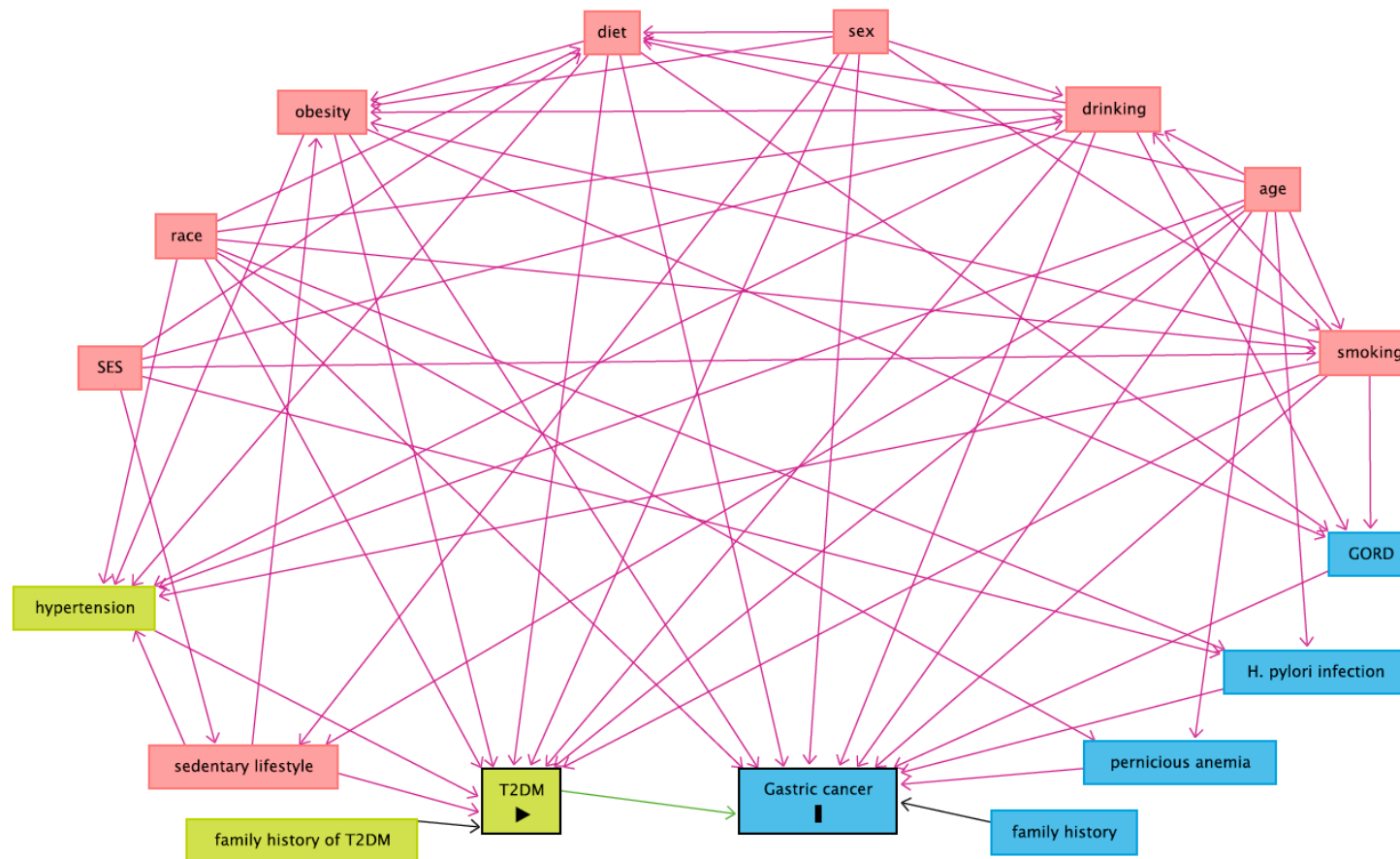


Figure B2-52 DAG of the association between T2DM and risk of stomach (gastric) cancer, including potential confounders. Drinking defined as alcohol intake.
Abbreviations. GORD/GERD: Gastroesophageal reflux disease; SES: socioeconomic status.

Testicular cancer

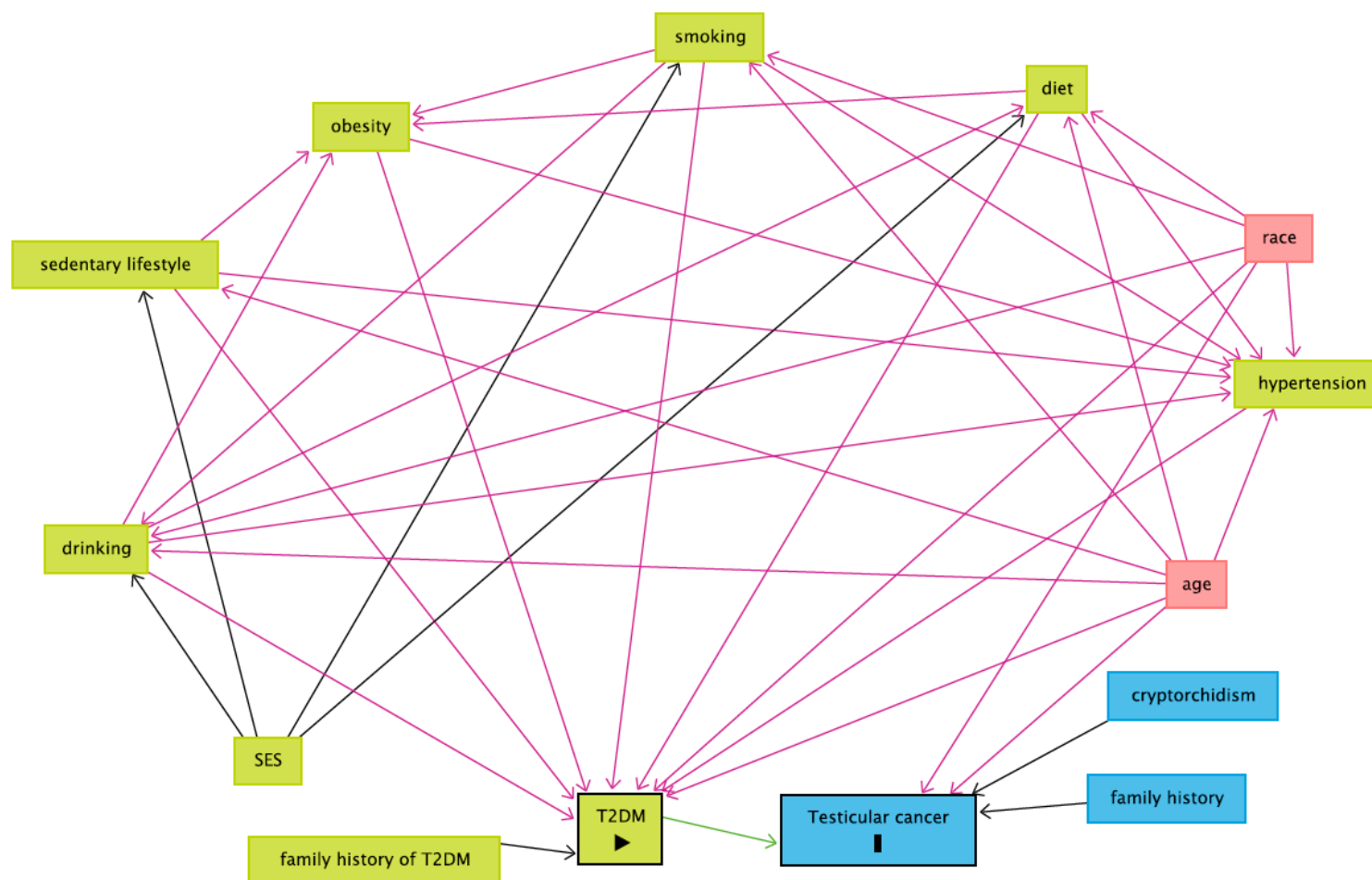


Figure B2-53 DAG of the association between T2DM and risk of testicular cancer, including potential confounders. Drinking defined as alcohol intake. SES: socioeconomic status.

Tongue cancer

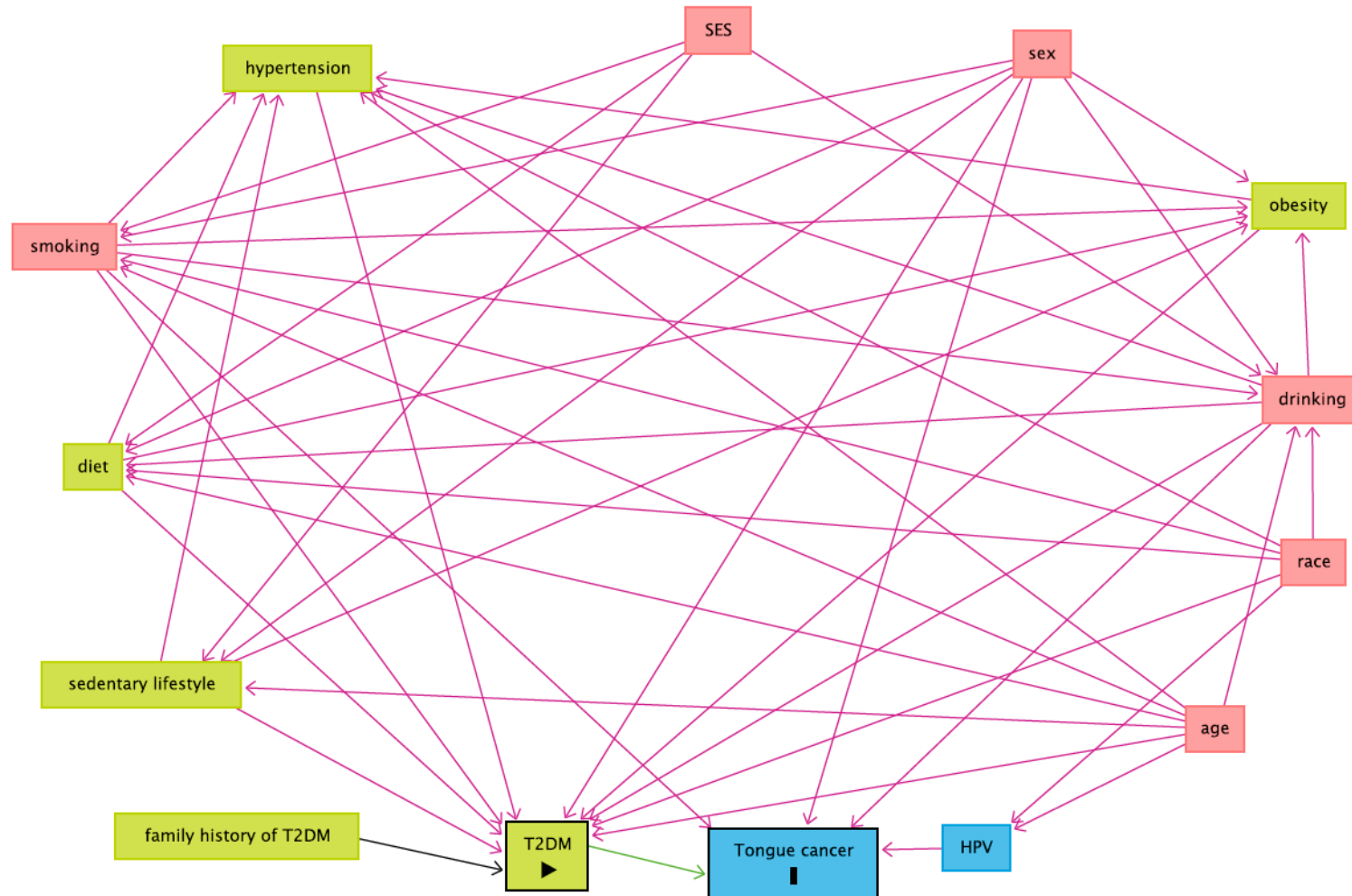


Figure B2-55 DAG of the association between T2DM and risk of tongue cancer, including potential confounders. Drinking defined as alcohol intake.
Abbreviations. HPV: human papillomavirus; SES: socioeconomic status.

Urinary tract cancer (UTC)

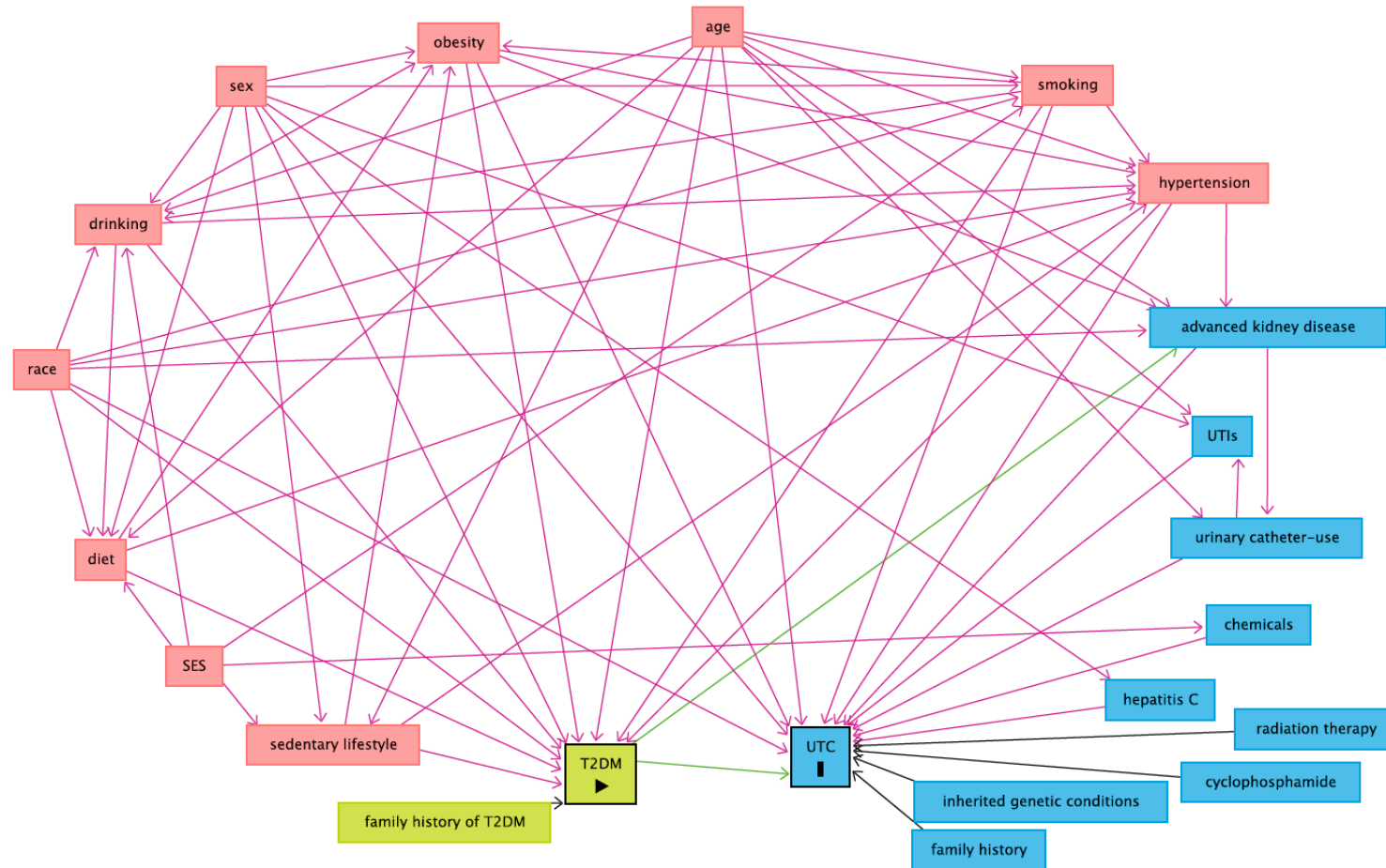


Figure B2-56 DAG of the association between T2DM and risk of UTC, including potential confounders. Chemicals: amylamines, Polycyclic aromatic hydrocarbons (PAH). Inherited genetic conditions: von Hippel-Lindau disease, Hereditary papillary renal cell carcinoma, Hereditary leiomyoma-renal cell carcinoma, Birt-Hogg-Dube (BHD) syndrome, Familial renal cancer, Cowden syndrome, Tuberous sclerosis. Drinking defined as alcohol intake. *Abbreviations.* SES: socioeconomic status; UTC: urinary tract cancer.

Appendix C. Meta-analyses

This appendix provides supplementary information for Chapter 6: Meta-analyses.

C1. Unadjusted (association) meta-analyses

C1.1 Overall results

Table C1-1 Pooled unadjusted RR estimates of the association between T2DM and risk of cancer risk

Type of cancer	No. Studies	RR (95% CI)	I ²	Forest plot
Bladder	13	1.51 (1.15, 1.98)	99%	
Brain/NS	4	1.07 (0.78, 1.45)	68%	
Breast	31	1.20 (1.00, 1.44)	99%	
BTC	2	1.34 (0.86, 2.09)	95%	
Cervix uteri	6	1.00 (0.88, 1.13)	12%	
Colon	20	1.51 (1.24, 1.85)	97%	
CRC	22	1.58 (1.13, 2.20)	100%	
Endometrium	13	2.37 (1.83, 3.08)	94%	
Esophagus	10	1.32 (0.99, 1.77)	89%	
Gallbladder	5	1.84 (1.07, 3.16)	79%	
GI	3	1.33 (0.96, 1.82)	93%	
H&N	3	1.52 (1.04, 2.23)	95%	
HCC	6	2.70 (1.88, 3.88)	99%	
Kidney	10	1.48 (1.08, 2.02)	97%	
Leukemia	3	1.10 (0.46, 2.65)	94%	
Liver	12	2.52 (1.93, 3.29)	98%	
Lung	15	1.46 (1.02, 2.09)	99%	
Lymphoma	3	1.12 (0.66, 1.90)	62%	
Melanoma	3	1.07 (0.59, 1.94)	90%	
NHL	7	1.49 (1.18, 1.87)	89%	
Non-melanoma	2	0.87 (0.62, 1.21)	86%	
Oral	3	1.51 (0.64, 3.56)	94%	
Ovary	6	1.24 (1.05, 1.46)	36%	
Pancreas	22	2.45 (1.97, 3.05)	97%	
Prostate	21	1.09 (0.91, 1.29)	99%	
Rectosigmoid	2	1.01 (0.43, 2.36)	73%	
Rectum	16	1.20 (1.04, 1.39)	76%	
Skin	2	1.12 (0.55, 2.30)	77%	
Small intestine	2	1.88 (0.32, 10.99)	78%	
Stomach	9	1.30 (0.93, 1.80)	97%	
Thyroid	5	1.17 (0.96, 1.42)	70%	

Statistical significance ($p < 0.05$) for effect estimates is indicated in blue. I² indicates heterogeneity estimate. *Abbreviations.* Brain/NS: Brain/Nervous system; BTC: Biliary tract cancer; CI: confidence interval; CRC: colorectal cancer; GI: gastrointestinal; H&N: head and neck; HCC: hepatocellular carcinoma; NHL: non-Hodgkin lymphoma; RR: pooled relative risk; UTC: urinary tract cancer.

As shown in the table, the point estimate is above one for nearly all cancers (30 out of 31 types), except for non-melanoma skin cancer (RR = 0.87; 95% CI, 0.62–1.21)

In total, statistically significant associations ($p < 0.05$) were observed for 14 cancer types, including bladder (RR, 1.51; 95% CI, 1.15, 1.98), colon (RR, 1.51; 95% CI, 1.24, 1.85), CRC (RR, 1.58; 95% CI 1.13, 2.20), endometrial (RR, 2.37; 95% CI 1.83, 3.08), H&N (RR, 1.52; 95% CI 1.04, 2.23), HCC (RR, 2.70; 95% CI 1.88, 3.88), kidney (RR, 1.48; 95% CI 1.08, 2.02), liver (RR, 2.52; 95% CI 1.93, 3.29), lung (RR, 1.46; 95% CI 1.02, 2.09), NHL (RR, 1.49; 95% CI, 1.18, 1.87), ovarian (RR, 1.24; 95% CI 1.05, 1.46), pancreatic (RR, 2.45; 95% CI 1.97, 3.05), and rectal cancer (RR, 1.20; 95% CI 1.04, 1.39). These findings suggest that T2DM is strongly associated with an increased risk of liver, biliary tract, and pancreatic cancer.

Substantial (I^2 , 61–75%) and considerable ($I^2 > 75\%$) heterogeneity was observed in all pooled unadjusted estimates, excluding the ones related to cervix uteri ($I^2=12\%$) and ovarian cancer ($I^2=36\%$).

The unadjusted estimates of the association between T2DM and the risk of ten specific cancer types were derived from a single study; a significant association was found in five (myeloma, pharyngeal, ECC, ICC, and esophagogastric junction carcinoma) (Table C1-2).

Table C1-2 Unadjusted effect estimates of the association between T2DM and cancer risk, based on a single study for each cancer type

Type of cancer	Effect estimate (95% CI)	Forest plot
Anal	0.91 (0.20, 4.10)	
ECC	1.36 (1.10, 1.67)	
EJA	1.51 (1.26, 1.81)	
ICC	1.74 (1.45, 2.08)	
Large intestine	1.20 (0.55, 2.62)	
Larynx	1.40 (0.32, 6.02)	
Myeloma	0.82 (0.71, 0.96)	
Pharynx	10.46 (3.19, 34.33)	
Soft tissue	2.03 (0.91, 3.97)	
UTC	1.90 (0.83, 4.40)	

Studies on ECC, EJA, and ICC used a case-control design, while all others were cohort studies. The outcome measure was RR for myeloma, HR for anal, soft tissue, and UTC, and OR for the remaining cancers. Statistical significance ($p < 0.05$) for effect estimates is indicated in blue. *Abbreviations.* CI: confidence interval; ECC: extrahepatic cholangiocarcinoma; EJA: esophagogastric junction adenocarcinoma; ICC: intrahepatic cholangiocarcinoma; HR: hazard ratio; OR: odds ratio; RR: relative risk.

C1.2 Subgroup analyses

Study design

Subgroup analyses by study design (cohort and case-control studies) were conducted for 14 types of cancer (Table C1-3).

Table C1-3 Subgroup analyses by study design (cohort and case control) from the unadjusted (association) meta-analyses of T2DM on cancer risk

Type of cancer	Effect estimate (95% CI)	p	I ²	p subgroup
Brain/NS				
Cohort	1.27 (0.61, 2.64)	0.52	78%	0.41
Case-control	0.92 (0.72, 1.17)	0.50	N/A	
Breast				
Cohort	1.09 (0.88, 1.36)	0.42	99%	0.20
Case-control	1.48 (0.98, 2.23)	0.06	96%	
Colon				
Cohort	1.55 (1.24, 1.94)	0.0001	98%	0.35
Case-control	1.37 (1.22, 1.55)	<0.00001	0%	
CRC				
Cohort	1.62 (1.11, 2.37)	0.01	100%	0.33
Case-control	1.32 (1.14, 1.54)	0.0002	51%	
Endometrium				
Cohort	2.22 (1.63, 3.02)	<0.00001	95%	0.45
Case-control	2.74 (1.75, 4.29)	<0.0001	88%	
Esophagus				
Cohort	1.30 (0.92, 1.85)	0.14	85%	0.85
Case-control	1.36 (1.02, 1.81)	0.03	60%	
H&N				
Cohort	1.97 (0.63, 6.18)	0.24	97%	0.3
Case-control	1.07 (0.92, 1.25)	0.39	N/A	
HCC				
Cohort	2.80 (1.04, 7.58)	0.04	100%	0.89
Case-control	2.59 (1.74, 3.85)	<0.00001	92%	
Kidney				
Cohort	1.51 (1.08, 2.11)	0.02	97%	0.44
Case-control	1.22 (0.81, 1.86)	0.34	N/A	
Lymphoma				
Cohort	0.76 (0.25, 2.31)	0.63	69%	0.25
Case-control	1.50 (1.08, 2.09)	0.02	N/A	
Pancreas				
Cohort	2.44 (1.91, 3.11)	<0.00001	97%	0.89
Case-control	2.49 (2.16, 2.86)	<0.00001	0%	
Prostate				
Cohort	1.16 (0.92, 1.47)	0.21	99%	0.003
Case-control	0.81 (0.77, 0.85)	<0.00001	25%	
Rectum				
Cohort	1.11 (0.96, 1.29)	0.15	68%	0.01
Case-control	1.65 (1.25, 2.18)	0.0004	65%	
Thyroid				
Cohort	1.25 (1.07, 1.47)	0.006	56%	0.009
Case-control	0.71 (0.48, 1.06)	0.09	N/A	

Unadjusted RR estimates in each subgroup. P-values in the third column indicate if the subgroup-specific effects are significant ($p < 0.05$). I², indicate heterogeneity in each subgroup; values are N/A for subgroups in which only 1 study was found, and therefore no subgroup I² value could be calculated. P subgroup represents the p-value for the test of subgroup differences (statistical significance [$p < 0.05$] is indicated in blue). Effect estimate: cohort studies (RR/OR) and case-control studies (OR). *Abbreviations.* Brain/NS: Brain/Nervous system; CI: confidence interval; CRC: colorectal cancer; H&N: head and neck; HCC: hepatocellular carcinoma; N/A: not applicable; OR: odds ratio; RR: relative risk.

The test of subgroup differences was statistically significant ($p < 0.05$) only for prostate ($p = 0.003$), rectal ($p = 0.01$), and thyroid cancer ($p = 0.009$). Cohort studies examining the association between T2DM and prostate cancer suggested that T2DM was associated with an increased risk (RR 1.16; 95% CI, 0.92-1.47), whereas case-control studies indicated a decreased risk (OR 0.81; 95% CI, 0.77-0.85). A similar pattern was observed for thyroid cancer, with cohort studies suggesting that T2DM was associated with an increased risk (RR 1.25; 95% CI, 1.07-1.47) and case-control studies indicating that T2DM was associated with a decreased risk (OR 0.71; 95% CI, 0.48-1.06).

Lastly, both study designs assessing the association between T2DM and rectal cancer suggested an increased risk. However, while cohort studies reported an RR of 1.11 (95% CI, 0.96-1.29), case-control studies indicated a stronger and statistically significant association (OR 1.65; 95% CI, 1.25-2.18), resulting in non-overlapping confidence intervals.

Among the 14 cancer types analysed, only six RRs derived from case-control studies were larger than those obtained from cohort studies. This pattern reflects the impact of matching in case-control studies, where selecting controls with similar characteristics to cases (e.g., age, sex) reduces the effect of confounders that might otherwise inflate risk estimates. Consequently, case-control studies may yield lower risk estimates than cohort studies, as matching helps account for confounding factors that could otherwise increase risk estimates in cohort studies. In contrast, cohort studies involve populations where risk factors are unevenly distributed between exposed and unexposed groups, potentially leading to higher unadjusted effect estimates.

Sex

Subgroup analyses by sex were conducted for 28 types of cancer (Table C1-4). The test of subgroup differences was statistically significant ($p < 0.05$) only for skin cancer ($p = 0.009$), although these findings are based on data reported from a single study. These results suggest that the association between T2DM and cancer risk is comparable between males and females.

Table C1-4 Subgroup analyses by sex from the unadjusted (association) meta-analyses of T2DM on cancer risk

Type of cancer	RR (95% CI)	p	I ²	p subgroup
Anal				
Male	1.67(0.17,16.07)	0.66	N/A	0.53
Female	0.63 (0.08, 5.01)	0.66	N/A	
Bladder				
Male	1.14 (1.04, 1.24)	0.006	0%	0.15
Female	1.82 (0.97, 3.43)	0.06	89%	
Bone				
Male	1.31 (0.60,2.84)	0.49	N/A	0.96
Female	1.35 (0.45,4.02)	0.59	N/A	
Brain/NS				
Male	0.71 (0.36, 1.43)	0.34	0%	0.8
Female	0.82 (0.38, 1.77)	0.61	0%	
Colon				
Male	1.54 (1.20, 1.99)	0.0007	93%	0.94
Female	1.53 (1.22, 1.91)	0.0002	91%	
CRC				
Male	1.44 (1.30, 1.59)	<0.00001	59%	0.99
Female	1.44 (1.31, 1.57)	<0.00001	40%	
Esophagus				
Male	1.23 (0.99, 1.55)	0.07	0%	0.55
Female	0.96 (0.43, 2.13)	0.91	30%	
Gallbladder				
Male	1.15 (0.87, 1.50)	0.33	0%	0.21
Female	1.78 (0.94, 3.37)	0.08	56%	
GI				
Male	1.69 (1.43, 2.01)	<0.00001	N/A	0.72
Female	1.61(1.32, 1.98)	<0.00001	N/A	
H&N				
Male	0.95 (0.81, 1.13)	0.57	0%	0.91
Female	1.00 (0.42, 2.39)	1.00	79%	
HCC				
Male	3.58 (2.83, 4.52)	<0.00001	67%	0.93
Female	3.49 (2.11, 5.77)	<0.00001	73%	
Kaposi's sarcoma				
Male	2.84 (1.41,5.71)	0.003	N/A	0.93
Female	3.03 (0.91,10.14)	0.07	N/A	
Kidney				
Male	1.50 (0.13, 17.03)	0.74	91%	0.76
Female	2.26 (0.74, 6.91)	0.15	96%	
Large intestine				
Male	1.44 (0.56,3.70)	0.45	N/A	0.54
Female	0.85 (0.20,3.51)	0.82	N/A	
Liver				
Male	2.37 (1.49, 3.77)	0.0003	78%	0.50
Female	2.84 (2.19, 3.68)	<0.00001	32%	
Lung				
Male	2.09 (1.14, 3.83)	0.02	97%	0.84
Female	2.28 (1.24, 4.20)	0.008	96%	
Melanoma				
Male	0.89 (0.43,1.83)	0.75	N/A	0.11
Female	0.16 (0.02,1.18)	0.07	N/A	
NHL				
Male	2.21 (1.62, 3.00)	<0.00001	0%	0.86
Female	2.30 (1.65, 3.20)	<0.00001	0%	
Non-melanoma				
Male	0.70 (0.52,0.95)	0.02	N/A	0.86
Female	0.73 (0.51,1.05)	0.09	N/A	
Oral				
Male	1.43 (0.26, 7.85)	0.68	68%	0.3
Female	3.58 (2.57, 4.98)	<0.00001	N/A	

Type of cancer	RR (95% CI)	p	I ²	p subgroup
Pancreas				
Male	1.65 (0.97, 2.79)	0.06	82%	0.95
Female	1.61 (1.26, 2.07)	0.0002	69%	
Pharynx				
Male	12.69 (2.84,56.80)	0.0009	N/A	0.56
Female	5.64 (0.59,54.78)	0.14	N/A	
Rectum				
Male	1.15 (0.96, 1.39)	0.14	47%	0.79
Female	1.20 (0.98, 1.47)	0.09	34%	
Skin				
Male	1.04 (0.76,1.43)	0.81	N/A	0.009
Female	0.45 (0.27,0.77)	0.004	N/A	
Small intestine				
Male	1.41 (0.31, 6.46)	0.66	49%	0.54
Female	2.97 (0.47, 18.91)	0.25	48%	
Stomach				
Male	1.06 (0.80, 1.39)	0.69	0%	0.71
Female	0.96 (0.62, 1.48)	0.85	0%	
Thyroid				
Male	0.72 (0.25,2.04)	0.54	N/A	0.17
Female	1.61 (0.96,2.69)	0.07	N/A	
UTC				
Male	1.17(0.96,1.42)	0.11	N/A	0.31
Female	1.43 (1.03,2.00)	0.04	N/A	

Unadjusted RR estimates in each subgroup. P-values in the third column indicate if the subgroup-specific effects are significant ($p < 0.05$). I^2 , indicate heterogeneity in each subgroup; values are N/A for subgroups in which only 1 study was found, and therefore no subgroup I^2 value could be calculated. P subgroup represents the p-value for the test of subgroup differences (statistical significance [$p < 0.05$] is indicated in blue). *Abbreviations.* Brain/NS: Brain/Nervous system; CI: confidence interval; CRC: colorectal cancer; GI: gastrointestinal; H&N: head and neck; HCC: hepatocellular carcinoma; NHL: non-Hodgkin lymphoma; N/A: not applicable; RR: relative risk; UTC: urinary tract cancer.

Geographic region

Subgroup analyses by geographic region were conducted for 30 types of cancer (Table C1-5). The test of subgroup differences was statistically significant ($p < 0.05$) for 16 types of cancer: breast ($p=0.02$), BTC ($p<0.00001$), CRC ($p=0.01$), gallbladder($p=0.04$), GI ($p<0.00001$), HCC ($p<0.00001$), kidney ($p=0.01$), leukemia ($p=0.05$), lung ($p=0.0004$), melanoma ($p<0.0001$), NHL ($p=0.008$), non-melanoma ($p=0.008$), prostate ($p<0.0001$), rectosigmoid ($p=0.05$), skin ($p=0.04$), and small intestine cancer ($p=0.03$).

The subgroup analysis conducted for most cancer types revealed that a single study assessed the association of interest in one or more regions. Multiple studies were only found for the associations between diabetes and the risk of breast, CRC, and prostate cancer in specific regions. Therefore, additional research is needed in various regions to validate these findings.

The results indicate that women from America and Asia with T2DM are at a higher risk of developing breast cancer, while a decreased risk is observed in the European population, albeit without statistical significance. As for CRC, all three populations exhibit a statistically significant increased risk of this cancer, with different magnitudes of effects across regions.

The pooled estimates for prostate cancer suggest that men with T2DM from Asia and Europe have a comparable decreased risk of developing this cancer. The decreased risk is not as strong and is not statistically significant in the American population.

Table C1-5 Subgroup analyses by geographic region (America, Asia, Europe) from the unadjusted (association) meta-analyses of T2DM on cancer risk

Type of cancer	RR (95% CI)	p	I ²	p subgroup
Bladder				
America	2.05 (1.24, 3.38)	0.005	97%	0.07
Asia	1.54 (0.90, 2.62)	0.11	99%	
Europe	1.18 (1.10, 1.28)	<0.00001	21%	
Brain/NS				
America	0.92 (0.72, 1.17)	0.5	N/A	0.65
Asia	1.50 (0.53, 4.23)	0.45	89%	
Europe	0.82 (0.30,2.27)	0.7	N/A	
Breast				
America	1.20 (0.93, 1.54)	0.16	98%	0.02
Asia	1.62 (1.23, 2.13)	0.0007	99%	
Europe	0.70 (0.40, 1.21)	0.20	96%	
BTC				
America	NR	N/A	N/A	<0.00001
Asia	1.08 (0.98,1.19)	0.12	N/A	
Europe	1.69 (1.43,2.01)	<0.00001	N/A	
Cervix uteri				
America	1.13 (0.85, 1.48)	0.4	N/A	0.76
Asia	1.03 (0.84, 1.27)	0.75	26%	
Europe	0.60 (0.08,4.45)	0.62	N/A	
Colon				
America	1.49 (1.35, 1.64)	<0.00001	54%	0.20
Asia	2.17 (1.21, 3.90)	0.009	99%	
Europe	1.31 (1.08, 1.58)	0.006	86%	
CRC				
America	1.61 (1.00, 2.58)	0.05	100%	0.01
Asia	1.92 (1.33, 2.77)	0.0005	98%	
Europe	1.15 (1.05, 1.25)	0.001	68%	
Endometrium				
America	2.59 (1.36, 4.91)	0.004	96%	0.56
Asia	1.93 (1.41, 2.63)	<0.0001	93%	
Europe	2.59 (1.43, 4.67)	0.002	88%	
Esophagus				
America	1.51 (0.97, 2.33)	0.07	N/A	0.49
Asia	1.15 (0.76, 1.72)	0.51	88%	
Europe	1.33 (0.98, 1.81)	0.06	29%	
Gallbladder				
America	NR	N/A	N/A	0.04
Asia	2.14 (1.25, 3.67)	0.006	81%	
Europe	0.43 (0.11,1.77)	0.24	N/A	
GI				
America	NR	N/A	N/A	<0.00001
Asia	1.12 (1.06, 1.19)	<0.0001	N/A	

Type of cancer	RR (95% CI)	p	I ²	p subgroup
Europe	1.66 (1.46, 1.89)	<0.00001	N/A	
HCC				
America	1.41 (1.33, 1.49)	<0.00001	N/A	<0.00001
Asia	3.11 (1.46, 6.63)	0.003	100%	
Europe	3.12 (2.57, 3.78)	<0.00001	0%	
Kidney				
America	1.34 (1.01, 1.77)	0.04	N/A	0.01
Asia	2.11 (1.41, 3.18)	0.0003	98%	
Europe	0.73 (0.41, 1.30)	0.29	55%	
Leukemia				
America	NR	N/A	N/A	0.05
Asia	1.55 (0.59, 4.08)	0.38	97%	
Europe	0.29 (0.07, 1.17)	0.08	N/A	
Liver				
America	3.19 (2.01, 5.06)	<0.00001	69%	0.52
Asia	2.36 (1.75, 3.18)	<0.00001	99%	
Europe	2.87 (1.73, 4.74)	<0.00001	0%	
Lung				
America	1.11 (1.04, 1.19)	0.003	N/A	0.0004
Asia	1.79 (1.12, 2.87)	0.01	100%	
Europe	0.76 (0.62, 0.94)	0.01	0%	
Lymphoma				
America	NR	N/A	N/A	0.63
Asia	1.21 (0.72, 2.01)	0.47	N/A	
Europe	0.85 (0.22, 3.20)	0.81	80%	
Melanoma				
America	1.90 (1.42, 2.54)	<0.0001	N/A	<0.0001
Asia	0.93 (0.76, 1.14)	0.47	N/A	
Europe	0.60 (0.31, 1.17)	0.13	N/A	
NHL				
America	1.06 (0.98, 1.14)	0.12	N/A	0.008
Asia	1.84 (0.82, 4.11)	0.14	93%	
Europe	1.69 (1.23, 2.31)	0.001	56%	
Non-melanoma				
America	NR	N/A	N/A	0.008
Asia	1.01 (0.93, 1.10)	0.81	N/A	
Europe	0.72 (0.57, 0.91)	0.006	N/A	
Oral				
America	NR	N/A	N/A	0.12
Asia	1.90 (0.78, 4.61)	0.16	96%	
Europe	0.33 (0.05, 2.41)	0.27	N/A	
Ovary				
America	1.17 (0.95, 1.44)	0.15	17%	0.65
Asia	1.45 (0.96, 2.18)	0.07	70%	
Europe	1.28 (0.46, 3.54)	0.63	N/A	
Pancreas				
America	2.17 (1.53, 3.08)	<0.0001	86%	0.83
Asia	2.37 (1.81, 3.08)	<0.00001	97%	
Europe	2.74 (1.74, 4.31)	<0.0001	95%	
Prostate				
America	0.94 (0.67, 1.32)	0.73	99%	<0.0001
Asia	2.35 (1.40, 3.97)	0.001	99%	
Europe	0.76 (0.71, 0.82)	<0.00001	75%	
Rectosigmoid				
America	NR	N/A	N/A	0.05
Asia	1.41 (1.24, 1.61)	<0.00001	N/A	
Europe	0.58 (0.23, 1.41)	0.23	N/A	
Rectum				
America	1.34 (1.03, 1.74)	0.03	80%	0.11
Asia	0.90 (0.68, 1.18)	0.43	N/A	
Europe	1.13 (0.98, 1.30)	0.09	40%	
Skin				

Type of cancer	RR (95% CI)	p	I ²	p subgroup
America	NR	N/A	N/A	0.04
Asia	1.75 (0.90, 3.41)	0.1	N/A	
Europe	0.83 (0.68 ,1.02)	0.08	N/A	
Small intestine				
America	NR	N/A	N/A	0.03
Asia	0.75(0.22, 2.53)	0.65	N/A	
Europe	4.57 (1.45,14.36)	0.009	N/A	
Stomach				
America	NR	N/A	N/A	0.08
Asia	1.48 (1.00, 2.19)	0.05	98%	
Europe	0.95 (0.70, 1.28)	0.73	0%	
Thyroid				
America	1.42 (1.00, 2.03)	0.05	N/A	0.26
Asia	1.12 (0.90, 1.40)	0.31	75%	
Europe	NR	N/A	N/A	

Unadjusted RR estimates in each subgroup. P-values in the third column indicate if the subgroup-specific effects are significant ($p < 0.05$). I², indicate heterogeneity in each subgroup; values are N/A for subgroups in which only 1 study or no study (NR) was found, and therefore no subgroup I² value could be calculated. P subgroup represents the p-value for the test of subgroup differences (statistical significance [$p < 0.05$] is indicated in blue). For pancreatic cancer, the study published by Davis et. al 2022 was based on an Australian cohort (HR 2.06 [1.19, 3.47]); the test for subgroup differences for this type of cancer considers these data but they are not shown in the table (Davis & Davis, 2022). *Abbreviations.* Brain/NS: Brain/Nervous system; BTC: Biliary tract cancer; CI: confidence interval; CRC: colorectal cancer; GI: gastrointestinal; HCC: hepatocellular carcinoma; NHL: non-Hodgkin lymphoma; N/A: not applicable; NR: not reported; RR: relative risk.

Menopausal status

The subgroup analysis by menopausal status showed no statistically significant difference ($p \geq 0.05$) in the risk of developing breast cancer associated with T2DM between premenopausal and postmenopausal women ($p = 0.09$). These results suggest that menopausal status does not significantly modify the association between T2DM and breast cancer risk (Table C1-6).

Table C1-6 Subgroup analyses by menopausal status (premenopausal and postmenopausal) from the unadjusted (association) meta-analyses of T2DM on cancer risk

Menopausal status	RR (95% CI)	p	I ²	p subgroup
Premenopausal	0.73(0.44, 1.19)	0.20	73%	0.09
Postmenopausal	1.14(0.98, 1.32)	0.08	71%	

Unadjusted RR estimates in each subgroup. P-values in the third column indicate if the subgroup-specific effects are significant ($p < 0.05$). I², indicate heterogeneity in each subgroup. P subgroup represents the p-value for the test of subgroup differences (statistical significance [$p < 0.05$] is indicated in blue).

Abbreviations. CI: confidence interval; RR: relative risk.

C2. Forest plots

The forest plots were generated with Review Manager software (RevMan, version 5.4.1)(The Cochrane Collaboration, 2022).

The following graphs show the forest plots from the unadjusted (association) meta-analyses and the adjusted (causal effect) meta-analyses of T2DM and each cancer type, classified by histological type.

C2.1 Adjusted (causal effect) meta-analyses

Lymphoma

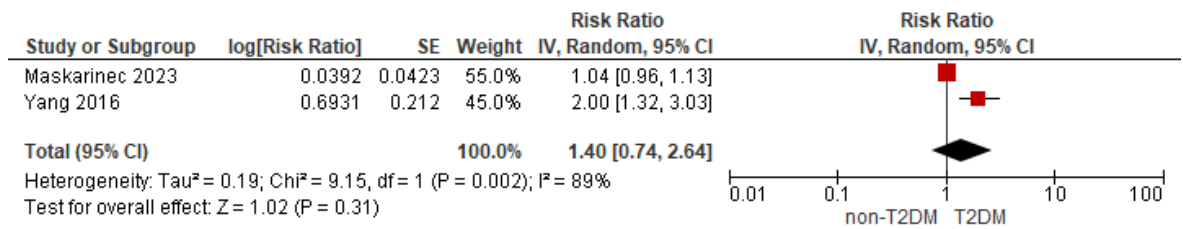


Figure C2-1 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of NHL. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant effect ($p \geq 0.05$) between T2DM and risk of NHL (summary RR 1.40; 95% CI, 0.74-2.64). Test for heterogeneity among studies: $I^2 = 89\%$. *Abbreviations.* CI: confidence interval; NHL: Non-Hodgkin lymphoma; RR: risk ratio.

Carcinoma

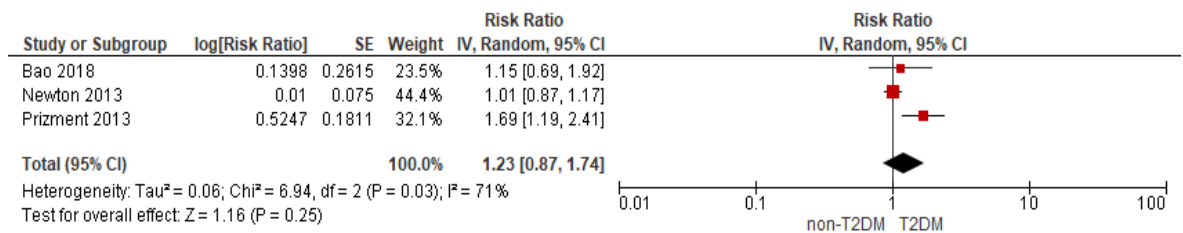


Figure C2-2 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of bladder cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant effect ($p \geq 0.05$) between T2DM and risk of bladder cancer (summary RR 1.23; 95% CI, 0.87-1.74). Test for heterogeneity among studies: $I^2 = 71\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

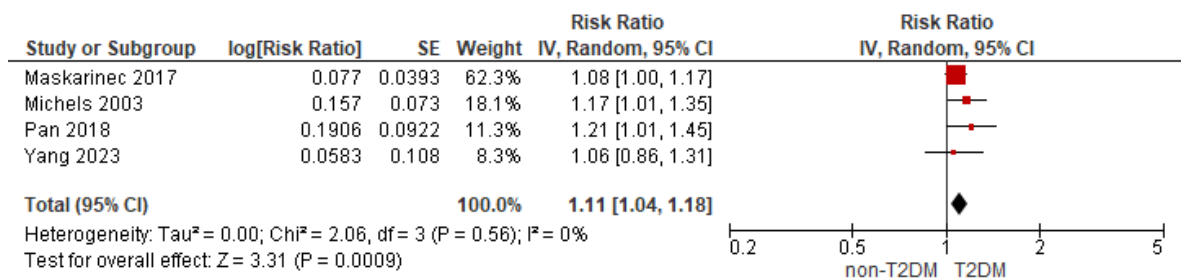


Figure C2-3 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of breast cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant effect ($p < 0.05$) between T2DM and risk of breast cancer (summary RR 1.11; 95% CI, 1.04-1.18). Test for heterogeneity among studies: $I^2 = 0\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

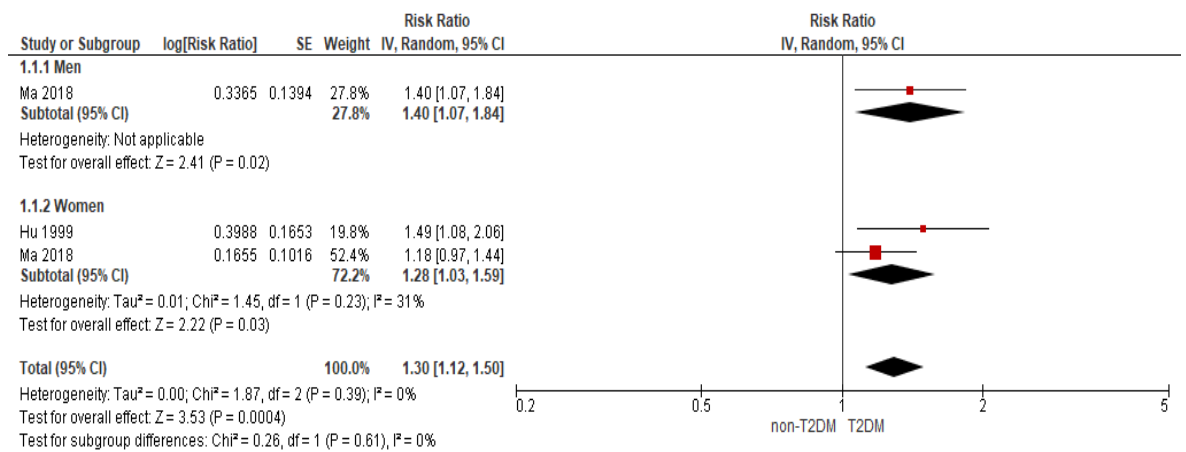


Figure C2-4 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of colon cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant effect ($p < 0.05$) between T2DM and risk of colon cancer (summary RR 1.30; 95% CI, 1.12-1.50). Test for heterogeneity among studies: $I^2 = 0\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

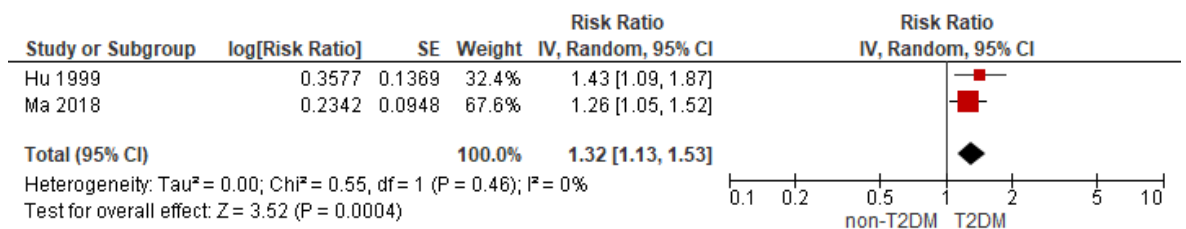


Figure C2-5 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of CRC. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant effect ($p < 0.05$) between T2DM and risk of CRC (summary RR 1.32; 95% CI, 1.13-1.53). Test for heterogeneity among studies: $I^2 = 0\%$. *Abbreviations.* CI: confidence interval; CRC: colorectal cancer; RR: risk ratio.

