Cancer Clinical Trials:

Identifying & Testing Approaches

to Impact Evaluation

Catherine Ruth Hanna

BA (Hons), BMBCh, MRCP, FRCR

This thesis is submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

Institute of Cancer Sciences

College of Medical, Veterinary and Life Sciences

University of Glasgow

May 2021

"If you think research is expensive, try disease."

Mary Lasker, American philanthropist and advocate for cancer research, 1900-1994.

Abstract

Most cancer research is performed with the aim of generating new knowledge that leads to benefits such as improved treatments, higher cure rates, and better cancer prevention. Evaluating these downstream effects of research, often referred to as research impact, is of increasing importance to all cancer research stakeholders. There is currently no consensus surrounding the optimal way to approach this evaluation. The work in this thesis aimed to address this gap by first identifying which approaches to impact assessment have been applied previously for cancer research, and in particular for cancer clinical trials, and secondly to test a number of these approaches within the context of a case study of one cancer clinical trial. The Short Course Oncology Treatment (SCOT) trial was chosen for the purposes of the case study.

SCOT was a phase III randomised controlled trial (RCT) which tested the non-inferiority of shortening adjuvant treatment for patients with colorectal cancer (CRC) from the standard of 6 months to 3 months of doublet chemotherapy. The trial met its pre-specified non-inferiority end-point but showed unexpected differences in outcome based on the treatment regimen used and stage of disease. Specifically, for patients receiving CAPOX (capecitabine and oxaliplatin), non-inferiority for 3 months versus 6 months of treatment was met, but this was not the case for those treated with FOLFOX (5-fluorouracil and oxaliplatin). Similarly, non-inferiority was met for patients with small tumours with a small nodal burden (low-risk stage III), but not for those with more extensive disease and/or a higher nodal burden (high-risk stage III disease). SCOT was the largest contributor to a collaboration of six trials addressing the same research question, called the International Duration Evaluation of Adjuvant therapy (IDEA).

A systematic literature review was used to identify methods, frameworks, and categories of impact frequently used to perform research impact assessment (Chapter 3). This review was also used to identify previous impact assessments specific to cancer research. Fourteen empirical examples were identified, published between the years 1996 to 2015. These included assessment of

research at the cancer project, programme and research centre level. One example specifically assessed the impact of a phase III cancer RCT. The methods for impact analysis included across these examples included surveys, interviews, bibliometric searching of journals and clinical guidelines, economic approaches and documentary analysis. The categories of impact most commonly used were policy, clinical practice, health and economic impact. The Payback framework and the Canadian Academy of Health Sciences (CAHS) framework were utilised to collect data and communicate the results of impact assessment in two of these examples.

A second approach was adopted to identify ways that have previously been used to assess the impact of cancer clinical trials in particular (Chapter 4). The research impact case studies submitted to the United Kingdom government's research performance exercise for universities in 2014 were screened to find examples of assessments of the impact of cancer clinical trials. In total, 46 case studies describing 110 clinical trials were identified. Many of these trials were phase III trials that met their primary endpoint, but earlier phase trials and those with negative findings were also impactful. Policy impact was the most commonly described downstream effect. There was a gap within these case studies in the use of real world evidence to demonstrate the impact of cancer trials on clinical practices and health.

A number of the approaches to impact assessment identified in the literature review and in the analysis of the REF 2014 case studies were then tested to evaluate the impact of the SCOT Trial. The methods used for this assessment included surveys of clinician prescribing practices (Chapter 5), economic evaluation of the budget impact of trial results implementation (Chapter 6), and interrogation of real world data to explore the clinical practice and potential health benefits attributable to the SCOT trial at both a local (Chapter 7) and national (Chapter 8) level.

A clinician survey performed in April 2019 demonstrated a high level of awareness of SCOT trial results (Chapter 5), with 98% of those who were aware of the trial indicating they had changed their clinical practice based on the trial results. This impact on practice was driven mainly by shortening of treatment to 3 months for patients with low-risk stage III CRC (SCOT non-inferiority met),

whereas most clinicians reported they still used 6 months of doublet chemotherapy for patients with high-risk stage III disease (SCOT non-inferiority not met). This finding aligned with the post-hoc subgroup analysis performed for both the SCOT trial and IDEA collaboration. When shortening treatment for this subgroup of patients, clinicians mainly used CAPOX, whereas there was a more even split between using CAPOX and FOLFOX when 6 months of treatment was still used. A follow up survey in August 2020 was performed using a subset of respondents to the first survey and showed an increase in the use of shorter (3 months) treatment for patients with stage III disease with one high-risk feature, compared to responses in April 2019.

The results of the first clinician survey were applied within a budget impact analysis (Chapter 6) to estimate the economic impact of implementing SCOT trial results in the six countries that recruited to the trial. It was estimated that implementation of SCOT trial findings could translate to over $150 million USD savings over five years for those 6 healthcare systems (Australia, Denmark, New Zealand, Spain, Sweden, UK). Adopting a societal perspective by including money lost because patients did not work when receiving longer treatment, as well as travel costs to hospital, increased this impact to $340 million USD. Adding the monetised quality adjusted life-year (QALY) gains from implementation to this calculation ($456 million USD) meant that the gains from implementation of SCOT were vastly in excess of the original investment to conduct the SCOT trial ($8.8 million USD).

The final analysis conducted as part of the SCOT case study involved examination of individual patient level chemotherapy prescribing data. Using local (one health board in Scotland) level data, five different approaches were tested to evaluate the impact of the SCOT trial. In this instance, the change in practice was obvious even using simple descriptive statistics. Out of the other methods tested, interrupted time series analysis (ITSA) was the additional method that added the most value; the strengths of the ITSA were the ability to visualise the trends in prescribing pre and post-SCOT, as well as the counterfactual situation. Focusing on patients prescribed doublet chemotherapy (as per the SCOT trial), there was a significant decrease (85% to 31%) in the proportion of patients receiving over 3 months of treatment after the SCOT trial

results were published (χ^2 p<0.001) compared to before this time-point. There was no significant change in a comparator group of patients who received monotherapy (76% pre-SCOT versus 77% post-SCOT (χ^2 p=0.774)).

In order to evaluate this impact at a national level, it was first necessary to establish, for the first time, linkage of chemotherapy prescribing data at a pan-Scottish level. This process presented several challenges relating to data access, resource and infrastructure. Analysis of this data demonstrated a reduction in the proportion of patients receiving over 3 months of treatment across cancer networks in Scotland, although this change was less marked for patients treated in the Northern cancer network because 3 months of treatment was used proportionally more in the pre-SCOT period, compared to in the West or South-East of the country. The change in practice across the country was driven by changes for patients receiving CAPOX specifically, rather than FOLFOX or monotherapy, again fitting with the SCOT and IDEA subgroup analyses. Change was also greater for patients with low-risk rather than high-risk stage III disease, mirroring the clinician survey results. Following these in-depth analyses across Chapters 5-8, results from the survey, economic evaluation and administrative data interrogation were combined and summarised using a number of different impact frameworks that had been identified in Chapter 3, including the Payback framework.

This study has demonstrated how cancer research impact has been assessed in the past and has tested how impact analysis can be performed specifically for a cancer clinical trial. Evaluating the impact of the SCOT trial demonstrated its rapid and significant effects on new knowledge, future research, policy, clinical practice, and monetary savings for the health service. This assessment also allowed reflection on the pathway to these impacts occurring, as well as on how future trials could be designed to maximise impact. The study has highlighted challenges that currently exist to accessing real world data to investigate cancer trial impact. Further research to understand which impacts from clinical trials are meaningful to patients and trialists would be useful. More investment by funders and governments to support access to healthcare datasets that can be used to assess clinical practice change in response to trials would make impact assessments more straightforward in future.

Table of Contents

Abstract 3
Table of Contents 7
List of Tables 12
List of Figures 14
Grants 18
Publications and presentations 18
Acknowledgements 21
Contributions 23
Preface 25
Author's Declaration 26
Abbreviations 27
1 Chapter 1: Introduction 33
1.1 Rationale 33
1.2 Research impact evaluation 34
1.2.1 Definition of research 34
1.2.2 History of research impact evaluation 34
1.2.3 Rationale for conducting research impact evaluation 38
1.2.4 Definition of research impact 38
1.2.5 Definitions relevant to research impact evaluation 40
1.2.6 Overlap between implementation science and impact evaluation .. 41
1.2.7 Cancer research, clinical trials and research impact assessment ... 42
1.2.8 Summary 44
1.3 The SCOT trial 45
1.3.1 Background 45
1.3.2 Clinical trial evidence prior to SCOT 45
1.3.3 SCOT: the clinical question and trial design 49
1.3.4 SCOT trial results 50
1.3.5 IDEA Collaboration 56
1.3.6 Summary 62
1.4 Thesis aims and objectives 62
2 Chapter 2: Materials and methods 64
2.1 Overview of reviews 64
2.1.1 Description of approach used 64

2.1.2 Systematic review data extraction 65
2.2 Content Analysis 65
2.2.1 Overview 65
2.2.2 Coding manual development 66
2.3 Survey Design 69
2.3.1 Search for previous surveys 69
2.3.2 Survey development 69
2.3.3 Survey validity 70
2.3.4 Survey piloting 71
2.3.5 Development of follow up survey 71
2.3.6 Survey analysis 72
2.4 Cost-utility analysis 72
2.4.1 Cost-utility analysis from a multi-national perspective 72
2.4.2 Unit costs for multi-national cost-utility analysis 73
2.5 Budget impact analysis 76
2.6 Evaluation of chemotherapy prescribing data 76
2.6.1 Data access: Local 76
2.6.2 Data access: National 78
2.6.3 Description of the datasets used 88
2.7 Ethics and governance approvals 91
2.8 Software used 91
3 Chapter 3: Systematic review of the literature on approaches to research impact assessment 92
3.1 Introduction 92
3.2 Methods 93
3.2.1 Part I: Data sources and search strategy 94
3.2.2 Part II: Data sources and search strategy 96
3.2.3 Data extraction and analysis 97
3.3 Results 97
3.3.1 Search results 97
3.3.2 Part One: Identification and analysis of literature reviews describing approaches to research impact assessment 99
3.3.3 Part Two: Empirical examples of cancer research impact evaluation 122
3.4 Discussion 134
3.5 Conclusion 147
4 Chapter 4: Lessons on cancer trial impact assessment from the REF 2014 case studies 148
4.1 Introduction 148
4.2 Methods 150

4.2.1 Overview of approach....150
4.2.2 Identification of a relevant sample for the purposes of analysis...150
4.2.3 Content analysis....151
4.3 Results....152
4.3.1 Identification of impact case studies....152
4.3.2 Description of case studies....153
4.3.3 Description of cancer trials....154
4.3.4 Categories of cancer trial impact....157
4.3.5 Dissemination and knowledge transfer....159
4.3.6 Methods of evidencing cancer trial impact....160
4.3.7 Researchers and research users enhancing cancer trial impact....160
4.4 Discussion....160
4.5 Conclusion....171
Linking Part I and Part II....172
5 Chapter 5: Impact on clinician attitudes and self-reported practice....176
5.1 Introduction....176
5.2 Methods....179
5.3 Results....184
5.3.1 Survey response rate and respondent characteristics....184
5.3.2 Clinician awareness of clinical studies....186
5.3.3 Trial impact on current clinical practice....187
5.3.4 Influences and barriers to practice change....190
5.3.5 Timing of practice change....191
5.3.6 Change in practice over time....193
5.4 Discussion....200
5.5 Conclusion....207
6 Chapter 6: Assessing the economic value of implementation of the SCOT trial findings....209
6.1 Introduction....209
6.2 Methods....209
6.2.1 Cost-utility analysis....210
6.2.2 Budget impact analysis....215
6.2.3 Calculation of QALYs gained from implementation....222
6.2.4 The value of implementation of the SCOT trial....222
6.3 Results....222
6.3.1 Cost-effectiveness of 3 months versus 6 months of treatment....222
6.3.2 Analysis of budget impact....235
6.3.3 QALY gain from SCOT trial implementation....239
6.3.4 Estimating the return on the investment in the SCOT trial....241

6.4 Discussion 241
6.5 Conclusion 246
7 Chapter 7: Impact of the SCOT trial on local practice 248
7.1 Introduction 248
7.2 Methods 249
7.2.1 Cohort derivation 249
7.2.2 Statistical methods used to analyse chemotherapy datasets 252
7.3 Results 256
7.3.1 Cohort derivation 256
7.3.2 Descriptive analysis of cohort 258
7.3.3 Description of practice change 260
7.3.4 Regression analysis 264
7.3.5 Segmented regression 269
7.3.6 Interrupted time series analysis 272
7.3.7 Median dose delivered 278
7.4 Discussion 278
7.5 Conclusion 283
8 Chapter 8: Impact on national practice 284
8.1 Introduction 284
8.2 Methods 285
8.3 Results 286
8.3.1 Feasibility of using administrative healthcare data to define a national cohort 286
8.3.2 Description of overall cohort 288
8.3.3 Practice change cohort 289
8.3.4 A description of practice change 292
8.3.5 Logistic regression: overall cohort 297
8.3.6 Interrupted time-series analysis 302
8.4 Discussion 305
8.5 Conclusion 315
9 Chapter 9: Discussion 316
9.1 Summary of this research 316
9.2 Finding a framework for SCOT trial impact evaluation 317
9.3 Lessons learned for future trial design from the SCOT impact case study 321
9.4 Do the methods used in this thesis work for assessing clinical cancer trial impact? 323
9.4.1 Surveys 323
9.4.2 Health economic approaches 324
9.4.3 Administrative data analysis 325

9.4.4 Documentary analysis 325
9.4.5 Case studies 325
9.4.6 Methods not used in this thesis 326
9.5 Reflection on cancer trial impact evaluation from the perspective of different cancer trial stakeholders 327
9.5.1 Funders 327
9.5.2 Research institutions 328
9.5.3 The pharmaceutical industry 329
9.5.4 Cancer researchers 330
9.5.5 Clinical Trials Units and the wider trials community 330
9.5.6 Patients 331
9.5.7 Unanswered questions across stakeholder groups 332
9.6 Strengths and limitations 334
9.7 Future directions: Research questions arising and gaps identified from this work 336
9.8 Final reflections 339
Appendices 341
10 Appendix 1 CRC chemotherapy regimens 341
11 Appendix 2 Review Search Terms 342
11.1 Search Terms for systematic review 342
11.2 Search strategy Ovid Medline 342
12 Appendix 3 Supplementary survey material 343
13 Appendix 4 Supplementary material for economic evaluation 360
14 Appendix 5 Supplementary material relevant to analysis of administrative healthcare data 377
14.1 Comparison of methods of cohort derivation for the GG&C cohort 379
15 Appendix 6 Comparison of impact frameworks 412
15.1 Payback framework 412
15.2 Research utilisation ladder 416
15.3 Research Impact Framework 418
15.4 Montague and Valentim (modified CAHS) 421
15.5 Weiss's logic model (modified United way framework) 426
15.6 Healthcare research impact matrix 428
16 Appendix 7 Funding and information governance approvals 430
List of References 433

List of Tables

Table 1-1 Terms relevant to the research impact literature 41
Table 1-2 Three year DFS for overall SCOT trial study group and by regimen and stage subgroups..... 55
Table 1-3 Trials contributing data to IDEA Collaboration. 57
Table 2-1 Unit costs used for this analysis..... 75
Table 3-1 Literature reviews identified in overview of reviews.....100
Table 3-2 Methods for data collection and analysis for the purposes of research impact assessment 110
Table 3-3 Examples of frameworks for research impact assessment 117
Table 3-4 Empirical examples of cancer research impact assessment 123
Table 3-5 Recommendations for cancer research impact assessment.....144
Table 4-1 Case study description.....154
Table 4-2 Cancer clinical trial description.....155
Table 4-3 Impact categories identified in REF 2014 case studies 158
Table 4-4 Lessons learned from this study regarding cancer trial impact evaluation and how these apply to the SCOT trial case study.....167
Table 5-1 Location of survey respondents according to country of recruitment to IDEA collaboration trials.....185
Table 5-2 Proportion of clinicians choosing 3 months of doublet chemotherapy 190
Table 6-1 Average yearly wage.....221
Table 6-2 SCOT trial patient characteristics dependent on location 223
Table 6-3 Survival times for each trial arm.....226
Table 6-4 Regression output for UK utilities.....226
Table 6-5 Regression output for yearly costs from UK perspective 229
Table 6-6 Country specific cost-utility analysis results 231
Table 6-7 Budget impact analysis base case and scenario analysis in country specific currency 236
Table 6-8 Budget impact scenario analysis results 239
Table 6-9 QALY gain over 5 years from a country specific perspective and the monetised value of that QALY gain 240
Table 6-10 Calculation of economic value of investment in the SCOT trial 241
Table 7-1 Description of whole cohort receiving adjuvant chemotherapy in GGC 258
Table 7-2 Results of univariate and multivariate linear regression 265
Table 7-3 Change in treatment duration post-SCOT for patients prescribed doublet treatment dependent on regimen and risk stage.....266
Table 7-4 Change in treatment duration post-SCOT for patients prescribed monotherapy dependent on age and risk stage 266
Table 7-5 Results of univariate and multivariate logistic regression. 267
Table 7-6 Change in odds of receiving over 3 months of treatment post-SCOT for patients prescribed doublet treatment dependent on regimen and risk stage..268
Table 7-7 Change in odds of receiving over 3 months of treatment post-SCOT for patients prescribed monotherapy dependent on regimen and risk stage 268
Table 7-8 Comparison of outcomes from time series versus counterfactual models 277
Table 7-9 Median chemotherapy doses received 278
Table 8-1 High-risk disease features for patients with stage II disease 289
Table 8-2 Cohort characteristics 290

Table 8-3 Proportion of patients receiving 3 months or less of doublet chemotherapy post-SCOT 296
Table 8-4 Univariate and multivariate analysis of factors affecting the likelihood of patients receiving over 3 months of treatment.....298
Table 8-5 The proportion of patients receiving over 3 months of treatment pre versus post-SCOT dependent on patient, disease, and treatment related factors 299
Table 8-6 Subgroup analysis 300
Table 8-7 Challenges of building a linked, administrative healthcare dataset for CRC on a national scale and recommendations for repeating this process for other tumour types in future. 312
Table 9-1 Recommendations for cancer trial impact assessment 333
Table 10-1 Chemotherapy regimens used to treat CRC in the adjuvant setting 341
Table 11-1 Search terms Ovid Medline 342
Table 12-1 Clinical scenarios included in both surveys 343
Table 12-2 Locations of respondents to first survey 343
Table 12-3 Practice details for respondents to first survey 344
Table 12-4 Location and practice details for respondents of the follow- up survey 344
Table 12-5 Change in clinician (n=106) preferences in August 2020. 357
Table 13-1 Types of economic evaluation. 360
Table 13-2 Types of multi-country perspective economic evaluations 361
Table 13-3 Pros and cons of different types of multi-county economic evaluations 361
Table 13-4 Types of productivity loss calculation 362
Table 13-5 Time-point of data collection.....363
Table 13-6 Variables and values used to populate the budget impact framework 364
Table 13-7 Utility regression (Australia) 366
Table 13-8 Utility regression (Denmark) 367
Table 13-9 Utility regression (New Zealand) 367
Table 13-10 Utility regression (Spain) 368
Table 13-11 Utility regression (Sweden) 368
Table 13-12 Yearly cost regression (Australia) 369
Table 13-13 Yearly cost regression (Denmark) 369
Table 13-14 Yearly cost regression (New Zealand) 370
Table 13-15 Yearly cost regression (Spain) 371
Table 13-16 Yearly cost regression (Sweden) 371
Table 13-17 Budget impact analysis results in country-specific currency.....375
Table 14-1 Datasets request as part of PBPP application 1718-0026 377
Table 14-2 Variables used for GG&C analysis 382
Table 14-3 Demographic, disease and treatment related variables for patients receiving doublet chemotherapy at cycle one (GG&C analysis) 384
Table 14-4 Demographic, disease and treatment related variables for patients receiving single agent chemotherapy at cycle one (GG&C analysis).....385
Table 14-5 Results of Durbin-Watson tests for autocorrelation 391
Table 14-6 Variables used for national analysis.....393
Table 14-7 OS and CRC survival estimates.....406
Table 14-8 Durbin-Watson test statistics for national level ITSAs.....411
Table 15-1 Montague and Valentim's framework applied to the SCOT trial 423

List of Figures

Figure 1-1 Clinical trials investigating the optimal treatment for patients with CRC in the adjuvant setting 46
Figure 1-2 Kaplan-Meier curve showing 3 year DFS by study group 51
Figure 1-3 Forest plot showing DFS by stratification factors and timing of randomisation. 52
Figure 1-4 Kaplan-Meier curves showing DFS by study group for patients receiving A CAPOX and B FOLFOX 53
Figure 1-5 Kaplan-Meier curves showing DFS by study group for patients receiving A T1-3N1 disease and B T4N2 disease 54
Figure 1-6 Three year DFS for the IDEA stage III collaboration by regimen and disease risk stage 59
Figure 1-7 Five-year overall survival with 3 versus 6 months of adjuvant treatment from the IDEA collaboration 60
Figure 1-8 Subgroup analyses from the updated IDEA stage III collaboration 61
Figure 2-1 Access and linkage of GG&C datasets 78
Figure 2-2 Flowchart outlining the main stages to access and link data 80
Figure 2-3 Cohort definition and transfer of datasets to PHS 83
Figure 2-4 Datasets included in data release one and two 86
Figure 2-5 Timeline for transfer of datasets to PHS 87
Figure 2-6 Direct costs and resource use 88
Figure 3-1 PRISMA diagram for review 98
Figure 3-2 Categories of impact identified from literature reviews of approaches to research impact assessment. 107
Figure 3-3 The Payback Framework logic model 121
Figure 4-1 Different search strategies applied to REF2014 case studies 152
Figure 4-2 REF 2014 case studies search results 153
Figure 4-3 Most frequent journal of publication for cancer trials identified from REF 2014 case studies 157
Figure 4-4 Guidelines most frequently reference in REF 2014 case studies 159
Figure 5-1 Disease characteristics for stage II patient scenarios 181
Figure 5-2 Survey development and dissemination 183
Figure 5-3 Location of survey respondents 184
Figure 5-4 Studies identified by UK and international survey respondents 186
Figure 5-5 Self-reported practice change for scenarios describing patients aged under 70 years old 188
Figure 5-6 Self-reported practice change for scenarios describing patients aged 70 years old and over. 189
Figure 5-7 Mechanisms of dissemination of trial results that most influenced clinician practice change 191
Figure 5-8 Timing of practice change (stage III disease) 192
Figure 5-9 Timing of practice change (stage II) 192
Figure 5-10 Change in practice between April 2019 - August 2020 for patients aged under 70 years old 195
Figure 5-11 Treatment choices for patients aged under 70 with stage II disease 197
Figure 5-12 Use of doublet chemotherapy for patients with stage II disease ... 198

Figure 5-13 Clinician attitudes to using 3 months of treatment as standard of care for the adjuvant treatment of patients 199
Figure 6-1 Budget impact analysis analytical framework 217
Figure 6-2 Eligible population and extent of practice change 219
Figure 6-3 Kaplan-Meier curves 224
Figure 6-4 Survival curves partitioned by health state 225
Figure 6-5 Fully pooled results for quality of life calculated by applying country specific utilities to EQ-5D responses 228
Figure 6-6 Mean undiscounted costs per patient 230
Figure 6-7 INMB over a range of WTP thresholds for each country 232
Figure 6-8 Country specific CEACs and cost-effectiveness planes 235
Figure 6-9 Budget impact analysis one-way sensitivity analysis 238
Figure 6-10 Monetised QALY gain from implementation of SCOT findings 240
Figure 7-1 GG&C cohort definition using ChemoCare dataset 257
Figure 7-2 Scottish Index of Multiple Deprivation categories (deciles) for Glasgow City and surrounding areas 259
Figure 7-3 Histogram demonstrating the distribution of treatment duration (weeks) for GG&C cohort 260
Figure 7-4 Average treatment duration by year 261
Figure 7-5 Change in proportion of patients receiving over 3 months of treatment 262
Figure 7-6 Duration of treatment by year dependent on regimen prescribed 263
Figure 7-7 Change in proportion of patients who received over 3 months of treatment pre June 2017 versus post June 2017 264
Figure 7-8 Segmented linear regression (doublet) 269
Figure 7-9 Segmented linear regression (monotherapy) 270
Figure 7-10 Segmented logistic regression (doublet) 271
Figure 7-11 Segmented logistic regression (monotherapy) 271
Figure 7-12 Time series plot for patients receiving doublet chemotherapy (average monthly duration of treatment) 272
Figure 7-13 ITSA doublet (linear regression) 273
Figure 7-14 Time series plot for patients receiving monotherapy (average monthly duration of treatment) 274
Figure 7-15 ITSA monotherapy (linear regression) 275
Figure 7-16 ITSA doublet (logistic regression) 276
Figure 7-17 ITSA monotherapy (logistic regression) 276
Figure 8-1 National cohort derivation using national, linked datasets 288
Figure 8-2 Kaplan-Meier curve showing OS from date of first chemotherapy cycle 291
Figure 8-3 Kaplan-Meier curve showing cancer specific survival (CRC) from date of first chemotherapy 292
Figure 8-4 Change in proportion of patients receiving over 3 months of treatment pre- versus post-SCOT. 292
Figure 8-5 Median duration of treatment per year from 2013-2018 293
Figure 8-6 Treatment regimen and duration of chemotherapy delivered pre- versus post-SCOT 296
Figure 8-7 Adjusted odds of receiving over 3 months of treatment for patients starting adjuvant treatment after June 2017 302
Figure 8-8 Change in monthly proportion of patients receiving over 3 months of treatment (overall cohort) 303
Figure 8-9 Change in monthly proportion of patients receiving over 3 months of treatment (SCOT eligible) 304

Figure 8-10 Change in monthly proportion of patients receiving over 3 months of treatment (SCOT ineligible).....305
Figure 9-1 The Payback Framework modified to consider how the impact of the SCOT trial.....318
Figure 9-2 Potential pathways to impact for the SCOT trial 320
Figure 12-1 Regimen and duration of treatment chosen for six scenarios describing patients aged under 70 years old 346
Figure 12-2 Regimen and duration of treatment chosen for six scenarios describing patients aged 70 and over 348
Figure 12-3 Scenario 1: T3N1 disease. Details of regimens used by the clinicians (43% of total) who selected different treatment regimens for patients dependent on age 348
Figure 12-4 Scenario 2: T4N1 disease. Details of regimens used by the clinicians (49% of total) who selected different treatment regimens for patients dependent on age. 349
Figure 12-5 Scenario 3: T4N1 disease. Details of regimens used by the clinicians (44% of total) who selected different treatment regimens for patients dependent on age. 349
Figure 12-6 Scenario 4: T4N2 disease. Details of regimens used by the clinicians (45% of total) who selected different treatment regimens for patients dependent on age. 350
Figure 12-7 Scenario 5: T3-4N0 MSI Low disease. Details of regimens used by the clinicians (38% of total) who selected different treatment regimens for patients dependent on age. 351
Figure 12-8 Scenario 6: T3-4N0 MSI high disease. Details of regimens used by the clinicians (28% of total) who selected different treatment regimens for patients dependent on age. 352
Figure 12-9 Switch in individual treatment choices for scenarios describing patients with stage III disease aged 70 and over 354
Figure 12-10 Treatment choices for patients aged 70 and over with stage II disease with a comparison of clinician choices between April 2019 and August 2020 (n=106) 355
Figure 12-11 Change in individual clinician opinions between April 2019-August 2020 for the group of clinicians who answered both surveys (n=106). 356
Figure 12-12 Clinicians' estimation of the proportion of patients of working age who continue to work full time during their adjuvant treatment for CRC.....360
Figure 13-1 An explanation of types of unit costs. 362
Figure 13-2 Subgroup analyses CEACs and cost-effectiveness planes.....374
Figure 14-1 Comparison of GG&C cohort definition using three approaches 381
Figure 14-2 Treatment switch 386
Figure 14-3 Time series to check for seasonality (logistic regression for GG&C doublet chemotherapy analysis) 387
Figure 14-4 Time series to check for seasonality (logistic regression for GG&C single agent chemotherapy analysis) 387
Figure 14-5 Plot of residuals (linear regression doublet chemotherapy).....388
Figure 14-6 Plot of partial residuals (linear regression doublet chemotherapy) 388
Figure 14-7 Plot of residuals (linear regression monotherapy) 389
Figure 14-8 Plot of partial residuals (linear regression monotherapy) 389
Figure 14-9 Plot of residuals (logistic regression doublet chemotherapy) 390
Figure 14-10 Plot of partial residuals (logistic regression doublet chemotherapy) 390
Figure 14-11 Plot of residuals (logistic regression monotherapy) 391

Figure 14-12 Plot of partial residuals (logistic regression monotherapy) 391
Figure 14-13 ITSA with interruption at January 2013 392
Figure 14-14 ITSA with interruption at June 2015 392
Figure 14-15 OS Kaplan-Meier curve by risk stage 398
Figure 14-16 CRC Kaplan-Meier curve by risk stage 398
Figure 14-17 OS Kaplan-Meier curve by SCOT eligibility 399
Figure 14-18 CRC Kaplan-Meier curve by SCOT eligibility 399
Figure 14-19 OS Kaplan-Meier curve by regimen.....400
Figure 14-20 CRC Kaplan-Meier curve by regimen 400
Figure 14-21 OS Kaplan-Meier curve by age group 401
Figure 14-22 CRC Kaplan-Meier curve by age group.....401
Figure 14-23 OS Kaplan-Meier curve by disease site 402
Figure 14-24 CRC Kaplan-Meier curve by disease site.....402
Figure 14-25 OS Kaplan-Meier curve by SIMD.....403
Figure 14-26 CRC Kaplan-Meier curve by SIMD 403
Figure 14-27 OS Kaplan-Meier curve by Charlson group 404
Figure 14-28 CRC Kaplan-Meier curve by Charlson group.....404
Figure 14-29 OS Kaplan-Meier curve by location.....405
Figure 14-30 CRC Kaplan-Meier curve by location 405
Figure 14-31 OS Kaplan-Meier curve by sex 406
Figure 14-32 CRC Kaplan-Meier curve by sex.....406
Figure 14-33 Time series for national cohort to check for seasonality 408
Figure 14-34 Plot of residuals for time series (national overall cohort).....409
Figure 14-35 Plot of partial residuals for time series (national overall cohort) .409
Figure 14-36 Plot of residuals for time series (national SCOT eligible cohort) ..410
Figure 14-37 Plot of partial residuals for time series (national SCOT eligible cohort) 410
Figure 14-38 Plot of residuals for time series (national SCOT ineligible cohort) 411
Figure 14-39 Plot of partial residuals for time series (national SCOT ineligible cohort) 411
Figure 15-1 Payback framework categories of impact. 412
Figure 15-2 Landry's ladder of knowledge utilisation 417
Figure 15-3 Five main impact areas included in the Research Impact Framework 419
Figure 15-4 Framework for research impact assessment. 423
Figure 15-5 Montague and Valentim's framework applied to the SCOT trial.....425
Figure 15-6 Anthony Weiss's logic model.....427
Figure 15-7 Cruz-Rivera's impact matrix 429

Grants

The work in this thesis was performed during a three year Cancer Research United Kingdom (CRUK) funded Clinical Trials Fellowship which commenced on 8th January 2018 (Grant ID: C61974/A2429).

Specific funding to access national level administrative healthcare data was granted by the Beatson Cancer Charity (Ref 17-18-139).

Publications and presentations

Publications

Hanna, C. R. , Boyd, K. A. and Jones, R. J. (2020) Evaluating cancer research impact: lessons and examples from existing reviews on approaches to research impact assessment. *Health Research Policy and Systems*, **19,** 36 (2021). https://doi.org/10.1186/s12961-020-00658-x

Hanna, C. R. , Gatting, L. P., Boyd, K. A. , Robb, K. A. and Jones, R. J. (2020) Evidencing the impact of cancer trials: insights from the 2014 UK Research Excellence Framework. *Trials*, 21, 486. (doi: 10.1186/s13063-020-04425-9) (PMID:32503612) (PMCID:PMC7275320)

Hanna, C. , Lemmon, E., Ennis, H., Jones, R. , Hay, J., Halliday, R., Clark, S., Morris, E. and Hall, P. (2021) Creation of the first national linked colorectal cancer dataset in Scotland: prospects for future research and a reflection on lessons learned : Creation of Scottish colorectal cancer dataset for research purposes. *International Journal of Population Data Science*, 6(1), 10. (doi: 10.23889/ijpds.v6i1.1654)

Hanna, C.R., Robles-Zurita, J.A., Briggs, A., Harkin, A., Kelly, C., McQueen, J., Allan, K., Pearson, S., Hollander, H., Glimelius, B., Salazar, R., Segelov, E., Saunders, M., Iveson, T., Jones, R.J., Boyd, K.A. Three versus six months of adjuvant doublet chemotherapy for patients with colorectal cancer: a multi-country cost-effectiveness and budget impact analysis. *Clinical Colorectal Cancer*. April 2021. https://doi.org/10.1016/j.clcc.2021.04.001

Hanna, C.R., Boyd, K.A., Wincenciak, J., Graham, J., Iveson, T., Jones, R.J., Wilson, R. *Cancer Treatment and Research Communications.* August 2021. https://doi.org/10.1016/j.ctarc.2021.100445

Abstracts

Self-reported prescribing practices in the setting of adjuvant treatment for colorectal cancer. Hanna CR, Boyd KA, Jones RJ. Poster presentation. European Society for Medical Oncology Gastrointestinal (ESMO GI) conference 2020.

Real-life implementation of the SCOT trial findings: An interrupted time series analysis. Hanna C, Robles-Zurita JA, McLoone P, Kelly C, Boyd KA, Jones RJ. Poster presentation. American Society of Clinical Oncology (ASCO) conference 2020.

Three versus six months of adjuvant chemotherapy for colorectal cancer: A multi-country cost-effectiveness and budget impact analysis. Hanna C, Robles-Zurita JA, Jones RJ, Boyd KA. Poster presentation. ASCO conference 2020.

The implications of changing clinical practice in colorectal cancer: multi-national budget impact assessment. Hanna C, Robles-Zurita JA, Jones RJ, Boyd KA. European Health Economics Association (EuHEA) conference 2020. Accepted for oral presentation but due to COVID-19 the conference was held virtually and all oral presentations were poster presentations.

Do clinical trials have an impact on real-life practice? An initial review of adjuvant chemotherapy prescribing practices in colorectal cancer pre-Short Course Oncology Trial (SCOT) publication. Hanna C, Worrall C, Jones RJ, Boyd KA, Paul J. Poster presentation. National Cancer Research Institute (NCRI) conference 2018.

Oral Presentations

The impact of cancer trials. Oral presentation. Scottish Cancer Clinical Trials Units (CaCTUS) conference, Glasgow 22nd March 2019.

The value of audit data to better understand the real life impact of cancer trials. 13th June 2019, Perth. Scottish National Cancer Audit Meeting.

The impact of the SCOT trial. Oral presentation. Glasgow CRUK Clinical Trials Unit Scientific Advisory Board, Glasgow, 13th January 2020.

The impact of the SCOT trial. Beatson West of Scotland Radiotherapy Research Group presentation, Glasgow. 4th February 2020.

Creation of the first national linked colorectal cancer dataset in Scotland: prospects for future research and a reflection on lessons learned. Webinar for International Journal of Population Data Science (virtual). 29th April 2021. (Oral presentation with Dr Elizabeth Lemmon, Dr Peter Hall and Mr Steve Clark). Recording available on request.

Cancer Clinical Trials: Identifying and testing approaches to impact evaluation. Beatson West of Scotland Education Meeting. 30th April 2021.

Cancer Trial Impact: Identifying and testing approaches to evaluation. Invited webinar CRUK Clinical Trials Units (virtual, 25th May 2021)

The impact of the SCOT trial. Invited talk at the Scottish Radiotherapy Research Forum (SCORFF) 2021 (virtual, 3rd June 2021)

Acknowledgements

I am greatly indebted to my doctoral supervisors Dr Kathleen A. Boyd and Professor Robert J Jones, for their encouragement, advice, and support throughout this period of study. Thank you for guiding me through the original CRUK Trials Fellowship application in 2017, and all of your help over the past 3.5 years.

Thank you to Dr Jose A Robles-Zurita for your supervision, advice and support, in particular for your patience and kindness in helping me understand the trickier components of health economics analyses and STATA code.

I would like to thank everyone at the CRUK Clinical Trials Unit (CTU) in Glasgow. In particular, thank you to Mrs Caroline Kelly and Mr Jim Paul for their time, advice, and input. Special thanks to Andrea Harkin, in particular for helping with the original grant application, for finding me a place to work, and for making me feel part of the CTU team.

I would like to thank everyone at the Health Economics and Health Technology Assessment team at the University of Glasgow. It was great to be a part of such a fun and dynamic group.

Thank you also to Mr Philip McLoone for his guidance on statistical analysis of real world datasets and to Christine Crearie who was the data analyst in Glasgow who helped me to access the Greater, Glasgow and Clyde chemotherapy prescribing data.

I would like to thank Professor Steven Hanney for his kindness and enthusiastic support. A chance meeting at an impact seminar in 2018 meant that I had a wonderful mentor when conducting this research.

Thank you to Dr Katie Robb and Lauren Gatting from the University of Glasgow. It was brilliant to work with you. Thank you for being so open to collaboration. Thank you also to Professor Tricia Greenhalgh who was happy to discuss her work with me at a conference and who gave me confidence to proceed with the work outlined in Chapter 4.

Thank you to Professor Nicol Keith, Professor Anthony Chalmers and Professor Richard Wilson for your advice and support over the past 3.5 years.

I am extremely grateful to Dr Peter Hall and Dr Elizabeth Lemmon for their support, guidance, and friendship. Thank you for welcoming me into your excellent team and for providing a wonderful opportunity for collaboration.

I would like to acknowledge the grant funders, CRUK and the Beatson Cancer Charity, who funded my fellowship and assisted with funding access to national level data. Thank you to Short Course Oncology Trial (SCOT) management group for their support and to all the patients who participated in the SCOT trial.

Finally, I would like to thank my family: my parents Aidan and Noelle, my siblings Rory, Timothy, and Fiona, my sister-in-law Catherine, and my nieces Holly and Lily, for listening to me talk about this work and showing an interest, feigned or otherwise! Finally, special thanks to my husband Liam, for your encouragement, your patience, for bringing me endless cups of tea, and maintaining the belief that I could get there in the end.

Contributions

Chapter 3: The search strategy was developed in collaboration with Lorraine MacLeod, Librarian at the Beatson West of Scotland Cancer Centre. The screening of titles for the purposes of the review was carried out by both myself and Lorraine MacLeod. The results of this study have been published in the Journal of Health Research Policy and Systems. Professor Steven Hanney (Brunel University) gave comments on an earlier draft of this manuscript prior to publication.

Chapter 4: The concept for this study was based on work by Professor Tricia Greenhalgh, who provided advice at the outset of the project. The coding manual for the analysis was developed jointly by Lauren Gatting (Doctor of Philosophy (PhD) student at the Institute of Health and Wellbeing, University of Glasgow), (supervised by Dr Katie Robb) and I. The analysis was undertaken by myself, with coding of a small number of case studies by Lauren Gatting to assist in verifying the validity of the analysis.

Chapter 5: The survey for this study was developed by myself in three piloting stages, which required input from colleagues who carried out the pilot.

Chapter 6: The cost-utility analysis (CUA) as part of this study had already been undertaken from a UK perspective by Dr Jose A. Robles-Zurita and Dr Kathleen A. Boyd. During my studies, I wanted to gain a better understanding of CUA and to use the most up to date outcome data and unit costs available for a budget impact assessment. For this reason, I repeated the within trial CUA, in a large part using the previous statistical code and the same methodology as had been employed previously. The main changes I made were to conduct the analysis from the perspectives of each country that participated in the Short Course Oncology Treatment (SCOT) trial, to use updated overall survival data, perform alternative subgroup analyses, calculate quality of life using country specific Euroqol 5-dimension (EQ-5D) weights, and to update the model used for costs so that it extended to ten years (in line with updated survival data). The within trial data used in this analysis was collected for the purposes of the SCOT trial, the initial results of which were published in April 2018.

Chapter 7: Christine Crearie (ChemoCare data analyst West of Scotland) and Finlay MacKay (data analyst QPI Greater, Glasgow and Clyde) both provided the data used in this study. I specified a number of different cohort derivation rules and Christine Crearie implemented a number of data extractions based on these different approaches. Mr Philip McLoone, Mrs Caroline Kelly, and Mr Jim Paul provided statistical advice at the outset of the study.

Chapter 8: The team at the electronic Data Research and Innovation Service (eDRIS), in particular Joy Hay, were instrumental in helping to obtain access to data for the purposes of this analysis. Dr Elizabeth Lemmon helped with data cleaning and discussions around cohort analysis, and both Dr Elizabeth Lemmon and Dr Peter Hall provided analytical advice.

Preface

The work for this thesis was ceased for four months whilst I returned to full time clinical work to support the National Health Service efforts against the COVID-19 pandemic (April-July 2020). A no-cost extension was applied to my CRUK grant to allow completion of this work over a full three-year period (end-date June 2021).

Author's Declaration

Chapter 1: This chapter was produced for this thesis and has not appeared anywhere else.

Chapter 2: Parts of this chapter relating to data access have been published (Hanna et al 2021). I undertook writing and drafting of all versions of that manuscript and contributions by co-authors were in the form of comments on drafts.

Chapter 3: A version of this chapter has been published (Hanna et al. 2020). I undertook writing and drafting of all versions of the manuscript, contributions by collaborating authors were in the form of comments on drafts.