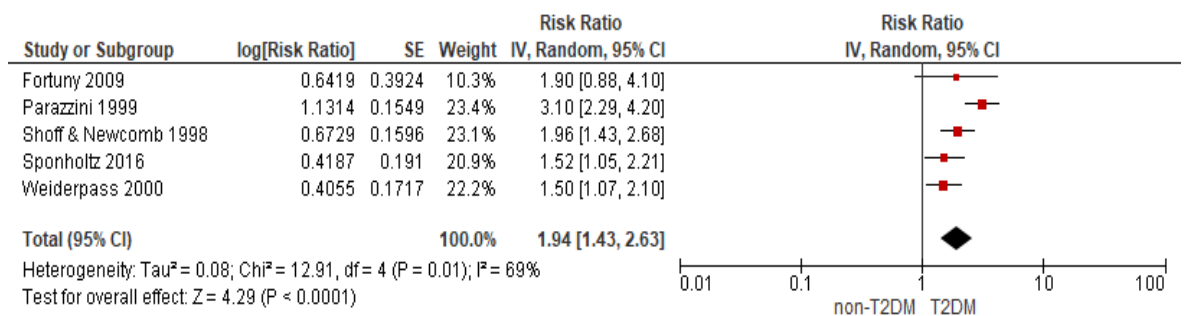


Figure C2-6 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of endometrial cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant effect ($p < 0.05$) between T2DM and risk of endometrial cancer (summary RR 1.94; 95% CI, 1.43-2.63). Test for heterogeneity among studies: $I^2 = 69\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

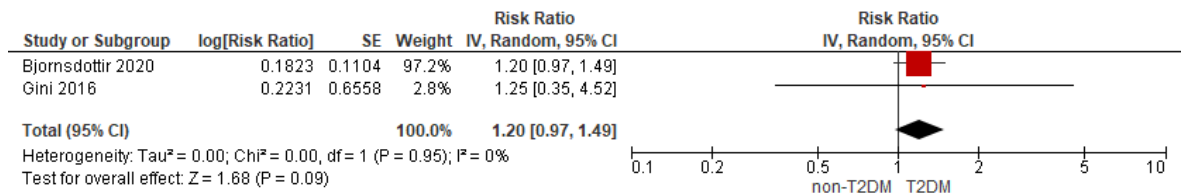


Figure C2-7 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of eye cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant effect ($p \geq 0.05$) between T2DM and risk of eye cancer (summary RR 1.20; 95% CI, 0.97-1.49). Test for heterogeneity among studies: $I^2 = 0\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

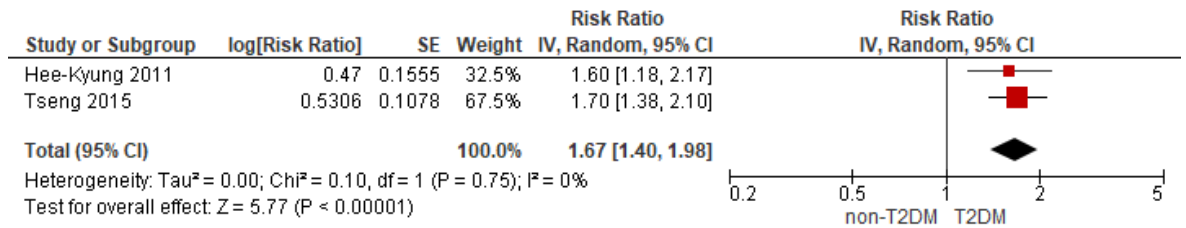


Figure C2-8 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of kidney cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant effect ($p < 0.05$) between T2DM and risk of kidney cancer (summary RR 1.67; 95% CI, 1.40-1.98). Test for heterogeneity among studies: $I^2 = 0\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

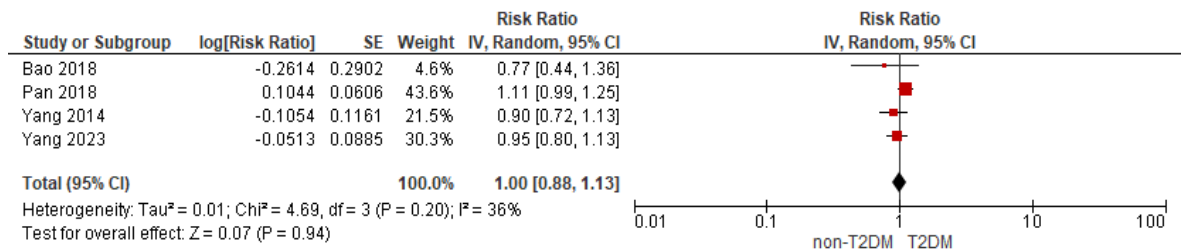


Figure C2-9 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of lung cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant effect ($p \geq 0.05$) between T2DM and risk of lung cancer (summary RR 1.00; 95% CI, 0.88-1.13). Test for heterogeneity among studies: $I^2 = 36\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

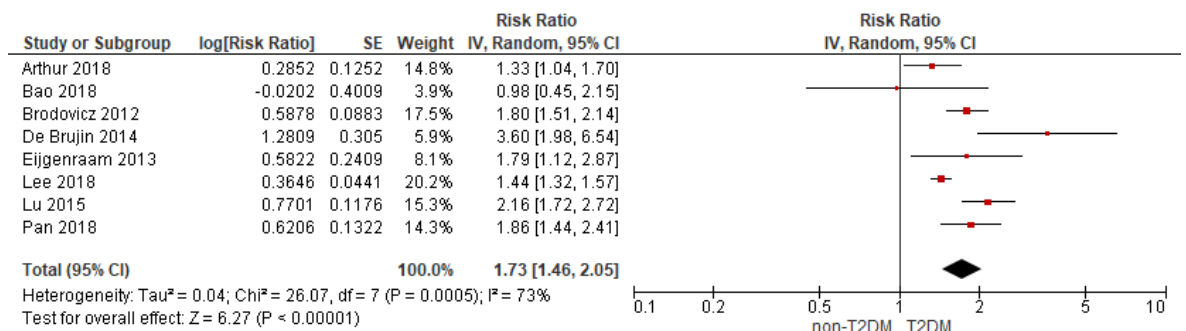


Figure C2-10 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of pancreatic cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant effect ($p < 0.05$) between T2DM and risk of pancreatic cancer (summary RR 1.73; 95% CI, 1.46-2.05). Test for heterogeneity among studies: $I^2 = 73\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

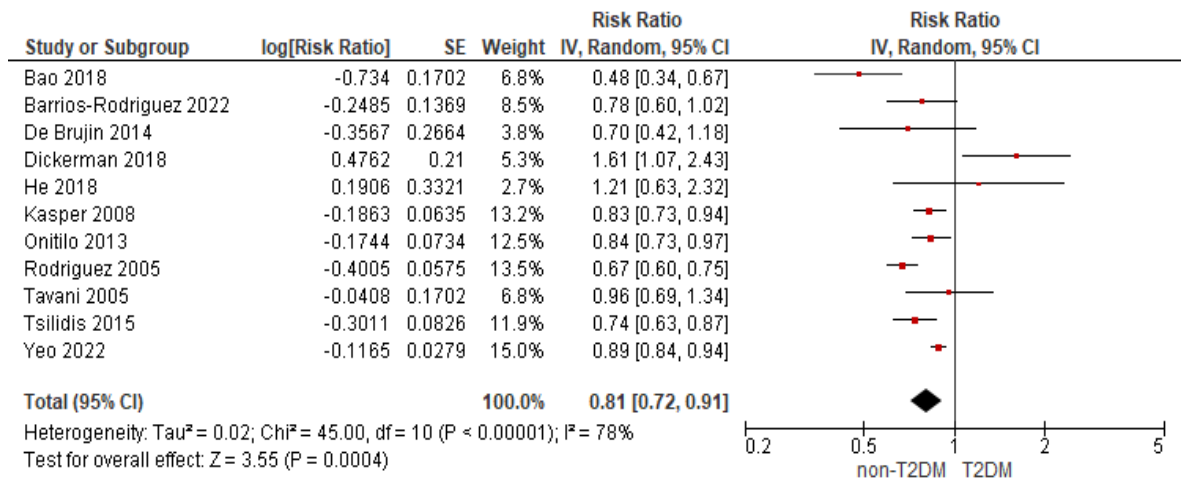


Figure C2-11 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of prostate cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant effect ($p < 0.05$) between T2DM and risk of prostate cancer (summary RR 0.81; 95% CI, 0.72-0.91). Test for heterogeneity among studies: $I^2 = 78\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

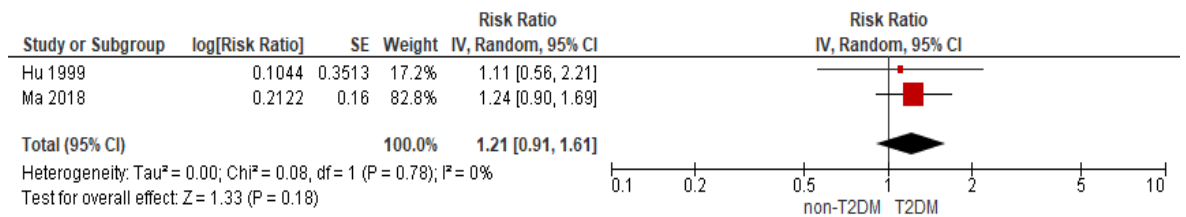


Figure C2-12 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of rectal cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant effect ($p > 0.05$) between T2DM and risk of rectal cancer (summary RR 1.21; 95% CI, 0.91-1.61). Test for heterogeneity among studies: $I^2 = 0\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

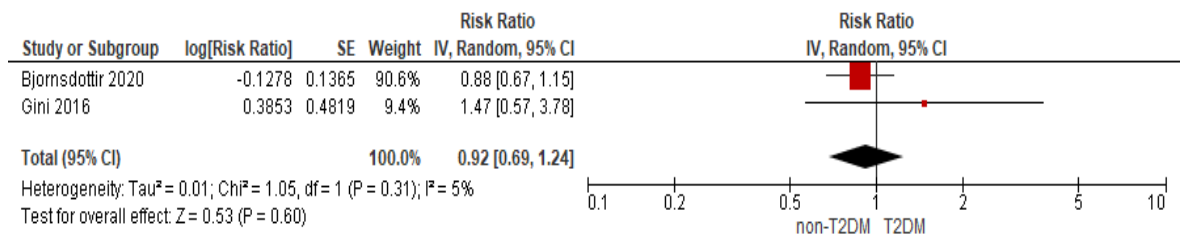


Figure C2-13 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of salivary gland cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant effect ($p > 0.05$) between T2DM and risk of salivary gland cancer (summary RR 0.92; 95% CI, 0.69-1.24). Test for heterogeneity among studies: $I^2 = 5\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

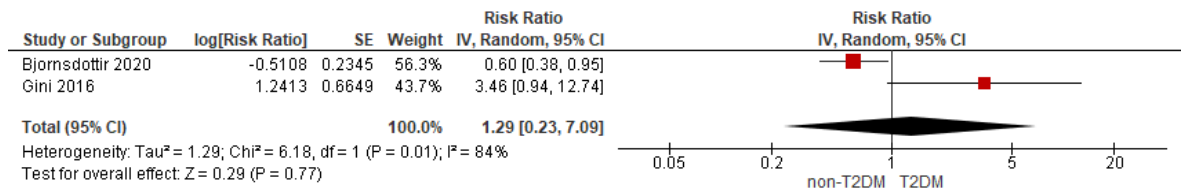


Figure C2-14 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of testicular cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant effect ($p \geq 0.05$) between T2DM and risk of testicular cancer (summary RR 1.29; 95% CI, 0.23-7.09). Test for heterogeneity among studies: $I^2 = 84\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

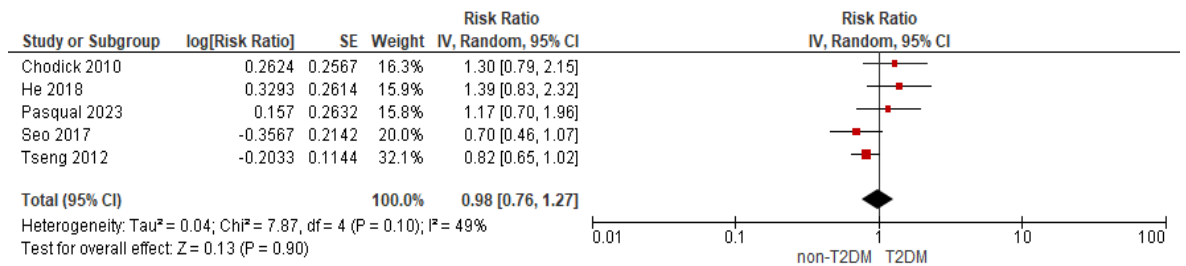


Figure C2-15 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of thyroid cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant effect ($p \geq 0.05$) between T2DM and risk of thyroid cancer (summary RR 0.98; 95% CI, 0.76-1.27). Test for heterogeneity among studies: $I^2 = 49\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

Sarcoma

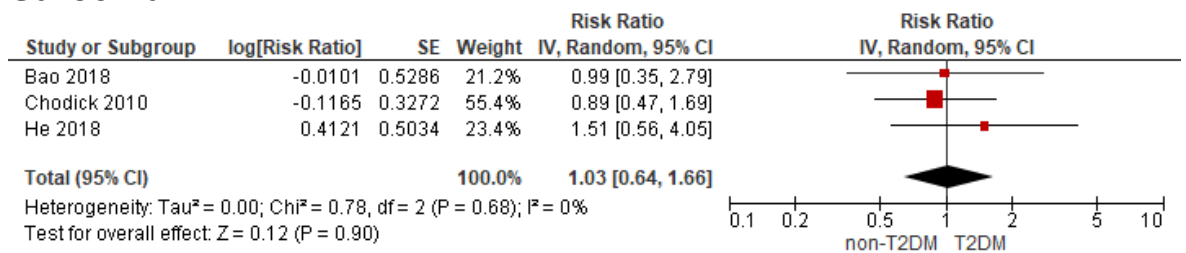


Figure C2-16 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of brain/nervous system cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant effect ($p \geq 0.05$) between T2DM and risk of brain/nervous system cancer (summary RR 1.03; 95% CI, 0.64-1.66). Test for heterogeneity among studies: $I^2 = 0\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

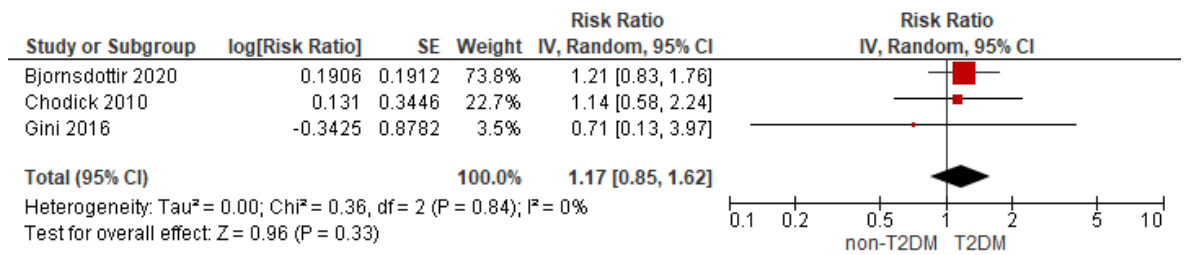


Figure C2-17 Forest plot from the adjusted (causal effect) meta-analysis of T2DM and risk of bone cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant effect ($p \geq 0.05$) between T2DM and risk of bone cancer (summary RR 1.17; 95% CI, 0.85-1.62). Test for heterogeneity among studies: $I^2 = 0\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

C2.2 Unadjusted (association) meta-analyses

Lymphoma

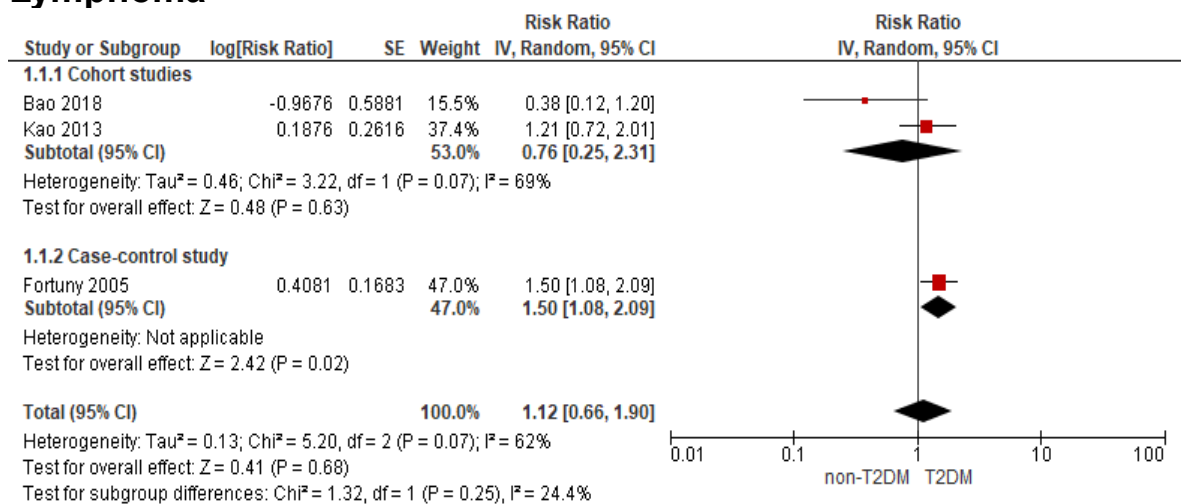


Figure C2-18 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of lymphoma. Studies are sub-grouped according to design. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant association ($p \geq 0.05$) between T2DM and risk of lymphoma (summary RR 1.12; 95% CI, 0.66-1.90). Test for heterogeneity among studies: $I^2 = 62\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

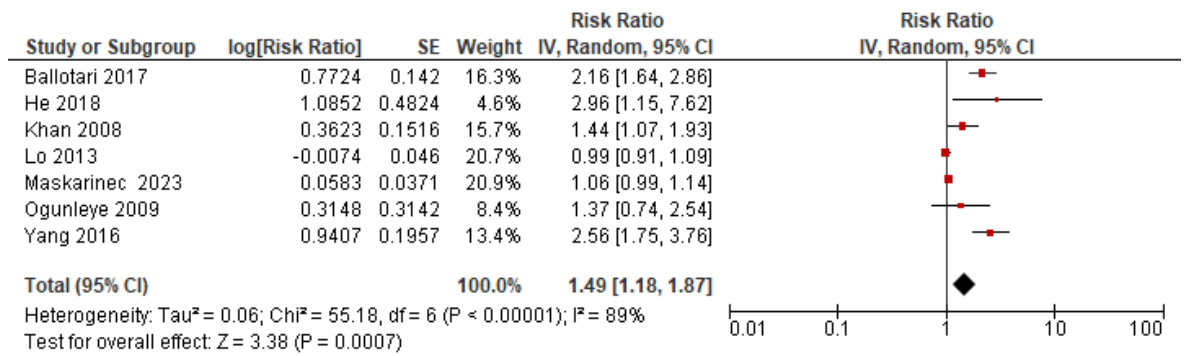


Figure C2-19 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of NHL. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant association ($p < 0.05$) between T2DM and risk of NHL (summary RR 1.49; 95% CI, 1.18-1.87). Test for heterogeneity among studies: $I^2 = 89\%$. *Abbreviations.* CI: confidence interval; NHL: Non-Hodgkin lymphoma; RR: risk ratio.

Leukemia

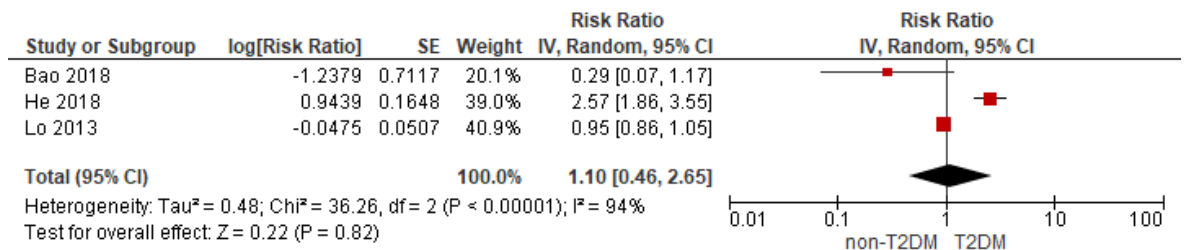


Figure C2-20 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of leukemia. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant association ($p \geq 0.05$) between T2DM and risk of leukemia (summary RR 1.10; 95% CI, 0.46-2.65). Test for heterogeneity among studies: $I^2 = 94\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

Carcinoma

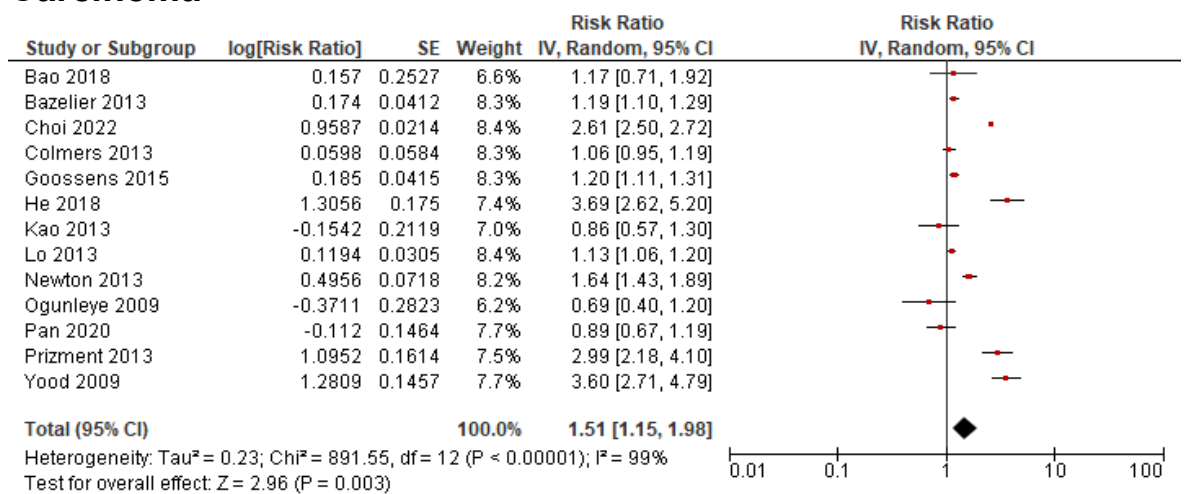


Figure C2-21 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of bladder cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant association ($p < 0.05$) between T2DM and risk of bladder cancer (summary RR 1.51; 95% CI, 1.15-1.98). Test for heterogeneity among studies: $I^2 = 99\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

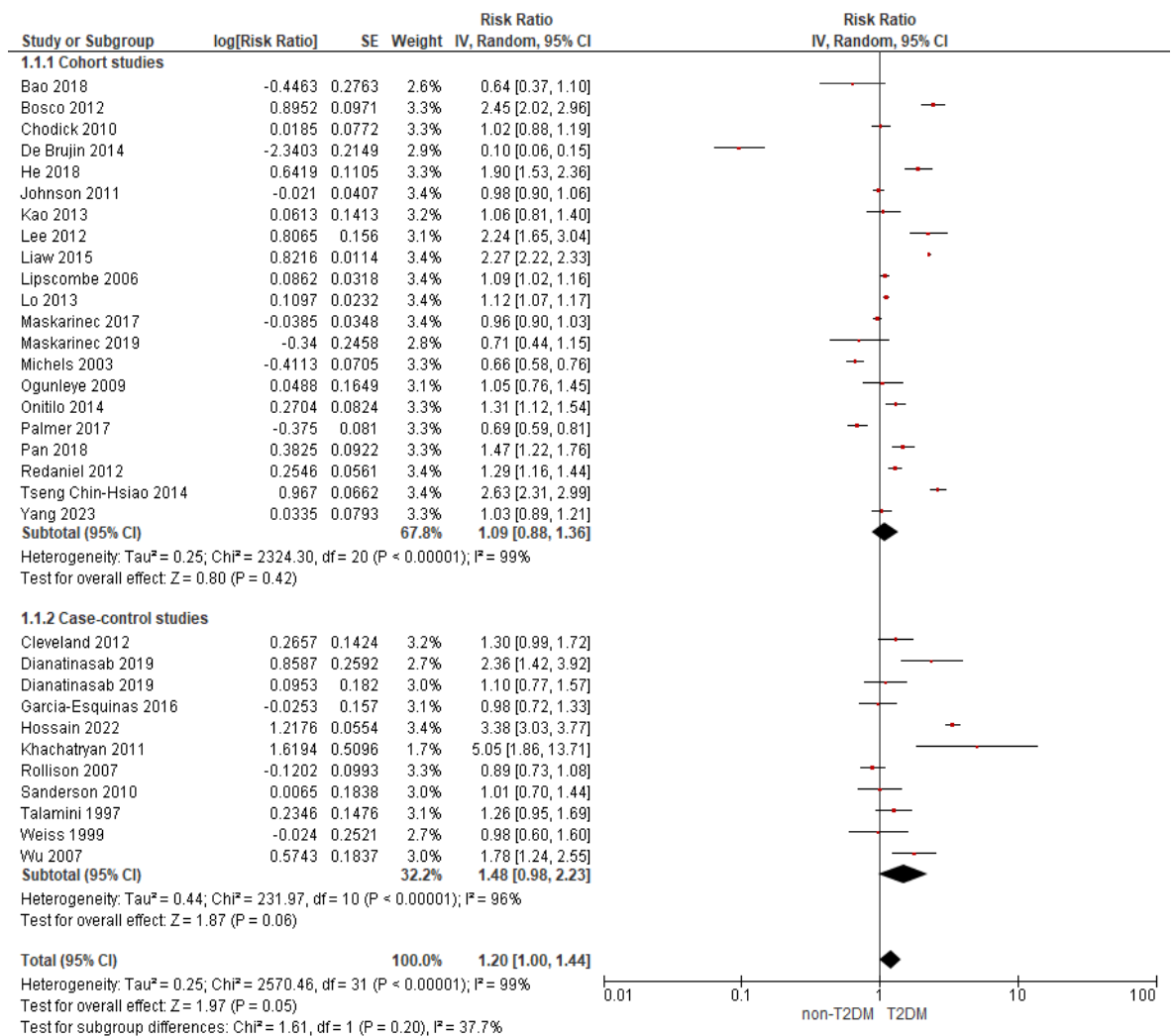


Figure C2-22 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of breast cancer. Studies are sub-grouped according to design. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant association between ($p < 0.05$) T2DM and risk of breast cancer (summary RR 1.20; 95% CI, 1.00-1.44). Test for heterogeneity among studies: $I^2 = 99\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio. Outcomes for invasive ductal carcinoma and invasive lobular carcinoma were reported separately in one study (Dianatinasab et al., 2019).

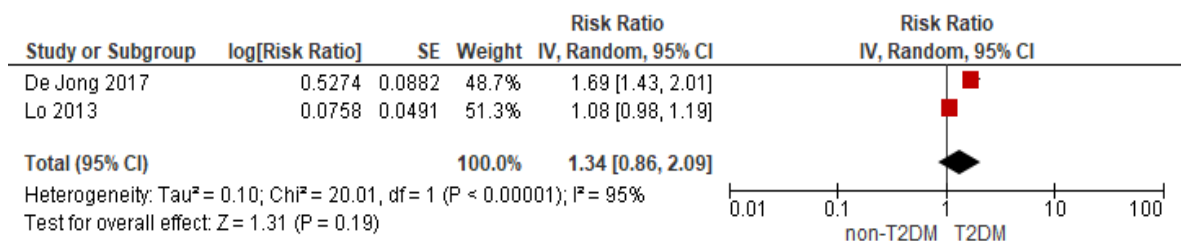


Figure C2-23 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of BTC. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant association ($p \geq 0.05$) between T2DM and risk of BTC (summary RR 1.34; 95% CI, 0.86-2.09). Test for heterogeneity among studies: $I^2 = 95\%$. *Abbreviations.* BTC: biliary tract cancer; CI: confidence interval; RR: risk ratio.

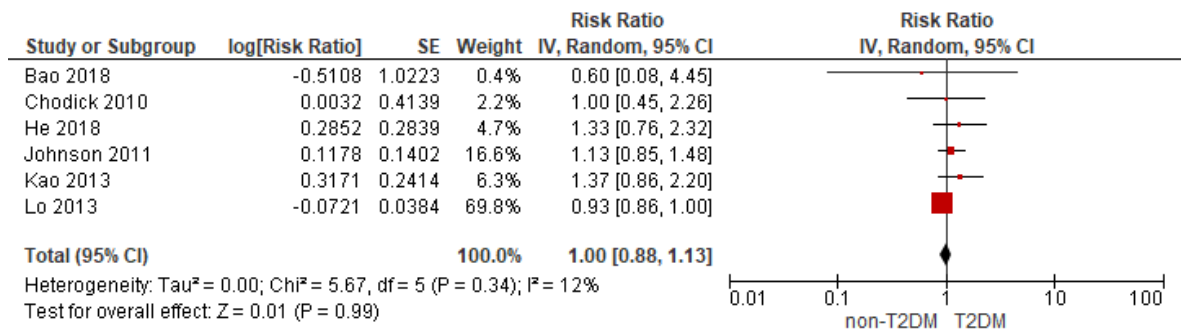


Figure C2-24 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of cervix uteri cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant association ($p \geq 0.05$) between T2DM and risk of cervix uteri cancer (summary RR 1.00; 95% CI, 0.88-1.13). Test for heterogeneity among studies: $I^2 = 12\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

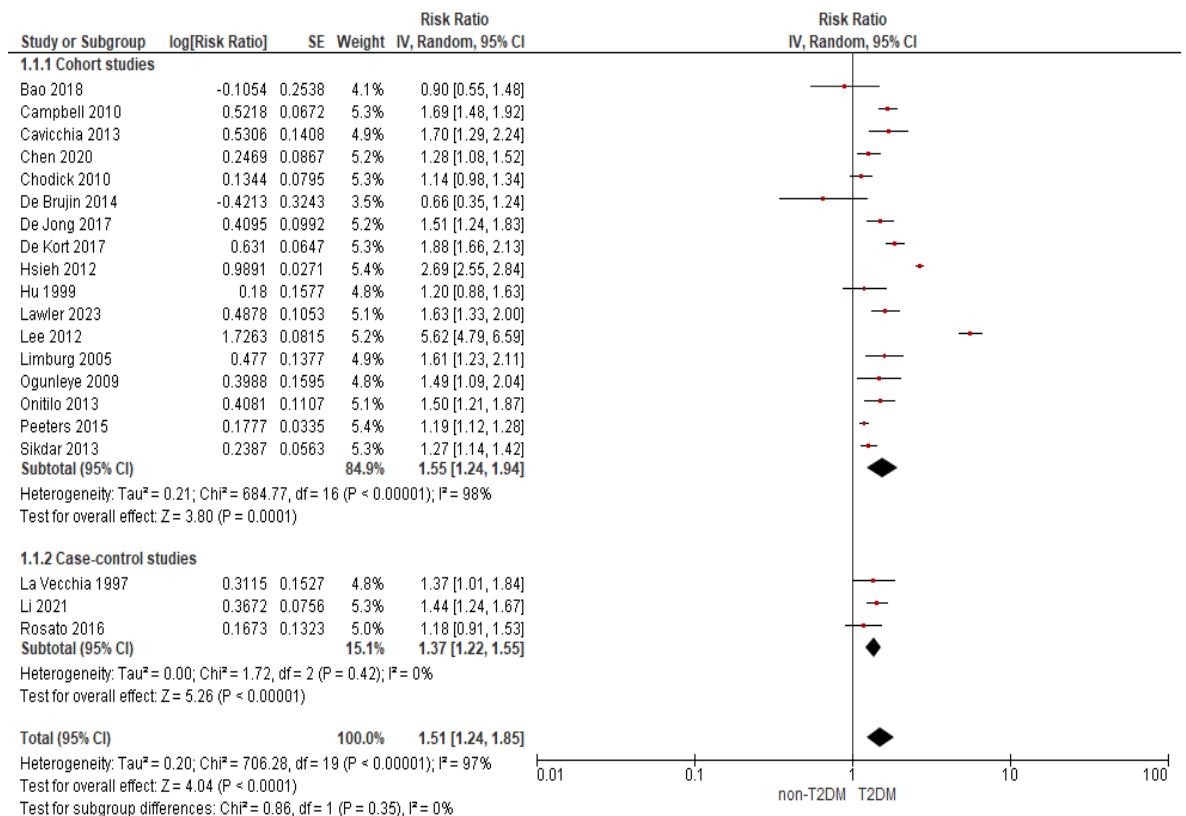


Figure C2-25 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of colon cancer. Studies are sub-grouped according to design. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant association ($p < 0.05$) between T2DM and risk of colon cancer (summary RR 1.51; 95% CI, 1.24-1.85). Test for heterogeneity among studies: $I^2 = 97\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

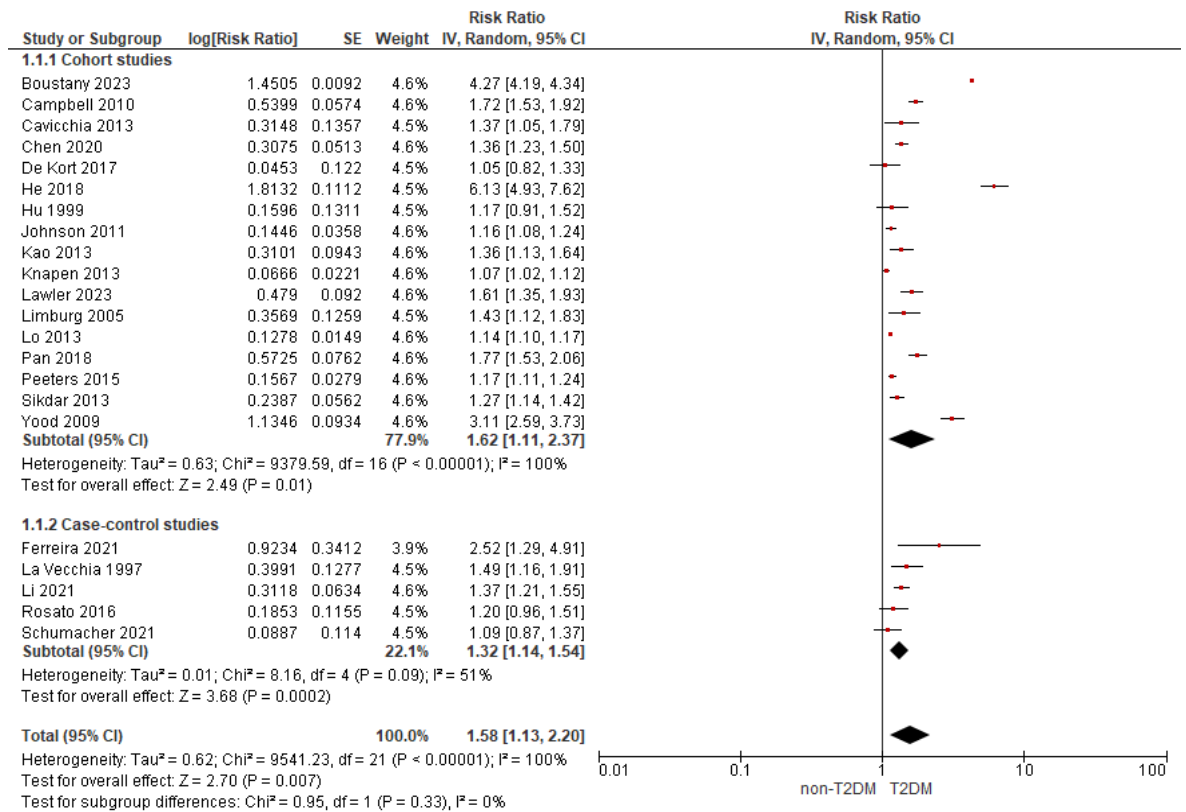


Figure C2-26 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of CRC. Studies are sub-grouped according to design. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant association ($p < 0.05$) between T2DM and risk of CRC (summary RR 1.58; 95% CI, 1.13-2.20). Test for heterogeneity among studies: $I^2 = 100\%$. *Abbreviations.* CI: confidence interval; CRC: colorectal cancer; RR: risk ratio.

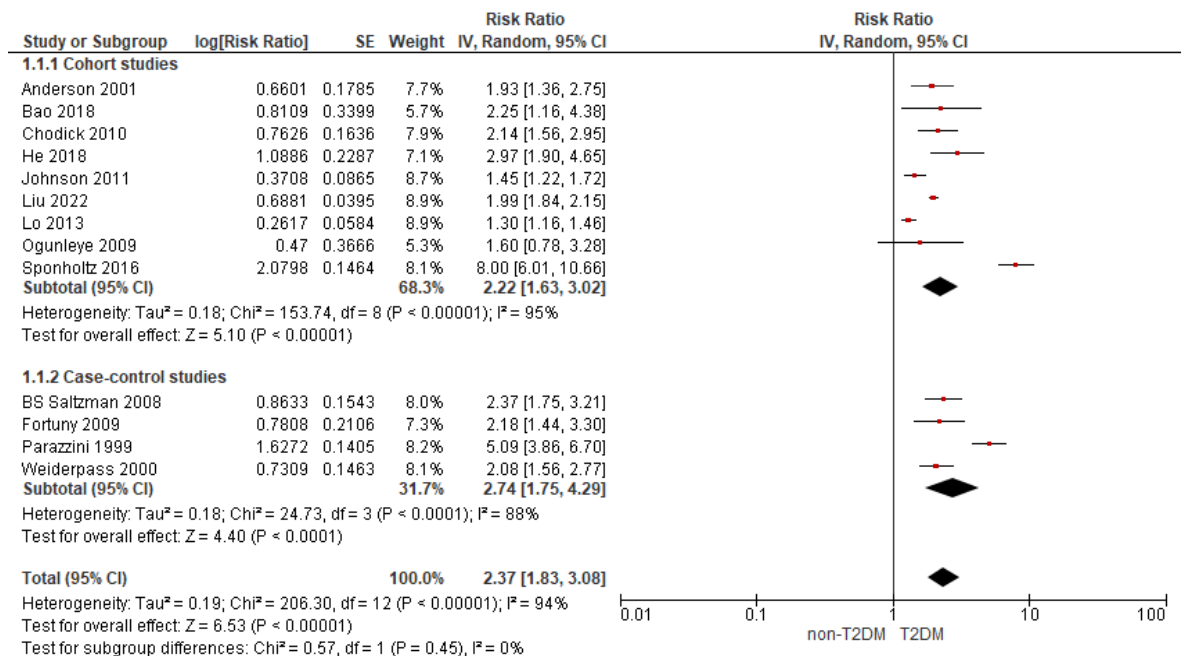


Figure C2-27 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of endometrial cancer. Studies are sub-grouped according to design. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant association ($p < 0.05$) between T2DM and risk of endometrial cancer (summary RR 2.37; 95% CI, 1.83-3.08). Test for heterogeneity among studies: $I^2 = 94\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

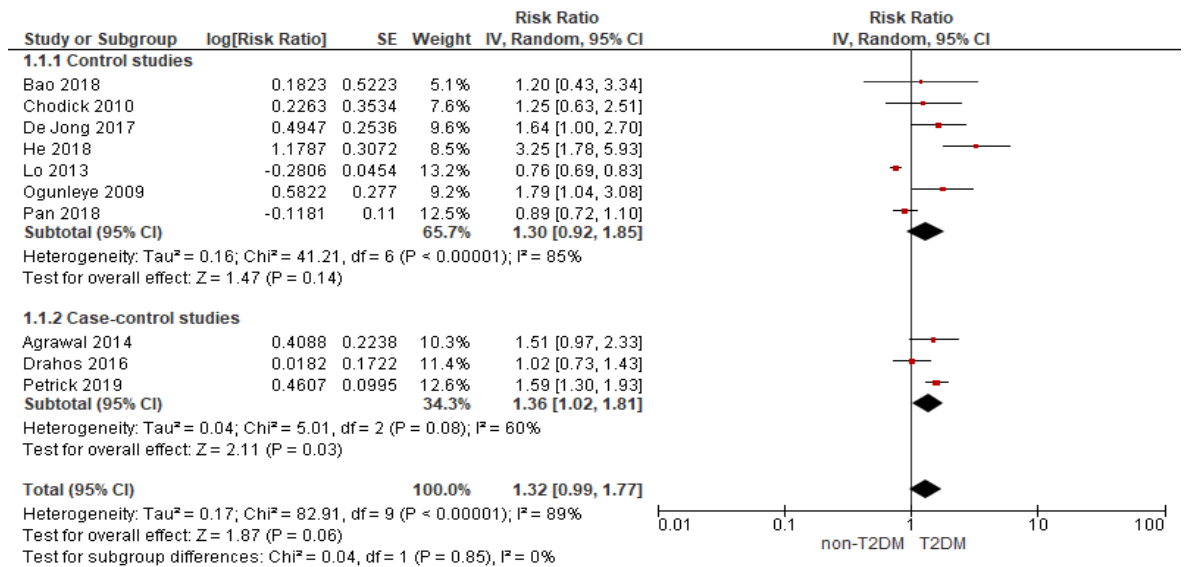


Figure C2-28 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of esophageal cancer. Studies are sub-grouped according to design. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant association ($p \geq 0.05$) between T2DM and risk of esophageal cancer (summary RR 1.32; 95% CI, 0.99-1.77). Test for heterogeneity among studies: $I^2 = 89\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

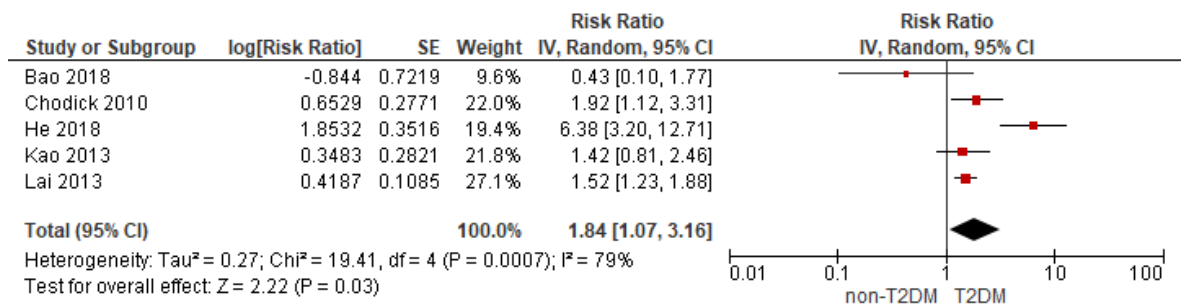


Figure C2-29 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of gallbladder cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant association ($p < 0.05$) between T2DM and risk of gallbladder cancer (summary RR 1.84; 95% CI, 1.07-3.16). Test for heterogeneity among studies: $I^2 = 79\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

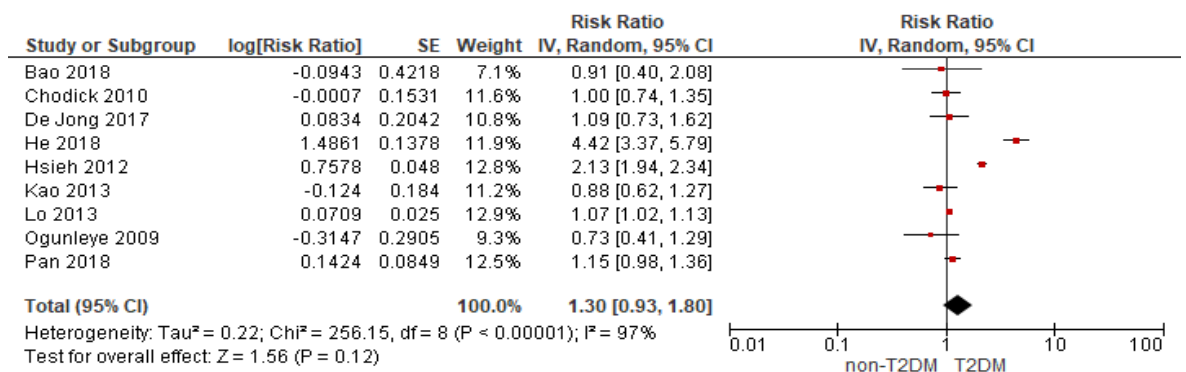


Figure C2-30 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of gastric cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant association ($p \geq 0.05$) between T2DM and risk of gastric cancer (summary RR 1.30; 95% CI, 0.93-1.80). Test for heterogeneity among studies: $I^2 = 97\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

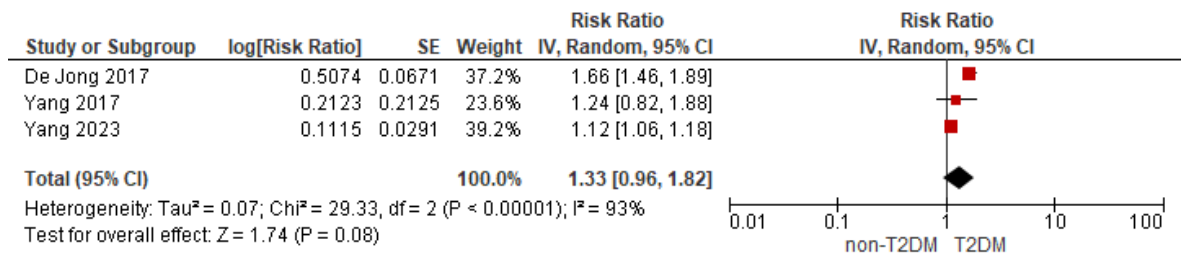


Figure C2-31 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of GI cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant association ($p \geq 0.05$) between T2DM and risk of GI cancer (summary RR 1.33; 95% CI, 0.96-1.82). Test for heterogeneity among studies: $I^2 = 93\%$. *Abbreviations.* CI: confidence interval; GI: gastrointestinal; RR: risk ratio.

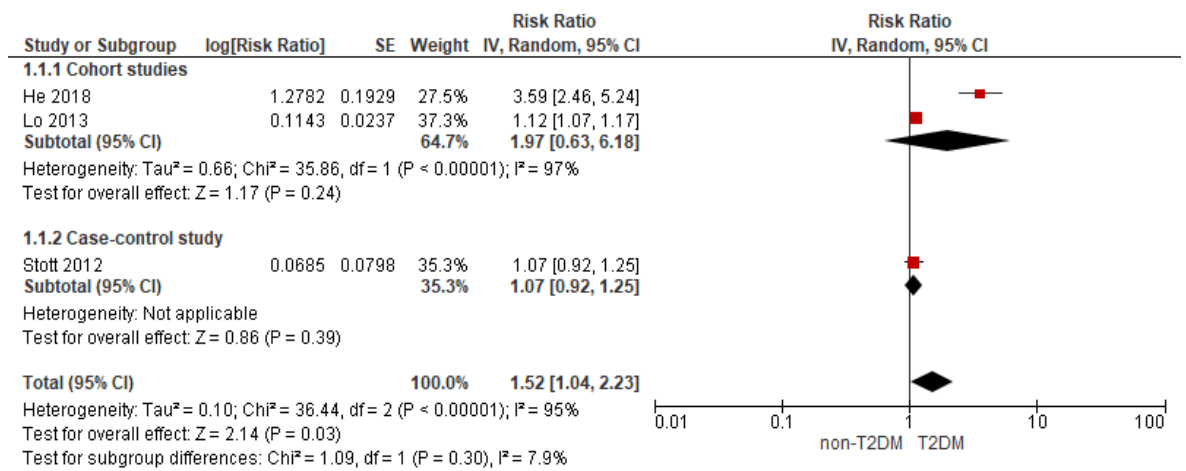


Figure C2-32 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of H&N cancer. Studies are sub-grouped according to design. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant association ($p < 0.05$) between T2DM and risk of H&N cancer (summary RR 1.52; 95% CI, 1.04-2.23). Test for heterogeneity among studies: $I^2 = 95\%$. *Abbreviations.* CI: confidence interval; H&N: head and neck; RR: risk ratio.

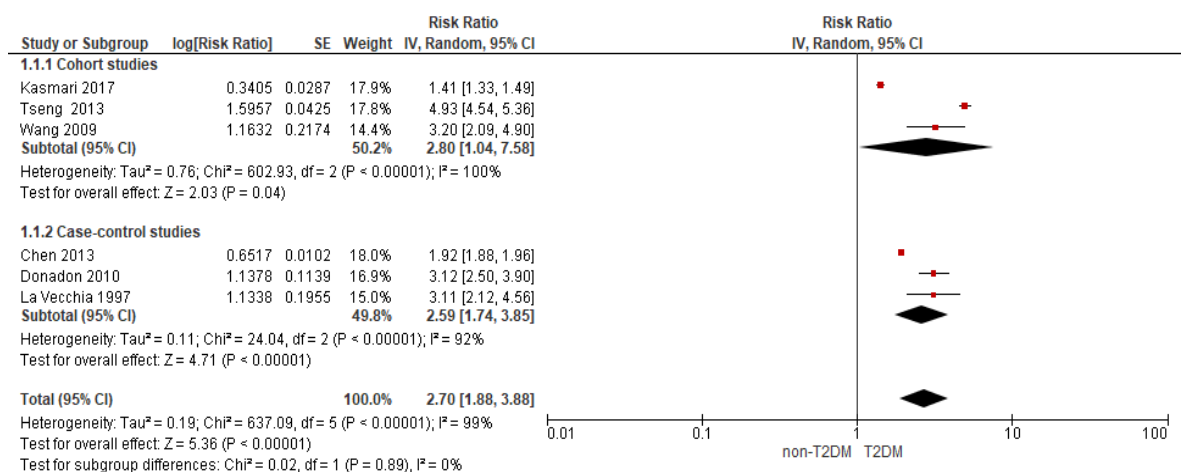


Figure C2-33 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of HCC. Studies are sub-grouped according to design. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant association ($p < 0.05$) between T2DM and risk of HCC (summary RR 2.70; 95% CI, 1.88-3.88). Test for heterogeneity among studies: $I^2 = 99\%$. *Abbreviations.* CI: confidence interval; HCC: hepatocellular carcinoma; RR: risk ratio.