Chapter 4: A version of this chapter has been published (Hanna et al. 2020). I undertook writing and drafting of all versions of the manuscript, contributions by collaborating authors were in the form of comments on drafts.

Chapter 5: This chapter was produced for this thesis and has not appeared anywhere else.

Chapter 6: A version of this chapter has been published (Hanna et al. 2021). I undertook writing and drafting of all versions of the manuscript, contributions by collaborating authors were in the form of comments on drafts.

Chapter 7: This chapter was produced for this thesis and has not appeared anywhere else.

Chapter 8: This chapter was produced for this thesis and has not appeared anywhere else.

Chapter 9: This chapter was produced for this thesis and has not appeared anywhere else.

I declare that I am the sole author of this thesis. The work presented here is my own, unless otherwise acknowledged. This thesis has not been submitted for consideration of another degree in this or any other university.

Abbreviations

5-FU 5-Fluorouracil

A&E Accident and Emergency

AcaDME Acute, Cancer, Deaths and Mental health

ACCENT Adjuvant Colon Cancer End Points

ACHIEVE Adjuvant Chemotherapy for Colon Cancer Oncology Group

AJCC American Joint Committee on Cancer

ASCO American Society of Clinical Oncology

BCG Bacille Clamette-Guerin

BIA Budget impact analysis

BNF British National Formulary

CAHS Canadian Academy of Health Sciences

CALGB/SWOG Cancer and Leukemia Group B/South-West Research Group

CAPOX Capecitabine and oxaliplatin

CC Christine Crearie

CEAC Cost-effectiveness acceptability curve

CH Catherine Hanna

CHEERS Consolidated Health Economic Evaluation Reporting Standards

CHI Community Health Index

CHILIS CHI Indexing and Linkage Service

CHOICE Choosing interventions that are cost-effective

CI Confidence interval

CMOP Cancer Medicines Outcomes Project

CORECT-R Colorectal Repository

COVID-19 Coronavirus-19

CRC Colorectal Cancer

CRUK Cancer Research UK

CUA Cost utility analysis

DESCRIBE Definitions, Evidence and Structures to Capture Research Impacts and Benefits

DFS Disease free survival

DRG Disease Research Group

eDRIS Electronic data research and innovation service

eMIT Electronic market information tool

EudraCT European Clinical Trials Database

EPCC Edinburgh Parallel Computing Centre

EQ-5D EuroQol 5 dimensions

EQ-5D-3L EuroQol 5 dimensions 3 levels

ERA Excellence in Research Australia

FACT/GOG-Ntx-4 Functional Assessment of Cancer Therapy/Gynecologic Oncology Group - Neurotoxicity

FLOX 5-fluorouracil and oxaliplatin

FOLFOX 5-fluorouracil and oxaliplatin

GDP Gross Domestic Product

GDPR General Data Protection Regulation

GERCOR Groupe Coopérateur Multdisciplinaire en Oncologie

GG&C Greater Glasgow and Clyde

GP General Practitioner

HDU High Dependency Unit

HR Hazard ratio

HTA Health Technology Assessment

ICC Intra-class correlation

ICU Intensive Care Unit

IDEA International Duration Evaluation of Adjuvant Therapy

IMPACT Multicenter Pooled Analysis of Colon Cancer Trials

INMB Incremental net monetary benefit

INT Intergroup

IRR Internal Rate of Return

ISD Information Services Division

ISRIA International School of Research Impact Assessment

ITSA Interrupted time series analysis

LEV Levasimole

LG Lauren Gatting

LV Leucovorin

MOF 5-fluorouracil, vincristine and semustine

MORIA Measurement of Research Impact and Assessment

MOSAIC Multicenter international study of oxaliplatin/5-fluorouracil-LV in the adjuvant treatment of colon cancer

MRC Medical Research Council

NCI National Cancer Institute

NCT National Clinical Trial

NCRI National Cancer Research Institute

NHS National Health Service

NICE National Institute for Health and Care Excellence

NoSCAN North of Scotland Cancer Network

NRS National Records Scotland

NSABP National Surgical Adjuvant Breast and Bowel Project

NSH National Safe Haven

NSS National Services Scotland

OS Overall survival

PBPP Public Benefit and Privacy Panel for Health and Social Care

PETACC Pan-European trials in Adjuvant Colon Cancer

PhD Doctor of Philosophy

PHS Public Health Scotland

PIS Prescribing Information System

PLICS Patient Level Costing System

PRISMA Preferred reporting items for systematic reviews and meta-analysis

QALY Quality Adjuvant Life Year

QPI Quality Performance Indicator

QUASAR QUick And Simple and Reliable

RAE Research Assessment Exercise

RCUK Research Councils UK

RDS Research Data Scotland

REF Research Excellence Framework

RFS Recurrence free survival

RQF Research Quality Framework

SACT Systemic Anti-Cancer Therapy

SAFFA Short Adjuvant Fluorouracil and Folinic Acid

SCAN South East Scotland Cancer Network

SCOT Short Course Oncology Treatment

SCRIS Scottish Cancer Registry and Intelligence Service

SEER Surveillance, Epidemiology and End Results

SICSAG Scottish Intensive Care Society Audit Group

SIMD Scottish Index of Multiple Deprivation

SMR Scottish Morbidity Records

TNM Tumour, Nodal and Metastases

TOSCA Three or Six Adjuvant Colon

ToT Time on treatment

UK United Kingdom

UMIN University Medical Information Network

US United States

USA United States of America

WHO World Health Organisation

WoSCAN West of Scotland Cancer Network

WTP Willingness to pay

X-ACT Xeloda in Adjuvant Colon Cancer Therapy

1 Chapter 1: Introduction

1.1 Rationale

Research to improve outcomes for patients with cancer and to prevent cancer occurring is a priority. Cancer is highly prevalent, with up to one in two people likely to have a diagnosis of malignancy in their lifetime (1, 2). Despite current treatments, mortality and morbidity remain high (1). Cancer research attracts substantial public and charity sector funding. The National Cancer Institute (NCI) in the United States of America (USA) had a 2020 budget of over $6 billion United States dollars (USD) for cancer research and the same organisation spends over $800 million each year specifically on developing and performing cancer trials (3). As well as public and charity funds, there is also huge financial investment into cancer research, and into cancer trials in particular, from private pharmaceutical companies, and the altruistic investment of time and effort from patients who participate in these clinical trials. In the United Kingdom (UK) in 2018, over 25,000 patients were recruited to trials run by one charity (Cancer Research UK (CRUK)) alone (4). Many of these patients will experience adverse effects from novel treatments tested within trials, without gaining personal benefits in terms of improved outcomes (5).

We live in an era of increasing austerity. In particular, the effect of the coronavirus-19 (COVID-19) pandemic on healthcare systems and economies worldwide means that now, more than ever, there is scrutiny around how public investments will benefit society; this scrutiny applies equally to investment in research. It is reasonable to question if research is leading to the real world benefits expected and to query the opportunity cost of not spending the same money directly within other public sectors such as health and social care, the environment, or education (6-8). These wider, real-life benefits from research on society are often referred to as ‘research impact’.

In the UK, the concept of assessing the impact from research has been put under the spotlight within the academic community because of a government funding allocation process known as the Research Excellence Framework (REF)(9). The REF is an assessment exercise used by the UK government to allocate core funding to

universities. The demonstration of wider impacts from research the universities perform is a substantial component of that assessment.

The impact from cancer research and cancer clinical trials in particular, will be important to patients who participate in these trials and the charities and companies who invest in them. It can be assumed that charities will want to understand why investment in research, rather than other endeavours such as direct care for the populations they support, is worthwhile. Even pharmaceutical companies will have a limited budget and will be mindful of the value of investing in one research endeavour over another. In light of the REF, cancer trial impact will also be of interest to universities who sponsor and support cancer clinical trials. In order to understand the impact of any research, it is first necessary to know how to evaluate this impact. It is only with a clearer understanding of approaches to assessment that the impact of cancer research and cancer trials can be maximised.

1.2 Research impact evaluation

1.2.1 Definition of research

Research can be defined as 'systematic investigation or inquiry aimed at contributing to knowledge of a theory or topic by careful consideration, observation or study of a subject.' (Oxford Dictionary (10))

Clinical trials performed within the field of oncology are a type of research and for the purposes of this thesis are referred to as cancer trials. Cancer trials are clinical trials using the recruitment of human subjects for the purposes of investigating the diagnosis or screening of malignancy, or treatment and management of individuals with a diagnosis of malignancy.

1.2.2 History of research impact evaluation

The interest in evaluating research impact has been rising, partly driven by the actions of national bodies and governments. Traditionally, assessments of research output and quality were directed at measuring scholarly and academic influence, and the ability to measure this aspect of research output was helped

by the rise of bibliometrics in the 1960s and the discovery of the internet (11). In contrast to the concept of assessing scholarly merit alone, the term research impact, referring to the wider impact of research on society, emerged in the 1990s and 2000s (12).

The academic community have been conducting research to explore which approaches are suitable for evaluating and communicating research impact outside government driven assessment exercises. Specifically, there has been a rise in academic literature investigating research impact assessment over the past decade. The results of the Definitions, Evidence and Structures to Capture Research Impacts and Benefits (DESCRIBE) (12) study identified that the research impact literature went through a rapid phase of development from 2010 onwards and entered the end of an initial phase of understanding the concepts around research impact assessment around 2013. This coincided with the formation of the International School of Research Impact Assessment (ISRIA), set up by a small network of researchers located in the UK, Canada and Spain. The purpose of ISRIA was to teach individuals to assess research investments and activities, with the aim of optimising returns from research. The organisation achieved this aim by holding a series of international workshops from 2013-2017 and published a set of guidelines in 2018 to summarise the findings from these workshops (13). Even in 2018, the authors of this report recognised that the practice of research impact assessment was still in its formative stages, without any accepted standards or recommendations around how research impact assessment should be performed.

Although the research evaluation community have been assessing research impact since the 1990s, it was not until the 2000s that the interest in assessing the wider impact of research on society arose, driven by national governments and research institutions. At that time, the Australian Government developed a Research Quality Framework (RQF) to assess the quality of research performed within academic institutions in Australia (14). This framework assessed both academic outputs and broader impact, defined as ‘the recognition by qualified end-users that quality research has been successfully applied.’ The plans for the RQF specified that research impact would be evaluated via impact statements generated by universities, which would be assessed by a panel of experts.

Although the RFQ was scheduled to be implemented in 2009, this was abandoned by a change in government in 2007.

Although not implemented, the development of the RQF signalled a shift in focus by national institutions globally, with an increased emphasis on the returns to society from the research they were funding and performing. In the UK, the Research Assessment Exercise (RAE) had been used since 1986 to measure the quality of research performed by higher education institutions. The RAE focused on traditional measures of academic merit such as publications and the corresponding impact factors of the journals of publication, the university environment, and research infrastructure. In 2008, the UK government piloted an assessment of research impact, with the definition and the methods for evaluation explicitly influenced by the Australian RQF, as well as by work performed by UK academics (Brunel University) who were exploring ways to evaluate impact using a case-study approach (15). Formal adoption of impact assessment into the UK national evaluation in 2014 led to the establishment of what is now known as the REF (16).

In the REF 2014, impact was evaluated using peer-reviewed case studies under four main disciplines of research. The first (panel A) included clinical medicine and biological sciences, the second (panel B) included physical and earth sciences, the third (panel C) included economics, law and politics, and the last (panel D) included the arts, languages, and music. The impact case studies generated by universities were read and scored from 1-4 by expert reviewers. The higher the score allocated to each impact case study, the more research funding that the higher education institution submitting the case study received from the government. Although the individual scores given to each impact case study have not been published, several institutions have analysed their own submissions to better understand the indicators of impact and approaches to impact evaluation that led to their case studies achieving higher scores (17-19). In 2014, impact case studies were worth 20% of £2 billion annual funding allocated to UK universities. The assessment is being repeated in 2021, with a higher proportion of funds (25%) being given to universities that successfully demonstrated the impact of their research.

Although their 2007 government rejected impact assessment, the Australian government recently performed its inaugural engagement and impact assessment (EI) (20). The main rationale behind the EI was to serve as a mechanism to show the government and the public how investment in university research in Australia is being translated into tangible benefits beyond academia, such as economic, environment, social, and cultural impacts. A major similarity between the EI and the REF is the use of impact narrative case studies that are scored by a panel of experts. In the EI, a score is allocated to the case study not only for the impact evidenced, but also for the efforts made by the university to translate their research to impact. Other countries have also made efforts to evaluate impact at a national level. For example, in Canada, the Canadian Academy of Health Sciences (CAHS) framework was developed in 2009 (21) by an expert panel and sponsored by multiple Canadian and international research foundations with the intention that it would be used by all funders of health research in Canada for evaluation of their health research impacts. More recently, a white paper from the Canadian Health Services and Policy Research Alliance (22) provided updated guidance on how research impact should be approached.

Research impact is also important in other sectors. For example, private pharmaceutical companies prioritise the intellectual property rights associated with patent policy, market exclusivity to new medications, stakeholder investment, and drug sales as key drivers of their work and investment. As an example, the first line of the annual report from a well-known pharmaceutical company (AstraZeneca) in 2020 mentioned “double-digit revenue growth to leverage improved profitability and cash generation” (23). Despite this focus on monetary gains, the mission statement of the same pharmaceutical company states that their aim is to “create a meaningful difference in the lives of patients”. It is therefore clear that wider, longer terms impacts on health are important to the pharmaceutical industry. It is not yet clear how AstraZeneca, or other pharmaceutical companies, are evaluating their goal of improving health, or what other impacts they are prioritising.

Third sector institutions such as charities are also facing increasing expectations to demonstrate the impact of their work and there are now several companies offering guidance and services to help charities with this endeavour (24, 25).

Despite this guidance, it is again not clear from looking at public facing information from charities, how or when this impact evaluation is being performed. A recent review of the academic literature on this topic identified only thirteen published examples of impact assessment of charity and public research investment (26). If these evaluations are not being communicated to either the public or the academic community, this raises the question as to whether they are being undertaken. Improved transparency in this regard would provide greater insight into how the third sector are approaching the topic of research impact evaluation.

1.2.3 Rationale for conducting research impact evaluation

Undertaking an assessment of research impact requires time and investment, therefore it is useful to consider why conducting such an assessment may be worthwhile. The purpose of any research impact assessment will also dictate the type of evaluation methodology that is used. The rationale for research impact assessment has commonly been divided into four main categories, known collectively as the "4 As" (27). These are: advocacy, accountability, analysis and allocation.

First, research funders and providers may want to use research impact evaluation to *advocate* for ongoing or increased investment for their work. Secondly, researchers and institutions may want to show *accountability* to their stakeholders by demonstrating that research conducted using funder investment has been worthwhile and aligns with the mission statement for that investor. Another reason to perform impact evaluation is driven by a desire to *analyse* if, why and how, impact occurs. Finally, research impact assessment can be used to guide prospective *allocation* of funding for future research endeavours. Of course, these reasons for conducting impact assessment are not mutually exclusive and impact evaluation could be performed for more than one reason.

1.2.4 Definition of research impact

Although research impact is a term frequently used in the literature, to date there is no consensus around what this encompasses, and several definitions exist. Penfield et al (28) have identified that different definitions may be relevant

depending on the purpose of evaluating research impact, the target audience, and users of the research in question. In their review of the literature, Alla et al (29) found 108 definitions of research impact, most of which (76%) had been constructed by research institutions and funders. These definitions included those based on bibliometric evaluation and use-based definitions.

Bibliometric based definitions identified by Alla et al (29) focused on quantifiable effects from research, usually related to how the research is cited in the academic literature. These definitions are often considered to reflect a narrow interpretation of research impact that aligns with a traditional concept of scholarly influence. The research governance definitions of research impact identified by Alla and colleagues mainly focused on the contribution or benefit of research to society and the economy. Examples included definitions from the Research Councils UK (RCUK) (30), and the UK REF assessment (31).

The RCRUK (30) defines research impact as:

> **"Academic impact**
>
> The demonstrable contribution that excellent research makes to academic advances, across and within disciplines, including significant advances in understanding, methods, theory, and application.
>
> **Economic and societal impacts**
>
> The demonstrable contribution that excellent research makes to society and the economy. Economic and societal impacts embrace all the extremely diverse ways in which research-related knowledge and skills benefit individuals, organisations, and nations by:
>
> fostering global economic performance, and specifically the economic competitiveness of the United Kingdom,
>
> increasing the effectiveness of public services and policy,
>
> enhancing quality of life, health and creative output."

The definition used for the purposes of the REF (31) is:

> "an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia".

An important distinction between these two definitions is that the former includes academic influence from research whereas the REF definition specifically excludes academic impact and only includes wider, downstream changes. Secondly, the RCUK definition focuses on the contribution of research to impacts on society and the economy, whereas the REF definition focuses more on the outcomes or effects from that research (29). This demonstrates a difference in the importance placed on whether impact occurs and the extent of impact (REF), versus the process through which impact takes place and the efforts adopted by institutions and researchers to maximise that impact (RCUK). This also aligns with the different perspectives adopted by the UK versus Australian governments in their national research performance assessment exercises.

Lastly, use-based definitions identified in Alla et al's review focused on the way research is utilised and how research findings are adopted by stakeholders. In this context, Lavis et al (32) have provided a classification to explain how research can be used. They distil the use of research into three main processes: instrumental use, conceptual use, and symbolic use. Instrumental use describes acting on research findings in a specific way to solve a particular problem and often constitutes the most obvious example of research impact. Conceptual use is the indirect use of research as a form of enlightenment. Symbolic use of research is not to inform decision making, but instead to justify a position of action or inaction that has already been taken. Although all types of research use could contribute to research impact as defined by the RCUK and the REF, examples of instrumental use of research to produce specific impacts will most closely align with the outputs based definition from the REF. Indeed, a criticism of the REF definition is that it prioritises examples of direct impact over indirect processes, which may be symbolic or conceptual.

1.2.5 Definitions relevant to research impact evaluation

The ISRIA guidelines (13) include several terms that are used throughout the research impact literature and will be used in this thesis. These are outlined in Table 1-1.

Table 1-1 Terms relevant to the research impact literature

Term	Definition
Indicator of impact	A description of what is being evaluated to indicate research impact. For example, number of lives saved, influence on patient views and attitudes, improvement in health literacy. Indicators of impact can be quantitative or qualitative measures. The term impact metric is often used to describe purely quantitative indicators of impact.
Methods for impact evaluation	Ways to collect or analyse data for the purposes of a research impact evaluation.
Categories	Types of impacts evaluated that are classified under common headings, for example, environmental impact or health impact.
Domains	Domains can be used as an alternative to categories of impact. Alla et al (29) use the term domains in a broader sense when they are describing definitions of research impact and the category of impact is just one of those domains. Other domains include the avenues of impact which are the processes through which research has impact, and the levels of impact, for example, individual/global/local.
Research impact framework	A schema or structure that guides how to collect, analyse and communicate data for the purposes of a research impact assessment. The constituents of a framework will depend on the philosophical assumptions underlying the approach to knowledge, and how impact occurs.
Outputs	Often used to refer to the immediate results or direct benefits arising from research, for example, the results from a research project.
Outcomes	Often used to refer to benefits or changes resulting from research that do not occur immediately and imply a longer-term process.

1.2.6 Overlap between implementation science and impact evaluation

Implementation science can be defined as 'the scientific study of methods to promote the systematic uptake of research findings and other evidence based practices into routine practice, and hence, to improve the quality and effectiveness of health services (33)'. The study of implementation has similarities to research impact evaluation, especially evaluation of the impact of health research, because a major part of how health research impacts society is through implementation of research findings (34). This is especially relevant for clinical trials that are an applied form of research.

One of the major reasons for conducting both impact evaluation and implementation studies is to understand if the expected downstream effects from research and research funding are being realised. There will also be overlap between the two disciplines in terms of the methods. Where they differ

is that investigation of if, and how, research has been implemented may or may not be one component of an impact evaluation. In contrast, within implementation science, the core aim of any assessment is to evaluate in detail the barriers to and facilitators for the translation of evidence into practice. The funder of a research project may use impact assessment to understand the value of their investment whereas the stakeholder for an implementation study may be a healthcare provider who is trying to run a health service as efficiently as possible by ensuring practice is evidence based.

This overlap between implementation science and the field of impact assessment was recently recognised formally at an international convention of researchers who are interested in impact assessment (35). The summit was called: 'In the Trenches: Implementation to Impact International Summit', and one of the key aims was to advance the science of implementation and impact by engaging in interdisciplinary dialogue. In the conference proceedings, the authors highlight that one way to achieve research impact is by implementation of research findings.

1.2.7 Cancer research, clinical trials and research impact assessment

Cancer research covers a vast array of activities, including basic laboratory studies, prospective or retrospective clinical projects at an individual patient level, or population based research. Cancer trials are an important example of applied, clinical research and as such, are an important mechanism through which new knowledge regarding the optimal management of patients will emerge.

Cancer trials are commonly described by the phase of the trial, the types of intervention being evaluating or the funding source. Phase I trials describe the early investigation of the use of novel therapies in humans, whereas phase II trials investigate the efficacy of treatments in the relevant disease area. Phase III trials are often large, randomised trials that compare the efficacy of novel treatments compared to a standard of care, and phase IV trials investigate the effectiveness of a therapy in practice after it has been shown to be at least as efficacious as the current standard of care and approved for use in the patient population.

The primary outcome investigated within cancer trials will often align with the phase of the trial. For example, commonly assessed endpoints for a phase I trial include measurement of the safety of a novel drug or a recommended dose of the drug to take forward for further testing in phase II trials (36). In phase II trials, the primary outcome often relates to efficacy of a treatment in a particular patient population without a specific comparator treatment (36). Phase III trials, which nearly always randomly allocate patients to treatments, commonly have primary outcomes that compare the efficacy of a treatment to a recognised standard treatment. The endpoint of a phase III trial can assess whether the novel treatment being tested is better than the current standard, which may include no active treatment/placebo (superiority trial), if the new treatment is equivalent (equivalence trial) or if the new therapy is not significantly worse than the standard of care (non-inferiority trial).

For superiority trials, the null hypothesis is that there is no difference between the treatments being tested. In equivalence trials, the difference between two treatments is pre-defined as "Δ" and the objective of the trial is to demonstrate that the treatments are equivalent, with the difference in confidence intervals for the treatment effect not exceeding -Δ and +Δ. For non-inferiority trials, the aim is to demonstrate that a treatment is not unacceptably worse than the standard of care. The treatment may be inferior, but as long as not unacceptably so. The acceptable margin of inferiority is defined in advance as Δ, and non-inferiority can be claimed if the lower boundary of the confidence interval of the treatment effect does not exceed -Δ (37). Often the treatment being compared to the standard of care in a non-inferiority trial has other benefits such as shorter treatment and/or reduced toxicity or cost.

The types of intervention being assessed in a cancer trial can vary from those directed at screening for malignancy, improved diagnostics, cancer treatments, assessment of the prevention of cancer, the natural history of how cancer develops, or improving the quality of life or palliation for patients with a diagnosis of malignancy (38). Cancer treatments can include investigational medicinal products, radiotherapy, surgery, or medical devices. Finally, the funding received to support trial set-up and delivery can, for example, be provided by

pharmaceutical companies, charities, academic institutions or a combination of these sources.

Cancer research funders, such as charities and government bodies, have shown interest in evaluating the impact of cancer trials. In 2014, CRUK commissioned a project to estimate the return to the UK economy from the £15 billion spent on cancer research between the years 1970 to 2009 (39). The results of this report were disseminated specifically to Members of Parliament (40). Despite interest in the topic, there is minimal accompanying guidance from cancer funders into how the evaluation of the impact of cancer research can be performed on a routine basis.

1.2.8 Summary

In summary, several definitions of research impact exist and the approach to evaluation will depend on the definition used, the stakeholder conducting the evaluation, the type of research being assessed, and the rationale for conducting impact evaluation. Acknowledging this issue, for the purposes of this thesis, research impact includes academic and wider impacts, neutral (effect) or positive (benefit) impacts, and those occurring through direct or indirect processes.

There is transparency needed in the methods used to assess impact in order for the value of the impact to be analysed and understood. Although research impact assessment has emerged formally as a recognised type of evaluation over the past decade, there is no standardised approach to assessment. One of the aims of this study is to reflect on approaches that may be suitably applied in particular to the evaluation of the impact of cancer trials. One large clinical trial (the Short Course Oncology Treatment (SCOT) trial, European Union Drug Regulating Authorities Clinical Trials Database (EudraCT) number 2007-003957-10 and National Clinical Trials (NCT) number 00749450) (41)) has been selected as a case study to explore and test various approaches to impact evaluation. This trial is discussed in more detail below.

1.3 The SCOT trial

1.3.1 Background

Although surgery is performed with curative intent in patients with stage II and stage III CRC, 40-50% of patients with stage III colorectal cancer (CRC) and at least 20% of patients with stage II CRC will experience disease relapse due to clinically occult micro metastatic disease present at the time of surgery (42). Adjuvant chemotherapy is used after surgery to reduce the risk of relapse, improve survival, and increase the chance of cure. The SCOT trial investigated the optimal duration of adjuvant chemotherapy for patients with a diagnosis of high-risk stage II or stage III CRC. Patients with stage II disease were defined as having high-risk disease if they had one or more of the following disease features: T4 disease, tumour obstruction with or without perforation of the primary tumour preoperatively, fewer than ten lymph nodes harvested, poorly differentiated histology, perineural invasion, or extramural venous or lymphatic vascular invasion. These features were in line with findings from the Multicenter international study of oxaliplatin/5-fluorouracil-LV in the adjuvant treatment of colon cancer (MOSAIC) trial (National Clinical Trial (NCT) number 00275210) (43), which identified a group of patients with stage II disease who benefited most from doublet chemotherapy in the adjuvant setting. Individuals included in the SCOT trial were fit (World Health Organisation (WHO) performance status 0 or 1) and aged over 18 years (44).

1.3.2 Clinical trial evidence prior to SCOT

1.3.2.1 Overview of previous studies

At the time the SCOT trial was developed, 6 months of adjuvant fluoropyrimidine-oxaliplatin doublet chemotherapy was a recognised standard of care for patients with CRC based on the results of several previous clinical trials (43, 45-58) which are summarised in Figure 1-1. The common medications and drug doses used within adjuvant CRC treatment trials are outlined in Appendix 1.

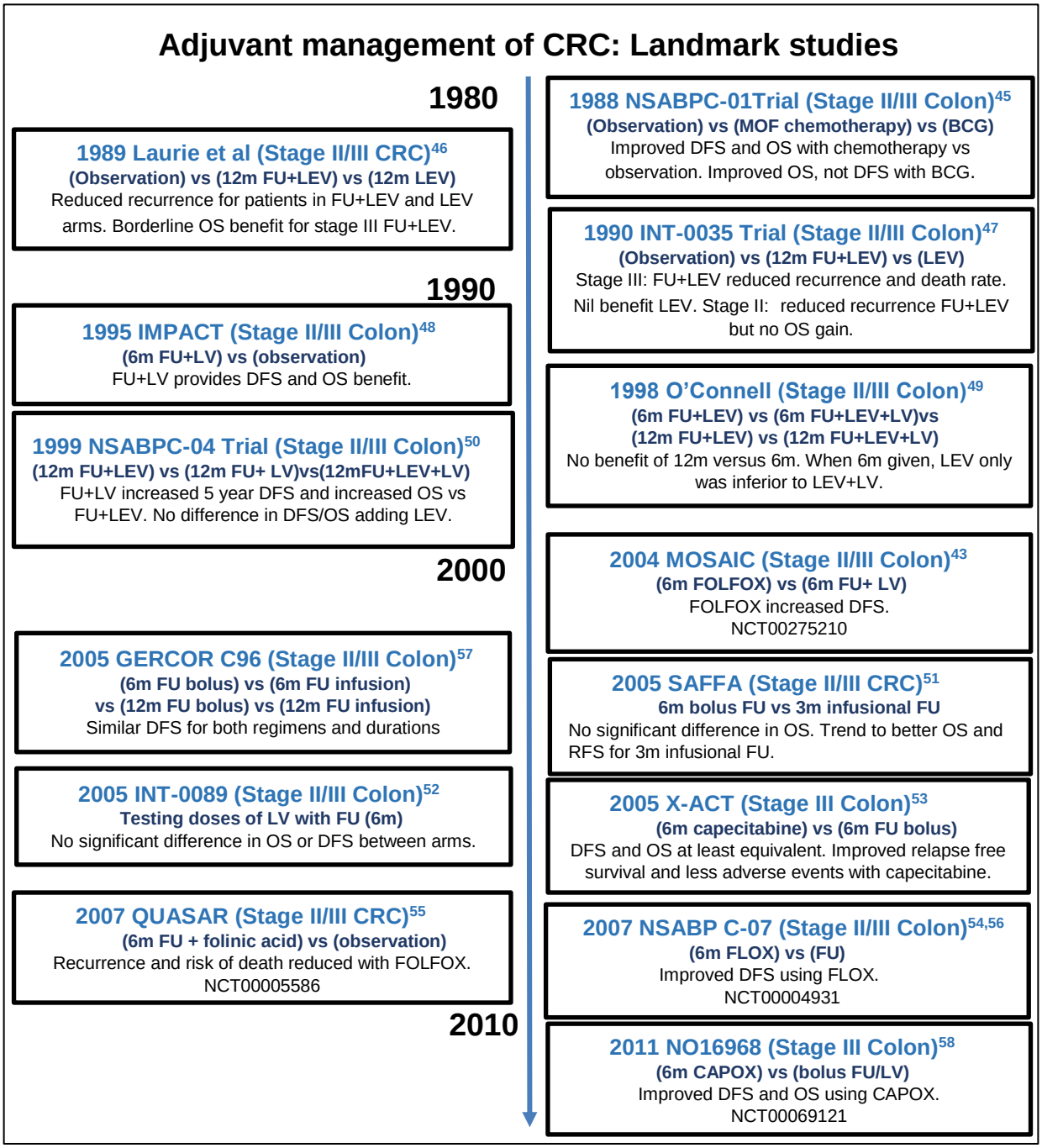


Figure 1-1 Clinical trials investigating the optimal treatment for patients with CRC in the adjuvant setting Abbreviations: FU, fluorouracil; LEV, levasimole; CRC, colorectal cancer; DFS, disease free survival; OS, overall survival; vs, versus; MOF, 5-fluorouracil, vincristine and semustine chemotherapy; BCG, Bacillus Clamette-Guerin; FOLFOX, 5-fluorouracil and oxaliplatin chemotherapy; CAPOX, capecitabine and oxaliplatin chemotherapy; m, months; LV, leucovorin; NSABP, National Surgical Adjuvant Breast and Bowel Project; MOSAIC, Multi-center International Study of Oxaliplatin/5FU-LV in the Adjuvant Treatment of Colon Cancer, RFS, recurrence free survival, IMPACT, Multi-Center Pooled Analysis of Colon Cancer Trials; INT, intergroup; X-ACT, Xeloda in Adjuvant Colon Cancer Therapy; SAFFA, Short Adjuvant Fluorouraci and Folinic Acid; QUASAR, Quick And Simple And Reliable; GERCOR, Groupe Cooperateur Multidisciplinaire en Oncologie.

In the 1980s, several therapeutic agents were trialled in the adjuvant setting for treatment of CRC (59), and by the end of the decade, it had been demonstrated that one year of intravenous 5-fluorouracil (5-FU), in combination with levamisole, offered survival benefits compared to observation alone (47). In the 1990s, several trials demonstrated that 5-FU/low-dose leucovorin was as effective as 5-FU/high dose leucovorin (52), 6 months of 5-FU with folinic acid/leucovorin was at least as effective as one year of treatment (48, 49) and that adding levamisole did not confer any extra benefit (49, 50).

During this time, bolus administration was used to deliver 5-FU and leucovorin, and monthly (Mayo) versus weekly (Roswell Park) bolus regimens were shown to be equally effective (59). In 2000, an oral form of fluoropyrimidine (capecitabine) was shown to be at least as effective as, and less toxic compared to bolus 5-FU when used in the adjuvant setting for treatment of colon cancer (53). Around the same time, there was increasing evidence that using two weekly infusional (FULV2), rather than bolus FU/LV, had comparable efficacy and less toxicity to the bolus regimens which had been traditionally used (57, 60).

Lastly, an important trial in relation to the hypothesis tested in SCOT was the UK Short Adjuvant Fluorouracil and Folinic Acid (SAFFA) trial, which compared 3 months of infusional 5-FU to 6 months of bolus Mayo clinic regimen (51). This trial was not large enough to test for non-inferiority but did suggest a trend for improved disease free survival (DFS) and overall survival (OS), and less toxicity, with the shorter infusional regimen. The authors of this trial concluded that the probability of 3 months of infusional 5-FU being inferior to 6 months of bolus treatment was highly unlikely and that the shorter duration of treatment merited further exploration.

In the 2000s, additional agents, such as irinotecan, bevacizumab, and cetuximab, were added to fluoropyrimidine monotherapy, but failed to significantly improve survival outcomes (61). However, the addition of oxaliplatin to 6 months of fluoropyrimidine was tested in three trials with success (43, 54, 58, 62). Each of these trials demonstrated that using doublet fluoropyrimidine-oxaliplatin chemotherapy conferred DFS and OS benefits compared to fluoropyrimidine monotherapy alone for patients with stage III colon cancer.

1.3.2.2 Previous evidence specific to patients with stage II colon cancer

The SCOT trial was set-up in the late 2000s. At that time, it was known that adjuvant fluoropyrimidine monotherapy conferred a survival benefit for patients with stage II and stage III colon cancer. Although it was also known that 6 months of fluoropyrimidine-oxaliplatin doublet treatment improved survival for patients with stage III disease, it was still not clear if doublet treatment conferred survival benefit for those with stage II disease; the long-term outcomes from important trials such as MOSAIC were awaited (56, 63). Whilst these trial results were still outstanding, patients with stage II disease were included within the SCOT trial. Subsequently, longer-term follow up from MOSAIC (reported in 2015) did not confirm a statistically significant 5-year DFS or 6-year OS improvement for patients with stage II disease with the addition of oxaliplatin versus fluoropyrimidine alone (63).

1.3.2.3 Previous evidence specific to patients with rectal cancer

In addition to patients with colon cancer, individuals with a diagnosis of rectal cancer were recruited to SCOT. A Cochrane review published in 2012, which pooled data from over 21 clinical trials (n>10,000), reported a significant DFS and OS benefit from giving fluoropyrimidine based adjuvant chemotherapy to patients with rectal cancer (64). Important limitations to this review included the heterogeneity of disease stages included and the fact that these trials were conducted over a long time-period, during which surgical resection for rectal cancer significantly improved. Also, several trials included in the Cochrane review combined adjuvant chemotherapy with radiotherapy, making it difficult to ascertain the benefit of chemotherapy treatment alone.

1.3.2.4 Previous evidence specific to older patients

There are no randomised trials specifically investigating the benefits of adjuvant chemotherapy for elderly patients with CRC. Pooled analyses of colon cancer trials have shown the same benefit from using adjuvant fluoropyrimidine monotherapy in patients above or below 70 years. A systematic review (65) which included the MOSAIC (66), NSABP C-07 (56) and XELOXA (62) trials demonstrated increased toxicity from doublet treatment, but no benefit from adding oxaliplatin to fluoropyrimidine therapy for patients over 70. In contrast,

other pooled analyses including more recent trials (67-69) have shown the benefit from adding oxaliplatin is present regardless of age, although less so for those aged 70 and over. Overall, the benefits for older patients receiving chemotherapy are similar to younger patients with regards time to recurrence, but are less for DFS and OS, implying that elderly patients may have the same recurrence reduction but cannot derive the benefit in terms of survival due to competing co-morbidities and poorer survival (70).

1.3.3 SCOT: the clinical question and trial design

The SCOT trial was a phase III, multi-centre, international, non-inferiority trial that compared 3 versus 6 months of adjuvant chemotherapy for patients with a diagnosis of colorectal cancer (CRC). When the SCOT trial was designed, it was recognised that the side effects of 6 months of treatment were often debilitating, and could be permanent. In particular, oxaliplatin chemotherapy caused sustained peripheral neuropathy in a substantial (>15%) proportion of patients, many of whom would be cured from their primary cancer (71-74). It was hypothesised that shorter treatment that is not unacceptably worse in terms of efficacy, may lead to reduced toxicity. As well as the clinical question, the trial included an economic analysis built into the study design to ensure that potential cost-effectiveness of the treatments tested were assessed. The full trial details are published in peer reviewed journals (41, 75, 76).

The primary endpoint for SCOT was non-inferiority of 3-year DFS. DFS was defined as the time from randomisation to relapse, development of a new CRC, or death from any cause. The estimated 3-year DFS for the standard 6 months of treatment was 78%, based on the use of 6 months of 5-fluorouracil-oxaliplatin (FOLFOX) chemotherapy within the MOSAIC trial. The pre-defined 3-year DFS difference between the two treatments (assuming shorter treatment would have worse efficacy) that was deemed clinically acceptable was 2.5%, that is, 3 months would not be unacceptably worse if at least half of the benefit from adding oxaliplatin to fluoropyrimidine monotherapy demonstrated in the MOSAIC trial was maintained (3-year DFS for monotherapy was 73%). This maximally acceptable drop in efficacy corresponded to a hazard ratio for death of 1.13 and the planned recruitment was 9,500 patients. Secondary endpoints included OS, toxicity, and cost-effectiveness. The fluoropyrimidine-oxaliplatin doublet permitted was either

CAPOX (capecitabine-oxaliplatin) or mFOLFOX6, and the choice of which regimen to use was at the discretion of the treating clinician, on the assumption that there was no difference in efficacy between these two regimens that deliver the same drugs in different formulations (77).

1.3.4 SCOT trial results

1.3.4.1 Patient information

The SCOT trial recruited 6,088 patients from 244 centres in six countries between March 2008 and November 2013. This was less than the target of 9,500 patients due to slow accrual and therefore the study had 66% power rather than 90% power for rejecting the null hypothesis. The majority of patients were recruited from the UK (n=5,244), followed by Denmark (n=311), Spain (n=237), Australia (n=197), Sweden (n=83) and New Zealand (n=16). Median follow up in both groups was 37 months (Inter quartile range (IQR) 36-49) and 787 patients had died at the time of cut-off for the primary analysis in December 2016. In total, 23 patients did not consent for their data to be used after randomisation, therefore analysis was conducted on 6,055 patients, 3,035 in the 3 month arm and 3,030 in the 6 month arm, on an intention to treat basis. Baseline demographics were balanced, with median age 65 years (IQR 58-70) and a female to male ratio of 39%:61% in both trial arms. Overall, approximately two thirds of patients received CAPOX and one third received FOLFOX according to clinician preference.

1.3.4.2 Main effectiveness results

The SCOT trial met its primary endpoint and demonstrated that 3 months of adjuvant doublet chemotherapy was non-inferior to using 6 months of treatment in the overall trial population. In the 3-month arm, 3-year DFS was 76.7% (95% confidence interval (CI) 75.1-78.2) compared to 77.1% (95% CI 75.6-78.6) in the 6-month arm (0.4% absolute difference). The pre-specified non-inferiority margin of an upper CI of 1.13 was met (HR 1.006 (0.909-1.114)), test for non-inferiority p=0.012) (41). The Kaplan-Meier curve for DFS is shown in Figure 1-2.

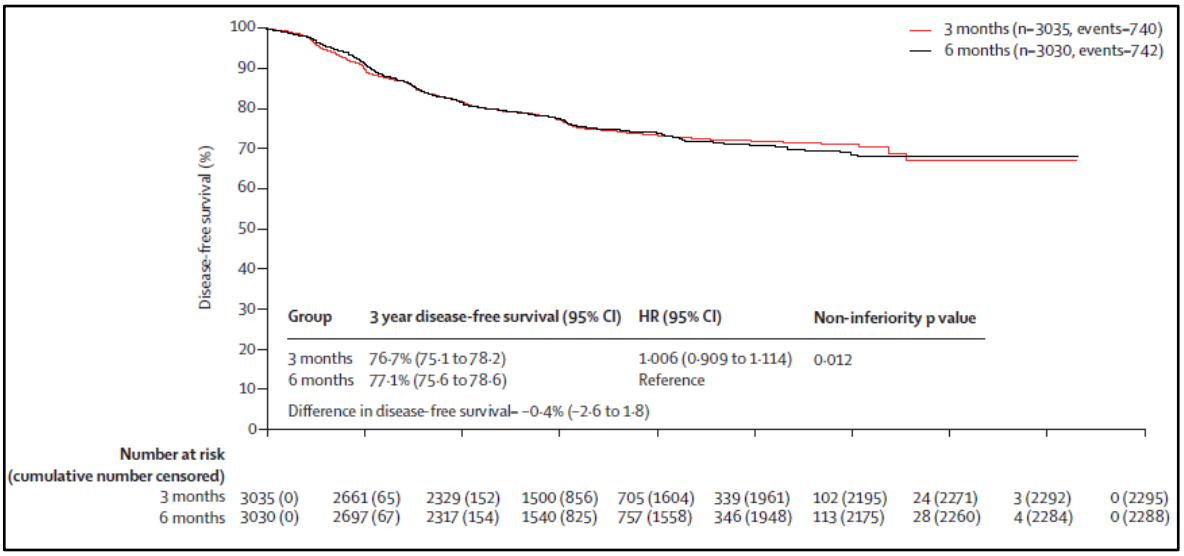


Figure 1-2 Kaplan-Meier curve showing 3 year DFS by study group This diagram has been taken directly from the original SCOT trial results publication (41).