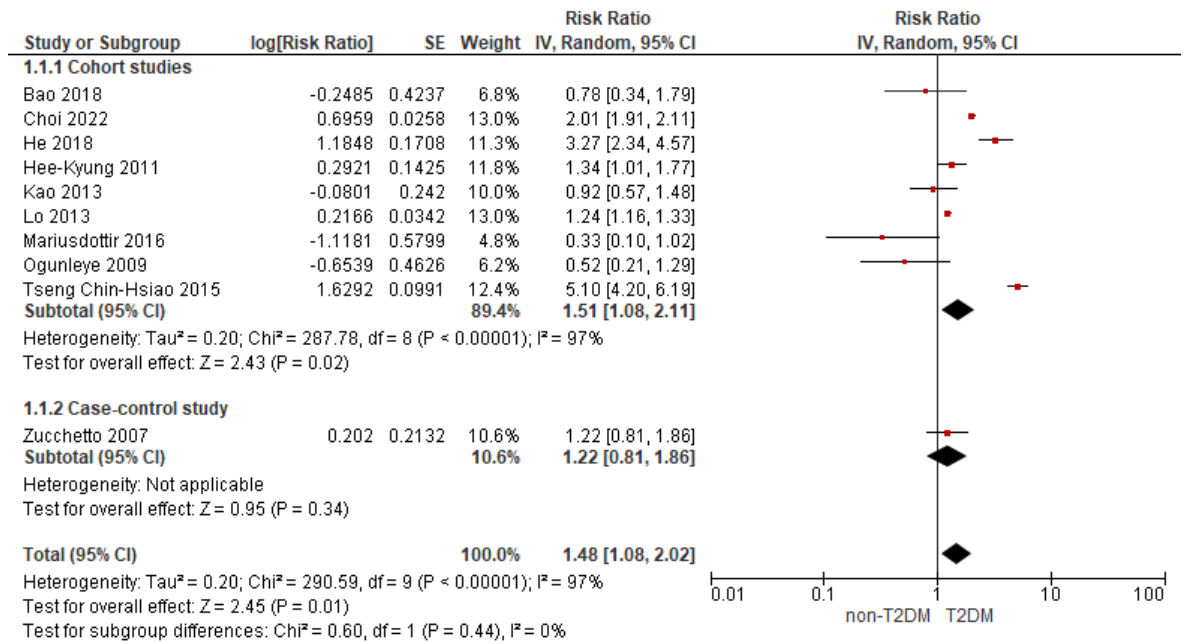


Figure C2-34 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of kidney cancer. Studies are sub-grouped according to design. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant association ($p < 0.05$) between T2DM and risk of kidney cancer (summary RR 1.48; 95% CI, 1.08-2.02). Test for heterogeneity among studies: $I^2 = 97\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

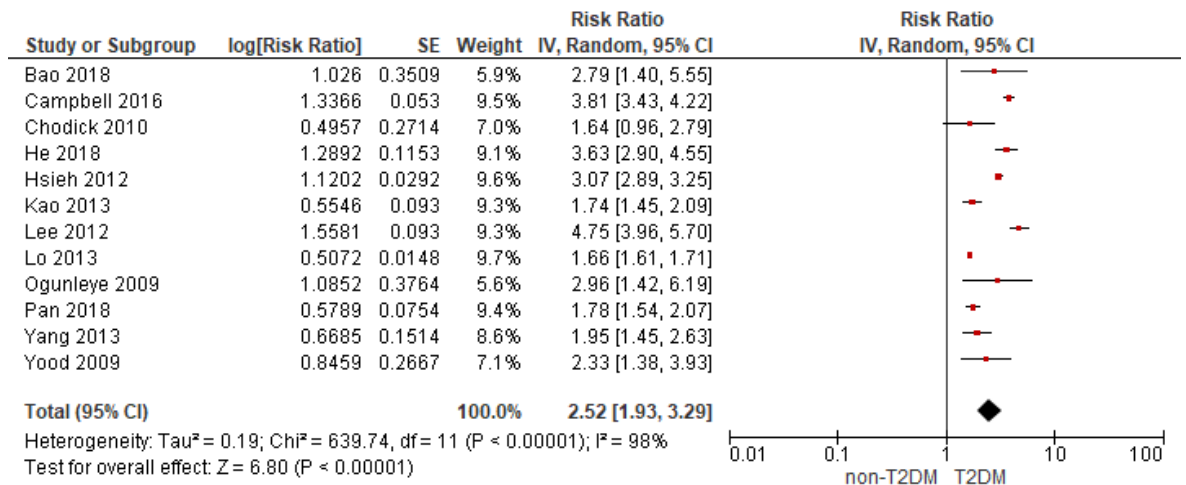


Figure C2-35 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of liver cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant association ($p < 0.05$) between T2DM and risk of liver cancer (summary RR 2.52; 95% CI, 1.93-3.29). Test for heterogeneity among studies: $I^2 = 98\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

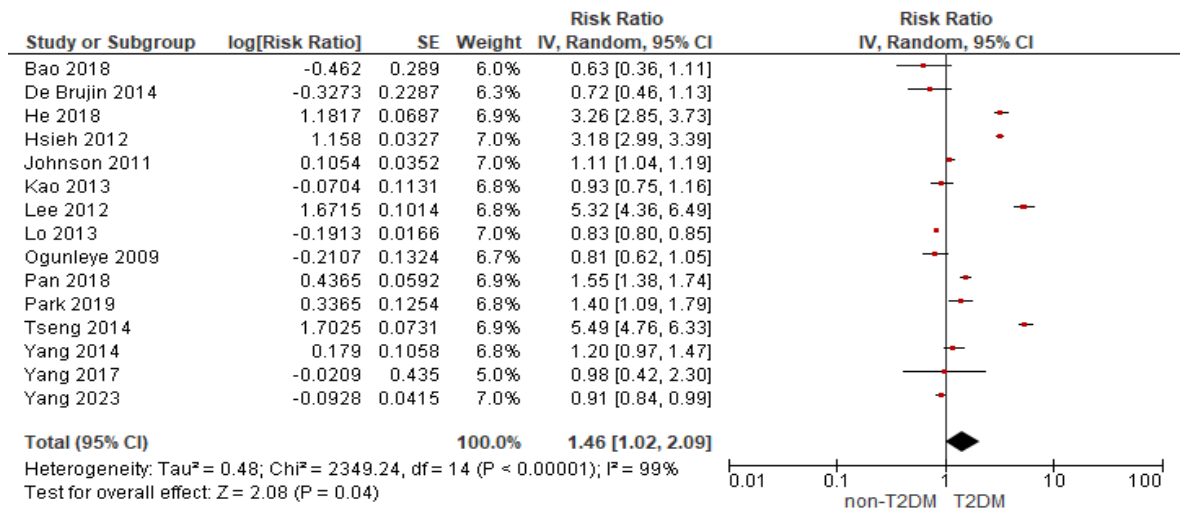


Figure C2-36 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of lung cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant association ($p < 0.05$) between T2DM and risk of lung cancer (summary RR 1.46; 95% CI, 1.02-2.09). Test for heterogeneity among studies: $I^2 = 99\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

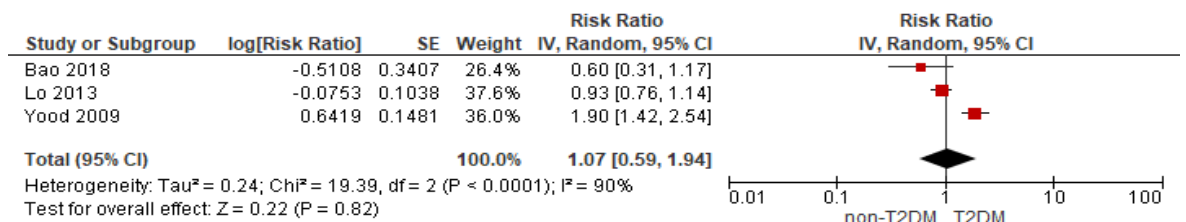


Figure C2-37 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of melanoma. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant association ($p \geq 0.05$) between T2DM and risk of melanoma (summary RR 1.07; 95% CI, 0.59-1.94). Test for heterogeneity among studies: $I^2 = 90\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

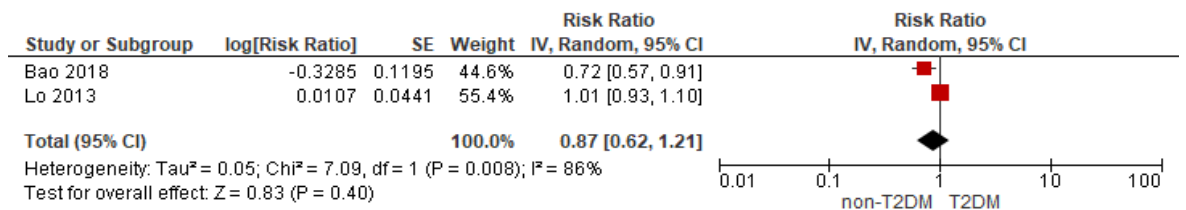


Figure C2-38 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of non-melanoma. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant association ($p \geq 0.05$) between T2DM and risk of non-melanoma (summary RR 0.87; 95% CI, 0.62-1.21). Test for heterogeneity among studies: $I^2 = 86\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

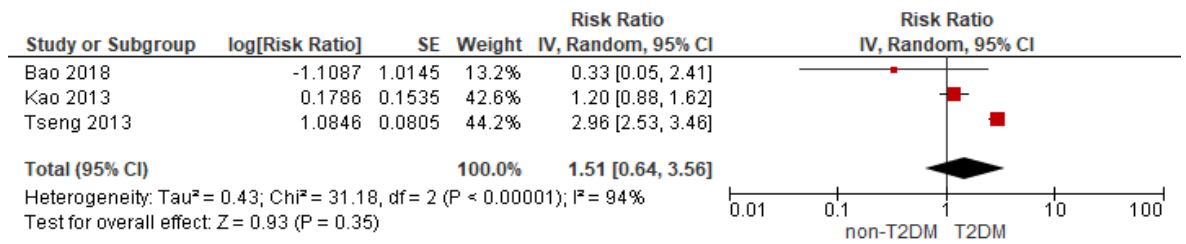


Figure C2-39 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of oral cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant association ($p \geq 0.05$) between T2DM and risk of oral cancer (summary RR 1.51; 95% CI, 0.64-3.56). Test for heterogeneity among studies: $I^2 = 94\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

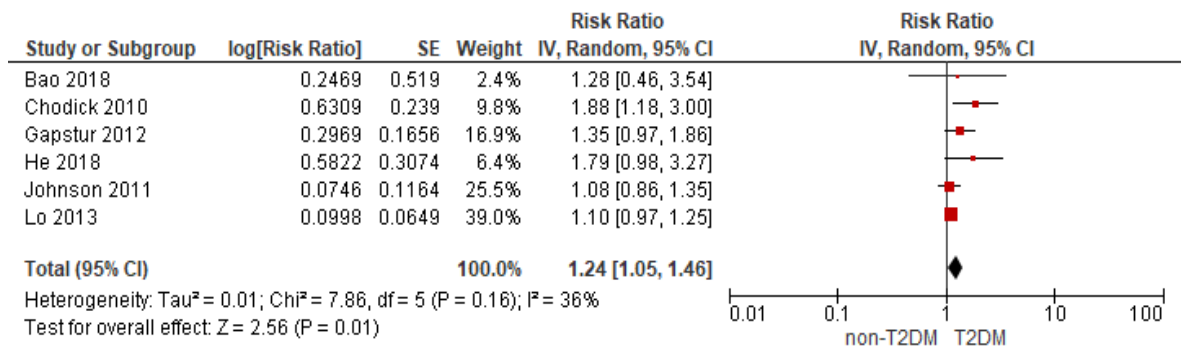


Figure C2- Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of ovarian cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant association ($p < 0.05$) between T2DM and risk of ovarian cancer (summary RR 1.24; 95% CI, 1.05-1.46). Test for heterogeneity among studies: $I^2 = 36\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

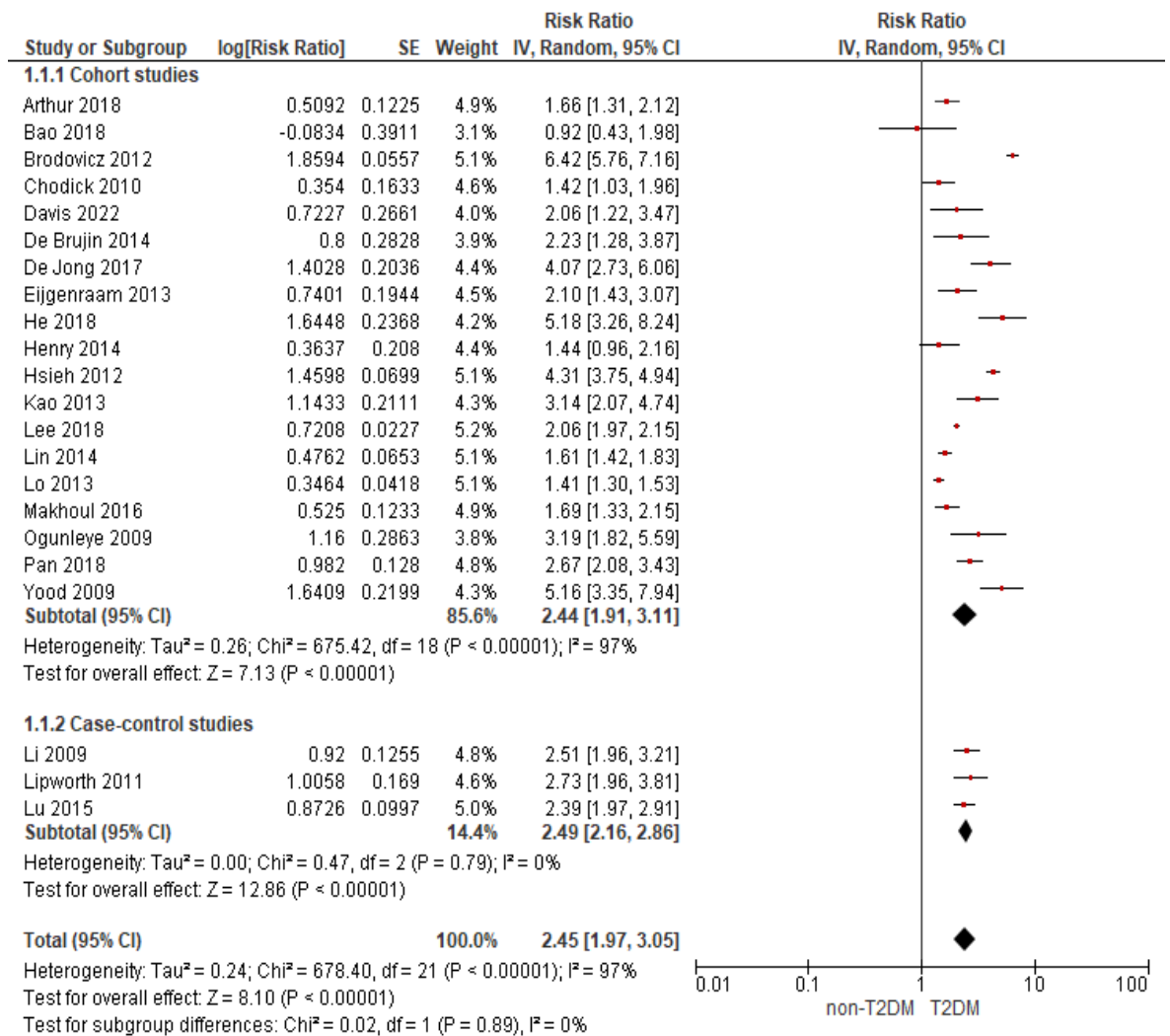


Figure C2-41 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of pancreatic cancer. Studies are sub-grouped according to design. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant association ($p < 0.05$) between T2DM and risk of pancreatic cancer (summary RR 2.45; 95% CI, 1.97-3.05). Test for heterogeneity among studies: $I^2 = 97\%$. Abbreviations: CI: confidence interval; RR: risk ratio.

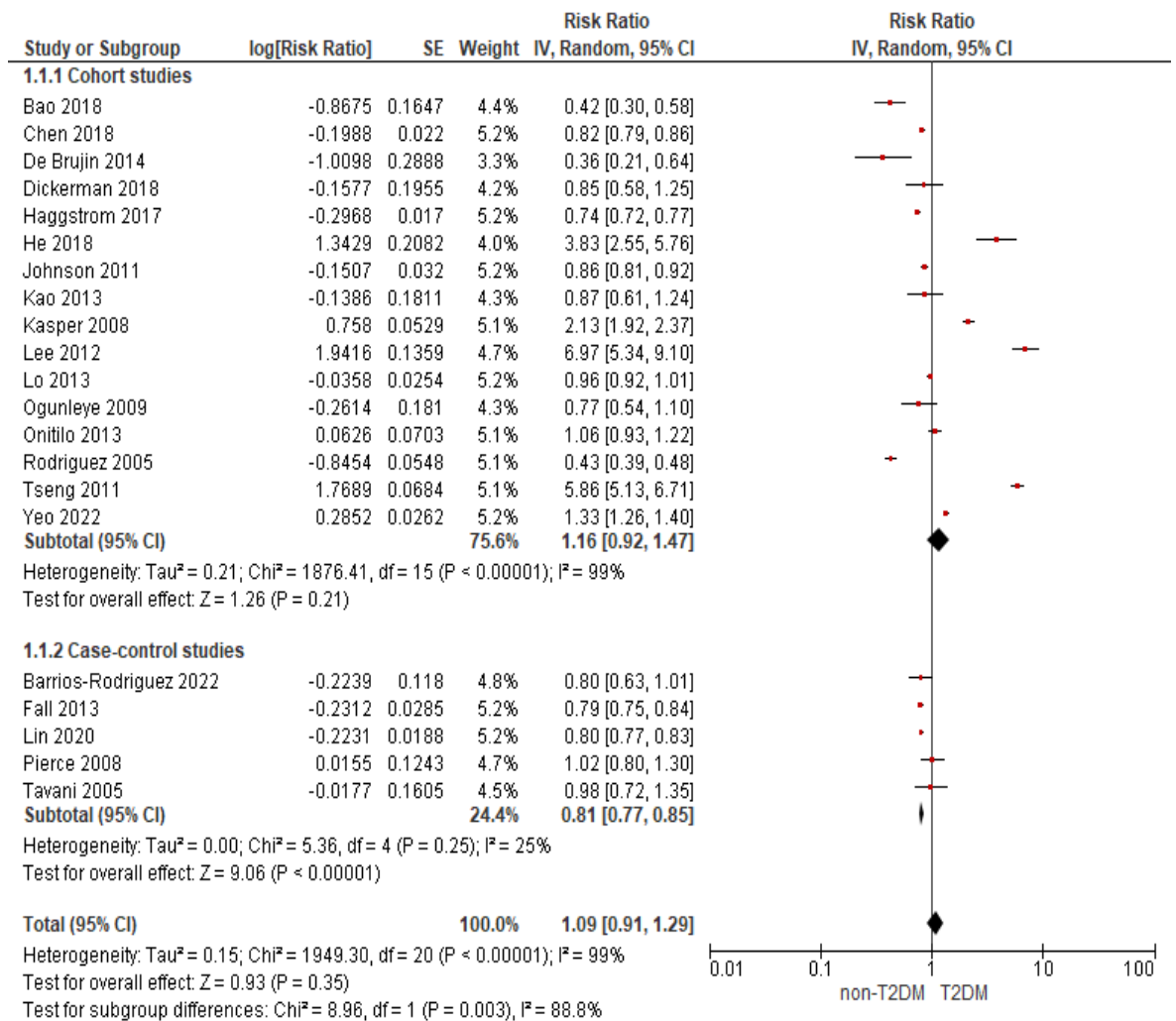


Figure C2-42 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of prostate cancer. Studies are sub-grouped according to design. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant association ($p \geq 0.05$) between T2DM and risk of prostate cancer (summary RR 1.09; 95% CI, 0.91-1.29). Test for heterogeneity among studies: $I^2 = 99\%$. Abbreviations. CI: confidence interval; RR: risk ratio.

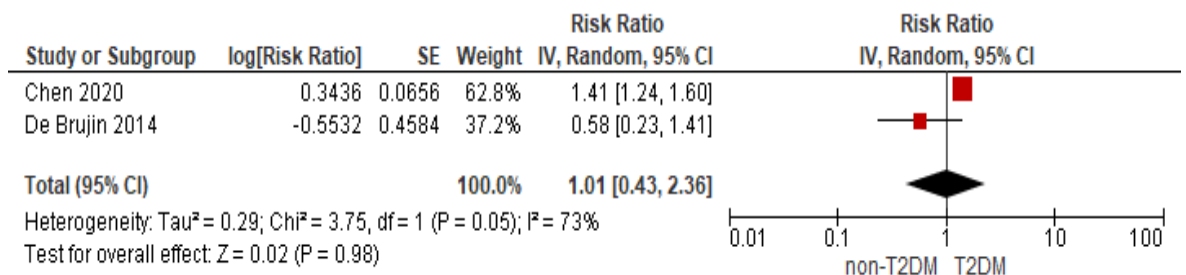


Figure C2-43 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of rectosigmoid cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant association ($p \geq 0.05$) between T2DM and risk of rectosigmoid cancer (summary RR 1.01; 95% CI, 0.43-2.36). Test for heterogeneity among studies: $I^2 = 73\%$. Abbreviations. CI: confidence interval; RR: risk ratio.

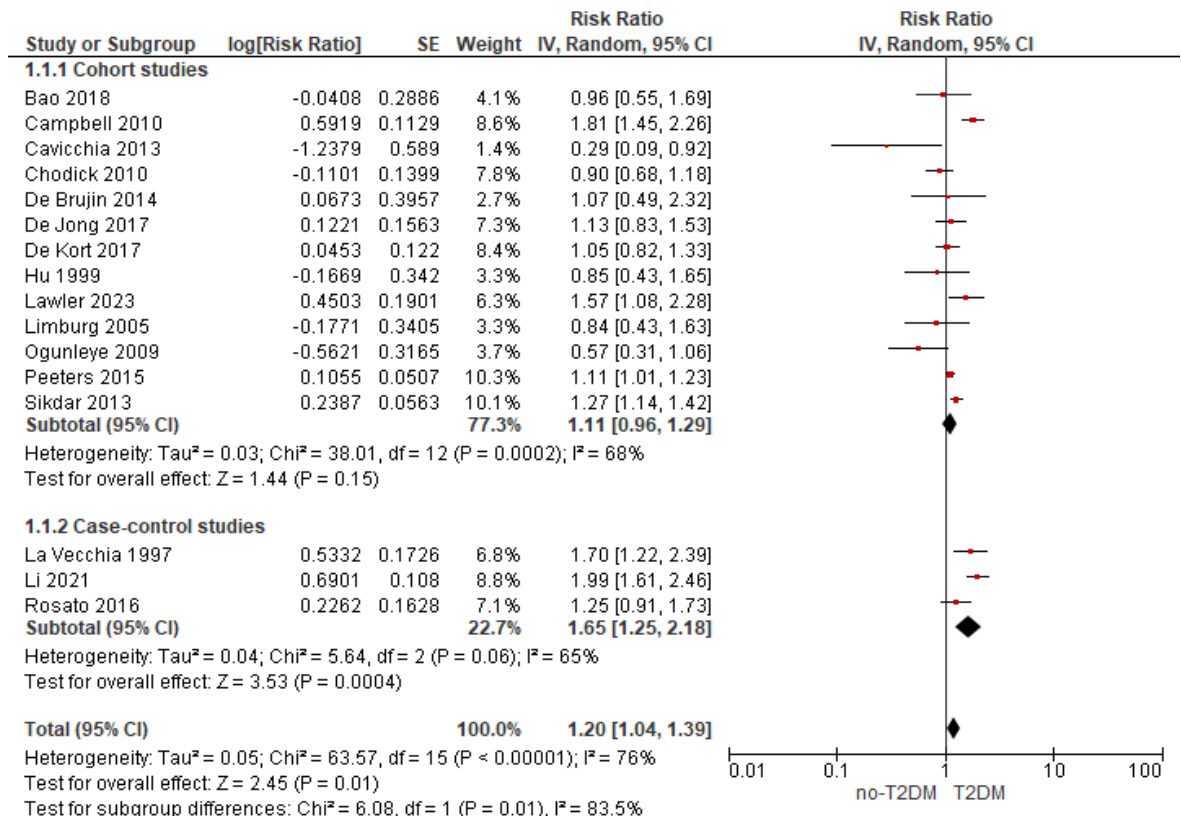


Figure C2-44 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of rectal cancer. Studies are sub-grouped according to design. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows a statistically significant association ($p < 0.05$) between T2DM and risk of rectal cancer (summary RR 1.20; 95% CI, 1.04-1.39). Test for heterogeneity among studies: $I^2 = 76\%$. Abbreviations. CI: confidence interval; RR: risk ratio.

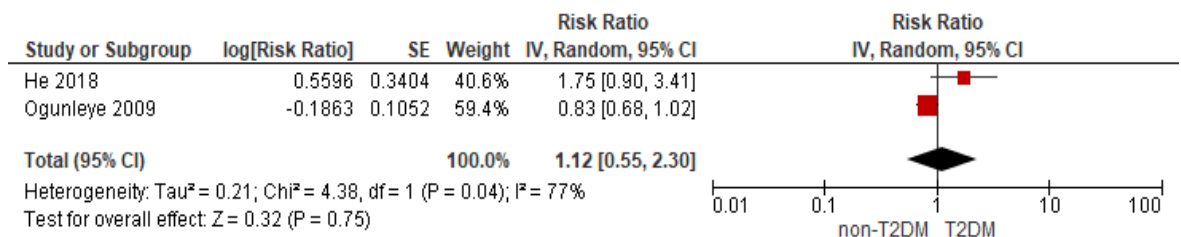


Figure C2-45 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of skin cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant association ($p \geq 0.05$) between T2DM and risk of skin cancer (summary RR 1.12; 95% CI, 0.55-2.30). Test for heterogeneity among studies: $I^2 = 77\%$. Abbreviations. CI: confidence interval; RR: risk ratio.

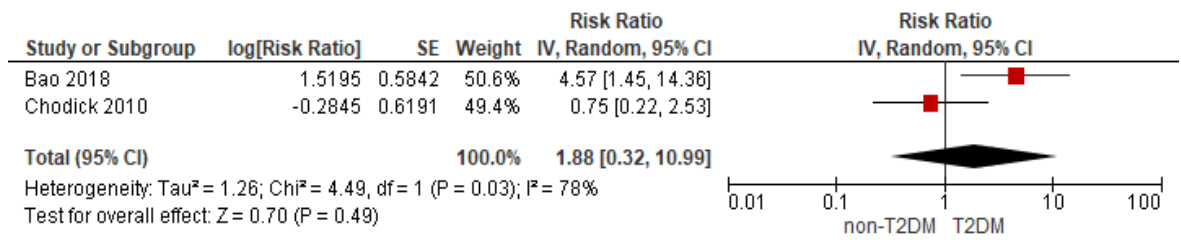


Figure C2-46 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of small intestine cancer. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant association ($p \geq 0.05$) between T2DM and risk of small intestine cancer (summary RR 1.88; 95% CI, 0.32-10.99). Test for heterogeneity among studies: $I^2 = 78\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

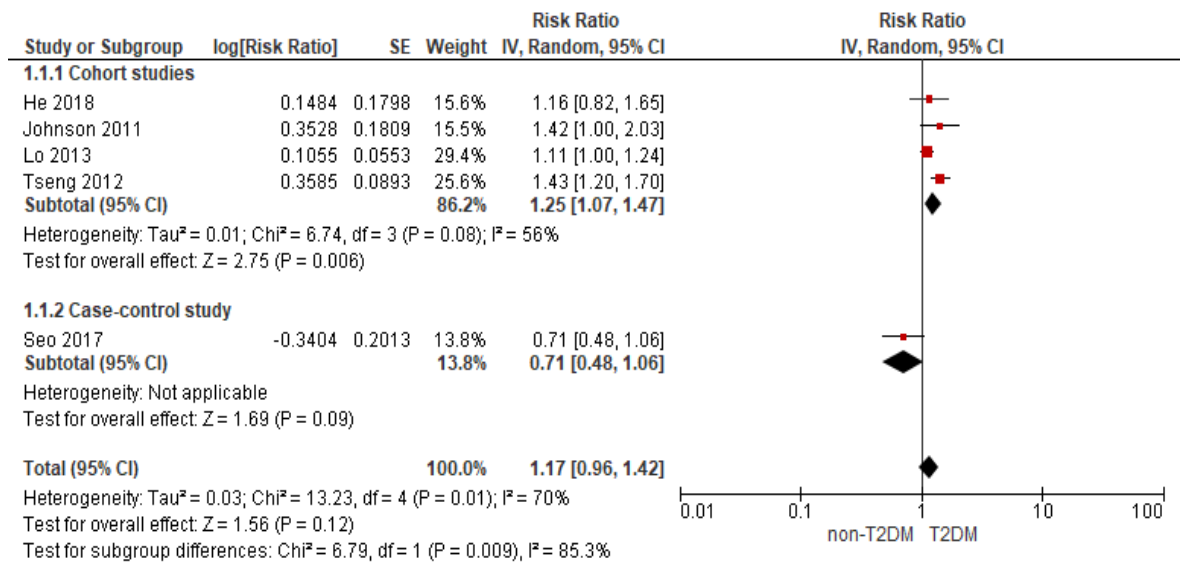


Figure C2-47 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of thyroid cancer. Studies are sub-grouped according to design. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant association ($p \geq 0.05$) between T2DM and risk of thyroid cancer (summary RR 1.17; 95% CI, 0.96-1.42). Test for heterogeneity among studies: $I^2 = 70\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

Sarcoma

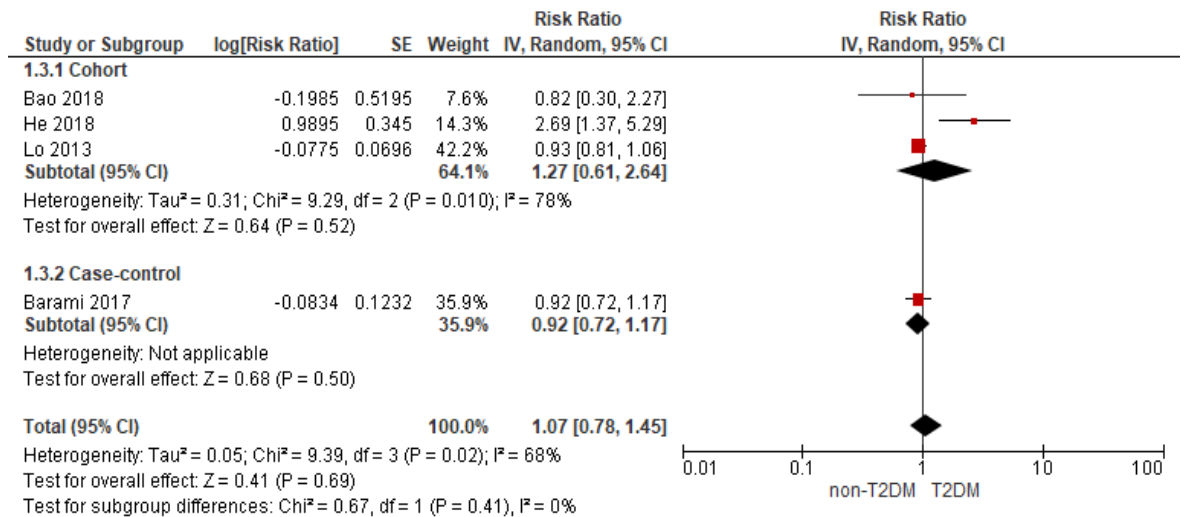


Figure C2-48 Forest plot from the unadjusted (association) meta-analysis of T2DM and risk of brain/nervous system cancer. Studies are sub-grouped according to design. The diamond denotes the pooled RR. Boxes represent study-specific RR. The size of the box is proportional to the weight of each study in the meta-analysis. Horizontal lines represent the 95% CI. The pooled RR shows no statistically significant association ($p \geq 0.05$) between T2DM and risk of brain/nervous system cancer (summary RR 1.07; 95% CI, 0.78-1.45). Test for heterogeneity among studies: $I^2 = 68\%$. *Abbreviations.* CI: confidence interval; RR: risk ratio.

C3. Publication bias (funnel plots and Egger tests)

Egger's test was used to assess potential publication bias in meta-analyses with at least ten studies via funnel plot asymmetry.

The funnel tests and egger tests were performed with R software (version 4.2.2, package "meta") (Schwarzer et al., 2015).

C3.1 Unadjusted (association) meta-analyses

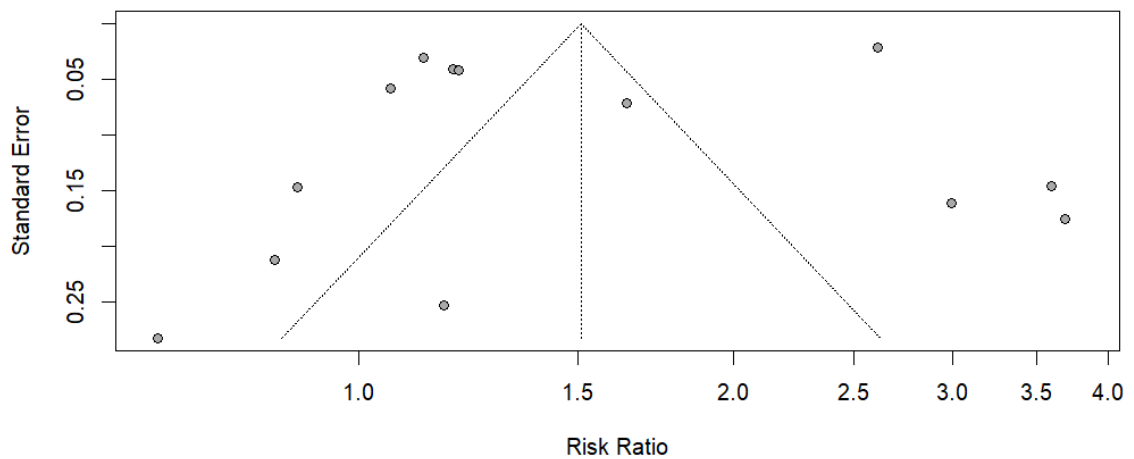


Figure C3-1 Funnel plot from the unadjusted (association) meta-analysis of T2DM and bladder cancer risk. Each point represents a separate study for the indicated association. The funnel plot shows no evidence of publication bias. Egger's test for a regression intercept gave a p-value of 0.434, indicating no evidence of publication bias.

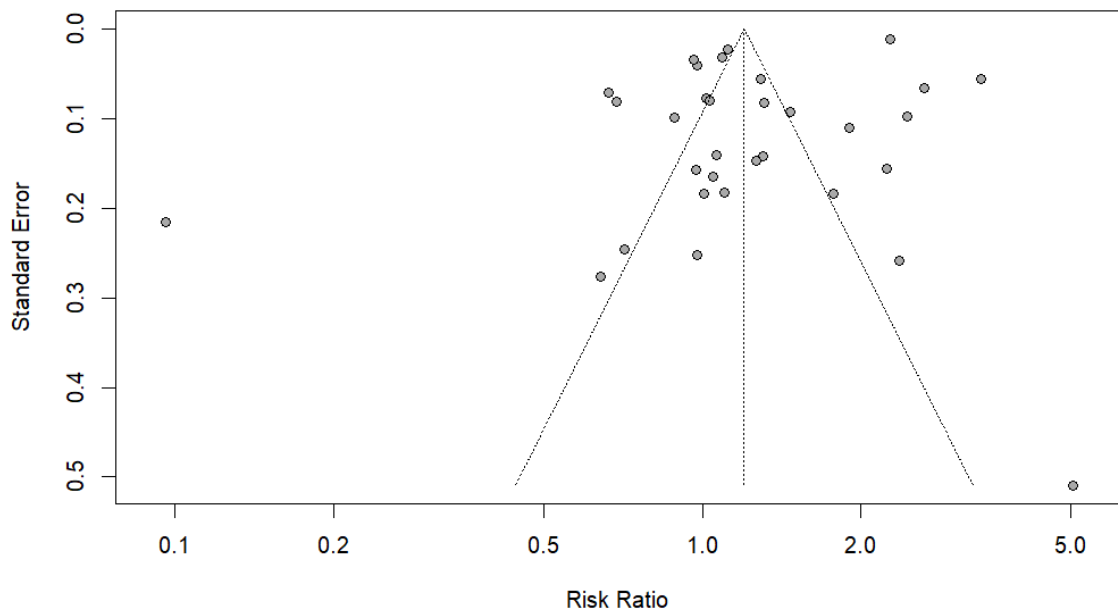


Figure C3-2 Funnel plot from the unadjusted (association) meta-analysis of T2DM and breast cancer risk. Each point represents a separate study for the indicated association. The funnel plot shows possible evidence of publication bias. Egger's test for a regression intercept gave a p-value of 0.014, indicating possible evidence of publication bias.

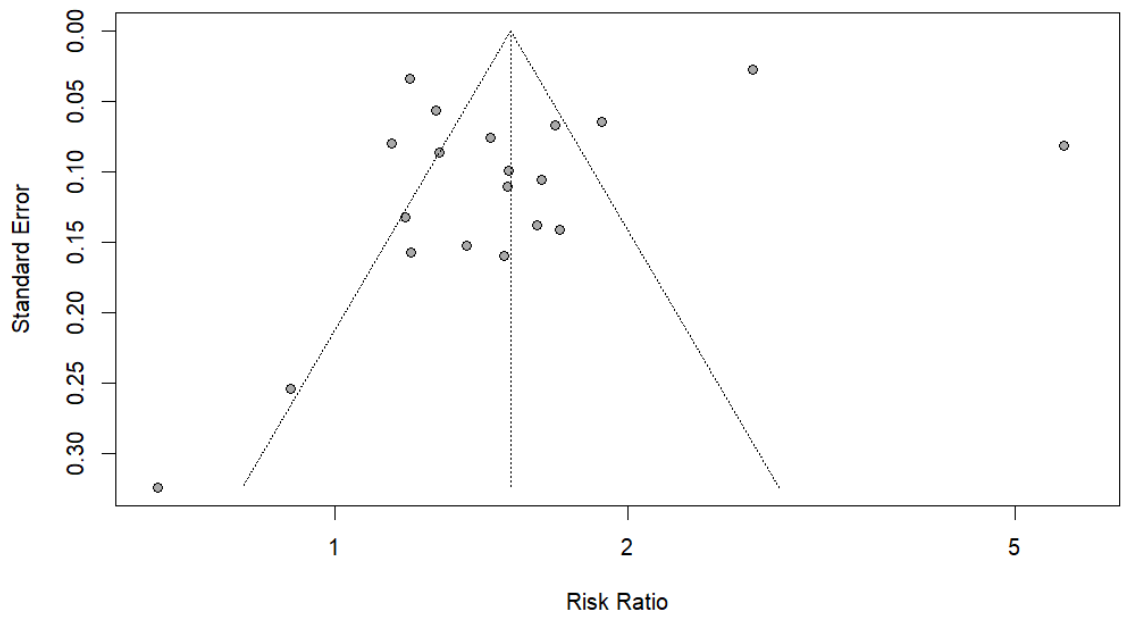


Figure C3-3 Funnel plot from the unadjusted (association) meta-analysis of T2DM and colon cancer risk. Each point represents a separate study for the indicated association. The funnel plot shows no evidence of publication bias. Egger's test for a regression intercept gave a p-value of 0.198, indicating no evidence of publication bias.

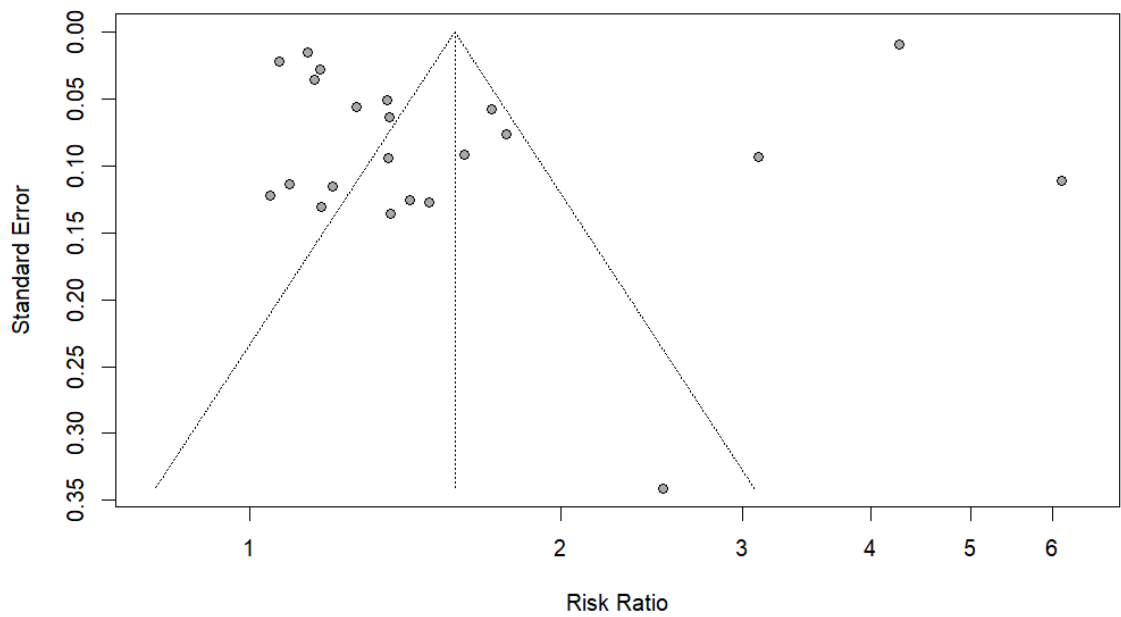


Figure C3-4 Funnel plot from the unadjusted (association) meta-analysis of T2DM and CRC risk. Each point represents a separate study for the indicated association. The funnel plot shows no evidence of publication bias. Egger's test for a regression intercept gave a p-value of 0.076, indicating no evidence of publication bias.

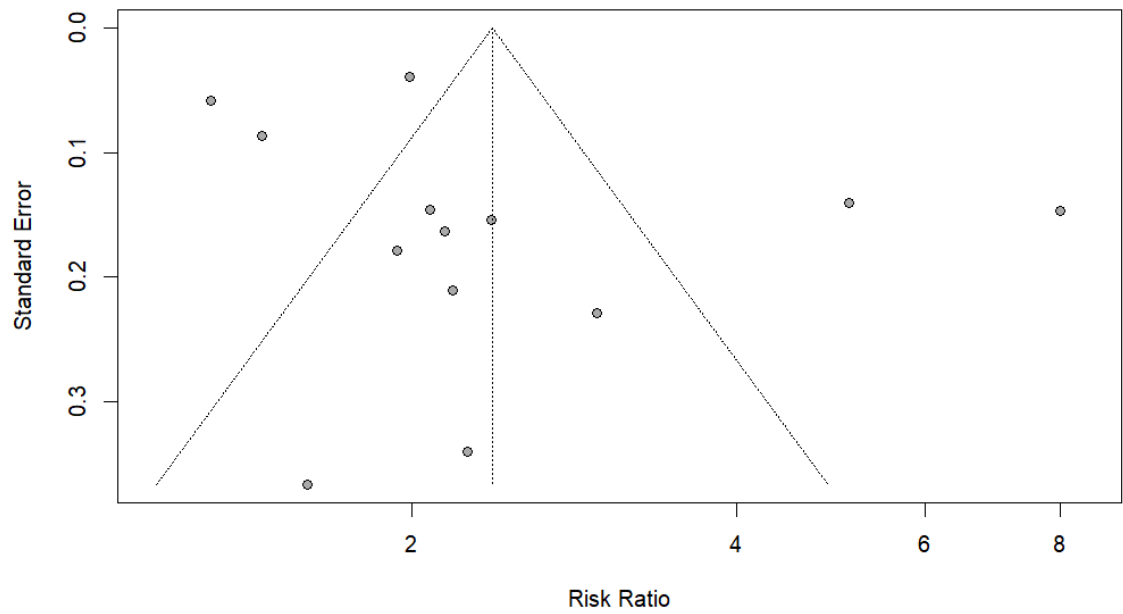


Figure C3-5 Funnel plot from the unadjusted (association) meta-analysis of T2DM and endometrial cancer risk. Each point represents a separate study for the indicated association. The funnel plot shows no evidence of publication bias. Egger's test for a regression intercept gave a p-value of 0.204, indicating no evidence of publication bias.

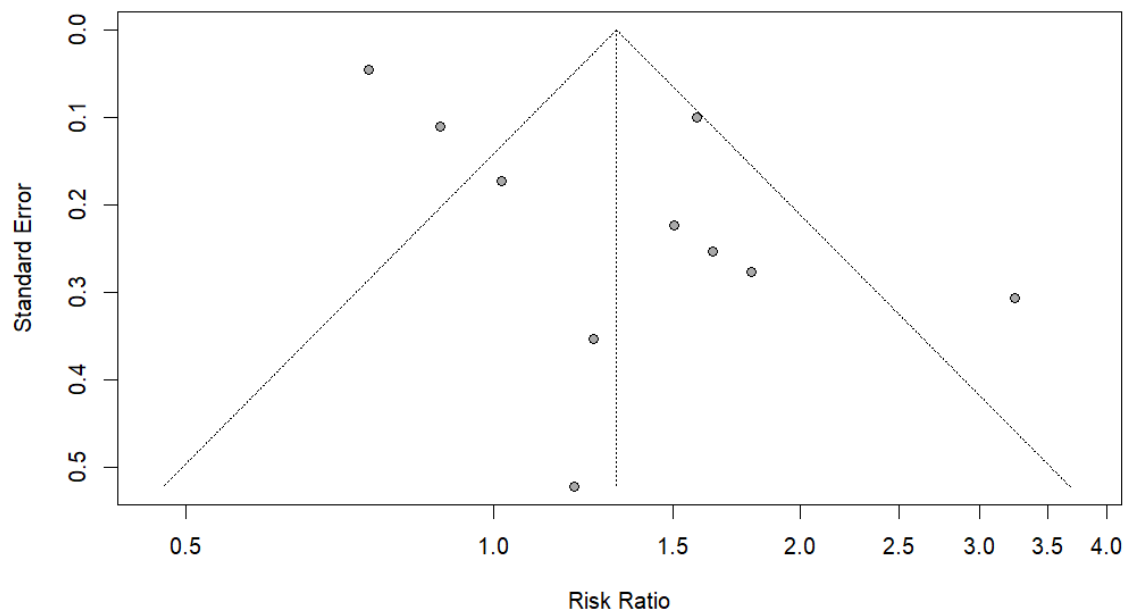


Figure C3-6 Funnel plot from the unadjusted (association) meta-analysis of T2DM and esophageal cancer risk. Each point represents a separate study for the indicated association. The funnel plot shows possible evidence of publication bias. Egger's test for a regression intercept gave a p-value of 0.019, indicating possible evidence of publication bias.

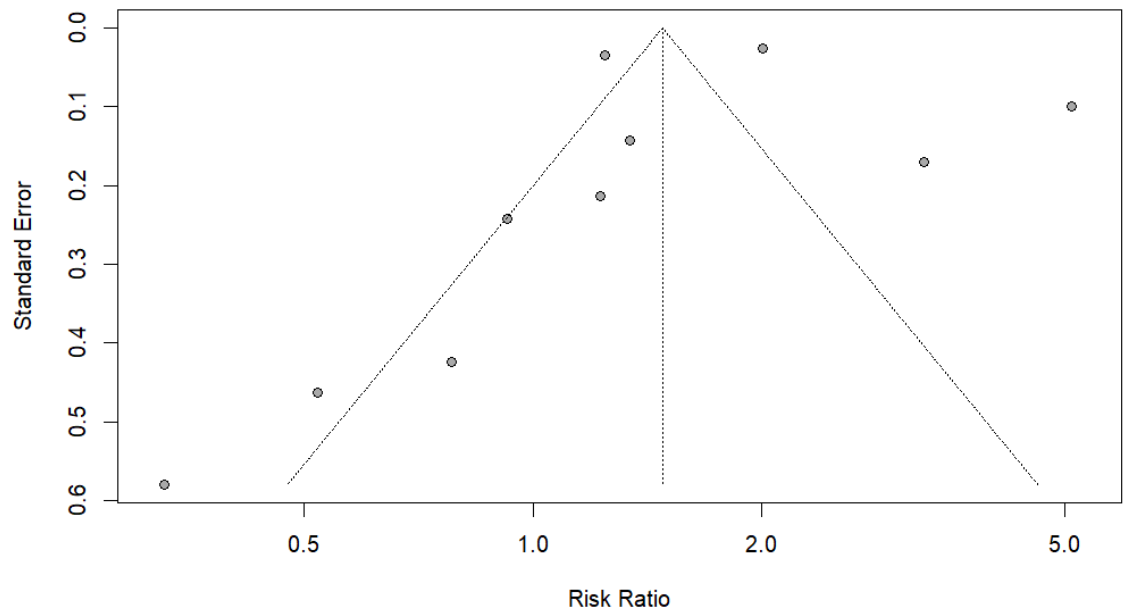


Figure C3-7 Funnel plot from the unadjusted (association) meta-analysis of T2DM and kidney cancer risk. Each point represents a separate study for the indicated association. The funnel plot shows no evidence of publication bias. Egger's test for a regression intercept gave a p-value of 0.802, indicating no evidence of publication bias.

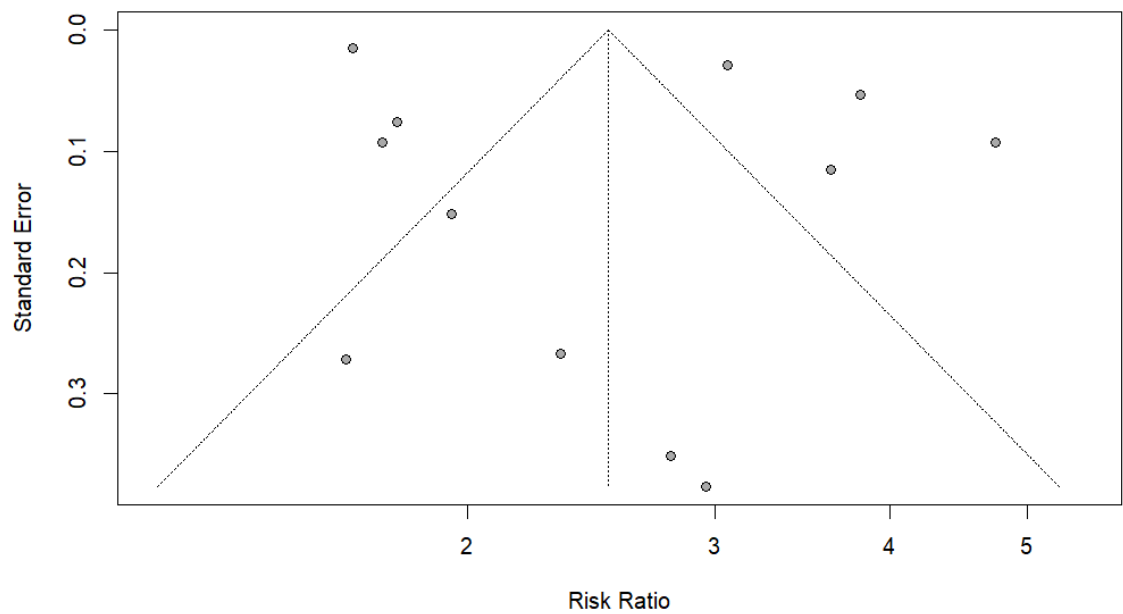


Figure C3-8 Funnel plot from the unadjusted (association) meta-analysis of T2DM and liver cancer risk. Each point represents a separate study for the indicated association. The funnel plot shows no evidence of publication bias. Egger's test for a regression intercept gave a p-value of 0.182, indicating no evidence of publication bias.

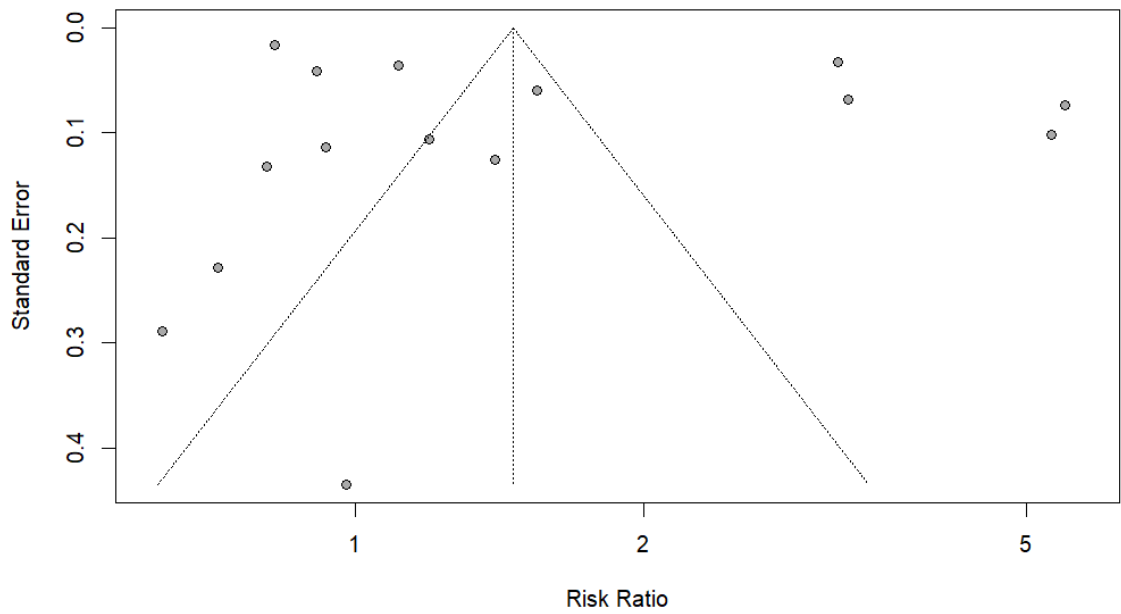


Figure C3-9 Funnel plot from the unadjusted (association) meta-analysis of T2DM and lung cancer risk. Each point represents a separate study for the indicated association. The funnel plot shows no evidence of publication bias. Egger's test for a regression intercept gave a p-value of 0.199, indicating no evidence of publication bias.

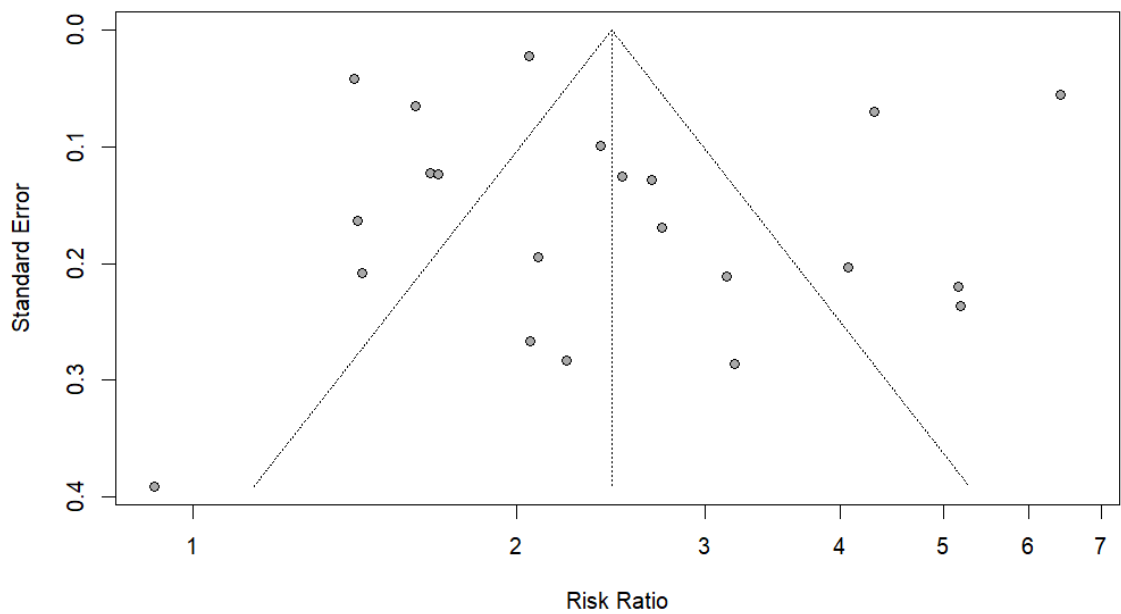


Figure C3-10 Funnel plot from the unadjusted (association) meta-analysis of T2DM and pancreatic cancer risk. Each point represents a separate study for the indicated association. The funnel plot shows no evidence of publication bias. Egger's test for a regression intercept gave a p-value of 0.459, indicating no evidence of publication bias.

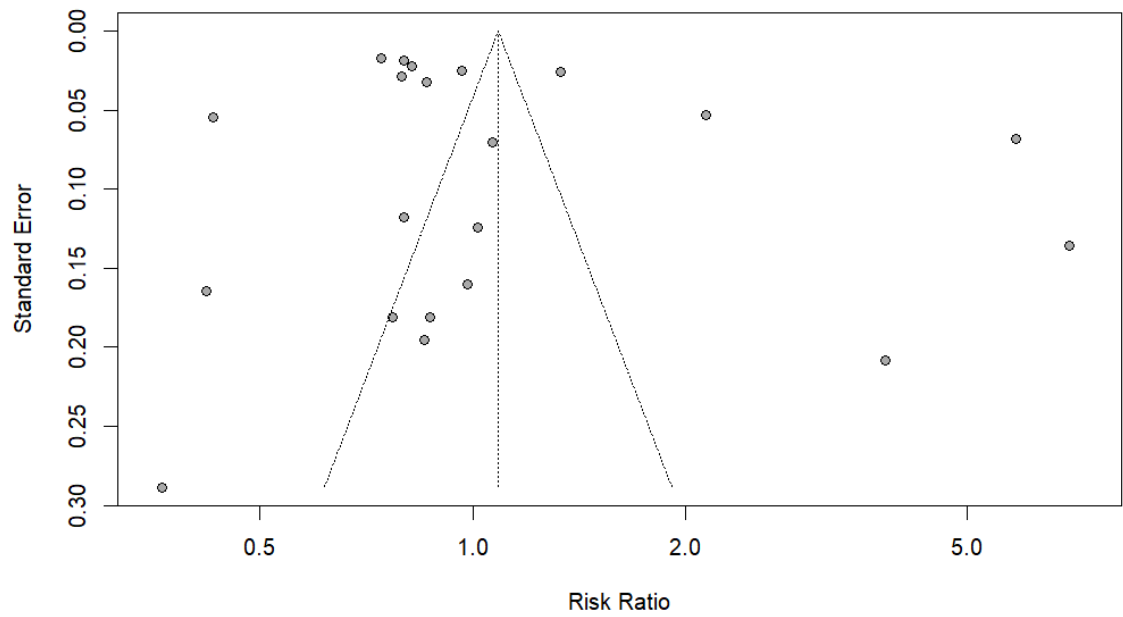


Figure C3-11 Funnel plot from the unadjusted (association) meta-analysis of T2DM and prostate cancer risk. Each point represents a separate study for the indicated association. The funnel plot shows no evidence of publication bias. Egger's test for a regression intercept gave a p-value of 0.187, indicating no evidence of publication bias.