1.3.4.3 Toxicity results

Side effects for patients receiving 3 months of treatment were significantly reduced compared to patients in the 6-month arm. The frequency of grade 3–5 diarrhoea (p=0.033), neutropenia (p=0.031), pain (p=0.014), hand-foot syndrome (p=0.031), and sensory neuropathy (p<0.0001) was significantly higher in the 6 month group than in the 3 month group. In particular, the percentage of patients suffering from Grade 2 or greater peripheral neuropathy during and up to one month after treatment was less than half (25% vs 58%). Information from neuropathy specific questionnaires completed by patients throughout follow-up demonstrated that higher symptoms of neuropathy in the 6 versus 3-month arm lasted up to 5 years post-treatment (Figure 7 in (41)).

1.3.4.4 Subgroup analysis

Analysis of 3-year DFS by stratification factors (Figure 1-3) revealed unexpected heterogeneity in outcome dependent on the choice of chemotherapy regimen (interaction p=0.069).

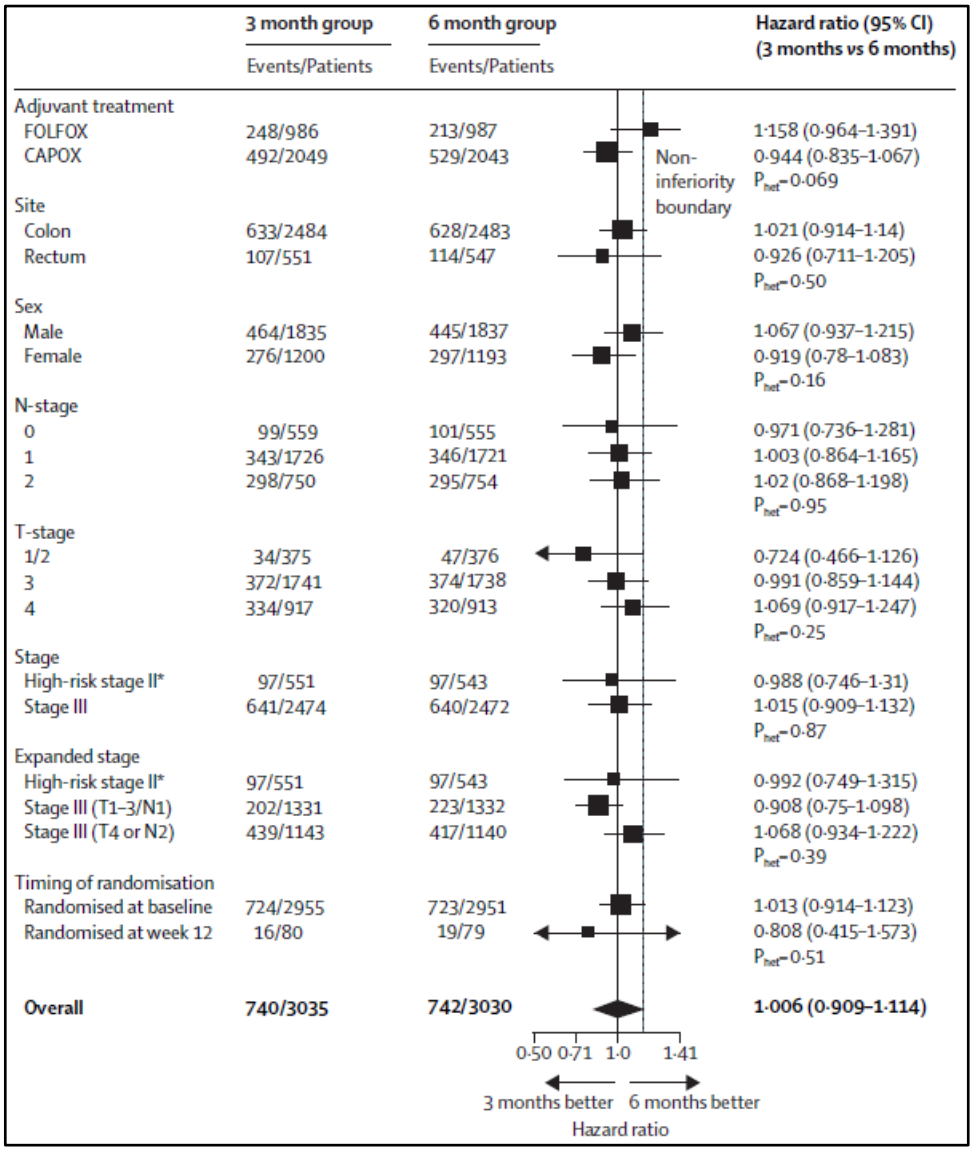


Figure 1-3 Forest plot showing DFS by stratification factors and timing of randomisation. This diagram has been taken directly from the original SCOT trial results publication (41).

Consequently the authors performed a post-hoc analysis to investigate this difference. The Kaplan-Meier curves for DFS for patients who received CAPOX compared to those who received FOLFOX are shown in Figure 1-4. Non-inferiority using 3-months versus 6-months was demonstrated for CAPOX (p=0.0020) but not for FOLFOX. Explanations put forward for this unexpected difference include the different oxaliplatin doses and dose density of fluoropyrimidine used in the two regimens, (78) and/or intrinsic differences in the tumour microenvironment that may dictate response to a specific regimen (79, 80).

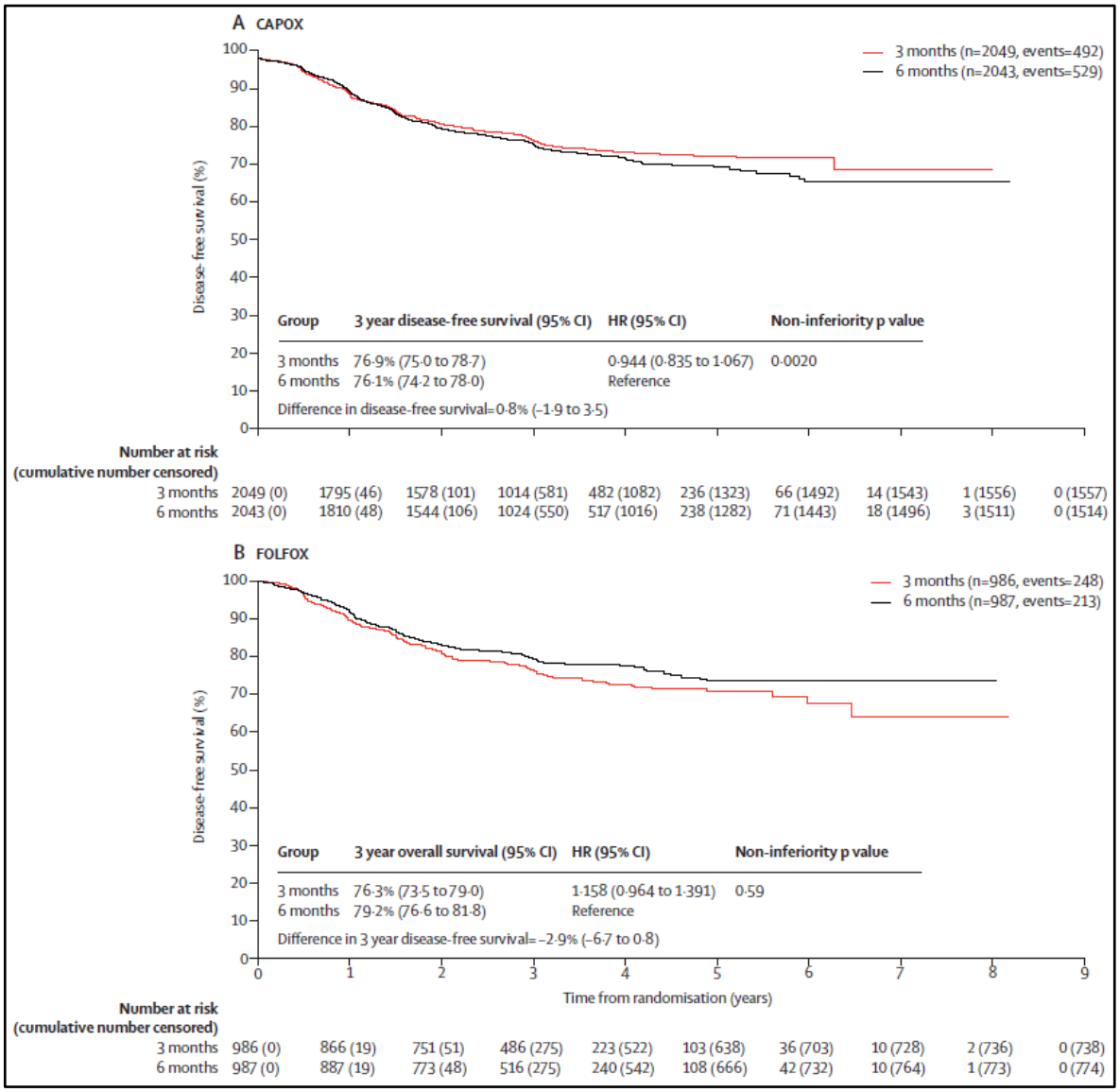


Figure 1-4 Kaplan-Meier curves showing DFS by study group for patients receiving A CAPOX and B FOLFOX This diagram has been taken directly from the original SCOT trial results publication (41).

A post-hoc analysis was also performed by separating stage III disease into two risk groups: low-risk (T1-3N1 disease) and high-risk (T4 or N2 or both). For patients in the low-risk group, 3 months of treatment was non-inferior to 6 months, but this was not the case for patients with high-risk stage III disease Figure 1-5.

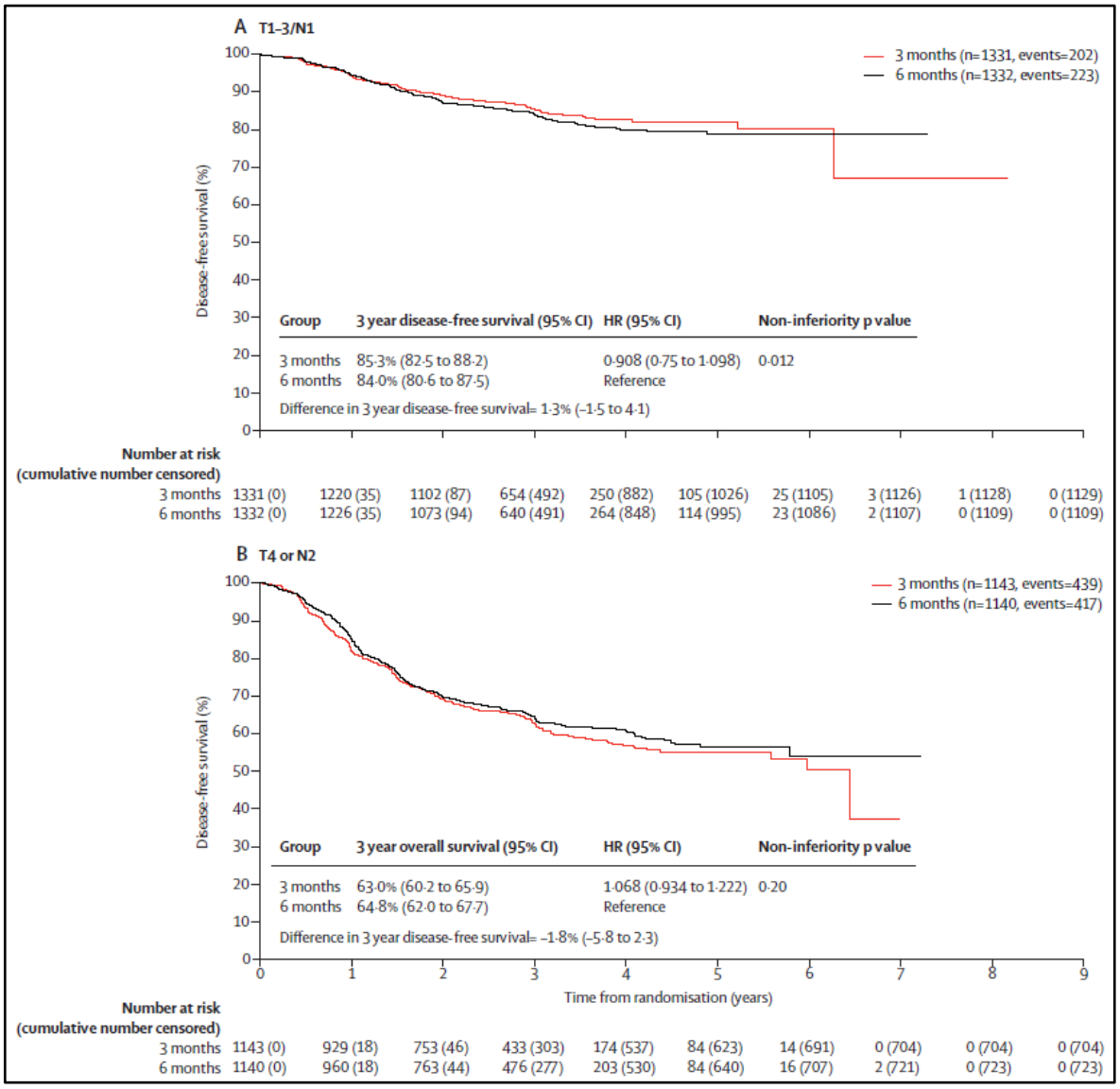


Figure 1-5 Kaplan-Meier curves showing DFS by study group for patients receiving A T1-3N1 disease and B T4N2 disease This diagram has been taken directly from the original SCOT trial results publication (41).

Although not yet published in full, SCOT trial results specifically for patients with stage II disease with high-risk features were presented at a special session at ESMO 2019. Non-inferiority was not met for this subgroup in the overall population. The clinical difference in 3 year DFS for 3 versus 6 months of treatment was small when CAPOX was used (1.4% difference). Table 1-2 outlines the 3 year DFS results for the SCOT overall study group and for the subgroups discussed above.

Table 1-2 Three year DFS for overall SCOT trial study group and by regimen and stage subgroups *Non-inferiority of 3 versus 6 months met according to pre-specified boundary for the SCOT trial. NR: Not recorded.

	3 months (95% CI)	6 months (95% CI)	HR (95% CI)	Non-inferiority p value
Overall*	76.7% (75.1%-78.2%)	77.1% (75.6%-78.6%)	1.006 (0.909-1.114)	0.012
CAPOX*	76.9% (75.0%-78.7%)	76.1% (74.2%-78.0%)	0.944 (0.835-1.067)	0.002
FOLFOX	76.3% (73.5%-79.0%)	79.2% (76.6%-81.8%)	1.158 (0.964-1.391)	0.590
T1-3N1 ("Low Risk Stage III")*	85.3% (82.5%-88.2%)	84.0% (80.6%-87.5%)	0.908 (0.750-1.098)	0.012
T4 or N2 ("High Risk Stage III")	63.0% (60.2%-65.9%)	64.8% (62.0%-67.7%)	1.068 (0.934-1.222)	0.200
CAPOX Low-risk Stage III (T1-3N1)	86.3% (84.0%-88.7%)	82.9% (80.4%-85.5%)	NR	NR
FOLFOX Low-risk Stage III (T1-3N1	83.4% (79.9%-86.9%)	86.3% (83.0%-89.6%)	NR	NR
CAPOX High-risk Stage III (T4 or N2)	62.1% (58.6%-65.6%)	63.4% (59.9%-66.9%)	NR	NR
FOLFOX High-risk Stage III (T4 or N2)	65.0% (60.0%-69.9%)	67.7% (62.9%-72.5%)	NR	NR
Stage II with high-risk features	84.3% (81.2%-87.3%)	86.1% (83.2-89.1%)	0.949 (0.730-1.223)	NR
CAPOX Stage II with high-risk features	84.3% (80.7%-90.1%)	85.7% (82.2%-89.2%)	0.923 (0.680-1.254)	NR
FOLFOX Stage II with high-risk features	84.1% (78.2%-90.1%)	87.2% (82.0%-92.3%)	1.059 (0.638-1.757)	NR

1.3.4.5 SCOT Cost-effectiveness analysis

The cost-effectiveness of 3 versus 6 months of adjuvant chemotherapy for patients with CRC was explored in a within-trial analysis as part of the SCOT trial using patient level data (75). Compared to 6 months, 3 months was the dominant strategy by being cheaper and providing a (non-statistically significant) quality adjusted life year (QALY) gain. The incremental net monetary benefit (INMB) at a willingness to pay (WTP) threshold of £30,000 was £7246. Sub-group analysis revealed that whereas 3 months of CAPOX was cost-effective at a range of willingness to pay thresholds, there was more uncertainty regarding the cost-effectiveness of 3 months versus 6 months of FOLFOX.

1.3.5 IDEA Collaboration

The SCOT trial findings were not only reported independently, but they also made a significant contribution to the International Duration Evaluation of Adjuvant therapy (IDEA) collaboration (81). The IDEA collaboration was a pre-planned pooling of results from six international trials that all investigated the use of shorter duration of adjuvant doublet chemotherapy for patients with stage III colon cancer. Four of the trials, including SCOT, also recruited patients with stage II disease, and these results were also pooled and reported separately. Table 1-3 (adapted from (81)) shows the details for each of the six trials. The primary endpoint for all six trials was DFS and the IDEA collaboration used a modified intention to treat method for the primary analysis. Assuming a 3-year DFS of 72% in the 6 month group and specifying a one-sided type 1 error rate of 0.025, the predefined upper limit of the 95% CI for the hazard ration of 3-year DFS was 1.12.

Table 1-3 Trials contributing data to IDEA Collaboration. Adapted from (81) Abbreviations: ACHIEVE, Adjuvant Chemotherapy for Colon Cancer with High Evidence; CALGB/SWOG, Cancer and Leukemia Group B/South-West Oncology Group; IDEA, International Duration Evaluation of Adjuvant Therapy; SCOT, Short Course Oncology Treatment; TOSCA, Three or Six Colon Adjuvant; HORG, Hellenic Oncology Research Group; NCT, National Clinical Trial number; UMIN, University Medical Information Network; EudraCT, European Clinical Trials database.

Name of trial	Clinical Trials Number	Countries	Stage of disease	Site of disease	Regimens	Regimen split	Included in stage III IDEA	Number of patients contributed to IDEA stage III	Included in IDEA stage II	Number of patients contributed to IDEA stage II
ACHIEVE	UMIN 000008543	Japan	III	Colon	CAPOX, FOLFOX	CAPOX 75.1% FOLFOX 24.9%	Yes	1291	No	0
ACHIEVE 2	UMIN 000013036	Japan	II	Colon	CAPOX, FOLFOX	CAPOX 84% FOLFOX 16%	No	0	Yes	514
CALGB/SWOG 80702	NCT 01150045	USA, Canada	III	Colon	FOLFOX (plus celecoxib)	FOLFOX 100%	Yes	2440	No	0
HORG	NCT 01308086	Greece	II/III	Colon	CAPOX, FOLFOX	CAPOX 58.2% FOLFOX 41.8%	Yes	708	Yes	413
IDEA France	EudraCT 2009-010384-16	France	III	Colon	CAPOX, FOLFOX	CAPOX 10% FOLFOX 90%	Yes	2010	No	0
SCOT	NCT 00749450	Australia, Denmark, New Zealand, Spain, Sweden, UK	II/III	Colon, Rectum	CAPOX, FOLFOX	CAPOX 66.5% FOLFOX 33.5%	Yes	3983	Yes	1078 (included 130 rectal patients)
TOSCA	NCT00646607	Italy	II/III	Colon	CAPOX, FOLFOX	CAPOX 35% FOLFOX 65%	Yes	2402	Yes	1268

The stage III IDEA collaboration did not meet its pre-specified non-inferiority endpoint (HR 1.07, 95% CI 1.00-1.15, p=0.11 for non-inferiority of 3-month therapy, p=0.045 for superiority of 6-month therapy). Nevertheless, the clinical difference in 3-year DFS between 3 months vs 6 months was small (0.9%). The 3-year DFS in the 3-month arm was 74.6% (95% CI 73.5%-75.7%) compared to 75.5% (95% CI 74.4%-76.7%) in the 6-month arm.

As in the SCOT trial, toxicity was less for patients in the 3-month group, and the same regimen and disease stage differences that were reported in the SCOT trial were reported. Specifically, non-inferiority of 3 months versus 6 months of treatment was proven for patients with low-risk stage III colon cancer (regardless of regimen), those patients who were prescribed CAPOX chemotherapy, and subsequently for those individuals who were prescribed CAPOX and had low-risk disease. Overall, for patients receiving FOLFOX chemotherapy (regardless of risk stage) or for those with high-risk stage III disease (regardless of regimen used), 3 months of treatment was inferior to 6 months of chemotherapy. For patients receiving FOLFOX who had low-risk disease specifically and for those receiving CAPOX who had high-risk disease specifically, non-inferiority of the shorter treatment duration was not proven. These regimen and disease risk subgroup differences are outlined in Figure 1-6 below.

3 yr DFS rate (%) and HR by regimen and risk group		Regimen								
		CAPOX			FOLFOX			CAPOX/FOLFOX combined		
		3 yr DFS, % (95% CI)		HR (95% CI)	3 yr DFS, % (95% CI)		HR (95% CI)	3 yr DFS, % (95% CI)		HR (95% CI)
		3 m	6 m		3 m	6 m		3 m	6 m	
Risk group	Low-risk (T1-3 N1) ~60%	85.0 (83.1-86.9)	83.1 (81.1-85.2)	0.85 (0.71-1.01)	81.9 (80.2-83.6)	83.5 (81.9-85.1)	1.10 (0.96-1.26)	83.1 (81.8-84.4)	83.3 (82.1-84.6)	1.01 (0.90-1.12)
	High-risk (T4 and / or N2) ~40%	64.1 (61.3-67.1)	64.0 (61.2-67.0)	1.02 (0.89-1.17)	61.5 (58.9-64.1)	64.7 (62.2-67.3)	1.20 (1.07-1.35)	62.7 (60.8-64.4)	64.4 (62.6-66.4)	1.12 (1.03-1.23)
	Risk groups combined	75.9 (74.2-77.6)	74.8 (73.1-76.6)	0.95 (0.85-1.06)	73.6 (72.2-75.1)	76.0 (74.6-77.5)	1.16 (1.06-1.26)	P-value interaction test: Regimen: 0.0061 Risk group: 0.11		

Non-inferior | Not proven | Inferior

Figure 1-6 Three year DFS for the IDEA stage III collaboration by regimen and disease risk stage This figure has been taken directly from the Supplementary material of the IDEA collaboration publication in the New England Journal of Medicine 2018 (81).

The early response by the academic community to the IDEA collaboration finding was to recognise the complexity of the results and the unexpected difference between regimens that was revealed. In a special European Society for Medical Oncology (ESMO) plenary session in September 2017, a panel of eleven academics, including nine of the chief investigators from the IDEA trials, debated what they referred to as 'hard road' to interpret the data (78). At this plenary session, patient preference was discussed by referring to patients as having a 'fighter' or 'fatalist' attitude to risk, with fighters more likely to accept small improvements in survival regardless of toxicity, whereas fatalists were less likely to accept longer, more toxic treatment if the additional benefit offered compared to shorter treatment was small. In general, there was a consensus that results relating to using 3 months of CAPOX for low-risk stage III disease were practice changing, whereas if using FOLFOX, then 6 months of treatment should remain the current standard when treating high-risk disease. The conclusions around using 3 months of CAPOX for high-risk disease and around continuing to use 6 months of FOLFOX for low-risk disease were less clear and in these situations, patient attitude to risk made the most difference to clinicians when making practice decisions. National

Comprehensive Cancer Network (NCCN) guidelines (USA), updated soon after publication of the IDEA results (82) allowed for a certain amount of flexibility and variability in practice depending on the risk of disease, regimen selected and patient preference.

Updated survival analysis from the stage III IDEA analysis, presented at the virtual America Society of Clinical Oncology (ASCO) conference 2020 (83), showed very little difference in 5-year OS (0.4%) between using 3 versus 6 months of treatment in the overall trial population (see Kaplan-Meier 5-year OS curve below).

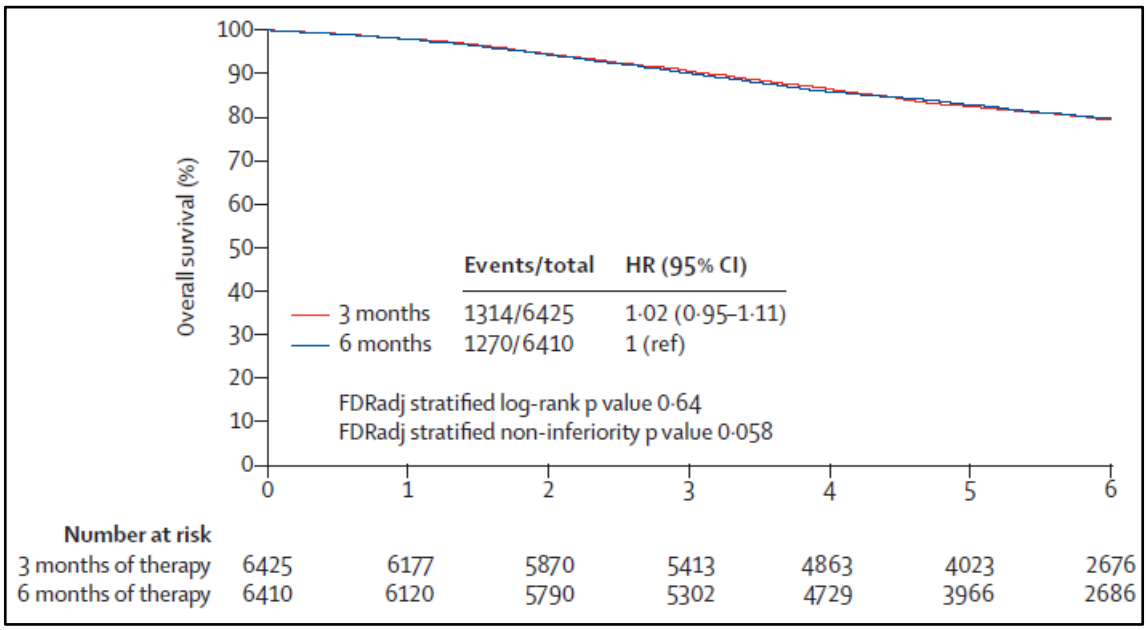


Figure 1-7 Five-year overall survival with 3 versus 6 months of adjuvant treatment from the IDEA collaboration This diagram has been taken directly from the publication describing the updated IDEA collaboration stage III results (84). 5 year OS for 3 month arm was 82.4% (95% CI 81.4%-83.3%) versus 82.8% (95% CI 81.8%-83.8%) for the 6 month arm (HR 1.11 (0.95-1.11), non-inferiority p=0.058.

Although the 95% non-inferiority HR margin met the IDEA pre-specified non-inferiority boundary (1.11), the statistical conclusion was to reject the null hypothesis after a multiplicity adjustment. Figure 1-8 shows a summary of the updated 5-year OS results for regime and disease risk subgroups in the IDEA collaboration. The general consensus following the presentation of these results was that the OS curves for 3 versus 6 months of treatment were virtually inseparable, with minimal clinical difference between the approaches, strengthening the argument to use shorter treatment for most patients (85, 86). The take home message was that 5-year OS was already very high for patients with stage III disease and using 3 versus 6 months was unlikely to make a significant

difference for most patients. The main rationale for still using 6 months of treatment would be when prescribing FOLFOX chemotherapy for patients with high-risk stage III disease.

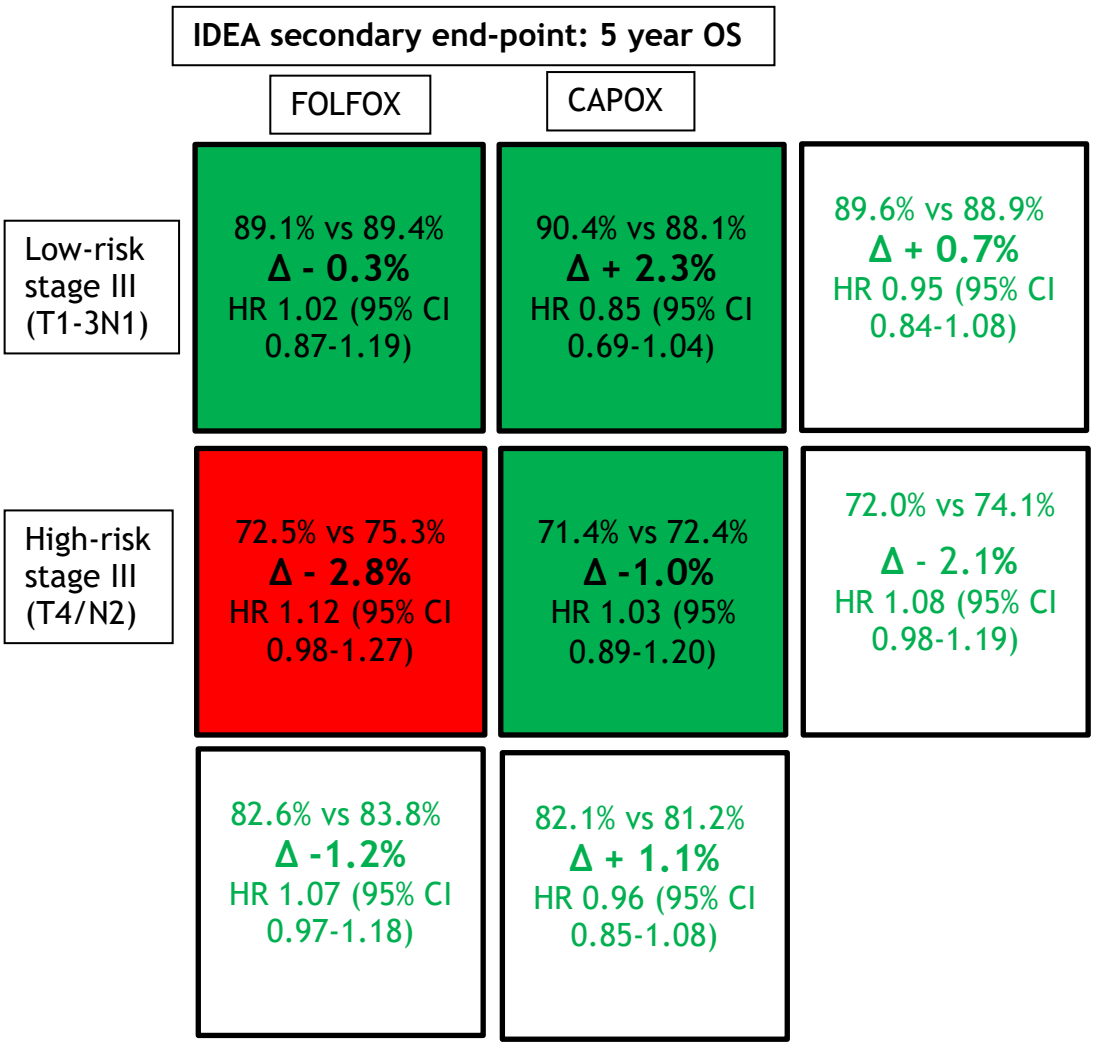


Figure 1-8 Subgroup analyses from the updated IDEA stage III collaboration This figure has been created using information from the updated IDEA stage III publication (84)

The stage II IDEA collaboration results, combining data from the four trials which recruited patients with stage II disease (see Table 1-3), were presented at ASCO 2019 (87) and reported in full in 2021 (88). Five-year DFS was similar in both groups (80.7% in 3 month arm versus 83.9% in 6 month arm, 3.2% absolute difference), but the non-inferiority margin was not met statistically (HR 1.17, 80% CI 1.05-1.31, p=0.3851). Again, toxicity was significantly reduced for the shorter arm. Overall, grade 3-5 toxicity was reduced approximately by half (40% in 6 month arm versus 26% in 3 month arm, p<0.0001) and grade 2-4 neurotoxicity was reduced by one third (36% versus 13%, p<0.0001). Between regimen differences in treatment effect were also demonstrated. Specifically, 5 year DFS was 81.7% in

the 3 month arm versus 82.0% in the 6 month arm for patients receiving CAPOX. For patients receiving FOLFOX, 5 year DFS was 79.2% in the 3 month arm versus 86.5% in the 6 month arm.

1.3.6 Summary

In summary, the SCOT trial illustrated that 3 months of adjuvant doublet chemotherapy is non-inferior (3 year-DFS), significantly less toxic, and more cost-effective compared to 6 months of treatment. SCOT was the largest contributor to the IDEA collaboration and has contributed to novel findings for stage III, high-risk stage II and both colon and rectal cancer patients.

The SCOT trial was chosen as a case study for the purposes of this thesis because it was a large phase III trial that provided both clinical and health economic results. The preliminary results of SCOT (89) were reported 6 months prior to the commencement of this study, providing a good opportunity to evaluate early impact. The preliminary results indicated that the trial had met its primary endpoint and therefore there was a hypothesis that these findings had the potential to impact on clinical practice. In addition, the SCOT trial was developed and run by the Glasgow Clinical Trials Unit (CTU), which was supporting the Clinical Trials Fellowship associated with this research.

1.4 Thesis aims and objectives

The aim of this thesis was to explore existing approaches to research impact assessment, to identify which approaches have been used previously for cancer trial impact evaluation, and to test a number of these approaches within the context of a case study. To meet this aim, the specific objectives were:

Part I: Exploring methods of research impact evaluation and identifying approaches relevant to cancer trials.

- Identify approaches to impact assessment used across research disciplines, find examples of the application of these approaches to evaluate cancer research, and explore which of these approaches would be best suited to evaluate the impact of cancer clinical trials (Chapter 3).

- Analyse REF 2014 impact case studies to identify examples of the evaluation of cancer clinical trial impact and establish the methods used and the types of impacts evaluated in these case studies (Chapter 4).

Part II: Case study of the SCOT trial to test approaches to impact assessment.

Part IIa: An analysis of estimated impact of SCOT

- Assess the impact of the SCOT trial findings on clinician attitudes and self-reported prescribing practice change (Chapter 5).

- Assess the potential economic impact of implementation of the SCOT trial findings on countries involved in SCOT (Chapter 6).

Part IIb: An analysis of real world implementation of SCOT trial results

- Evaluate real life practice change post-SCOT using local prescribing data (Chapter 7).

- Assess the feasibility of using national data to assess practice change (Chapter 8).

Chapter 2 outlines a number of the materials and methods developed and used throughout this study. The methods utilised to perform each specific part of the analysis are explained in detail within each chapter (Chapter 3-8). Chapter 9 reflects on the work performed, considers how the findings can be used, and suggests how the results can help to direct future research in this field.

2 Chapter 2: Materials and methods

Several distinct methods were used to achieve the aim of this thesis. This Chapter provides a description of a number of these methods, with signposting to good practice guidelines, and when relevant, describes the development of, or access to, materials and data required to perform the analyses. There is also a description of any regulatory approvals attained and software versions used. A more specific description of why specific methods were chosen and the details of the particular analyses that were performed are provided in each results chapter (Chapters 3-8).

2.1 Overview of reviews

2.1.1 Description of approach used

In Chapter 3, a systematic review of the literature was performed to identify approaches to assessment of research impact. An overview of reviews, also known as an umbrella review, was the approach taken, given the large number of studies and many reviews that have already been undertaken in this area over the past decade. Overviews of reviews are generally used to summarise broad issues and current knowledge on a topic, to signpost the reader to evidence, summarise existing research, and highlight where an absence of evidence may exist (90).

In carrying out this overview of reviews, existing guidelines on the methodology for conducting an overview of previously published literature reviews were followed (91, 92). The approach taken aligned with guidance by using a clearly defined research question, adopting a systematic approach to searching for relevant review articles, and reporting results of the search using the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) (93). The methodology for this overview differed from the guidelines for a traditional umbrella review by including both systematic and narrative reviews, rather than systematic reviews only (94).

2.1.2 Systematic review data extraction

A data extraction form was designed to collect details for each review. Information on the year of publication, primary author location, relevant research discipline, aims of the review as detailed by the authors, and the search strategy used (if any) to review the literature were collected.

A prior scoping review of the literature (unpublished) performed in the initial stages of this study revealed three themes that were relevant to describing approaches to impact assessment. Findings from each review were summarised under these themes. These were (i) categorisation of impact into different types depending on who or what is affected by the research (the individuals, institutions, or parts of society, the environment), and how they are affected (for example health, monetary gain, sustainability) (ii) methods of data collection and analysis for impact assessment, (iii) frameworks to organise and communicate research impact. There was also space available on the data extraction form to document other key findings.

A separate data extraction form was developed to extract data from any empirical examples of cancer research impact assessment identified from the review articles. The information extracted included: year of publication, location of primary authors, research discipline, aims of the evaluation as described by the authors, research unit under assessment, and the rationale for impact evaluation. Data was extracted using the same three themes as outlined above, and the approaches used in these studies were compared to those identified from the literature reviews.

2.2 Content Analysis

2.2.1 Overview

Content analysis (95) was the methodology used for the study reported in Chapter 4. Content analysis is defined as ‘a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use’ (Krippendorf 2004) (96). An earlier definition from Breleson (1952) describes content analysis as: ‘a research technique for the objective, systematic and quantitative description of the manifest content of communication (97).’ The data

analysed for the purposes of a content analysis is usually non-numerical and can be in the form of written, oral, or visual data. It is now generally accepted that content analysis can be used to analyse both the manifest content of a document, as well as the latent content. Manifest content refers to the inherent meaning that can be deduced directly from the text or other medium, whereas latent content refers to the interpretation of meanings that may be implied but not explicitly stated in the source (95).

In order to perform a content analysis for the purposes of this study it was necessary to develop a coding manual. A coding manual is a statement of instructions to researchers that includes all the possible options for each dimension being coded (95). The content manual for this study was developed using the methodology outlined in the social sciences research methods primary text by Bryman (Fourth Edition) (95).

2.2.2 Coding manual development

2.2.2.1 Source documents for analysis

The coding manual was developed to analyse the impact case studies from the REF exercise in 2014. A typical case study contained an initial section that included a title and information on the submitting institution, research subject area(s), the 'Unit of Assessment' and the 'Summary Impact Type'. The Units of Assessment were 36 subject areas, each with a separate REF expert review panel. The Summary Impact Types were eight categories of impact, assigned to each case study by text analysis after submission to the REF. These categories were technological, economic, health, political, legal, cultural, societal and environmental (9). The next section within each case study was a short summary of the impact, a description of the underpinning research on which the impact described was based, and a list of references that representing the findings of that research. The main part of the case study in which the higher education institutions described the impact attributed to the research they listed was the 'Details of the impact' section, followed by a reference list of sources to corroborate the claims made in this text. For the purposes of this study, the title section, list of research references and the 'Details of the impact' section were analysed. It was expected that the researcher using the coding manual would look

up the full publication for any clinical trials listed in the research references section.

2.2.2.2 Units of analysis

There were two main units of analysis and the coding manual was developed to allow coding of each unit separately. The first unit of analysis was the characteristics of the case studies and clinical trial(s) described in each case study and the second unit of analysis was the content of the case study itself (that is, the approaches to impact assessment used by universities).

2.2.2.3 Dimensions used in the coding manual

Dimensions for the first version of the coding manual were chosen by the primary researcher (CH). For the first unit of analysis (characteristics of case studies and clinical trials), these dimensions included the main headings within the case studies (university, type of research) as well as characteristics of clinical trials deemed important to the researcher (for example, type of malignancy investigated, funding source for trial). For the second unit of analysis (approaches to impact assessment used by universities), categories of impact and methods to evaluate impact were used based on results from Chapter 3 of the thesis.

Each dimension within a case study could take a number of different pre-specified options; these options were the codes used for the purposes of the content analysis. These codes were initially populated using information from the previous literature (98-100), the researcher's own experience, from information contained on the REF 2014 webpage, or from the results of Chapter 3 (Table 3-2 and Figure 3-2).

The initial aim was that all codes would apply to manifest messages included in the REF case studies to minimise any bias that may be introduced through subjective interpretation on the part of the coder. During development of the codes, it became clear that coding for some of the dimensions would require analysis of a latent message within the text. This was particularly relevant when coding categories of impact described by the higher education institutions. On occasions, there were sub-headings used within the text of case studies, such as 'impact on practice' or 'impact on clinical guidelines' that manifest the message

being portrayed by the author of the case study. However, on many occasions, the case study text was not explicit regarding the type of impact being described and it was necessary for the researcher to make a judgement about the message being conveyed within the text. To reduce bias in this coding as much as possible, detailed examples and descriptions of categories of impact were developed through iteration of the coding manual.

The first version of the coding manual was developed by the primary researcher (CH) alone. Although all dimensions for the initial coding manual were decided in advance of reading the case studies, some codes used in the final analysis were emergent from the data and finalised through the iterative process of coding manual development. For example, some of the pre-specified codes did not meet the criteria for being mutually exclusive and exhaustive, and were developed and refined by reading the case studies and ensuring all options were included and that no options were overlapping. The primary researcher read and re-read the 46 case studies and adjusted the coding manual deductively so that is was fit for purpose.

A second researcher (Lauren Gatting, (LG)) was involved in the coding manual development from this stage onward. LG is a PhD student at the Institute of Health and Wellbeing, University of Glasgow. The primary researcher (CH), having identified the need for this collaboration to meet the requirements of checking validity of coding for the purposes of content analysis, sought out, and initiated contact with LG.

The primary researcher coded all case studies and LG coded two randomly selected case studies using the initial coding manual. Random number generation (Microsoft Excel®) was used to select case studies for double coding. Both researchers then discussed the ease of use of the manual and the appropriateness and relevance of the codes. After discussion, several changes were made; in particular, the manual was divided into ‘Part A’ (case study and trial characteristics) and ‘Part B’ (impact evaluation). More detailed descriptions and examples for each impact category relevant to cancer trials were also included.

Version 2 of the coding manual was used for double coding of two further case studies. Again, the primary researcher manually coded all case studies and the

second researcher coded two randomly selected case studies. On this occasion, an inter-rater reliability score (101) was calculated as an indication of the agreement between coders using coding manual version 2. Rather than continuing with this coding manual, it was agreed by both researchers that a further iteration to the manual would be beneficial. The manual was changed from Microsoft Word® to Microsoft Excel®, with two separate spreadsheets for Part A and Part B. In Part B, indicators (see Table 1-1 Chapter 1 for definition) of impact were provided for each category to make coding more straightforward. These changes were made to produce the final version of the coding manual, which can be found at the following link: http://researchdata.gla.ac.uk/1135/.