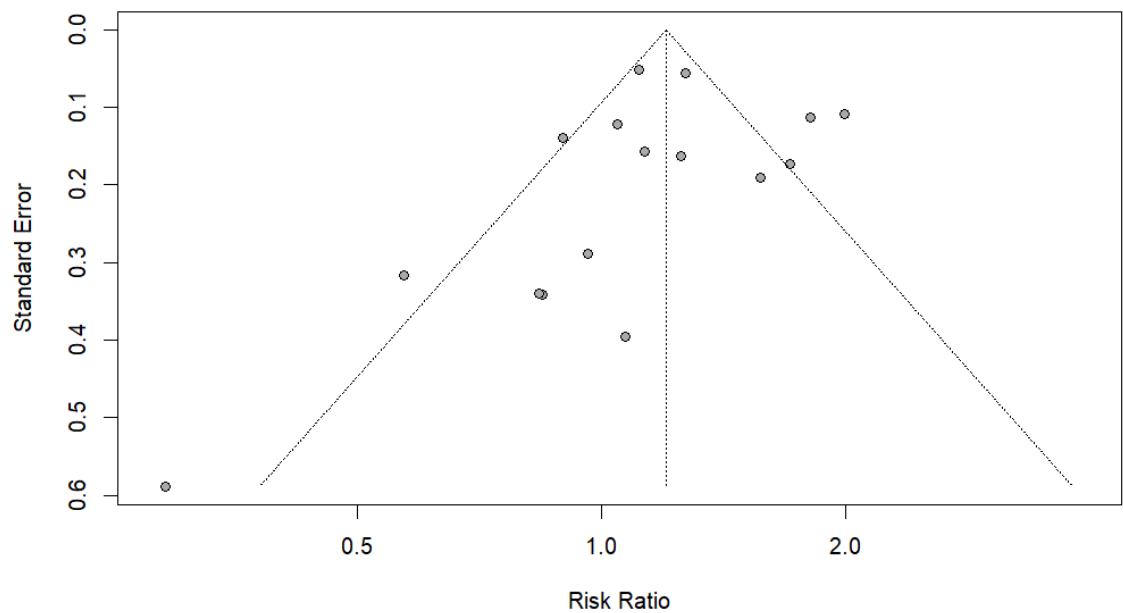


Figure C3-12 Funnel plot from the unadjusted (association) meta-analysis of T2DM and rectal cancer risk. Each point represents a separate study for the indicated association. The funnel plot shows no evidence of publication bias. Egger's test for a regression intercept gave a p-value of 0.492, indicating no evidence of publication bias.

Appendix D. Burden of disease

This appendix provides supplementary information for Chapter 7: Burden of Disease.

Table D1-1 PIF, projected number of cancer cases, DALYs loss and economic value of DALYs loss (monetised DALYs) attributable to T2DM in 2025, by country

No.	Region	Country	HDI	Number of cancer cases	PIF	Estimated number of attributable cancer cases			Estimated number of DALYs loss			Estimated monetised DALYs loss in millions (INT. \$)		
						Total	Male	Female	Total	Male	Female	Total	Male	Female
1	MENA	Afghanistan	Low	6,111	2.37%	145	42	103	2,817	902	1,915	\$1.45	\$0.47	\$0.99
2	Europe	Albania	High	1,920	2.68%	51	21	31	838	434	404	\$17.06	\$8.84	\$8.22
3	MENA	Algeria	High	21,241	1.01%	214	65	149	3,680	1,711	1,970	\$50.43	\$23.44	\$26.99
4	Africa	Angola	Medium	4,765	0.48%	23	5	18	810	332	478	\$3.25	\$1.33	\$1.91
5	SACA	Argentina	Very high	51,316	0.98%	503	193	309	7,685	3,363	4,322	\$254.48	\$111.37	\$143.12
6	Europe	Armenia	High	3,245	1.19%	39	13	26	515	191	325	\$10.97	\$4.06	\$6.91
7	WP	Australia	Very high	46,614	1.30%	606	223	383	5,002	2,351	2,651	\$394.17	\$185.26	\$208.92
8	Europe	Austria	Very high	13,958	0.83%	116	44	72	1,571	761	809	\$132.40	\$64.19	\$68.20
9	Europe	Azerbaijan	High	4,641	0.98%	46	14	31	1,323	612	712	\$25.24	\$11.67	\$13.58
10	NAC	Bahamas	Very high	331	2.02%	7	1	6	117	43	74	\$6.17	\$2.26	\$3.91
11	MENA	Bahrain	Very high	543	2.37%	13	4	9	246	125	120	\$18.08	\$9.23	\$8.85
12	SEA	Bangladesh	Medium	23,041	2.77%	638	198	440	18,551	6,396	12,155	\$98.90	\$34.10	\$64.80
13	NAC	Barbados	High	427	3.69%	16	4	11	217	73	145	\$4.33	\$1.45	\$2.89
14	Europe	Belarus	Very high	13,667	1.34%	183	58	125	1,974	969	1,005	\$56.92	\$27.94	\$28.98
15	Europe	Belgium	Very high	25,248	0.44%	112	41	71	1,223	534	689	\$96.98	\$42.31	\$54.67
16	NAC	Belize	Medium	147	2.33%	3	1	3	102	40	62	\$0.61	\$0.24	\$0.37
17	SEA	Bhutan	Medium	52	2.99%	2	1	1	76	30	46	\$0.63	\$0.25	\$0.38
18	SACA	Bolivia	Medium	3,770	1.26%	48	10	37	819	245	574	\$4.68	\$1.40	\$3.28
19	Europe	Bosnia and Herzegovina	High	3,880	2.32%	90	34	56	1,472	753	719	\$30.63	\$15.66	\$14.97
20	Africa	Botswana	Medium	369	0.80%	3	0	3	177	71	106	\$1.96	\$0.78	\$1.17
21	SACA	Brazil	High	178,185	1.80%	3,213	960	2,253	43,888	18,103	25,785	\$835.16	\$344.49	\$490.67
22	WP	Brunei Darussalam	Very high	337	2.59%	9	3	6	174	81	93	\$16.68	\$7.76	\$8.93
23	Europe	Bulgaria	High	10,300	1.84%	189	64	126	3,256	1,590	1,665	\$111.76	\$54.60	\$57.16
24	Africa	Burkina Faso	Low	3,555	0.02%	1	0	1	13	4	9	\$0.01	\$0.00	\$0.01
25	Africa	Burundi	Low	1,366	0.96%	13	3	10	381	132	249	\$0.10	\$0.03	\$0.07
26	Africa	Cabo Verde	Medium	154	0.03%	0	0	0	1	0	1	\$0.01	\$0.00	\$0.00
27	WP	Cambodia	Medium	3,871	1.29%	50	14	36	1,832	744	1,088	\$6.75	\$2.74	\$4.01
28	Africa	Cameroon	Medium	6,001	0.58%	35	7	28	1,165	472	693	\$3.09	\$1.25	\$1.84
29	NAC	Canada	Very high	69,889	1.80%	1,260	425	835	12,557	5,926	6,631	\$909.56	\$429.24	\$480.31

No.	Region	Country	HDI	Number of cancer cases	PIF	Estimated number of attributable cancer cases			Estimated number of DALYs loss			Estimated monetised DALYs loss in millions (INT. \$)		
						Total	Male	Female	Total	Male	Female	Total	Male	Female
30	Africa	Central African Republic	Low	798	0.65%	5	1	4	131	49	82	\$0.04	\$0.02	\$0.03
31	Africa	Chad	Low	2,805	0.70%	20	5	15	384	131	253	\$0.19	\$0.07	\$0.13
32	SACA	Chile	Very high	16,263	2.87%	466	189	277	7,716	3,282	4,434	\$278.25	\$118.35	\$159.89
33	WP	China	High	1,124,789	2.62%	29,517	11,167	18,350	412,923	227,914	185,009	\$10,469.59	\$5,778.73	\$4,690.86
34	SACA	Colombia	High	35,291	1.78%	629	185	445	6,937	2,688	4,249	\$138.15	\$53.53	\$84.62
35	Africa	Comoros	Medium	91	1.77%	2	0	1	114	46	69	\$0.22	\$0.09	\$0.13
36	Africa	Congo	Medium	674	0.65%	4	1	3	354	155	199	\$1.03	\$0.45	\$0.58
37	SACA	Costa Rica	Very high	4,016	1.97%	79	24	55	1,271	562	708	\$41.79	\$18.49	\$23.29
38	Africa	Côte d'Ivoire	Medium	5,382	0.02%	1	0	1	20	9	11	\$0.08	\$0.04	\$0.04
39	Europe	Croatia	Very high	7,951	0.94%	75	28	46	914	468	447	\$48.73	\$24.93	\$23.79
40	SACA	Cuba	High	13,479	1.90%	256	62	194	3,340	1,182	2,157	\$-	\$-	\$-
41	Europe	Cyprus	Very high	1,756	1.94%	34	13	21	381	147	234	\$25.76	\$9.92	\$15.84
42	Europe	Czechia	Very high	22,089	1.79%	395	155	240	4,408	2,272	2,135	\$281.92	\$145.33	\$136.59
43	WP	Democratic People's Republic of Korea	N/A	13,743	1.92%	264	92	172	4,925	2,240	2,685	\$-	\$-	\$-
44	Africa	Democratic Republic of the Congo	Low	11,301	0.74%	84	21	63	2,305	877	1,428	\$1.01	\$0.38	\$0.62
45	Europe	Denmark	Very high	12,987	1.00%	130	48	83	1,380	649	731	\$122.71	\$57.70	\$65.01
46	Africa	Djibouti	Low	291	1.01%	3	1	2	87	34	54	\$0.18	\$0.07	\$0.11
47	SACA	Dominican Republic	High	6,436	1.88%	121	39	83	2,377	852	1,525	\$66.35	\$23.79	\$42.56
48	SACA	Ecuador	High	8,619	0.73%	63	18	45	1,006	363	643	\$13.90	\$5.01	\$8.89
49	MENA	Egypt	High	36,827	4.13%	1,521	446	1,075	31,358	14,244	17,114	\$564.40	\$256.37	\$308.03
50	SACA	El Salvador	Medium	2,834	1.03%	29	8	21	569	210	358	\$3.74	\$1.38	\$2.36
51	Africa	Equatorial Guinea	Medium	270	0.67%	2	0	1	88	42	46	\$0.71	\$0.34	\$0.37
52	Africa	Eritrea	Low	799	0.85%	7	2	5	202	65	136	\$0.12	\$0.04	\$0.08
53	Europe	Estonia	Very high	2,387	1.61%	38	14	25	435	195	241	\$25.75	\$11.52	\$14.24
54	Africa	Eswatini	Medium	130	0.46%	1	0	0	58	29	29	\$0.37	\$0.19	\$0.19
55	Africa	Ethiopia	Low	26,429	0.59%	156	39	117	2,536	858	1,677	\$2.88	\$0.97	\$1.90
56	WP	Fiji	High	477	3.83%	18	3	16	319	102	217	\$5.44	\$1.75	\$3.69
57	Europe	Finland	Very high	11,663	1.27%	148	50	98	1,500	675	825	\$111.64	\$50.26	\$61.38
58	Europe	France	Very high	134,755	1.04%	1,402	504	897	13,305	6,367	6,938	\$956.47	\$457.73	\$498.73
59	Africa	Gabon	High	523	0.87%	5	1	3	165	84	81	\$3.26	\$1.66	\$1.60
60	Europe	Georgia	Very high	3,668	1.13%	41	10	32	724	287	437	\$21.01	\$8.32	\$12.69
61	Europe	Germany	Very high	161,860	1.56%	2,527	915	1,611	30,170	14,641	15,529	\$2,450.84	\$1,189.37	\$1,261.47
62	Africa	Ghana	Medium	6,563	0.12%	8	1	7	349	111	238	\$1.37	\$0.44	\$0.93

No.	Region	Country	HDI	Number of cancer cases	PIF	Estimated number of attributable cancer cases			Estimated number of DALYs loss			Estimated monetised DALYs loss in millions (INT. \$)		
						Total	Male	Female	Total	Male	Female	Total	Male	Female
63	Europe	Greece	Very high	21,320	1.44%	306	114	192	2,548	1,187	1,362	\$124.35	\$57.90	\$66.45
64	WP	Guam	N/A	118	5.27%	6	2	4	89	43	46	\$-	\$-	\$-
65	SACA	Guatemala	Medium	4,249	2.59%	110	34	76	2,595	1,013	1,582	\$15.40	\$6.01	\$9.39
66	Africa	Guinea	Low	1,109	0.02%	0	0	0	8	2	6	\$0.01	\$0.00	\$0.01
67	Africa	Guinea-Bissau	Low	244	0.01%	0	0	0	2	1	1	\$0.00	\$0.00	\$0.00
68	NAC	Guyana	High	364	2.31%	8	2	7	229	74	155	\$20.63	\$6.68	\$13.95
69	NAC	Haiti	Low	3,313	1.78%	59	19	40	1,382	302	1,080	\$1.25	\$0.27	\$0.97
70	SACA	Honduras	Medium	2,628	0.83%	22	8	14	451	118	333	\$1.85	\$0.48	\$1.36
71	Europe	Hungary	Very high	20,947	1.62%	339	126	213	4,300	2,142	2,158	\$239.86	\$119.49	\$120.36
72	Europe	Iceland	Very high	590	1.10%	6	2	4	57	33	25	\$4.88	\$2.77	\$2.11
73	SEA	India	Medium	287,659	1.65%	4,750	1,348	3,402	139,116	53,147	85,969	\$765.38	\$292.40	\$472.98
74	WP	Indonesia	High	113,116	1.91%	2,166	580	1,585	62,591	23,791	38,800	\$1,069.82	\$406.64	\$663.17
75	MENA	Iran	High	35,781	1.79%	642	241	401	11,583	5,528	6,055	\$230.04	\$109.79	\$120.25
76	MENA	Iraq	Medium	13,050	1.91%	249	73	176	5,588	2,193	3,394	\$39.70	\$15.58	\$24.11
77	Europe	Ireland	Very high	8,887	0.32%	28	10	18	260	125	135	\$47.50	\$22.87	\$24.62
78	Europe	Israel	Very high	10,172	1.99%	202	72	130	2,122	991	1,131	\$143.35	\$66.96	\$76.40
79	Europe	Italy	Very high	131,261	1.35%	1,778	615	1,163	16,530	7,536	8,994	\$1,089.75	\$496.79	\$592.96
80	NAC	Jamaica	High	2,638	2.70%	71	15	56	878	285	593	\$11.65	\$3.78	\$7.87
81	WP	Japan	Very high	286,005	1.53%	4,369	1,778	2,591	39,000	19,912	19,088	\$2,468.94	\$1,260.57	\$1,208.37
82	MENA	Jordan	High	4,677	3.19%	149	46	104	2,773	1,360	1,414	\$37.54	\$18.40	\$19.14
83	Europe	Kazakhstan	Very high	10,560	1.53%	161	50	111	2,949	1,238	1,711	\$121.93	\$51.19	\$70.75
84	Africa	Kenya	Medium	12,585	0.46%	58	12	46	895	291	605	\$3.42	\$1.11	\$2.31
85	MENA	Kuwait	Very high	1,843	5.74%	106	37	69	2,170	1,397	773	\$139.71	\$89.96	\$49.76
86	Europe	Kyrgyzstan	Medium	1,822	1.60%	29	8	21	508	206	302	\$1.80	\$0.73	\$1.07
87	WP	Lao People's Democratic Republic	Medium	2,057	0.92%	19	6	13	530	198	332	\$2.99	\$1.12	\$1.87
88	Europe	Latvia	Very high	3,526	1.38%	49	17	32	459	184	275	\$23.53	\$9.45	\$14.07
89	MENA	Lebanon	High	3,794	1.63%	62	22	39	1,153	450	703	\$18.27	\$7.13	\$11.14
90	Africa	Lesotho	Low	301	0.59%	2	0	1	101	40	61	\$0.09	\$0.04	\$0.06
91	Africa	Liberia	Low	829	0.02%	0	0	0	5	2	3	\$0.00	\$0.00	\$0.00
92	MENA	Libya	High	2,656	1.67%	44	16	28	923	433	490	\$25.88	\$12.15	\$13.73
93	Europe	Lithuania	Very high	5,146	1.46%	75	23	52	667	283	383	\$41.98	\$17.85	\$24.13
94	Europe	Luxembourg	Very high	1,062	1.08%	11	4	7	122	55	66	\$20.70	\$9.42	\$11.28
95	Africa	Madagascar	Low	4,724	0.48%	23	5	18	723	246	478	\$0.40	\$0.14	\$0.26
96	Africa	Malawi	Low	2,189	0.86%	19	3	16	723	281	442	\$0.34	\$0.13	\$0.21
97	WP	Malaysia	Very high	18,473	4.35%	803	244	559	14,277	5,859	8,418	\$667.46	\$273.89	\$393.57
98	SEA	Maldives	High	151	1.49%	2	1	2	46	21	25	\$1.82	\$0.84	\$0.98
99	Africa	Mali	Low	3,929	0.02%	1	0	0	11	5	7	\$0.01	\$0.00	\$0.01
100	Europe	Malta	Very high	919	1.81%	17	6	11	172	85	87	\$13.49	\$6.66	\$6.83
101	Africa	Mauritania	Medium	873	0.02%	0	0	0	4	1	3	\$0.02	\$0.01	\$0.01

No.	Region	Country	HDI	Number of cancer cases	PIF	Estimated number of attributable cancer cases			Estimated number of DALYs loss			Estimated monetised DALYs loss in millions (INT. \$)		
						Total	Male	Female	Total	Male	Female	Total	Male	Female
102	SEA	Mauritius	Very high	1,229	4.87%	60	17	43	901	379	522	\$33.77	\$14.21	\$19.56
103	NAC	Mexico	High	65,417	4.17%	2,727	847	1,880	42,642	19,219	23,423	\$1,029.56	\$464.03	\$565.54
104	WP	Mongolia	High	667	1.85%	12	5	8	392	181	211	\$6.34	\$2.93	\$3.41
105	Europe	Montenegro	Very high	804	2.18%	18	7	11	273	131	141	\$9.55	\$4.61	\$4.94
106	MENA	Morocco	Medium	19,086	1.51%	288	84	204	5,441	1,886	3,556	\$32.47	\$11.25	\$21.22
107	Africa	Mozambique	Low	2,981	0.25%	7	1	6	369	123	246	\$0.17	\$0.06	\$0.11
108	WP	Myanmar	Medium	12,961	1.23%	160	48	112	6,050	2,337	3,713	\$17.67	\$6.82	\$10.84
109	Africa	Namibia	Medium	847	0.94%	8	2	6	166	54	112	\$1.05	\$0.35	\$0.71
110	SEA	Nepal	Medium	3,874	1.64%	63	20	44	2,067	683	1,384	\$6.20	\$2.05	\$4.15
111	Europe	Netherlands	Very high	38,904	0.72%	281	109	172	3,312	1,574	1,738	\$293.83	\$139.64	\$154.18
112	WP	New Zealand	Very high	9,028	1.26%	114	39	75	1,170	533	637	\$76.64	\$34.90	\$41.73
113	SACA	Nicaragua	Medium	2,510	1.98%	50	13	36	761	321	441	\$3.31	\$1.39	\$1.92
114	Africa	Niger	Low	3,471	0.70%	24	7	17	393	132	261	\$0.20	\$0.07	\$0.13
115	Africa	Nigeria	Low	42,523	0.30%	129	31	98	3,359	1,381	1,978	\$5.81	\$2.39	\$3.42
116	Europe	North Macedonia	High	2,278	1.40%	32	9	23	319	106	213	\$7.20	\$2.40	\$4.81
117	Europe	Norway	Very high	10,755	0.51%	55	21	35	523	251	272	\$53.09	\$25.49	\$27.60
118	MENA	Oman	Very high	1,209	2.84%	34	17	17	514	245	269	\$26.07	\$12.43	\$13.63
119	MENA	Pakistan	Low	43,366	5.69%	2,469	618	1,851	122,187	38,979	83,208	\$240.71	\$76.79	\$163.92
120	SACA	Panama	Very high	2,604	1.85%	48	14	34	626	251	375	\$31.58	\$12.66	\$18.92
121	WP	Papua New Guinea	Medium	2,647	3.44%	91	18	74	2,062	449	1,613	\$5.26	\$1.14	\$4.11
122	SACA	Paraguay	High	4,274	1.45%	62	20	42	795	332	463	\$12.86	\$5.37	\$7.48
123	SACA	Peru	High	17,515	0.89%	155	52	103	2,187	865	1,322	\$36.51	\$14.43	\$22.07
124	WP	Philippines	Medium	56,874	1.32%	751	211	540	10,807	4,157	6,650	\$74.09	\$28.50	\$45.59
125	Europe	Poland	Very high	64,130	1.68%	1,080	314	766	14,858	7,384	7,474	\$854.72	\$424.76	\$429.96
126	Europe	Portugal	Very high	17,346	2.09%	363	141	222	4,515	2,204	2,311	\$250.30	\$122.16	\$128.13
127	SACA	Puerto Rico	Very high	4,538	3.13%	142	45	97	1,901	812	1,089	\$98.67	\$42.14	\$56.53
128	MENA	Qatar	Very high	563	4.32%	24	11	13	516	307	209	\$82.38	\$49.01	\$33.37
129	WP	Republic of Korea	Very high	67,525	1.48%	1,001	426	575	11,119	6,016	5,102	\$788.91	\$426.89	\$362.02
130	Europe	Republic of Moldova	High	4,422	1.25%	55	19	36	591	300	291	\$11.24	\$5.71	\$5.53
131	Europe	Romania	Very high	28,930	1.40%	404	153	251	4,804	2,288	2,516	\$260.58	\$124.12	\$136.46
132	Europe	Russian Federation	Very high	202,867	1.29%	2,619	751	1,867	32,402	13,717	18,685	\$1,380.98	\$584.63	\$796.35
133	Africa	Rwanda	Low	1,900	0.79%	15	3	12	622	210	411	\$0.59	\$0.20	\$0.39
134	NAC	Saint Lucia	High	129	2.51%	3	1	2	75	33	42	\$1.42	\$0.63	\$0.79
135	WP	Samoa	High	146	2.24%	3	1	2	26	9	18	\$0.18	\$0.06	\$0.12
136	Africa	Sao Tome and Principe	Medium	30	0.65%	0	0	0	11	5	6	\$0.03	\$0.01	\$0.02

No.	Region	Country	HDI	Number of cancer cases	PIF	Estimated number of attributable cancer cases			Estimated number of DALYs loss			Estimated monetised DALYs loss in millions (INT. \$)		
						Total	Male	Female	Total	Male	Female	Total	Male	Female
137	MENA	Saudi Arabia	Very high	10,583	4.91%	520	201	319	9,753	5,411	4,342	\$768.77	\$426.54	\$342.23
138	Africa	Senegal	Low	2,912	0.22%	6	2	5	178	63	115	\$0.24	\$0.09	\$0.16
139	Europe	Serbia	Very high	14,125	2.18%	308	102	206	4,163	1,875	2,288	\$136.82	\$61.62	\$75.20
140	Africa	Sierra Leone	Low	1,443	0.01%	0	0	0	6	2	4	\$0.00	\$0.00	\$0.00
141	WP	Singapore	Very high	9,699	2.90%	281	106	174	2,296	1,117	1,180	\$380.06	\$184.81	\$195.24
142	Europe	Slovakia	Very high	10,080	1.37%	138	51	86	1,670	842	828	\$87.27	\$44.00	\$43.27
143	Europe	Slovenia	Very high	4,355	1.31%	57	23	34	589	302	287	\$38.76	\$19.87	\$18.89
144	WP	Solomon Islands	Medium	191	3.94%	8	1	7	312	65	248	\$0.41	\$0.09	\$0.33
145	Africa	Somalia	N/A	3,113	0.93%	29	7	22	431	155	276	\$-	\$-	\$-
146	Africa	South Africa	High	27,911	2.18%	610	152	457	9,555	3,967	5,588	\$153.58	\$63.76	\$89.82
147	Africa	South Sudan	Low	1,516	0.99%	15	4	11	445	169	275	\$0.07	\$0.03	\$0.04
148	Europe	Spain	Very high	90,375	2.52%	2,277	932	1,346	24,254	12,115	12,138	\$1,475.16	\$736.88	\$738.27
149	SEA	Sri Lanka	High	6,638	2.25%	149	34	115	4,005	1,536	2,468	\$57.77	\$22.16	\$35.61
150	MENA	State of Palestine	High	1,904	1.81%	34	12	23	576	208	368	\$3.86	\$1.40	\$2.46
151	MENA	Sudan	Low	9,511	2.95%	281	58	223	6,621	2,785	3,836	\$8.25	\$3.47	\$4.78
152	NAC	Suriname	High	346	2.59%	9	3	6	209	96	113	\$3.97	\$1.82	\$2.14
153	Europe	Sweden	Very high	18,078	0.94%	169	57	112	1,850	855	994	\$147.83	\$68.36	\$79.47
154	Europe	Switzerland	Very high	15,437	0.76%	118	41	76	1,163	549	614	\$123.69	\$58.41	\$65.28
155	MENA	Syrian Arab Republic	Medium	10,119	3.00%	304	90	213	2,850	1,147	1,702	\$-	\$-	\$-
156	Europe	Tajikistan	Medium	1,548	1.47%	23	6	16	748	291	457	\$2.27	\$0.88	\$1.39
157	WP	Thailand	Very high	48,143	2.14%	1,029	317	712	15,959	6,871	9,088	\$460.09	\$198.09	\$262.01
158	WP	Timor-Leste	Medium	223	1.53%	3	1	2	111	40	71	\$0.21	\$0.08	\$0.14
159	Africa	Togo	Low	1,525	0.02%	0	0	0	7	3	4	\$0.01	\$0.00	\$0.00
160	NAC	Trinidad and Tobago	Very high	1,388	3.50%	49	12	36	653	229	424	\$25.97	\$9.12	\$16.85
161	MENA	Tunisia	High	5,852	1.92%	112	37	76	1,797	762	1,035	\$24.35	\$10.32	\$14.03
162	Europe	Turkey	Very high	62,914	4.03%	2,538	930	1,608	32,520	16,141	16,379	\$1,674.18	\$830.95	\$843.23
163	Europe	Turkmenistan	High	1,888	1.37%	26	8	18	405	167	238	\$8.23	\$3.38	\$4.84
164	Africa	Uganda	Low	4,850	0.61%	30	6	24	1,181	473	708	\$1.13	\$0.45	\$0.68
165	Europe	Ukraine	High	51,540	1.37%	708	203	505	10,566	5,184	5,382	\$138.33	\$67.87	\$70.47
166	MENA	United Arab Emirates	Very high	2,259	3.53%	80	25	54	3,965	2,797	1,168	\$437.02	\$308.28	\$128.74
167	Europe	United Kingdom	Very high	132,793	1.34%	1,781	624	1,157	17,033	7,652	9,381	\$1,171.28	\$526.17	\$645.12
168	Africa	United Republic of Tanzania	Low	7,190	2.37%	171	51	120	6,308	2,409	3,899	\$6.71	\$2.56	\$4.15
169	NAC	United States of America	Very high	583,650	2.78%	16,230	5,368	10,861	184,816	98,469	86,346	\$17,878.38	\$9,525.55	\$8,352.83
170	SACA	Uruguay	Very high	5,490	2.27%	125	53	71	1,820	860	960	\$65.36	\$30.88	\$34.49
171	Europe	Uzbekistan	High	9,010	1.42%	128	40	88	4,310	1,909	2,402	\$47.29	\$20.94	\$26.35

No.	Region	Country	HDI	Number of cancer cases	PIF	Estimated number of attributable cancer cases			Estimated number of DALYs loss			Estimated monetised DALYs loss in millions (INT. \$)		
						Total	Male	Female	Total	Male	Female	Total	Male	Female
172	WP	Vanuatu	Medium	58	3.34%	2	0	2	59	19	40	\$0.10	\$0.03	\$0.07
173	SACA	Venezuela	Medium	19,669	2.21%	435	132	303	6,697	2,784	3,914	\$30.80	\$12.80	\$18.00
174	WP	Viet Nam	High	41,616	1.12%	466	100	366	11,481	5,383	6,098	\$185.37	\$86.91	\$98.46
175	MENA	Yemen	Low	5,091	0.64%	33	10	22	705	248	457	\$0.42	\$0.15	\$0.27
176	Africa	Zambia	Medium	1,672	2.08%	35	9	25	2,105	828	1,277	\$4.80	\$1.89	\$2.91
177	Africa	Zimbabwe	Medium	3,109	0.02%	1	0	0	21	5	15	\$0.03	\$0.01	\$0.02

The total number of cancer cases are based on Globocan projections. The total number of DALYs loss are based on GBD Compare tool and World Population Prospects, reported by United Nations. Monetised DALYs could not be calculated for five countries (Cuba, the Democratic People's Republic of Korea, Guam, Somalia, and the Syrian Arab Republic) due to unavailable data on GDP projections and HDI classification.

Abbreviations. INT \$: international dollars at constant 2017 prices; MENA: Middle East and North Africa; NAC: North America and Caribbean; SACA: South and Central America; SEA: South-East Asia; WP: Western Pacific.

Table D1-2 Estimates used to monetise the DALYs per country. Monetised DALYs were calculated based on GDP per capita and adjusted for each country's HDI tier.

No.	Region	Country	HDI	GDP per capita (INT \$)	Cost /DALY (INT \$)
1	MENA	Afghanistan	Low	\$1,517.016	\$515.786
2	Europe	Albania	High	\$16,691.145	\$20,363.197
3	MENA	Algeria	High	\$11,231.566	\$13,702.511
4	Africa	Angola	Medium	\$5,983.376	\$4,008.862
5	SACA	Argentina	Very high	\$22,680.318	\$33,113.264
6	Europe	Armenia	High	\$17,442.683	\$21,280.073
7	WP	Australia	Very high	\$53,975.512	\$78,804.248
8	Europe	Austria	Very high	\$57,740.709	\$84,301.435
9	Europe	Azerbaijan	High	\$15,638.758	\$19,079.285
10	NAC	Bahamas	Very high	\$36,253.146	\$52,929.593
11	MENA	Bahrain	Very high	\$50,360.215	\$73,525.914
12	SEA	Bangladesh	Medium	\$7,957.329	\$5,331.410
13	NAC	Barbados	High	\$16,339.438	\$19,934.114
14	Europe	Belarus	Very high	\$19,751.640	\$28,837.394
15	Europe	Belgium	Very high	\$54,320.316	\$79,307.661
16	NAC	Belize	Medium	\$8,923.397	\$5,978.676
17	SEA	Bhutan	Medium	\$12,357.461	\$8,279.499
18	SACA	Bolivia	Medium	\$8,539.242	\$5,721.292
19	Europe	Bosnia and Herzegovina	High	\$17,058.538	\$20,811.416
20	Africa	Botswana	Medium	\$16,532.994	\$11,077.106
21	SACA	Brazil	High	\$15,597.704	\$19,029.199
22	WP	Brunei Darussalam	Very high	\$65,800.301	\$96,068.439
23	Europe	Bulgaria	High	\$28,138.845	\$34,329.391
24	Africa	Burkina Faso	Low	\$2,356.155	\$801.093
25	Africa	Burundi	Low	\$769.392	\$261.593
26	Africa	Cabo Verde	Medium	\$8,541.889	\$5,723.066
27	WP	Cambodia	Medium	\$5,504.329	\$3,687.900
28	Africa	Cameroon	Medium	\$3,955.272	\$2,650.032
29	NAC	Canada	Very high	\$49,614.043	\$72,436.503
30	Africa	Central African Republic	Low	\$949.420	\$322.803
31	Africa	Chad	Low	\$1,477.722	\$502.425
32	SACA	Chile	Very high	\$24,698.260	\$36,059.460
33	WP	China	High	\$20,782.642	\$25,354.823
34	SACA	Colombia	High	\$16,322.288	\$19,913.191
35	Africa	Comoros	Medium	\$2,831.164	\$1,896.880
36	Africa	Congo	Medium	\$4,333.709	\$2,903.585
37	SACA	Costa Rica	Very high	\$22,523.723	\$32,884.636
38	Africa	Côte d'Ivoire	Medium	\$6,161.897	\$4,128.471
39	Europe	Croatia	Very high	\$36,496.981	\$53,285.592
40	SACA	Cuba	High	N/A	\$-
41	Europe	Cyprus	Very high	\$46,311.704	\$67,615.088
42	Europe	Czechia	Very high	\$43,809.616	\$63,962.039
43	WP	Democratic People's Republic of Korea	N/A	N/A	\$-
44	Africa	Democratic Republic of the Congo	Low	\$1,282.611	\$436.088
45	Europe	Denmark	Very high	\$60,906.415	\$88,923.366
46	Africa	Djibouti	Low	\$6,131.656	\$2,084.763
47	SACA	Dominican Republic	High	\$22,880.086	\$27,913.705
48	SACA	Ecuador	High	\$11,326.000	\$13,817.720
49	MENA	Egypt	High	\$14,753.270	\$17,998.989
50	SACA	El Salvador	Medium	\$9,813.849	\$6,575.279
51	Africa	Equatorial Guinea	Medium	\$11,961.321	\$8,014.085
52	Africa	Eritrea	Low	\$1,821.154	\$619.192
53	Europe	Estonia	Very high	\$40,511.139	\$59,146.263
54	Africa	Eswatini	Medium	\$9,653.365	\$6,467.755
55	Africa	Ethiopia	Low	\$3,335.756	\$1,134.157
56	WP	Fiji	High	\$13,961.284	\$17,032.766
57	Europe	Finland	Very high	\$50,974.920	\$74,423.383
58	Europe	France	Very high	\$49,237.620	\$71,886.925
59	Africa	Gabon	High	\$16,247.880	\$19,822.414
60	Europe	Georgia	Very high	\$2,423.130	\$823.864
61	Europe	Germany	Very high	\$19,883.587	\$29,030.037
62	Africa	Ghana	Medium	\$55,640.774	\$81,235.530
63	Europe	Greece	Very high	\$5,838.482	\$3,911.783
64	WP	Guam	N/A	\$33,421.165	\$48,794.901
65	SACA	Guatemala	Medium	N/A	\$-
66	Africa	Guinea	Low	\$8,860.089	\$5,936.260
67	Africa	Guinea-Bissau	Low	\$2,791.220	\$949.018
68	NAC	Guyana	High	\$2,649.859	\$900.952
69	NAC	Haiti	Low	\$73,908.996	\$90,168.975
70	SACA	Honduras	Medium	\$2,650.403	\$901.137

No.	Region	Country	HDI	GDP per capita (INT \$)	Cost /DALY (INT \$)
71	Europe	Hungary	Very high	\$6,115.708	\$4,097.524
72	Europe	Iceland	Very high	\$38,209.489	\$55,785.854
73	SEA	India	Medium	\$58,156.494	\$84,908.481
74	WP	Indonesia	High	\$8,211.532	\$5,501.726
75	MENA	Iran	High	\$16,278.492	\$19,859.760
76	MENA	Iraq	Medium	\$10,603.168	\$7,104.123
77	Europe	Ireland	Very high	\$125,078.102	\$182,614.029
78	Europe	Israel	Very high	\$46,262.326	\$67,542.996
79	Europe	Italy	Very high	\$45,154.147	\$65,925.055
80	NAC	Jamaica	High	\$10,880.132	\$13,273.761
81	WP	Japan	Very high	\$43,360.776	\$63,306.718
82	MENA	Jordan	High	\$11,096.036	\$13,537.164
83	Europe	Kazakhstan	Very high	\$28,315.158	\$41,340.131
84	Africa	Kenya	Medium	\$5,705.597	\$3,822.750
85	MENA	Kuwait	Very high	\$44,108.054	\$64,397.759
86	Europe	Kyrgyzstan	Medium	\$5,278.405	\$3,536.531
87	WP	Lao People's Democratic Republic	Medium	\$8,429.228	\$5,647.583
88	Europe	Latvia	Very high	\$35,101.676	\$51,248.447
89	MENA	Lebanon	High	\$12,986.913	\$15,844.034
90	Africa	Lesotho	Low	\$2,714.894	\$923.064
91	Africa	Liberia	Low	\$1,553.345	\$528.137
92	MENA	Libya	High	\$22,990.208	\$28,048.054
93	Europe	Lithuania	Very high	\$43,122.541	\$62,958.910
94	Europe	Luxembourg	Very high	\$116,441.081	\$170,003.978
95	Africa	Madagascar	Low	\$1,624.189	\$552.224
96	Africa	Malawi	Low	\$1,388.778	\$472.185
97	WP	Malaysia	Very high	\$32,020.985	\$46,750.638
98	SEA	Maldives	High	\$32,295.810	\$39,400.888
99	Africa	Mali	Low	\$2,250.231	\$765.079
100	Europe	Malta	Very high	\$53,613.775	\$78,276.112
101	Africa	Mauritania	Medium	\$6,494.854	\$4,351.552
102	SEA	Mauritius	Very high	\$25,662.642	\$37,467.457
103	NAC	Mexico	High	\$19,790.596	\$24,144.527
104	WP	Mongolia	High	\$13,258.310	\$16,175.138
105	Europe	Montenegro	Very high	\$23,994.463	\$35,031.916
106	MENA	Morocco	Medium	\$8,906.543	\$5,967.384
107	Africa	Mozambique	Low	\$1,365.950	\$464.423
108	WP	Myanmar	Medium	\$4,358.168	\$2,919.973
109	Africa	Namibia	Medium	\$9,472.288	\$6,346.433
110	SEA	Nepal	Medium	\$4,477.012	\$2,999.598
111	Europe	Netherlands	Very high	\$60,772.781	\$88,728.260
112	WP	New Zealand	Very high	\$44,851.776	\$65,483.593
113	SACA	Nicaragua	Medium	\$6,490.927	\$4,348.921
114	Africa	Niger	Low	\$1,463.081	\$497.448
115	Africa	Nigeria	Low	\$5,089.592	\$1,730.461
116	Europe	North Macedonia	High	\$18,536.479	\$22,614.504
117	Europe	Norway	Very high	\$69,566.420	\$101,566.973
118	MENA	Oman	Very high	\$34,764.951	\$50,756.828
119	MENA	Pakistan	Low	\$5,794.127	\$1,970.003
120	SACA	Panama	Very high	\$34,572.903	\$50,476.438
121	WP	Papua New Guinea	Medium	\$3,805.425	\$2,549.635
122	SACA	Paraguay	High	\$13,255.396	\$16,171.583
123	SACA	Peru	High	\$13,685.300	\$16,696.066
124	WP	Philippines	Medium	\$10,232.963	\$6,856.085
125	Europe	Poland	Very high	\$39,402.436	\$57,527.557
126	Europe	Portugal	Very high	\$37,970.466	\$55,436.880
127	SACA	Puerto Rico	Very high	\$35,549.244	\$51,901.896
128	MENA	Qatar	Very high	\$109,342.928	\$159,640.675
129	WP	Republic of Korea	Very high	\$48,598.295	\$70,953.511
130	Europe	Republic of Moldova	High	\$15,592.520	\$19,022.874
131	Europe	Romania	Very high	\$37,155.526	\$54,247.068
132	Europe	Russian Federation	Very high	\$29,191.838	\$42,620.083
133	Africa	Rwanda	Low	\$2,787.821	\$947.859
134	NAC	Saint Lucia	High	\$15,490.441	\$18,898.338
135	WP	Samoa	High	\$5,454.167	\$6,654.084
136	Africa	Sao Tome and Principe	Medium	\$4,032.932	\$2,702.064
137	MENA	Saudi Arabia	Very high	\$53,987.648	\$78,821.966
138	Africa	Senegal	Low	\$4,046.282	\$1,375.736
139	Europe	Serbia	Very high	\$22,509.969	\$32,864.555
140	Africa	Sierra Leone	Low	\$1,798.000	\$611.320
141	WP	Singapore	Very high	\$113,370.175	\$165,520.456
142	Europe	Slovakia	Very high	\$35,792.282	\$52,256.732

No.	Region	Country	HDI	GDP per capita (INT \$)	Cost /DALY (INT \$)
143	Europe	Slovenia	Very high	\$45,055.130	\$65,780.490
144	WP	Solomon Islands	Medium	\$1,978.410	\$1,325.535
145	Africa	Somalia	N/A	\$1,142.720	\$-
146	Africa	South Africa	High	\$13,174.246	\$16,072.580
147	Africa	South Sudan	Low	\$434.857	\$147.851
148	Europe	Spain	Very high	\$41,659.042	\$60,822.201
149	SEA	Sri Lanka	High	\$11,824.179	\$14,425.498
150	MENA	State of Palestine	High	\$5,487.769	\$6,695.078
151	MENA	Sudan	Low	\$3,665.303	\$1,246.203
152	NAC	Suriname	High	\$15,562.965	\$18,986.817
153	Europe	Sweden	Very high	\$54,739.804	\$79,920.114
154	Europe	Switzerland	Very high	\$72,826.240	\$106,326.310
155	MENA	Syrian Arab Republic	Medium	N/A	\$-
156	Europe	Tajikistan	Medium	\$4,535.027	\$3,038.468
157	WP	Thailand	Very high	\$19,746.279	\$28,829.567
158	WP	Timor-Leste	Medium	\$2,894.838	\$1,939.541
159	Africa	Togo	Low	\$2,381.771	\$809.802
160	NAC	Trinidad and Tobago	Very high	\$27,222.878	\$39,745.402
161	MENA	Tunisia	High	\$11,106.223	\$13,549.592
162	Europe	Turkey	Very high	\$35,261.856	\$51,482.310
163	Europe	Turkmenistan	High	\$16,647.579	\$20,310.046
164	Africa	Uganda	Low	\$2,817.531	\$957.961
165	Europe	Ukraine	High	\$10,731.440	\$13,092.357
166	MENA	United Arab Emirates	Very high	\$75,486.697	\$110,210.578
167	Europe	United Kingdom	Very high	\$47,100.246	\$68,766.359
168	Africa	United Republic of Tanzania	Low	\$3,128.622	\$1,063.731
169	NAC	United States of America	Very high	\$66,257.748	\$96,736.312
170	SACA	Uruguay	Very high	\$24,595.503	\$35,909.434
171	Europe	Uzbekistan	High	\$8,993.917	\$10,972.579
172	WP	Vanuatu	Medium	\$2,463.910	\$1,650.820
173	SACA	Venezuela	Medium	\$6,864.359	\$4,599.121
174	WP	Viet Nam	High	\$13,234.182	\$16,145.702
175	MENA	Yemen	Low	\$1,739.561	\$591.451
176	Africa	Zambia	Medium	\$3,399.804	\$2,277.869
177	Africa	Zimbabwe	Medium	\$2,174.713	\$1,457.058

Cost per DALY was calculated using HDI-based thresholds: 0.34 times GDP per capita for low HDI countries, 0.67 for medium, 1.22 for high, and 1.46 for very high HDI countries (Daroudi et al., 2021). Projected 2025 GDP per capita values were obtained from *the World Economic Outlook Database (International Monetary Fund)* and the *United Nations Development Programme (2024)*. Monetised DALYs could not be calculated for five countries (Cuba, the Democratic People's Republic of Korea, Guam, Somalia, and the Syrian Arab Republic) due to unavailable data on GDP projections and HDI classification.

Abbreviations. INT \$: international dollars at constant 2017 prices; MENA: Middle East and North Africa; NAC: North America and Caribbean; SACA: South and Central America; SEA: South-East Asia; WP: Western Pacific.

References

1. Abudawood, M. (2019). Diabetes and cancer: A comprehensive review. *Journal of Research in Medical Sciences*, 24(1), 94. https://doi.org/10.4103/jrms.JRMS_242_19
2. Agache, A., Mustăţea, P., Mihalache, O., Bobirca, F. T., Georgescu, D. E., Jauca, C. M., Birligea, A., Doran, H., & Pătraşcu, T. (2018). Diabetes Mellitus as a Risk-factor for Colorectal Cancer Literature Review - Current Situation and Future Perspectives. *Chirurgia*, 113(5), 603. <https://doi.org/10.21614/chirurgia.113.5.603>
3. Agrawal, A., Dave, A., & Jaiswal, A. (2023). Type 2 Diabetes Mellitus in Patients With Polycystic Ovary Syndrome. *Cureus*. <https://doi.org/10.7759/cureus.46859>
4. Agrawal, S., Patel, P., Agrawal, A., Makhijani, N., Markert, R., & Deidrich, W. (2014). Metformin Use and the Risk of Esophageal Cancer in Barrett Esophagus. *Southern Medical Journal*, 107(12), 774-779. <https://doi.org/10.14423/SMJ.0000000000000212>
5. Alkerwi, A., Vernier, C., Sauvageot, N., Crichton, G. E., & Elias, M. F. (2015). Demographic and socioeconomic disparity in nutrition: application of a novel Correlated Component Regression approach. *BMJ Open*, 5(5), e006814. <https://doi.org/10.1136/bmjopen-2014-006814>
6. Alsheredah, N., & Akhtar, S. (2018). Diet, obesity and colorectal carcinoma risk: Results from a national cancer registry-based middle-eastern study. *BMC Cancer*, 18(1), 1-10. <https://doi.org/10.1186/s12885-018-5132-9>
7. American Academy of Orthopaedic Surgeons. (2023). OrthoInfo. <https://orthoinfo.aaos.org/en/diseases--conditions/>
8. American Cancer Society. (2021). Cancer Types. <https://www.cancer.org/cancer/types.html>
9. American College of Gastroenterology. (2024). American College of Gastroenterology, Advancing Gastroenterology, Improving Patient Care. <https://gi.org/topics/barretts-esophagus/>
10. American College of Obstetricians and Gynecologists. (2024). Diseases and conditions. <https://www.acog.org/womens-health/>
11. American Diabetes Association. (2024). American Standards of Medical Care in Diabetes American STANDARDS OF MEDICAL CARE IN DIABETES American Diabetes Association— 2024. *The Journal of Clinical and Applied Research and Education*, 46(Supplement 1), 1-298.
12. American Society of Clinical Oncology's (ASCO). (2023). Types of Cancer. <https://www.cancer.net/cancer-types>
13. American Society of Colon and Rectal Surgeons. (2020). Diseases and conditions. <https://fascrs.org/patients/diseases-and-conditions/a-z/>
14. Anastasi, E., Filardi, T., Tartaglione, S., Lenzi, A., Angeloni, A., & Morano, S. (2018). Linking type 2 diabetes and gynecological cancer: an introductory overview. *Clinical Chemistry and Laboratory Medicine (CCLM)*, 56(9), 1413-1425. <https://doi.org/10.1515/cclm-2017-0982>
15. Anderson, K. E., Anderson, E., Mink, P. J., Ching Ping Hong, Kushi, L. H., Sellers, T. A., Lazovich, D., & Folsom, A. R. (2001). Diabetes and

- endometrial cancer in the Iowa Women's Health Study. *Cancer Epidemiology Biomarkers and Prevention*, 10(6), 611-616.
16. Andrén-Sandberg, Å. (2012). Diagnosis and management of gallbladder polyps. *North American Journal of Medical Sciences*, 4(5), 203-211. <https://doi.org/10.4103/1947-2714.95897>
 17. Anothaisintawee, T., Wiratkapun, C., Lerdsitthichai, P., Kasamesup, V., Wongwaisayawan, S., Srinakarin, J., Hirunpat, S., Woodtichartpreecha, P., Boonlikit, S., Teerawattananon, Y., & Thakkinstian, A. (2013). Risk Factors of Breast Cancer. *Asia Pacific Journal of Public Health*, 25(5), 368-387. <https://doi.org/10.1177/1010539513488795>
 18. Antwi, S. O., Oberg, A. L., Shivappa, N., Bamlet, W. R., Chaffee, K. G., Steck, S. E., Hébert, J. R., & Petersen, G. M. (2016). Pancreatic cancer: associations of inflammatory potential of diet, cigarette smoking and long-standing diabetes. *Carcinogenesis*, 37(5), 481-490. <https://doi.org/10.1093/carcin/bgw022>
 19. Anwar, M. A., Kheir, W. A., Eid, S., Fares, J., Liu, X., Eid, A. H., & Eid, A. A. (2014). Colorectal and Prostate Cancer Risk in Diabetes: Metformin, an Actor behind the Scene. *Journal of Cancer*, 5(9), 736-744. <https://doi.org/10.7150/jca.9726>
 20. Apostoli, P., Magoni, M., Bergonzi, R., Carasi, S., Indelicato, A., Scarcella, C., & Donato, F. (2005). Assessment of reference values for polychlorinated biphenyl concentration in human blood. *Chemosphere*, 61(3), 413-421. <https://doi.org/10.1016/j.chemosphere.2005.02.034>
 21. Arcidiacono, B., Iiritano, S., Nocera, A., Possidente, K., Nevolo, M. T., Ventura, V., Foti, D., Chiefari, E., & Brunetti, A. (2012). Insulin Resistance and Cancer Risk: An Overview of the Pathogenetic Mechanisms. *Experimental Diabetes Research*, 2012, 1-12. <https://doi.org/10.1155/2012/789174>
 22. Arthur, R., Kabat, G. C., Kim, M. Y., Ho, G. Y. F., Chlebowski, R. T., Pan, K., & Rohan, T. E. (2019). Adiposity, history of diabetes, and risk of pancreatic cancer in postmenopausal women. *Annals of Epidemiology*, 29, 23-29.e1. <https://doi.org/10.1016/j.annepidem.2018.09.005>
 23. Asiri, A., Al Qarni, A., & Bakillah, A. (2024). The Interlinking Metabolic Association between Type 2 Diabetes Mellitus and Cancer: Molecular Mechanisms and Therapeutic Insights. *Diagnostics*, 14(19), 2132. <https://doi.org/10.3390/diagnostics14192132>
 24. Au Yeung, S. L., & Schooling, C. M. (2019). Impact of glycemic traits, type 2 diabetes and metformin use on breast and prostate cancer risk: a Mendelian randomization study. *BMJ Open Diabetes Research & Care*, 7(1), e000872. <https://doi.org/10.1136/bmjdr-2019-000872>
 25. Australian Cancer Research Foundation. (2017). Types of cancer. <https://www.acrf.com.au/support-cancer-research/types-of-cancer/>
 26. Baden, R., Rockstroh, J. K., & Buti, M. (2014). Natural history and management of hepatitis C: Does sex play a role? *Journal of Infectious Diseases*, 209(SUPPL. 3), 81-86. <https://doi.org/10.1093/infdis/jiu057>
 27. Baffy, G. (2012). Editorial: Hepatocellular Carcinoma in Type 2 Diabetes: More Than Meets the Eye. *American Journal of Gastroenterology*, 107(1), 53-55. <https://doi.org/10.1038/ajg.2011.390>

28. Balbi, M., Donadon, V., Ghersetti, M., Grazioli, S., Della Valentina, G., Gardenal, R., Dal Mas, M., Casarin, P., Zanette, G., Miranda, C., & Cimarosti, P. (2010). Alcohol and HCV Chronic Infection Are Risk Cofactors of Type 2 Diabetes Mellitus for Hepatocellular Carcinoma in Italy. *International Journal of Environmental Research and Public Health*, 7(4), 1366-1378. <https://doi.org/10.3390/ijerph7041366>
29. Ballotari, P., Vicentini, M., Manicardi, V., Gallo, M., Chiatamone Ranieri, S., Greci, M., & Giorgi Rossi, P. (2017). Diabetes and risk of cancer incidence: results from a population-based cohort study in northern Italy. *BMC Cancer*, 17(1), 703. <https://doi.org/10.1186/s12885-017-3696-4>
30. Bansal, D., Bhansali, A., Kapil, G., Undela, K., & Tiwari, P. (2013). Type 2 diabetes and risk of prostate cancer: a meta-analysis of observational studies. *Prostate Cancer and Prostatic Diseases*, 16(2), 151-158. <https://doi.org/10.1038/pcan.2012.40>
31. Bao, C., Pedersen, N. L., Yang, R., Marseglia, A., Xu, W., Wang, Y., Qi, X., & Xu, W. (2018). Diabetes in midlife and risk of cancer in late life: A nationwide Swedish twin study. *International Journal of Cancer*, 143(4), 793-800. <https://doi.org/10.1002/ijc.31365>
32. Bao, C., Yang, X., Xu, W., Luo, H., Xu, Z., Su, C., & Qi, X. (2013). Diabetes mellitus and incidence and mortality of kidney cancer: a meta-analysis. *Journal of Diabetes and Its Complications*, 27(4), 357-364. <https://doi.org/10.1016/j.jdiacomp.2013.01.004>
33. Barami, K., Lyon, L., & Conell, C. (2017). Type 2 Diabetes Mellitus and Glioblastoma Multiforme-Assessing Risk and Survival: Results of a Large Retrospective Study and Systematic Review of the Literature. *World Neurosurgery*, 106, 300-307. <https://doi.org/10.1016/j.wneu.2017.06.164>
34. Baron, J. A., Weiderpass, E., Newcomb, P. A., Stampfer, M., Titus-Ernstoff, L., Egan, K. M., & Greenberg, E. R. (2001). Metabolic disorders and breast cancer risk (United States). *Cancer Causes & Control: CCC*, 12(10), 875-880. <https://doi.org/10.1023/a:1013796112348>
35. Barrios-Rodríguez, R., García-Esquinas, E., Pérez-Gómez, B., Castaño-Vinyals, G., Llorca, J., de Larrea-Baz, N. F., Olmedo-Requena, R., Vanaclocha-Espi, M., Alguacil, J., Fernández-Tardón, G., Fernández-Navarro, P., Cecchini, L., Lope, V., Gómez-Acebo, I., Aragonés, N., Kogevinas, M., Pollán, M., & Jiménez-Moleón, J. J. (2022). Prostate cancer genetic propensity risk score may modify the association between this tumour and type 2 diabetes mellitus (MCC-Spain study). *Prostate Cancer and Prostatic Diseases*, 25(4), 694-699. <https://doi.org/10.1038/s41391-021-00446-w>
36. Bartosch-Härlid, A., & Andersson, R. (2010). Diabetes Mellitus in Pancreatic Cancer and the Need for Diagnosis of Asymptomatic Disease. *Pancreatology*, 10(4), 423-428. <https://doi.org/10.1159/000264676>
37. Batabyal, P., vander Hoorn, S., Christophi, C., & Nikfarjam, M. (2014). Association of diabetes mellitus and pancreatic adenocarcinoma: a meta-analysis of 88 studies. *Annals of Surgical Oncology*, 21(7), 2453-2462. <https://doi.org/10.1245/s10434-014-3625-6>
38. Baur, D. M., Klotzsche, J., Hamnvik, O.-P. R., Sievers, C., Pieper, L., Wittchen, H.-U., Stalla, G. K., Schmid, R. M., Kales, S. N., & Mantzoros, C. S. (2011). Type 2 diabetes mellitus and medications for type 2 diabetes

- mellitus are associated with risk for and mortality from cancer in a German primary care cohort. *Metabolism*, 60(10), 1363-1371.
<https://doi.org/10.1016/j.metabol.2010.09.012>
39. Beard, E., Brown, J., West, R., Kaner, E., Meier, P., Boniface, S., & Michie, S. (2019). Associations between socio-economic factors and alcohol consumption: A population survey of adults in England. *PLOS ONE*, 14(4), e0216378. <https://doi.org/10.1371/journal.pone.0216378>
 40. Ben, Q., Xu, M., Ning, X., Liu, J., Hong, S., Huang, W., Zhang, H., & Li, Z. (2011). Diabetes mellitus and risk of pancreatic cancer: A meta-analysis of cohort studies. *European Journal of Cancer*, 47(13), 1928-1937.
<https://doi.org/10.1016/j.ejca.2011.03.003>
 41. Benavente, Y., Casabonne, D., Costas, L., Robles, C., Alonso, E., de la Banda, E., Gonzalez-Barca, E., Marcos-Gragera, R., Llorca, J., Tardón, A., Monleon, J. J., Aymerich, M., Campo, E., Gimeno-Vázquez, E., Castaño-Vinyals, G., Aragonés, N., Pollán, M., Kogevinas, M., & de Sanjosé, S. (2018). Established and suggested exposures on CLL/SLL etiology: Results from the CLL-MCC-Spain study. *Cancer Epidemiology*, 52, 106-111.
<https://doi.org/10.1016/j.canep.2017.12.012>
 42. Berbudi, A., Rahmadika, N., Tjahjadi, A. I., & Ruslami, R. (2019). Type 2 Diabetes and its Impact on the Immune System. *Current Diabetes Reviews*, 16(5), 442-449. <https://doi.org/10.2174/1573399815666191024085838>
 43. Berger, S. M., Gislason, G., Moore, L. L., Andersson, C., Torp-Pedersen, C., Denis, G. V., & Schmiegelow, M. D. (2016). Associations between metabolic disorders and risk of cancer in Danish men and women - a nationwide cohort study. *BMC Cancer*, 16(1), 133. <https://doi.org/10.1186/s12885-016-2122-7>
 44. Bernard, L., Reix, N., Benabu, J.-C., Gabriele, V., & Mathelin, C. (2016). Cancer du sein et diabète de type 2: des interactions complexes. *Gynécologie Obstétrique & Fertilité*, 44(12), 701-711.
<https://doi.org/10.1016/j.gyobfe.2016.09.004>
 45. Bjornsdottir, H. H., Rawshani, A., Rawshani, A., Franzén, S., Svensson, A. M., Sattar, N., & Gudbjörnsdottir, S. (2020). A national observation study of cancer incidence and mortality risks in type 2 diabetes compared to the background population over time. *Scientific Reports*, 10(1), 1-12.
<https://doi.org/10.1038/s41598-020-73668-y>
 46. Blanc-Lapierre, A., Spence, A., Karakiewicz, P. I., Aprikian, A., Saad, F., & Parent, M.-É. (2015). Metabolic syndrome and prostate cancer risk in a population-based case-control study in Montreal, Canada. *BMC Public Health*, 15(1), 913. <https://doi.org/10.1186/s12889-015-2260-x>
 47. Bloom, D. E., Cafiero, E., Jané-Llopis, E., Abrahams-Gessel, S., Bloom, L. R., Fathima, S., Feigl, A. B., Gaziano, T., Hamandi, A., Mowafi, M., Pandya, A., Prettnner, K., Rosenberg, L., Seligman, B., Stein, A. Z., & Weinstein, C. (2011). The Global Economic Burden of Noncommunicable Diseases (World Economic Forum, Ed.; Issue September).
http://www3.weforum.org/docs/WEF_Harvard_HE_GlobalEconomicBurdenNonCommunicableDiseases_2011.pdf
 48. Bodmer, M., Becker, C., Meier, C., Jick, S. S., & Meier, C. R. (2012). Use of Antidiabetic Agents and the Risk of Pancreatic Cancer: A Case-Control

- Analysis. *American Journal of Gastroenterology*, 107(4), 620-626.
<https://doi.org/10.1038/ajg.2011.483>
49. Bogumil, D., Cortessis, V. K., Wilkens, L. R., Le Marchand, L., Haiman, C. A., Maskarinec, G., & Setiawan, V. W. (2023). Interethnic Differences in Bladder Cancer Incidence and the Association between Type 2 Diabetes and Bladder Cancer in the Multiethnic Cohort Study. *Cancer Research Communications*, 3(5), 755-762. <https://doi.org/10.1158/2767-9764.CRC-22-0288>
 50. Bonora, Enzo., & DeFronzo, R. A. (2020). Diabetes Complications, Comorbidities and Related Disorders.
<http://link.springer.com/10.1007/978-3-030-36694-0>
 51. Bonovas, S., Filioussi, K., & Tsantes, A. (2004). Diabetes mellitus and risk of prostate cancer: a meta-analysis. *Diabetologia*, 47(6), 1071-1078.
<https://doi.org/10.1007/s00125-004-1415-6>
 52. Bosco, J. L. F., Palmer, J. R., Boggs, D. A., Hatch, E. E., & Rosenberg, L. (2012). Cardiometabolic factors and breast cancer risk in U.S. black women. *Breast Cancer Research and Treatment*, 134(3), 1247-1256.
<https://doi.org/10.1007/s10549-012-2131-4>
 53. Bosdriesz, J. R., Stel, V. S., van Diepen, M., Meuleman, Y., Dekker, F. W., Zoccali, C., & Jager, K. J. (2020). Evidence-based medicine—When observational studies are better than randomized controlled trials. *Nephrology*, 25(10), 737-743. <https://doi.org/10.1111/nep.13742>
 54. Bosetti, C., Rosato, V., Li, D., Silverman, D., Petersen, G. M., Bracci, P. M., Neale, R. E., Muscat, J., Anderson, K., Gallinger, S., Olson, S. H., Miller, A. B., Bas Bueno-de-Mesquita, H., Scelo, G., Janout, V., Holcatova, I., Laggiou, P., Serraino, D., Lucenteforte, E., ... Ghadirian, P. (2014a). Diabetes, antidiabetic medications, and pancreatic cancer risk: an analysis from the International Pancreatic Cancer Case-Control Consortium. *Annals of Oncology: Official Journal of the European Society for Medical Oncology*, 25(10), 2065-2072. <https://doi.org/10.1093/annonc/mdu276>
 55. Bosetti, C., Rosato, V., Polesel, J., Levi, F., Talamini, R., Montella, M., Negri, E., Tavani, A., Zucchetto, A., Franceschi, S., Corrao, G., & la Vecchia, C. (2012). Diabetes mellitus and cancer risk in a network of case-control studies. *Nutrition and Cancer*, 64(5), 643-651.
<https://doi.org/10.1080/01635581.2012.676141>
 56. Bours, M. J. L. (2020). A nontechnical explanation of the counterfactual definition of confounding. *Journal of Clinical Epidemiology*, 121, 91-100.
<https://doi.org/10.1016/j.jclinepi.2020.01.021>
 57. Boustany, A., Onwuzo, S., Almomani, A., Asaad, I., A., B., S., O., & A., A. (2023). Epidemiology and risk of colorectal cancer in patients with a history of *Helicobacter pylori* infection: a population-based study. *ANNALS OF GASTROENTEROLOGY*, 36(2), 203-207.
<https://doi.org/10.20524/aog.2023.0783>
 58. Bowker, S. L., Richardson, K., Marra, C. A., & Johnson, J. A. (2011). Risk of Breast Cancer After Onset of Type 2 Diabetes. *Diabetes Care*, 34(12), 2542-2544. <https://doi.org/10.2337/dc11-1199>
 59. Boyle, P., Boniol, M., Koechlin, A., Robertson, C., Valentini, F., Coppens, K., Fairley, L.-L., Boniol, M., Zheng, T., Zhang, Y., Pasterk, M., Smans, M., Curado, M. P., Mullie, P., Gandini, S., Bota, M., Bolli, G. B., Rosenstock, J.,

- & Autier, P. (2012). Diabetes and breast cancer risk: a meta-analysis. *British Journal of Cancer*, 107(9), 1608-1617. <https://doi.org/10.1038/bjc.2012.414>
60. Braga, C., La Vecchia, C., Negri, E., & Franceschi, S. (1997). Attributable risks for hepatocellular carcinoma in Northern Italy. *European Journal of Cancer*, 33(4), 629-634. [https://doi.org/10.1016/S0959-8049\(96\)00500-X](https://doi.org/10.1016/S0959-8049(96)00500-X)
 61. Brinton, L. A., Berman, M. L., Mortel, R., Twiggs, L. B., Barrett, R. J., Wilbanks, G. D., Lannom, L., & Hoover, R. N. (1992). Reproductive, menstrual, and medical risk factors for endometrial cancer: Results from a case-control study. *American Journal of Obstetrics and Gynecology*, 167(5), 1317-1325. [https://doi.org/10.1016/S0002-9378\(11\)91709-8](https://doi.org/10.1016/S0002-9378(11)91709-8)
 62. Brinton, L. A., Carreon, J. D., Gierach, G. L., McGlynn, K. A., & Gridley, G. (2010). Etiologic factors for male breast cancer in the U.S. Veterans Affairs medical care system database. *Breast Cancer Research and Treatment*, 119(1), 185-192. <https://doi.org/10.1007/s10549-009-0379-0>
 63. Britt, J. P., & Bonci, A. (2013). Alcohol and tobacco: How smoking may promote excessive drinking. *Neuron*, 79(3), 406-407. <https://doi.org/10.1016/j.neuron.2013.07.018>
 64. Brodovicz, K. G., Kou, T. D., Alexander, C. M., O'Neill, E. A., Engel, S. S., Girman, C. J., & Goldstein, B. J. (2012). Impact of diabetes duration and chronic pancreatitis on the association between type 2 diabetes and pancreatic cancer risk. *Diabetes, Obesity and Metabolism*, 14(12), 1123-1128. <https://doi.org/10.1111/j.1463-1326.2012.01667.x>
 65. Bronsveld, H. K., Peeters, P. J. H. L., de Groot, M. C. H., de Boer, A., Schmidt, M. K., & De Bruin, M. L. (2017). Trends in breast cancer incidence among women with type-2 diabetes in British general practice. *Primary Care Diabetes*, 11(4), 373-382. <https://doi.org/10.1016/j.pcd.2017.02.001>
 66. Bugianesi, E. (2005). Review article: steatosis, the metabolic syndrome and cancer. *Alimentary Pharmacology and Therapeutics*, 22(s2), 40-43. <https://doi.org/10.1111/j.1365-2036.2005.02594.x>
 67. Bullard, K. M., Cowie, C. C., Lessem, S. E., Saydah, S. H., Menke, A., Geiss, L. S., Orchard, T. J., Rolka, D. B., & Imperatore, G. (2018). Prevalence of diagnosed diabetes in adults by diabetes type—United States, 2016. *MMWR. Morbidity and Mortality Weekly Report*, 67(12), 359-361. <https://doi.org/10.15585/mmwr.mm6712a2>
 68. Campbell, P. T., Deka, A., Jacobs, E. J., Newton, C. C., Hildebrand, J. S., McCullough, M. L., Limburg, P. J., & Gapstur, S. M. (2010). Prospective study reveals associations between colorectal cancer and type 2 diabetes mellitus or insulin use in men. *Gastroenterology*, 139(4), 1138-1146. <https://doi.org/10.1053/j.gastro.2010.06.072>
 69. Campbell, P. T., Newton, C. C., Freedman, N. D., Koshiol, J., Alavanja, M. C., Beane Freeman, L. E., Buring, J. E., Chan, A. T., Chong, D. Q., Datta, M., Gaudet, M. M., Gaziano, J. M., Giovannucci, E. L., Graubard, B. I., Hollenbeck, A. R., King, L., Lee, I.-M., Linet, M. S., Palmer, J. R., ... McGlynn, K. A. (2016). Body Mass Index, Waist Circumference, Diabetes, and Risk of Liver Cancer for U.S. Adults. *Cancer Research*, 76(20), 6076-6083. <https://doi.org/10.1158/0008-5472.CAN-16-0787>
 70. Canadian Cancer Society. (2023). Cancer types. <https://cancer.ca/en/cancer-information/cancer-types>

71. Cancer Australia. (2023). Cancer types.
<https://www.canceraustralia.gov.au/cancer-types>
72. Cancer Council NSW. (2023). Types of cancer.
<https://www.cancer.org.au/cancer-information/types-of-cancer>
73. Cancer Research UK. (2023). About Cancer.
<https://www.cancerresearchuk.org/about-cancer>
74. Cantiello, F., Cicione, A., Salonia, A., Autorino, R., de Nunzio, C., Briganti, A., Gandaglia, G., Dell'Oglio, P., Capogrosso, P., & Damiano, R. (2015). Association between metabolic syndrome, obesity, diabetes mellitus and oncological outcomes of bladder cancer: A systematic review. *International Journal of Urology*, 22(1), 22-32. <https://doi.org/10.1111/iju.12644>
75. Carreras-Torres, R., Johansson, M., Gaborieau, V., Haycock, P. C., Wade, K. H., Relton, C. L., Martin, R. M., Davey Smith, G., & Brennan, P. (2017). The Role of Obesity, Type 2 Diabetes, and Metabolic Factors in Pancreatic Cancer: A Mendelian Randomization Study. *JNCI: Journal of the National Cancer Institute*, 109(9). <https://doi.org/10.1093/jnci/djx012>
76. Castillo, J. J., Mull, N., Reagan, J. L., Nemr, S., & Mitri, J. (2012). Increased incidence of non-Hodgkin lymphoma, leukemia, and myeloma in patients with diabetes mellitus type 2: a meta-analysis of observational studies. *Blood*, 119(21), 4845-4850. <https://doi.org/10.1182/blood-2011-06-362830>
77. Castillo-Otiniano, C. C., & Yan-Quiroz, E. F. (2019). Blood hypertension and diabetes mellitus as risk factors for breast cancer. *Revista Del Cuerpo Medico Del Hospital Nacional Almanzor Aguinaga Asenjo*, 12(1), 41-45. %3CGo
78. Cavanaugh, J. (2014). Handbook of Epidemiology. In W. Ahrens & I. Pigeot (Eds.), *Journal of the American Statistical Association* (Vol. 101, Issue 473). Springer New York. <https://doi.org/10.1007/978-0-387-09834-0>
79. Cavicchia, P. P., Adams, S. A., Steck, S. E., Hussey, J. R., Liu, J., Daguisé, V. G., & Hebert, J. R. (2013). Racial disparities in colorectal cancer incidence by type 2 diabetes mellitus status. *Cancer Causes & Control*, 24(2), 277-285. <https://doi.org/10.1007/s10552-012-0095-7>
80. Centers for Disease Control and Prevention (CDC). (2023). Types of Cancer.
<https://www.cdc.gov/cancer/>
81. Cerhan, J. R., Wallace, R. B., Folsom, A. R., Potter, J. D., Sellers, T. A., Zheng, W., & Lutz, C. T. (1997). Medical history risk factors for non-Hodgkin's lymphoma in older women. *Journal of the National Cancer Institute*, 89(4), 314-318. <https://doi.org/10.1093/jnci/89.4.314>
82. Ceylan, B., & Özerdoğan, N. (2015). Menopoza girme yaşını etkileyen faktörler ve menopozda yaşam kalitesinin belirlenmesi. *Türk Jinekoloji ve Obstetrik Dernegi Dergisi*, 12(1), 43-49.
<https://doi.org/10.4274/tjod.79836>
83. Chakraborty, S. (2014). Analysis of NHANES 1999-2002 data reveals noteworthy association of alcohol consumption with obesity. *Annals of Gastroenterology*, 27(3), 250-257.
84. Chao, C., & Page, J. H. (2008). Type 2 Diabetes Mellitus and Risk of Non-Hodgkin Lymphoma: A Systematic Review and Meta-Analysis. *American Journal of Epidemiology*, 168(5), 471-480.
<https://doi.org/10.1093/aje/kwn160>

85. Chen, C. B., Eurich, D. T., Majumdar, S. R., & Johnson, J. A. (2018). Risk of prostate cancer across different racial/ethnic groups in men with diabetes: a retrospective cohort study. *Diabetic Medicine*, 35(1), 107-111. <https://doi.org/10.1111/dme.13536>
86. Chen, C. H., Lin, C. L., Hsu, C. Y., & Kao, C. H. (2020). Insulin enhances and metformin reduces risk of colorectal carcinoma in type-2 diabetes. *Qjm*, 113(3), 194-200. <https://doi.org/10.1093/qjmed/hcz253>
87. Chen, C.-H., Lin, C.-L., Hsu, C.-Y., & Kao, C.-H. (2019). Insulin enhances and metformin reduces risk of colorectal carcinoma in type-2 diabetes. *QJM: An International Journal of Medicine*, 113(3), 194-200. <https://doi.org/10.1093/qjmed/hcz253>
88. Chen, H.-P., Shieh, J.-J., Chang, C.-C., Chen, T.-T., Lin, J.-T., Wu, M.-S., Lin, J.-H., & Wu, C.-Y. (2013). Metformin decreases hepatocellular carcinoma risk in a dose-dependent manner: population-based and in vitro studies. *Gut*, 62(4), 606-615. <https://doi.org/10.1136/gutjnl-2011-301708>
89. Chen, J., Han, Y., Xu, C., Xiao, T., & Wang, B. (2015). Effect of type 2 diabetes mellitus on the risk for hepatocellular carcinoma in chronic liver diseases. *European Journal of Cancer Prevention*, 24(2), 89-99. <https://doi.org/10.1097/CEJ.0000000000000038>
90. Chen, J., Yuan, S., Fu, T., Ruan, X., Qiao, J., Wang, X., Li, X., Gill, D., Burgess, S., Giovannucci, E. L., & Larsson, S. C. (2023). Gastrointestinal Consequences of Type 2 Diabetes Mellitus and Impaired Glycemic Homeostasis: A Mendelian Randomization Study. *Diabetes Care*, 46(4), 828-835. <https://doi.org/10.2337/dc22-1385>
91. Chen, M., Luo, Y., Yang, S., Xu, L., Li, N., Li, H., & Qu, S. (2018). Association of diabetes mellitus with prostate cancer grade and prostate-specific antigen in Chinese biopsy population. *Diabetes Research and Clinical Practice*, 141(301), 80-87. <https://doi.org/10.1016/j.diabres.2018.04.022>
92. Chen, S., Phuc, P. T., Nguyen, P., Burton, W., Lin, S., Lin, W., Lu, C. Y., Hsu, M., Cheng, C., & Hsu, J. C. (2023). A novel prediction model of the risk of pancreatic cancer among diabetes patients using multiple clinical data and machine learning. *Cancer Medicine*, 12(19), 19987-19999. <https://doi.org/10.1002/cam4.6547>
93. Chen, Y., Cheng, X., Sun, C., Kim, N. H., Kailas, S., Qureshi, F., Karadsheh, Z., Wu, Y., Hu, L., Zhou, Z., Bhan, C., Kim, K. Y., Manem, R., Cheng, C., & Zhou, Q. (2022). Is metformin use associated with a reduced risk of oesophageal cancer? A systematic review and meta-analysis. *Postgraduate Medical Journal*, 98(1165), 866-870. <https://doi.org/10.1136/postgradmedj-2021-140432>
94. Chen-Sankey, J. C., Mead-Morse, E. L., Le, D., Rose, S. W., Quisenberry, A. J., Delnevo, C. D., & Choi, K. (2021). Cigar-Smoking Patterns by Race/Ethnicity and Cigar Type: A Nationally Representative Survey Among U.S. Adults. *American Journal of Preventive Medicine*, 60(1), 87-94. <https://doi.org/10.1016/j.amepre.2020.07.005>
95. Chi, Y., Wang, X., Jia, J., & Huang, T. (2022). Smoking Status and Type 2 Diabetes, and Cardiovascular Disease: A Comprehensive Analysis of Shared Genetic Etiology and Causal Relationship. *Frontiers in Endocrinology*, 13(February), 1-16. <https://doi.org/10.3389/fendo.2022.809445>

96. Chiquete, E., Nu, P., Panduro, A., Perspectiva, C., La, R. D. E., Mellitus, D., La, C., & Guadalajara, G. (2001). Perspectiva histórica de la diabetes mellitus . Comprendiendo la enfermedad. *Diabetes*, 111, 5-10.
97. Chodick, G., Heymann, A. D., Rosenmann, L., Green, M. S., Flash, S., Porath, A., Kokia, E., & Shalev, V. (2010). Diabetes and risk of incident cancer: a large population-based cohort study in Israel. *Cancer Causes & Control*, 21(6), 879-887. <https://doi.org/10.1007/s10552-010-9515-8>
98. Choi, Y. H., Park, Y.-M., Hur, K. J., Ha, U.-S., Hong, S.-H., Lee, J. Y., Kim, S. W., Han, K., Ko, S. H., & Park, Y. H. (2022). Duration of diabetes mellitus and risk of kidney and bladder cancer: a longitudinal nationwide cohort study. *American Journal of Cancer Research*, 12(8), 4050-4061. <https://doi.org/10.1097/JU.0000000000002079.01>
99. Cleveland, R. J., North, K. E., Stevens, J., Teitelbaum, S. L., Neugut, A. I., & Gammon, M. D. (2012). The association of diabetes with breast cancer incidence and mortality in the Long Island Breast Cancer Study Project. *Cancer Causes & Control*, 23(7), 1193-1203. <https://doi.org/10.1007/s10552-012-9989-7>
100. Cochrane Scientific Committee. (2017). Review of the development of the risk of bias tool for nonrandomised studies for interventions - ROBINS-I (Recommendation statement/report).
101. Cofield, S. S., Corona, R. V., & Allison, D. B. (2010). Use of Causal Language in Observational Studies of Obesity and Nutrition. *Obesity Facts*, 3(6), 353-356. <https://doi.org/10.1159/000322940>
102. Collier, A., Ghosh, S., McGlynn, B., & Hollins, G. (2012). Prostate Cancer, Androgen Deprivation Therapy, Obesity, the Metabolic Syndrome, Type 2 Diabetes, and Cardiovascular Disease. *American Journal of Clinical Oncology*, 35(5), 504-509. <https://doi.org/10.1097/COC.0b013e318201a406>
103. Colmers, I. N., Majumdar, S. R., Yasui, Y., Bowker, S. L., Marra, C. A., & Johnson, J. A. (2013). Detection bias and overestimation of bladder cancer risk in type 2 diabetes: A matched cohort study. *Diabetes Care*, 36(10), 3070-3075. <https://doi.org/10.2337/dc13-0045>
104. Coman, L. I., Coman, O. A., Bădăraș, I. A., Păunescu, H., & Ciocîrlan, M. (2021). Association between Liver Cirrhosis and Diabetes Mellitus: A Review on Hepatic Outcomes. *Journal of Clinical Medicine*, 10(2), 262. <https://doi.org/10.3390/jcm10020262>
105. Consumer Financial Protection Bureau. (2014). Using publicly available information to proxy for unidentified race and ethnicity: A methodology and assessment. https://files.consumerfinance.gov/f/201409_cfpb_report_proxy-methodology.pdf.
106. Cooper, C., Booth, A., Varley-Campbell, J., Britten, N., & Garside, R. (2018). Defining the process to literature searching in systematic reviews: A literature review of guidance and supporting studies. *BMC Medical Research Methodology*, 18(1), 1-14. <https://doi.org/10.1186/s12874-018-0545-3>
107. Cortellini, A., D'Alessio, A., Cleary, S., Buti, S., Bersanelli, M., Bordi, P., Tonini, G., Vincenzi, B., Tucci, M., Russo, A., Pantano, F., Russano, M., Stucci, L. S., Sergi, M. C., Falconi, M., Zarzana, M. A., Santini, D., Spagnolo, F., Tanda, E. T., ... Pinato, D. J. (2023). Type 2 Diabetes Mellitus and Efficacy Outcomes from Immune Checkpoint Blockade in Patients with

- Cancer. *Clinical Cancer Research*, 29(14), 2714-2724.
<https://doi.org/10.1158/1078-0432.CCR-22-3116>
108. Cosmin Stan, M., & Paul, D. (2024). Diabetes and Cancer: A Twisted Bond. *Oncology Reviews*, 18(May), 1-9. <https://doi.org/10.3389/or.2024.1354549>
 109. Craig, E. R., Londoño, A. I., Norian, L. A., & Arend, R. C. (2016). Metabolic risk factors and mechanisms of disease in epithelial ovarian cancer: A review. *Gynecologic Oncology*, 143(3), 674-683.
<https://doi.org/10.1016/j.ygyno.2016.10.005>
 110. Crawley, D., Chamberlain, F., Garmo, H., Rudman, S., Zethelius, B., Holmberg, L., Adolfsson, J., Stattin, P., Carroll, P., & van Hemerijck, M. (2018). A systematic review of the literature exploring the interplay between prostate cancer and type two diabetes mellitus. *Ecancermedicalscience*, 12. <https://doi.org/10.3332/ecancer.2018.802>
 111. Crichton, J., Hickman, M., Campbell, R., Batista-Ferrer, H., & Macleod, J. (2015). Socioeconomic factors and other sources of variation in the prevalence of genital chlamydia infections: A systematic review and meta-analysis. *BMC Public Health*, 15(1), 729. <https://doi.org/10.1186/s12889-015-2069-7>
 112. Cui, F., Zhang, L., Yu, C., Hu, S., & Zhang, Y. (2016). Estimation of the Disease Burden Attributable to 11 Risk Factors in Hubei Province, China: A Comparative Risk Assessment. *International Journal of Environmental Research and Public Health*, 13(10).
<https://doi.org/10.3390/ijerph13100944>
 113. Cui, Y., & Andersen, D. K. (2012). Diabetes and pancreatic cancer. *Endocrine-Related Cancer*, 19(5), F9-F26. <https://doi.org/10.1530/ERC-12-0105>
 114. Dąbrowski, M. (2021). Diabetes, Antidiabetic Medications and Cancer Risk in Type 2 Diabetes: Focus on SGLT-2 Inhibitors. *International Journal of Molecular Sciences*, 22(4), 1680. <https://doi.org/10.3390/ijms22041680>
 115. Dare, S., Mackay, D. F., & Pell, J. P. (2015). Relationship between smoking and obesity: A cross-sectional study of 499,504 middle-aged adults in the UK general population. *PLoS ONE*, 10(4), 1-12.
<https://doi.org/10.1371/journal.pone.0123579>
 116. Daroudi, R., Akbari Sari, A., Nahvijou, A., & Faramarzi, A. (2021). Cost per DALY averted in low, middle- and high-income countries: evidence from the global burden of disease study to estimate the cost-effectiveness thresholds. *Cost Effectiveness and Resource Allocation*, 19(1), 7.
<https://doi.org/10.1186/s12962-021-00260-0>
 117. Davis, T. M. E., & Davis, W. A. (2022). The relationship between pancreatic cancer and type 2 diabetes: the Fremantle Diabetes Study Phase I. *Internal Medicine Journal*, 52(7), 1258-1262. <https://doi.org/10.1111/imj.15846>
 118. Davtalab Esmaeili, E., Asadollahi, K., Delpisheh, A., Sayehmiri, K., & Azizi, H. (2019). Metabolic Syndrome and Risk of Colorectal Cancer: A Case-Control Study. *International Journal of Cancer Management*, 12(10), e84627. <https://doi.org/10.5812/ijcm.84627>
 119. de Bruijn, K. M. J., Arends, L. R., Hansen, B. E., Leeftang, S., Ruiter, R., & van Eijck, C. H. J. (2013). Systematic review and meta-analysis of the association between diabetes mellitus and incidence and mortality in

- breast and colorectal cancer. *British Journal of Surgery*, 100(11), 1421-1429. <https://doi.org/10.1002/bjs.9229>
120. De Bruijn, K. M. J., Ruiter, R., de Keyser, C. E., Hofman, A., Stricker, B. H., & van Eijck, C. H. J. (2014). Detection bias may be the main cause of increased cancer incidence among diabetics: Results from the Rotterdam Study. *European Journal of Cancer*, 50(14), 2449-2455. <https://doi.org/10.1016/j.ejca.2014.06.019>
 121. De Cassai, A., Boscolo, A., Zarantonello, F., Pettenuzzo, T., Sella, N., Geraldini, F., Munari, M., & Navalesi, P. (2023). Enhancing study quality assessment: an in-depth review of risk of bias tools for meta-analysis—a comprehensive guide for anesthesiologists. *Journal of Anesthesia, Analgesia and Critical Care*, 3(1), 44. <https://doi.org/10.1186/s44158-023-00129-z>
 122. de Jong, R. G. P. J., Burden, A. M., de Kort, S., van Herk-Sukel, M. P. P., Vissers, P. A. J., Janssen, P. K. C., Haak, H. R., Masclee, A. A. M., de Vries, F., & Janssen-Heijnen, M. L. G. (2017). Impact of detection bias on the risk of gastrointestinal cancer and its subsites in type 2 diabetes mellitus. *European Journal of Cancer*, 79, 61-71. <https://doi.org/10.1016/j.ejca.2017.03.039>
 123. de Jong, R. G. P. J., Peeters, P. J. H. L., Burden, A. M., de Bruin, M. L., Haak, H. R., Masclee, A. A. M., de Vries, F., & Janssen-Heijnen, M. L. G. (2018). Gastrointestinal cancer incidence in type 2 diabetes mellitus; results from a large population-based cohort study in the UK. *Cancer Epidemiology*, 54(December 2017), 104-111. <https://doi.org/10.1016/j.canep.2018.04.008>
 124. de Kort, S., Masclee, A. A. M., Sanduleanu, S., Weijenberg, M. P., van Herk-Sukel, M. P. P., Oldenhof, N. J. J., van den Bergh, J. P. W., Haak, H. R., & Janssen-Heijnen, M. L. (2017). Higher risk of colorectal cancer in patients with newly diagnosed diabetes mellitus before the age of colorectal cancer screening initiation. *Scientific Reports*, 7(April), 1-8. <https://doi.org/10.1038/srep46527>
 125. de Kort, S., Simons, C. C. J. M., van den Brandt, P. A., Goldbohm, R. A., Arts, I. C. W., de Bruine, A. P., Janssen-Heijnen, M. L. G., Sanduleanu, S., Masclee, A. A. M., & Weijenberg, M. P. (2016). Diabetes mellitus type 2 and subsite-specific colorectal cancer risk in men and women: Results from the Netherlands Cohort Study on diet and cancer. *European Journal of Gastroenterology and Hepatology*, 28(8), 896-903. <https://doi.org/10.1097/MEG.0000000000000626>
 126. De Mesquita, H. B. B., Maisonneuve, P., Moerman, C. J., & Walker, A. M. (1992). Aspects of medical history and exocrine carcinoma of the pancreas: A population-based case-control study in The Netherlands. *International Journal of Cancer*, 52(1), 17-23. <https://doi.org/10.1002/ijc.2910520105>
 127. De Rosa, S., Arcidiacono, B., Chiefari, E., Brunetti, A., Indolfi, C., & Foti, D. P. (2018). Type 2 Diabetes Mellitus and Cardiovascular Disease: Genetic and Epigenetic Links. *Frontiers in Endocrinology*, 9(JAN), 1-13. <https://doi.org/10.3389/fendo.2018.00002>
 128. De Socio, G. V., Pasqualini, M., Ricci, E., Maggi, P., Orofino, G., Squillace, N., Menzaghi, B., Madeddu, G., Taramasso, L., Francisci, D., Bonfanti, P., Vichi, F., Dell'omo, M., Pieroni, L., Santoro, C., Quirino, T., Farinazzo, M., Magne, F., Molteni, C., ... Guastavigna, M. (2020). Smoking habits in HIV-

- infected people compared with the general population in Italy: A cross-sectional study. *BMC Public Health*, 20(1), 1-10.
<https://doi.org/10.1186/s12889-020-08862-8>
129. de Souza, A., Irfan, K., Masud, F., & Saif, M. W. (2016). Diabetes Type 2 and Pancreatic Cancer: A History Unfolding. *Journal of the Pancreas*, 17(2), 144-148. %3CGo
 130. Dekkers, O. M., Vandenbroucke, J. P., Cevallos, M., Renehan, A. G., Altman, D. G., & Egger, M. (2019). COSMOS-E: Guidance on conducting systematic reviews and meta-analyses of observational studies of etiology. *PLOS Medicine*, 16(2), e1002742.
<https://doi.org/10.1371/journal.pmed.1002742>
 131. del Carmen, M. G., Halpern, E. F., Kopans, D. B., Moy, B., Moore, R. H., Goss, P. E., & Hughes, K. S. (2007). Mammographic Breast Density and Race. *American Journal of Roentgenology*, 188(4), 1147-1150.
<https://doi.org/10.2214/AJR.06.0619>
 132. Deng, L., Gui, Z., Zhao, L., Wang, J., & Shen, L. (2012). Diabetes Mellitus and the Incidence of Colorectal Cancer: An Updated Systematic Review and Meta-Analysis. *Digestive Diseases and Sciences*, 57(6), 1576-1585.
<https://doi.org/10.1007/s10620-012-2055-1>
 133. DeVito, N. J., & Goldacre, B. (2019). Catalogue of bias: publication bias. *BMJ Evidence-Based Medicine*, 24(2), 53-54.
<https://doi.org/10.1136/bmjebm-2018-111107>
 134. Dianatinasab, M., Fararouei, M., Daneshi, N., Rezaian, S., Mohammadianpanah, M., Chaman, R., & Ghiasvand, R. (2019). Heterogeneity in risk factors for ductal and lobular breast carcinomas: A case-control study. *International Journal of Cancer*, 145(11), 2917-2925.
<https://doi.org/10.1002/ijc.32182>
 135. Diaz-Algorri, Y., Lozada, M. E., Lopez, S. M., Bertran-Rodriguez, C. E., Gonzalez-Hernandez, C. M., Gonzalez, D., Perez-Cardona, C. M., Hernandez, J., Pedrosa, C., Toro, D. H., Gonzalez-Pons, M., Cruz-Correa, M., Díaz-Algorri, Y., Lozada, M. E., López, S. M., Bertrán-Rodríguez, C. E., González-Hernández, C. M., González, D., Pérez-Cardona, C. M., ... Cruz-Correa, M. (2015). Type 2 diabetes mellitus and colorectal neoplasia risk in Hispanics: a case-control study. *Journal of Diabetes and Its Complications*, 29(4), 502-507. <https://doi.org/10.1016/j.jdiacomp.2015.01.010>
 136. Dickerman, B. A., Batista, J. L., Markt, S. C., Torfadottir, J. E., Valdimarsdottir, U. A., Steingrimsdottir, L., Aspelund, T., Fall, K., Wilson, K. M., Mucci, L. A., Giovannucci, E., Gudnason, V., Sigurdardottir, L. G., & Tryggvadottir Barbra A.; ORCID: <http://orcid.org/0000-0003-2843-687X>, L. A. O.-D. (2018). Midlife metabolic factors and prostate cancer risk in later life. *International Journal of Cancer*, 142(6), 1166-1173.
<https://doi.org/http://dx.doi.org/10.1002/ijc.31142>
 137. Diemer, E. W., Hudson, J. I., & Javaras, K. N. (2021). More (Adjustment) Is Not Always Better: How Directed Acyclic Graphs Can Help Researchers Decide Which Covariates to Include in Models for the Causal Relationship between an Exposure and an Outcome in Observational Research. *Psychotherapy and Psychosomatics*, 90(5), 289-298.
<https://doi.org/10.1159/000517104>

138. Digitale, J. C., Martin, J. N., & Glymour, M. M. (2022). Tutorial on directed acyclic graphs. *Journal of Clinical Epidemiology*, 142, 264-267.
<https://doi.org/10.1016/j.jclinepi.2021.08.001>
139. Disney-Hogg, L., Sud, A., Law, P. J., Cornish, A. J., Kinnersley, B., Ostrom, Q. T., Labreche, K., Eckel-Passow, J. E., Armstrong, G. N., Claus, E. B., Il'yasova, D., Schildkraut, J., Barnholtz-Sloan, J. S., Olson, S. H., Bernstein, J. L., Lai, R. K., Swerdlow, A. J., Simon, M., Hoffmann, P., ... Houlston, R. S. (2018). Influence of obesity-related risk factors in the aetiology of glioma. *British Journal of Cancer*, 118(7), 1020-1027.
<https://doi.org/10.1038/s41416-018-0009-x>
140. Donadon, V. (2010). Glycated hemoglobin and antidiabetic strategies as risk factors for hepatocellular carcinoma. *World Journal of Gastroenterology*, 16(24), 3025. <https://doi.org/10.3748/wjg.v16.i24.3025>
141. Donadon, V., Balbi, M., Mas, M. D., Casarin, P., & Zanette, G. (2010). Metformin and reduced risk of hepatocellular carcinoma in diabetic patients with chronic liver disease. *Liver International*, 30(5), 750-758.
<https://doi.org/10.1111/j.1478-3231.2010.02223.x>
142. Dong, W., Zhang, D.-L., Wang, Z.-H., Lv, C.-Z., Zhang, P., & Zhang, H. (2022). Different types of diabetes mellitus and risk of thyroid cancer: A meta-analysis of cohort studies. *FRONTIERS IN ENDOCRINOLOGY*, 13.
<https://doi.org/10.3389/fendo.2022.971213>
143. Drahos, J., Li, L., Jick, S. S., & Cook, M. B. (2016). Metabolic syndrome in relation to Barrett's esophagus and esophageal adenocarcinoma: Results from a large population-based case-control study in the Clinical Practice Research Datalink. *Cancer Epidemiology*, 42, 9-14.
<https://doi.org/10.1016/j.canep.2016.02.008>
144. Dulskas, A., Patasius, A., Kaceniene, A., Linkeviciute-Ulinskiene, D., Zabuliene, L., & Smailyte, G. (2020). A Cohort Study of Antihyperglycemic Medication Exposure and Gastric Cancer Risk. *Journal of Clinical Medicine*, 9(2), 435. <https://doi.org/10.3390/jcm9020435>
145. Dulskas, A., Patasius, A., Linkeviciute-Ulinskiene, D., Zabuliene, L., Urbonas, V., & Smailyte, G. (2020). Positive effect of metformin treatment in colorectal cancer patients with type 2 diabetes: national cohort study. *European Journal of Cancer Prevention*, 29(4), 289-293.
<https://doi.org/10.1097/CEJ.0000000000000547>
146. Eibl, G., Cruz-Monserrate, Z., Korc, M., Petrov, M. S., Goodarzi, M. O., Fisher, W. E., Habtezion, A., Lugea, A., Pandol, S. J., Hart, P. A., Andersen, D. K., & Consortium Study Chronic, P. (2018). Diabetes Mellitus and Obesity as Risk Factors for Pancreatic Cancer. *Journal of the Academy of Nutrition and Dietetics*, 118(4), 555-567.
<https://doi.org/10.1016/j.jand.2017.07.005>
147. Eijgenraam, P., Heinen, M. M., Verhage, B. A. J., Keulemans, Y. C., Schouten, L. J., & van den Brandt, P. A. (2013). Diabetes type II, other medical conditions and pancreatic cancer risk: a prospective study in The Netherlands. *British Journal of Cancer*, 109(11), 2924-2932.
<https://doi.org/10.1038/bjc.2013.629>
148. Ekholuenetale, M., Onuoha, H., Ekholuenetale, C. E., Barrow, A., & Nzoputam, C. I. (2021). Socioeconomic Inequalities in Human Immunodeficiency Virus (HIV) Sero-Prevalence among Women in Namibia:

- Further Analysis of Population-Based Data. *International Journal of Environmental Research and Public Health*, 18(17), 9397. <https://doi.org/10.3390/ijerph18179397>
149. Elangovan, A., Skeans, J., Landsman, M., Ali, S. M. J., Elangovan, A. G., Kaelber, D. C., Sandhu, D. S., & Cooper, G. S. (2021). Colorectal Cancer, Age, and Obesity-Related Comorbidities: A Large Database Study. *Digestive Diseases and Sciences*, 66(9), 3156-3163. <https://doi.org/10.1007/s10620-020-06602-x>
 150. El-Serag, H. (2001). The role of diabetes in hepatocellular carcinoma: a case-control study among United States veterans. *The American Journal of Gastroenterology*, 96(8), 2462-2467. [https://doi.org/10.1016/S0002-9270\(01\)02617-X](https://doi.org/10.1016/S0002-9270(01)02617-X)
 151. ERN Board of Member States. (2020). European Reference Networks. <https://rare-liver.eu/healthcare-professionals/disease-specific-information/>
 152. Fall, K., Garmo, H., Gudbjörnsdottir, S., Stattin, P., & Zethelius, B. (2013). Diabetes Mellitus and Prostate Cancer Risk; A Nationwide Case-Control Study within PCBaSe Sweden. *Cancer Epidemiology, Biomarkers & Prevention*, 22(6), 1102-1109. <https://doi.org/10.1158/1055-9965.EPI-12-1046>
 153. Fang, H., Shan, S., Zhou, Y., & Zhong, L. (2018). Diabetes mellitus and the risk of gastrointestinal cancer in women compared with men: a meta-analysis of cohort studies. *BMC Cancer*, 18(1), 422. <https://doi.org/10.1186/s12885-018-4351-4>
 154. Fang, H., Yao, B., Yan, Y., Xu, H., Liu, Y., Tang, H., Zhou, J., Cao, L., Wang, W., Zhang, J., Zhao, L., Chen, X., Zhang, F., & Zhao, Y. (2013). Diabetes mellitus increases the risk of bladder cancer: An updated meta-analysis of observational studies. *Diabetes Technology and Therapeutics*, 15(11), 914-922. <https://doi.org/10.1089/dia.2013.0131>
 155. Feng, L.-H., Bu, K.-P., Ren, S., Yang, Z., Li, B.-X., & Deng, C.-E. (2020). Nomogram for Predicting Risk of Digestive Carcinoma Among Patients with Type 2 Diabetes. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy*, Volume 13, 1763-1770. <https://doi.org/10.2147/DMSO.S251063>
 156. Feola, T., Puliani, G., Sesti, F., Modica, R., Centello, R., Minotta, R., Cannavale, G., Di Meglio, S., Di Vito, V., Lauretta, R., Appetecchia, M., Colao, A., Lenzi, A., Isidori, A. M., Faggiano, A., & Giannetta, E. (2022). Risk factors for gastroenteropancreatic neuroendocrine neoplasms (GEP-NENs): a three-centric case-control study. *Journal of Endocrinological Investigation*, 45(4), 849-857. <https://doi.org/10.1007/s40618-021-01715-0>
 157. Ferlay, J., Colombet, M., Soerjomataram, I., Parkin, D. M., Piñeros, M., Znaor, A., & Bray, F. (2024). Global Cancer Observatory: Cancer Today. International Agency for Research on Cancer. <https://doi.org/10.1002/ijc.33588>
 158. Ferlay J, Laversanne M, Ervik M, et al. (n.d.). Global cancer observatory: cancer tomorrow. In Lyon, France: International Agency for Research on Cancer. International Agency for Research on Cancer. Retrieved August 22, 2023, from <https://gco.iarc.fr/tomorrow/en>
 159. Ferreira, A., Chodankar, S., Vaz, F., D'souza, D., & Kulkarni, M. (2021). Risk factors for colorectal cancer in Goa, India: A Hospital-based Case-

- Control study. *Indian Journal of Community Medicine*, 46(3), 474.
https://doi.org/10.4103/ijcm.IJCM_848_20
160. Ferrís-i-Tortajada, J., García-i-Castell, J., Berbel-Tornero, O., & Ortega-García, J. A. (2011). Factores de riesgo constitucionales en el cáncer de próstata. *Actas Urológicas Españolas*, 35(5), 282-288.
<https://doi.org/10.1016/j.acuro.2010.12.009>
 161. Fischborn, A. R., Andreis, J. D., Wambier, L. M., Pedroso, C. M., Claudino, M., & Nobre Franco, G. C. (2023). Performance of panoramic radiography compared with computed tomography in the evaluation of pathological changes in the maxillary sinuses: a systematic review and meta-analysis. *Dentomaxillofacial Radiology*, 52(5).
<https://doi.org/10.1259/dmfr.20230067>
 162. Foong, R. E., & Vicendese, D. (2025). Considerations for Causal Inference Studies. *Respirology*. <https://doi.org/10.1111/resp.70018>
 163. Fortuny, J., Benavente, Y., Bosch, R., García-Villanueva, M., de Sevilla, A. F., & de Sanjosé, S. (2005). Type 2 diabetes mellitus, its treatment and risk for lymphoma. *European Journal of Cancer*, 41(12), 1782-1787.
<https://doi.org/10.1016/j.ejca.2005.04.032>
 164. Fortuny, J., Sima, C., Bayuga, S., Zauber, A. G., Wilcox, H., Pulick, K., Faulkner, S., & Olson, S. H. (2009). Risk of endometrial cancer in relation to medical conditions and medication use. *Cancer Epidemiology Biomarkers and Prevention*, 18(5), 1448-1456.
<https://doi.org/http://dx.doi.org/10.1158/1055-9965.EPI-08-0936>
 165. Franzen, S., Svensson, A.-M., Björnsdóttir, H. H., Rawshani, A., Gudbjörnsdóttir, S., & Sattar, N. (2018). Cancer incidence and mortality among 457,473 persons with type 2 diabetes compared to 2,287,365 matched controls in Sweden: An observational study. *Diabetologia*, 61(Supplement 1), S591.
<http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed19&NEWS=N&AN=624031437>
 166. Friberg, E., Orsini, N., Mantzoros, C. S., & Wolk, A. (2007). Diabetes mellitus and risk of endometrial cancer: a meta-analysis. *Diabetologia*, 50(7), 1365-1374. <https://doi.org/10.1007/s00125-007-0681-5>
 167. Fröhlich, E., & Wahl, R. (2012). Do antidiabetic medications play a specific role in differentiated thyroid cancer compared to other cancer types? *Diabetes, Obesity and Metabolism*, 14(3), 204-213.
<https://doi.org/10.1111/j.1463-1326.2011.01491.x>
 168. Fussey, J. M., Beaumont, R. N., Wood, A. R., Vaidya, B., Smith, J., Tyrrell, J., J.M., F., R.N., B., A.R., W., B., V., J., S., & Tyrrell J. AO - Fussey, J. M. O. <https://orcid.org/0000-0001-9971-4334>. (2020). Does obesity cause thyroid cancer? A mendelian randomization study. *Journal of Clinical Endocrinology & Metabolism*, 105(7), dgaa250.
<https://doi.org/10.1210/clinem/dgaa250>
 169. Gallagher, E. J., & LeRoith, D. (2013). Diabetes, antihyperglycemic medications and cancer risk. *Current Opinion in Endocrinology, Diabetes & Obesity*, 20(5), 485-494.
<https://doi.org/10.1097/01.med.0000433065.16918.83>
 170. Gapstur, S. M., Patel, A. v, Diver, W. R., Hildebrand, J. S., Gaudet, M. M., Jacobs, E. J., & Campbell, P. T. (2012). Type II diabetes mellitus and the

- incidence of epithelial ovarian cancer in the cancer prevention study-II nutrition cohort. *Cancer Epidemiology Biomarkers and Prevention*, 21(11), 2000-2005. <https://doi.org/http://dx.doi.org/10.1158/1055-9965.EPI-12-0867>
171. Garawi, F., Devries, K., Thorogood, N., & Uauy, R. (2014). Global differences between women and men in the prevalence of obesity: Is there an association with gender inequality? *European Journal of Clinical Nutrition*, 68(10), 1101-1106. <https://doi.org/10.1038/ejcn.2014.86>
 172. García-Esquinas, E., Guinó, E., Castaño-Vinyals, G., Pérez-Gómez, B., Llorca, J., Altzibar, J. M., Peiró-Pérez, R., Martín, V., Moreno-Iribas, C., Tardón, A., Caballero, F. J., Puig-Vives, M., Guevara, M., Villa, T. F., Salas, D., Amiano, P., Dierssen-Sotos, T., Pastor-Barriuso, R., Sala, M., ... Pollán, M. (2016). Association of diabetes and diabetes treatment with incidence of breast cancer. *Acta Diabetologica*, 53(1), 99-107. <https://doi.org/10.1007/s00592-015-0756-6>
 173. Garcia-Huidobro, D., & Michael Oakes, J. (2017). Squeezing observational data for better causal inference: Methods and examples for prevention research. *International Journal of Psychology*, 52(2), 96-105. <https://doi.org/10.1002/ijop.12275>
 174. Garg, S. K., Maurer, H., Reed, K., & Selagamsetty, R. (2014). Diabetes and cancer: two diseases with obesity as a common risk factor. *Diabetes, Obesity & Metabolism*, 16(2), 97-110. <https://doi.org/10.1111/dom.12124>
 175. Gaskell, A. L., & Sleight, J. W. (2020). An Introduction to Causal Diagrams for Anesthesiology Research. *Anesthesiology*, 132(5), 951-967. <https://doi.org/10.1097/ALN.0000000000003193>
 176. Gasparini, S., Neri, S., Brigo, F., Cianci, V., Mammi, A., Pascarella, A., Manzo, L., Benna, P., Striano, P., Martino, T., Romoli, M., Muccioli, L., Nilo, A., Valente, M., Cagnetti, C., Labate, A., Gambardella, A., Pisani, F., Casciato, S., ... Ferlazzo, E. (2022). Late epileptic seizures following cerebral venous thrombosis: a systematic review and meta-analysis. *Neurological Sciences*, 43(9), 5229-5236. <https://doi.org/10.1007/s10072-022-06148-y>
 177. Ge, Z., Ben, Q., Qian, J., Wang, Y., & Li, Y. (2011). Diabetes mellitus and risk of gastric cancer. *European Journal of Gastroenterology & Hepatology*, 23(12), 1127-1135. <https://doi.org/10.1097/MEG.0b013e32834b8d73>
 178. Gini, A., Bidoli, E., Zanier, L., Clagnan, E., Zanette, G., Gobbato, M., De Paoli, P., & Serraino, D. (2016). Cancer among patients with type 2 diabetes mellitus: A population-based cohort study in northeastern Italy. *Cancer Epidemiology*, 41, 80-87. <https://doi.org/10.1016/j.canep.2016.01.011>
 179. Giovannucci, E. (2001). Insulin, Insulin-Like Growth Factors and Colon Cancer: A Review of the Evidence. *The Journal of Nutrition*, 131(11), 3109S-3120S. <https://doi.org/10.1093/jn/131.11.3109S>
 180. Giovannucci, E. (2007). Metabolic syndrome, hyperinsulinemia, and colon cancer: a review. *American Journal of Clinical Nutrition*, 86(3), 836S--842S. %3CGo
 181. Giovannucci, E., Harlan, D. M., Archer, M. C., Bergenstal, R. M., Gapstur, S. M., Habel, L. A., Pollak, M., Regensteiner, J. G., & Yee, D. (2010).