Throughout the process of manual development and content analysis, the primary researcher performed the first two stages of coding all case studies using paper copies of printed source data. The third and final stage of coding was performed using Nvivo®. The second researcher (LG) used Nvivo® software for all coding. The results of coding at each stage were transferred into a Microsoft Excel® for analysis.

2.3 Survey Design

A survey was used for the analysis reported in Chapter 5. General texts (95, 102-105) on survey design were consulted in the development of the survey used.

2.3.1 Search for previous surveys

As per recognised good practice (105), a scoping review of the literature was performed to identify any existing, validated surveys that could be used to investigate clinicians attitudes and self-reported practice, in general or specifically in response to clinical trial evidence. No validated, pre-existing surveys were identified.

2.3.2 Survey development

In the development of survey content, a mixture of closed and open questions were used. More open questions were used in the pilot phases of the survey to get broad feedback from respondents on the face and content validity of the survey questions. Key open questions were kept in the final survey to allow

respondents to expand on answers about practice change where they felt appropriate.

The final survey began with a statement of purpose, followed by a confidentiality statement (106). This was important to ensure proper understanding by the respondents of the rationale for the survey, and to inform participants how their personal data would be used. All participants confirmed at the start of the survey that they prescribed adjuvant CRC chemotherapy. Section 1 asked respondents about their awareness of clinical trials and guidelines. Section 2 asked about participants' current practice, and how their practice may have changed in response to clinical trial findings. Section 3 used a Likert scale (107) to explore attitudes to using 3 months of doublet chemotherapy, and the final section asked about respondents their place of work and occupation. As per good practice guidelines, personal questions were left to the end of the so that respondents knew what information they had provided in the survey before entering more personal details (104).

2.3.3 Survey validity

Face validity (103), which assesses the extent to which a survey analyses what it has set out to achieve, was assessed in the pilot phase by asking respondents if they felt that the survey asked them about their current clinical practice, practice change in response to trials, and their attitudes to using shorter adjuvant treatment for CRC. Content validity (103), which assesses if there is sufficient content in the survey to explore the issue in question, was analysed by asking pilot respondents if the survey adequately addressed all relevant aspects of this topic. Finally, criterion (or concurrent) validity considers the extent to which questionnaire results agree with an independent or gold-standard measure of the same variable (103). As there was no gold standard for clinicians' interpretation or opinion towards the results of the clinical trials, the possibility of testing the validity of the responses on current practice by assessing the chemotherapy prescribing records of five local clinicians who answered the pilot survey was considered. However, although this would have been technically possible, on further consideration, it was felt that this would be questioning the truthfulness of clinician's responses and would raise issues about potential comparisons between clinicians. Also, if the respondents' replies did not

correlate with their actual prescribing it would be difficult to know if this was an individual issue or if this would be mirrored when assessing a larger group of health care professionals. One measure of reliability is the internal consistency of the answers to the survey. This was addressed in the pilot survey by using questions that queried the same concept or opinion but used different wording. The aim was to test this in the final survey using Cronbach's alpha statistic (108).

2.3.4 Survey piloting

Piloting was performed to test survey layout, wording and comprehension (109). The initial pilot survey was sent to a six colleagues, including clinicians who were representative of the intended sample of respondents, and five responses were received. One of the main pieces of feedback was that duplicate questions, which were included to assess internal consistency, were cumbersome, adding irrelevant length to the survey, and were likely to decrease response rate. On consultation with the expert in survey design, it was decided to omit this measure of reliability.

Pilot 2 was sent to twenty individuals, including lay people, academics, and potential participants, four of whom had answered the first survey; thirteen people replied. The survey was piloted for a third and final time by emailing to seven potential participants, six of whom replied. Feedback from all pilot rounds and subsequent changes made are available at this link: http://researchdata.gla.ac.uk/1135/. The first pilot was written in Microsoft Word®, whereas versions 2 and 3 were constructed using OnlineSurveys®. The final survey is available to view at this link: https://glasgow-research.onlinesurveys.ac.uk/uk-survey-adjuvant-treatment-for-patients-with-colorectal-2

2.3.5 Development of follow up survey

If clinicians provided their email addresses for the purposes of sending them the study results or contacting them in the future with a follow up survey, this list of email addresses was downloaded separately and stored within a different file on a password protected, University of Glasgow One Drive account.

A second survey was developed with the aim of following-up responders to the first survey to explore if, and how, clinicians' self-reported practice in response to clinical trials and their attitudes to using a shorter duration of adjuvant CRC chemotherapy had changed over time. A pilot for the second survey was sent to four individuals, two of whom replied. Both of these individuals were clinicians, one treated patients with CRC as part of their clinical duties. The second survey is available to view at this link: https://glasgow-research.onlinesurveys.ac.uk/follow-up-survey-adjuvant-treatment-for-colon-cancer-gene

2.3.6 Survey analysis

Descriptive statistics were used to analyse the survey results using Microsoft Excel® and STATA. Diagrams were drawn using Microsoft Excel®, Microsoft Word®, Online maps® and Google® Charts. Where appropriate, if comparing the proportions of different responses between group of clinicians, Chi-squared or Fisher's exact tests were calculated (110).

2.4 Cost-utility analysis

For the purposes of the study reported in Chapter 6, one of the methods used to conduct the analysis was a within trial cost-utility analysis. Other types of economic evaluation and how they compare to a cost-utility analysis are outlined in Table 13-1 Appendix 4. This analysis adhered to good practice economic methodology (Consolidated Health Economic Evaluation Reporting Standards (CHEERS) (111) and the statistical code developed by researchers (Dr Jose Antonio Robles-Zurita and Dr Kathleen Anne Boyd) at the University of Glasgow for the initial SCOT study evaluation (75, 76) was adapted for this updated analysis. The rationale for repeating this analysis and the approaches to estimate outcomes and costs within the cost-utility calculation are outlined in Chapter 6.

2.4.1 Cost-utility analysis from a multi-national perspective

The objective of the cost-utility analysis reported in Chapter 6 was to conduct the analysis from the perspective of all six countries that recruited to SCOT. As

different countries' regulatory bodies often set different requirements in their health technology assessments (HTAs), there is not one recognised approach for conducting an economic evaluation alongside a clinical trial that has recruited patients from multiple countries (112). Prior to commencing this analysis, the main features of each of the healthcare systems in these countries were reviewed and the HTA guidelines specific to each country were compared and contrasted (information available by following this link: http://researchdata.gla.ac.uk/1135/).

Multi-country cost-effectiveness analyses can be described according to the source of the effectiveness data and resource data used (see Table 13-2 Appendix 4). Fully pooled analyses use patient level data on clinical effectiveness and resource use from all patients in the trial. Partially split analyses use clinical effectiveness estimates from patients in all countries but resource use data from only one country, or a selection of countries, (with unit costs from that country or selection of countries). Finally, analyses that are fully-split use clinical effectiveness and resource use estimates from the same group of patients in one country only. This has the advantage of maintaining the patient level link between clinical and cost-effectiveness but small, country-specific patient numbers often limits the statistical robustness of this approach.

Multi-country cost-effective analyses can also be defined by the method of costing used. Single country costing is when unit costs relevant to one country are applied to all patients in a trial regardless of the location for that patient. The alternative is multi-country costing, in which country specific costs are applied to patients from that specific country. The limitations and benefits of different combinations of approaches for pooling data and applying unit costs are provided in Table 13-3 Appendix 4.

2.4.2 Unit costs for multi-national cost-utility analysis

Country specific unit costs were used within the cost-utility analysis to calculate the cost of resource use of interest. For chemotherapy unit costs, as far as possible, prices that reflected the drug tariff price were used. Figure 13-1 Appendix 4 shows how the tariff price compares to other categories of unit costs

available. The source of chemotherapy medication costs used are shown in Table 2-1.

Gathering country specific information on hospital unit costs was challenging. Unit costs were often not available and when they were, they were only in the national language, which in many cases was not English. Often, resource allocation and pricing is at a regional rather than national level (Sweden, Spain) and these regional level unit costs were not publicly available (Sweden). When national unit costs were available they were often based on diagnosis-related codes (DRGs). Although DRGs are a commonly used method of calculating the cost of resource use based on the diagnosis of the individual patient, for the purposes of this analysis costs based on DRG coding were not applicable to the patient level data collected in the SCOT trial. Finally, although in several instances individual resource unit costs were found for a particular type of hospital stay and country, it was not clear how comparable these would be between countries in terms of the component costs included.

For all of these reasons, one main source of unit cost information, the WHO CHOICE (CHOosing Interventions that are Cost-Effective) project, was used (113, 114). The ratios of unit costs for each non-UK country as compared to those for the UK were calculated from the CHOICE study (113) (Table 2-1). The Scottish Information Services Division (ISD) cost book (2019) (115) unit costs for hospitalisation resource use were used as the source of UK costs which were adjusted for other countries using the WHO ratios. In order to calculate the unit cost per day or per night of stay, the total cost was divided by the average length of stay. WHO ratios for tertiary public hospital stays were used to adjust unit costs for ICU and oncology inpatient stays. Ratios for secondary public hospital stays were used to make adjustments for general medical stays, and outpatient unit costs were adjusted using ratios for public, tertiary hospital outpatient visits. A consumer price inflation (CPI) rate was used to convert unit costs from 2010 to 2019 prices. Health specific purchasing power parity (PPP) was used to convert country specific currencies to USD (116). USD and 2019 were the currency and year used for the base case analysis.

Table 2-1 Unit costs used for this analysis

	AUSTRALIA	DENMARK	NEW ZEALAND	SPAIN	SWEDEN	UK
Medication unit costs (all in USD)						
Oxaliplatin	0.3200	0.3504	0.3925	0.1135	0.5263	0.1164
Capecitabine	0.0009	0.0006	0.0009	0.0006	0.0008	0.0005
5-fluorouracil (bolus)	0.0150	0.0117	0.0051	0.0030	0.0031	0.0010
5-fluorouracil (infusion)	0.0150	0.0117	0.0051	0.0033	0.0031	0.0010
Information source	www.pbs.gov.au	Personal communication with Amgros December 2019 (amgros.dk/en/)	www.pharmac.govt.nz	FU: https://es.vidal-consult.com (free trial provided by company on request) Oxaliplatin: Pharmacy Service, Hospital General Universitario de Valencia, Spain (Communication). Contacted in September 23, 2020. Capecitabine: Framework agreement for the supply of the drug "capecitabine" to the hospitals of the Madrid Health Service, 2016. Consulted, September 24th, 2020.	www.tlv.se ("Decision drugs" tab and AIP used)	https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit
Hospitalisation unit costs (all in USD)						
Inpatient ratio	1.22	1.18	0.81	0.87	1.35	1.00
Outpatient ratio	1.15	1.13	0.86	0.91	1.10	1.00
ICU	4065	3938	2709	2917	4512	3340
HDU	1739	1685	1159	1248	1931	1429
GM	899	871	599	646	998	739
IN_clinonc	1869	1811	1246	1342	2075	1536
IN_clinonc_with_tx	1412	1443	993	1069	1654	1224
OUT_clinonc	521	509	388	410	496	452
DAY_clinonc	1566	1530	1166	1231	1489	1358
DAY_clinonc_with_tx	819	800	610	644	778	710

2.5 Budget impact analysis

Budget impact analysis was a second method used for the purposes of the study reported in Chapter 6. It is increasingly recognised that understanding not only the value of new treatments, but also the real world consequences of using these treatments in practice are important from a health technology and policy perspective, and therefore several countries have published country specific budget impact guidelines (117). Important information from these guidelines were summarised prior to embarking on this analysis, and this information is available by following this link: http://researchdata.gla.ac.uk/1135/. After comparing and contrasting these guidelines, a decision was made to use the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) guidelines (2014) (118) for the purposes of this study to ensure consistency of approach rather than following guidelines from one of the countries.

In order to perform the budget impact analysis, a cost-calculator was built. This can be viewed at: http://researchdata.gla.ac.uk/1135/. Face validity of the calculator was checked by asking a clinician and a health economist from each country to review the approach used. The proformas used for these checks are available at the same link.

2.6 Evaluation of chemotherapy prescribing data

Chemotherapy prescribing data was assessed on a local (Greater Glasgow and Clyde (GG&C) health board) (Chapter 7) and national (Scotland) (Chapter 8) level. Below is a description of how these data sets were acquired. A description of how these datasets were analysed and the results of these analyses are provided in the relevant results chapters.

2.6.1 Data access: Local

2.6.1.1 Acquisition of datasets

Acquisition of datasets for the purposes of analysis of local chemotherapy prescribing data is outlined in Figure 2-1. Extracting and linking most of the GG&C data was performed by a data analyst (Christine Crearie (CC)). GG&C

chemotherapy prescribing data was linked with additional, key data variables, such as disease staging. This linkage was possible because CC had access to a system called (**A**cute hospital discharges, **Ca**ncer Registrations, **D**eath records and **Me**ntal health) AcaDME. ACaDMe is part of a National Health Service (NHS) National Services Scotland (NSS) data warehouse that stores information from the Scottish Medical Registry (SMR) for acute admissions (SMR01), mental health (SMR04) and cancer (SMR06), and NRS deaths. Demographic and staging data were provided to the primary researcher in a separate data file to the dataset containing drug dose data.

Cancer quality performance data was acquired from a separate data provider. This data was extracted by a Quality Performance Indicator (QPI) data analyst (Finlay MacKay). The three datasets (chemotherapy drug dose, demographics from AcaDME, and QPI data) were linked by the primary researcher (CH) using a unique identifier called the Community Health Index (CHI) number, which was available for each patient. This unique identifier (CHI) has been assigned to every patient on first registration to the healthcare service in Scotland since the 1970s. A CHI number consists of a patient's date of birth plus four additional digits.

There were no direct research costs incurred to access this local data and the datasets were available within one month of approval to access the data. Patient identifiable information was sent to the primary researcher (CH) via NHS email and this data was stored on a NHS computer.

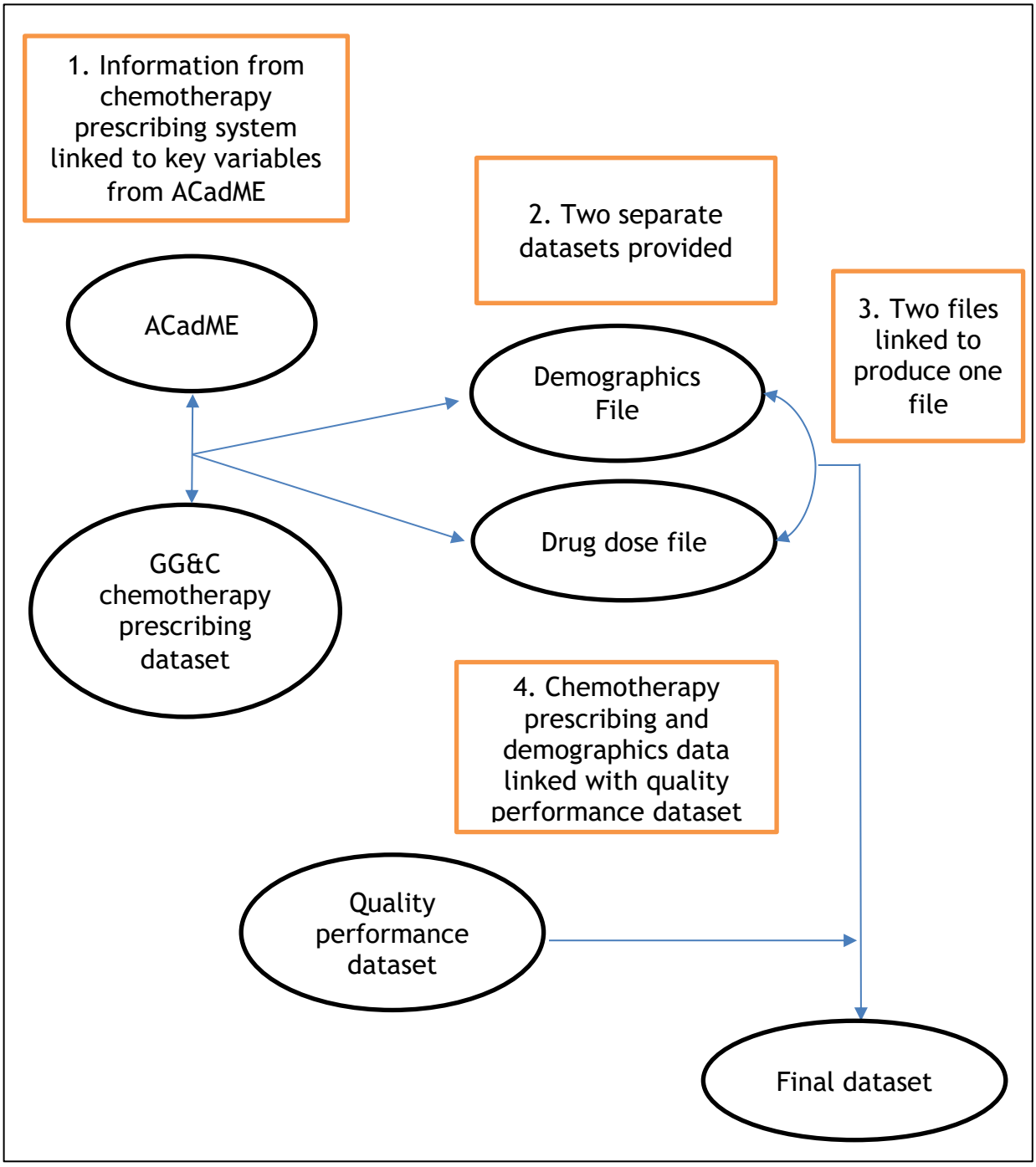


Figure 2-1 Access and linkage of GG&C datasets

2.6.2 Data access: National

2.6.2.1 National data

The data used for analysis at a national level represented the first example of a dataset that included granular chemotherapy prescribing data covering the Scottish population linked with other administrative datasets, which was made available to researchers. This dataset was acquired for the purposes of this study and the Scottish arm of a UK wide programme called the Cancer Research UK

(CRUK) funded COloRECTal Repository (CORECT-R), which is aiming to quantify the characteristics of, and any variation in, CRC and its management in the UK (119, 120). The data used within the study described in this thesis was a subset of the whole dataset that was acquired, but the PhD researcher (CH) was heavily involved in the process of acquiring data for both projects because of this overlap. The process of acquisition of the whole, larger dataset is described below, followed by a more specific focus on the datasets used for the purposes of this project. There were four main stages (Figure 2-2) in accessing and linking data on a national level.

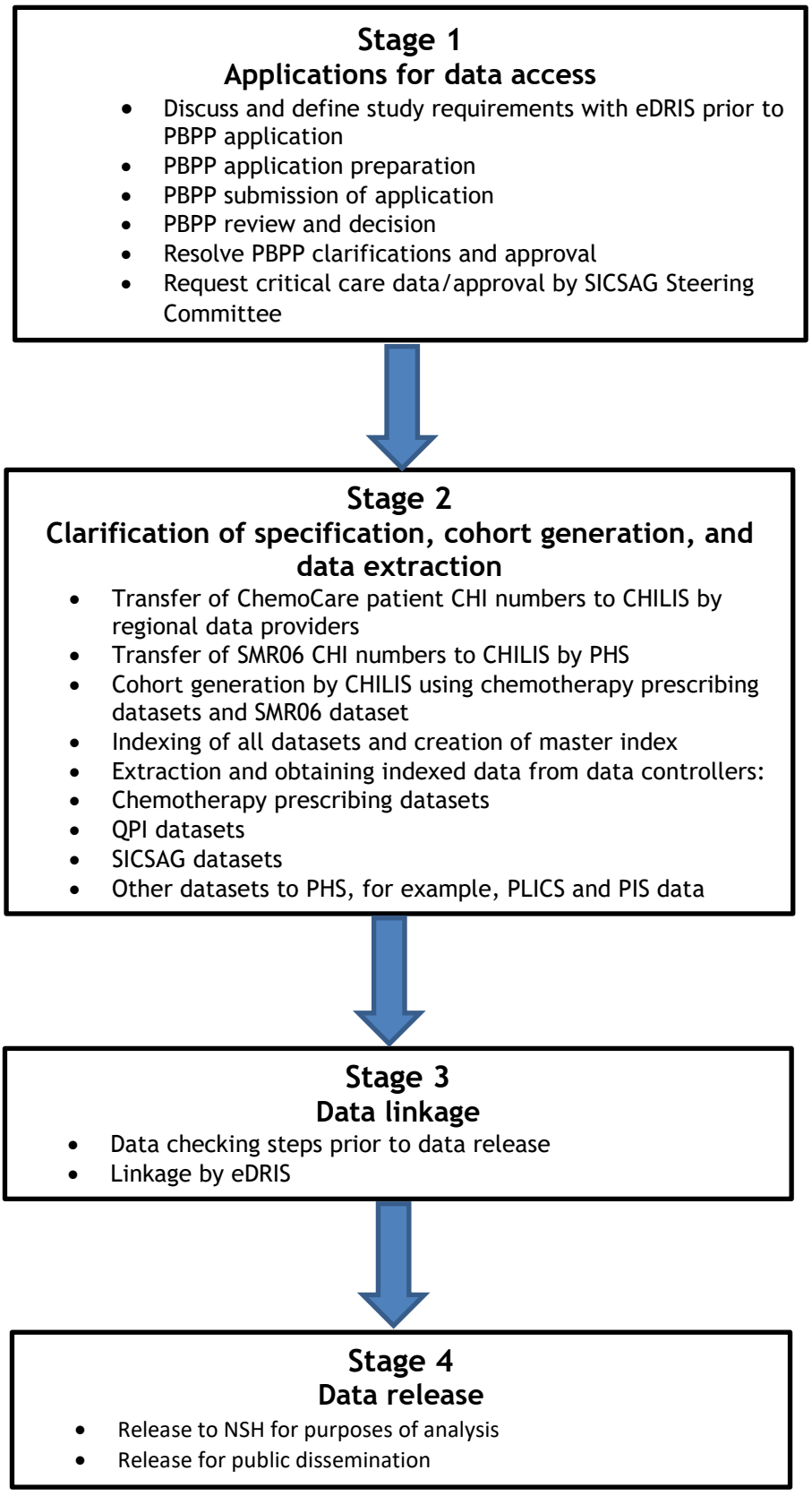


Figure 2-2 Flowchart outlining the main stages to access and link data Abbreviations: NSH, National Safe Haven; eDRIS, electronic Data Research and Innovation Service; PHS, Public Health Scotland; QPI, Quality Performance Indicator; SICSAG, Scottish Intensive Care Society Audit Group; CHILIS, CHI Indexing and Linkage Service; SMR, Scottish Morbidity Record; PBPP, Public Benefit and Privacy Panel for Health and Social Care.

Stage 1

The first stage in accessing data was to define the study requirements in order to apply to the Public Benefit and Privacy Panel (PBPP) for Health and Social Care in Scotland (121). PBPP have responsibility for weighing up the benefits to the public from granting access to healthcare data against the risk that the sharing of the data poses to an individual's privacy. All applications to PBPP go to a Tier 1 panel for proportionate review. Some applications will be referred on for further review by a Tier 2 Committee.

Separate PBPP applications for this project (PBPP reference number 1718-0263) and the CORECT-R project (PBPP reference number 1718-0026) were developed, but it was recognised in advance that this project would use a subset of data acquired for the larger initiative. The PBPP application for this project was written by the primary researcher (CH), whereas the PBPP application for the CORECT-R project was written by Dr Holly Ennis (CORECT-R project manager in Scotland) and Dr Peter Hall (CORECT-R principal investigator in Scotland). The primary researcher (CH) for this project was named on both applications. A list of the datasets requested for the larger application are outlined in Table 14-1 in Appendix 5.

First contact with an eDRIS co-ordinator was made in January 2018 and both PBPP applications were submitted in parallel in April 2018. PBPP approval was granted in June 2018 for this project and in October 2018 after Tier 2 review by a full panel of PBPP committee members for the CORECT-R submission. A substantial amendment for project 1718-0026 was necessary (written by Dr Holly Ennis) and this was approved by the PBPP committee in February 2020.

Stage 2

The second stage was acquisition of datasets for transfer into the National Safe Haven (NSH). The NSH is a research platform operated by Edinburgh Parallel Computing Centre (EPCC) on behalf of Public Health Scotland (PHS). The NSH provides a secure analytical environment where data controllers can allow administrative data to be used for research purposes when it is not practical to obtain individual patient consent, whilst protecting patent privacy and identity.

eDRIS were the principal department of PHS responsible for overseeing data transfer.

Datasets were divided into those held by PHS and those held by data controllers external to PHS (Table 14-1 in Appendix 5). Figure 2-3 shows the data transfer process that occurred to transfer data to the NSH. A trusted Third-Party indexing team (CHI Indexing and Linkage Service (CHILIS)) facilitated this transfer for the cohort generation and indexing of datasets. Specifically, this meant that no identifiable data was sent directly from data controllers external to PHS to the eDRIS team. Instead, patient identifiers were replaced with a unique patient identifier and the data was subsequently considered pseudonymised because the link between unique identifiers and CHI numbers was held by a Trusted Third Party (CHILIS). In addition, under the General Data Protection Regulation (GDPR), health data is considered sensitive category personal data and therefore cannot be considered fully anonymised. The cohort of patients included in the final dataset (Figure 2-3 “Master Cohort List”) was defined using a combination of Cancer Registry and chemotherapy prescribing data.

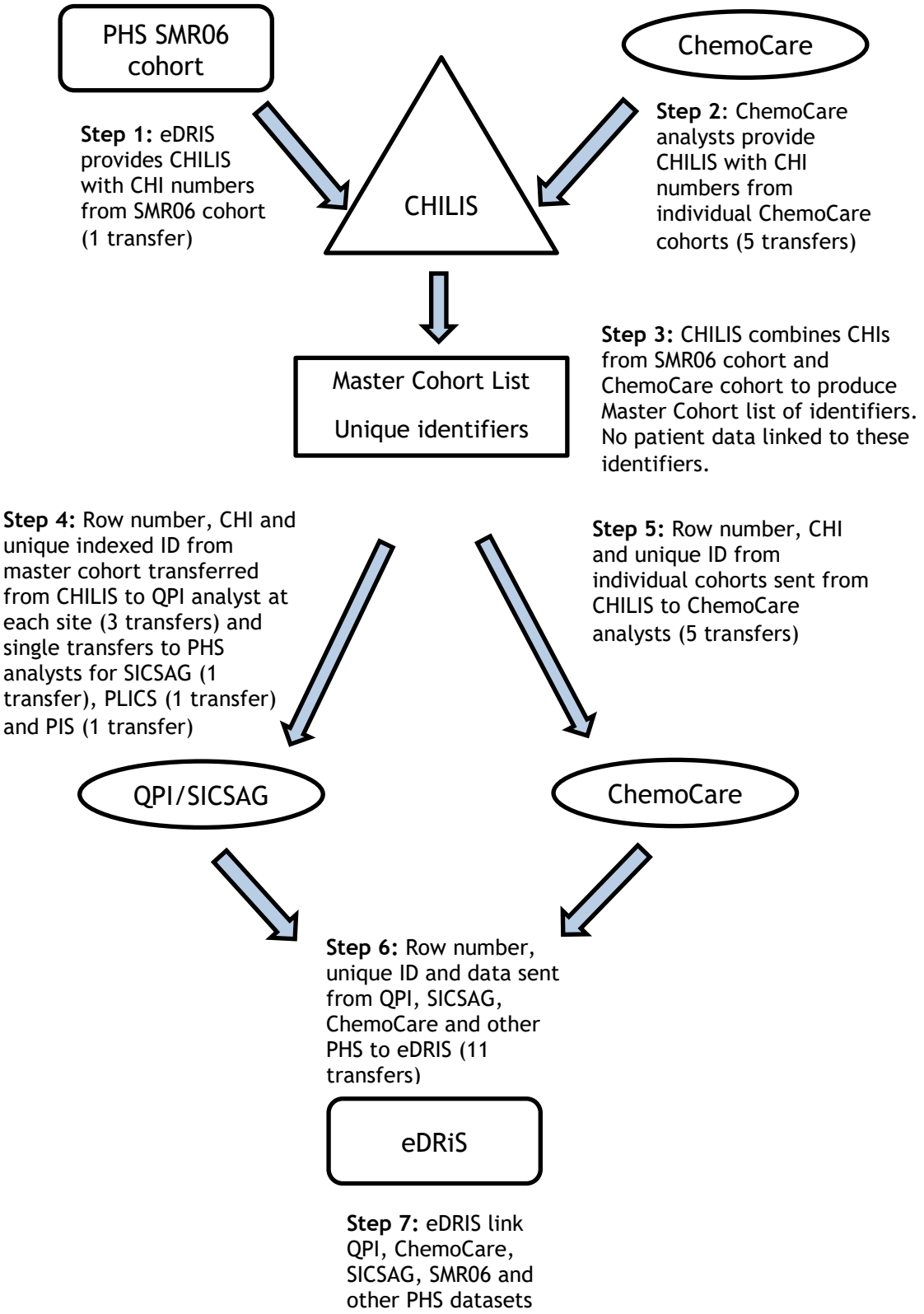


Figure 2-3 Cohort definition and transfer of datasets to PHS Abbreviations: PHS, Public Health Scotland; SMR, Scottish Morbidity Record; QPI, Quality Performance Indicator; CHILIS, CHI Indexing and Linkage Service; eDRIS, electronic Data Research and Innovation Service; CHI, Community Heath Index; SICSAG, Scottish Intensive Care Society Audit Group; PLICS, patient level information costing system.

Stage 2 involved initiating and continuing a dialogue and discussion with the relevant data controllers/data providers in the NHS Boards and other analytical teams in PHS. The primary researcher for this study (CH) played a major role in liaising with data providers from the five ChemoCare sites and three QPI locations, for the purposes of both projects (CORECT-R and this project).

The transfer of ChemoCare and QPI datasets to PHS required the use of a secure transfer platform. In total, 28 successful secure transfers were performed to transfer data from external data providers to PHS. The final master cohort contained information on all patients aged 18+ who had a CRC diagnoses between January 2006 and April 2018 in Scotland.

Stage 3

Each dataset that was to be linked and subsequently released to the research team for analysis was checked by eDRIS to confirm it matched the approved specification. Deterministic linkage of pseudonymised datasets was performed by the eDRIS team within their NSH using individual unique identifiers. Essentially each of the unique indexed identifiers supplied in Step 4 of Figure 2-3 was replaced with the master index in Step 7 so each patient had the same unique identifier across all datasets. The linkage process, with the number of patients per dataset, is outlined in Figure 2-4 below. This outlines all of the datasets that were linked by December 2020.

Stage 4

After linkage was performed, the pseudonymised dataset was transferred to the researcher-facing NSH. Access to data within the NSH was limited to the project team named on the most recently approved PBPP application. Prior to accessing the data, each named person demonstrated up to date, approved information governance training, and completed an eDRIS User Agreement.

The initial plan from eDRIS had been that the subset of data required and approved for this study would be identified by the primary researcher in the large project (CORECT-R) NSH space. This data would then be transferred by eDRIS to a separate NSH folder that could be accessed by the PhD student and supervisors. Due to

costs and time delays already incurred, a decision was made by eDRIS in November 2020 not to proceed with a separate NSH transfer and instead all analysis was performed in the CORECT-R project space. A requirement for working with this linked data was that all outputs had to undergo a disclosure-controlled release, which requires two eDRIS employees to check the outputs.

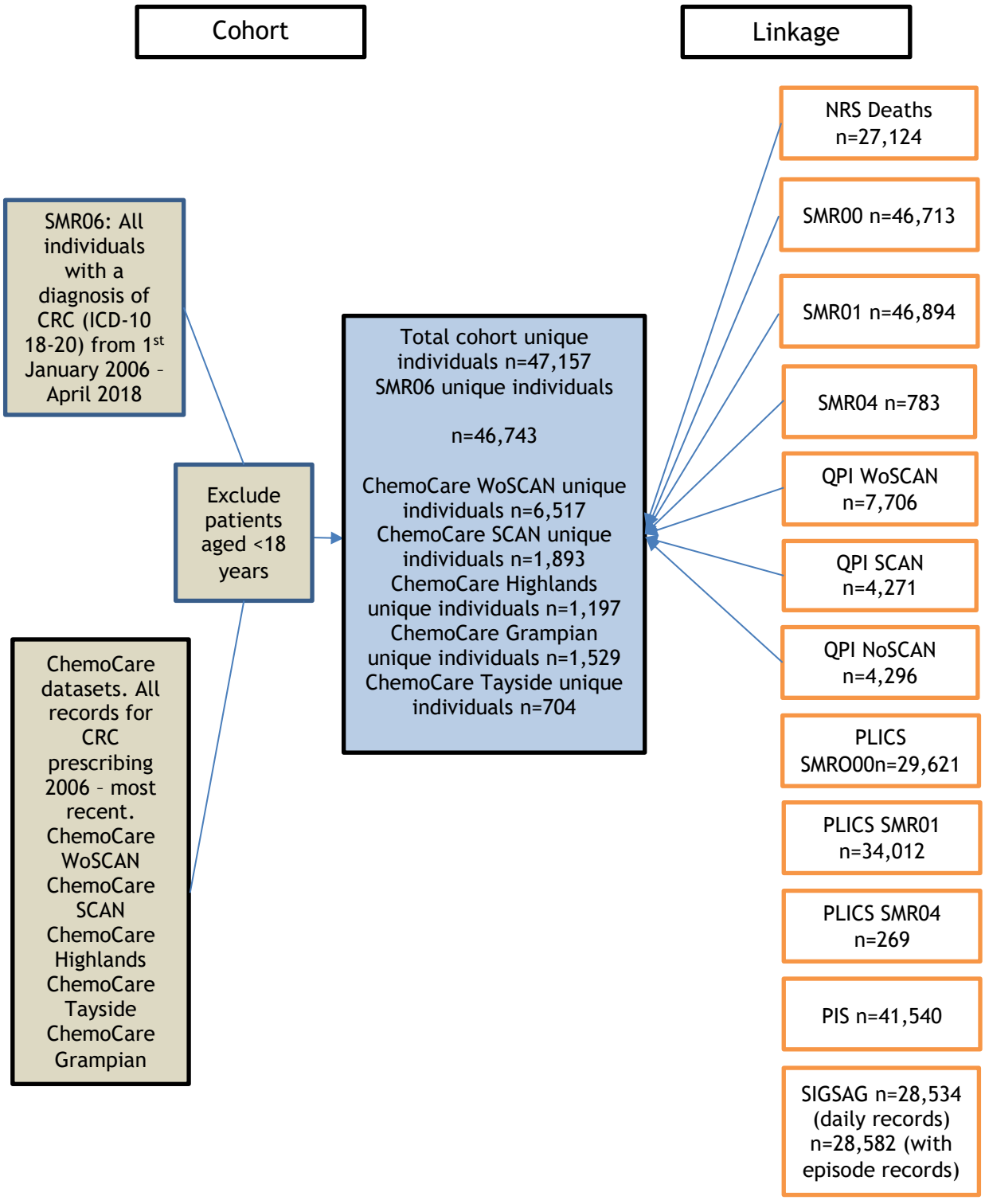


Figure 2-4 Datasets included in data release one and two Abbreviations: ICD, International Classification of Disease; CRC, Colorectal cancer; WoSCAN, West of Scotland Cancer Network; SCAN, South East Scotland Cancer Network; NoSCAN, North of Scotland Cancer Network; SMR, Scottish Morbidity Record; QPI, Quality Performance Indicator; NRS, National Registry Scotland; SICSAG, Scottish Intensive Care Society Audit Group; PIS, Prescribing Information System; PLICS, patient level information costing system.

Figure 2-5 summarises the steps involved in the process of national level data access and the times at which each step occurred.

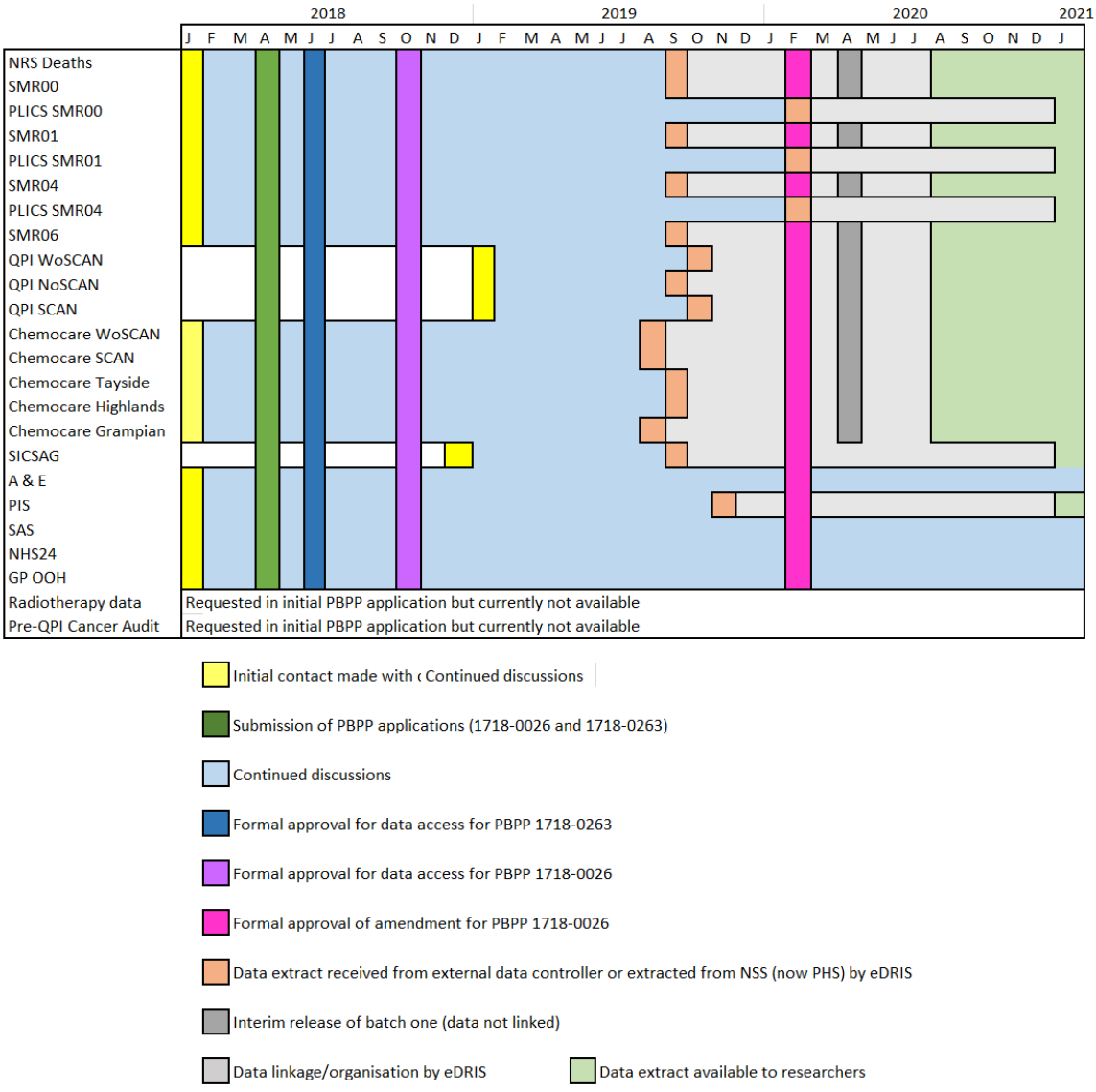


Figure 2-5 Timeline for transfer of datasets to PHS As of January 2020, 32 individual data files were available. There was also a demographics file that contained all patients in the master cohort which was provided to research team with release one of the data. PIS datasets were provided as nine separate data files, one for each year (2010-2019). ChemoCare Grampian and Highlands data were provided each as three separate files. ChemoCare Grampian provided an additional file with information regarding body surface area, height, and weight. SICSAG information consisted of two files (episodes and daily information). Abbreviations: WoSCAN, West of Scotland Cancer Network; SCAN, South East Scotland Cancer Network; NoSCAN, North of Scotland Cancer Network; SMR, Scottish Morbidity Record; QPI, Quality Performance Indicator; NRS, National Registry Scotland; PIS, Prescribing Information System; A&E, Accident and Emergency; SAS, Scottish Ambulance Service; PLICS, Patient Level Information Costing System; GP OOH, General Practice Out of Hours; SICSAG, Scottish Intensive Care Society Audit Group.

Estimated direct costs and resource use incurred during this process

An estimation of the costs and resources required to achieve data access are outlined in Figure 2-6. These costs were linked with the number of datasets being linked and, in particular, the number of datasets external to PHS that were used.