- Diabetes and cancer: a consensus report. *Diabetes Care*, 33(7), 1674-1685.
<https://doi.org/10.2337/dc10-0666>
182. Giraldi, L., Vecchioni, A., Carioli, G., Bilotta, M., La Rosa, S., Imperatori, A., Volante, M., Brizzi, M. P., Inzani, F., Petrone, G., Schinzari, G., Bianchi, A., Margaritora, S., Alfieri, S., La Vecchia, C., Boccia, S., & Rindi, G. (2021). Risk factors for pancreas and lung neuroendocrine neoplasms: a case-control study. *Endocrine*, 71(1), 233-241.
<https://doi.org/10.1007/s12020-020-02464-5>
 183. Glaser, S. L., Lin, R. J., Stewart, S. L., Ambinder, R. F., Jarrett, R. F., Brousset, P., Pallesen, G., Gulley, M. L., Khan, G., O'Grady, J., Hummel, M., Preciado, M. V., Knecht, H., Chan, J. K. C., & Claviez, A. (1997). Epstein-Barr virus-associated Hodgkin's disease: Epidemiologic characteristics in international data. *International Journal of Cancer*, 70(4), 375-382. [https://doi.org/10.1002/\(SICI\)1097-0215\(19970207\)70:4<375::AID-IJC1>3.0.CO;2-T](https://doi.org/10.1002/(SICI)1097-0215(19970207)70:4<375::AID-IJC1>3.0.CO;2-T)
 184. Glass, T. A., Goodman, S. N., Hernán, M. A., & Samet, J. M. (2013). Causal Inference in Public Health. *Annual Review of Public Health*, 34(1), 61-75.
<https://doi.org/10.1146/annurev-publhealth-031811-124606>
 185. Gligorijević, N., Dobrijević, Z., Šunderić, M., Robajac, D., Četić, D., Penezić, A., Miljuš, G., & Nedić, O. (2022). The Insulin-like Growth Factor System and Colorectal Cancer. *Life*, 12(8), 1274.
<https://doi.org/10.3390/life12081274>
 186. Gong, Y., Wei, B., Yu, L., & Pan, W. (2015). Type 2 diabetes mellitus and risk of oral cancer and precancerous lesions: A meta-analysis of observational studies. *Oral Oncology*, 51(4), 332-340.
<https://doi.org/10.1016/j.oraloncology.2015.01.003>
 187. Gong, Y., Yang, Y.-S., Zhang, X.-M., Su, M., Wang, J., Han, J.-D., & Guo, M.-Z. (2012). ABO blood type, diabetes and risk of gastrointestinal cancer in northern China. *World Journal of Gastroenterology*, 18(6), 563-569.
<https://doi.org/10.3748/wjg.v18.i6.563>
 188. Goodwin, P. J. (2021). Diabetes and Cancer: Unraveling the Complexity. *JNCI: Journal of the National Cancer Institute*, 113(4), 347-348.
<https://doi.org/10.1093/jnci/djaa142>
 189. Goossens, M. E., Zeegers, M. P., Bazelier, M. T., De Bruin, M. L., Buntinx, F., & de Vries, F. (2015). Risk of bladder cancer in patients with diabetes: a retrospective cohort study. *BMJ Open*, 5(6), e007470.
<https://doi.org/10.1136/bmjopen-2014-007470>
 190. Gordis, L. (2014). *Epidemiology* (Elsevier Saunders, Ed.; 5th ed.).
 191. Goto, A., Yamaji, T., Sawada, N., Momozawa, Y., Kamatani, Y., Kubo, M., Shimazu, T., Inoue, M., Noda, M., Tsugane, S., & Iwasaki, M. (2020). Diabetes and cancer risk: A Mendelian randomization study. *International Journal of Cancer*, 146(3), 712-719. <https://doi.org/10.1002/ijc.32310>
 192. Gouveri, E., Papanas, N., & Maltezos, E. (2011). The female breast and diabetes. *The Breast*, 20(3), 205-211.
<https://doi.org/10.1016/j.breast.2011.02.019>
 193. Graff, R. E., Sanchez, A., Tobias, D. K., Rodríguez, D., Barrisford, G. W., Blute, M. L., Li, Y., Sun, Q., Preston, M. A., Wilson, K. M., & Cho, E. (2018). Type 2 Diabetes in Relation to the Risk of Renal Cell Carcinoma

- Among Men and Women in Two Large Prospective Cohort Studies. *Diabetes Care*, 41(7), 1432-1437. <https://doi.org/10.2337/dc17-2518>
194. GREENLAND, S. (1987). QUANTITATIVE METHODS IN THE REVIEW OF EPIDEMIOLOGIC LITERATURE¹. *Epidemiologic Reviews*, 9(1), 1-30. <https://doi.org/10.1093/oxfordjournals.epirev.a036298>
 195. Greenland, S., Pearl, J., & Robins, J. M. (2002). An overview of relations among causal modeling methods. *International Journal of Epidemiology*, 31(5), 1030-1037. <https://doi.org/10.1093/ije/31.5.1030>
 196. Griffith, M., Tajik, M., & Wing, S. (2007). Patterns of agricultural pesticide use in relation to socioeconomic characteristics of the population in the rural U.S. South. *International Journal of Health Services*, 37(2), 259-277. <https://doi.org/10.2190/R410-7263-4811-4415>
 197. Gu, J., Yan, S., Wang, B., Shen, F., Cao, H., Fan, J., & Wang, Y. (2016). Type 2 diabetes mellitus and risk of gallbladder cancer: a systematic review and meta-analysis of observational studies. *Diabetes/Metabolism Research and Reviews*, 32(1), 63-72. <https://doi.org/10.1002/dmrr.2671>
 198. Guo, J., Liu, C., Pan, J., & Yang, J. (2022). Relationship between diabetes and risk of gastric cancer: A systematic review and meta-analysis of cohort studies. *Diabetes Research and Clinical Practice*, 187(April), 109866. <https://doi.org/10.1016/j.diabres.2022.109866>
 199. Guo, J.-Z., Wu, Q.-J., Liu, F.-H., Gao, C., Gong, T.-T., & Li, G. (2022). Review of Mendelian Randomization Studies on Endometrial Cancer. *FRONTIERS IN ENDOCRINOLOGY*, 13. <https://doi.org/10.3389/fendo.2022.783150>
 200. Guraya, S. Y. (2015). Association of type 2 diabetes mellitus and the risk of colorectal cancer: A meta-analysis and systematic review. *World Journal of Gastroenterology*, 21(19), 6026-6031. <https://doi.org/10.3748/wjg.v21.i19.6026>
 201. Haber, N. A., Wieten, S. E., Rohrer, J. M., Arah, O. A., Tennant, P. W. G., Stuart, E. A., Murray, E. J., Pilleron, S., Lam, S. T., Riederer, E., Howcutt, S. J., Simmons, A. E., Leyrat, C., Schoenegger, P., Booman, A., Dufour, M.-S. K., O'Donoghue, A. L., Baglini, R., Do, S., ... Fox, M. P. (2022). Causal and Associational Language in Observational Health Research: A Systematic Evaluation. *American Journal of Epidemiology*, 191(12), 2084-2097. <https://doi.org/10.1093/aje/kwac137>
 202. Hadaegh, F., Asgari, S., Hashemi, P., Baghbani-Oskouei, A., Eskandari, F., Azizi, F., & Tohidi, M. (2020). All-cancer incidence in Tehranian adults: more than a decade of follow-up-results from the Tehran Lipid and Glucose Study. *Public Health*, 181, 189-195. <https://doi.org/10.1016/j.puhe.2020.01.003>
 203. Häggström, C., Van Hemelrijck, M., Zethelius, B. B., Robinson, D., Grundmark, B., Holmberg, L., Gudbjörnsdottir, S., Garmo, H., Stattin, P. P., Haggstrom, C., Van Hemelrijck, M., Zethelius, B. B., Grundmark, B., Holmberg, L., Garmo, H., & Gudbjörnsdottir, S. (2017). Prospective study of Type 2 diabetes mellitus, anti-diabetic drugs and risk of prostate cancer. *International Journal of Cancer*, 140(3), 611-617. <https://doi.org/10.1002/ijc.30480>
 204. Halder, S. K., & Melkani, G. C. (2025, March 1). The interplay of genetic predisposition, circadian misalignment, and metabolic regulation in

- obesity. *Current Obesity Reports*, 14(1), Article 21.
<https://doi.org/10.1007/s13679-025-00613-3>
205. Hammerton, G., & Munafò, M. R. (2021). Causal inference with observational data: the need for triangulation of evidence. *Psychological Medicine*, 51(4), 563-578. <https://doi.org/10.1017/S0033291720005127>
 206. Han, M. (2020). The Dose-Response Relationship between Alcohol Consumption and the Risk of Type 2 Diabetes among Asian Men: A Systematic Review and Meta-Analysis of Prospective Cohort Studies. *Journal of Diabetes Research*, 2020, 1032049.
<https://doi.org/10.1155/2020/1032049>
 207. Han, M. A., Leung, G., Storman, D., Xiao, Y., Srivastava, A., Talukdar, J. R., El Dib, R., Morassut, R. E., Zeraatkar, D., Johnston, B. C., & Guyatt, G. (2022). Causal language use in systematic reviews of observational studies is often inconsistent with intent: a systematic survey. *Journal of Clinical Epidemiology*, 148, 65-73. <https://doi.org/10.1016/j.jclinepi.2022.04.023>
 208. Haq, Z., Mirza, F. N., Abdi, P., Diaz, M. J., & Libby, T. J. (2024). Metformin Use and Risk of Non-Melanoma Skin Cancer: A Propensity-Matched Case-Control Study. *Journal of Drugs in Dermatology*, 23(12), 1089-1095.
<https://doi.org/10.36849/JDD.8249>
 209. Hara, N. (2012). Prostate Carcinogenesis with Diabetes and Androgen-Deprivation-Therapy-Related Diabetes: An Update. *Experimental Diabetes Research*, 2012, 1-8. <https://doi.org/10.1155/2012/801610>
 210. Hardefeldt, P. J., Edirimanne, S., & Eslick, G. D. (2012). Diabetes increases the risk of breast cancer: a meta-analysis. *Endocrine-Related Cancer*, 19(6), 793-803. <https://doi.org/10.1530/ERC-12-0242>
 211. Harden, F. A., Toms, L. M. L., Paepke, O., Ryan, J. J., & Müller, J. F. (2007). Evaluation of age, gender and regional concentration differences for dioxin-like chemicals in the Australian population. *Chemosphere*, 67(9), S318-S324. <https://doi.org/10.1016/j.chemosphere.2006.05.146>
 212. Harding, J. L., Shaw, J. E., Peeters, A., Cartensen, B., & Magliano, D. J. (2015). Cancer Risk Among People With Type 1 and Type 2 Diabetes: Disentangling True Associations, Detection Bias, and Reverse Causation. *Diabetes Care*, 38(2), 264-270. <https://doi.org/10.2337/dc14-1996>
 213. Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). Doing Meta-Analysis with R. Chapman and Hall/CRC.
<https://doi.org/10.1201/9781003107347>
 214. Hassan, M. M., Hwang, L.-Y., Hatten, C. J., Swaim, M., Li, D., Abbruzzese, J. L., Beasley, P., & Patt, Y. Z. (2002). Risk factors for hepatocellular carcinoma: Synergism of alcohol with viral hepatitis and diabetes mellitus. *Hepatology*, 36(5), 1206-1213. <https://doi.org/10.1053/jhep.2002.36780>
 215. Hatia, R. I., Eluri, M., Hawk, E. T., Shalaby, A., Karatas, E., Shalaby, A., Abdelhakeem, A., Abdel-Wahab, R., Chang, P., Rashid, A., Jalal, P. K., Amos, C. I., Han, Y., Armaghany, T., Shroff, R. T., Li, D., Javle, M., & Hassan, M. M. (2023). Independent of Primary Sclerosing Cholangitis and Cirrhosis, Early Adulthood Obesity Is Associated with Cholangiocarcinoma. *Cancer Epidemiology, Biomarkers & Prevention*, 32(10), 1338-1347.
<https://doi.org/10.1158/1055-9965.EPI-23-0388>
 216. He, S., Wang, J., Shen, X., Qian, X., An, Y., Gong, Q., Zhang, B., Chen, B., Zhang, L., Chen, X., Chen, Y., & Li, G. (2022). Cancer and its predictors in

- Chinese adults with newly diagnosed diabetes and impaired glucose tolerance (IGT): a 30-year follow-up of the Da Qing IGT and Diabetes Study. *British Journal of Cancer*, 127(1), 102-108.
<https://doi.org/10.1038/s41416-022-01758-x>
217. He, X., Shi, L., & Wu, J. (2018). Retrospective database analysis of cancer risk in patients with type 2 diabetes mellitus in China. *Current Medical Research and Opinion*, 34(6), 1089-1098.
<https://doi.org/10.1080/03007995.2017.1421527>
 218. Hemminki, K., Försti, A., Sundquist, K., & Li, X. (2016). Cancer of unknown primary is associated with diabetes. *European Journal of Cancer Prevention: The Official Journal of the European Cancer Prevention Organisation (ECP)*, 25(3), 246-251.
<https://doi.org/10.1097/CEJ.0000000000000165>
 219. Henry, S. (2014). Duration of diabetes and pancreatic cancer in a case-control study in the Midwest and the Iowa Women's Health Study (IWHs) cohort. *JOP*, 14(3), 243-249. <https://doi.org/10.6092/1590-8577/1317>
 220. Hense, H.-W., Kajüter, H., Wellmann, J., & Batzler, W. U. (2011). Cancer incidence in type 2 diabetes patients - first results from a feasibility study of the D2C cohort. *Diabetology & Metabolic Syndrome*, 3(1), 15.
<https://doi.org/10.1186/1758-5996-3-15>
 221. Hernan, M. A. (2002). Causal Knowledge as a Prerequisite for Confounding Evaluation: An Application to Birth Defects Epidemiology. *American Journal of Epidemiology*, 155(2), 176-184. <https://doi.org/10.1093/aje/155.2.176>
 222. Hernan, M. A. (2004). A definition of causal effect for epidemiological research. *Journal of Epidemiology & Community Health*, 58(4), 265-271.
<https://doi.org/10.1136/jech.2002.006361>
 223. Hernán, M. A. (2018). The C-Word: Scientific Euphemisms Do Not Improve Causal Inference From Observational Data. *American Journal of Public Health*, 108(5), 616-619. <https://doi.org/10.2105/AJPH.2018.304337>
 224. Hernán MA, R. J. (2020). Causal Inference: What If: Vol. [1] A. M. (Chapman & Hall/CRC, Ed.). <https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/>
 225. Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2019). Cochrane Handbook for Systematic Reviews of Interventions. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions*. Wiley. <https://doi.org/10.1002/9781119536604>
 226. Hippisley-Cox, J., & Coupland, C. (2015). Development and validation of risk prediction algorithms to estimate future risk of common cancers in men and women: prospective cohort study. *BMJ Open*, 5(3), e007825-e007825. <https://doi.org/10.1136/bmjopen-2015-007825>
 227. Hitchman, S. C., Fong, G. T., Zanna, M. P., Thrasher, J. F., Chung-Hall, J., & Siahpush, M. (2014). Socioeconomic status and smokers' number of smoking friends: Findings from the International Tobacco Control (ITC) Four Country Survey. *Drug and Alcohol Dependence*, 143(1), 158-166.
<https://doi.org/10.1016/j.drugalcdep.2014.07.019>
 228. Hjalgrim, H., Frisch, M., Ekbom, A., Kyvik, K. O., Melbye, M., & Green, A. (1997). Cancer and diabetes--a follow-up study of two population-based

- cohorts of diabetic patients. *Journal of Internal Medicine*, 241(6), 471-475.
<https://doi.org/10.1111/j.1365-2796.1997.tb00004.x>
229. Hoes, J., Boef, A. G. C., Knol, M. J., de Melker, H. E., Mollema, L., van der Klis, F. R. M., Rots, N. Y., & van Baarle, D. (2018). Socioeconomic Status Is Associated With Antibody Levels Against Vaccine Preventable Diseases in the Netherlands. *Frontiers in Public Health*, 6(July), 1-9.
<https://doi.org/10.3389/fpubh.2018.00209>
 230. Hong, T., Qin, N., Zhao, X., Wang, C., Jiang, Y., Ma, H., & Dai, J. (2021). Investigation of Causal Effect of Type 2 Diabetes Mellitus on Lung Cancer: A Mendelian Randomization Study. *Frontiers in Genetics*, 12.
<https://doi.org/10.3389/fgene.2021.673687>
 231. Hossain, F. M., Danos, D. M., Fu, Q., Wang, X., Scribner, R. A., Chu, S. T., Horswell, R. L., Price-Haywood, E. G., Collins-Burow, B. M., Wu, X.-C., Ochoa, A. C., & Miele, L. (2022). Association of Obesity and Diabetes With the Incidence of Breast Cancer in Louisiana. *American Journal of Preventive Medicine*, 63(1), S83-S92. <https://doi.org/10.1016/j.amepre.2022.02.017>
 232. Howards, P. P., Schisterman, E. F., Poole, C., Kaufman, J. S., & Weinberg, C. R. (2012). "Toward a Clearer Definition of Confounding" Revisited With Directed Acyclic Graphs. *American Journal of Epidemiology*, 176(6), 506-511. <https://doi.org/10.1093/aje/kws127>
 233. Hsieh, M. H., Sun, L.-M., Lin, C.-L., Hsieh, M.-J., Hsu, C.-Y., & Kao, C.-H. (2018). Development of a prediction model for pancreatic cancer in patients with type 2 diabetes using logistic regression and artificial neural network models. *Cancer Management and Research*, Volume 10, 6317-6324.
<https://doi.org/10.2147/CMAR.S180791>
 234. Hsieh, M.-C., Lee, T.-C., Cheng, S.-M., Tu, S.-T., Yen, M.-H., & Tseng, C.-H. (2012). The Influence of Type 2 Diabetes and Glucose-Lowering Therapies on Cancer Risk in the Taiwanese. *Experimental Diabetes Research*, 2012, 1-6. <https://doi.org/10.1155/2012/413782>
 235. Hsieh, M.-H., Sun, L.-M., Lin, C.-L., Hsieh, M.-J., Hsu, C., & Kao, C.-H. (2019). The Performance of Different Artificial Intelligence Models in Predicting Breast Cancer among Individuals Having Type 2 Diabetes Mellitus. *Cancers*, 11(11), 1751. <https://doi.org/10.3390/cancers11111751>
 236. Hsieh, M.-H., Sun, L.-M., Lin, C.-L., Hsieh, M.-J., Sun, K., Hsu, C.-Y., Chou, A.-K., & Kao, C.-H. (2018). Development of a Prediction Model for Colorectal Cancer among Patients with Type 2 Diabetes Mellitus Using a Deep Neural Network. *Journal of Clinical Medicine*, 7(9), 277.
<https://doi.org/10.3390/jcm7090277>
 237. Hsieh, S.-H., Lin, J.-D., Chiou, W.-K., & Wang, M.-H. (2014). Association of body weight with the risk for malignancies in hospitalized patients with or without diabetes mellitus in Taiwan. *Journal of Investigative Medicine*, 62(1), 37-42.
<https://doi.org/http://dx.doi.org/10.2310/JIM.0000000000000004>
 238. Hu, F. B., Manson, J. A. E., Liu, S., Hunter, D., Colditz, G. A., Michels, K. B., Speizer, F. E., & Giovannucci, E. (1999). Prospective study of adult onset diabetes mellitus (type 2) and risk of colorectal cancer in women. *Journal of the National Cancer Institute*, 91(6), 542-547.
<https://doi.org/10.1093/jnci/91.6.542>

239. Hu, Y., Wu, K., Zhang, X., Ma, Y., Yuan, C., Hu, F. B., Giovannucci, E., Wang, M., Tabung, F. K., Tobias, D., & Song, M. (2021). Incident Type 2 Diabetes Duration and Cancer Risk: A Prospective Study in Two US Cohorts. *Journal of the National Cancer Institute*, 113(4), 381-389. <https://doi.org/http://dx.doi.org/10.1093/jnci/djaa141>
240. Hua, S. (2019). Physiological and Pharmaceutical Considerations for Rectal Drug Formulations. *Frontiers in Pharmacology*, 10(October), 1-16. <https://doi.org/10.3389/fphar.2019.01196>
241. Huang, D. Q., Mathurin, P., Cortez-Pinto, H., & Loomba, R. (2023). Global epidemiology of alcohol-associated cirrhosis and HCC: trends, projections and risk factors. *Nature Reviews Gastroenterology & Hepatology*, 20(1), 37-49. <https://doi.org/10.1038/s41575-022-00688-6>
242. Huang, L., Feng, X., Yang, W., Li, X., Zhang, K., Feng, S., Wang, F., & Yang, X. (2022). Appraising the Effect of Potential Risk Factors on Thyroid Cancer: A Mendelian Randomization Study. *The Journal of Clinical Endocrinology & Metabolism*, 107(7), e2783-e2791. <https://doi.org/10.1210/clinem/dgac196>
243. Huang, W., Ren, H., Ben, Q., Cai, Q., Zhu, W., & Li, Z. (2012). Risk of esophageal cancer in diabetes mellitus: a meta-analysis of observational studies. *Cancer Causes & Control*, 23(2), 263-272. <https://doi.org/10.1007/s10552-011-9874-9>
244. Huang, Y., Cai, X., Qiu, M., Chen, P., Tang, H., Hu, Y., & Huang, Y. (2014). Prediabetes and the risk of cancer: a meta-analysis. *Diabetologia*, 57(11), 2261-2269. <https://doi.org/10.1007/s00125-014-3361-2>
245. Husain, K., Ansari, R. A., & Ferder, L. (2014). Alcohol-induced hypertension: Mechanism and prevention. *World Journal of Cardiology*, 6(5), 245. <https://doi.org/10.4330/wjc.v6.i5.245>
246. Hussain, S., & Chowdhury, T. A. (2019). The Impact of Comorbidities on the Pharmacological Management of Type 2 Diabetes Mellitus. *Drugs*, 79(3), 231-242. <https://doi.org/10.1007/s40265-019-1061-4>
247. Huxley, R., Ansary-Moghaddam, A., Berrington de González, A., Barzi, F., & Woodward, M. (2005). Type-II diabetes and pancreatic cancer: a meta-analysis of 36 studies. *British Journal of Cancer*, 92(11), 2076-2083. <https://doi.org/10.1038/sj.bjc.6602619>
248. Idilbi, N. M., Barchana, M., Milman, U., & Carel, R. S. (2013). Incidence of cancer among diabetic and non-diabetic adult Israeli Arabs. *The Israel Medical Association Journal: IMAJ*, 15(7), 342-347. <http://www.ncbi.nlm.nih.gov/pubmed/23943977>
249. Ikeda, M. L. R., Barcellos, N. T., Alencastro, P. R., Wolff, F. H., Moreira, L. B., Gus, M., Brandão, A. B. M., Fuchs, F. D., & Fuchs, S. C. (2016). Alcohol Drinking Pattern: A Comparison between HIV-Infected Patients and Individuals from the General Population. *PLOS ONE*, 11(6), e0158535. <https://doi.org/10.1371/journal.pone.0158535>
250. Inelmen, E. M., Sergi, G., & Enzi, G. (2007). When are indwelling urinary catheters appropriate in elderly patients? *Geriatrics*, 62(10), 18-22. <http://www.ncbi.nlm.nih.gov/pubmed/17922564>
251. Institute for Health Metrics and Evaluation (IHME). (2019). Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019 (GBD

- 2019) Results. IHME, University of Washington.
<https://vizhub.healthdata.org/gbd-results/>
252. International Diabetes Federation. (2021). IDF Diabetes Atlas. In International Diabetes Federation (Ed.), IDF official website (10th ed.).
 253. Isayama, H., Tazuma, S., Kokudo, N., Tanaka, A., Tsuyuguchi, T., Nakazawa, T., Notohara, K., Mizuno, S., Akamatsu, N., Serikawa, M., Naitoh, I., Hirooka, Y., Wakai, T., Itoi, T., Ebata, T., Okaniwa, S., Kamisawa, T., Kawashima, H., Kanno, A., ... Takikawa, H. (2018). Clinical guidelines for primary sclerosing cholangitis 2017. *Journal of Gastroenterology*, 53(9), 1006-1034. <https://doi.org/10.1007/s00535-018-1484-9>
 254. J., W., L., T., F., Z., X., C., Wu, J., Tang, L., Zheng, F., Chen, X., & Li, L. (2023). A review of the last decade: pancreatic cancer and type 2 diabetes. *Archives of Physiology and Biochemistry*, 1-9.
<https://doi.org/https://dx.doi.org/10.1080/13813455.2023.2252204> PT - Review
 255. Jackson, J. M., DeFor, T. A., Crain, A. L., Kerby, T. J., Strayer, L. S., Lewis, C. E., Whitlock, E. P., Williams, S. B., Vitolins, M. Z., Rodabough, R. J., Larson, J. C., Habermann, E. B., & Margolis, K. L. (2014). Validity of diabetes self-reports in the Women's Health Initiative. *Menopause*, 21(8), 861-868. <https://doi.org/10.1097/GME.0000000000000189>
 256. Jamal, M. M., Yoon, E. J., Vega, K. J., Hashemzadeh, M., & Chang, K. J. (2009a). Diabetes mellitus as a risk factor for gastrointestinal cancer among American veterans. *World Journal of Gastroenterology*, 15(42), 5274-5278. <https://doi.org/10.3748/wjg.15.5274>
 257. James Everhart, MD, MPH, David Wright, P. (1995). Diabetes Mellitus as a Risk Factor for Pancreatic Cancer. *Journal of the American Medical Association*, 273(20), 1605-1609.
 258. Janssen, J. A. M. J. L. (2021). Hyperinsulinemia and Its Pivotal Role in Aging, Obesity, Type 2 Diabetes, Cardiovascular Disease and Cancer. *International Journal of Molecular Sciences*, 22(15), 7797. <https://doi.org/10.3390/ijms22157797>
 259. Jian Gang, P., Mo, L., Lu, Y., Runqi, L., & Xing, Z. (2015). Diabetes mellitus and the risk of prostate cancer: an update and cumulative meta-analysis. *Endocrine Research*, 40(1), 54-61. <https://doi.org/10.3109/07435800.2014.934961>
 260. Jiang, Y., Ben, Q., Shen, H., Lu, W., Zhang, Y., & Zhu, J. (2011). Diabetes mellitus and incidence and mortality of colorectal cancer: a systematic review and meta-analysis of cohort studies. *European Journal of Epidemiology*, 26(11), 863-876. <https://doi.org/10.1007/s10654-011-9617-y>
 261. Jing, W., Jin, G., Zhou, X., Zhou, Y., Zhang, Y., Shao, C., Liu, R., & Hu, X. (2012). Diabetes mellitus and increased risk of cholangiocarcinoma. *European Journal of Cancer Prevention*, 21(1), 24-31. <https://doi.org/10.1097/CEJ.0b013e3283481d89>
 262. Jo, J. K., Song, H. K., Heo, Y., Kim, M. J., & Kim, Y. J. (2023). Risk analysis of metformin use in prostate cancer: a national population-based study. *The Aging Male*, 26(1). <https://doi.org/10.1080/13685538.2022.2156497>

263. Joh, H.-K., Willett, W. C., & Cho, E. (2011). Type 2 Diabetes and the Risk of Renal Cell Cancer in Women. *Diabetes Care*, 34(7), 1552-1556.
<https://doi.org/10.2337/dc11-0132>
264. Johansson, M., Carreras-Torres, R., Scelo, G., Purdue, M. P., Mariosa, D., Muller, D. C., Timpson, N. J., Haycock, P. C., Brown, K. M., Wang, Z., Ye, Y., Hofmann, J. N., Foll, M., Gaborieau, V., Machiela, M. J., Colli, L. M., Li, P., Garnier, J. G., Blanche, H., ... Others. (2019). The influence of obesity-related factors in the etiology of renal cell carcinoma-A mendelian randomization study. *Plos Medicine*, 16(1).
<https://doi.org/10.1371/journal.pmed.1002724>
265. Johns Hopkins Medicine. (2023). Cancer.
<https://www.hopkinsmedicine.org/health/cancer>
266. Johnson, J. A., Bowker, S. L., Richardson, K., & Marra, C. A. (2011). Time-varying incidence of cancer after the onset of type 2 diabetes: evidence of potential detection bias. *Diabetologia*, 54(9), 2263-2271.
<https://doi.org/10.1007/s00125-011-2242-1>
267. Johnson, J. A., Carstensen, B., Witte, D., Bowker, S. L., Lipscombe, L., Renehan, A. G., & Diabetes and Cancer Research Consortium. (2012). Diabetes and cancer (1): evaluating the temporal relationship between type 2 diabetes and cancer incidence. *Diabetologia*, 55(6), 1607-1618.
<https://doi.org/10.1007/s00125-012-2525-1>
268. Joost, H.-G. (2014). Diabetes and cancer: Epidemiology and potential mechanisms. *Diabetes and Vascular Disease Research*, 11(6), 390-394.
<https://doi.org/10.1177/1479164114550813>
269. Jordan, S., Lim, L., Vilainerun, D., Banks, E., Sripaiboonkij, N., Seubsman, S., Sleight, A., & Bain, C. (2009). Breast cancer in the Thai Cohort Study: An exploratory case-control analysis. *The Breast*, 18(5), 299-303.
<https://doi.org/10.1016/j.breast.2009.09.004>
270. Joung, K. H., Jeong, J.-W., & Ku, B. J. (2015). The Association between Type 2 Diabetes Mellitus and Women Cancer: The Epidemiological Evidences and Putative Mechanisms. *BioMed Research International*, 2015, 1-12. <https://doi.org/10.1155/2015/920618>
271. Kajüter, H., Geier, A. S., Wellmann, I., Krieg, V., Fricke, R., Heidinger, O., & Hense, H. W. (2014). Kohortenstudie zur Krebsinzidenz bei Patienten mit Diabetes mellitus Typ 2: Record Linkage von kryptografierten Daten einer externen Kohorte mit Daten des Epidemiologischen Krebsregisters Nordrhein-Westfalen. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 57(1), 52-59. <https://doi.org/10.1007/s00103-013-1880-5>
272. Kallergi, A., Chambre, C., Duchemann, B., Fysekidis, M., & Bihan, H. (2015). Diabetes Mellitus and Colorectal Cancer Risk. *Current Colorectal Cancer Reports*, 11(4), 209-216. <https://doi.org/10.1007/s11888-015-0274-5>
273. Kang, J., Jeong, S.-M., Shin, D. W., Cho, M., Cho, J. H., & Kim Dong Wook; ORCID: <http://orcid.org/0000-0001-8128-8920>, J. A. O.-S. (2021). The Associations of Aspirin, Statins, and Metformin With Lung Cancer Risk and Related Mortality: A Time-Dependent Analysis of Population-Based Nationally Representative Data. *Journal of Thoracic Oncology*, 16(1), 76-88.
<https://doi.org/http://dx.doi.org/10.1016/j.jtho.2020.08.021>

274. Kao, C.-H., Sun, L.-M., Chen, P.-C., Lin, M.-C., Liang, J.-A., Muo, C.-H., Chang, S.-N., & Sung, F.-C. (2013). A population-based cohort study in Taiwan--use of insulin sensitizers can decrease cancer risk in diabetic patients? *Annals of Oncology*, 24(2), 523-530. <https://doi.org/10.1093/annonc/mds472>
275. Kao, W. H. L. (2001). Alcohol Consumption and the Risk of Type 2 Diabetes Mellitus: Atherosclerosis Risk in Communities Study. *American Journal of Epidemiology*, 154(8), 748-757. <https://doi.org/10.1093/aje/154.8.748>
276. Kasmari, A. J., Welch, A., Liu, G., Leslie, D., McGarrity, T., & Riley, T. (2017). Independent of Cirrhosis, Hepatocellular Carcinoma Risk Is Increased with Diabetes and Metabolic Syndrome. *American Journal of Medicine*, 130(6), 746.e1-746.e7. <https://doi.org/10.1016/j.amjmed.2016.12.029>
277. Kasper, J. S., & Giovannucci, E. (2006). A meta-analysis of diabetes mellitus and the risk of prostate cancer. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, 15(11), 2056-2062. <https://doi.org/10.1158/1055-9965.EPI-06-0410>
278. Kasper, J. S., Liu, Y., & Giovannucci, E. (2009). Diabetes mellitus and risk of prostate cancer in the health professionals follow-up study. *International Journal of Cancer*, 124(6), 1398-1403. <https://doi.org/10.1002/ijc.24044>
279. Katzke, V. A., Kaaks, R., Kneisel, J., & Kühn, T. (2020). Diabetes und Krebs. *Der Diabetologe*, 16(7), 675-684. <https://doi.org/10.1007/s11428-020-00677-w>
280. Kezios, K. L. (2021). Is the Way Forward to Step Back? Documenting the Frequency With Which Study Goals Are Misaligned With Study Methods and Interpretations in the Epidemiologic Literature. *Epidemiologic Reviews*, 43(1), 4-18. <https://doi.org/10.1093/epirev/mxab008>
281. Khachatryan, L., Scharpf, R., & Kagan, S. (2011). Influence of Diabetes Mellitus Type 2 and Prolonged Estrogen Exposure on Risk of Breast Cancer Among Women in Armenia. *Health Care for Women International*, 32(11), 953-971. <https://doi.org/10.1080/07399332.2011.569041>
282. Khan, A. E., Gallo, V., Linseisen, J., Kaaks, R., Rohrmann, S., Raaschou-Nielsen, O., Tjonneland, A., Johnsen, H. E., Overvad, K., Bergmann, M. M., Boeing, H., Benetou, V., Psaltopoulou, T., Trichopoulou, A., Masala, G., Mattiello, A., Grioni, S., Tumino, R., Vermeulen, R. C. H., ... Riboli, E. (2008). Diabetes and the risk of non-Hodgkin's lymphoma and multiple myeloma in the European Prospective Investigation into Cancer and Nutrition. *Haematologica*, 93(6), 842-850. <https://doi.org/10.3324/haematol.12297>
283. Khoshsirat S, Peyvandi AA, Oroei M, Peyvandi H, R. A. (2020). The Estimation of Thyroid Cancer Risk based on Type 2 Diabetes Mellitus and ABO Blood Group. *Turkish Journal of Oncology*, 35(1), 64-69. <https://doi.org/10.5505/tjo.2020.2066>
284. Kilvert, A., & Fox, C. (2020). Hyperinsulinaemia and cancer risk: cause and effect? *Practical Diabetes*, 37(6), 223. <https://doi.org/10.1002/pdi.2310>
285. Kim, H. J., Lee, S., Chun, K. H., Jeon, J. Y., Han, S. J., Kim, D. J., Kim, Y. S., Woo, J.-T., Nam, M.-S., Baik, S. H., Ahn, K. J., & Lee, K. W. (2018).

- Metformin reduces the risk of cancer in patients with type 2 diabetes. *Medicine*, 97(8), e0036. <https://doi.org/10.1097/MD.00000000000010036>
286. Kim, J., Hyun, H.-J., Choi, E.-A., Yoo, J. W., Lee, S., Jeong, N., Shen, J. J., You, H.-S., Kim, Y., & Kang, H.-T. (2020). Diabetes, Metformin, and Lung Cancer: Retrospective Study of the Korean NHIS-HEALS Database. *Clinical Lung Cancer*, 21(6), e551-e559. <https://doi.org/10.1016/j.clcc.2020.04.001>
 287. Kim, M.-J., Lim, N.-K., Choi, S.-J., & Park, H.-Y. (2015). Hypertension is an independent risk factor for type 2 diabetes: the Korean genome and epidemiology study. *Hypertension Research: Official Journal of the Japanese Society of Hypertension*, 38(11), 783–789. <https://doi.org/10.1038/hr.2015.72>
 288. Kim, S. K., Jang, J. Y., Kim, D. L., Rhyu, Y. A., Lee, S. E., Ko, S. H., Han, K., & Song, K. H. (2020). Site-specific cancer risk in patients with type 2 diabetes: A nationwide population-based cohort study in Korea. *Korean Journal of Internal Medicine*, 35(3), 641-651. <https://doi.org/10.3904/kjim.2017.402>
 289. Kincius, M., Patasius, A., Linkeviciute-Ulinskiene, D., Zabuliene, L., & Smailyte, G. (2020). Reduced risk of prostate cancer in a cohort of Lithuanian diabetes mellitus patients. *The Aging Male*, 23(5), 1333-1338. <https://doi.org/10.1080/13685538.2020.1766013>
 290. Knapen, L., Dittrich, S., Vries, F., Starup-Linde, J., Vestergaard, P., Henry, R., Stolk, L., Neef, C., & Bazelier, M. (2013). Use of Biguanides and the Risk of Colorectal Cancer: A Register-Based Cohort Study. *Current Drug Safety*, 8(5), 349-356. <https://doi.org/10.2174/15680266113136660068>
 291. Koh, W. P., Wang, R., Jin, A., Yu, M. C., & Yuan, J. M. (2013). Diabetes mellitus and risk of hepatocellular carcinoma: Findings from the Singapore Chinese Health Study. *British Journal of Cancer*, 108(5), 1182-1188. <https://doi.org/10.1038/bjc.2013.25>
 292. Kramer, H. U., Schottker, B., Raum, E., Brenner, H., Krämer, H. U., Schöttker, B., Raum, E., & Brenner, H. (2012). Type 2 diabetes mellitus and colorectal cancer: Meta-analysis on sex-specific differences. *European Journal of Cancer*, 48(9), 1269-1282. <https://doi.org/10.1016/j.ejca.2011.07.010>
 293. Krasnodebski, P., Mrozikiewicz-Rakowska, B., Nehring, P., Stelmasiak, P., Nyckowski, J., Stankiewicz, R., & Karnafel, W. (2010). Risk factors of colon neoplasms in patients with type 2 diabetes. *Diabetologia Doswiadczalna i Kliniczna*, 10(1), 34-40. <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed11&NEWS=N&AN=359348979>
 294. Kusunoki, Y., & Hayashi, T. (2008). Long-lasting alterations of the immune system by ionizing radiation exposure: Implications for disease development among atomic bomb survivors. *International Journal of Radiation Biology*, 84(1), 1-14. <https://doi.org/10.1080/09553000701616106>
 295. La Torre, G., Sferrazza, A., Gualano, M. R., de Waure, C., Clemente, G., De Rose, A. M., Nicolotti, N., Nuzzo, G., Siliquini, R., Boccia, A., & Ricciardi, W. (2014). Investigating the Synergistic Interaction of Diabetes, Tobacco Smoking, Alcohol Consumption, and Hypercholesterolemia on the

- Risk of Pancreatic Cancer: A Case-Control Study in Italy. *BioMed Research International*, 2014, 1-6. <https://doi.org/10.1155/2014/481019>
296. La Vecchia, C., Negri, E., Decarli, A., & Franceschi, S. (1997a). Diabetes mellitus and colorectal cancer risk. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, 6(12), 1007–1010. <https://doi.org/10.1007/s11888-015-0274-5>
 297. La Vecchia, C., Negri, E., Decarli, A., & Franceschi, S. (1997b). Diabetes mellitus and the risk of primary liver cancer. *International Journal of Cancer*, 73(2), 204-207. [https://doi.org/10.1002/\(sici\)1097-0215\(19971009\)73:2<204:aid-ijc7>3.0.co;2-#](https://doi.org/10.1002/(sici)1097-0215(19971009)73:2<204:aid-ijc7>3.0.co;2-#)
 298. La Vecchia, C., Negri, E., Franceschi, S., D'Avanzo, B., & Boyle, P. (1994). A case-control study of diabetes mellitus and cancer risk. *British Journal of Cancer*, 70(5), 950-953. <https://doi.org/10.1038/bjc.1994.427>
 299. Lai, H.-C., Chang, S.-N., Lin, C.-C., Chen, C.-C., Chou, J.-W., Peng, C.-Y., Lai, S.-W., Sung, F.-C., & Li, Y.-F. (2013). Does diabetes mellitus with or without gallstones increase the risk of gallbladder cancer? Results from a population-based cohort study. *Journal of Gastroenterology*, 48(7), 856-865. <https://doi.org/10.1007/s00535-012-0683-z>
 300. Lai, M.-S., Hsieh, M.-S., Chiu, Y.-H., & Chen, T. H.-H. (2006). Type 2 diabetes and hepatocellular carcinoma: A cohort study in high prevalence area of hepatitis virus infection. *Hepatology*, 43(6), 1295-1302. <https://doi.org/10.1002/hep.21208>
 301. Lann, D., & LeRoith, D. (2008). The Role of Endocrine Insulin-Like Growth Factor-I and Insulin in Breast Cancer. *Journal of Mammary Gland Biology and Neoplasia*, 13(4), 371-379. <https://doi.org/10.1007/s10911-008-9100-x>
 302. Larsson, S. C., & Burgess, S. (2022). Appraising the causal role of smoking in multiple diseases: A systematic review and meta-analysis of Mendelian randomization studies. *EBioMedicine*, 82, 104154. <https://doi.org/10.1016/j.ebiom.2022.104154>
 303. Larsson, S. C., Giovannucci, E., & Wolk, A. (2005). Diabetes and Colorectal Cancer Incidence in the Cohort of Swedish Men. *Diabetes Care*, 28(7), 1805-1807. <https://doi.org/10.2337/diacare.28.7.1805>
 304. Larsson, S. C., Mantzoros, C. S., & Wolk, A. (2007). Diabetes mellitus and risk of breast cancer: A meta-analysis. *International Journal of Cancer*, 121(4), 856-862. <https://doi.org/10.1002/ijc.22717>
 305. Larsson, S. C., Orsini, N., Brismar, K., & Wolk, A. (2006). Diabetes mellitus and risk of bladder cancer: a meta-analysis. *Diabetologia*, 49(12), 2819-2823. <https://doi.org/10.1007/s00125-006-0468-0>
 306. Larsson, S. C., Orsini, N., & Wolk, A. (2005). Diabetes Mellitus and Risk of Colorectal Cancer: A Meta-Analysis. *JNCI: Journal of the National Cancer Institute*, 97(22), 1679-1687. <https://doi.org/10.1093/jnci/dji375>
 307. Larsson, S. C., & Wolk, A. (2011). Diabetes mellitus and incidence of kidney cancer: a meta-analysis of cohort studies. *Diabetologia*, 54(5), 1013-1018. <https://doi.org/10.1007/s00125-011-2051-6>
 308. Lawler, T., Walts, Z. L., Steinwandel, M., Lipworth, L., Murff, H. J., Zheng, W., & Warren Andersen, S. (2023). Type 2 Diabetes and Colorectal Cancer Risk. *JAMA Network Open*, 6(11), e2343333. <https://doi.org/10.1001/jamanetworkopen.2023.43333>

309. Lawrence, Y. R., Morag, O., Benderly, M., Boyko, V., Novikov, I., Dicker, A. P., Goldbourt, U., Behar, S., Barchana, M., & Wolf, I. (2013). Association between metabolic syndrome, diabetes mellitus and prostate cancer risk. *Prostate Cancer and Prostatic Diseases*, 16(2), 181-186. <https://doi.org/10.1038/pcan.2012.54>
310. Leahy, S., Nolan, A., O'Connell, J., & Kenny, R. A. (2014). Obesity in an Ageing Society: Implications for health, physical function and health service utilisation (Issue July). <https://doi.org/10.38018/TildaRe.2014-01>
311. Lederer, D. J., Bell, S. C., Branson, R. D., Chalmers, J. D., Marshall, R., Maslove, D. M., Ost, D. E., Punjabi, N. M., Schatz, M., Smyth, A. R., Stewart, P. W., Suissa, S., Adjei, A. A., Akdis, C. A., Azoulay, É., Bakker, J., Ballas, Z. K., Bardin, P. G., Barreiro, E., ... Vincent, J.-L. (2019). Control of Confounding and Reporting of Results in Causal Inference Studies. Guidance for Authors from Editors of Respiratory, Sleep, and Critical Care Journals. *Annals of the American Thoracic Society*, 16(1), 22-28. <https://doi.org/10.1513/AnnalsATS.201808-564PS>
312. Lee, B. S., Park, E.-C., Park, S. W., Nam, C. M., & Roh, J. (2015). Hepatitis B virus infection, diabetes mellitus, and their synergism for cholangiocarcinoma development: a case-control study in Korea. *World Journal of Gastroenterology*, 21(2), 502-510. <https://doi.org/10.3748/wjg.v21.i2.502>
313. Lee, D. Y., Yu, J. H., Park, S., Han, K., Kim, N. H., Yoo, H. J., Choi, K. M., Baik, S. H., Kim, N. H., & Seo, J. A. (2018). The influence of diabetes and antidiabetic medications on the risk of pancreatic cancer: a nationwide population-based study in Korea. *Scientific Reports*, 8(1), 9719. <https://doi.org/10.1038/s41598-018-27965-2>
314. Lee, J. Y., Jeon, I., Lee, J. M., Yoon, J. M., & Park, S. M. (2013). Diabetes mellitus as an independent risk factor for lung cancer: A meta-analysis of observational studies. *European Journal of Cancer*, 49(10), 2411-2423. <https://doi.org/10.1016/j.ejca.2013.02.025>
315. Lee, J.-Y., Jeon, I., Kim, J. W., Song, Y.-S., Yoon, J.-M., & Park, S. M. (2013). Diabetes Mellitus and Ovarian Cancer Risk: A Systematic Review and Meta-Analysis of Observational Studies. *International Journal of Gynecologic Cancer*, 23(3), 402-412. <https://doi.org/10.1097/IGC.0b013e31828189b2>
316. Lee, M. Y., Lin, K. der, Hsiao, P. J., & Shin, S. J. (2012). The association of diabetes mellitus with liver, colon, lung, and prostate cancer is independent of hypertension, hyperlipidemia, and gout in Taiwanese patients. *Metabolism: Clinical and Experimental*, 61(2), 242-249. <https://doi.org/10.1016/j.metabol.2011.06.020>
317. Lee, Y. H. (2020). Overview of Mendelian Randomization Analysis. *Journal of Rheumatic Diseases*, 27(4), 241-246. <https://doi.org/10.4078/jrd.2020.27.4.241>
318. Lega, I. C., & Lipscombe, L. L. (2020). Review: Diabetes, Obesity, and Cancer-Pathophysiology and Clinical Implications. *Endocrine Reviews*, 41(1), bnz014. <https://doi.org/http://dx.doi.org/10.1210/endrev/bnz014>
319. Lega, I. C., Wilton, A. S., Austin, P. C., Fischer, H. D., Johnson, J. A., & Lipscombe, L. L. (2016). The temporal relationship between diabetes and

- cancer: A population-based study. *Cancer*, 122(17), 2731–2738.
<https://doi.org/10.1002/cncr.30095>
320. Li, D. (2012). Diabetes and pancreatic cancer. *Molecular Carcinogenesis*, 51(1), 64–74. <https://doi.org/10.1002/mc.20771>
 321. Li, D., Yeung, S. J., Hassan, M. M., Konopleva, M., & Abbruzzese, J. L. (2009). Antidiabetic Therapies Affect Risk of Pancreatic Cancer. *Gastroenterology*, 137(2), 482–488.
<https://doi.org/10.1053/j.gastro.2009.04.013>
 322. Li, H., Chen, Y., Hu, L., Yang, W., Gao, Z., Liu, M., Tao, H., & Li, J. (2023). Will metformin use lead to a decreased risk of thyroid cancer? A systematic review and meta-analyses. *European Journal of Medical Research*, 28(1), 392. <https://doi.org/10.1186/s40001-023-01287-0>
 323. Li, H., & Qian, J. (2017). Association of diabetes mellitus with thyroid cancer risk. *Medicine*, 96(47), e8230.
<https://doi.org/10.1097/MD.00000000000008230>
 324. Li L, Huang HL, Zhang XY, Zhang XY, L. Z. (2011). [Association of metabolic syndrome and colorectal cancer] (pp. 948–951). *Zhonghua Wei Chang Wai Ke Za Zhi*. <https://doi.org/10.3760/cma.j.issn.1671-0274.2011.12.010>
 325. Li, X., Li, F., Wang, J., van Giessen, A., & Feenstra, T. L. (2023). Prediction of complications in health economic models of type 2 diabetes: a review of methods used. *Acta Diabetologica*, 60(7), 861–879.
<https://doi.org/10.1007/s00592-023-02045-8>
 326. Li, Y., Bian, X., Wei, S., He, M., & Yang, Y. (2019). The relationship between pancreatic cancer and type 2 diabetes: cause and consequence. *Cancer Management and Research*, Volume 11, 8257–8268.
<https://doi.org/10.2147/CMAR.S211972>
 327. Li, Y. R., Ro, V., & Tchou, J. C. (2018). Obesity, Metabolic Syndrome, and Breast Cancer: From Prevention to Intervention. *Current Surgery Reports*, 6(3). <https://doi.org/10.1007/s40137-018-0204-y>
 328. Li, Z., Chen, H., Fritz, C. D. L., Zheng, X., Zong, X., Nickel, K. B., Tipping, A., Nguyen, L. H., Chan, A. T., Giovannucci, E. L., Colditz, G. A., Olsen, M. A., Campbell, P. T., Davidson, N. O., Fields, R. C., & Cao, Y. (2022). Type 2 Diabetes and Risk of Early-Onset Colorectal Cancer. *Gastro Hep Advances*, 1(2), 186–193. <https://doi.org/10.1016/j.gastha.2021.10.009>
 329. Liang, Y., Teng, J. A., Qin, J., Wang, R. S., & Tang, C. Y. (2008). Risk analysis on type II diabetes mellitus patients complicating endometrioid carcinoma. *Chinese Journal of Cancer Prevention and Treatment*, 15(19), 1448–1450.
 330. Liao, C., Zhang, D., Mungo, C., Andrew Tompkins, D., & Zeidan, A. M. (2014). Is diabetes mellitus associated with increased incidence and disease-specific mortality in endometrial cancer? A systematic review and meta-analysis of cohort studies. *Gynecologic Oncology*, 135(1), 163–171.
<https://doi.org/10.1016/j.ygyno.2014.07.095>
 331. Liao, S., Li, J., Wei, W., Wang, L., Zhang, Y., Li, J., Wang, C., & Sun, S. (2011). Association between diabetes mellitus and breast cancer risk: a meta-analysis of the literature. *Asian Pacific Journal of Cancer Prevention: APJCP*, 12(4), 1061–1065. <http://www.ncbi.nlm.nih.gov/pubmed/21790252>
 332. Liaw, Y.-P., Ko, P.-C., Jan, S.-R., Huang, J.-Y., Nfor, O. N., Lung, C.-C., Chiang, Y.-C., Yeh, L.-T., Chou, M.-C., Tsai, H.-D., & Hsiao, Y.-H. (2015).