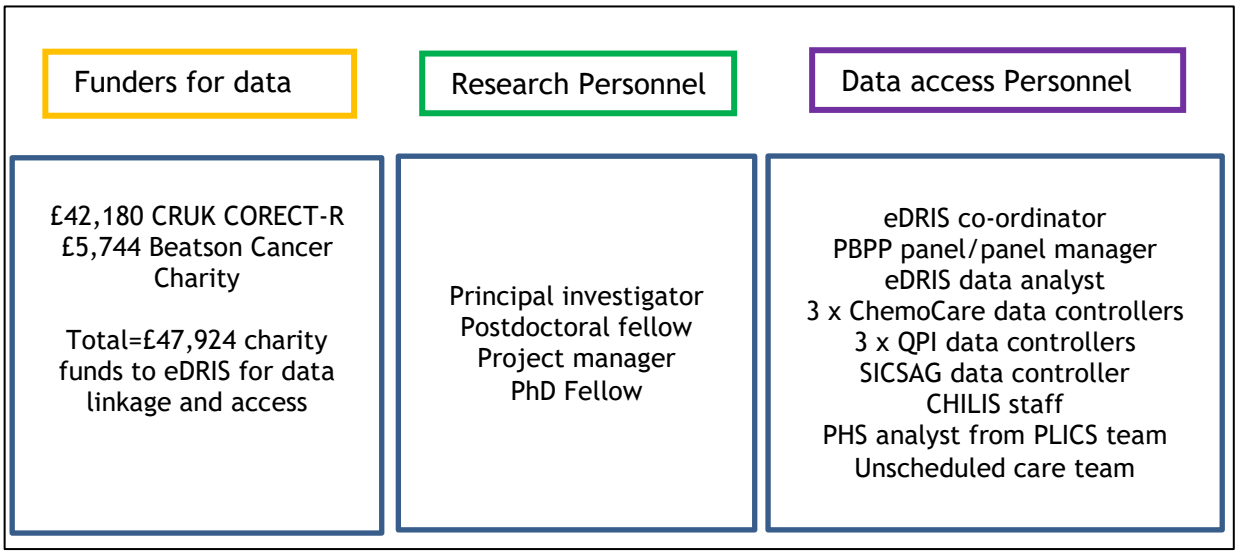


Figure 2-6 Direct costs and resource use Abbreviations: eDRIS: electronic Data Research and Innovation Service, PhD: Doctor of Philosophy, QPI: Quality Performance Indicators, SICSAG: Scottish Intensive Care Society, CHILIS: CHI Indexing and Linkage Service, PHS: Public Health Scotland, PLICS: Patient Level Information Costing System, PBPP: Public Benefit and Privacy Panel for Health and Social Care.

2.6.3 Description of the datasets used

Below is a general description of the datasets used for the purposes of the analyses reported in Chapters 7 and 8. An outline of which datasets and variables were used for each specific study are provided in more detail in each results chapter.

2.6.3.1 ChemoCare

ChemoCare is an electronic chemotherapy prescribing platform which is used for the majority of systemic anticancer therapy prescriptions in Scotland. There are five separate instances of ChemoCare used in Scotland: one each for the West of Scotland Cancer Network (WoSCAN) and the South-East Scotland Cancer Network (SCAN) and three in the Northern Cancer Alliance (NCA) (Grampian, the Highlands, and Tayside). Reports from ChemoCare for audit and research purposes are generated by software called CRYSTAL and compiled in Microsoft Excel®.

2.6.3.2 Quality Performance Indicators and Cancer Audit data

QPI data collection has existed in Scotland since 2013. There are typically 10-15 individual QPIs defined for each tumour type and these have been developed collaboratively by representatives from the three regional cancer networks, PHS, and Healthcare Improvement Scotland under the supervision of the National Cancer Quality Steering Group. NHS boards are required to report their activity against QPIs as part of a mandatory national cancer quality programme (122). Prior to 2013, a similar data collection process known as the Cancer Audit existed but unfortunately the availability of pre-2013 Cancer Audit data is not consistent across regional health boards.

2.6.3.3 Scottish Morbidity Record 06: Scottish Cancer Registry

SMR 06 collects patient level information relevant to the diagnosis and management of tumours. Cancer diagnoses are coded within the registry using the International Classification of Diseases, 10th edition (ICD-10). A patient may have two records within Registry if they have a diagnosis of more than one cancer. In 1997, a new electronic system for data capture was launched alongside an expansion of the variables included within the dataset. Information on stage and grade of tumour, as well as treatment information is now available. This includes an indication of if the patient received chemotherapy and/or radiotherapy and the intention of that therapy, but does not include information on the individual names or doses of systemic therapy or radiotherapy. Data is collected annually.

2.6.3.4 National Records of Scotland Deaths Data

Death record collection within National Records of Scotland (NRS) contains information on date, cause, and place of death.

2.6.3.5 SMR01: The General/Acute and Inpatient Daycase dataset

The SMR 01 dataset comprises patient level episode data on hospital inpatient and day case discharges from acute specialities in Scotland. Each patient hospital admission creates a new SMR01 record and an individual patient can have several records for the same admission if they transfer between hospitals, treating

consultant or facility (such as an intensive care unit admission). Data is collected continuously throughout the year.

2.6.3.6 Deprivation index and Charlson co-morbidity index

Additional information was provided as derived variables included within a number of the above above-named datasets. An indication of the socio-economic demographic of each patient, based on their residential post-code, was provided using the Scottish Index of Multiple Deprivation (SIMD). SIMD is calculated using a patient's postcode and a publicly available file, which has been developed by the Scottish ISD to identify over 6000 small area concentrations of multiple deprivation across Scotland in a consistent way. It combines 38 indicators of deprivation across seven domains: income, employment, health, education, skills and training, housing, geographic access and crime. The small data zones are grouped into bands (vigintiles, deciles and quintiles) with the first band, for example, quintile 1, representing the most deprived areas in Scotland. SIMD look up tables are updated approximately every four years. For the purposes of this analysis, SIMD quintiles were used based on SIMD codes from 2016.

Information on patient co-morbidity was provided for the national dataset as the indicators required to calculate the Charlson co-morbidity index within SMR-01 (123). The index was calculated by summing the total number of co-morbidities associated with hospital admissions (excluding cancer) from the year prior to diagnosis (QPI diagnosis) until the time of death or censoring. Quan weights (124) were applied to produce the final Charlson score for each patient.

For the GG&C analysis, datasets to perform this analysis were all received as Microsoft Excel® 2016 files and Microsoft Access® was used for data linkage. All datasets were converted to STATA data files for analysis. For the national analysis, datasets were all received in the NSH as Microsoft Excel® files and were converted to STATA data files for analysis.

2.7 Ethics and governance approvals

Ethical approval for using a survey to evaluate clinician practice was granted by the University of Glasgow Medical Veterinary and Life Sciences College Ethics Committee (project number 200180056).

Approval to access and use records from patients within GG&C health board for research purposes was granted from the local Caldicott Guardian. Separate approval was granted to access data collected for quality performance purposes for patients who received adjuvant chemotherapy in GG&C.

Approval from the Public Benefit and Privacy Panel for Health and Social Care (PBPP reference 1718-0263) was granted to perform the work reported in Chapter 8. The primary researcher (CH) was also listed on a separate PBPP application (PBPP reference 1718-0026).

Appendix 7 contains the approval letters for ethical approval that was specifically sought for the purposes of the work in this thesis.

The original ethical approval to conduct the SCOT trial and to collect and analyse information for the purposes of clinical and economic outcomes for the trial was granted by the West Glasgow Research Ethics (REC) Committee in January 2008 (REC reference number 08/S0703/136). Approval was sought from the SCOT trial management group (TMG) to use the data for the purpose of the work performed in this thesis.

2.8 Software used

The following versions of software were used throughout this study: NVivo® version 11 (125), Microsoft Excel® 2016 (126), Microsoft Word® 2016 (127), STATA® version 14 (128), Google® Charts® 2020 (129)(used in results Chapter 5 and 7), Online Surveys® (130).

3 Chapter 3: Systematic review of the literature on approaches to research impact assessment

3.1 Introduction

Despite the contemporary interest in research impact assessment described in Chapter 1, knowledge of how best to perform assessments and the infrastructure for, and experience in doing so, are lacking (8, 131, 132). This lack of clarity extends to cancer research. The enthusiasm regarding research impact assessment from cancer funders and researchers (133) has not been accompanied by instruction or reflection on which approaches would be suited to assessing the impact of cancer research specifically. Being cognisant of the discipline specific nature of impact assessment, and understanding the uniqueness of cancer research in approaching such evaluations, underpins the rationale for the study reported in this chapter.

In a 2016 survey of Australian cancer researchers, respondents indicated they felt a responsibility to deliver impactful research, but that evaluating and communicating this impact to stakeholders was difficult. In total, 80% of the respondents agreed that the researcher themselves should be contributing to doing an impact assessment, but acknowledged that this activity would take time away from research itself, teaching and writing. Respondents also suggested that the types of impact expected from research, and the approaches used, should be discipline specific (134). These results add weight to the rationale for this study, which was to consider approaches to impact assessment from a cancer research specific perspective.

As already discussed in Chapter 1, there is no single definition of what research impact encompasses, with potential differences in the evaluation approach depending on the definition. For the purposes of this study, the definition of research impact given by the RCUK is used (see Chapter 1 Section 1.2.4). This definition was chosen because it takes a broad perspective, which incorporates academic, economic and societal perspectives of research impact (135).

The aim of this study was to explore approaches to research impact assessment across research disciplines, identify approaches that have been used previously for cancer research, to explore whether these approaches are adequately capturing impact, and to consider which are relevant to cancer clinical trials. For the purposes of this study, cancer research included both basic science and applied research, research into any malignant disease, concerning paediatric or adult cancer, and including but not limited to studies spanning nursing, medical and/or public health elements of cancer research.

The study objectives were to:

(i) Identify existing literature reviews that report approaches to research impact assessment and summarise these approaches.

(ii) Use these same literature reviews to identify examples of cancer research impact evaluations, describe the approaches to evaluation used within these empirical examples, and compare them to those described in the broader literature.

3.2 Methods

This study was undertaken in two parts: (i) An overview of existing reviews of the literature on approaches to research impact assessment and (ii) A search of the reference lists of the reviews identified in part one to find empirical examples of the evaluation of the impact of cancer research.

An overview of reviews (discussed in Chapter 2) was performed because a scoping review of the literature had identified that a large number of reviews on the topic of approaches to research impact already existed. Rather than duplicate previous work, the aim of identifying and synthesising evidence from existing reviews was to provide a summary of the important approaches to impact evaluation used. The choice to include both systematic and narrative reviews in the study was purposeful because it was felt that both types of reviews (94) would make a useful contribution to understanding the important themes on the topic of impact evaluation and help to identify previous evaluations of impact that have been carried out.

As research impact assessment is a field that has not been extensively developed in oncology, it was felt that to perform a primary review of empirical studies that have evaluated the impact of cancer trials, or those describing the methodology to assess research impact within the field of oncology, was too narrow. There have been ad hoc examples of evaluating specific aspects of the impact of cancer trials, for example, how a trial result has been incorporated into guidelines (136) or how it has changed clinician's views (137) and practice (138-140). However, these studies are often not identified in the literature under the umbrella term of "impact" and it is rarely acknowledged how the outcomes sit within a wider analysis of all the potential impacts of that trial. A review of existing reviews was undertaken because of these anticipated challenges of conducting a primary review of empirical examples of cancer research impact evaluation, and to allow a critique of empirical studies in the context of lessons learnt from the wider literature.

3.2.1 Part I: Data sources and search strategy

For part one (overview of reviews), eleven publication databases and the grey literature (for example, reports not published in peer-reviewed journals) from January 1998 to May 2019 were searched to identify review articles that summarised approaches to research impact assessment.

The electronic databases searched were Medline, Embase, Health Management and Policy Database, Education Resources Information Centre, Cochrane, Cumulative Index of Nursing and Allied Health Literature, Applied Social Sciences Index and Abstract, Social Services Abstracts, Sociological Abstracts, Health Business Elite and Emerald. These were chosen in collaboration with a Beatson West of Scotland Cancer Centre specialist librarian (Lorraine MacLeod, LMacL) with the aim of identifying review articles that spanned across research disciplines. The search strategy specified that article titles must contain the word 'impact', as well as a second term indicating that the article described the evaluation of impact, such as 'model' or 'measurement' or 'method'. The search terms used, and an example of the search performed for the Ovid Medline database are outlined in Appendix 2.

The grey literature was searched using a proforma that provided a list of online grey literature sources. Keywords were inserted into the search function of websites and the first fifty results were screened. Title searches were performed by either a specialist librarian (LMacL) or the primary researcher (CH). All further screening of records was performed by the primary researcher (CH).

Articles were kept for final inclusion in the study by assessing each article against the following inclusion and exclusion criteria.

The following criteria were used to select relevant review articles for inclusion in the overview of reviews:

Inclusion criteria

"Review" or "overview" mentioned in abstract or methods

Review of methods to assess or evaluate the impact of research

Publication date between the years 1998 to 2019

Can include part of a document as long as that is a standalone review, for example a chapter in a PhD thesis or a literature review in a supplementary appendix

Exclusion criteria

Description of one framework for impact evaluation with no review of approaches to existing approaches to research impact evaluation

Primary, empirical examples analysing research impact

PowerPoint® presentations, visual or multimedia

Editorials, blogs, short opinion pieces, workshops, books

Instructions on how to use a specific impact assessment framework

Website address with no identifiable document

Documents focusing on one country (Reviews of methods from several different countries to evaluate research impact *are* permitted)

Articles only focusing on the *definition* of research impact

Articles only focusing on "health impact assessment" (not the same as the impact of health research)

Articles describing the development of a new framework with a brief mention of a scoping review in the methodology are excluded.

To meet these criteria, an article had to be a review of approaches to evaluate the impact of research. No restrictions were placed on the discipline, field, or scope (national/global) of research for this part of the study. Articles reviewing concepts and methodology of approaches to impact assessment were permitted, as well as reviews of empirical impact evaluations. If two articles drew primarily on the same review but contributed a different critique of the literature or methods to evaluate impact, both were kept. If a review article was part of a grey literature report, for example a thesis, but was also later published in a journal, the journal article only was kept. The reference list of final, included reviews were also searched to identify any additional, relevant articles.

3.2.2 Part II: Data sources and search strategy

For part two of the study, the reference lists from the literature reviews included in part one were manually screened (141) to identify empirical examples of assessment of the impact of cancer research using the inclusion and exclusion criteria outlined below. Summary tables and diagrams from the reviews were also searched using the words 'cancer', 'oncology' and 'trial'.

Inclusion criteria

Empirical examples of research impact assessment.

Research under evaluation is specifically:

Cancer research,

Clinical trials (one trial, multiple trials or a programme of trials), or

Cancer trials

Exclusion criteria

Primary evaluations of research impact when the research may include cancer research or clinical trials, but this is not the primary focus of the research or research programme being evaluated.

3.2.3 Data extraction and analysis

A data extraction form (see Chapter 2) was used to summarise information from the review articles identified. These lists were tabulated or presented graphically. A separate data extraction form (see Chapter 2) was used to summarise information from empirical examples of cancer research impact assessment. A narrative approach (142) was used to synthesise and describe the results from both parts of this review and to compare and contrast the approaches to cancer research with the approaches identified from literature reviews in part one.

3.3 Results

3.3.1 Search results

For part one, following an initial title screen, 800 abstracts were reviewed and 140 selected for full review. Out of 140 articles read in full, 27 met the inclusion criteria. A further 13 relevant articles were found through reference list searching from the included reviews (141), giving an overall number of 40 reviews for inclusion in part one of the study. For part two, 4,479 titles were screened and after removal of duplicates, 57 full articles were read and fourteen were deemed relevant. Figure 3-1 shows the search strategy for both parts of the study (93).

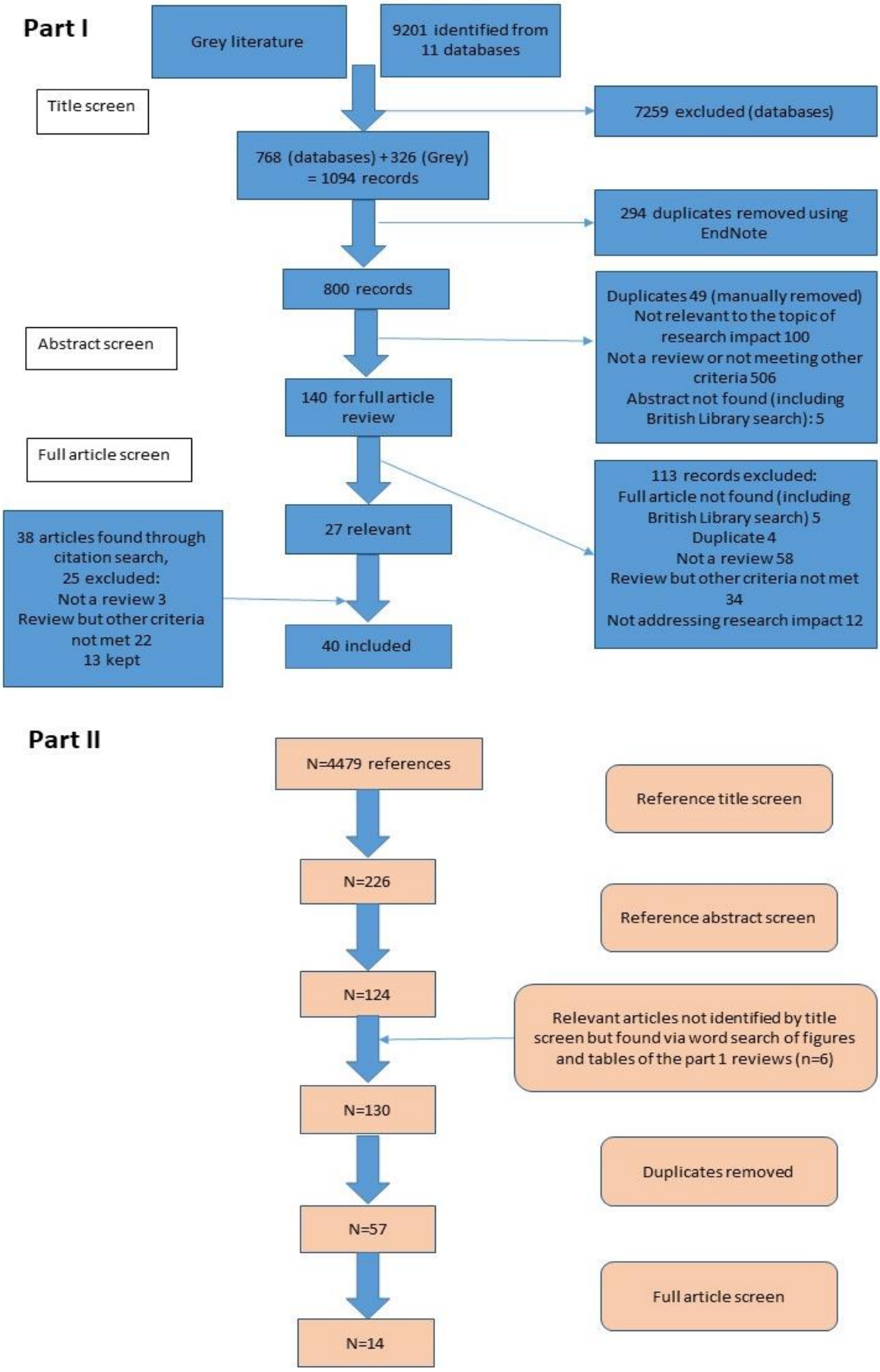


Figure 3-1 PRISMA diagram for review

3.3.2 Part One: Identification and analysis of literature reviews describing approaches to research impact assessment

3.3.2.1 Characteristics of included literature reviews

The characteristics of the forty reviews of the literature on approaches to research impact assessment that met the pre-specified inclusion criteria are outlined in Table 3-1. A large proportion (20/40; 50%) were written by primary authors based in the UK, followed by the USA (5/40; 13%) and Australia (5/40; 13%), with the remainder from Germany (3/40; 8%), Italy (3/40; 8%), the Netherlands (1/40; 3%), Canada (1/40; 3%), France (1/40; 3%) and Iran (1/40; 3%). All reviews were published since 2003, despite the search strategy dating from 1998. Raftery et al 2016 (143) was an update to Hanney et al 2007 (144) and both were reviews of studies assessing research impact relevant to a programme of HTA research. The narrative review article by Greenhalgh et al (145) was based on the same search strategy used by Raftery et al (143).

Table 3-1 Literature reviews identified in overview of reviews

ID	Author/Year /Location	Aims	Search methods	Timeline for literature search	Research discipline	Methods of data collection or analysis	Categories of impact	Frameworks for impact assessment	List of empirical examples of research impact assessment
1	Hanney et al 2003 (146). UK.	Review how health research is used in policy-making and the approaches to assess the policy impact of research.	Not detailed	NA	Healthcare	1	0	1	1
2	Buxton et al 2004 (100). UK.	Identify key studies that have estimated the economic value of the impact of health research to society.	Databases (11) and grey literature.	Unknown	Healthcare	1	0	0	1
3	Coryn et al 2007 (147). USA.	Describe, classify, and evaluate national models and mechanisms used to evaluate research in 16 countries.	Not detailed	NA	Not discipline specific	0	0	1	0
4	Hanney et al 2007 (144) (Chapter 2). UK	Review the literature describing the evaluation of the impact of programmes of health research.	Databases (13), citation analysis, expert consultation, advisory group consultation.	1990-2005	Healthcare	1	0	1	1
5	Brutsher et al 2008 (Part 1 of report) (148). UK	Present and discuss five key elements of research evaluation.	Not detailed.	NA	Not discipline specific.	1	1	1	0
6	Buxton et al 2008 (Chapter 2) (149). UK.	Review current practices of assessing the economic benefits of medical research.	Not detailed.	NA	Medical research	1	0	0	0

7	Boaz et al 2009 (150). UK.	Review methods for evaluating the impact of research on policy outcomes.	Databases (10 in total), web searches of 30 organisation websites, citation tracking, expert contacts.	1987-2007	Cross sector but specifically interested environmental research impact.	1	0	1	0
8	Marjanovic et al 2009 (151). UK	Historical overview of landmark studies in the research evaluation field to reflect on methodological developments.	Not detailed.	NA	Biomedical and health research.	1	0	1	0
9	Yazdizadeh et al 2010 (152). Iran.	Systematic review to identify methods used to assess the economic impact of healthcare research.	Databases (8), 21 relevant websites.	Unknown.	Healthcare	1	0	1	0
10	Banzi et al 2011 (153). Italy.	Review to identify the most common approaches to research impact assessment, categories of impact, and their respective indicators.	Databases (2), charity websites, citation screening.	1990-2009	Not discipline specific.	1	1	1	0
11	Hanney et al 2011 (154). UK.	Review of studies that have assessed economic impacts from health research in the field of nursing health research.	Databases (2), review of retrospective studies already known to the authors.	Unknown	Healthcare	1	0	0	1
12	Patel et al 2011 (155). UK.	Systematic review to identify indicators used to measure healthcare research performance.	Databases (4), citation screening.	1950-2010	Healthcare	1	0	1	0
13	Ruscio et al. 2012 (156). USA.	Evaluate 22 scholarly impact metrics.	Not detailed.	NA	Not discipline specific.	1	0	0	1
14	Bornmann et al 2013 (157). Germany.	Literature survey of existing research on practices employed in the assessment of societal impact of research.	Databases (2), internet search engines, citation screening.	Unknown	Not discipline specific.	1	0	1	1
15	Guthrie et al 2013 (158). UK	Identify and review frameworks in use for research evaluation, to identify the research evaluation tools applied to those frameworks to provide a guide to	Not detailed.	NA	Not discipline specific.	1	0	1	0

		developing a research evaluation framework that can be used in a range of circumstances.							
16	Smith et al 2013 (159). UK	Review the methods of assessing research impact relevant to academic promotion.	Databases (5), internet search engine, citation screening.	NA	Not discipline specific.	1	0	1	0
17	Carpenter et al 2014 (160). USA	Broad overview of widely available measures of academic productivity and impact using publication data and to highlight the uses of these metrics.	Not detailed.	NA	Science	1	0	0	0
18	Penfield et al 2014 (28). UK	Explore what is understood by the term research impact and provide a comprehensive overview of the literature to understand which methods and frameworks of impact assessment could be used for UK based impact assessment.	Not detailed.	NA	Not discipline specific	1	0	1	0
19	Milat et al 2015 (161). Australia.	Synthesise evidence that describes processes and conceptual models for assessing policy and practice impacts of public health research.	Databases (6)	1990-2013	Healthcare	1	0	1	0
20	Moed et al 2015 (162). The Netherlands.	Provide a broad overview of the wide array of metrics to assess research impact currently in use in academic and research.	Not detailed	NA	Not discipline specific.	1	1	1	0
21	Pollit et al (Appendix 2) 2015 (163). UK	Identify a wide range of potential impacts of research, investigate different ways of classifying impacts, produce a list of types and domains of impact.	Grey literature, academic literature focused on a limited set of key sources known to the study team.	Unknown	Not discipline specific	0	1	1	0
22	Thonon et al 2015 (164). France.	Identify indicators of impact that could be used to measure the output and outcome of medical research.	Databases, snowballing technique.	Unknown.	Biomedical research	1	1	0	0

23	Wouters et al 2015 (165). UK	Review literature of academic research looking at a range of impact indicators.	Not detailed	NA	Not discipline specific	1	0	0	0
24	Agarwal et al 2016 (166). USA	Provide a broad overview of evaluation metrics currently used in academic and research.	Not detailed	NA	Not discipline specific	1	0	0	0
25	Chikoore et al 2016 (167). UK	Explore the meaning of research impact, issues regarding how it can be evaluated, and challenges associated with assessment.	Not detailed	NA	Not discipline specific	1	0	1	0
26	Raftery et al 2016 (143). UK	Review published research studies on tools and approaches to assessing the impact of programmes of health research and specifically to update the previous 2007 systematic review (Hanney et al).	Databases (8), hand searching selected journals, citation screening, literature known to the research team, snowballing, bibliographic searches of other reviews and references.	2005-2014	Healthcare	1	1	1	1
27	Greenhalgh et al 2016 (145). UK	Review the strengths and limitations of six established approaches of measuring both the outcomes of research and the processes and activities through which this is achieved.	Search strategy based on Raftery et al 2016.	As above	Not discipline specific	1	0	1	1
28	Wimmer et al 2016 (168). USA	Review both traditional and novel impact evaluation tools, the impact metrics they calculate and to explore if and why these tools are relevant to the field of nursing research.	Not detailed.	NA	Healthcare	1	0	0	0
29	Bornmann et al 2017 (169). Germany	Review how impact is measured within science and beyond and the problems associated with impact evaluation.	Databases (3 in total) and other literature reviews.	Unknown	Science	1	0	0	0

30	Cruz Rivera et al 2017 (170). UK	Identify existing methodological frameworks used to measure healthcare research impact and to summarise the common themes and metrics in an impact matrix.	Databases (4 in total), internet search engines (including images) and communication with experts	Unknown	Healthcare	1	0	1	0
31	Deeming et al 2017 (171). Australia	List the stated objective for research impact frameworks, to identify existing frameworks and to evaluate whether these have the capabilities necessary to address pre-specified objectives.	One database and grey literature	2005-2015	Healthcare or medical	1	0	1	0
32	Peter et al 2017 (172). Canada	Identify approaches that have been used to understand the impacts of health research, identify ways that research impacts have been defined and measured, and provide recommendations for occupational science.	Traditional databases, author search from the assessment tools that were mentioned in the included reviews used as keywords to search traditional databases, reverse citation analysis and a forward citation search in the Scopus database.	No time restriction	Occupational science	1	0	1	0
33	Reale et al 2017 (173). Italy	Understand how impact assessment methods are used in social sciences and humanities and how far these approaches attempt to apply methods and instruments that take into account the distinctive features of this discipline.	Journal articles/databases, books, reports, working papers, CORDIC database, EU FP17 Flash-it project	2006-2012	Social sciences and humanities	1	1	1	0

34	Newson et al 2018 (174). Australia	Review the extent and nature of studies measuring the impact of health research on policy and compare forward and backward tracing approaches for assessment.	Electronic databases (5), references of included studies	1995-2016	Not discipline specific	1	0	1	1
35	Pedrini et al 2018 (175). Italy	Analyse the approaches to the assessment of healthcare research social impact with a focus on different stakeholders.	Databases (3)	2000-2016	Healthcare	1	0	0	1
36	Weisshuhn et al 2018 (176). Germany	Conduct a literature review to analyse how impacts of agricultural research are assessed.	Databases	2008-2016	Agricultural	1	1	0	0
37	Williams et al 2008 (14). UK	Systematic review of the evolution of research impact assessment approaches in Australia and the UK.	Public policy documents, newspaper commentary, academic literature	Unknown	Not discipline specific	0	0	1	0
38	Braithwaite et al 2019 (177). UK	Identify what is known about methods for assessing researchers' achievements for the purposes of producing a new assessment model.	Databases (all Web of Science)	2007-2017	Not discipline specific	1	0	0	0
39	Gomes et al 2019 (26) UK	Review empirical impact evaluation to understand the impact generated by publicly and charity-funded health research in the UK.	Databases, citation tracking, reference searching of included articles, hand searching of specific journals	2006-2017	Health research	1	0	1	1
40	Heyeres et al 2019 (178) Australia	Systematic review of studies that used a case study approach to assess research impact.	Databases (11), reference lists of impact case studies identified	2000-2018	Not discipline specific	1	0	1	1

Approximately half of the reviews identified (19/40; 48%) described approaches to evaluate research impact without focusing on a specific discipline and nearly the same amount (16/40; 40%) focused on evaluating the impact of health or biomedical research. Two reviews examined approaches to impact evaluation for environmental research and one focused on research within the social sciences and humanities. Finally, two reviews provided a critique of impact evaluation methods used by different countries at a national level (147, 148). None of these reviews focused specifically on cancer research.

Twenty-five reviews (25/40; 63%) specified search criteria and eleven of these included a PRISMA diagram. The articles that did not outline a search strategy were often expert reviews of the approaches to research impact assessment and the authors stated they had chosen the articles in their reviews included based on their prior knowledge of the topic. Most reviews were found by searching traditional publication databases, however seven (7/40; 18%) were from the grey literature. These included four reports written by an independent, not-for-profit research institution (Research and Development (RAND) Europe) (148, 149, 151, 158), one literature review which was part of a PhD thesis (167), a literature review informing a quantitative study (163), and a review that provided background information for a report to the UK government on the best use of impact metrics (165).

3.3.2.2 Key findings from the reviews: Approaches to research impact evaluation

(i) Categorisation of impact for the purpose of research impact assessment

Nine reviews attempted to categorise the type of research impact being assessed according to who, or what, was affected by research, and how they were affected. In Figure 3-2, colour coding was used to identify overlap between impact types identified in these nine reviews to produce a summary list of seven main impact categories.

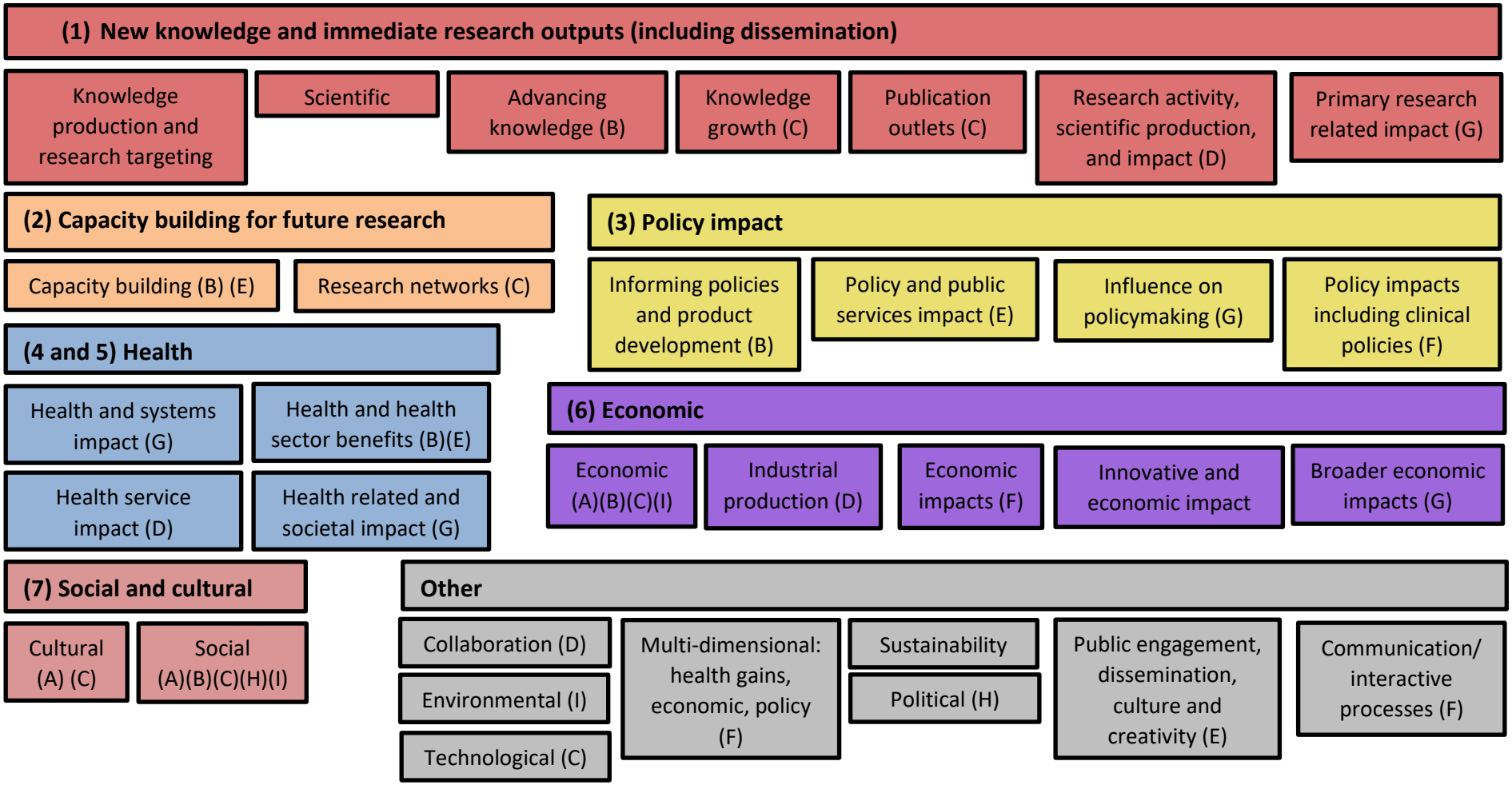


Figure 3-2 Categories of impact identified from literature reviews of approaches to research impact assessment. The 9 literature reviews have each been denoted with a letter A-I in this figure. A:Brutscher et al (2008) (148) B:Banzi et al (2011) (153) C:Moed et al (2012) (162) D:Pollitt et al (2016) (Supplementary material) (163) E:Thonon et al (2015) (164) F: Raftery et al (2016) (143) G:Cruz Rivera (2017) (170) H:Real et al (2017) (173) I:Weisshuhn (2018) (176).

The first category of impact refers to the immediate knowledge produced from research and the second focuses on the contribution research makes to driving innovation and building capacity for future activities within research institutions. The former is synonymous with the academic impact of research, and applied to cancer trials specifically would traditionally refer to the results of the clinical trial and resultant publications. The latter may refer to the act of securing funding to perform future research, providing knowledge that allows development of later phase clinical trials, or training cancer researchers.

The third category identified was the impact of research on policy. Three of the review articles included in this overview specifically focused on policy impact (146, 150, 174). In their review, Hanney et al (146) suggested that policy impact (of health research) falls into one of three sub-categories: impact on national health policies from the government, impact on clinical guidelines from professional bodies, and impact on local health service policies. Cancer clinical trials are often cited in clinical guidelines written by medical professional bodies and used as the evidence to support suggested guidance. Examples of such guidelines include those published by the National Institute for Health and Care Excellence (NICE) in the UK, or the National Comprehensive Cancer Network (NCCN) in the USA. Cruz Rivera and colleagues, in their review (170), specifically distinguished the impact of research on policy making from impact on clinical guidelines; they listed the latter under health impact. This shows that the lines between categories are subjective and will often blur.

Impact on health was the next category, and several of the reviews differentiated impacts on the health sector from health gain. Both of these impact categories will be relevant when assessing the impact of cancer clinical trials, given that cancer is a major burden for both healthcare systems and the patients they treat. Economic impact of research was the sixth category. Relevant to cancer trials research, there is likely to be close overlap between healthcare system and economic impacts because of the high cost of cancer care for healthcare services globally.

With regards to the economic impact of research, a key study identified from the previous literature was the review article authored by Buxton et al (100). The authors searched the literature for examples of studies that estimated the

value of the economic return on investment in health research and found four main approaches to perform these assessments, which were referenced in later reviews by other authors (149) (143, 152, 154). The four approaches were (i) measuring direct cost savings to the health-care system, (ii) estimating benefits to the economy from a healthy workforce, (iii) evaluating benefits to the economy from commercial development and, (iv) measuring the intrinsic value to society of the health gain from research. In a later review, (149) Buxton et al added an additional approach of estimating the 'spillover' contribution of research to the Gross Domestic Product (GDP) of a nation. The term 'spillover' is used by economists to describe how an organisational investment will not only benefit the organisation in question, but also other organisations in the same or other sectors of the economy, in the same or within other countries (179).

The final impact category identified from these literature reviews was social and cultural impact. This term social impact was commonly used in a specific way to refer to research improving human rights, well-being, employment, education, and social inclusion (153, 170). Two of the reviews which included this category focused on the impact of non-health related research (social sciences and agriculture), indicating that this type of impact may be less relevant, or less obvious, for health related disciplines such as oncology. A related term, societal impact, was used in a distinct way in the literature to describe any wider impact from research that is external to traditional academic benefits (157, 175). Lastly, other categories of impact identified that did not show significant overlap between the review articles included technological, environmental and political impacts, amongst others.

(ii) Methods for data collection and analysis

In total, thirty-six (36/40, 90%) of the reviews of approaches to research impact assessment discussed methods to collect or analyse the data required to conduct an impact evaluation. The common methods described, with strengths and weaknesses of each approach, and how they may be applied to assessing cancer trial impact are shown in Table 3-2.

Table 3-2 Methods for data collection and analysis for the purposes of research impact assessment

Method of data collection	ID number of review (from Table 1)	Explanation	Advantage	Disadvantage	Relevance to cancer clinical trials
Documentary analysis	1,4,7,10,14,15, 16,18,19,22,25, 26,27,34,36,39, 40	Umbrella term referring to the use of documents and data to analyse research impact (See (146) Additional File 1: Elements of a protocol for documentary analysis.)	Time efficient Likely to be cost-effective. Can be applied to a wide range of sources, e.g. publications, guidelines, conference proceedings, reports.	No standard methodology for analysis. Relies on the quality of the primary document.	Pre-existing documents/data that could be analysed to assess cancer trial impact may include national/international clinical guidelines, local treatment protocols, patient facing websites, funding information, drug or device patents, information on drug sales and research publications where the study was performed using information from the clinical trial in question.
Surveys	1,4,5,7,8,14,15, 16,18,26,27,31, 34,36,39,40	Includes online, telephone and paper/postal surveys.	Can be used for a wide range of stakeholders and across disciplines. Often cost effective. Can identify areas to focus on in an interview.	Relies on robust response rate and access to respondents Burdensome to complete. Decreased accuracy if incomplete data. Response bias	This method could be used to survey users of clinical trial information/results, such as clinicians, patients and policy makers. Alternatively, a survey could be used to ask the clinical trialists what impacts they perceive have arisen from their clinical trial.
Interviews	1,4,5,7,10,15,16, 19,25,26,27,32,34, 35,36,39,40	Can be structured, semi-structured or unstructured and open-ended. Can be conducted face-to-face, over the telephone or via video link. (See (146) Additional File	Provide personal perspectives from individuals e.g. researchers or users of research. Allow in depth analysis.	Time consuming. Interviewers need to be skilled so as not to introduce bias. Results may not be generalisable.	As for surveys, these could be directed to the direct users of clinical trials, or the trialists. Interviews are more likely to be useful if the aim of the assessment is to have an in depth

		2: Draft interview schedule for assessing research utilisation in policy-making).			understanding of how trial information/results are used, or, if interviewing trialists, to get a broad overview of potential impacts that are not pre-defined, as would be more common when using a survey.
Case studies or impact narratives	1,4,5,7,8,9,10, 15,16,18,19,25,26, 27,29,32,33,34, 35,39,40	A narrative description of the impact of research. Other methods may have been used to collect data that is used within the impact narrative.	Narratives often offer a wide perspective of research impact. Can provide evidence of a pathway to and a context for the research impact described.	Time consuming and costly. Case study selection may be biased towards only positive or high achieving research examples.	Individual cancer clinical trials or a small collection of cancer clinical trials assessing the same research question are good candidates for a case study or impact narrative, as usually the trial is developed with a specific, applied research question in mind.
Bibliometrics (including citation analysis)	5,7,8,10,12, 13,15,17,18,19, 20,22,23,24,25,26, 27,28,29,33,34, 35,38,39,40	Quantitative analysis of research activity, usually in the form of journal publications and citations.	Low cost and burden	Focus on outputs Not necessarily a measure of impact. Discriminate against researchers who have been active for less time. Not comparable between research disciplines. Often only considers citations in publication databases and does not include other data sources such as books or reports. Open to gaming.	This method would include citation analysis of publications describing cancer clinical trial results. This is an approach/metric that focuses only on academic impact and does not evaluate cancer trial impact as per the wider definition. Bibliometric software can be adapted to analyse other data sources, such as policy documents and non-academic websites, which may provide a better indication of wider impact/use of trial results.
Alternative metrics	17,18,22,23,25, 26,28,29,33,38	Quantitative analysis of research activity, usually based	Captures different routes of research dissemination	Disadvantages similar to bibliometrics.	Examples of software to carry out this approach to assessment are ‘Almetrics®’ or ‘Plum metrics®’.

		on non-academic, internet-related citations.	compared to traditional bibliometrics. Can be analysed in an automated fashion with a high degree of accuracy.	Mentions on social media or the lay new media may be more an indication of dissemination and influence than impact. Open to manipulation.	These are tools that track and collect online activity relating to publications such as those which report the results from cancer clinical trials. They could be applied to one or several trials (or the outputs from a specific trialist). Similar to bibliometrics, these scores will not describe the wider impact of a cancer trial, but they may give an indication of how and by whom the original publication is being used and read.
Peer review	1,5,7,12,14,15,16, 19,23,25,29,31,33, 35,38	Relies on advisory service by experts in the field to assess the impact of the research in question. Material can be reviewed by a group of peers (allows discussion) or individuals.	Credibility within the academic community. Can offer expert feedback for future improvement, flexible. Can be conducted at any time during an impact analysis.	Subjective/not always transparent. May be costly, requires facilitation, can be slow. Peers may be expert in their field but may not be expert at assessing research impact. Impractical to assess broad research area given peers generally experts in one field. Issues with time/cost and reporting bias around how information presented for review.	Peer/expert review could be used to assess the impact of a cancer clinical trial or group of trials. The limitation is that any expert review will be biased by the individual's own research expertise and understanding of research impact. This type of analysis could be guided by using an impact framework to make the assessment more objective.
Economic evaluation strategies	2,4,5,6,8,9,10, 11,12,15,18,20, 25,26,31,36,39	An umbrella term that incorporates many methods e.g. cost-benefit analysis, cost-utility analysis, cost-	Offers an estimation of the return on research investment which may be especially useful to funders.	Challenging to monetise broad impacts such as health and to account for	This method of impact assessment may be particularly relevant for funders of cancer trials. There is not one

		effectiveness, proportion of GDP.	A single monetary figure for the return on a specific research project means that a comparison with the return on the benefits from other research projects may be possible.	all research costs and benefits. Rather than the final monetary value, important to detail the methodology for calculation of the return on investment to allow comparison with other impact assessments.	recognised approach to evaluate the economic impact of cancer clinical trials and the impact may be positive or negative depending on the stakeholder perspective, for example, a pharmaceutical company (sales of drug) versus a public health service (cost of buying a drug). There may be economic benefits relating to the implementation of clinical trial results, but there may also be benefits if the clinical trial leads to future research investment/patients/drug sales.
Using a scale	1	Using a scale to assess the extent of research use within a certain category. For example, using a scale to report the level of research utilisation in policymaking. (See (146)Additional File 3: Draft scales of the level of research utilisation in health policy-making).	Relatively easy to carry out.	Disadvantages of any quantitative metric. A scale provides a superficial indication of the use of research findings and may lack context. Problems with ensuring a fair comparison between research items could be decreased by using the same person or same team to score the research projects.	A scale could be applied to cancer clinical trial but it is unlikely to provide worthwhile insight into the wider impact of a trial. A scale may be used in conjunction with other approaches to impact evaluation.
Benchmarking	5,14,	The act of comparing metrics, usually bibliometrics, in order to compare impact from different research studies.	May be useful on an institutional level as a tool to encourage improved research productivity.	Used alone this is not a measure of "impact". The benchmarking output requires careful interpretation in context.	Benchmarking has important limitations and should only be used to assess and compare clinical trials if the metric being used is comparable between trials, which will often not be the

					case. If the metrics being used are only bibliometric assessments of research publications, these are not a satisfactory measure of trial impact.
Workshop or focus group	7,34,35,40	Discussion amongst a group of people, usually facilitated by a researcher.	May be used as a tool to evaluate the impact of research on a group of stakeholders.	Time and cost intensive to run and analyse.	As for interviews, this approach could be used to ask research users, such as patients, or researchers, regarding their use of clinical trial information or their perception of the impact of a single or group of cancer clinical trials. Compared to interviews, individuals will be able to share and discuss their ideas with others. Compared to surveys, the topics for discussion and information gathering may deviate from a pre-defined list. In this way, unexpected impacts related to cancer trials may be discovered and explored.
Literature review or meta-analysis	7,16,26,35,36,40	An overview of the literature.	Usually straightforward to perform. May be used to understand the impact of current research in the context of other evidence on the same topic.	Measure of academic impact only.	Cancer clinical trial results will often be reported in academic journals and therefore will be amenable to evaluate via a literature review or meta-analysis. These are approaches focusing on scholarly impact only.
User or expert testimony	7,8,16,18	A statement from the user of the research or an expert in the field, which describes the	Straightforward to collect and demonstrates the impact of research directly from the stakeholder perspective.	One perspective only.	Applied to a cancer trial, the user or expert could be a patient or clinician. It is important to recognise the reason the

		impact of the research from their perspective.			testimony was given, and explore any bias, for example, was the expert involved in development of the cancer clinical trial.