- Implications of Type1/2 Diabetes Mellitus in Breast Cancer Development: A General Female Population-based Cohort Study. *Journal of Cancer*, 6(8), 734-739. <https://doi.org/10.7150/jca.12197>
333. Limburg, P. J., Anderson, K. E., Johnson, T. W., Jacobs, D. R., Lazovich, D., Hong, C. P., Nicodemus, K. K., & Folsom, A. R. (2005). Diabetes mellitus and subsite-specific colorectal cancer risks in the Iowa Women's Health Study. *Cancer Epidemiology Biomarkers & Prevention*, 14(1), 133-137. %3CGo
 334. Limburg, P. J., Vierkant, R. A., Fredericksen, Z. S., Leibson, C. L., Rizza, R. A., Gupta, A. K., Ahlquist, D. A., Melton, L. J., Sellers, T. A., & Cerhan, J. R. (2006). Clinically Confirmed Type 2 Diabetes Mellitus and Colorectal Cancer Risk: A Population-Based, Retrospective Cohort Study. *The American Journal of Gastroenterology*, 101(8), 1872-1879. <https://doi.org/10.1111/j.1572-0241.2006.00725.x>
 335. Lin, C.-C., Chiang, J.-H., Li, C.-I., Hsieh, T.-F., Liu, C.-S., Lin, W.-Y., Lee, Y.-D., & Li, T.-C. (2014). Independent and joint effect of type 2 diabetes and gastric and hepatobiliary diseases on risk of pancreatic cancer risk: 10-year follow-up of population-based cohort. *British Journal of Cancer*, 111(11), 2180-2186. <https://doi.org/10.1038/bjc.2014.525>
 336. Lin, C.-C., Chiang, J.-H., Li, C.-I., Liu, C.-S., Lin, W.-Y., Hsieh, T.-F., & Li, T.-C. (2014). Cancer risks among patients with type 2 diabetes: a 10-year follow-up study of a nationwide population-based cohort in Taiwan. *BMC Cancer*, 14(1), 381. <https://doi.org/10.1186/1471-2407-14-381>
 337. Lin, E., Garmo, H., Van Hemelrijck, M., Adolfsson, J., Stattin, P., Zethelius, B., & Crawley, D. (2020). Association of type 2 diabetes mellitus and antidiabetic medication with risk of prostate cancer: a population-based case-control study. *BMC Cancer*, 20(1), 551. <https://doi.org/10.1186/s12885-020-07036-4>
 338. Lin, L., Benard, V. B., Greek, A., Hawkins, N. A., Roland, K. B., & Saraiya, M. (2015). Racial and ethnic differences in human papillomavirus positivity and risk factors among low-income women in Federally Qualified Health Centers in the United States. *Preventive Medicine*, 81(1), 258-261. <https://doi.org/10.1016/j.ypmed.2015.08.027>
 339. Lin, L., Chu, H., Murad, M. H., Hong, C., Qu, Z., Cole, S. R., & Chen, Y. (2018). Empirical Comparison of Publication Bias Tests in Meta-Analysis. *Journal of General Internal Medicine*, 33(8), 1260-1267. <https://doi.org/10.1007/s11606-018-4425-7>
 340. Lind, M., Fahlén, M., Eliasson, B., & Odén, A. (2012). The relationship between the exposure time of insulin glargine and risk of breast and prostate cancer: An observational study of the time-dependent effects of antidiabetic treatments in patients with diabetes. *Primary Care Diabetes*, 6(1), 53-59. <https://doi.org/10.1016/j.pcd.2011.10.004>
 341. Ling, S., Brown, K., Miksza, J. K., Howells, L. M., Morrison, A., Issa, E., Yates, T., Khunti, K., Davies, M. J., & Zaccardi, F. (2021). Risk of cancer incidence and mortality associated with diabetes: A systematic review with trend analysis of 203 cohorts. *Nutrition, Metabolism and Cardiovascular Diseases*, 31(1), 14-22. <https://doi.org/10.1016/j.numecd.2020.09.023>
 342. Ling, S., Brown, K., Miksza, J. K., Howells, L., Morrison, A., Issa, E., Yates, T., Khunti, K., Davies, M. J., & Zaccardi, F. (2020). Association of Type 2

- Diabetes With Cancer: A Meta-analysis With Bias Analysis for Unmeasured Confounding in 151 Cohorts Comprising 32 Million People. *Diabetes Care*, 43(9), 2313-2322. <https://doi.org/10.2337/dc20-0204>
343. Lipscombe, L. L., Goodwin, P. J., Zinman, B., McLaughlin, J. R., & Hux, J. E. (2006). Diabetes mellitus and breast cancer: a retrospective population-based cohort study. *Breast Cancer Research and Treatment*, 98(3), 349-356. <https://doi.org/10.1007/s10549-006-9172-5>
 344. Lipsky, A. M., & Greenland, S. (2022). Causal Directed Acyclic Graphs. *JAMA*, 327(11), 1083. <https://doi.org/10.1001/jama.2022.1816>
 345. Lipworth, L., Zucchetto, A., Bosetti, C., Franceschi, S., Talamini, R., Serraino, D., McLaughlin, J. K., La Vecchia, C., & Negri, E. (2011). Diabetes mellitus, other medical conditions and pancreatic cancer: a case-control study. *Diabetes/Metabolism Research & Reviews*, 27(3), 255-261. <https://doi.org/10.1002/dmrr.1162>
 346. Liu, H.-S., Chen, C.-D., Lee, C.-C., Chen, Y.-C., & Cheng, W.-F. (2022). Age Specific Risks of Uterine Cancer in Type 2 Diabetes and Associated Comorbidities in Taiwan. *Cancers*, 14(19), 4912. <https://doi.org/10.3390/cancers14194912>
 347. Liu, P., Ioannidis, J. P. A., Ross, J. S., Dhruva, S. S., Luxkaranayagam, A. T., Vasiliou, V., & Wallach, J. D. (2019). Age-treatment subgroup analyses in Cochrane intervention reviews: a meta-epidemiological study. *BMC Medicine*, 17(1), 188. <https://doi.org/10.1186/s12916-019-1420-8>
 348. Liu, X., Hemminki, K., Försti, A., Sundquist, K., Sundquist, J., & Ji, J. (2015). Cancer risk in patients with type 2 diabetes mellitus and their relatives. *International Journal of Cancer*, 137(4), 903-910. <https://doi.org/10.1002/ijc.29440>
 349. Liu, Y., Jin, P.-P., Sun, X.-C., & Hu, T.-T. (2018). Thiazolidinediones and risk of colorectal cancer in patients with diabetes mellitus: A meta-analysis. *Saudi Journal of Gastroenterology*, 24(2), 75. https://doi.org/10.4103/sjg.SJG_295_17
 350. Liu, Y., & Zhang, W. (2017). Association between body mass index and endometriosis risk: A meta-analysis. *Oncotarget*, 8(29), 46928-46936. <https://doi.org/10.18632/oncotarget.14916>
 351. Liu, Y.-C., Nguyen, P.-A., Humayun, A., Chien, S.-C., Yang, H.-C., Asdary, R. N., Syed-Abdul, S., Hsu, M.-H., Moldovan, M., Yen, Y., Li, Y.-C. (Jack), Jian, W.-S., & Iqbal, U. (2019). Does long-term use of antidiabetic drugs changes cancer risk? *Medicine*, 98(40), e17461. <https://doi.org/10.1097/MD.00000000000017461>
 352. Lo, S. F., Chang, S. N., Muo, C. H., Chen, S. Y., Liao, F. Y., Dee, S. W., Chen, P. C., & Sung, F. C. (2013). Modest increase in risk of specific types of cancer types in type 2 diabetes mellitus patients. *International Journal of Cancer*, 132(1), 182-188. <https://doi.org/10.1002/ijc.27597>
 353. Long, X.-J., Lin, S., Sun, Y.-N., & Zheng, Z.-F. (2012). Diabetes Mellitus and Prostate Cancer Risk in Asian Countries: a Meta-analysis. *Asian Pacific Journal of Cancer Prevention*, 13(8), 4097-4100. <https://doi.org/10.7314/APJCP.2012.13.8.4097>
 354. López-Suárez, A. (2019). Burden of cancer attributable to obesity, type 2 diabetes and associated risk factors. *Metabolism*, 92, 136-146. <https://doi.org/10.1016/j.metabol.2018.10.013>

355. Lu, Y., Gentiluomo, M., Lorenzo-Bermejo, J., Morelli, L., Obazee, O., Campa, D., & Canzian, F. (2020). Mendelian randomisation study of the effects of known and putative risk factors on pancreatic cancer. *Journal of Medical Genetics*, 57(12), 820-828. <https://doi.org/10.1136/jmedgenet-2019-106200>
356. Lu, Y., Hajjar, A., Cryns, V. L., Trentham-Dietz, A., Gangnon, R. E., Heckman-Stoddard, B. M., & Alagoz, O. (2023). Breast cancer risk for women with diabetes and the impact of metformin: A meta-analysis. *CANCER MEDICINE*, 12(10), 11703-11718. <https://doi.org/10.1002/cam4.5545>
357. Lu, Y., Rodríguez, L. A. G., Malgerud, L., González-Pérez, A., Martín-Pérez, M., Lagergren, J., Bexelius, T. S., Rodríguez, L. A. G., Malgerud, L., González-Pérez, A., Martín-Pérez, M., & García Rodríguez, L. A. (2015). New-onset type 2 diabetes, elevated HbA1c, anti-diabetic medications, and risk of pancreatic cancer. *British Journal of Cancer*, 113(11), 1607-1614. <https://doi.org/10.1038/bjc.2015.353>
358. Luchini, C., Stubbs, B., Solmi, M., & Veronese, N. (2017). Assessing the quality of studies in meta-analyses: Advantages and limitations of the Newcastle Ottawa Scale. *World Journal of Meta-Analysis*, 5(4), 80. <https://doi.org/10.13105/wjma.v5.i4.80>
359. Luijken, K., Groenwold, R. H. H., van Smeden, M., Strohmaier, S., & Heinze, G. (2024). A comparison of full model specification and backward elimination of potential confounders when estimating marginal and conditional causal effects on binary outcomes from observational data. *Biometrical Journal*, 66(1), 1-14. <https://doi.org/10.1002/bimj.202100237>
360. Luo, S., Li, J.-Y., Zhao, L.-N., Yu, T., Zhong, W., Xia, Z.-S., Shan, T.-D., Ouyang, H., Yang, H.-S., & Chen, Q.-K. (2016). Diabetes mellitus increases the risk of colorectal neoplasia: An updated meta-analysis. *Clinics and Research in Hepatology and Gastroenterology*, 40(1), 110-123. <https://doi.org/10.1016/j.clinre.2015.05.021>
361. Luo, S., Schooling, C. M., Wong, I. C. K., & Au Yeung, S. L. (2020). Evaluating the impact of AMPK activation, a target of metformin, on risk of cardiovascular diseases and cancer in the UK Biobank: a Mendelian randomisation study. *Diabetologia*, 63(11), 2349-2358. <https://doi.org/10.1007/s00125-020-05243-z>
362. Luo, W., Cao, Y., Liao, C., & Gao, F. (2012). Diabetes mellitus and the incidence and mortality of colorectal cancer: a meta-analysis of 24 cohort studies. *Colorectal Disease: The Official Journal of the Association of Coloproctology of Great Britain and Ireland*, 14(11), 1307-1312. <https://doi.org/10.1111/j.1463-1318.2012.02875.x>
363. Lymphoma Action. (2023). Types of lymphoma. <https://lymphoma-action.org.uk/index.php/types-lymphoma>
364. Lymphoma Australia. (2023). Types of Lymphoma. <https://www.lymphoma.org.au/lymphoma-types/>
365. Ma, Y., Yang, W., Song, M., Smith-Warner, S. A., Yang, J., Li, Y., Ma, W., Hu, Y., Ogino, S., Hu, F. B., Wen, D., Chan, A. T., Giovannucci, E. L., & Zhang, X. (2018). Type 2 diabetes and risk of colorectal cancer in two large U.S. prospective cohorts. *British Journal of Cancer*, 119(11), 1436-1442. <https://doi.org/10.1038/s41416-018-0314-4>

366. Macías-Gómez, N. M., Peralta-Leal, V., Meza-Espinoza, J. P., Gutiérrez-Angulo, M., Durán-González, J., Ramírez-González, J. M., Gaspar-Del Toro, A., Norberto-Rodríguez, A., & Leal-Ugarte, E. (2015). Polymorphisms of the XRCC1 gene and breast cancer risk in the Mexican population. *Familial Cancer*, 14(3), 349-354. <https://doi.org/10.1007/s10689-015-9787-y>
367. Maddatu, J., Anderson-Baucum, E., & Evans-Molina, C. (2017). Smoking and the risk of type 2 diabetes. *Translational Research*, 184(3), 101-107. <https://doi.org/10.1016/j.trsl.2017.02.004>
368. Magliano, D. J., Davis, W. A., Shaw, J. E., Bruce, D. G., & Davis, T. M. E. (2012). Incidence and predictors of all-cause and site-specific cancer in type 2 diabetes: the Fremantle Diabetes Study. *European Journal of Endocrinology*, 167(4), 589-599. <https://doi.org/10.1530/EJE-12-0053>
369. Magruder, J. T., Elahi, D., & Andersen, D. K. (2011). Diabetes and Pancreatic Cancer. *Pancreas*, 40(3), 339-351. <https://doi.org/10.1097/MPA.0b013e318209e05d>
370. Maisonneuve, P., Lowenfels, A. B., Bueno-de-Mesquita, H. B., Ghadirian, P., Baghurst, P. A., Zatonski, W. A., Miller, A. B., Duell, E. J., Boffetta, P., & Boyle, P. (2010). Past Medical History and Pancreatic Cancer Risk: Results From a Multicenter Case-Control Study. *Annals of Epidemiology*, 20(2), 92-98. <https://doi.org/10.1016/j.annepidem.2009.11.010>
371. Makhoul, I., Yacoub, A., & Siegel, E. (2016). Type 2 diabetes mellitus is associated with increased risk of pancreatic cancer: A veteran administration registry study. *SAGE Open Medicine*, 4, 205031211668225. <https://doi.org/10.1177/2050312116682257>
372. Malaguarnera, R., Vella, V., Nicolosi, M. L., & Belfiore, A. (2017). Insulin Resistance: Any Role in the Changing Epidemiology of Thyroid Cancer? *Frontiers in Endocrinology*, 8. <https://doi.org/10.3389/fendo.2017.00314>
373. Malhotra, J., Malvezzi, M., Negri, E., La Vecchia, C., & Boffetta, P. (2016). Risk factors for lung cancer worldwide. *The European Respiratory Journal*, 48(3), 889-902. <https://doi.org/10.1183/13993003.00359-2016>
374. Man Chu, T. S. (2018). The Relationship between Type 2 Diabetes Mellitus and Risk of Prostate Cancer: Literature Review and Critical Appraisal. *Journal of Diabetes & Metabolism*, 09(03). <https://doi.org/10.4172/2155-6156.1000787>
375. Mantovani, A., & Targher, G. (2017). Type 2 diabetes mellitus and risk of hepatocellular carcinoma: spotlight on nonalcoholic fatty liver disease. *Annals of Translational Medicine*, 5(13), 270-270. <https://doi.org/10.21037/atm.2017.04.41>
376. Marengo, A., Rosso, C., & Bugianesi, E. (2016). Liver Cancer: Connections with Obesity, Fatty Liver, and Cirrhosis. *Annual Review of Medicine*, 67(1), 103-117. <https://doi.org/10.1146/annurev-med-090514-013832>
377. Mariusdottir, E., Ingimarsson, J. P., Jonsson, E., Einarsson, G. v, Aspelund, T., Gudnason, V., & Gudbjartsson, T. (2016). Occupation as a risk factor for renal cell cancer: a nationwide, prospective epidemiological study. *Scandinavian Journal of Urology*, 50(3), 181-185. <https://doi.org/10.3109/21681805.2016.1151460>
378. Marti, J. L., Jain, K. S., & Morris, L. G. T. (2015). Increased Risk of Second Primary Malignancy in Pediatric and Young Adult Patients Treated with

- Radioactive Iodine for Differentiated Thyroid Cancer. *Thyroid*, 25(6), 681-687. <https://doi.org/10.1089/thy.2015.0067>
379. Martins, L. C. G., Lopes, M. V. de O., Diniz, C. M., & Guedes, N. G. (2021). The factors related to a sedentary lifestyle: A meta-analysis review. *Journal of Advanced Nursing*, 77(3), 1188-1205. <https://doi.org/10.1111/jan.14669>
 380. Maskarinec, G., Brown, S. M., Lee, J., Bogumil, D., Walsh, C., Haiman, C. A., Setiawan, V. W., Shvetsov, Y. B., & Le Marchand, L. (2023). Association of Obesity and Type 2 Diabetes with Non-Hodgkin Lymphoma: The Multiethnic Cohort. *Cancer Epidemiology, Biomarkers & Prevention*, 32(10), 1348-1355. <https://doi.org/10.1158/1055-9965.EPI-23-0565>
 381. Maskarinec, G., Fontaine, A., Torfadottir, J. E., Lipscombe, L. L., Lega, I. C., Figueroa, J., & Wild, S. (2018). The Relation of Type 2 Diabetes and Breast Cancer Incidence in Asian, Hispanic and African American Populations—A Review. *Canadian Journal of Diabetes*, 42(1), 100-105. <https://doi.org/10.1016/j.jcjd.2017.02.005>
 382. Maskarinec, G., Haraldsdóttir, Á., Einarsdóttir, K., Aspelund, T., Tryggvadóttir, L., Harris, T. B., Gudnason, V., & Torfadóttir, J. E. (2019). Type 2 diabetes and obesity in midlife and breast cancer risk in the Reykjavik cohort. *Cancer Causes & Control*, 30(10), 1057-1065. <https://doi.org/10.1007/s10552-019-01213-y>
 383. Maskarinec, G., Jacobs, S., Park, S. Y., Haiman, C. A., Setiawan, V. W., Wilkens, L. R., & Marchand, L. Le. (2017). Type II diabetes, obesity, and breast cancer risk: The multiethnic cohort. *Cancer Epidemiology Biomarkers and Prevention*, 26(6), 854-861. <https://doi.org/10.1158/1055-9965.EPI-16-0789>
 384. Mathers, C. D. (2020). History of global burden of disease assessment at the World Health Organization. *Archives of Public Health*, 78(1), 77. <https://doi.org/10.1186/s13690-020-00458-3>
 385. Mathur, M. B., & VanderWeele, T. J. (2020). Sensitivity Analysis for Unmeasured Confounding in Meta-Analyses. *Journal of the American Statistical Association*, 115(529), 163-172. <https://doi.org/10.1080/01621459.2018.1529598>
 386. McCullough, M. L., Chantaprasopsuk, S., Islami, F., Rees-Punia, E., Um, C. Y., Wang, Y., Leach, C. R., Sullivan, K. R., & Patel, A. V. (2022). Association of Socioeconomic and Geographic Factors with Diet Quality in US Adults. *JAMA Network Open*, 5(6), E2216406. <https://doi.org/10.1001/jamanetworkopen.2022.16406>
 387. McFarland, M. S., & Cripps, R. (2010). Diabetes Mellitus and Increased Risk of Cancer: Focus on Metformin and the Insulin Analogs. *Pharmacotherapy*, 30(11), 1159-1178. <https://doi.org/10.1592/phco.30.11.1159>
 388. Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R., & Cheraghi-Sohi, S. (2014). PICO, PICOS and SPIDER: A comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Services Research*, 14(1). <https://doi.org/10.1186/s12913-014-0579-0>
 389. Mhaskar, R., Ricardo, I., Azliyati, A., Laxminarayan, R., Amol, B., Santosh, W., & Boo, K. (2013). Assessment of risk factors of helicobacter pylori

- infection and peptic ulcer disease. *Journal of Global Infectious Diseases*, 5(2), 60. <https://doi.org/10.4103/0974-777X.112288>
390. Miao, Z.-F., Xu, H., Xu, Y.-Y., Wang, Z.-N., Zhao, T.-T., Song, Y.-X., & Xu, H.-M. (2017). Diabetes mellitus and the risk of gastric cancer: a meta-analysis of cohort studies. *Oncotarget*, 8(27), 44881-44892. <https://doi.org/10.18632/oncotarget.16487>
 391. Michels, K. B., Solomon, C. G., Hu, F. B., Rosner, B. A., Hankinson, S. E., Colditz, G. A., & Manson, J. E. (2003). Type 2 Diabetes and Subsequent Incidence of Breast Cancer in the Nurses' Health Study. *Diabetes Care*, 26(6), 1752-1758. <https://doi.org/10.2337/diacare.26.6.1752>
 392. Miele, L., Bosetti, C., Turati, F., Rapaccini, G., Gasbarrini, A., la Vecchia, C., Boccia, S., & Grieco, A. (2015). Diabetes and Insulin Therapy, but Not Metformin, Are Related to Hepatocellular Cancer Risk. *Gastroenterology Research and Practice*, 2015. <https://doi.org/10.1155/2015/570356>
 393. Mikaeel, R. R., Symonds, E. L., Kimber, J., Smith, E., Horsnell, M., Uylaki, W., Rico, G. T., Hewett, P. J., Yong, J., Tonkin, D., Jesudason, D., Poplawski, N. K., Ruszkiewicz, A. R., Drew, P. A., Hardingham, J. E., Wong, S., Frank, O., Tomita, Y., Patel, D., ... Young, J. P. (2021). Young-onset colorectal cancer is associated with a personal history of type 2 diabetes. *Asia-Pacific Journal of Clinical Oncology*, 17(1), 131-138. <https://doi.org/10.1111/ajco.13428>
 394. Mitri, J., Castillo, J., & Pittas, A. G. (2008). Diabetes and Risk of Non-Hodgkin's Lymphoma. *Diabetes Care*, 31(12), 2391-2397. <https://doi.org/10.2337/dc08-1034>
 395. Mizuno, S., Nakai, Y., Ishigaki, K., Saito, K., Oyama, H., Hamada, T., Suzuki, Y., Inokuma, A., Kanai, S., Noguchi, K., Sato, T., Hakuta, R., Saito, T., Takahara, N., Kogure, H., Isayama, H., & Koike, K. (2020). Screening Strategy of Pancreatic Cancer in Patients with Diabetes Mellitus. *Diagnostics*, 10(8), 572. <https://doi.org/10.3390/diagnostics10080572>
 396. Mkuu, R. S., Hall, J. M., Galochkina, Z., Cho, H. D., Staras, S. A. S., Lee, J.-H., Guo, Y., Chakrabarti, C., Barrow, S. B., Ortega, S., Avery, D. M., Higginbotham, J., Lockhart, J., & Shenkman, E. A. (2023). Does the Intersectionality of Race/Ethnicity and Type 2 Diabetes Increase the Odds of a Cervical Cancer Diagnosis? A Nested Case-Control Study of a Florida Statewide Multisite EHR Database. *Healthcare*, 11(13), 1863. <https://doi.org/10.3390/healthcare11131863>
 397. Modesti, P. A., Reboldi, G., Cappuccio, F. P., Agyemang, C., Remuzzi, G., Rapi, S., Perruolo, E., Parati, G., Agostoni, P., Barros, H., Basu, S., Benetos, A., Ceriello, A., Prato, S. del, Kalyesubula, R., Kilama, M. O., O'Brien, E., Perlini, S., Picano, E., ... Okechukwu, O. S. (2016). Panethnic differences in blood pressure in Europe: A systematic review and meta-analysis. *PLoS ONE*, 11(1). <https://doi.org/10.1371/journal.pone.0147601>
 398. Mohd Hanafiah, K., Groeger, J., Flaxman, A. D., & Wiersma, S. T. (2013). Global epidemiology of hepatitis C virus infection: New estimates of age-specific antibody to HCV seroprevalence. *Hepatology*, 57(4), 1333-1342. <https://doi.org/10.1002/hep.26141>
 399. Molina, P. E., Happel, K. I., Zhang, P., Kolls, J. K., & Nelson, S. (2010). Focus on: Alcohol and the immune system. *Alcohol Research and Health*, 33(1-2), 97-108.

400. Molina-Montes, E., Coscia, C., Gómez-Rubio, P., Fernández, A., Boenink, R., Rava, M., Márquez, M., Molero, X., Löhr, M., Sharp, L., Michalski, C. W., Farré, A., Perea, J., O'Rorke, M., Greenhalf, W., Iglesias, M., Tardón, A., Gress, T. M., Barberá, V. M., ... Malats, N. (2021). Deciphering the complex interplay between pancreatic cancer, diabetes mellitus subtypes and obesity/BMI through causal inference and mediation analyses. *Gut*, 70(2), 319-329. <https://doi.org/10.1136/gutjnl-2019-319990>
401. Morimoto, M., Matsuo, T., & Mori, N. (2021). Management of porcelain gallbladder, its risk factors, and complications: A review. *Diagnostics*, 11(6). <https://doi.org/10.3390/diagnostics11061073>
402. Muller, W. A. (1998). Diabetes mellitus--long time survival. *Journal of Insurance Medicine (New York, N.Y.)*, 30(1), 17-27. <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed6&NEWS=N&AN=128300396>
403. Murad, M. H., Chu, H., Lin, L., & Wang, Z. (2018). The effect of publication bias magnitude and direction on the certainty in evidence. *BMJ Evidence-Based Medicine*, 23(3), 84-86. <https://doi.org/10.1136/bmjebm-2018-110891>
404. Murphy, N., Song, M., Papadimitriou, N., Carreras-Torres, R., Langenberg, C., Martin, R. M., Tsilidis, K. K., Barroso, I., Chen, J., Frayling, T. M., Bull, C. J., Vincent, E. E., Cotterchio, M., Gruber, S. B., Pai, R. K., Newcomb, P. A., Perez-Cornago, A., van Duijnhoven, F. J. B., Van Guelpen, B., ... Gunter, M. J. (2022). Associations Between Glycemic Traits and Colorectal Cancer: A Mendelian Randomization Analysis. *JNCI: Journal of the National Cancer Institute*, 114(5), 740-752. <https://doi.org/10.1093/jnci/djac011>
405. Nagel, J., & Göke, B. (2006). Diabetes mellitus Typ 2 - ein Risikofaktor für kolorektale Karzinome. *Zeitschrift Für Gastroenterologie*, 44(11), 1153-1165. <https://doi.org/10.1055/s-2006-927132>
406. Najafi, F., Rajati, F., Sarokhani, D., Bavandpour, M., & Moradinazar, M. (2023). The relationship between metformin consumption and cancer risk: An updated umbrella review of systematic reviews and meta-analyses. *International Journal of Preventive Medicine*, 14(1), 90. https://doi.org/10.4103/ijpvm.ijpvm_62_21
407. National Cancer Institute. (2023). Cancer types. <https://www.cancer.gov/types>
408. National Health Service. (2023). Health A to Z. <https://www.nhs.uk/conditions/>
409. National Institute for Health and Care Excellence (NICE). (2023). Health topics A to Z. <https://cks.nice.org.uk/topics/>
410. National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). (2022). Risk Factors for Type 2 Diabetes. <https://www.niddk.nih.gov/health-information/diabetes/overview/risk-factors-type-2-diabetes>
411. National Organization for Rare Disorders. (2023). NORD, National Organization for Rare Disorders. <https://rarediseases.org/rare-diseases/>
412. Ndjaboue, R., Ngueta, G., Rochefort-Brihay, C., Delorme, S., Guay, D., Ivers, N., Shah, B. R., Straus, S. E., Yu, C., Comeau, S., Farhat, I., Racine, C., Drescher, O., & Witteman, H. O. (2022). Prediction models of diabetes

- complications: a scoping review. *Journal of Epidemiology and Community Health*, 76(10), 896-904. <https://doi.org/10.1136/jech-2021-217793>
413. Nead, K. T., Sharp, S. J., Thompson, D. J., Painter, J. N., Savage, D. B., Semple, R. K., Barker, A., Perry, J. R. B., Attia, J., Dunning, A. M., Easton, D. F., Holliday, E., Lotta, L. A., O'Mara, T., McEvoy, M., Pharoah, P. D. P., Scott, R. J., Spurdle, A. B., Langenberg, C., ... Scott, R. A. (2015). Evidence of a Causal Association between Insulinemia and Endometrial Cancer: A Mendelian Randomization Analysis. *Journal of the National Cancer Institute*, 107(9), 1-7. <https://doi.org/10.1093/jnci/djv178>
 414. New York State Department of Health. (2023). AIDS Institute. <https://www.health.ny.gov/diseases/aids/general/resources/>
 415. Newby, P. K., Muller, D., Hallfrisch, J., Qiao, N., Andres, R., & Tucker, K. L. (2003). Dietary patterns and changes in body mass index and waist circumference in adults. *American Journal of Clinical Nutrition*, 77(6), 1417-1425. <https://doi.org/10.1093/ajcn/77.6.1417>
 416. Newton, C. C., Gapstur, S. M., Campbell, P. T., & Jacobs, E. J. (2013). Type 2 diabetes mellitus, insulin-use and risk of bladder cancer in a large cohort study. *International Journal of Cancer*, 132(9), 2186-2191. <https://doi.org/10.1002/ijc.27878>
 417. Nicodemus, K. K., Sweeney, C., & Folsom, A. R. (2004). Evaluation of dietary, medical and lifestyle risk factors for incident kidney cancer in postmenopausal women. *International Journal of Cancer*, 108(1), 115-121. <https://doi.org/10.1002/ijc.11532>
 418. Nicolucci, A. (2010). Epidemiological aspects of neoplasms in diabetes. *Acta Diabetologica*, 47(2), 87-95. <https://doi.org/10.1007/s00592-010-0187-3>
 419. Nodari, Y., Gentiluomo, M., Mohelnikova-Duchonova, B., Kreivenaite, E., Milanetto, A. C., Skieceviciene, J., Landi, S., Lawlor, R. T., Petrone, M. C., Arcidiacono, P. G., Lovecek, M., Gazouli, M., Bijlsma, M. F., Morelli, L., Kiudelis, V., Tacelli, M., Zanette, D. L., Soucek, P., Uzunoglu, F., ... Campa, D. (2023). Genetic and non-genetic risk factors for early-onset pancreatic cancer. *Digestive and Liver Disease*, 55(10), 1417-1425. <https://doi.org/10.1016/j.dld.2023.02.023>
 420. Noto, H., Osame, K., Sasazuki, T., & Noda, M. (2010). Substantially increased risk of cancer in patients with diabetes mellitus. *Journal of Diabetes and Its Complications*, 24(5), 345-353. <https://doi.org/10.1016/j.jdiacomp.2010.06.004>
 421. Noto, H., Tsujimoto, T., Sasazuki, T., & Noda, M. (2011). Significantly increased risk of cancer in patients with diabetes mellitus: a systematic review and meta-analysis. *Endocrine Practice: Official Journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists*, 17(4), 616-628. <https://doi.org/10.4158/EP10357.RA>
 422. Nouredin, M., & Rinella, M. E. (2015). Nonalcoholic Fatty Liver Disease, Diabetes, Obesity, and Hepatocellular Carcinoma. *Clinics in Liver Disease*, 19(2), 361-379. <https://doi.org/10.1016/j.cld.2015.01.012>
 423. Nussbaumer-Streit, B., Klerings, I., Wagner, G., Titscher, V., Greimel, J., Affengruber, L., Gartlehner, G., & Cochrane Rapid Reviews Methods Group. (2023). Rapid reviews methods series: Guidance on team considerations, study selection, data extraction and risk of bias assessment. *BMJ Evidence-*

- Based Medicine, 28(6), 418-423. <https://doi.org/10.1136/bmjebm-2022-112107>
424. Oberman, B., Khaku, A., Camacho, F., & Goldenberg, D. (2015). Relationship between obesity, diabetes and the risk of thyroid cancer. *American Journal of Otolaryngology - Head and Neck Medicine and Surgery*, 36(4), 535-541. <https://doi.org/10.1016/j.amjoto.2015.02.015>
 425. OECD, O. for E. C. and D. (2013). National Accounts at a Glance 2013. OECD. https://doi.org/10.1787/na_glance-2013-en
 426. Ogunleye, A. A., Ogston, S. A., Morris, A. D., & Evans, J. M. M. (2009). A cohort study of the risk of cancer associated with type 2 diabetes. *British Journal of Cancer*, 101(7), 1199-1201. <https://doi.org/10.1038/sj.bjc.6605240>
 427. Ogunmola, J. O., Oladosu, O., & Olamoyegun, A. M. (2014). Relationship between socioeconomic status and HIV infection in a rural tertiary health center. *HIV/AIDS - Research and Palliative Care*, 6, 61. <https://doi.org/10.2147/HIV.S59061>
 428. Ogurtsova, K., Guariguata, L., Barengo, N. C., Ruiz, P. L.-D., Sacre, J. W., Karuranga, S., Sun, H., Boyko, E. J., & Magliano, D. J. (2022). IDF diabetes Atlas: Global estimates of undiagnosed diabetes in adults for 2021. *Diabetes Research and Clinical Practice*, 183, 109118. <https://doi.org/10.1016/j.diabres.2021.109118>
 429. Ohkuma, T., Peters, S. A. E., & Woodward, M. (2018). Sex differences in the association between diabetes and cancer: a systematic review and meta-analysis of 121 cohorts including 20 million individuals and one million events. *Diabetologia*, 61(10), 2140-2154. <https://doi.org/10.1007/s00125-018-4664-5>
 430. Olarte Parra, C., Bertizzolo, L., Schroter, S., Dechartres, A., & Goetghebeur, E. (2021). Consistency of causal claims in observational studies: a review of papers published in a general medical journal. *BMJ Open*, 11(5), e043339. <https://doi.org/10.1136/bmjopen-2020-043339>
 431. Ong, K. L., Stafford, L. K., McLaughlin, S. A., Boyko, E. J., Vollset, S. E., Smith, A. E., Dalton, B. E., Duprey, J., Cruz, J. A., Hagins, H., Lindstedt, P. A., Aali, A., Abate, Y. H., Abate, M. D., Abbasian, M., Abbasi-Kangevari, Z., Abbasi-Kangevari, M., Abd ElHafeez, S., Abd-Rabu, R., ... Vos, T. (2023). Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. *The Lancet*, 402(10397), 203-234. [https://doi.org/10.1016/S0140-6736\(23\)01301-6](https://doi.org/10.1016/S0140-6736(23)01301-6)
 432. Onitilo, A. A., Berg, R. L., Engel, J. M., Glurich, I., Stankowski, R. V., Williams, G., & Doi, S. A. (2013). Increased Risk of Colon Cancer in Men in the Pre-Diabetes Phase. *PLoS ONE*, 8(8), e70426. <https://doi.org/10.1371/journal.pone.0070426>
 433. Onitilo, A. A., Engel, J. M., Glurich, I., Stankowski, R. v., Williams, G. M., & Doi, S. A. (2012). Diabetes and cancer I: risk, survival, and implications for screening. *Cancer Causes & Control*, 23(6), 967-981. <https://doi.org/10.1007/s10552-012-9972-3>
 434. Onitilo, A. A., Stankowski, R. V., Berg, R. L., Engel, J. M., Glurich, I., Williams, G. M., & Doi, S. A. R. (2014). Breast cancer incidence before and after diagnosis of type 2 diabetes mellitus in women: Increased risk in the

- prediabetes phase. *European Journal of Cancer Prevention*, 23(2), 76-83.
<https://doi.org/10.1097/CEJ.0b013e32836162aa>
435. Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews*, 5(1), 210. <https://doi.org/10.1186/s13643-016-0384-4>
 436. Overbeek, J. A., van der Heijden, A. A. W. A., Nijpels, G., Kuiper, J. G., Labots, M., Herings, R. M. C., Haug, U., Overbeek, J. A., Kuiper, J. G., Labots, M., Haug, U., Herings Jetty A.; ORCID: <http://orcid.org/0000-0003-0935-5193>, R. M. C. A. O.-O., van der Heijden, A. A. W. A., Labots, M., Haug, U., Herings, R. M. C., & Nijpels, G. (2019). Sex- and site-specific differences in colorectal cancer risk among people with type 2 diabetes. *International Journal of Colorectal Disease*, 34(2), 269-276.
<https://doi.org/http://dx.doi.org/10.1016/j.jval.2018.09.738>
 437. Pal, K., Horsfall, L., Sharma, M., Nazareth, I., & Petersen, I. (2021). Time trends in the incidence of clinically diagnosed type 2 diabetes and pre-diabetes in the UK 2009-2018: a retrospective cohort study. *BMJ Open Diabetes Research & Care*, 9(1). <https://doi.org/10.1136/bmjdr-2020-001989>
 438. Palmer, J. R., Castro-Webb, N., Bertrand, K., Bethea, T. N., & Denis, G. V. (2017). Type II Diabetes and Incidence of Estrogen Receptor Negative Breast Cancer in African American Women. *Cancer Research*, 77(22), 6462-6469. <https://doi.org/10.1158/0008-5472.CAN-17-1903>
 439. Pan, X. F., He, M., Yu, C., Lv, J., Guo, Y., Bian, Z., Yang, L., Chen, Y., Wu, T., Chen, Z., Pan, A., & Li, L. (2018). Type 2 Diabetes and Risk of Incident Cancer in China: A Prospective Study among 0.5 Million Chinese Adults. *American Journal of Epidemiology*, 187(7), 1380-1391.
<https://doi.org/10.1093/aje/kwx376>
 440. Pan, Y., Lee, C.-Y., Lee, L.-M., Wen, Y.-C., Huang, J.-Y., Yang, S.-F., & Hsiao, C.-H. (2020). Incidence of Bladder Cancer in Type 2 Diabetes Mellitus Patients: A Population-Based Cohort Study. *Medicina*, 56(9), 441. <https://doi.org/10.3390/medicina56090441>
 441. Pang, Y., Kartsonaki, C., Guo, Y., Bragg, F., Yang, L., Bian, Z., Chen, Y., Iona, A., Millwood, I. Y., Lv, J., Yu, C., Chen, J., Li, L., Holmes, M. v., & Chen, Z. (2017). Diabetes, plasma glucose and incidence of pancreatic cancer: A prospective study of 0.5 million Chinese adults and a meta-analysis of 22 cohort studies. *International Journal of Cancer*, 140(8), 1781-1788. <https://doi.org/10.1002/ijc.30599>
 442. Parazzini, F., La Vecchia, C., Negri, E., Luca Riboldi, G., Surace, M., Benzi, G., Maina, A., & Chiaffarino, F. (1999). Diabetes and endometrial cancer: An Italian case-control study. *International Journal of Cancer*, 81(4), 539-542. [https://doi.org/10.1002/\(SICI\)1097-0215\(19990517\)81:4<539: AID-IJC6>3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1097-0215(19990517)81:4<539: AID-IJC6>3.0.CO;2-Q)
 443. Park, H. J., Joh, H.-K., Choi, S., & Park, S. M. (2019). Type 2 diabetes mellitus does not increase the risk of lung cancer among never-smokers: a nationwide cohort study. *Translational Lung Cancer Research*, 8(6), 1073-1077. <https://doi.org/10.21037/tlcr.2019.11.01>
 444. Pasqual, E., O'Brien, K., Rinaldi, S., Sandler, D. P., & Kitahara, C. M. (2023). Obesity, obesity-related metabolic conditions, and risk of thyroid cancer in women: results from a prospective cohort study (Sister Study).

- The Lancet Regional Health - Americas, 23, 100537.
<https://doi.org/10.1016/j.lana.2023.100537>
445. Pathak, M., Dwivedi, S. N., Thakur, B., & Vishnubhatla, S. (2020). Methods of estimating the pooled effect size under meta-analysis: A comparative appraisal. *Clinical Epidemiology and Global Health*, 8(1), 105-112.
<https://doi.org/10.1016/j.cegh.2019.05.004>
 446. Patrascu Mihaita MD PhD, Diclescu Mircea MD PhD, N., Petrut MD, Costache Raluca MD PhD, B. S. M. D. P., Macadon Bogdan MD, Balaban Vasile MD PhD, P. A., MD, Calin Roxana MD, Raduta Ioana MD PhD, P. D. M. D., & Ionita Radu Florentina MD PhD, J. M. M. D. P. (2018). The combined phenotype of diabetes melitus and colorectal cancer. *Journal of Gastrointestinal and Liver Diseases*, 27(2).
<http://www.jgld.ro/wp/y2018/supplement2.pdf>
 447. Pearson-Stuttard, J., Papadimitriou, N., Markozannes, G., Cividini, S., Kakourou, A., Gill, D., Rizos, E. C., Monori, G., Ward, H. A., Kyrgiou, M., Gunter, M. J., & Tsilidis, K. K. (2021). Type 2 Diabetes and Cancer: An Umbrella Review of Observational and Mendelian Randomization Studies. *Cancer Epidemiology, Biomarkers & Prevention*, 30(6), 1218-1228.
<https://doi.org/10.1158/1055-9965.EPI-20-1245>
 448. Pearson-Stuttard, J., Zhou, B., Kontis, V., Bentham, J., Gunter, M. J., & Ezzati, M. (2018). Worldwide burden of cancer attributable to diabetes and high body-mass index: a comparative risk assessment. *The Lancet Diabetes & Endocrinology*, 6(6), e6-e15. [https://doi.org/10.1016/S2213-8587\(18\)30150-5](https://doi.org/10.1016/S2213-8587(18)30150-5)
 449. Peeters, P. J. H. L., Bazelier, M. T., Leufkens, H. G. M., de Vries, F., & De Bruin, M. L. (2015). The Risk of Colorectal Cancer in Patients With Type 2 Diabetes: Associations With Treatment Stage and Obesity. *Diabetes Care*, 38(3), 495-502. <https://doi.org/10.2337/dc14-1175>
 450. Pereira, S. P., Oldfield, L., Ney, A., Hart, P. A., Keane, M. G., Pandol, S. J., Li, D., Greenhalf, W., Jeon, C. Y., Koay, E. J., Almario, C. v, Halloran, C., Lennon, A. M., & Costello, E. (2020). Early detection of pancreatic cancer. *The Lancet. Gastroenterology & Hepatology*, 5(7), 698-710.
[https://doi.org/10.1016/S2468-1253\(19\)30416-9](https://doi.org/10.1016/S2468-1253(19)30416-9)
 451. Peters, N., Dichgans, M., Surendran, S., Argilés, J. M., López-Soriano, F. J., Busquets, S., Dittmann, K., Rodemann, H. P., Sindrilariu, A., Sunderkötter, C., Watanabe, H., Roden, D. M., Feuerstein, G., Ruffolo, R., Knöll, R., Sen-Chowdhry, S., Ward, D., McKenna, W. J., Mogensen, J., ... Sauve, R. S. (2009). Cholecystolithiasis. In *Encyclopedia of Molecular Mechanisms of Disease* (pp. 325-326). Springer Berlin Heidelberg.
https://doi.org/10.1007/978-3-540-29676-8_324
 452. Petrick, J. L., Li, N., Anderson, L. A., Bernstein, L., Corley, D. A., El Serag, H. B., Hardikar, S., Liao, L. M., Liu, G., Murray, L. J., Rubenstein, J. H., Schneider, J. L., Shaheen, N. J., Thrift, A. P., van den Brandt, P. A., Vaughan, T. L., Whiteman, D. C., Wu, A. H., Zhao, W. K., ... Cook, M. B. (2019). Diabetes in relation to Barrett's esophagus and adenocarcinomas of the esophagus: A pooled study from the International Barrett's and Esophageal Adenocarcinoma Consortium. *Cancer*, 125(23), 4210-4223.
<https://doi.org/10.1002/cncr.32444>

453. Petrie, J. R., & Boyle, J. G. (2022). Diabetes mellitus. In S. H. Ralston, I. Penman, & M. W. J. Strachan (Eds.), *Davidson's principles and practice of medicine* (24th ed., pp. 703-754). Elsevier.
454. Pierce, B. L. (2012). Why are diabetics at reduced risk for prostate cancer? A review of the epidemiologic evidence. *Urologic Oncology: Seminars and Original Investigations*, 30(5), 735-743.
<https://doi.org/10.1016/j.urolonc.2012.07.008>
455. Pierce, B. L., Plymate, S., Ostrander, E. A., & Stanford, J. L. (2008). Diabetes mellitus and prostate cancer risk. *The Prostate*, 68(10), 1126-1132. <https://doi.org/10.1002/pros.20777>
456. Pizzato, M., Turati, F., Rosato, V., & la Vecchia, C. (2019). Exploring the link between diabetes and pancreatic cancer. *Expert Review of Anticancer Therapy*, 19(8), 681-687. <https://doi.org/10.1080/14737140.2019.1642109>
457. Poppe, L., Steen, J., Loh, W. W., Crombez, G., De Block, F., Jacobs, N., Tennant, P. W. G., Cauwenberg, J. Van, & Paepe, A. L. De. (2024). How to develop causal directed acyclic graphs for observational health research: a scoping review. *Health Psychology Review*, 1-21.
<https://doi.org/10.1080/17437199.2024.2402809>
458. Prizment, A. E., Anderson, K. E., Yuan, J. M., & Folsom, A. R. (2013). Diabetes and risk of bladder cancer among postmenopausal women in the Iowa women's health study. *Cancer Causes and Control*, 24(3), 603-608.
<https://doi.org/10.1007/s10552-012-0143-3>
459. Public Health England. (2019). Guidance, Health matters: preventing STIs. <https://www.gov.uk/government/publications/health-matters-preventing-stis/health-matters-preventing-stis>
460. Qi, L., Qi, X. L., Xiong, H. Y., Liu, Q., Li, J. X., Zhang, Y., Ma, X. Y., Wu, N., Liu, Q. Y., & Feng, L. G. (2014). Type 2 Diabetes Mellitus and Risk of Malignant Melanoma: A Systematic Review and Meta-Analysis of Cohort Studies. *Iranian Journal of Public Health*, 43(7), 857-866. %3CGo
461. Qiu, B., Chen, H., Zhang, E., Ma, F., An, G., Zong, Y., Shang, L., Zhang, Y., & Zhu, H. (2024). A machine learning prediction model for cancer risk in patients with type 2 diabetes based on clinical tests. *Technology and Health Care*, 32(3), 1431-1443. <https://doi.org/10.3233/THC-230385>
462. Rare Cancers Australia Ltd. (2023). Cancer knowledgebase.
<https://knowledge.rarecancers.org.au/knowledgebase/cancer-types>
463. Rau, H.-H., Hsu, C.-Y., Lin, Y.-A., Atique, S., Fuad, A., Wei, L.-M., & Hsu, M.-H. (2016). Development of a web-based liver cancer prediction model for type II diabetes patients by using an artificial neural network. *Computer Methods and Programs in Biomedicine*, 125, 58-65.
<https://doi.org/10.1016/j.cmpb.2015.11.009>
464. Raychaudhuri, S., Dieli-Conwright, C. M., Cheng, R. K., Barac, A., Reding, K. W., Vasbinder, A., Cook, K. L., Nair, V., Desai, P., & Simon, M. S. (2022). A review of research on the intersection between breast cancer and cardiovascular research in the Women's Health Initiative (WHI). *Frontiers in Oncology*, 12(March), 1-18. <https://doi.org/10.3389/fonc.2022.1039246>
465. Redaniel, M. T. M., Jeffreys, M., May, M. T., Ben-Shlomo, Y., & Martin, R. M. (2012). Associations of type 2 diabetes and diabetes treatment with breast cancer risk and mortality: A population-based cohort study among

- British women. *Cancer Causes and Control*, 23(11), 1785-1795.
<https://doi.org/http://dx.doi.org/10.1007/s10552-012-0057-0>
466. Reeves, H. L., Zaki, M. Y. W., & Day, C. P. (2016). Hepatocellular Carcinoma in Obesity, Type 2 Diabetes, and NAFLD. *Digestive Diseases and Sciences*, 61(5), 1234-1245. <https://doi.org/10.1007/s10620-016-4085-6>
 467. Reinholdt, K., Thomsen, L. T., Munk, C., Dehlendorff, C., Carstensen, B., Jørgensen, M. E., & Kjaer, S. K. (2021). Incidence of HPV-related Anogenital Intraepithelial Neoplasia and Cancer in Men With Diabetes Compared With the General Population. *Epidemiology*, 32(5), 705-711. <https://doi.org/10.1097/EDE.0000000000001375>
 468. Reis, A. M., Quintal, C., & Lourenço. (2018). Killing two birds with one stone? Association between tobacco and alcohol consumption. *Public Health*, 154, 136-143. <https://doi.org/10.1016/j.puhe.2017.10.019>
 469. Ren, H.-B., Yu, T., Liu, C., & Li, Y.-Q. (2011). Diabetes mellitus and increased risk of biliary tract cancer: systematic review and meta-analysis. *Cancer Causes & Control*, 22(6), 837-847. <https://doi.org/10.1007/s10552-011-9754-3>
 470. Rennert, G., Rennert, H. S., Gronich, N., Pinchev, M., & Gruber, S. B. (2020). Use of metformin and risk of breast and colorectal cancer. *Diabetes Research and Clinical Practice*, 165, 108232. <https://doi.org/10.1016/j.diabres.2020.108232>
 471. Rhodes, R. E., Mark, R. S., & Temmel, C. P. (2012). Adult Sedentary Behavior. *American Journal of Preventive Medicine*, 42(3), e3-e28. <https://doi.org/10.1016/j.amepre.2011.10.020>
 472. Ribeiro, C. M., Beserra, B. T. S., Silva, N. G., Lima, C. L., Rocha, P. R. S., Coelho, M. S., Neves, F. D. A. R., & Amato, A. A. (2020). Exposure to endocrine-disrupting chemicals and anthropometric measures of obesity: A systematic review and meta-analysis. *BMJ Open*, 10(6). <https://doi.org/10.1136/bmjopen-2019-033509>
 473. Rodbard, H. W., Blonde, L., Braithwaite, S. S., Brett, E. M., Cobin, R. H., Handelsman, Y., Hellman, R., Jellinger, P. S., Jovanovic, L. G., Levy, P., Mechanick, J. I., Zangeneh, F., & AACE Diabetes Mellitus Clinical Practice Guidelines Task Force. (2007). American Association of Clinical Endocrinologists medical guidelines for clinical practice for the management of diabetes mellitus. *Endocrine Practice: Official Journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists*, 13 Suppl 1(SUPPL. 1), 1-68. <https://doi.org/10.4158/EP.13.S1.1>
 474. Rodriguez, C., Patel, A. V., Mondul, A. M., Jacobs, E. J., Thun, M. J., & Calle, E. E. (2005). Diabetes and risk of prostate cancer in a prospective cohort of US men. *American Journal of Epidemiology*, 161(2), 147-152. <https://doi.org/10.1093/aje/kwh334>
 475. Rollison, D. E., Giuliano, A. R., Sellers, T. A., Laronga, C., Sweeney, C., Risendal, B., Baumgartner, K. B., Byers, T., & Slattery, M. L. (2007). Population-based case-control study of diabetes and breast cancer risk in Hispanic and non-Hispanic white women living in US southwestern states. *American Journal of Epidemiology*, 167(4), 447-456. <https://doi.org/10.1093/aje/kwm322>

476. Rosato, V., Tavani, A., Gracia-Lavedan, E., Guinó, E., Castaño-Vinyals, G., Villanueva, C. M., Kogevinas, M., Polesel, J., Serraino, D., Pisa, F. E., Barbone, F., Moreno, V., La Vecchia, C., & Bosetti, C. (2016). Type 2 Diabetes, Antidiabetic Medications, and Colorectal Cancer Risk: Two Case-Control Studies from Italy and Spain. *Frontiers in Oncology*, 6(OCT). <https://doi.org/10.3389/fonc.2016.00210>
477. Rosato, V., Zucchetto, A., Bosetti, C., Dal Maso, L., Montella, M., Pelucchi, C., Negri, E., Franceschi, S., & La Vecchia, C. (2011). Metabolic syndrome and endometrial cancer risk. *Annals of Oncology*, 22(4), 884-889. <https://doi.org/10.1093/annonc/mdq464>
478. Rose, D. P., Haffner, S. M., & Baillargeon, J. (2007). Adiposity, the Metabolic Syndrome, and Breast Cancer in African-American and White American Women. *Endocrine Reviews*, 28(7), 763-777. <https://doi.org/10.1210/er.2006-0019>
479. Rysz, J., Franczyk, B., Lawinski, J., Olszewski, R., & Gluba-Brzozka, A. (2020). The Role of Metabolic Factors in Renal Cancers. *International Journal of Molecular Sciences*, 21(19). <https://doi.org/10.3390/ijms21197246>
480. S., L., K., B., J.K., M., L., H., A., M., E., I., T., Y., K., K., M.J., D., Ling, S., Brown, K., Miksza, J. K., Howells, L., Morrison, A., Issa, E., Yates, T., Khunti, K., Davies, M. J., & Zaccardi, F. (2020). Association of type 2 diabetes with cancer: A meta-analysis with bias analysis for unmeasured confounding in 151 cohorts comprising 32 million people. *Diabetes Care*, 43(9), 2313-2322. <https://doi.org/https://dx.doi.org/10.2337/dc20-0204> PT - Article
481. Saarela, K., Tuomilehto, J., Sund, R., Keskimäki, I., Hartikainen, S., & Pukkala, E. (2019). Cancer incidence among Finnish people with type 2 diabetes during 1989-2014. *European Journal of Epidemiology*, 34(3), 259-265. <https://doi.org/10.1007/s10654-018-0438-0>
482. Sacerdote, C., & Ricceri, F. (2018). Epidemiological dimensions of the association between type 2 diabetes and cancer: A review of observational studies. *Diabetes Research and Clinical Practice*, 143, 369-377. <https://doi.org/10.1016/j.diabres.2018.03.002>
483. Saed, L., Varse, F., Baradaran, H. R., Moradi, Y., Khateri, S., Friberg, E., Khazaei, Z., Gharahjeh, S., Tehrani, S., Sioofy-Khojine, A. B., & Najmi, Z. (2019). The effect of diabetes on the risk of endometrial Cancer: An updated a systematic review and meta-analysis. *BMC Cancer*, 19(1), 1-10. <https://doi.org/10.1186/s12885-019-5748-4>
484. Safaee, A., Fatemi, R., Pourhoseingholi, M. A., Moghimi Dehkordi, B., Vahedi, M., Pourhoseingholi, A., Malek Nemati, F., & Zali, M. R. (2009). Positive association between diabetes mellitus and risk of colorectal cancer. *Iranian Journal of Cancer Prevention*, 2(4), 189-193. <http://www.ijcp.ir/download.asp?code=IJCP0924189>
485. Saka Herrán, C., Jané-Salas, E., Estrugo Devesa, A., & López-López, J. (2018). Protective effects of metformin, statins and anti-inflammatory drugs on head and neck cancer: A systematic review. *Oral Oncology*, 85(August), 68-81. <https://doi.org/10.1016/j.oraloncology.2018.08.015>
486. Salinas-Martínez, A. M., Flores-Cortés, L. I., Cardona-Chavarría, J. M., Hernández-Gutiérrez, B., Abundis, A., Vázquez-Lara, J., & González-