A mixed methods approach to impact analysis was often advocated in these review articles, in particular the triangulation of surveys, interviews (of researchers or research users), and documentary analysis (144, 146, 150, 174). A large number of reviews cautioned against the use of quantitative metrics, specifically bibliometrics, alone (28, 146, 155, 156, 159, 160, 165, 166, 177, 180). Concerns were that these metrics were not designed to be comparable between research programmes (162), their use incentivised researchers to focus on quantity rather than quality (156), and they could be gamed and used in the wrong context to make decisions about researcher funding, employment and promotion (180) (159, 160).

Several reviews explained that the methods for data collection and analysis chosen for impact evaluation depended on the unit of research and the rationale (4 As, see Chapter 1) for the analysis (148, 150, 151, 154, 158, 161, 176). Regarding cancer research, the unit of analysis could be, for example, a single clinical trial or a programme of trials, research performed at a cancer centre, or research funded by a specific institution or charity.

(iii) Using a framework within a research impact evaluation

Applied to research impact evaluation, a framework (see Table 1-1 Chapter 1) provides a way of collecting and organising data, facilitating a more objective and structured evaluation than would be possible with an *ad hoc* analysis. In total, twenty-seven (68%) articles discussed the use of a framework in this context. Table 3-3 lists and provides a short description of the frameworks mentioned in three or more of the included reviews. Although several of the frameworks identified were designed to assess the impact of health research, none were specifically developed to assess the impact of cancer research. Table 3-3 also outlines which of these frameworks may be relevant to assessing the impact of cancer clinical trials.

Table 3-3 Examples of frameworks for research impact assessment Abbreviations: REF, Research Excellence Framework; RAE, Research Assessment Exercise; SIAMPI, Social Impact Assessment Methods for research; RIF, Research Impact Framework; RQF, Research Quality Framework; ERA, Excellence in Research Australia; MORIA, Measurement of Research Impact and Assessment; SEP, Strategy Evaluation Policy.

Framework	Main level intended for evaluation	Description	Relevance to cancer clinical trials
Payback Framework	Various	Developed at Brunel University in 1990s to evaluate the impact of health services research.	This framework has been applied to health research and could be used to assess the impact of a cancer trial or a programme of trials.
Social Impact Assessments Methods for research (SIAMPI) and other frameworks	Initial case studies aimed at a centre/institution level of assessment.	Developed through a collaboration between the UK's Economic and Social Research Council and researchers in the Netherlands. The focus is on social impact and there is an assumption is that "productive interactions" between researchers and stakeholders are important. The aim of this framework is learning from research impact evaluations rather than accounting or judging impact.	Applied to cancer trials, an assessment could be made of direct, indirect and funding productive interactions. Examples of direct impacts could include use of cancer trial results by clinicians or patients, indirect impacts could be citation in policy guidelines, and funding interactions could be increased funding for a trials unit, or for future trials that occurs because of the success of a previous cancer clinical trial. If using this framework, the aim of the assessment would be to learn about impact from previous trials to maximise trial impact going forward.
Monetary/economic framework	Various	Any method that attempts to evaluate the opportunity costs of research and its outputs.	Several review articles referred to using an economic framework but it is not clear how these could be applied to a cancer clinical trial. This is a gap in the current literature.
Research Impact framework (RIF)	Individual researcher or project	Developed by researchers at the London School of Hygiene and Tropical Medicine for healthcare researchers. Includes four categories of impact with sub-categories and indicators within each area. Based initially on a review of the literature and other assessment exercises and developed by interviewing researchers and applying the categories to research projects at their centre.	A framework developed for health research, with categories and indicators of impact that are relevant, although not specific to, cancer clinical trials. This framework could be used at the level of evaluating a single clinical trial.

UK REF and RAE	National frameworks to evaluate institutions	Introduced by the UK government in 2014 to allocate funding to universities. The RAE was the previous national assessment effort by the UK government used prior to the REF; impact was not explicitly included.	These frameworks were intended for a national funding allocation process. Alone, they could not be applied to a cancer trial because the approach relies on scoring and comparing many different impact case studies in order to allocate funding. The method of using a case study (see methods above) used within the REF framework would be applicable to a cancer trial but it is not clear how the scoring method used in the REF would compare between trials.
Canadian Academy of Health Sciences (CAHS)	Various	Developed by a panel of experts to provide a framework for assessing Canadian healthcare research and based on the Payback Framework.	There are four different versions/pillars of this framework (biochemical research, clinical research, health services research and public/population health research), with the applied clinical research framework being most relevant to cancer trials.
Australian national frameworks RQF, Excellence in Research Australia (ERA), Measurement of Research Impact and Assessment (MORIA)	National	The aim of the RQF was to use it to assess the quality of publicly funded research using case studies and peer review to assess impact. The RQF was never employed. Instead, the ERA was introduced in Australia, with a focus on quantitative metrics. First used in 2010, then 2012, 2015 and 2018. MORIA was developed for use at the grant review stage.	These are national assessment exercises, designed to evaluate the impact of a large body of research at an institutional level. They are therefore less relevant to assessing individual cancer trials, but trials may be used as examples of research within these assessments.
Weiss Logic model	Various. Developed to assess medical research.	Developed to assess medical research by psychiatrist Anthony Weiss.	This framework focuses on implementation of clinical research findings and therefore is most relevant to later phase cancer clinical trials. It is also not clear how the framework would apply to any form of research that did not have findings that could be readily implemented into practice, for example, early phase clinical trials or trials that do not meet their primary endpoint.
Netherlands Royal Academy of Arts and Sciences Strategy Evaluation Policy (SEP)	National, institution or programme level. Uses self-evaluation and intermittent external review of the institution	Used to assess research performance generally, not just the impact of research. There is overlap with ERiC. SEP is the national assessment process but there was initially no framework or specific methodology specified to carry out the assessment; the ERiC was used for this purpose.	This framework has been formulated to evaluate research units, rather than specific research projects. It could therefore by applied at the level of a cancer clinical trials unit. The framework relies on a group of experts to assess a research

	at set intervals. Overlap with the SIAMPI approach.		unit's research over the previous 6 years. This framework has similarities to the assessments performed by cancer research funders for their clinical trials units (for example the quinquennial review of clinical trials units undertaken by CRUK).
Lavis exchange model	Various	A conceptual model describing producer push, user pull, and the exchange model.	This model focuses on knowledge exchange as a technical exercise that links research to action. The model describes the theory of how knowledge exchange occurs and is limited in the guidance offered on how to practically evaluate the impact of research. The theory of the exchange of knowledge between trialists and cancer trial users, such as clinicians or patients, could be used to guide data collection within another framework.
Research Utilisation ladder	Various. Could be used at the project or individual researcher level.	Focuses on the role of the researcher in creating impact from their research. The ladder consists of six stages from the transmission of research results to those results being used by others in a different context to that of the original research.	This framework focuses on the researcher/clinical trialist and is most suitable to use if assessing the efforts made by trialists to increase the impact of their clinical trial, rather than providing a way to evaluate the overall impact of the trial in question.
HTA Quebec model	Programme of HTA research.	Developed to assess the impact of research for the HTA programme in Quebec. Effectiveness of research is assessed by the ability to impact on decision makers. Uses case studies, documentary analysis, and interviews.	This framework has been developed for health technology assessment purposes. Although it could be applied to assess the wider impact of cancer clinical trials, this was not the intended use of the framework.
Becker model	Focused at the research study level.	Framework for tracking diffusion of research outputs and activities to locate indicators of impact that demonstrate evidence of biochemical research impact. The pathways of diffusion are: advancement of knowledge, clinical implementation, community benefit, legislation and policy, and economic benefit.	This framework was developed to assess biomedical research generally. It is suitable for both basic and applied research. Although not specific to cancer clinical trials, the main strength is the list of impact indicators which could be used to assess the impact of a clinical trial.
Banzi	Various	This approach to impact evaluation is based on their 2011 umbrella literature review. They summarise indicators of impact under five main headings: advancing knowledge, capacity building, informing decision making, health	There is overlap with the CAHS framework, from which the categories and indicators of impact are developed. These are not specific to cancer clinical trials, but could be applied to trials either

		benefits, broad socio-economic benefits. The impact categories and relative indicators are adapted from the CAHS framework.	using the Banzi publication or the original CAHS framework.
Balanced scorecard	Various	A strategy performance management tool that is not specific to research impact. Focuses on financial dimensions, customer, business process, learning, and growth.	This is more appropriately applied at the organisational level, rather than the individual clinical trial level.
Canadian Institutes of Health Research framework (CIHR)	Various	Developed in 2005 by Canadian and international experts and based on the Payback framework. Pre-dated and provided a basis for the CAHS model.	The CAHS model is a more up to date version that could be applied to cancer clinical trials (clinical research pillar).
Program Assessment Rating Tool (PART)	Research programme level	This approach was developed for evaluating the research performance of all USA government research programmes during the time of the Bush administration (2003). It uses a survey format and focuses on impacts and efficiency.	This is most suited to assessing research at the programme level, and therefore would be most relevant to a programme of cancer clinical trial work. The framework focuses on the strategy of the programme and is not specific to assessing the wider impact of research.

The most frequently described framework was the Payback Framework, developed by Buxton and Hanney in 1996 (181), and many of the other frameworks identified reported that they were developed by adapting key elements of the Payback framework. This framework was originally developed to evaluate the impact of health services research. It consists of a modified logic model that explains the process of how impact from research occurs, alongside five categories of impact. A logic model is a graphic or description that represents the theory of how the critical aspects of an intervention are arranged to show how the intervention produces change (182). Figure 3-3 is a diagrammatic representation taken from the original manuscript describing the Payback Framework (15). The authors explain that the Payback Framework is a research tool that can be used to inform surveys, interviews, or documentary analysis and may be particularly useful in structuring data collection when the impact of more than one research project or programme is being assessed to help with consistency and communication of the results of the evaluation.

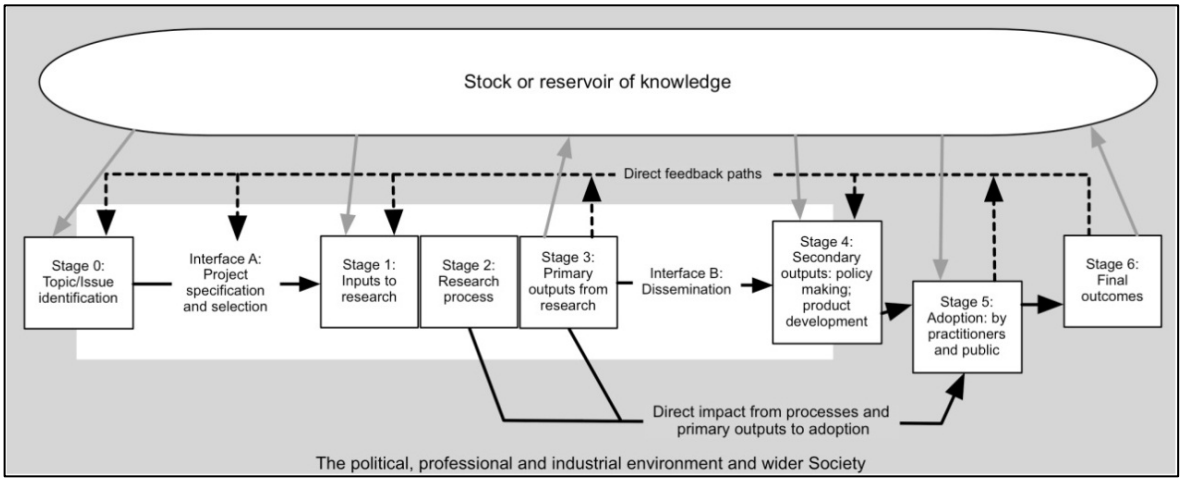


Figure 3-3 The Payback Framework logic model Figure taken directly from (15).

In the original Payback Framework, these impacts were described as benefits. However, iterations of the framework by other authors, for example by Klautzer et al (183) for use in the social sciences, have changed the term ‘benefits’ to ‘impacts’ in order to signify that recognising both the positive and negative effects from research is important.

(iv) Additional findings from the included reviews

The challenges of research impact evaluation were commonly discussed in these reviews. Several mentioned that the time lag (169) (28, 149, 152, 157, 158, 161, 170, 172) between research completion and impact occurring will influence when an impact evaluation is carried out; too early and impact will not have occurred, too late and it is difficult to link impact to the research in question. This overlaps with the challenge of attributing impact to a particular piece of research (28, 100, 151-153, 157, 158, 161, 170, 173, 175). Many authors argued that the ability to show attribution in an assessment was inversely related to the time since the research was carried out (28, 149, 150, 158, 172).

3.3.3 Part Two: Empirical examples of cancer research impact evaluation

3.3.3.1 Study characteristics

Fourteen empirical impact evaluations relevant to cancer research were identified by searching the reference lists of the review articles. These studies were published between the years 1994 to 2015 by primary authors located in the UK (7/14; 50%), USA (2/14; 14%), Italy (2/14; 14%), Canada (2/14; 14%), and Brazil (1/14; 14%). Table 3-4 lists these studies with the rationale for each assessment, the unit of analysis of cancer research evaluated, the main findings from each evaluation, if the research included cancer trials and/or how the approach may be relevant to cancer trial assessment.

Table 3-4 Empirical examples of cancer research impact assessment Abbreviations: QALY, Quality adjusted life year; UK, United Kingdom; USD, United State dollars; GDP, Gross Domestic Product; SIGN, Scottish Intercollegiate Guidelines Network; NICE, National Institute for Health and Care Excellence; BBC, British Broadcasting Agency; USA, United States of America; CAHS, Canadian Academy of Health Sciences; NBCF, National Breast Cancer Foundation; CRC, colorectal cancer.

Primary author/ year/ location	Unit of analysis	Main reason for assessment	Categories assessed	Methods used	Frame work	Main findings	Strengths	Limitations	Relevance to cancer clinical trials
Brown ML et al (184) 1994. USA.	Cancer clinical trials	Accountability	Based on principles of economic evaluation. Economic Health	Cost-effectiveness analysis followed by evaluation of the social return on research investment. Human capital approach used to value return on investment.	Nil	Cost-effectiveness of treatment estimated at $2094 USD per year of life saved. The net present value of the return on the $10.84 million investment in the trial estimated at $1.66 billion USD.	The cost of conducting research incorporated into analysis of the value of the research investment. Time-period for assessment of the costs and benefits extended to 2020 to capture downstream effects.	Assumptions made about adoption of trial results into practice rather than an assessment of actual practice change.	Highly relevant given that the focus was on a programme of phase III RCTs assessing colorectal cancer treatments. This could be used by a funding or research institute to assess the impact of the clinical trials they fund/support to reflect on the impact of trials already performed.
Ugolini D et al (185) 1997. Italy.	Cancer research centre	Allocation	Academic	Bibliometric assessment using an journal impact factor based metric	Nil	Most publications from the research centre that were identified and analysed scored highly (8-10/10) on the normalised journal impact factor score.	An early (1997) attempt to evaluate the academic impact of a cancer institute.	Metric, quantitative based approach only looking at a narrow interpretation of impact. Using the impact factor of a journal to assess the quality of individual research articles and for allocation of resources.	The journal impact factor is a misnomer with regards to assessing the wider impact of research. This approach would not be recommended to assess the impact of cancer clinical trials.
Ugolini D et al (186) 2002. Italy.	European countries	Analysis		Bibliometric method using the number of occurrences of an oncological publication in a journal (by author country of origin) compared to country population/GDP and	Nil	The UK made the highest contribution to European cancer publication output (21.12%), whereas Sweden performed best in the metric of number of publication occurrences versus country population and the Netherlands was ranked first for the mean	Comparison of two methods to understand how to assess the impact of cancer researchers' work.	Quantitative metric used that does not assess the quality of the individual articles or the contribution of the authors to the work. The authors acknowledge the limitations of using bibliometrics alone.	This approach is better suited to a large programme of research, rather than individual studies, such as clinical cancer trials. The main reason for performing an analysis using this approach is to make between country

			Academic	the "mean impact factor" of the occurrences of the publication by author.		impact factor of the occurrences.			comparisons. The metric used is narrow and does not adequately address the wider impact of research; it would not be recommended as an approach to evaluate cancer trials.
Coyle D et al (187) 2003. Canada.	Cancer clinical trial	Allocation	Economic Health (potential)	Assessment of "time to payback" for two alternative hypothetical cancer trials. This is the number of years until the returns from conducting a clinical trial outweigh the costs.	Nil	An effectiveness clinical trial would be worthwhile to look for a 5% improvement in survival from more intensive follow up (if a 5% improvement is considered likely). An equivalence trial would not be a worthwhile investment.	Ex-ante evaluation shows how impact assessment can prevent investment in cancer trials that are unlikely to be worthwhile.	An evaluation of potential rather than realised impact. Any error in the assumptions made and results obtained in these types of analysis means that some potentially impactful trials will not be performed.	This is highly relevant to cancer trials, however it is an approach to evaluate the potential impact of trials that are yet to be performed. The main rationale for using this approach would be to prioritise which cancer trials to develop based on cost-effectiveness of carrying out the trials. This approach would therefore be relevant to cancer trial funders who are making funding allocation decisions.
Lewison G et al (188) 2008. UK.	Any cancer research as cited in guidelines	Analysis	Policy (guidelines)	Identification of 43 UK guidelines from NICE, SIGN and Clinical Evidence. Bibliometric software used to analyse guideline citations.	Nil	UK papers were cited more frequently in cancer clinical guidelines than expected from their presence in the world oncology literature. The publications were generally more clinical than basic.	Outlines and executes a method for evaluating the impact of cancer research on guidelines.	Authors highlight that small clinical trials with negative outcomes are unlikely to be cited in guidelines, and discuss limitation of using clinical guideline impact as a surrogate for practice change.	Although this analysis included both basic cancer research and applied research such as clinical trials, the authors showed that it is mainly clinical research that is cited in these guidelines. Consequently, this approach to impact evaluation would be relevant to assess the policy impact of a programme of specific trials. If evaluating one cancer clinical trial, this could be performed using manual analysis of the relevant guidelines rather than the software used in this analysis.

Lewison G et al (189) 2008. UK	Any cancer research as cited in the media	Analysis	Media impact (dissemination)	Search of BBC archive health section (1998-2006). Percentage of stories focusing on cancer compared with UK's burden of disease (WHO 2002). Journal/author, research level (basic versus clinical), potential citation rate, actual citation rate/ funding of any cited paper recorded.	Nil	Research on breast, cervical and skin cancer are over-reported in the media compared to their burden of disease whereas lung cancer is under-reported. New and improved drugs are the research topic most cited. UK research was over-cited in the UK media compared to its place in world oncology research.	Novel attempt to investigate the impact of research in the media. Methodology clearly explained.	Only one media website archive used for analysis.	This approach to assessment would be relevant to cancer trials, both early and late phase. It is likely the type of trials reported in the media will be those either deemed to be of contemporary interest or those actively disseminated by trialists or pharmaceutical/device companies, and any analysis must recognise this bias.
Saad et al (190) 2009. Brazil.	Individual research projects presented in abstract form at a conferenc e.	Analysis	Academic Dissemination	Bibliometric analysis of a sample of abstracts.		Over 50% of abstracts were from the USA. Clinical trials were more likely than "other" types of research to be presented in poster or oral form (vs publication only).		Analysed only a sample (10%) of all abstracts, no comparison between the two time periods sampled and no information on the abstracts that were submitted but rejected.	This approach could be applied to clinical trials reported at a conference but it is assessing a narrow, academic definition of research impact.
Lewison G et al (191) 2010. UK.	Cancer research performed by one country	Analysis	Academic	Quantification of the number of cancer publications compared to national wealth/disease burden of cancer in Russia. Citation scores for Russian cancer publications compared with the citations to cancer papers worldwide in the same years.	Nil	Russia publishes one sixth as many cancer papers as its wealth and disease burden suggest. Collaboration in cancer research is dominated by former socialist states. Russian cancer research is incorporated into UK clinical guidelines, but rarely into UK media.	Multiple indicators used to map some important impacts of cancer research.	Mainly quantitative indicators used. No assessment of broader aspects of societal impact. Only looked at the impact on UK clinical guidelines and media rather that in Russia (likely to be much higher).	The aim of this analysis was to compare between country outputs for cancer research. This could be restricted to cancer clinical trials only from different countries, however this type of assessment focuses mainly on academic publications rather than downstream outcomes and impacts.

Lakdawalla DN et al 2010 (192). USA.	National cancer research	Accountability	Economic (societal perspective)	Quantification of gains in cancer survival using willingness to pay estimates derived from the literature and a comparison with the cost in cancer spending.	Nil	Cancer survival increased by 3.9 years from 1988-2000; equates to 23 million additional life years. Authors estimate that an average life year is worth 82,000 USD, therefore the value of this survival is monetised at 1.9 trillion USD in social value. 18-year time lag from research investment to survival outcome attributable research used.	Attempts to monetise a complex concept of willingness to pay for cancer survival. Detailed description of analysis.	Complex methods, willingness to pay values from previous literature. The actual investment costs in different cancer types and different research types (basic versus trials) not outlined.	This analysis would not be suitable to apply to a single clinical trial. Also, this top down approach is not transferrable to analysis of a pre-defined group of clinical trials alone because it relies initially on a population health change to have occurred, and then works down to identify the research (including clinical trials) that contributed to the health change.
Montague S et al 2010 (193). Canada.	One cancer trial	Analysis	Multiple (based on CAHS)	Clinical version of the CAHS to guide impact evaluation. Indicators of impact taken from this logic model framework and combined with a Bennett's hierarchy. Logic model is horizontal, hierarchy is vertical.	CAHS model, Bennett hierarchy	Impacts in all categories of the CAHS can be identified for this trial. The authors produce an impact timeline to show the pathway to impact in chronological order.	Use of conceptual framework means that impact of this cancer trial is communicated in a transparent and organised fashion. Use of the program chain of events helps demonstrate the processes through which impact has occurred. Incorporated the funding initially allocated to perform the cancer trial.	The theory of action hierarchy makes assumptions that prior events influence events higher up the chain. No estimation of the economic consequences of conducting the trial. Mainly desk analysis used.	This is highly relevant to cancer clinical trials and is a suitable approach to apply to a single clinical trial, as has been performed in this analysis. The use of a framework to report the results of the analysis is a strength of the evaluation.
Sullivan R et al 2010 (194). UK.	Research from cancer centres in the UK	Analysis	Academic Health sector Dissemination	Academic papers linked with cancer centres in the UK identified via database review. Journals categorised by "research level" (basic versus clinical). Potential citation impact/actual citation impact calculated. Calculation of the number of papers from UK cancer	Nil	UK cancer centres heterogeneous in terms of their overall research output and type of research performed. More focus on basic/fundamental research compared to applied research. Heterogeneity in proportion of papers from UK cancer centres that are cited in guidelines/media. Does not correlate with the size of centre or the	The authors combine publication citation impact with other impacts (on guidelines and media) to give a broader overview of cancer centre research impact.	Mainly quantitative measures of impact. The authors mention the potential future use of the Research Impact Framework but they do not use this or any other framework in their current analysis.	This approach could be applied to academic cancer clinical trial publications but it is a narrow interpretation of research impact and therefore of limited usefulness if wanting to evaluate wider impact.

				centres reported in UK guidelines and in the media.		conventional citation impact.			
Donovan C et al 2014 (195). UK.	Programme of cancer research projects	Accountability	Multiple (based on Payback framework)	Desk analysis of data held by the charity, survey of chief investigators, 16 case studies, bibliometrics and international benchmarking using bibliometrics. The case studies used document and archival analysis, citation analysis, citations in guidelines and interviews with principal investigators.	Payback	153 responses to a survey sent to 242 NBCF-funded researchers showed the research performed had impacts on drug development, higher degree attainment, capacity building for future research, policy, and health gain. Differences in impact between basic and applied cancer research.	Transparent methodology using a recognised framework and a mixture of quantitative and qualitative methods to gain in depth insight into impact of the research.	No consideration of the economic value of the return on the investment. Impact from the perspective of the researchers, no other stakeholders approached to comment on perceived impact of the funded research.	This analysis assessed both basic and applied cancer research, but it could be applied to clinical trials alone. The approach would rely on the researchers responsible for developing and running the clinical trials answering a survey and being available to provide information on costs and potential outcomes from the trials. It would be less suitable to evaluate the impact of a single trial.
Glover M et al 2014 (39). UK.	National cancer research	Accountability	Economic	Estimated research spending 1970-2009 (£15 billion), estimated NMB using monetised value for QALY, estimated cost of delivering benefit, estimated the proportion of NMB attributable to UK research and time lapse between funding and health gain and the internal rate of return from cancer research on health.	Nil	Time lag between research spending and impact on health gain estimated as 15 years. Overall return on public spending on cancer research estimated as 10.1%.	Considered time lag. Sensitivity analysis performed where possible and acknowledged areas of uncertainty. Acknowledged the purely quantitative nature of this assessment and accompanying case studies were performed (Guthrie et al, see below).	Assumptions made to perform this economic analysis. The authors outline the difficulty in differentiating the impact of smoking cessation in their calculation of the impact of cancer research overall.	This economic approach to evaluate impact uses a bottom up approach (versus the top down approach used by Lakdawalla) and therefore could be applied to a group of cancer clinical trials or a single trial. It could be improved by making a clearer link between the research/trials performed and the clinical practice and health changes attributed to the research in question.
Guthrie S et al 2015 (196). UK.	Six cancer research topics which included clinical trials	Accountability	Health	Case study approach using mainly desk/documentary analysis	Nil	Individual narratives for each case study.	The use of case studies enables the reader to understand the process through which impact has occurred and the time lines involved.	Requires in depth documentary analysis to contextualise and explain case study specific impacts.	This case study approach could be applied to one or a number of clinical cancer trials. It would be most useful when understanding the process of impact occurring is the focus of the investigation.

3.3.3.2 Approaches to cancer research impact evaluation used in empirical studies

(i) Categories of impact evaluated in cancer research impact assessments

Several of the empirical studies focused on academic impact. For example, Ugolini and colleagues evaluated scholarly outputs from one cancer research centre in Italy (185) and in a second study looked at the academic impact of cancer research from European countries (186). Saed et al (190) used submissions to an international cancer conference (ASCO) to evaluate the dissemination of cancer research to the academic community, and Lewison and colleagues (188, 189, 191, 194) assessed academic, as well as policy impact and dissemination of cancer research findings to the lay media.

The category of health impact was also commonly evaluated, with particular focus on the assessment of survival gains. Life years gained or deaths averted (184), life expectancy gains (187), years of extra survival (192) and QALYs were all used as indicators of the health impact attributable to cancer research. Lakdawalla and colleagues (192) considered the impact of research investigating both cancer screening and cancer treatments, and concluded that survival gains were 80% attributable to treatment improvement. In contrast, Glover and colleagues (39) acknowledged the importance of improved cancer therapies but also highlighted the significant advances from research around smoking cessation, as well as cervical cancer and bowel cancer screening. Several of these studies that assessed health impact, also used the information on health gains to assess the economic impact of the same research (39, 184, 187, 192).

Finally, two studies (193, 195) performed multi-dimensional research impact assessments, which incorporated nearly all of the seven categories of impact identified from the previous literature (Figure 3-2). In their assessment of the impact of research funded by one breast cancer charity in Australia, Donovan and colleagues (195) evaluated academic, capacity building, policy, health, and wider economic impacts. Montague and Valentim (193) assessed the impact of one randomised clinical trial which investigated the use of a hormonal medication as an adjuvant treatment for patients with breast cancer. In their

study, they assessed the dissemination of research findings, academic impact, capacity building for future trials and international collaborations, policy citation, and the health impact of decreased breast cancer recurrence attributable to the clinical trial.

(ii) Methods for cancer research impact evaluation

Methods of data collection and analysis for the purposes of research impact assessment used in these studies mostly aligned with the categories of impact that were being assessed. For example, studies assessing academic impact used traditional bibliometric searching of publication databases and associated metrics. In their evaluation, Ugolini et al (185) applied a normalised journal impact factor to publications from a cancer research centre as an indicator of the research quality and productivity. The journal impact factor is an index that reflects the yearly average number of citations for articles published in the previous two years by a given journal. This analysis was adjusted for the number of employees within each department and the scores were used to apportion 20% of future research funding. The same bibliometric method of analysis was used in a second study by the same authors to compare and contrast national level, cancer research efforts across Europe (186). They assessed the quantity and the mean impact factor of the journals for publications from each country and compared this to the location-specific population and GDP. A similar approach was used for the manual assessment of 10% of cancer research abstracts submitted to ASCO between 2001-2003 and 2006-2008 (190). These authors examined if the location of authors affected the likelihood of the abstract being presented orally, as a face-to-face poster or online only.

Lewison and colleagues, who performed four of the studies identified, (188, 189, 191, 194) used a different bibliometric method of publication citation count to analyse the dissemination, academic, and policy impact of cancer research. The authors also assigned a research level to publications to differentiate if the research was a basic science or clinical cancer study by coding the words in the title of each article or the journal in which the paper was published. The cancer research types assessed by these authors included cancer research at a national level for two different countries (UK and Russia) and research performed by cancer centres in the UK.

To assess policy impact these authors extracted journal publications from cancer clinical guidelines and for media impact they looked at publications cited in articles stored within an online repository from a well-known UK media organisation (British Broadcasting Corporation). Most of the cancer research publications contained in guidelines and cited in the UK media were clinical studies whereas a much higher proportion published by UK cancer centres were basic science studies. These authors also identified that funders of cancer research played an important role as commentators to explain the importance of the research in the lay media. The top ten most frequent commentators (commenting on >19 media articles out of 725) were all representatives from the charity CRUK.

Within these empirical examples of cancer research impact evaluation, a combination of clinical trial findings and documentary analysis of large data repositories were used to estimate health system or health impact from cancer research. In their study, Montague and Valentim (193) cited the effect size for a decrease in cancer recurrence reported from a clinical trial and implied the same health gains would be expected in real life for patients with breast cancer living in Canada. In their study of the impact of charitable and publicly funded cancer research in the UK, Glover et al (39) used CRUK and Office for National Statistics cancer incidence data. They also utilised national hospital databases listing episodes of radiotherapy delivered, the number of cancer surgeries performed, and systemic anti-cancer treatments prescribed, to evaluate changes in real world practice attributable to cancer research. In their USA perspective study, Lakdawalla et al (192) used the population-based Surveillance, Epidemiology and End Results Program (SEER) database to evaluate the number of patients likely to be affected by the implementation of cancer research findings (192). Survival calculations from clinical trials were also applied to population incidence estimates to predict the scale of survival gain attributable to cancer research (184, 192).

The methods of data collection and analysis used for economic evaluations aligned with the categories of assessment identified by Buxton in their 2004 literature review (100). For example, three studies (39, 187, 192) estimated direct healthcare cost savings from implementation of cancer research. This was

particularly relevant in one assessment of the potential impact of a clinical trial testing the equivalence of using less intensive follow up for patients following cancer surgery (187). These authors assessed the number of years it would take ("years to payback") of implementing the hypothetical clinical trial findings to outweigh the money spent developing and running the trial. The return on investment calculation was performed by estimating direct cost savings to the healthcare system by using less intensive follow up without any detriment to survival. There was an implicit assumption that the direct healthcare cost savings would be a way of paying back investment a future clinical trial from a research funder, even if the investment and cost savings were not from the same budget.

The second of Buxton's categories was an estimation of productivity loss avoided from research using the human capital approach. As described in Chapter 2, this method uses the national average income to value survival gains from patients surviving longer who are of working age. This approach was used in two studies (184, 192) and in both, estimates of average income (USA) were utilised. Buxton's fourth category, an estimation of an individual's willingness to pay for a statistical life, was used in two assessments (187, 192), and Glover and colleagues (39) adapted this method, placing a monetary value on the opportunity cost of QALYs forgone in the UK health service within a fixed budget (197). One of the studies that used this method identified that there may be differences in how patients diagnosed with distinct cancer types value the impact of research on cancer specific survival (192). In particular, individuals with pancreatic cancer seemed to be willing to spend up to 80% of their annual income for the extra survival attributable to implementation of cancer research findings, whereas this fell to below 50% for breast cancer and CRC. Only one of the studies specifically calculated Buxton's third category of benefits to the economy from commercial development (192). These authors assessed the gain to commercial companies from sales of on-patent pharmaceuticals and concluded that economic gains to commercial producers were small relative to gains from research experienced by cancer patients. Glover et al (39) did mention GDP spillover in their study but they did not carry out a cancer specific evaluation for this type of impact.

The cost estimates relevant to cancer treatments used in these impact evaluations came from documentary analysis, clinical trial publications, real-life data repositories, surveys, and population average income estimates. For example, in one study, cost information from a NCI trials database was supplemented by using a telephone phone survey to pharmacies, historical Medicare (USA national health insurance programme) documents and estimates of the average income from the 1986 US Bureau of the Census Consumer Income (184). In their study, Coyle et al (187) costed annual follow up and treatment for cancer recurrence based on the Ontario Health Insurance plan, a cost model relevant to an Ottawa hospital and cost estimates from Statistics Canada (198).

The data used to calculate the cost of performing cancer research was usually from funding bodies and research institutions. For example, charity reports and Canadian research institution documents were used to estimate that it costs the National Cancer Institute in Canada $1,500 per patient accrued to a clinical trial (187). Government research investment outgoings were used to calculate the $300 billion was spent on cancer research in the USA from 1971-2000, 25% of which was contributed by the NCI (192), and that the NCI spent over $10 million USD in the 1980s to generate the knowledge that adjuvant chemotherapy was beneficial to colorectal cancer patients (184). Charity and research institution spending reports, along with an estimation of the proportion of funds spent specifically on cancer research, were used to demonstrate £15 billion of UK charity and public money was spent on cancer research between the years 1970 to 2009 (39).

Lastly, the two studies (193, 195) which adopted a multi-category approach to impact assessment used the highest number and broadest range of methods identified from the previous literature (Table 3-2). These included surveys and semi-structured telephone interviews with clinicians, documentary analysis of funding and project reports, case studies, content analysis of media release, peer review, bibiliometrics, budget analysis, large data repository review, and observations of meetings.

(iii) Frameworks for cancer research impact evaluation

Only two of the empirical articles identified in this study used an impact framework. These were also the studies that performed a multi-category assessment and used the broadest range of methods within their analyses. Donovan et al (195) used the Payback framework to guide the questions asked in their researcher surveys and interviews when they evaluated the impact of research funded by the Australian National Breast Cancer Foundation. Montague and Valentim (193) used the Canadian Academy Health Services (CAHS) Framework. Rather than using the framework in it is original form, they arranged impact indicators from the CAHS framework within a hierarchy to illustrate impacts occurring over time. The authors distinguished short term, intermediate and longer-term changes resulting from one clinical cancer trial, aligning with the concept of categorising impacts based on when they occur, which was described in one of the literature review identified in part one of this study (170).

(iv) Time-lags and impact attribution

Lastly, the challenges of time lags and attribution of impact were identified and addressed by several of these empirical studies. Lewison and colleagues tracked the citation of over 3,000 cancer publications in UK cancer clinical guidelines over time, (188) and in their analysis Donovan et al (195) explicitly acknowledged that the short time frame between their analysis and funding of the research projects under evaluations was likely to under-estimate the impact achieved. Glover et al (39) used bibliometric analysis of citations in clinical cancer guidelines to estimate the average time from publication to clinical practice change (eight years). They added seven years to account for the time between funding allocation and publication of research results giving an overall time lag from funding cancer research to impact of 15 years. The challenge of attribution was addressed in one study by using a time-line to describe impacts occurring at different time-points but linking back to the original research in question (193). The difficultly of estimating time lags and attributing impact to cancer research were both specifically addressed in a companion study (196) to the one conducted by Glover and colleagues. In this study, instead of quantifying the return on cancer research investment, qualitative methods of assessment were used. This approach identified factors that enhanced and accelerated the

process of impact occurring and helped to provide a narrative to link impacts to research.

3.4 Discussion

This study identified existing reviews of approaches to evaluate the impact of research and empirical examples of the assessment of cancer research impact. The cancer research impact assessment examples were performed over three decades, and mostly evaluated research performed in high-income countries. The empirical examples of cancer research impact highlighted the huge investment into cancer research that currently exists, and the desire by many research organisations and funders to understand the return on that investment. These examples included studies supported or commissioned by the National Breast Cancer Charity in Australia , CRUK, the Wellcome Trust, the NIHR and Academy of Medical Sciences in the UK (39), the Chief Scientist Office in Scotland and the Medical Research Council (188), Genentech Ltd and the RAND Cooperation in the USA (192).

Many of these evaluations acknowledged the contribution of both basic and applied cancer studies, and several studies categorised research publications based on this distinction. For example, Donovan and colleagues (195) successfully used the same approach to assess the impact of both basic and applied research. They found that basic research was more likely to have knowledge and product based impacts and to take longer to accrue impacts compared to applied, clinical studies that more frequently affected policy and health outcomes.

Looking across the empirical examples identified, conclusions can also be drawn about what type of cancer studies were deemed most impactful by these authors. In their study (2008) on the impact of cancer research within the media, Lewison et al (189) found that the largest proportion of cancer studies cited in the media concerned novel treatments. The authors reflect that this is likely to be explained by the high research activity in the area of drug discovery and testing, and the advocacy of pharmaceutical and biotechnology companies that rely on media coverage to leverage support from their stakeholders. There were also studies cited in the media that focused on lifestyle choices preventing

cancer although no mention of research focusing on screening or cancer diagnosis. Breast cancer research also featured highly in several of these studies (39, 192, 193, 195), in particular, research on the treatment of breast cancer. Rather than concluding that breast cancer research per se is more impactful that research into other cancers, it is more likely this relates to the high incidence of breast cancer (most common cancer in women) and the new treatments that were tested clinically at the end of the 20th century.