- Guajardo, E. E. (2014). Prediabetes, Diabetes, and Risk of Breast Cancer: A Case-Control Study. *Archives of Medical Research*, 45(5), 432-438. <https://doi.org/10.1016/j.arcmed.2014.06.004>
487. Saltzman, B. S., Doherty, J. A., Hill, D. A., Beresford, S. A., Voigt, L. F., Chen, C., & Weiss, N. S. (2007). Diabetes and Endometrial Cancer: An Evaluation of the Modifying Effects of Other Known Risk Factors. *American Journal of Epidemiology*, 167(5), 607-614. <https://doi.org/10.1093/aje/kwm333>
488. Samson, M. E., Adams, S. A., Orekoya, O., & Hebert, J. R. (2016). Understanding the Association of Type 2 Diabetes Mellitus in Breast Cancer Among African American and European American Populations in South Carolina. *Journal of Racial and Ethnic Health Disparities*, 3(3), 546-554. <https://doi.org/10.1007/s40615-015-0173-0>
489. Samuel, S. M., Varghese, E., Varghese, S., & Büsselberg, D. (2018). Challenges and perspectives in the treatment of diabetes associated breast cancer. *Cancer Treatment Reviews*, 70, 98-111. <https://doi.org/10.1016/j.ctrv.2018.08.004>
490. Sanderson, M., Peltz, G., Perez, A., Johnson, M., Vernon, S. W., Fernandez, M. E., & Fadden, M. K. (2010). Diabetes, physical activity and breast cancer among Hispanic women. *Cancer Epidemiology*, 34(5), 556-561. <https://doi.org/10.1016/j.canep.2010.06.001>
491. Sandhu, M. S. (2001). Self reported non-insulin dependent diabetes, family history, and risk of prevalent colorectal cancer: population based, cross sectional study. *Journal of Epidemiology & Community Health*, 55(11), 804-805. <https://doi.org/10.1136/jech.55.11.804>
492. Sauer, B., & VanderWeele, T. J. (2013). Supplement 2 - Use of Directed Acyclic Graphs. *Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide*, May, 177-184.
493. Sayyid, R. K., & Fleshner, N. E. (2016). Diabetes Mellitus Type 2: A Driving Force for Urological Complications. *Trends in Endocrinology & Metabolism*, 27(5), 249-261. <https://doi.org/10.1016/j.tem.2016.02.006>
494. Schettgen, T., Gube, M., Alt, A., Fromme, H., & Kraus, T. (2011). Pilot study on the exposure of the German general population to non-dioxin-like and dioxin-like PCBs. *International Journal of Hygiene and Environmental Health*, 214(4), 319-325. <https://doi.org/10.1016/j.ijheh.2011.04.002>
495. Schintler, L. A., & McNeely, C. L. (Eds.). (2022). *Encyclopedia of Big Data*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-32010-6>
496. Schlesinger, S., Aleksandrova, K., Pischon, T., Jenab, M., Fedirko, V., Trepo, E., Overvad, K., Roswall, N., Tjønneland, A., Boutron-Ruault, M. C., Fagherazzi, G., Racine, A., Kaaks, R., Grote, V. A., Boeing, H., Trichopoulou, A., Pantzalis, M., Kritikou, M., Mattiello, A., ... Nöthlings, U. (2013). Diabetes mellitus, insulin treatment, diabetes duration, and risk of biliary tract cancer and hepatocellular carcinoma in a European cohort. *Annals of Oncology*, 24(9), 2449-2455. <https://doi.org/10.1093/annonc/mdt204>
497. Schmid, D., Behrens, G., Jochem, C., Keimling, M., & Leitzmann, M. (2013). Physical activity, diabetes, and risk of thyroid cancer: a systematic

- review and meta-analysis. *European Journal of Epidemiology*, 28(12), 945-958. <https://doi.org/10.1007/s10654-013-9865-0>
498. Schott, S., Schneeweiss, A., & Sohn, C. (2010). Breast Cancer and Diabetes Mellitus. *Experimental and Clinical Endocrinology & Diabetes*, 118(10), 673-677. <https://doi.org/10.1055/s-0030-1254116>
 499. Schumacher, A. J., Chen, Q., Attaluri, V., McLemore, E. C., & Chao, C. R. (2021). Metabolic Risk Factors Associated with Early-Onset Colorectal Adenocarcinoma: A Case-Control Study at Kaiser Permanente Southern California. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, 30(10), 1792-1798. <https://doi.org/10.1158/1055-9965.EPI-20-1127>
 500. Schwarzer, G., Carpenter, J. R., & Rücker, G. (2015). *Meta-Analysis with R*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-21416-0>
 501. Sekula, P., Del Greco M, F., Pattaro, C., & Köttgen, A. (2016). Mendelian Randomization as an Approach to Assess Causality Using Observational Data. *Journal of the American Society of Nephrology: JASN*, 27(11), 3253-3265. <https://doi.org/10.1681/ASN.2016010098>
 502. Sellers, T. A., Sprafka, J. M., Gapstur, S. M., Rich, S. S., Potter, J. D., Ross, J. A., McGovern, P. G., Nelson, C. L., & Folsom, A. R. (1994). Does body fat distribution promote familial aggregation of adult onset diabetes mellitus and postmenopausal breast cancer? *Epidemiology*, 5(1), 102-108. <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed5&NEWS=N&AN=24021121>
 503. Seo, Y.-G., Choi, H.-C., An, A. R., Park, D. J., Park, Y. J., Lee, K. E., Park, S. K., Hwang, Y., & Cho, B. (2017). The Association between Type 2 Diabetes Mellitus and Thyroid Cancer. *Journal of Diabetes Research*, 2017, 1-8. <https://doi.org/10.1155/2017/5850879>
 504. Setiawan, V. W., Stram, D. O., Porcel, J., Chari, S. T., Maskarinec, G., Le Marchand, L., Wilkens, L. R., Haiman, C. A., Pandol, S. J., & Monroe, K. R. (2019). Pancreatic Cancer Following Incident Diabetes in African Americans and Latinos: The Multiethnic Cohort. *JNCI: Journal of the National Cancer Institute*, 111(1), 27-33. <https://doi.org/10.1093/jnci/djy090>
 505. Shackley, D. C., Whytock, C., Parry, G., Clarke, L., Vincent, C., Harrison, A., John, A., Provost, L., & Power, M. (2017). Variation in the prevalence of urinary catheters: A profile of National Health Service patients in England. *BMJ Open*, 7(6), 1-8. <https://doi.org/10.1136/bmjopen-2016-013842>
 506. Sheng, B.-W., Zhang, J.-Q., Chen, M., & Ma, M. (2023). The inverse association between fasting blood glucose and the occurrence of gallbladder cancer in type 2 diabetes mellitus patients: a case-control study. *Journal of Cancer Research and Clinical Oncology*, 149(12), 10387-10398. <https://doi.org/10.1007/s00432-023-04944-3>
 507. Shimoyama, S. (2013). Diabetes mellitus carries a risk of gastric cancer: A metaanalysis. *World Journal of Gastroenterology*, 19(40), 6902-6910. <https://doi.org/10.3748/wjg.v19.i40.6902>
 508. Shlomai, G., Neel, B., LeRoith, D., & Gallagher, E. J. (2016). Type 2 Diabetes Mellitus and Cancer: The Role of Pharmacotherapy. *Journal of*

- Clinical Oncology, 34(35), 4261-4269.
<https://doi.org/10.1200/JCO.2016.67.4044>
509. Shoff, S. M., & Newcomb, P. A. (1998). Diabetes, Body Size, and Risk of Endometrial Cancer. *American Journal of Epidemiology*, 148(3), 234-240.
<https://doi.org/10.1093/oxfordjournals.aje.a009630>
 510. Siddiqui, A. A. (2011). Metabolic Syndrome and Its Association With Colorectal Cancer: A Review. *American Journal of the Medical Sciences*, 341(3), 227-231. <https://doi.org/10.1097/MAJ.0b013e3181df9055>
 511. Sikdar, K. C., Walsh, S. J., Roche, M., Jiang, Y., Syrowatka, A., & Collins, K. D. (2013). Diabetes and sex-specific colorectal cancer risks in Newfoundland and Labrador: A population-based retrospective cohort study. *Canadian Journal of Public Health*, 104(2), 101-107.
<https://doi.org/10.1007/bf03405668>
 512. Silveira, E. A., Mendonça, C. R., Delpino, F. M., Elias Souza, G. V., Pereira de Souza Rosa, L., de Oliveira, C., & Noll, M. (2022). Sedentary behavior, physical inactivity, abdominal obesity and obesity in adults and older adults: A systematic review and meta-analysis. *Clinical Nutrition ESPEN*, 50, 63-73. <https://doi.org/10.1016/j.clnesp.2022.06.001>
 513. Simon, D., & Balkau, B. (2010). Diabetes mellitus, hyperglycaemia and cancer. *Diabetes & Metabolism*, 36(3), 182-191.
<https://doi.org/10.1016/j.diabet.2010.04.001>
 514. Simon, T. G., King, L. Y., Chong, D. Q., Nguyen, L. H., Ma, Y., VoPham, T., Giovannucci, E. L., Fuchs, C. S., Meyerhardt, J. A., Corey, K. E., Khalili, H., Chung, R. T., Zhang, X., & Chan, A. T. (2018). Diabetes, metabolic comorbidities, and risk of hepatocellular carcinoma: Results from two prospective cohort studies. *Hepatology*, 67(5), 1797-1806.
<https://doi.org/10.1002/hep.29660>
 515. Soltani, S., Abdollahi, S., Aune, D., & Jayedi, A. (2021). Body mass index and cancer risk in patients with type 2 diabetes: a dose-response meta-analysis of cohort studies. *Scientific Reports*, 11(1), 2479.
<https://doi.org/10.1038/s41598-021-81671-0>
 516. Song, J. W., & Chung, K. C. (2010). Observational Studies: Cohort and Case-Control Studies. *Plastic and Reconstructive Surgery*, 126(6), 2234-2242.
<https://doi.org/10.1097/PRS.0b013e3181f44abc>
 517. Song, S., Wang, B., Zhang, X., Hao, L., Hu, X., Li, Z., & Sun, S. (2015). Long-Term Diabetes Mellitus Is Associated with an Increased Risk of Pancreatic Cancer: A Meta-Analysis. *PLOS ONE*, 10(7), e0134321.
<https://doi.org/10.1371/journal.pone.0134321>
 518. Soper, M. (2002). Effects of cigarette smoke on the immune system. *Nat Rev Immunol* 2002;2:372-377. *Nat Rev Immunol*, 2(May), 372-377.
 519. Sowers, J. R., Epstein, M., & Frohlich, E. D. (2001). Diabetes, hypertension, and cardiovascular disease: an update. *Hypertension (Dallas, Tex.: 1979)*, 37(4), 1053-1059. <https://doi.org/10.1161/01.hyp.37.4.1053>
 520. Sponholtz, T. R., Palmer, J. R., Rosenberg, L., Hatch, E. E., Adams-Campbell, L. L., & Wise, L. A. (2016). Body Size, Metabolic Factors, and Risk of Endometrial Cancer in Black Women. *American Journal of Epidemiology*, 183(4), 259-268. <https://doi.org/10.1093/aje/kwv186>
 521. Sripa, B., Kaewkes, S., Sithithaworn, P., Mairiang, E., Laha, T., Smout, M., Pairojikul, C., Bhudhisawasdi, V., Tesana, S., Thinkamrop, B., Bethony, J.

- M., Loukas, A., & Brindley, P. J. (2007). Liver fluke induces cholangiocarcinoma. *PLoS Medicine*, 4(7), 1148-1155. <https://doi.org/10.1371/journal.pmed.0040201>
522. Stacey, D., Hill, S., Little, M., & Seguin, M. (2010). Causal language and research on interventions: Suggestions for researchers. *Psychotherapy and Psychosomatics*, 79(3), 139-145. <https://doi.org/10.1159/000322940>
 523. Standerwick, K., Davies, C., Tucker, L., & Sheron, N. (2007). Binge drinking, sexual behaviour and sexually transmitted infection in the UK. *International Journal of STD & AIDS*, 18(12), 810-813. <https://doi.org/10.1258/095646207782717027>
 524. Stanford Medicine. (2023). Conditions & Treatments. <https://stanfordhealthcare.org/>
 525. Starup-Linde, J., Karlstad, Ø., Eriksen, S., Vestergaard, P., Bronsveld, H., Vries, F., Andersen, M., Auvinen, A., Haukka, J., Hjellvik, V., Bazelier, M., Boer, A., Furu, K., & De Bruin, M. (2013). CARING (CAncer Risk and INsulin analoGues): The Association of Diabetes Mellitus and Cancer Risk with Focus on Possible Determinants - A Systematic Review and a Meta-Analysis. *Current Drug Safety*, 8(5), 296-332. <https://doi.org/10.2174/15748863113086660071>
 526. Steinbrecher, A., Morimoto, Y., Heak, S., Ollberding, N. J., Geller, K. S., Grandinetti, A., Kolonel, L. N., & Maskarinec, G. (2011). The Preventable Proportion of Type 2 Diabetes by Ethnicity: The Multiethnic Cohort. *Annals of Epidemiology*, 21(7), 526-535. <https://doi.org/10.1016/j.annepidem.2011.03.009>
 527. Stott-Miller, M., Chen, C., Chuang, S., Lee, Y. A., Boccia, S., Brenner, H., Cadoni, G., Dal Maso, L., La Vecchia, C., Lazarus, P., Levi, F., Matsuo, K., Morgenstern, H., Müller, H., Muscat, J., Olshan, A. F., Purdue, M. P., Serraino, D., Vaughan, T. L., ... Schwartz, S. M. (2012). History of Diabetes and Risk of Head and Neck Cancer: A Pooled Analysis from the International Head and Neck Cancer Epidemiology Consortium. *Cancer Epidemiology, Biomarkers & Prevention*, 21(2), 294-304. <https://doi.org/10.1158/1055-9965.EPI-11-0590>
 528. Strain, T., Brage, S., Sharp, S. J., Richards, J., Tainio, M., Ding, D., Benichou, J., & Kelly, P. (2020). Use of the prevented fraction for the population to determine deaths averted by existing prevalence of physical activity: a descriptive study. *The Lancet Global Health*, 8(7), e920-e930. [https://doi.org/10.1016/S2214-109X\(20\)30211-4](https://doi.org/10.1016/S2214-109X(20)30211-4)
 529. Suarez Arbelaiz, M. C., Nackeeran, S., Shah, K., Blachman-Braun, R., Bronson, I., Towe, M., Bhat, A., Marcovich, R., Ramasamy, R., & Shah, H. N. (2023). Association between body mass index, metabolic syndrome and common urologic conditions: a cross-sectional study using a large multi-institutional database from the United States. *Annals of Medicine*, 55(1). <https://doi.org/10.1080/07853890.2023.2197293>
 530. Suchanek, S., Grega, T., Ngo, O., Vojtechova, G., Majek, O., Minarikova, P., Brogyuk, N., Bunganic, B., Seifert, B., Dusek, L., & Zavoral, M. (2016a). How significant is the association between metabolic syndrome and prevalence of colorectal neoplasia? *World Journal of Gastroenterology*, 22(36), 8103. <https://doi.org/10.3748/wjg.v22.i36.8103>

531. Suh, S., & Kim, K.-W. (2019). Diabetes and Cancer: Cancer Should Be Screened in Routine Diabetes Assessment. *Diabetes & Metabolism Journal*, 43(6), 733. <https://doi.org/10.4093/dmj.2019.0177>
532. Sun, L., & Yu, S. (2012). Diabetes Mellitus Is an Independent Risk Factor for Colorectal Cancer. *Digestive Diseases and Sciences*, 57(6), 1586-1597. <https://doi.org/10.1007/s10620-012-2059-x>
533. Sun, X., Zhang, Q., Kadier, K., Hu, P., Liu, X., Liu, J., Yan, Y., Sun, C., Yau, V., Lowe, S., Meng, M., Liu, Z., & Zhou, M. (2023). Association between diabetes status and breast cancer in US adults: findings from the US National Health and Nutrition Examination Survey. *Frontiers in Endocrinology*, 14(June), 1059303. <https://doi.org/10.3389/fendo.2023.1059303>
534. Suzuki, E., Shinozaki, T., & Yamamoto, E. (2020). Causal diagrams: Pitfalls and tips. *Journal of Epidemiology*, 30(4), 153-162. <https://doi.org/10.2188/jea.JE20190192>
535. Szablewski, L. (2014). Diabetes mellitus: influences on cancer risk. *Diabetes/Metabolism Research and Reviews*, 30(7), 543-553. <https://doi.org/10.1002/dmrr>
536. T. Bazelier, M., de Vries, F., Vestergaard, P., Leufkens, H., & De Bruin, M. (2013). Use of Thiazolidinediones and Risk of Bladder Cancer: Disease or Drugs? *Current Drug Safety*, 8(5), 364-370. <https://doi.org/10.2174/15748863113086660069>
537. Tabassum, I., Mahmood, H., & Faheem, M. (2016). Type 2 Diabetes Mellitus as a Risk Factor for Female Breast Cancer in the Population of Northern Pakistan. *Asian Pacific Journal of Cancer Prevention: APJCP*, 17(7), 3255-3258. <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed17&NEWS=N&AN=614380224>
538. Talamini, R., Franceschi, S., Favero, A., Negri, E., Parazzini, F., & La Vecchia, C. (1997). Selected medical conditions and risk of breast cancer. *British Journal of Cancer*, 75(11), 1699-1703. <https://doi.org/10.1038/bjc.1997.289>
539. Tan, J., You, Y., Guo, F., Xu, J., Dai, H., & Bie, P. (2017). Association of elevated risk of pancreatic cancer in diabetic patients: A systematic review and meta-analysis. *Oncology Letters*, 13(3), 1247-1255. <https://doi.org/10.3892/ol.2017.5586>
540. Tang, Y.-L., Zhu, L.-Y., Li, Y., Yu, J., Wang, J., Zeng, X.-X., Hu, K.-X., Liu, J.-Y., & Xu, J.-X. (2017). Metformin Use Is Associated with Reduced Incidence and Improved Survival of Endometrial Cancer: A Meta-Analysis. *BioMed Research International*, 2017, 1-9. <https://doi.org/10.1155/2017/5905384>
541. Tarantino, G., Crocetto, F., di Vito, C., Creta, M., Martino, R., Pandolfo, S. D., Pesce, S., Napolitano, L., Capone, D., & Imbimbo, C. (2021). Association of NAFLD and Insulin Resistance with Non Metastatic Bladder Cancer Patients: A Cross-Sectional Retrospective Study. *Journal of Clinical Medicine*, 10(2), 346. <https://doi.org/10.3390/jcm10020346>
542. Tavani, A., Gallus, S., Bertuzzi, M., Dal Maso, L., Zucchetto, A., Negri, E., Franceschi, S., Ramazzotti, V., Montella, M., & La Vecchia, C. (2005).

- Diabetes Mellitus and the Risk of Prostate Cancer in Italy. *European Urology*, 47(3), 313-317. <https://doi.org/10.1016/j.eururo.2004.10.027>
543. Tavani, A., Gallus, S., Bosetti, C., Tzonou, A., Lagiou, P., Negri, E., Trichopoulos, D., & La Vecchia, C. (2002). Diabetes and the risk of prostate cancer. *European Journal of Cancer Prevention*, 11(2), 125-128. <https://doi.org/10.1097/00008469-200204000-00003>
 544. Tennant, P. W. G., Murray, E. J., Arnold, K. F., Berrie, L., Fox, M. P., Gadd, S. C., Harrison, W. J., Keeble, C., Ranker, L. R., Textor, J., Tomova, G. D., Gilthorpe, M. S., & Ellison, G. T. H. (2021). Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. *International Journal of Epidemiology*, 50(2), 620-632. <https://doi.org/10.1093/ije/dyaa213>
 545. Textor, J., van der Zander, B., Gilthorpe, M. S., Liskiewicz, M., & Ellison, G. T. (2016). Robust causal inference using directed acyclic graphs: the R package “dagitty”. *International Journal of Epidemiology*, 45(6), 1887-1894. <https://doi.org/10.1093/ije/dyw341>
 546. The American Society of Retina Specialists. (2024). The Foundation. <https://www.asrs.org/patients/retinal-diseases/42/choroidal-nevus>
 547. The British Diabetic Association. (2023). Diabetes risk factors. <https://www.diabetes.org.uk/diabetes-the-basics/types-of-diabetes/type-2/diabetes-risk-factors>
 548. The Cochrane Collaboration. (2022). Review Manager (RevMan) (5.4.1). revman.cochrane.org
 549. The relationship between diabetes mellitus and gastric cancer and the potential benefits of metformin: An extensive review of the literature. (2021). *Biomolecules*, 11(7), 1022. <https://doi.org/https://dx.doi.org/10.3390/biom11071022> PT - Review
 550. Tian, T., Zhang, L., Ma, X., Zhou, J., & Shen, J. (2012). Diabetes Mellitus and Incidence and Mortality of Gastric Cancer: A Meta-Analysis. *Experimental and Clinical Endocrinology & Diabetes*, 120(04), 217-223. <https://doi.org/10.1055/s-0031-1297969>
 551. Tolman, K. G., Fonseca, V., Tan, M. H., & Dalpiaz, A. (2004). Narrative Review: Hepatobiliary Disease in Type 2 Diabetes Mellitus. *Annals of Internal Medicine*, 141(12), 946. <https://doi.org/10.7326/0003-4819-141-12-200412210-00011>
 552. Tomic, D., Shaw, J. E., & Magliano, D. J. (2022). The burden and risks of emerging complications of diabetes mellitus. *Nature Reviews Endocrinology*, 18(9), 525-539. <https://doi.org/10.1038/s41574-022-00690-7>
 553. Tseng, C. (2013). Diabetes is not an independent risk factor for hepatocellular carcinoma. *Diabetes/Metabolism Research and Reviews*, 29(7), 515-524. <https://doi.org/10.1002/dmrr.2422>
 554. Tseng, C. (2014). Diabetes and breast cancer in Taiwanese women: a detection bias? *European Journal of Clinical Investigation*, 44(10), 910-917. <https://doi.org/10.1111/eci.12323>
 555. Tseng, C.-H. (2011a). Diabetes and risk of bladder cancer: a study using the National Health Insurance database in Taiwan. *Diabetologia*, 54(8), 2009-2015. <https://doi.org/10.1007/s00125-011-2171-z>

556. Tseng, C.-H. (2011b). Diabetes and Risk of Prostate Cancer. *Diabetes Care*, 34(3), 616-621. <https://doi.org/10.2337/dc10-1640>
557. Tseng, C.-H. (2012a). Diabetes and non-Hodgkin's lymphoma: analyses of prevalence and annual incidence in 2005 using the National Health Insurance database in Taiwan. *Annals of Oncology*, 23(1), 153-158. <https://doi.org/10.1093/annonc/mdr334>
558. Tseng, C.-H. (2012b). Thyroid cancer risk is not increased in diabetic patients. *PloS One*, 7(12), e53096. <https://doi.org/10.1371/journal.pone.0053096>
559. Tseng, C.-H. (2013). Oral cancer in Taiwan: is diabetes a risk factor? *Clinical Oral Investigations*, 17(5), 1357-1364. <https://doi.org/10.1007/s00784-012-0820-3>
560. Tseng, C.-H. (2014). Diabetes but Not Insulin Increases the Risk of Lung Cancer: A Taiwanese Population-Based Study. *PLoS ONE*, 9(7), e101553. <https://doi.org/10.1371/journal.pone.0101553>
561. Tseng, C.-H. (2015). Type 2 Diabetes Mellitus and Kidney Cancer Risk: A Retrospective Cohort Analysis of the National Health Insurance. *PloS One*, 10(11), e0142480. <https://doi.org/10.1371/journal.pone.0142480>
562. Tshikuka, J. G., Boitshwarelo, M., Hamda, S. G., Tapera, R., Bukonda, N. K. Z., & Onokoko, M. E. (2024). The Prevalence of Overweight/Obesity and Related Comorbidities among HIV-Positive and HIV-Negative Patients Attending the Same Clinics in Gaborone, Botswana. *Advances in Infectious Diseases*, 14(03), 523-540. <https://doi.org/10.4236/aid.2024.143038>
563. Tsilidis, K. K., Allen, N. E., Appleby, P. N., Rohrmann, S., Nothlings, U., Arriola, L., Gunter, M. J., Chajes, V., Rinaldi, S., Romieu, I., Murphy, N., Riboli, E., Tzoulaki, I., Kaaks, R., Lukanova, A., Boeing, H., Pischon, T., Dahm, C. C., Overvad, K., ... Key, T. J. (2015). Diabetes mellitus and risk of prostate cancer in the European Prospective Investigation into Cancer and Nutrition. *International Journal of Cancer*, 136(2), 372-381. <https://doi.org/10.1002/ijc.28989>
564. Tsilidis, K. K., Kasimis, J. C., Lopez, D. S., Ntzani, E. E., & Ioannidis, J. P. A. (2015). Type 2 diabetes and cancer: umbrella review of meta-analyses of observational studies. *BMJ*, 350(jan02 1), g7607-g7607. <https://doi.org/10.1136/bmj.g7607>
565. Turati, F., Polesel, J., Di Maso, M., Montella, M., Libra, M., Grimaldi, M., Tavani, A., Serraino, D., La Vecchia, C., & Bosetti, C. (2015). Diabetes mellitus and the risk of bladder cancer: an Italian case-control study. *British Journal of Cancer*, 113(1), 127-130. <https://doi.org/10.1038/bjc.2015.178>
566. Ulcickas Yood, M., Oliveria, S. A., Campbell, U. B., & Koro, C. E. (2009). Incidence of cancer in a population-based cohort of patients with type 2 diabetes. *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, 3(1), 12-16. <https://doi.org/10.1016/j.dsx.2008.10.003>
567. Ullrich, S. E., Alcalay, J., Applegate, L. A., & Kripke, M. L. (2007). Immunosuppression in Phototherapy (Gregory Bock & Sara Harnett, Ed.; pp. 131-147). <https://doi.org/10.1002/9780470513842.ch9>
568. Undzyte, G., Patasius, A., Linkeviciute-Ulinskiene, D., Zabuliene, L., Stukas, R., Dulskas, A., & Smailyte, G. (2020). Increased kidney cancer risk in diabetes mellitus patients: a population-based cohort study in Lithuania.

- The Aging Male, 23(5), 1241-1245.
<https://doi.org/10.1080/13685538.2020.1755249>
569. United Nations, Department of Economic and Social Affairs, P. D. (2022). World Population Prospects 2022, Online Edition (Online Edi). United Nations. <https://population.un.org/wpp>
 570. United Nations Development Programme. (2024). Human Development Index (HDI) Data Center. <https://hdr.undp.org/data-center/human-development-index#/indicies/HDI>
 571. University of Florida Health. (2022). UFHealth. <https://ufhealth.org/conditions-and-treatments/>
 572. University of Oxford. Medical Sciences Division. (2024). Nuffield Department of Women's and Reproductive Health. <https://www.wrh.ox.ac.uk/research>
 573. University of Rochester Medical Center (URMC). (2023). Search Adult and Children's Health Encyclopedia. <https://www.urmc.rochester.edu/encyclopedia.aspx>
 574. Valent, F. (2015). Diabetes mellitus and cancer of the digestive organs: An Italian population-based cohort study. *Journal of Diabetes and Its Complications*, 29(8), 1056-1061.
<https://doi.org/10.1016/j.jdiacomp.2015.07.017>
 575. VanderWeele, T. J. (2020). Optimal approximate conversions of odds ratios and hazard ratios to risk ratios. *Biometrics*, 76(3), 746-752.
<https://doi.org/10.1111/biom.13197>
 576. Vatsaba, T. S., & Sokolova, L. K. (2019). BREAST CANCER IN WOMEN WITH DIABETES MELLITUS TYPE 2. *World of Medicine and Biology*, 68(2), 30-34.
<https://doi.org/10.26724/2079-8334-2019-2-68-30-34>
 577. Viera, A. J. (2008). Odds Ratios and Risk Ratios: What's the Difference and Why Does It Matter? *Southern Medical Journal*, 101(7), 730-734.
<https://doi.org/10.1097/SMJ.0b013e31817a7ee4>
 578. Vigneri, P., Frasca, F., Sciacca, L., Pandini, G., & Vigneri, R. (2009). Diabetes and cancer. *Endocrine-Related Cancer*, 16(4), 1103-1123.
<https://doi.org/10.1677/ERC-09-0087>
 579. Wainwright, P., Scorletti, E., & Byrne, C. D. (2017). Type 2 Diabetes and Hepatocellular Carcinoma: Risk Factors and Pathogenesis. *Current Diabetes Reports*, 17(4), 20. <https://doi.org/10.1007/s11892-017-0851-x>
 580. Walker, E. J., Ko, A. H., Holly, E. A., & Bracci, P. M. (2015). Metformin use among type 2 diabetics and risk of pancreatic cancer in a clinic-based case-control study. *International Journal of Cancer*, 136(6), E646-E653.
<https://doi.org/10.1002/ijc.29120>
 581. Walker, J. J., Brewster, D. H., Colhoun, H. M., Fischbacher, C. M., Leese, G. P., Lindsay, R. S., McKnight, J. A., Philip, S., Sattar, N., Stockton, D. L., & Wild, S. H. (2013). Type 2 diabetes, socioeconomic status and risk of cancer in Scotland 2001-2007. *Diabetologia*, 56(8), 1712-1715.
<https://doi.org/10.1007/s00125-013-2937-6>
 582. Wallner, L. P., Morgenstern, H., McGree, M. E., Jacobson, D. J., St. Sauver, J. L., Jacobsen, S. J., & Sarma, A. V. (2011). The effects of metabolic conditions on prostate cancer incidence over 15 years of follow-up: results from the Olmsted County Study. *BJU International*, 107(6), 929-935.
<https://doi.org/10.1111/j.1464-410X.2010.09703.x>

583. Wang, C., Wang, X., Gong, G., Ben, Q., Qiu, W., Chen, Y., Li, G., & Wang, L. (2012). Increased risk of hepatocellular carcinoma in patients with diabetes mellitus: A systematic review and meta-analysis of cohort studies. *International Journal of Cancer*, 130(7), 1639-1648. <https://doi.org/10.1002/ijc.26165>
584. Wang, C.-S., Yao, W.-J., Chang, T.-T., Wang, S.-T., & Chou, P. (2009). The Impact of Type 2 Diabetes on the Development of Hepatocellular Carcinoma in Different Viral Hepatitis Statuses. *Cancer Epidemiology, Biomarkers & Prevention*, 18(7), 2054-2060. <https://doi.org/10.1158/1055-9965.EPI-08-1131>
585. Wang, F., Gupta, S., & Holly, E. A. (2006). Diabetes Mellitus and Pancreatic Cancer in a Population-Based Case-Control Study in the San Francisco Bay Area, California. *Cancer Epidemiology, Biomarkers & Prevention*, 15(8), 1458-1463. <https://doi.org/10.1158/1055-9965.EPI-06-0188>
586. Wang, J.-Y., Chao, T.-T., Lai, C.-C., Wang, C.-Y., Wu, V.-C., Wang, S.-M., Lin, H.-I., Tang, S.-H., & Lin, H.-W. (2013). Risk of Colorectal Cancer in Type 2 Diabetic Patients: A Population-based Cohort Study. *Japanese Journal of Clinical Oncology*, 43(3), 258-263. <https://doi.org/10.1093/jjco/hys228>
587. Wang, L., Wang, L., Zhang, J., Wang, B., & Liu, H. (2017). Association between diabetes mellitus and subsequent ovarian cancer in women. *Medicine*, 96(16), e6396. <https://doi.org/10.1097/MD.00000000000006396>
588. Wang, L., Zhong, L., Xu, B., Chen, M., & Huang, H. (2020). Diabetes mellitus and the risk of ovarian cancer: A systematic review and meta-analysis of cohort and case-control studies. *BMJ Open*, 10(12). <https://doi.org/10.1136/bmjopen-2020-040137>
589. Wang, P., Kang, D., Cao, W., Wang, Y., & Liu, Z. (2012). Diabetes mellitus and risk of hepatocellular carcinoma: a systematic review and meta-analysis. *Diabetes/Metabolism Research and Reviews*, 28(2), 109-122. <https://doi.org/10.1002/dmrr.1291>
590. Wang, Y., Liu, X., Yan, P., Bi, Y., Liu, Y., & Zhang, Z. J. (2020). Association between type 1 and type 2 diabetes and risk of non-Hodgkin's lymphoma: A meta-analysis of cohort studies. *Diabetes & Metabolism*, 46(1), 8-19. <https://doi.org/10.1016/j.diabet.2019.04.006>
591. Wang, Y., O'Neil, A., Jiao, Y., Wang, L., Huang, J., Lan, Y., Zhu, Y., & Yu, C. (2019). Sex differences in the association between diabetes and risk of cardiovascular disease, cancer, and all-cause and cause-specific mortality: a systematic review and meta-analysis of 5,162,654 participants. *BMC Medicine*, 17(1), N.PAG-N.PAG. <https://doi.org/10.1186/s12916-019-1355-0>
592. Wang, Y., Sun, Y., Tang, J., Zhou, W., Liu, X., Bi, Y., & Zhang, Z. J. (2019). Does diabetes decrease the risk of glioma? A systematic review and meta-analysis of observational studies. *Annals of Epidemiology*, 30, 22-29.e3. <https://doi.org/10.1016/j.annepidem.2018.11.010>
593. Wang, Y., Wang, B., Yan, S., Shen, F., Cao, H., Fan, J., Zhang, R., & Gu, J. (2016). Type 2 diabetes and gender differences in liver cancer by considering different confounding factors: a meta-analysis of cohort studies. *Annals of Epidemiology*, 26(11), 764-772. <https://doi.org/10.1016/j.annepidem.2016.09.006>

594. Wang, Z., Yao, G., Yan, B., & Zhanghuang, C. (2023). Association between type 2 diabetes mellitus and prostate cancer. *Metabolism and Target Organ Damage*, 3(3), 15. <https://doi.org/10.20517/mtod.2023.34>
595. Weiderpass, E., Persson, I., Adami, H. O., Magnusson, C., Lindgren, A., & Baron, J. A. (2000). Body size in different periods of life, diabetes mellitus, hypertension, and risk of postmenopausal endometrial cancer (Sweden). *Cancer Causes and Control*, 11(2), 185-192. <https://doi.org/10.1023/A:1008946825313>
596. Weiss, H. A., Brinton, L. A., Potischman, N. A., Brogan, D., Coates, R. J., Gammon, M. D., Malone, K. E., & Schoenberg, J. B. (1999). Breast cancer risk in young women and history of selected medical conditions. *International Journal of Epidemiology*, 28(5), 816-823. <https://doi.org/10.1093/ije/28.5.816>
597. Welzel, T. M., Graubard, B. I., El-Serag, H. B., Shaib, Y. H., Hsing, A. W., Davila, J. A., & McGlynn, K. A. (2007). Risk Factors for Intrahepatic and Extrahepatic Cholangiocarcinoma in the United States: A Population-Based Case-Control Study. *Clinical Gastroenterology and Hepatology*, 5(10), 1221-1228. <https://doi.org/10.1016/j.cgh.2007.05.020>
598. Weyand, C. M., & Goronzy, J. J. (2016). Aging of the immune system: Mechanisms and therapeutic targets. *Annals of the American Thoracic Society*, 13(December), S422-S428. <https://doi.org/10.1513/AnnalsATS.201602-095AW>
599. Wild, S. H., Walker, J. J., McAllister, D. A., Read, S. H., Morling, J. R., Colhoun, H. M., Farran, B., McGurnaghan, S., McCrimmon, R., Sattar, N., & Byrne, C. D. (2018). Cardiovascular disease, cancer, and mortality among people with type 2 diabetes and alcoholic or nonalcoholic fatty liver disease hospital admission. *Diabetes Care*, 41(2), 341-347. <https://doi.org/http://dx.doi.org/10.2337/dc17-1590>
600. Williamson, E. J., Aitken, Z., Lawrie, J., Dharmage, S. C., Burgess, J. A., & Forbes, A. B. (2014). Introduction to causal diagrams for confounder selection. *Respirology*, 19(3), 303-311. <https://doi.org/10.1111/resp.12238>
601. Wojciechowska, J., Krajewski, W., Bolanowski, M., Kręcicki, T., & Zatoński, T. (2016). Diabetes and Cancer: a Review of Current Knowledge. *Experimental and Clinical Endocrinology & Diabetes*, 124(05), 263-275. <https://doi.org/10.1055/s-0042-100910>
602. Wolf, I., Sadetzki, S., Catane, R., Karasik, A., & Kaufman, B. (2005). Diabetes mellitus and breast cancer. *The Lancet Oncology*, 6(2), 103-111. [https://doi.org/10.1016/S1470-2045\(05\)01736-5](https://doi.org/10.1016/S1470-2045(05)01736-5)
603. World Economic Outlook, October 2023. (2023). International Monetary Fund. <https://doi.org/10.5089/9798400235801.081>
604. World Health Organization. (1994). Management of diabetes mellitus standards of care and clinical practice. World Health Organization Guidelines, DIN6, 1-35.
605. World Health Organization. (1999). Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications. Part 1, Diagnosis and classification of diabetes mellitus. World Health Organization, 59. <https://apps.who.int/iris/handle/10665/66040>
606. World Health Organization & International Diabetes Federation. (2006). Definition and diagnosis of diabetes mellitus and intermediate

- hyperglycaemia: report of a WHO/IDF consultation Organization. In World Health Organization. <https://apps.who.int/iris/handle/10665/43588>
607. Wu, A. H., Yu, M. C., Tseng, C.-C., Stanczyk, F. Z., & Pike, M. C. (2007). Diabetes and risk of breast cancer in Asian-American women. *Carcinogenesis*, 28(7), 1561-1566. <https://doi.org/10.1093/carcin/bgm081>
 608. Wu, L., Yu, C., Jiang, H., Tang, J., Huang, H.-L., Gao, J., & Zhang, X. (2013). Diabetes Mellitus and the Occurrence of Colorectal Cancer: An Updated Meta-Analysis of Cohort Studies. *Diabetes Technology & Therapeutics*, 15(5), 419-427. <https://doi.org/10.1089/dia.2012.0263>
 609. Xiao, B., Velez Edwards, D. R., Lucas, A., Drivas, T., Gray, K., Keating, B., Weng, C., Jarvik, G. P., Hakonarson, H., Kottyan, L., Elhadad, N., Wei, W.-Q., Luo, Y., Kim, D., Ritchie, M., Verma, S. S., & Regeneron Genetics Center *. (2023). Inference of Causal Relationships Between Genetic Risk Factors for Cardiometabolic Phenotypes and Female-Specific Health Conditions. *Journal of the American Heart Association*, 12(5), e026561. <https://doi.org/10.1161/JAHA.121.026561>
 610. Xu, B., Zhou, X., Li, X., Liu, C., & Yang, C. (2017). Diabetes mellitus carries a risk of esophageal cancer. *Medicine*, 96(35), e7944. <https://doi.org/10.1097/MD.0000000000007944>
 611. Xu, H., Jiang, H., Ding, G., Zhang, H., Zhang, L., Mao, S., & Ding, Q. (2013). Diabetes mellitus and prostate cancer risk of different grade or stage: A systematic review and meta-analysis. *Diabetes Research and Clinical Practice*, 99(3), 241-249. <https://doi.org/10.1016/j.diabres.2012.12.003>
 612. Xu, X., Wu, J., Mao, Y., Zhu, Y., Hu, Z., Xu, X., Lin, Y., Chen, H., Zheng, X., Qin, J., & Xie, L. (2013). Diabetes Mellitus and Risk of Bladder Cancer: A Meta-Analysis of Cohort Studies. *PLoS ONE*, 8(3), e58079. <https://doi.org/10.1371/journal.pone.0058079>
 613. Xue, F., & Michels, K. B. (2007). Diabetes, metabolic syndrome, and breast cancer: a review of the current evidence. *American Journal of Clinical Nutrition*, 86(3), 823S-835S. %3CGo
 614. Xue, L., Wang, H., He, Y., Sui, M., Li, H., Mei, L., & Ying, X. (2022). Incidence and risk factors of diabetes mellitus in the Chinese population: a dynamic cohort study. *BMJ Open*, 12(11), e060730. <https://doi.org/10.1136/bmjopen-2021-060730>
 615. Yan, P., Wang, Y., Fu, T., Liu, Y., & Zhang, Z.-J. (2021). The association between type 1 and 2 diabetes mellitus and the risk of leukemia: a systematic review and meta-analysis of 18 cohort studies. *Endocrine Journal*, 68(3), 281-289. <https://doi.org/10.1507/endocrj.EJ20-0138>
 616. Yan, P., Wang, Y., Yu, X., Liu, Y., & Zhang, Z.-J. (2021). Type 2 diabetes mellitus and risk of head and neck cancer subtypes: a systematic review and meta-analysis of observational studies. *Acta Diabetologica*, 58(5), 549-565. <https://doi.org/10.1007/s00592-020-01643-0>
 617. Yang, W. S., J., G., Li, H. L., Gao, Y. T., Xiang, Y. B., Shu, X. O., Cai, H., Yang, G., Zheng, W., Ji, B. T., & Rothman, N. (2013). Prospective evaluation of type 2 diabetes mellitus on the risk of primary liver cancer in chinese men and women. *Annals of Oncology*, 24(6), 1679-1685. <https://doi.org/http://dx.doi.org/10.1093/annonc/mdt017>

618. Yang, W.-S., Chen, P.-C., Lin, H.-J., Su, T.-C., Hsu, H.-C., Chen, M.-F., Lee, Y.-T., & Chien, K.-L. (2017). Association between type 2 diabetes and cancer incidence in Taiwan: data from a prospective community-based cohort study. *Acta Diabetologica*, 54(5), 455-461. <https://doi.org/10.1007/s00592-017-0966-1>
619. Yang, W.-S., Li, H.-L., Xu, H.-L., Yang, G., Gao, Y.-T., Zheng, W., Shu, X.-O., & Xiang, Y.-B. (2016). Type 2 diabetes and the risk of non-Hodgkin's lymphoma. *European Journal of Cancer Prevention*, 25(2), 149-154. <https://doi.org/10.1097/CEJ.0000000000000150>
620. Yang, W.-S., Yang, Y., Yang, G., Chow, W.-H., Li, H.-L., Gao, Y.-T., Ji, B.-T., Rothman, N., Zheng, W., Shu, X.-O., & Xiang, Y.-B. (2014). Pre-existing type 2 diabetes and risk of lung cancer: a report from two prospective cohort studies of 133 024 Chinese adults in urban Shanghai. *BMJ Open*, 4(7), e004875-e004875. <https://doi.org/10.1136/bmjopen-2014-004875>
621. Yang, X., So, W.-Y., Ma, R. C. W., Ko, G. T. C., Kong, A. P. S., Wang, Q., Cockram, C. S., Chow, C.-C., Chan, J. C. N., & Tong, P. C. Y. (2008). Predicting values of lipids and white blood cell count for all-site cancer in type 2 diabetes. *Endocrine Related Cancer*, 15(2), 597-607. <https://doi.org/10.1677/ERC-07-0266>
622. Yang, X.-Q., Xu, C., Sun, Y., & Han, R.-F. (2013). Diabetes Mellitus Increases the Risk of Bladder Cancer: An Updated Meta-analysis. *Asian Pacific Journal of Cancer Prevention*, 14(4), 2583-2589. <https://doi.org/10.7314/APJCP.2013.14.4.2583>
623. Yang, Y. X., Hennessy, S., & Lewis, J. D. (2005). Type 2 diabetes mellitus and the risk of colorectal cancer. *Clinical Gastroenterology and Hepatology*, 3(6), 587-594. [https://doi.org/10.1016/S1542-3565\(05\)00152-7](https://doi.org/10.1016/S1542-3565(05)00152-7)
624. Yang, Z., Wu, Y., Xu, L., Zhu, Z., Li, T., Yu, L., Gao, K., Zhang, X., Shen, P., Lin, H., Shui, L., Tang, M., Jin, M., Sun, Y., Wang, J., & Chen, K. (2023). Age at diagnosis modifies associations of type 2 diabetes with cancer incidence and mortality: a retrospective matched-cohort study. *Diabetologia*, 66(8), 1450-1459. <https://doi.org/10.1007/s00125-023-05920-9>
625. Yarmolinsky, J., Relton, C. L., Lophatananon, A., Muir, K., Menon, U., Gentry-Maharaj, A., Walther, A., Zheng, J., Fasching, P., Zheng, W., Yin Ling, W., Park, S. K., Kim, B.-G., Choi, J.-Y., Park, B., Davey Smith, G., Martin, R. M., & Lewis, S. J. (2019). Appraising the role of previously reported risk factors in epithelial ovarian cancer risk: A Mendelian randomization analysis. *PLOS Medicine*, 16(8), e1002893. <https://doi.org/10.1371/journal.pmed.1002893>
626. Yau, S. T. Y., Hung, C. T., Leung, E. Y. M., Chong, K. C., Lee, A., & Yeoh, E. K. (2024). Point-Based Prediction Model for Bladder Cancer Risk in Diabetes: A Random Survival Forest-Guided Approach. *Journal of Clinical Medicine*, 14(1), 4. <https://doi.org/10.3390/jcm14010004>
627. Ye, F., Chen, L., & Zheng, X. (2023). Diabetes and the risk of cirrhosis and HCC: An analysis of the UK Biobank. *Hepatology Communications*, 7(11). <https://doi.org/10.1097/HC9.0000000000000280>
628. Yedjou, C. G., Sims, J. N., Miele, L., Noubissi, F., Lowe, L., Fonseca, D. D., Alo, R. A., Payton, M., & Tchounwou, P. B. (2019). Health and Racial

- Disparity in Breast Cancer (pp. 31-49). https://doi.org/10.1007/978-3-030-20301-6_3
629. Yee, L. D., Mortimer, J. E., Natarajan, R., Dietze, E. C., & Seewaldt, V. L. (2020). Metabolic Health, Insulin, and Breast Cancer: Why Oncologists Should Care About Insulin. *Frontiers in Endocrinology*, 11. <https://doi.org/10.3389/fendo.2020.00058>
 630. Yeo, T. P. (2015). Demographics, Epidemiology, and Inheritance of Pancreatic Ductal Adenocarcinoma. *Seminars in Oncology*, 42(1), 8-18. <https://doi.org/10.1053/j.seminoncol.2014.12.002>
 631. Yeo, Y., Shin, D. W., Lee, J., Han, K., Park, S. H., Jeon, K. H., Shin, J., Shin, A., & Park, J. (2021). Personalized 5-Year Prostate Cancer Risk Prediction Model in Korea Based on Nationwide Representative Data. *Journal of Personalized Medicine*, 12(1), 2. <https://doi.org/10.3390/jpm12010002>
 632. Yoon, J. M. (2013). Pre-existing diabetes mellitus increases the risk of gastric cancer: A meta-analysis. *World Journal of Gastroenterology*, 19(6), 936. <https://doi.org/10.3748/wjg.v19.i6.936>
 633. Yu, F., Guo, Y., Wang, H., Feng, J., Jin, Z., Chen, Q., Liu, Y., & He, J. (2016). Type 2 diabetes mellitus and risk of colorectal adenoma: a meta-analysis of observational studies. *BMC Cancer*, 16(1), 642. <https://doi.org/10.1186/s12885-016-2685-3>
 634. Yuan, S., Kar, S., Carter, P., Vithayathil, M., Mason, A. M., Burgess, S., & Larsson, S. C. (2020). Is Type 2 Diabetes Causally Associated With Cancer Risk? Evidence From a Two-Sample Mendelian Randomization Study. *Diabetes*, 69(7), 1588-1596. <https://doi.org/10.2337/db20-0084>
 635. Zabuliene, L., Kaceniene, A., Steponaviciene, L., Linkeviciute-Ulinskiene, D., Stukas, R., Arlauskas, R., Vanseviciute-Petkeviciene, R., & Smailyte, G. (2021). Risk of Endometrial Cancer in Women with Diabetes: A Population-Based Retrospective Cohort Study. *Journal of Clinical Medicine*, 10(16), 3453. <https://doi.org/10.3390/jcm10163453>
 636. Zamani, M., Ebrahimitabar, F., Zamani, V., Miller, W. H., Alizadeh-Navaei, R., Shokri-Shirvani, J., & Derakhshan, M. H. (2018). Systematic review with meta-analysis: the worldwide prevalence of *Helicobacter pylori* infection. *Alimentary Pharmacology & Therapeutics*, 47(7), 868-876. <https://doi.org/10.1111/apt.14561>
 637. Zanders, M. M. J., Vissers, P. A. J., Haak, H. R., & van de Poll-Franse, L. v. (2014). Colorectal cancer, diabetes and survival: Epidemiological insights. *Diabetes & Metabolism*, 40(2), 120-127. <https://doi.org/10.1016/j.diabet.2013.12.007>
 638. Zapata-Diomedí, B., Barendregt, J. J., & Veerman, J. L. (2018). Population attributable fraction: names, types and issues with incorrect interpretation of relative risks. *British Journal of Sports Medicine*, 52(4), 212-213. <https://doi.org/10.1136/bjsports-2015-095531>
 639. Zhang, C., Zhang, L., Tian, Y. et al. Association between metabolic syndrome and early-stage colorectal cancer. *BMC Cancer* 23, 1020 (2023). <https://doi.org/10.1186/s12885-023-11537-3>
 640. Zhang, D., Li, N., Xi, Y., Zhao, Y., & Wang, T. (2017). Diabetes mellitus and risk of ovarian cancer. A systematic review and meta-analysis of 15 cohort

- studies. *Diabetes Research and Clinical Practice*, 130, 43-52.
<https://doi.org/10.1016/j.diabres.2017.04.005>
641. Zhang, F., Yang, Y., Skrip, L., Hu, D., Wang, Y., Wong, C., Qiu, J., & Lei, H. (2012). Diabetes mellitus and risk of prostate cancer: an updated meta-analysis based on 12 case-control and 25 cohort studies. *Acta Diabetologica*, 49(S1), 235-246. <https://doi.org/10.1007/s00592-012-0439-5>
 642. Zhang, Z.-H., Su, P.-Y., Hao, J.-H., & Sun, Y.-H. (2013). The Role of Preexisting Diabetes Mellitus on Incidence and Mortality of Endometrial Cancer. *International Journal of Gynecological Cancer*, 23(2), 294-303. <https://doi.org/10.1097/IGC.0b013e31827b8430>
 643. Zhao, L., Zheng, Z., & Huang, P. (2016). Diabetes mellitus and the risk of glioma: a meta-analysis. *Oncotarget*, 7(4), 4483-4489. <https://doi.org/10.18632/oncotarget.6605>
 644. Zhao, S., Gu, J., Tian, Y., Wang, R., & Li, W. (2024). Low levels of sex hormone-binding globulin predict an increased breast cancer risk and its underlying molecular mechanisms. *Open Life Sciences*, 19(1). <https://doi.org/10.1515/biol-2022-0822>
 645. Zhu, K. (2004). History of Diabetes Mellitus and Risk of Prostate Cancer in Physicians. *American Journal of Epidemiology*, 159(10), 978-982. <https://doi.org/10.1093/aje/kwh139>
 646. Zhu, Z., Wang, X., Shen, Z., Lu, Y., Zhong, S., & Xu, C. (2013). Risk of bladder cancer in patients with diabetes mellitus: an updated meta-analysis of 36 observational studies. *BMC Cancer*, 13(1), 310. <https://doi.org/10.1186/1471-2407-13-310>
 647. Zucchetto, A., Dal Maso, L., Tavani, A., Montella, M., Ramazzotti, V., Talamini, R., Canzonieri, V., Garboglio, A., Negri, E., Franceschi, S., & La Vecchia, C. (2007). History of treated hypertension and diabetes mellitus and risk of renal cell cancer. *Annals of Oncology*, 18(3), 596-600. <https://doi.org/10.1093/annonc/mdl438>