Looking at how the categories of impact assessed varied across the fourteen articles identified, academic, health/healthcare, policy and economic impacts were commonly assessed. The studies assessing academic impact only were all published in the 1990s, whereas those studies looking at how publications were being used in the media and in guidelines were performed in the 2000s. Both of the articles which adopted a multi-category approach to impact assessment and those focusing on the economic returns on cancer research that had been performed were conducted from 2010 onwards. These findings are likely to reflect the broadening in the concept of research impact over time, demonstrating how the field of impact assessment has evolved.

Within the evaluations of health impact, there was a focus on survival, in particular in economic studies looking at the value of health gains. This reflects the high mortality rate of cancer as a disease entity and the importance of this as a measure of wider impact from cancer research specifically, contrasting with similar evaluations of musculoskeletal or mental health research, which have focused on improvements in morbidity (199, 200). Studies that analysed the number of patients diagnosed with cancer, or population-level survival gains, often used site-specific cancer incidence and other studies evaluated research relating to only one type of cancer (184, 187) (193, 195).

Impact under the category of policy change from cancer research was also commonly evaluated and this was done using clinical practice guidelines. Policy citation was often used as an indicator of the time lag between research being performed and being used in clinical practice (39). Reflecting on this approach, using cancer guidelines as a surrogate for clinical practice change and health service impact does have drawbacks. For example, guidelines can often be outdated, irrelevant, or simply not used by cancer clinicians. Furthermore, local

hospitals often have their own local clinical guidelines, which may take precedence over national documents. Lewison and colleagues (188) identified that the location where research has been conducted has a strong influence on the likelihood of clinical guideline citation. Studies performed in a certain country were more likely to be cited in guidelines from the same location. This makes sense if we consider many of the 'policy makers' for clinical guidelines are clinicians and researchers who may be more aware of research being conducted in their vicinity. The other aspects of policy impact described in the broader literature (146), such as impact on policy agenda setting and implementation, were rarely assessed.

The overview of reviews showed that the concept of valuing the economic return on research investment emerged as a field of study in the 1990s. Buxton's review (100) of key studies that estimated the economic value of the impact of health research on society was performed in 2004 and commissioned by the WHO. This was closely followed by the formation of the UK Evaluation Forum (2004), which was initiated by the Academy of Medical Sciences, the Medical Research Council, and the Wellcome Trust. The aim of the UK Evaluation Forum was to co-ordinate activity in determining the socio-economic benefits of UK medical research, and it's work led to one of the other literature reviews identified in this study (149) as well as informing the empirical example of estimating the value of cancer research performed by Glover et al (39). The other cancer specific studies which used this type of economic approach to evaluate impact that were identified in this study were performed by Ladawalla et al (192) and Brown et al (184) (see Table 3-4). All of these studies concluded that returns on medical research investment far exceeded expectations, with Glover and colleagues showing this was the case for UK cancer research specifically (39).

Lastly, although several of the empirical examples of cancer research impact assessment did address numerous categories of impact, there were no specific examples of social, environmental, or cultural impacts arising from cancer research. This is not surprising considering these are potentially less expected impacts from cancer research, and more in keeping with impacts that would be expected from arts or social science research. Nevertheless, it raises the question of whether these impacts may exist if they were actively sought out, or

if they could exist if researchers thought about those impacts in the early phases of their research. Examples of these more unexpected impacts from cancer research could include research targeted at improving social inclusion regarding access to cancer screening and treatment, or identifying and reducing socio-demographic inequalities where they exist. From a cultural point of view, cancer is a disease that affects all groups in society and there may be cultural differences with the approaches to care and treatment, especially around end of life care. There may also be an opportunity for cancer research to impact culture in terms of the arts, or for cancer researchers to work with those who work in the arts who may be able to disseminate cancer research findings in innovative ways so that it reaches a wider audience.

In general, the categorisation of research impact will be useful when considering how to approach the impact of the SCOT trial, and the categories of health, policy and economic value would be reasonable to investigate in greater detail. Regarding health gain, the likely benefit from SCOT will be the reduction in peripheral neuropathy within a real world population after the SCOT trial results are implemented in practice. Policy impact could be investigated by focusing on clinical practice guidelines and economic impact may be approached by considering a form of monetisation of the benefits from implementing SCOT, drawing on lessons from Lakdawalla (192), Brown (184) and Glover et al (39).

The methods of data collection and analysis (Table 3-2) commonly used in the empirical cancer research impact were bibliometrics, alternative metrics (media citation), documentary analysis, surveys, and economic approaches. All of these approaches may be suitable for assessing cancer trial impact, although if only bibliometrics or alternative metrics are used, the assessment is likely to provide information on academic impact alone. The method of collecting expert testimony from researchers was also utilised in the examples identified, but there were no obvious examples of testimony about the impact of cancer research from other stakeholders such as clinicians, cancer patients or their families.

Methods less commonly adopted in the empirical cancer research impact assessments were interviews, using a scale, and focus groups. This may have been due to the time and resource implications of using qualitative techniques

and more in depth analysis. It would be most appropriate to use these types of methods applied to a cancer clinical trial if the rationale for conducting the assessment was to understand the process through which impact occurred, or to perform a more in depth analysis if barriers to impact occurring had been identified. An example of when this was done was in the study by Guthrie et al (201). Using interviews and documentary analysis, these authors constructed cases studies to assess the impact of specific cancer research studies or programmes of work. Adopting this approach, they were able to identify ways that the impact of cancer research could be maximised going forward. For example, across their case studies the authors learned that having infrastructure and networks between researchers and policymakers could speed up impact on policy. Having a champion for the research in question, for example, the chief investigator, who pushes the research findings to research users through dissemination activities can reduce the time lag to impact occurring.

Despite the large number of framework examples identified from the previous literature (Table 3-3), only two of the empirical assessments, published after 2010, used an impact framework (193, 195). The first example used the Payback Framework to assess the impact of a programme of cancer research. This framework could be used to structure a survey or interviews with trialists or patients in order to understand the breadth and depth of impacts that are expected from a clinical cancer trial in development, or that have occurred from a completed cancer trial. In their review article, Hanney et al (146) provide a template for using the Payback Framework when conducting interviews, and there are previous examples of when the Payback Framework has been used to conduct surveys (195). The strengths of the Framework applied to cancer trial assessment are that it would encourage reflection on the original research question posed in the trial(s) in question, the inputs required to perform the trial(s), as well as the outputs and outcomes arising from the trial and its findings. The Framework also draws attention to the inevitable reservoir of knowledge that will exist whilst the clinical trial is being performed, which in turn influences the impact such a trial has on society. One drawback of the Payback Framework is that the list of categories and the logic model are separate entities, and do not overlap intuitively.

The second example of the use of a framework to assess cancer research impact identified from the overview of reviews was an iteration of the CAHS framework by Montague and Valentim (193). These authors provided a framework to structure the description of how impact from research is likely to occur, including a detailed list of impact categories and indicators. They authors chose a phase III cancer trial to demonstrate how their framework can be successfully applied, which is clearly very relevant when considering how to evaluate the impact of other cancer trials, such as SCOT.

Other frameworks identified from the literature (Table 3-3) that may be appropriate for the assessment of cancer trials impact include Anthony Weiss's logic model (202), the research impact framework (203), the research utilisation ladder (204) and the framework from Cruz Rivera et al (170). Their relevance to cancer trial assessment is discussed in more detail below.

The research utilisation ladder (204) is relevant to cancer trials because it focuses on the application of research results in practice, with the researcher self-evaluating their own research. A limitation of any self-assessment approach is that it may lead to an over-estimation of the extent to which research has been used. Overall, despite its relevance, this framework lacks the breadth of impact categories and the structure to aid data collection and communication offered by other frameworks on this list.

The Research Impact Framework (RIF) (99) consists of a comprehensive list of impact categories and indicators to evaluate the impact of health research. This has similarities to another of the frameworks identified (Table 3-3), the Becker model, although the latter was developed for biomedical research. There is no accompanying logic model or any indication of the how to prioritise the importance of the impacts described within the RIF. This could be seen as a disadvantage, however it also means there are no assumptions or restrictions regarding the process through which impact occurs, which may be direct or indirect. The categories and sub-categories of the RIF could be used to structure a cancer trial impact case study, to identify what type of data collection should occur and to build an impact narrative, either when a trial is being developed or once completed. It could also be performed for several trials simultaneously, which could then be qualitatively compared, for example, if making decisions

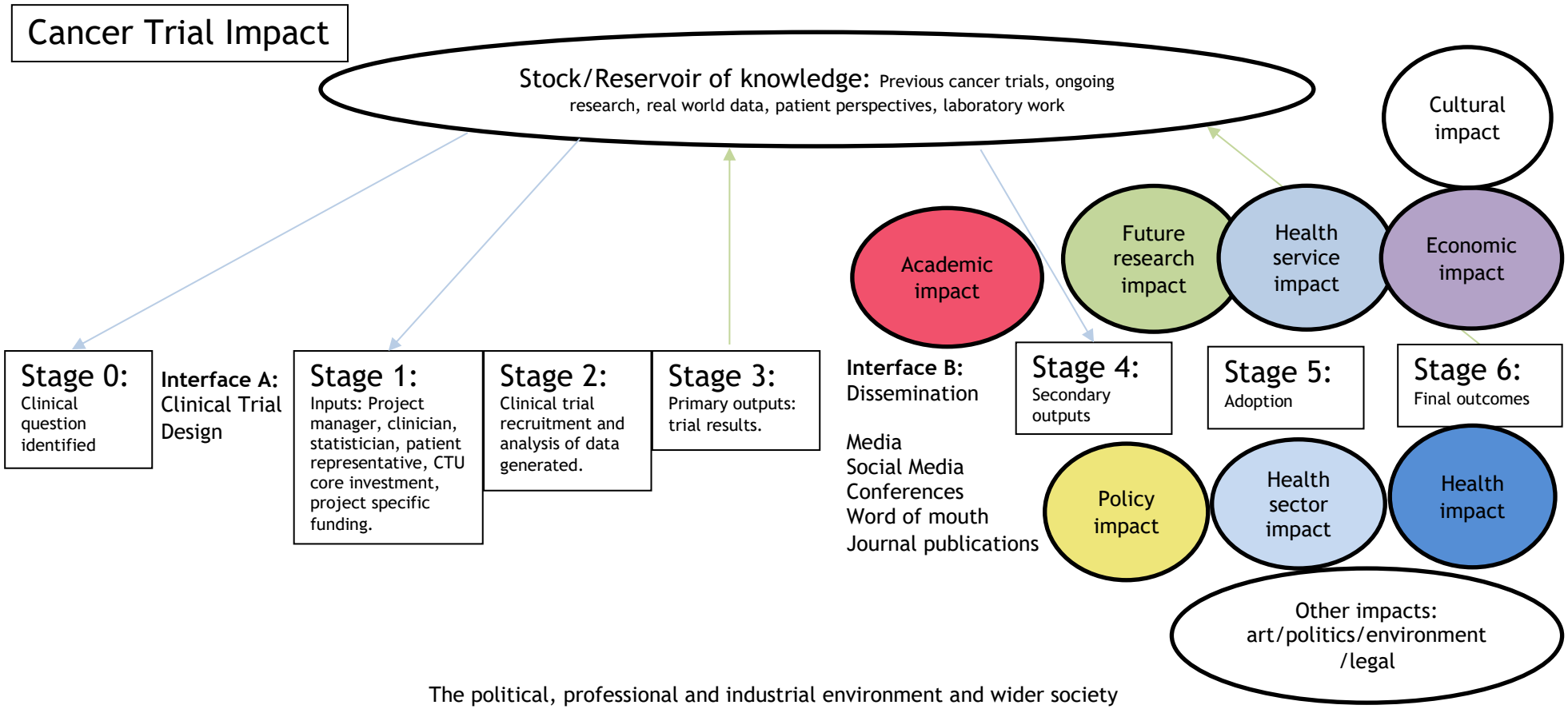
around trial investment or reflection on a collection of trials that have already been performed. This framework could also be applied to a survey or interview schedule to ask trialists, or users of cancer trial results, about the impact of that research. Overall, it is a good choice of framework to apply to cancer trial impact evaluation.

Weiss's medical research logic model (202) is highly relevant to the evaluation of the impact of cancer trials and is particularly well suited to evaluate the impact of a phase III trial that has met its primary endpoint. It is less obvious how it would be used to assess the impact of a trial that did not demonstrate efficacy or non-inferiority of an experimental treatment regimen, as the premise of the model is that the results of the trial are integrated into practice. In addition, it is not particularly relevant for the purposes of evaluating basic cancer research or earlier phase trials that are not expected to change practice. This model could be embedded or used alongside a wider impact evaluation framework that considers impacts that do not always rely on implementation of research findings.

Lastly, Cruz Rivera et al (170) have produced a comprehensive impact matrix that merges several aspects of both the Payback Framework and the RIF. Their matrix uses a linear model to order the impact categories, providing an indication of when different impacts may occur over time, alongside indicators of impact that can be utilised within each category. The limitations of the frameworks are that the indicators are not provided in any form of easily accessible list and there is no indication as to whether the research inputs and activities form part of this pathway. Finally, it is not straightforward to pick apart the distinction between their categories of health and health systems versus health related impacts. Improved health outcomes are grouped under a sub-category describing quality of care and service delivery, which are not necessarily synonymous. Overall, it is not clear that this matrix offers a better approach to impact assessment over that would be achieved by using the RIF or CAHS impact indicators alongside the Payback Framework or Montague and Valentim's framework.

This process of comparing and contrasting impact frameworks has indicated which components of these frameworks would be useful to assess cancer trial

impact. A modified logic model, such as the Payback or CAHS Framework, is a useful platform which could be used to structure data collection and build an impact narrative. The Payback Framework can be adapted to make it more relevant to cancer trial research by inserting additional categories of impact within the logic model, such as the seven identified from the overview of reviews. Putting this idea into action, below is a diagram of an adapted Payback Framework. It is recommended that a list of impact indicators, such as those in the RIF or the Becker model, are used alongside this type of logic model to provide specific examples of the types of impacts that can be evaluated within each impact category.



A modified Payback Framework

Reflecting on the findings from this overview of reviews and the examples of cancer research impact assessment that have been identified, Table 3-5 lists suggestions of how these findings can be used to guide the assessment of cancer research impact assessment in future, alongside an indication of how this advice could apply to assessing the impact of the SCOT trial in particular, to help guide analysis performed in later chapters (Chapters 5-8).

Table 3-5 Recommendations for cancer research impact assessment

Recommendation for cancer research impact assessment	Lesson from this the empirical examples of cancer research impact assessment identified in this study	Relevance to SCOT case study
Identify the unit of analysis relevant to the assessment, for example a cancer centre, cancer research programme, one clinical trial, one study	Many of these empirical examples chose to assess the impact of several studies. Montague and Valentim chose one clinical trial, showing it is possible to assess impact at the single study level. Focusing on one researcher alone is not encouraged.	The SCOT case study will focus on one phase III RCT but it will be important to consider other trials/studies that may influence the impact of SCOT, for example, the IDEA collaboration, and other clinical trials currently in set-up.
Identify the users of the cancer research, for example, patients/clinicians/funders.	The users of cancer research can be the target of surveys/interviews/focus groups to collect information on impact. Alternatively, the researchers themselves can be the focus of surveys or interviews to collect information on how they perceive their research is being used.	Relevant to the SCOT trial, the main users of the trial will be clinicians who prescribe adjuvant chemotherapy, patients who receive the chemotherapy and the healthcare systems that pay for the treatment and care for those patients.
Identify the rationale for the impact assessment	This rationale can be based around the 4 "As" and more than one may be relevant. The rationale for any assessment will help to inform the methods of assessment used. For example, if the rationale is to understand why and how impact occurs (analysis), the optimal methods may be interviews or case studies. If the rationale is mainly accountability, the main methods may be surveys and economic approaches.	The SCOT case study is being used to test different approaches to impact assessment. The most obvious rationale for this study, which has already completed, would be to analyse if/how/why impact has occurred and to demonstrate accountability for the resources invested to perform SCOT. The rationale of allocation of future funds and advocacy are less relevant for this case study.
Identify the main categories of impact relevant to the research in question.	Within the cancer research examples identified in this study, there was a strong focus on health, policy and economic impacts.	The categories of health, policy and economic impact are all relevant to the SCOT trial. Clinical practice impact is also highly relevant.
Decide on which methods are most suited to the assessment. Be mindful of the strengths and limitations of each method.	Surveys, interviews, documentary analysis, case studies and economic approaches were most commonly used in the empirical examples. Some methods and indicators of impact will be more suited to certain types of cancer research/clinical trials and it is necessary to be mindful of this if assessing more than one study. Care must be taken when making comparisons between the impact of studies. For example,	The methods used in the SCOT case study will be chosen from these commonly used approaches listed. Thought will be given to the likelihood that SCOT impacted on practice and the fact that it was a non-inferiority trial testing a reduction in treatment to

	cancer research funded by pharmaceutical companies or testing new treatments in a topical area are more likely to be cited in the lay media, with increased awareness amongst the public. Also, cancer trials that test a reduction in treatment or those investigating treatments that are off-patent are more likely to be directly cost saving if using economic approaches whereas trials testing expensive treatments for health gain will be more impactful if assessing e.g. survival improvement, the value of any health gain.	reduce toxicity, rather than an intensification of treatment with the aim of improving survival.
Consider if a framework would help to collect data and communicate the results of the assessment	There are six frameworks that have been identified from the overview of reviews that may be relevant to cancer research. Two that were used in the empirical examples were the Payback Framework and the modified CAHS by Montague and Valentim. A framework can also be used to structure surveys or interviews and this was done successfully in the study by Donovan and colleagues (2014).	The six frameworks described above will be tested for their relevance to SCOT specifically.
Consider the time and resources available to do any assessment and plan accordingly	Documentary analysis using a framework to guide data collection is a good initial approach if time and resources are limited. More in depth analyses can be performed within several categories of a framework and using more than one method if time and resources allow.	The time available to perform the SCOT case study analyses will be in line with the timeline for the overall doctoral study (approximately 3 years). The main resource will be researcher time. The researcher has expert clinical trials, statistical and health economic support for advice and access to SPSS and STATA software. Additional resources may require extra funding. Given these time and resources, it will be possible to do a number of in depth analyses, rather than using documentary analysis alone.

There are limitations of this study that must be acknowledged. Although the platform PROSPERO (205) was checked in advance to ensure a similar literature review was not being undertaken, the review protocol for this study was not registered in advance on the platform. This was an oversight and for any future studies this registration would ideally be performed. In addition, one researcher only (CH) screened the full articles for review, introducing potential selection bias into the review. Ideally there would have been two researchers available to screen the articles to improve reliability and validity of the final selection. This limitation was related to resource capacity and it was partly mitigated in this instance by using pre-defined inclusion and exclusion criteria in the screening process. The approaches to research impact identified from the previous literature were grouped under three main themes. Although there was space on the review proforma to collect other information deemed relevant to the research objective, if any other major themes regarding the topic exist, information on these themes was not systematically collected. Lastly, the literature search for review articles assessed the literature until April 2019. This meant that the empirical examples of cancer research impact assessment, by definition of their inclusion in those reviews, dated from before this time-point. There will undoubtedly be more recent examples of cancer research impact assessment that have been published which will not have been captured by this study.

The two-part approach to this study was taken because of the anticipated challenge of conducting a primary review of empirical examples of cancer research impact evaluation, and to allow a critique of empirical studies in the context of lessons learnt from the wider literature. Justification for the approach taken is provided by looking at the titles of the articles identified. In only 14% (2/14) was the word ‘impact’ included, suggesting that performing a search for empirical examples of cancer research impact evaluation using traditional publication databases would have been challenging. Furthermore, all the studies identified were included within reviews of approaches to research impact evaluation, which negated the subjective decision of whether the studies complied with a particular definition of research impact.

3.5 Conclusion

Impact assessment is a way of communicating to stakeholders such as research funders and patients, the merits of undertaking cancer research and learning from previous research to develop better studies that will have positive impacts on society in the future. This study is the first review to consider how to approach evaluation of the impact of cancer research specifically. The following chapter will look at publicly available research impact narratives to find examples of case studies describing the impact of cancer clinical trials.

4 Chapter 4: Lessons on cancer trial impact assessment from the REF 2014 case studies

4.1 Introduction

Cancer trials are an essential component of cancer research and a key step in translating basic research findings to the clinic. In contrast to 'blue skies' research, clinical trials are usually focused on answering a specific research question, often asking if a novel approach to treatment is superior, equivalent, or not significantly worse compared to the current standard of care. As a type of applied research, performed at large expense and involving human participants, it could be argued that, out of the many types of cancer research that exist, clinical trials in particular should have real world impacts for patients and society.

In Chapter 3, journal publications and the grey literature were both searched to identify key examples of when the impact of cancer research in general has been assessed. In order to understand how research impact assessment is relevant to cancer trials and how this assessment can be performed, it is useful to review previous examples of cancer trial impact evaluation specifically. As described in Chapter 1, the UK government allocates core research funding to higher education institutions based on an exercise known as the REF, and in 2014, 154 UK universities were assessed in this way (206). For the first time in the history of this national funding allocation exercise, the impact of research performed by each university was evaluated by using expert review of narrative case studies. In this chapter, the REF 2014 case study repository was used to identify and analyse examples of cancer trial impact evaluation.

Several authors have previously reflected on how universities evidenced the impact of their research in the REF 2014. Chowdhury, Koya and Philipson (17) reviewed 363 case studies in six disciplines from top or bottom ranking institutions and identified variables that predicted the average REF scores received by the institutions. For 92 case studies submitted under the discipline of Clinical Medicine, the number of publications in highly cited journals was the variable most consistently associated with higher REF scores. These authors also used automated word frequency analysis to identify themes of research

submitted under different disciplines. For Clinical Medicine, these included oncology, paediatrics, genetics, diabetes and heart disease research.

Terämä et al (207) used computational text mining of the REF 2014 case studies to understand how higher education institutions interpreted impact. By analysing 6,637 case studies, six classes of impact were identified (1 – Education; 2 – Public engagement; 3 – Environment & energy solutions; 4 – Enterprise; 5 - Policy; 6 – Clinical uses) and the class of impact described differed according to discipline. Similarly, a review of the REF 2014, commissioned by the Higher Education Funding Council for England, discovered that frameworks and taxonomies of impact were often context specific (208). Greenhalgh and Fahy (209) outlined 14 types of impact evidenced by higher education institutions in 162 REF impact case studies describing the impact of primary healthcare research. They found that an influence on guidelines was most commonly described, followed by impact on informing policy change and changes in clinical or public health practice.

The work reported in this chapter applies content analysis (see Chapter 2) to the REF 2014 case studies to understand how higher education institutions evidenced the impact of their cancer trials. The aim was to use this analysis in order to reflect on if, and how, impact assessment for cancer trials can be performed, and how impact evaluation can be improved, both for the REF 2021, and beyond.

To meet these aims, the objectives of this study were:

- to identify cancer trials included by higher education institutions in the REF 2014 case studies,

- to quantify and explore the characteristics of these trials and the types of impacts they were claimed to have had,

- to describe the types of evidence used by higher education institutions to substantiate those claims of impact,

- to identify any examples of researchers or research users making active attempts to maximise impact.

4.2 Methods

4.2.1 Overview of approach

Content analysis was used for the purposes of this study and this choice of approach was based on the work by Greenhalgh et al (209). In the initial phases of this study, Professor Greenhalgh provided an example of the coding manual used for a previous analysis (209). That coding manual was not used for the basis of the manual in this study, but it still provided a useful example for the primary researcher (CH) of how coding manual is organised. The development of the coding manual for this content analysis is outlined in Chapter 2.

4.2.2 Identification of a relevant sample for the purposes of analysis

The REF 2014 impact case studies were publicly available via the REF 2014 website (31). Case studies were deemed relevant if they described the impact of one or more clinical cancer trials, with the cancer trial(s) being the main focus of the research described. In order to identify this sample, initially a search of the case studies was performed by combining the terms 'cancer' and 'trial' in the website search function (31). Several combinations of search terms were tried with the aim of producing the most sensitive search. The REF 2014 website has a simple search function and it was therefore not possible to search the title only or to use truncated words to capture a number of key search terms.

This search identified case studies that included these words in any part of the submission (title, main text, or references). The case studies identified were read in full and the application of inclusion and exclusion criteria at this stage allowed the selection for final analysis. Inclusion criteria required that the case study focused on the impact of clinical trials that prospectively recruited adult patients with a diagnosis of malignancy, or individuals without a known diagnosis but where the aim of the trial was to investigate the prevalence of cancer in future years, diagnosis, or screening of cancer. All stages of cancer and clinical trials of all phases were included. Impact case studies were excluded if they described paediatric cancer trials. Case studies were also excluded if clinical trials were mentioned but were not the focus of the case study, for example if the case study

described basic science research and mentioned the development of a clinical trial as an example of the impact of that research.

4.2.3 Content analysis

The final version of the coding manual was used by the primary researcher (CH) to code all of the case studies and by LG to code five, randomly selected case studies. In Part 1 of the final coding manual used, the following information was recorded: (1) the institution responsible for the submission; (2) the Unit of Assessment and (3) the Summary Impact Type. For the clinical trials identified, the following key characteristics were extracted: (1) name; (2) phase of the trial; (3) type of cancer investigated; (4) focus of the trial (screening, diagnosis and treatment, other); (5) journal of publication cited in the case study; (6) category of funder; (7) primary endpoint; (8) if the primary endpoint was met. For the purposes of the final characteristic, trials were marked as positive if they met their pre-specified primary endpoint with statistical significance. For non-inferiority trials, if the experimental arm of the trial was deemed to be statistically non-inferior than the control arm at the level of significance pre-defined by the trialists, this was considered a positive result. For earlier phase trials such as phase I trials focusing on safety, if, for example, the authors set out to find a recommended phase II dose of a novel drug, and this was achieved and reporting in the trial findings, this was considered as having a positive result

Part 2 of the coding manual captured the following information for each impact case study: (1) all categories of impact described; (2) examples of dissemination of trial information and results; (3) methods used by institutions to evidence impact; (4) clinical guidelines cited; (5) examples of when researchers or research users acted to enhance trial impact (32). This information was collected by reading and manually coding the ‘Details of Impact’ section of each case study. Final testing of this coding manual provided an inter-rater reliability score of 81%. Analysis was performed using descriptive statistics.

4.3 Results

4.3.1 Identification of impact case studies

Out of 6,637 publicly available REF 2014 impact case studies, using the search word 'cancer' alone retrieved 494 case studies, the word 'trial' alone, 1120 and both together returned 234 case studies (Figure 4-1). Given that this search function returned all case studies that had the search terms anywhere within the case study, using both together improved the relevance of the search. The terms oncology (n=167), neoplasm (n=1) and malignant (n=63) were also used alone, and in combination with 'trial', but these searches did not identify any additional, relevant case studies compared to using the search strategy described above.

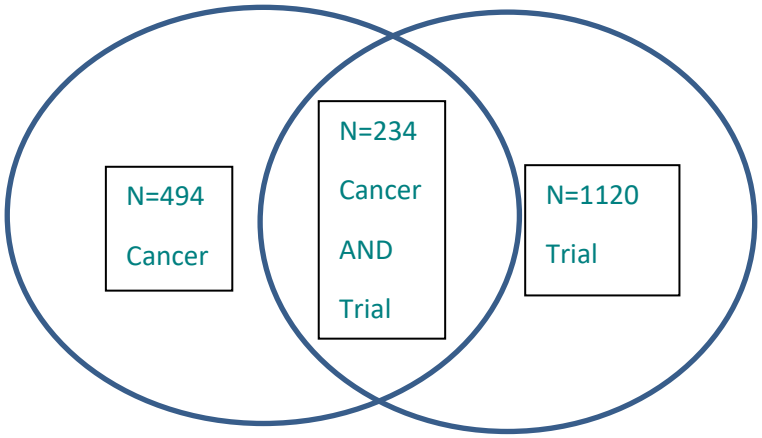


Figure 4-1 Different search strategies applied to REF2014 case studies

On reading the full submissions of these 234 case studies, 46 met the pre-defined inclusion criteria. Figure 4-2 presents the search results and details the reasons for exclusion.

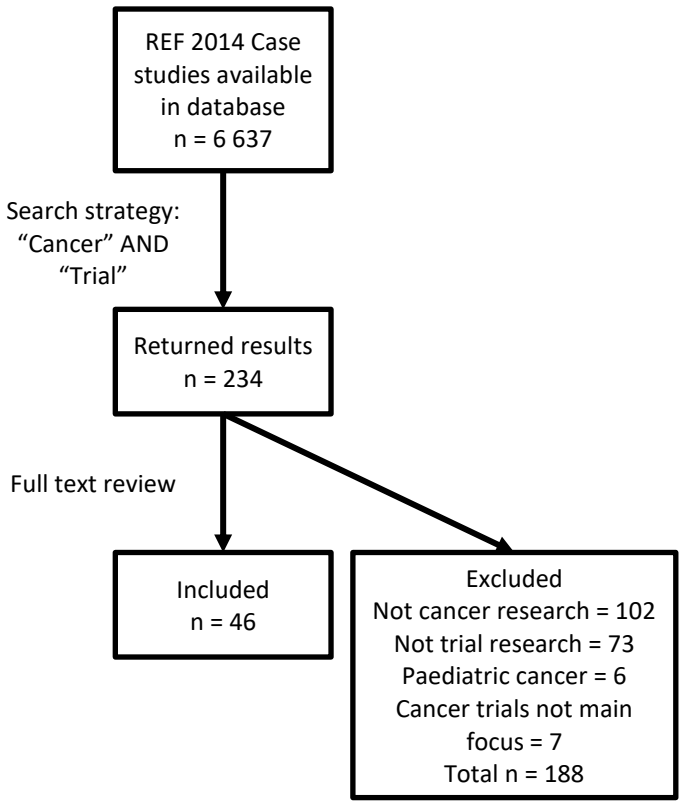


Figure 4-2 REF 2014 case studies search results

4.3.2 Description of case studies

The REF Unit of Assessment, Summary Impact Type, and name of institutions responsible for the submission for each of the final 46 case studies that met the study inclusion criteria are in Table 4-1. Nineteen (12%) out of 154 institutions participating in the REF 2014 submitted 46 case studies that specifically focused on cancer trials. Most of the higher education institutions were Russell Group Universities (89%; 16/19) (210), a members only association of 24 leading public research universities in the UK, whose member institutions submitted 68% of the highest ranked (4* outstanding) case studies in the REF 2014 (211). The majority (80%) of case studies describing cancer trials were defined as describing health impacts and most (83%) were in the field of Clinical medicine.

Table 4-1 Case study description (n=46 case studies)

Code	Number	Percentage*
REF impact type		
Health	37	80%
Technological	8	17%
Political	1	2%
REF unit of assessment		
Clinical Medicine	38	83%
Public Health, Health Services and Primary Care	4	9%
Allied Health Professions, Dentistry, Nursing and Pharmacy	2	4%
Psychology, Psychiatry and Neuroscience	1	2%
Biological Sciences	1	2%
Higher education institution (n=19)		
University College London	9	20%
Institute of Cancer Research	4	9%
University of Leeds	4	9%
University of Manchester	4	9%
Queen Mary University of London	4	9%
Imperial College London	3	7%
University of Cardiff	2	4%
University of Edinburgh	2	4%
University of Glasgow	2	4%
University of Nottingham	2	4%
University of Oxford	2	4%
University of Birmingham	1	2%
University of Bradford	1	2%
University of Bristol	1	2%
Cardiff University	1	2%
University of Cambridge	1	2%
King's College London	1	2%
Newcastle University	1	2%
University of Southampton	1	2%

*May not add to 100% due to rounding.

4.3.3 Description of cancer trials

The number of trials specifically cited in each case study ranged from 1-7. Overall, 106 individual trials were referenced 110 times. The majority of trials identified (68%) were phase III randomised clinical trials and most trials focused on the treatment of cancer (88%); trials investigating screening and diagnosis were much less common at 5% and 4% respectively. A large proportion of these trials recruited patients with a diagnosis of breast cancer (35%) (Table 4-2).

Table 4-2 Cancer clinical trial description

Code	Number	Percentage
Trials (n=110)*		
Trial focus		
Treatment of cancer or its side effects	97	88%
Screening	6	5%
Diagnosis	4	4%
Other (e.g. large observational trial to investigate cancer incidence)	3	3%
Phase of trial		
I	17	15%
II	14	13%
III	75	68%
Unknown/Other	4	4%
Diagnoses of patients recruited to the included clinical trials (n=110)**		
Breast	38	35%
Gastrointestinal (lower)	15	14%
Haematological malignancy	15	14%
Urological	13	12%
Gynaecological	10	9%
Thorax	8	7%
Central nervous system	4	4%
Head and neck (including thyroid)	3	3%
Multiple cancer types	3	3%
Gastrointestinal (upper)	1	1%
Main source of clinical trial funding (n=110)**		
Industry only	33	30%
Charity and Research	19	17%
Council/Government/University	16	15%
Research Council/Government/University only	14	13%
Unknown	13	12%
Charity and Industry	8	7%
Charity, Industry and Research	7	6%
Council/Government/University		
Charity only		

*Each clinical trial (n=106) was counted for each individual case study in which it was mentioned (a total of 106 trials mentioned in separate case studies 110 times).

The Arimidex, Tamoxifen, Alone or in Combination (ATAC) trial (212) was discussed in five separate case studies by four universities (213-217). The ATAC trial investigated the efficacy of an oral aromatase inhibitor compared to an oral anti-oestrogen, for the adjuvant endocrine treatment for postmenopausal women with hormone receptor positive, localised breast cancer. When the same university submitted more than one case study describing the same clinical trial, the first case study focused on the impact on clinical practice change worldwide

and the sales for the drug company responsible for the production of the aromatase inhibitor (215). In the second case study, the focus was on subsequent research which was possible because of knowledge generated by the clinical trial (216). Impacts described in the other three ATAC trial case studies included the provision of tumour specimens for translational research and investigation of novel biomarkers (217), citation of the trial results in guidelines, and subsequent impact on clinical practice and breast cancer relapse (214, 217).

The Prostate Testing for Cancer and Treatment (ProtecT) trial, (218) which was still recruiting at the time of the REF 2014 submissions, was described by two universities as an example of their work (219, 220). Both institutions outlined the collaborative approach to designing and performing this trial and the impact that the background work for the trial contributed to the concept of active monitoring for men with prostate cancer and on providing evidence to support a government decision not to introduce prostate cancer screening.

As shown in Table 4-2, there were often collaborative funding streams for these clinical trials from industry, the charity sector and government led research councils. Figure 4-3 shows that the journals of publication included both cancer specific journals and those aimed at a more generic clinical readership. The most common primary outcomes evaluated were overall or cancer specific survival (18%; 20/110) or a measure of disease recurrence or progression (18%; 20/110). Several trials used a co-primary endpoint (16%; 18/110). Although most trials (78%; 86/110) met their primary endpoint, one fifth of trials (20%; 22/110) did not and, for a minority of the trials (2%; 2/110) this was unclear.

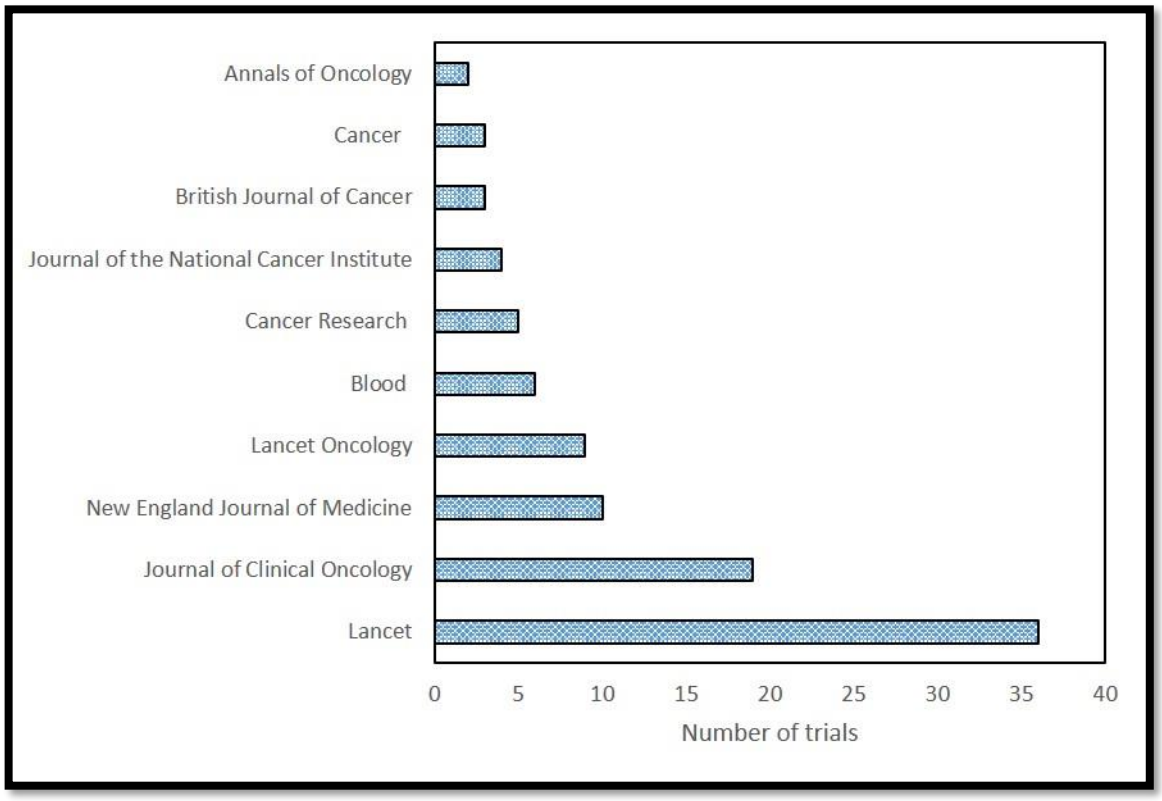


Figure 4-3 Most frequent journal of publication for cancer trials identified from REF 2014 case studies

4.3.4 Categories of cancer trial impact

The frequency with which different categories of impact were identified in the case studies are in Table 4-3.

Table 4-3 Impact categories identified in REF 2014 case studies

Category/Sub-category*	Case study references for this subcategory (number)	Case study references for this subcategory (percentage)
1. New knowledge and immediate research outputs	39	85%
1.1 New knowledge generated directly from clinical trial.	38	83%
1.2 New knowledge from clinical trial has contributed to a secondary analysis e.g. systematic review or meta-analysis.	3	7%
2. Capacity building for future research	24	52%
2.1 Clinical trial has contributed to the development (or intentional ceasing of the development) of further research, clinical trials and researchers.	19	41%
2.2 Clinical trial has led to collaboration and/or data sharing.	3	7%
2.3 Clinical trial has led to training of future clinicians and researchers.	5	11%
2.4 Clinical trial has led to innovation and novel infrastructure (other than health service related) e.g. the development of a novel technique or tool by a commercial company.	4	9%
3. Policy and guidelines	43	93%
3.1 Clinical trial has influence policy agenda setting.	7	15%
3.2 Clinical trial has led to a treatment approval (e.g. drug, device, procedure licensing, or marketing approval).	15	33%
3.3 Clinical trial contributed to clinical guidelines.	39	85%
3.4 Clinical trial contributed to other public policy e.g. government policy.	6	13%
3.5 Clinical trial has provided justification of the implementation of existing policy.	4	9%
4. Health sector (Health service)	16	35%
4.1 Clinical trial has influenced/benefited health service delivery.	16	35%
Health sector (Clinical practice)	37	80%
4.2 Clinical trial has changed clinical practice and actual clinical practice has been evaluated.	19	41%
4.3 Clinical trial has changed clinical practice and potential or estimated clinical practice has been evaluated.	30	65%
5. Improved Health for patients and public	32	70%
5.1 Clinical trial has contributed to improved health for patients (other than those in the trial) and actual health changes have been evaluated.	7	15%
5.2 Clinical trial has contributed to improved health for patients (other than those in the trial) and health changes have been estimated.	29	63%
6. Economic impact	25	54%
6.1 Clinical trial has led to direct cost savings for the health service.	12	26%
6.2 Clinical trial has shown benefit of a diagnostic or management strategy that is cost effective.	8	17%
6.3 Clinical trial has led to measured or estimated benefits for the macro economy e.g. sales of drug for a pharmaceutical company, setting up a new spin off company.	10	22%
6.4 Clinical trial has led to measured or estimated benefits to the macro economy from a healthy workforce e.g. patient returning to work earlier.	1	2%

*Sub-categories not mutually exclusive

Most case studies (93%) described the impact of cancer trials on policy, and in particular, the citation of trial results in clinical guidelines. A list of the ten clinical guidelines in which these trials were most cited is in Figure 4-4. None of the case studies referred to social or cultural impacts of clinical trials. One case study did explain that a clinical trial had changed 'culture and behaviour', but on reading the narrative this was coded as a change in the prescribing practice of clinicians (221). Another case study (222) discussed differences in cancer screening uptake between different socioeconomic groups which was partly identified by a clinical trial and has led to funding for a future trial to investigate and tackle this problem. There is potential for this subsequent trial to have substantial social impact if it successfully identifies ways to address this screening uptake imbalance.

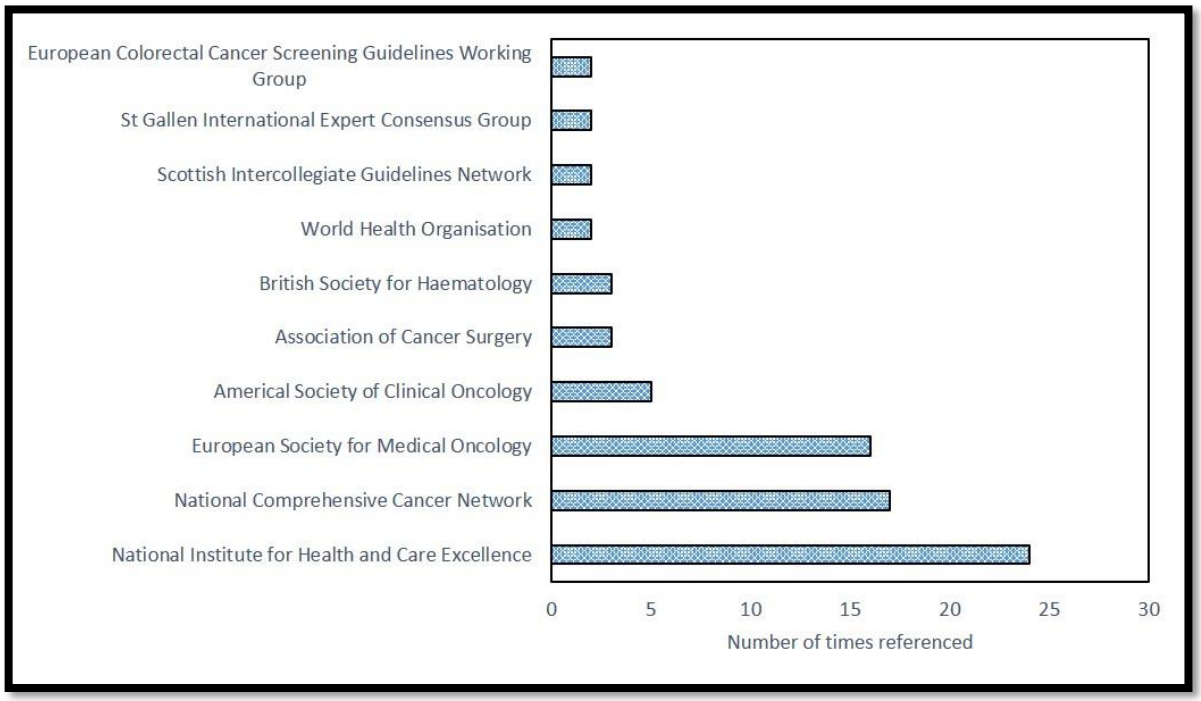


Figure 4-4 Guidelines most frequently reference in REF 2014 case studies

4.3.5 Dissemination and knowledge transfer

Overall, half (50%, 23/46) of case studies mentioned at least one type of dissemination. These examples were divided into a description of the publication of trial results in an academic journal (20% of case studies; 9/46), citation of the results publication in other academic articles (7%; 3/46) or other methods of communication (35%; 16/46) such as reports in the lay or social media, patient-facing websites and conference presentations.

4.3.6 Methods of evidencing cancer trial impact

Common methods used by higher education institutions to evidence the cancer trial impacts that were identified included: (1) identification of citations of trial publications in policy documents (78%; 36/46); (2) interrogation of real-life patient or population level data on clinical practice or health service use (52%; 24/46); (3) the use of expert or user testimony (30%; 13/46); and (4) surveys (both quantitative and qualitative) (15%; 7/46). Testimonies were only from researchers and funders, with none from policymakers or patients. Although many (70%; 32/46) case studies described the impact that cancer trials had on changing health outcomes, only seven (15%) described an actual, rather than predicted or estimated, change in health of patients (Table 4-3). Several (39%; 18/46) universities specifically quoted the monetary value of the funding linked to the research described in their case studies, totalling approximately £90 million. None incorporated this monetary value in an estimation of the economic return on research investment.

4.3.7 Researchers and research users enhancing cancer trial impact

A minority (15%; 7/46) of case studies mentioned that researchers actively enhanced the impact of a clinical trial. Examples included researchers interacting with policymakers to give advice on how to pilot implementation of clinical trial findings (223) and researchers making efforts to ensure trial findings are presented in the lay media, health blogs, and charity websites (221). There was also an example of researchers training clinicians in the selection of patients that would benefit from radiotherapy treatment that had been developed in the context of a clinical trial (224). There was one example of when a research user enhanced the impact of a cancer trial. This occurred when a patient used the results from a cancer trial to lobby the UK government to fund a novel drug to treat breast cancer for treatment of patients within the UK (225).

4.4 Discussion

There have been prior reviews of the REF 2014 case studies (17, 207, 209, 226, 227), but this is the first analysis that focuses specifically on cancer research or clinical trials. This study shows that UK universities recognise cancer trials as

impactful research undertaken at their institutions. The relatively small number of universities submitting cancer trial case studies implies that this type of research is concentrated at specific locations. Over half of the case studies described the impact of more than one trial, raising the question of whether it is feasible to expect a single, rather than a combination of trials or a programme of trials research, to lead to tangible impacts on patients and society. Conversely, a number of different universities described the impact of the same trial, illustrating the collaborative approach adopted at those institutions.

Trials recruiting patients with breast cancer constituted over a third of the cancer trials described as impactful by universities in the REF 2014; a much greater proportion than those recruiting patients, for example, with lung cancer (7%). This is not reflective of the actual, real world disease burden (228). For example, although breast cancer is the most common cancer (15% incidence) in men and women combined in the UK (229), lung cancer has the highest mortality rate and accounts for over one fifth of all cancer deaths (230). This over-representation of breast cancer research aligns with the findings from the literature review in Chapter 3, which showed that that many published examples of cancer impact assessment also focused on breast cancer. Rather than implying that breast cancer trials are intrinsically more impactful, the focus on breast cancer in these case studies is more likely to align with the landmark trials that reported results within the assessment REF 2014 eligible period (1993-2014). There were no cancer clinical trials identified in the REF 2014 that reported the benefits of modern immunotherapies, widely regarded as a major recent advance in cancer treatment. Again, it is likely that this reflects the publication dates of key trials investigating immunotherapy, which commenced around 2010 (231) and it will be interesting to explore if case studies describing immunotherapy trials are submitted to the REF 2021.

This study has demonstrated that, in the REF 2014, higher education institutions did not exclusively use clinical cancer trials that met their pre-specified primary endpoints as examples of impactful research. (This is distinct from trials which do not meet their primary objective, which may be, for example, to recruit sufficient patients to answer a specific research question). Those trials described in the REF 2014 which did not statistically meet their primary endpoint included

the LIBERATE trial (232) and the FOCUS2 trial (233). The LIBERATE trial closed early because it had recruited enough patients to demonstrate an increase in breast cancer recurrence occurred in patients being managed with hormone replacement therapy to treat symptoms following cancer treatment. The submitting university argued that the impact of this trial was a change in guidelines to prevent subsequent use of hormone replacement therapy for this group of patients. Another example was the FOCUS2 trial, (233) which tested the optimal treatment for elderly and frail patients with metastatic CRC. Although the trial did not meet its primary endpoint, it demonstrated the feasibility of recruiting patients from an often under researched patient cohort. It also provided important information around toxicity and quality of life that has subsequently been cited in clinical guidelines and changed clinical practice. This demonstrates that the pathway to impact is not solely dictated by practice-changing trials that demonstrate a novel treatment is beneficial compared to the standard of care, but that practice-affirming trials may be impactful by preventing harmful variation in practice (234, 235).

The fact that some higher education institutions used early phase trials as standalone examples of impactful research in the REF 2014 shows that robust examples of real-life impact do not only emerge from large, later phase trials. As an example, a portfolio of trials which demonstrated the safety, optimal dosing and blood brain barrier penetration of a drug to treat brain tumours, led to both direct (licensing of the drug) and indirect impacts (a phase III trial performed at another institution, subsequent introduction of the drug into routine practice and increased revenue for the pharmaceutical company) (236). Another case study described the impact of early phase trials investigating the use a targeted treatment for patients with BRCA associated breast and ovarian cancer. The significant improvement in outcomes for this sub-group of patients meant these trials directly influenced international guidelines for genetic testing and led to further research investment and collaboration with industry for that institution. Submission guidelines indicate that examples of indirect impact will be welcomed in the REF 2021 (237, 238).

The ten journals in which the clinical trials described in these case studies were most frequently published all had a journal impact factor over 5 and the top

three had a Journal Impact Factor above 25 (239). Also, one fifth of the case studies mentioned academic journal publication or citation of clinical cancer trials in other academic literature. This supports the findings from Chowdhury, Koya and Phillipson (17) that, although not an article level metric, and not a measure of impact, the research outputs underlying REF 2014 impact case studies were often published in journals with a high average citation count. This is likely to be driven in part by the REF requirement that the research described in the impact case studies is of high quality. Whether all high quality cancer research will necessarily be published in a high impact factor journal is debateable. For example, basic science projects, or studies in more niche areas of cancer research, such as radiation physics or investigation into rarer cancers, may be of high quality but will not be of sufficient interest to a wide readership to be included in major journals. Impact factor is not a useful term and should be considered a misnomer given that the term impact is now being used more frequently to refer to wider downstream effects from research rather than journal citation counts.

Clinical practice impact was commonly discussed in these case studies. This study showed that clinical practice impact in the REF 2014 was often estimated by stating the results of clinical trials and predicting the real world changes that would occur if the findings were implemented. This was the same approach to impact assessment used in Chapter 3 in the study by Montague and Valentim (193). In contrast, only a small number of case studies evidenced actual impact that had occurred using methods such as the analysis of national audit data (224), quantification of drug sales to indicate practice change, (217, 240) or referencing epidemiological studies to show improved health outcomes (221). Describing actual impact presents significant challenges in terms of timelines and planning but gives a much stronger indication of the real-life benefits from cancer trials compared to estimations of potential impacts. Although there is no specific guidance for the REF 2021 that indicates demonstrable impacts will be scored more highly than predicted impacts, informal conversations with individuals involved in REF 2021 case study development have highlighted that universities are looking for stronger examples of actual impact for their case studies for the upcoming assessment. In addition, there was not much focus within these case studies on how much investment, financial or otherwise, was

required to perform the research described, nor was there reflection on the return on any investment. For the REF 2021 and beyond, this approach may be a useful way for universities to evaluate the impact of their cancer trials and to attribute impact to specific trials and investments into those trials.

This study has shown that around half of case studies describing cancer clinical trials in the REF 2014 included a description of dissemination of research findings, and only a limited number of case studies described the process by which researchers or research users acted to maximise research impact. These findings do not imply that researchers or research users did not play an important role in the promotion, implementation, and the wider impact of cancer trial findings, but if this did occur, it was not identified by universities as important part of their impact narrative within these case studies. In addition, many examples of cancer trial impact from the REF 2014 case studies were predictions of impact. If these types of descriptions of impact are encouraged and rewarded by the government, there is no incentive for researchers and institutions to actually maximise the impact from cancer trials in real life.

The independent review of REF 2014 led by Lord Stern (241) suggested a bigger emphasis should be placed on public engagement for the next REF period, however, this guidance has not changed for the REF 2021 (242). This contrasts with Australia's new research assessment exercise (see Chapter 1), which includes engagement of the public by researchers as a key component that will be rewarded in their funding allocation framework. The REF also contrasts with an approach developed by several European nations and led by the Netherlands, which focuses on the productive interactions between science and society. Specifically, this consortium of nations has developed the SIAMPI framework (see Chapter 3 Table 3-3) which focuses on learning from impact assessment to develop and improve research institutions, and does not include a scorecard component. Government policy makers in the UK have a responsibility to drive an impact agenda that will improve future research and encourage institutions and individuals to increase trial impact. It would be beneficial if efforts to promote and maximise research impact this was explicitly incorporated into future REF guidance and rewarded in future REF assessments.

Lord Stern also suggested in his report that higher education institutions should be allowed to re-submit the same case study in 2021, recognising that the time lag (see Chapters 1 and 3) to impact that may occur. There is now guidance for the REF 2021 that resubmission of case studies will be allowed as long as new impacts from the same research have occurred. We know from this study that several institutions described the impact of the same clinical trial; this will be allowed in the REF 2021, but the REF guidance specifies that the contribution of the specific institution to the research in question is made clear. Finally, another difference that has been observed in the REF 2021 submission guidance compared to 2014 that is specific to Panel A (which includes Clinical Medicine), is that the experts reviewing submissions to this Panel will no longer prioritise the use of quantitative methods to assess and communicate impact when they are scoring the impact case studies. This may lead an increase in the use of more qualitative methods, such as qualitative surveys, testimonies, interviews, and focus groups, for submissions in 2021 compared to 2014.

This study has shown that the lack of publicly available scores for REF impact case studies means that it is impossible to tell what type of clinical cancer trial, and which type of impacts, were scored favourably by the REF expert reviewers. The lack of insight into these scores could be viewed as beneficial, in that the influence of the subjective view of REF expert reviewers and the government on setting the impact agenda is reduced. Alternatively, publishing the impact case study scores would increase transparency around the REF evaluation process and would stimulate conversations around what does constitute an impactful cancer trial. Currently, it is difficult to have this conversation as there is a gap in understanding around which case studies in 2014 scored most highly.

Lastly, policy impact has been identified in this study as an example of cancer trial impact that is important and relatively easy to measure. If policy impact is viewed as important and rewarded by the government via the REF, a straightforward way to maximise this impact is to recognise its perceived importance and for policy makers to mandate that major professional bodies update their guidance in a timelier manner. Trialists and patients could help the process of policy impact by lobbying policy makers and institutions to prioritise guideline updates.

Overall, the REF has several strengths. First, the inclusion of impact assessment within the REF exercise has brought impact assessment into the spotlight, and forced higher education institutions across the UK to consider the wider impact of the research they perform. It has helped to change the focus of research away from academic publications, and asked researchers to adopt a broader, societal view of research quality. Encouragingly, REF does not evaluate impact using a single metric, but rather asks for impact narratives which allows the impact of any research to be contextualised and discussed. The drawbacks of the REF in its current form include the lack of transparency around scoring and the inevitable bias and subjectivity that will accompany any expert review, especially as many of the individuals scoring the case studies will be experts in their fields, rather than experts in impact assessment. The aim of the REF is for allocation of resources, and there is little focus currently on learning from the impact case studies written for the exercise. In addition, creating the case studies comes at a huge cost for universities, in the hope that the direct return on investment in terms of funding awarding for a highly scoring case study will be worthwhile financially. The cost-effectiveness of performing the REF exercise itself would be improved with more focus on reflecting on the pathways to impact and maximising societal impact of future research.

The lessons learned from this study that are relevant to assessing the impact of cancer trials are outlined in Table 4-4, alongside a description of how these lessons will be considered when planning and performing the SCOT trial impact case study in part II of this doctoral study.

Table 4-4 Lessons learned from this study regarding cancer trial impact evaluation and how these apply to the SCOT trial case study

Lessons learned regarding cancer trial impact assessment	Considerations for the SCOT trial impact case study
"Negative" trials that do not meet their primary endpoint can be impactful (E.g. LIBERATE and FOCUS2).	SCOT did meet its primary endpoint and demonstrated non-inferiority of 3 versus 6 months of treatment. However, the IDEA collaboration did not show non-inferiority in the overall population. Consideration may be given to if different types of impact are relevant for "negative" versus "positive" trials.
Policy impact was commonly assessed in REF 2014 using clinical guideline citation.	Guideline citation for one trial such as SCOT is a relatively straightforward assessment exercise and therefore will not be the focus of an in-depth analysis. If policy impact was being assessed for more than one trial, this would require more time and resource. This type of analysis for a number of trials simultaneously may be helped by using software, such as that used in the study by Lewison et al (Chapter 3).
Clinical practice and health impact were often mentioned in these case studies by discussing trial results and predicting their effect on practice/health, but this was rarely evidenced using real world data.	This is a gap that could be addressed in the SCOT case study by investigating how feasible it is to demonstrate actual clinical practice/health impact attributable to cancer trials. Relevant to the SCOT trial, these impacts would be a change in prescribing adjuvant chemotherapy for CRC and a reduction in subsequent toxicity from treatment, especially peripheral neuropathy.
Many REF case studies described the impact of more than one trial and several clinical trials were described as research performed by several universities.	When assessing SCOT, one must be mindful of any/all CTUs that contributed to design and set-up, as well as all of the centres and countries that recruited patients. In addition, SCOT was part of a pre-planned pooling of data with that from other trials/countries in IDEA and the impact of SCOT will be influenced by the reservoir of knowledge that exists from previous and other current trials.
Those REF 2014 case studies which used headings and categories of impact within their narrative were easier to understand.	As decided in Chapter 3, the categories of impact that will be considered for the SCOT case study will be policy, clinical practice, health and economic impact.
None of the REF case studies mentioned the use of a framework to collect or communicate impact data.	Several frameworks have already been identified (Chapter 3), which will be tested for their relevance to the SCOT trial.
A list of indicators of impact relevant to cancer trials has been compiled as part of the coding manual used to perform content analysis for this study.	This list was partly based on the list of indicators included in the Research Impact Framework (RIF) described in Chapter 3, but it was altered to better reflect indicators relevant to cancer trials. This list included in the coding manual has been reviewed to identify which indicators are likely to be most relevant to the SCOT trial. This list can be found at this link:

	http://researchdata.gla.ac.uk/1135/ . These indicators will be used to guide data collection to build the SCOT clinical trial impact narrative.
Higher education institutions are likely to have picked their most impactful trials for this REF 2014 exercise. This means it is not clear how their approaches to impact assessment would compare if they were asked to describe the impact of all cancer trials performed at their institution, some of which are likely to be less impactful.	The SCOT trial was selected for this case study because it was a large study performed by Glasgow CTU. Consideration can be given to how these approaches may apply to trials that do not meet their primary endpoint or to smaller, earlier phase trials.
Impact assessment in these case studies was strongly influenced by when the trials recruited and published their results (with the dates stipulated in the REF guidelines) and the time-lag to impact occurring.	The SCOT case study for this research project will be done in the early years following the SCOT trial publication. It is important to recognise that the same case study performed in 10 years would look different because impact will change and mature over time and as new trial evidence emerges.

There are several limitations to this analysis. Firstly, as with any review of the REF 2014 impact case studies, these case studies were not specifically intended for this type of secondary analysis. Consequently, the data used was not collected to align with the research question, which was to explore existing methods of approaches to evaluate the impact of cancer trials. Secondly, although having content validity for this study, the list of indicators of impact used to code these case studies will not be an exhaustive list of cancer trial impacts. In addition to the timeline restrictions regarding which cancer trials could be presented within the REF 2014 case studies, it is also important to recognise that universities will have picked the small number of cancer trials they perceived were the best representation of their work in terms of demonstrating impact. This means that the majority of the cancer trials performed by these universities are not being presented or assessed within this impact analysis, therefore it is not clear if the same approach to assessment used in the REF 2014 can be used across a whole programme of trials, which will include trials deemed less impactful by universities.

The method of assessing inter-rater reliability that was used during development of the coding manual was a comparison of the number of identical responses between individuals, expressed as the percentage agreement between individuals. This is a recognised way to assess reliability and has the advantages of being simple to calculate and directly interpretable. The disadvantage of this measure is that it does not take into account the agreement that may occur due to chance and as such, may overestimate true agreement. This is more problematic with an increasing number of raters testing the assessment instrument. An alternative approach to estimate inter-rater reliability of the dichotomous, categorical data presented in this study would have been the Cohen's kappa (108). This is calculated by comparing the observed proportion of agreement between individuals to the predicted agreement based on chance. Both Cohen's kappa and the predicted chance agreement are calculated using specific formulae (243, 244). The disadvantage of this approach is the lack of agreement on what level of kappa is acceptable, especially with regards to healthcare research (244). If the content analysis for this study had used continuous codes, inter-rater reliability would have been best evaluated using intra-class correlation (ICC) (245).

The result of this content analysis was a quantification of patterns identified in the impact case studies. An alternative approach would have been to adopt the similarly named but distinct approach of ethnographic content analysis or a qualitative analytical approach such as the application of grounded theory or thematic analysis. These methodologies emphasise both the role of the researcher and the context within which the documents are generated when constructing the meaning of the message in the documents being analysed. Usually, categories and themes emerge from the primary data (95). Quantitative content analysis was felt to be most appropriate for this study because the key objectives were to identify, describe, and quantify the characteristics of the cancer trials used by higher education institutions, and to what extent different categories of cancer trial impact were described in a systematic and structured way. The documents being analysed were secondary data sources and some aspects of the data were single words or phrases, rather than text, which would not have been amenable to purely qualitative analysis. In addition, the aim was to collect this information in a systematic and structured way rather than to explore any deeper meaning or themes around these trial characteristics. Finally, the categories of impact and methods for impact analysis of interest had already been generated from a review of the previous literature (Chapter 3), therefore the aim was not to explore new categories or methods emerging from the case studies.

For the purposes of this study, the focus was on evaluating research impact. Going forward, it would also be useful to make an assessment of the investment, both economic and non-monetary, into cancer trials. This would allow the impact of trials to be contextualised in terms of the investment provided from funders, and burden for patients from participating in these trials (246). In addition, a binary assessment to indicate if the primary endpoint was met for each trial was used in this analysis. In future, this could be evaluated in greater detail by also looking at secondary endpoints or widening the evaluation to explore if a trial met its objective to recruit sufficient patients to answer a clinical question.

It will be useful to repeat this exercise using the REF 2021 case studies to identify which cancer trials conducted during the REF 2021 time period (underpinning research in the period 1st January 2000 to 31st December 2020) are

regarded as most impactful by higher education institutions, and to understand how the methods of impact assessment have changed. Although not coded for the purposes of this study, a comparison of the dates of both the initial publication of clinical trial findings and the impact evidenced would be useful. This would improve understanding of the time taken to achieve impact from UK cancer trials. In one of the empirical studies identified in Chapter 3, the time lag from funding to clinical practice was estimated to be in the order of 15 years for cancer research (39). A better understanding of time lags specifically for cancer trials would provide insight into when an analysis of the return in cancer trial investment should ideally be performed and may identify opportunities to speed up impact in some scenarios (196).

If an analysis is repeated using the 2021 submissions, it will be important to realise that any changes in approach to impact assessment and the impacts from cancer trials described in the case studies may relate to multiple factors. These include the clinical trials themselves, changes in submission requirements for the REF 2021 and lessons learnt from 2014 by expert reviewers regarding how they value impact and how they carry out this assessment.

4.5 Conclusion

This study has identified previous examples of cancer trial impact evaluation and described the strengths and weaknesses of different approaches to the evaluation of those trials. This study will be helpful to institutions in any country that conducts cancer trials, in particular in the UK as higher education institutions prepare their REF 2021 submissions. This study will also allow cancer trial funders to contextualise responses received when trialists describe the actual or potential impact of their work. Reflecting on the lessons learned regarding research impact assessment, how it has been performed previously for cancer research (Table 3-5 Chapters 3), and specifically for clinical cancer trials (Table 4-4 Chapter 4), the next section of this study will apply some of these lessons to assess the impact of the SCOT trial (Chapters 5-8).

Linking Part I and Part II

Part I of this thesis includes Chapters 3 and 4. The results from these Chapters were used to inform the analyses undertaken in Part II (results Chapters 5-8), which describe in-depth analyses of the impact of the SCOT trial. In Chapter 5 surveys were used to explore clinical practice impact, Chapter 6 focused on economic impact and the analyses presented in Chapters 7 and 8 look at if, and how, the clinical practice and health impact of the SCOT trial can be assessed. Below is a description of how the results from Chapter 3, the overview of reviews and identification of examples of cancer research impact assessment, and Chapter 4, a review of REF case studies of cancer trial impact, informed the direction taken for the analyses reported in the subsequent Chapters.

The findings of Chapter 3 demonstrated that policy, clinical practice, health, and economic impacts have been frequently assessed in previous evaluations of cancer research impact. These were also the categories of impact commonly described in the REF 2014 cancer trials case studies (Chapter 4). Given that SCOT was a phase III RCT that demonstrated non-inferiority between a new approach to treatment compared to the standard of care, these categories were all deemed relevant to SCOT and candidates for further investigation in this study.

Policy impact of cancer research was often evaluated the REF case studies (Chapter 4) by assessing citation of research results within clinical guidelines written by medical professional bodies. This approach was also used in one of the study be Lewison et al (188) identified in Chapter 3. Although very relevant to the SCOT trial, this approach to impact assessment was not chosen to test in further detail in this study because it was considered that manually identifying the SCOT trial publication citation within guidelines could be performed in a relatively straightforward manner. If assessing the policy impact of a programme of trials, rather than a single trial, this endeavour would be more suitable for an in depth analysis and the work by Lewison and colleagues demonstrates one way this could be done using specific bibliometric software.

The next category, clinical practice impact, was considered very relevant to the SCOT trial. Both surveys and interviews were identified as potential methods to

test clinical practice impact and relevant groups to survey/interview were users of the SCOT results, such as clinicians and patients, or alternatively, the SCOT trialists themselves. The previous work by Donovan et al (195) provided a good example of how to use a survey to ask cancer researchers how their work has impacted across a number of impact categories. Guthrie et al (196) showed that using informal interviews of researchers could provide information to build narrative case studies describing the impact of cancer research. In contrast, there were few examples identified in Chapter 3 and 4 of using surveys or interviews of research users, such as patients or clinicians, and this may have been linked to the time and resources available to perform these previous analyses.

For the purpose of this study, a survey of clinicians was chosen as the preferred method to take forward for testing, with the aim that this approach could quantify the extent of implementation of the SCOT trial results. Surveys(195)Surveys rather than interviews were used initially to assess a high number and broad range of clinicians from a number of different locations. It was acknowledged that interviews of a small number of clinicians could be used in an additional analysis, if barriers to practice change were identified within the survey. The results of the survey are presented in Chapter 5.

The results of the first survey presented in Chapter 5 led to consideration of how to evaluate the economic impact of the practice change identified. The work by Brown et al (184), Glover et al (39) and Lakdawalla et al (192) stood out in Chapter 3 as previous approaches to the assess the value of cancer research which could be tested specifically for the SCOT trial. Brown et al (184) estimated the social return on the research investment by the National Institute of Health (USA) to perform clinical trials investigating adjuvant treatment for colon cancer. The authors made assumptions to predict adoption of the clinical trial results into practice and calculated the number of deaths averted due to adoption of this new research evidence. In order to calculate the return on research investment ($1.66 billion USD), the cost of treatment (($10.8 million USD between 1978-1990), including an estimate of lost earnings during treatment, was subtracted from the value of deaths averted (valued by calculating the net value of average remaining lifetime earnings ($2.01 billion

USD)). The strengths of this analysis were the use of within trial and real world cost information, whereas the limitation of this evaluation was the uncertainty surrounding adoption of trial findings.

Lakdawalla et al (192) analysed the value of gains in cancer survival attributable to cancer research funded by the US public and private sectors. These authors used a top down approach to estimate by how much life expectancy for cancer survival gains, the authors calculated that the life-year gains due to cancer research during this period amounted to $1.9 trillion USD of additional social value. The limitation of this approach was the lack of transparency around which individual research projects are being considered in this calculation.

Glover et al (2015) (39) estimated the internal rate of return on cancer research conducted in the UK between the years 1991 and 2010 by identifying the most impactful research findings and calculating the investment required to produce these findings. The benefit of this approach over that by Ladawalla was that it was clearer which studies were included in their calculation. The authors monetised research benefits by identifying the QALY gain achieved (£25,000/QALY) within the UK from implementation of the research findings and included an estimation for the return on investment from this research to the UK GDP. The health service cost of implementing research findings and the original cost of conducting the research were subtracted from the predicted monetised health gains to calculate a net monetary benefit attributable to UK cancer research. The limitation of Glover's approach was that only healthcare service costs of implementation were considered, whereas Lakdawalla et al (192) and Brown et al (184) considered broader societal costs and benefits. Also, the spillover benefit to the UK GDP used in Glover's study was taken directly from the previous literature, with no attempt to calculate this for cancer research specifically. Lastly, the measure of benefit (QALY) was potentially biased towards research that, when implemented, resulted in survival gains, rather than disinvestment in a futile treatment, affirmation that current practice is optimal, or treatments that improve quality of life but not survival. Neither of the studies by Lakdawalla (192) or Glover (39), focused specifically on clinical trials.

Acknowledging the strengths and limitations of these analyses, an adaptation and combination of the approaches by Glover et al and Brown et al (184) were selected to test the impact of the SCOT trial. Specifically, the healthcare service budget impact of implementation of the SCOT trial findings was tested. In addition, impact from a societal perspective and the additional value of potential QALY gain from implementation were investigated. This approach was performed from the perspective of all six countries that recruited to SCOT. The results of this analysis are presented in Chapter 6.

The results of the REF 2014 case studies analysis reported in Chapter 4 highlighted an apparent gap in using real world data to assess clinical practice and health impact from cancer trials. There were a small number of examples in the REF case studies of using hospital level data to explore the impact of clinical trials, but more commonly, the impact on practice and health was predicted by describing how many patients may affected if the trial results were used. The method of assessing cancer trial impact using administrative data was tested in this study by using chemotherapy prescribing records. This was first carried out using local data to explore how this could be approached statistically. The feasibility of accessing appropriate data and carrying out an analysis was then tested at the national, Scotland-wide level. The results of these studies are presented in Chapters 7 and 8.

Lastly, the findings of Chapter 3 highlighted that using a framework to carry out and/or present the results of an impact analysis could be beneficial. Six different frameworks identified from the overview of reviews in Chapter 3 were tested for their relevance to the SCOT trial by incorporating results from Chapters 5-8 (Appendix 6). Documentary analysis was also used to populate the categories within these frameworks which had not been the focus of in-depth analysis.

5 Chapter 5: Impact on clinician attitudes and self-reported practice

5.1 Introduction

The dissemination of cancer trial findings to clinicians, patients, and policy makers influences the downstream impact that may be achieved from that trial. The subsequent implementation of trial findings by those stakeholders into real life practice is a necessary step if trial results are going to translate into health gains at a population level; clinicians play an instrumental role in this process. Although clinicians involve patients in the decision making process surrounding their management options, ultimately the doctor has responsibility for choosing and prescribing the patient's treatment. For these reasons, clinicians were identified as key users of the SCOT trial findings and the downstream effect of the SCOT trial findings on their clinical practice was considered a key impact worth exploring.

As outlined in Chapter 1, SCOT was a randomised phase III RCT which demonstrated that using 3 months of doublet chemotherapy (CAPOX or FOLFOX according to clinician choice) was non-inferior to 6 months for patients with stage II (with high risk features) and stage III CRC in the adjuvant setting. Important sub-group analyses showed that non-inferiority was met for patients receiving CAPOX, but not for those receiving FOLFOX, and for those with low-risk stage III disease but not for those with high-risk stage III disease. A separately presented analysis at ESMO 2019 demonstrated that non-inferiority was not met statistically for patients with stage II CRC with high-risk features.

The SCOT trial was the biggest contributor to the IDEA collaboration (81), which was a study that pooled the results of six international randomised phase trials (ACHIEVE, HORG, IDEA-France, SCOT, SWOG/CALGB 70802, TOSCA). All of these trials assessed the comparison of 3 versus 6 months of doublet chemotherapy for stage III colon cancer. This collaboration did not meet the pre-specified non-inferiority endpoint to show 3 months was not unacceptably worse than 6 months of treatment, although the clinical difference between the two approaches was small (3-year DFS difference of 0.9%, 5 year OS different of 0.4%). The subgroup findings from the SCOT trial were mirrored in this larger analysis.

Four of the trials in the IDEA collaboration, ACHIEVE2, HORG, TOSCA and SCOT, also recruited patients with stage II colon cancer and one (SCOT) recruited patients with rectal cancer. A separate IDEA collaboration analysis of patients with stage II disease was undertaken using the results from these four trials only (247). This analysis also did not meet the pre-defined non-inferiority endpoint but again the clinical difference between 3 versus 6 months of treatment was relatively small (5 year DFS difference 3.2%). Non-inferiority was met for patients with stage II disease receiving CAPOX (0.3% absolute difference) but not for FOLFOX (7.3% absolute difference).

Previous surveys have been performed to investigate practice change following publication of the findings from the IDEA collaboration (81) and/or contributory trials. One of these surveys (248) was performed by Iveson et al in September 2017, 4 months after the SCOT and IDEA findings were publicised at ASCO 2017, but before the full publications in peer-reviewed journals in March/April 2018. This survey was answered by 458 clinicians, a large percentage of whom were from Japan (51%), followed by the USA (17%) and the UK (10%). The responses to this survey indicated that even in the short time since dissemination of results as an abstract, most (90%) clinicians felt that 3 months of treatment was could be used for 'some' patients with stage III colon cancer. Overall, clinicians preferred to use 3 months (56%) versus 6 months (44%) of treatment for patients with low-risk stage III colon cancer (in keeping with subgroup analysis showing non-inferiority), and most used CAPOX (70%) in this context. In contrast, the majority (88%) still used 6 months to treat high-risk stage III disease (again in keeping with subgroup analysis) and there was a split between using CAPOX (59%) versus FOLFOX (41%) in this context. There were important between country differences in the choice of regimen. Across all scenarios, individuals from the USA preferred FOLFOX (73%), whereas those from Japan and Europe preferred CAPOX (77% and 65% respectively).

A later study led by a Canadian group of researchers (249), surveyed 145 clinicians regarding their perspectives towards the IDEA collaboration findings and how these results impacted on their prescribing patterns. Over half of their respondents were from South America, with only a small proportion (<12%) from Europe. Almost all of their respondents (98%) were aware of the IDEA

collaboration findings although over one fifth indicated that they found it challenging to communicate the results of IDEA, and specifically the concept of a non-inferiority trial to their patients. Interestingly, slightly less (75% versus 90%) respondents to this survey, compared to the earlier survey by Iveson et al, indicated that the IDEA collaboration findings supported the use of 3 months of adjuvant treatment in 'some' patients. In total, 71% specified that they had changed their clinical practice in response to the IDEA findings. A high proportion (81%) of the respondents to the survey preferred to use FOLFOX pre-IDEA but this changed, with a majority (55%) preferring CAPOX in the post-IDEA period. Approximately 68% of respondents tailored the duration of adjuvant treatment delivered in line with stage III risk stratification, giving 3 months to patients with T1-3N1 disease, and 6 months to those with T4 and/or N2 disease. Nearly a third of respondents indicated that they still gave 6 months of treatment to all patients, whereas a very small minority (<1%) indicated that they had changed practice to give 3 months of treatment to all patients post-IDEA.

Finally, practice change in France in response to results from the IDEA-France trial (one of the contributors to the IDEA collaboration) was investigated in an online survey by Ouali et al (n=213) (250). These authors disseminated a survey to clinicians from January to March 2019 using mailing lists for three medical professional organisations in France. The results of this survey indicated a strong distinction in approach to treatment based on stage III risk stratification. For patients with low-risk stage III disease, 81% of respondents preferred to use 3 months of treatment and 74% indicated they used CAPOX. For high-risk disease, almost all (99.5%) respondents agreed with the statement that adjuvant treatment should be given for 6 months and 94% preferred FOLFOX in this context. The preference for CAPOX for low-risk disease was particularly interesting in this clinician cohort because there was a strong preference within the IDEA-France trial for using FOLFOX (90%) rather than CAPOX (10%). The results of the all three of these surveys had not been published when the survey used for the purposes of this thesis was developed and disseminated.

The aim of the work in this chapter was to explore if clinicians were aware of the SCOT trial findings, if those findings made an impact on real life practice and if that impact changed over time.

To meet this aim, the objectives were:

(i) To explore clinicians' awareness of recent clinical trial publications and their attitudes to using 3 months of adjuvant chemotherapy for CRC

(ii) To document trial impact on clinical practice and if there were any barriers to change

(iii) To explore the timing of any practice change

(iv) To assess if clinicians' attitudes and self-reported practice changed over time.

5.2 Methods

A survey was selected for the purposes of this study to allow elicitation of responses from a large number of participants from dispersed geographical locations. An online platform was chosen to allow rapid and widespread distribution and to reduce research costs associated with postage within the UK and internationally. Online Surveys® was used for development and piloting of the survey (outlined in Chapter 2). The final survey included four sections: i) Clinical studies and guidelines ii) Current practice iii) Attitudes towards using 3 months of adjuvant doublet chemotherapy for CRC (the experimental arm for SCOT/IDEA), and iv) personal clinical practice details.

Current practice at the time of survey completion was investigated using twelve hypothetical patient scenarios, six in which patients were aged under 70 years old and six in which patients were aged 70 years or older. This age cut off was used because of the increased uncertainty around using doublet chemotherapy for elderly patients due to a lack of randomised evidence specifically in this age group (see Chapter 1). This differentiation based on age was also suggested in

feedback from clinicians who piloted the survey, that patient age was likely to affect clinical practice. A list of the patient scenarios used in the survey is provided in Appendix 3.

Scenarios relevant to stage II disease were separated by molecular disease characteristics. Specifically, scenarios describing stage II patients with tumours deficient in mismatch repair (MMR) proteins (dMMR) which are characterised by high levels of micro-satellite instability (MSI-H), were distinguished from patients with tumours proficient in MMR proteins, also known as micro-satellite stable (MSS) tumours (251). Patients with stage II dMMR CRC tumours have better survival (252) but appear to respond less well to fluoropyrimidine chemotherapy compared to patients with pMMR tumours. MMR status does not appear to predict response to oxaliplatin-based treatment (252).

All survey respondents who answered the first survey were asked if they would be willing to be contacted again and those who agreed were sent a follow up survey in August 2020. The same questions from the first survey regarding the acceptability of 3 months of doublet chemotherapy were included. The same patient scenarios were used except it was specified that patients had colon cancer rather than CRC. This was based on feedback from respondents completing the first survey that it would be more straightforward to answer questions based on patients with colon cancer alone, given then higher uncertainty regarding the use of adjuvant chemotherapy for patients with rectal cancer (see Chapter 1). In addition, stage II scenarios were separated into T3N0 and T4N0 (Figure 5-1). These changes were made in response to feedback from clinicians completing the first survey that T stage (T3 versus T4) was an important determinant of treatment choices for patients with stage II disease.

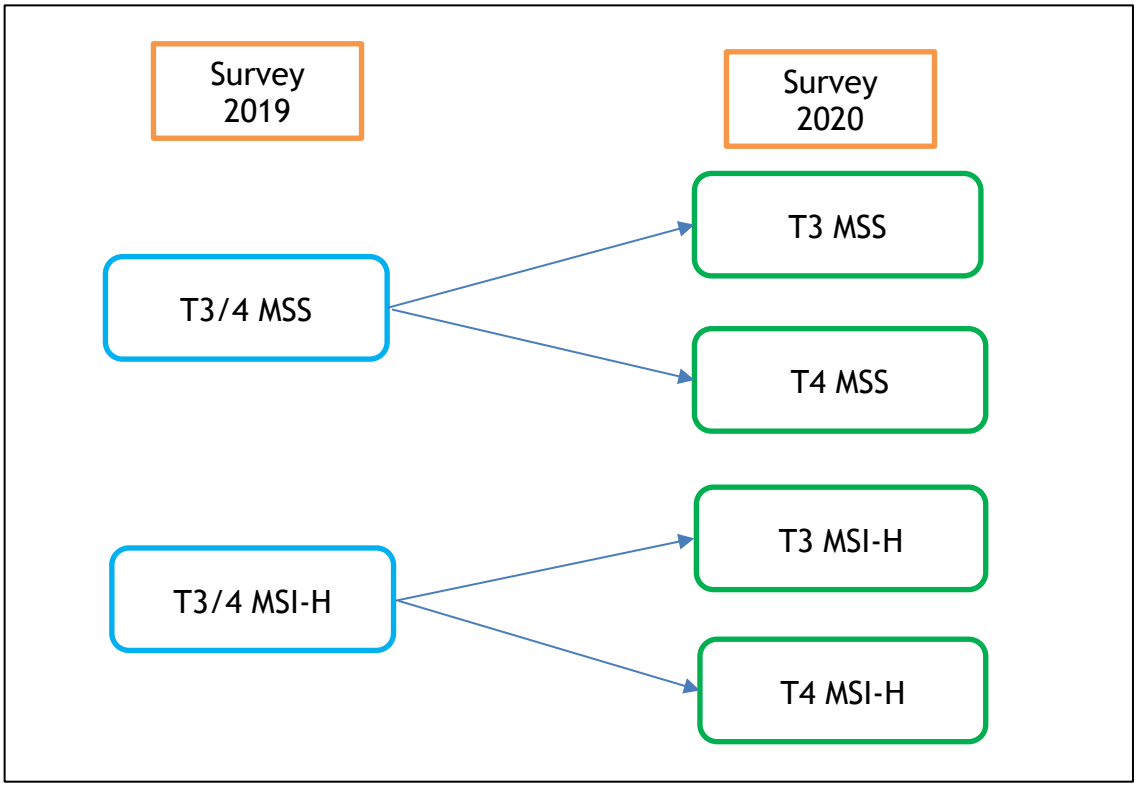


Figure 5-1 Disease characteristics for stage II patient scenarios

In the second survey, clinicians were asked to respond to patient scenarios initially disregarding the impact of COVID-19, and then asked to repeat the questions indicating changes in their practice due to the pandemic. In these answers, respondents were asked to indicate enduring changes that were likely to be maintained in their future practice, rather than temporary changes they made during the first peak of the pandemic. Responses to the second survey were linked to responses to the first survey from the same participants.

In the UK, a list of colorectal oncologists who prescribe adjuvant chemotherapy (n=247) was compiled for the purposes of this study by asking National Cancer Research Institute (NCRI) network leads in England and one oncology consultant in each of Scotland, Wales and Northern Ireland for a current list of their colorectal oncologists. The intention was that this list would be representative of the whole population of consultant CRC oncologists in the UK and therefore be a generalisable sample of all individuals who were responsible for chemotherapy prescribing in the adjuvant setting in the UK (107). It was recognised that this wider population would include other professional groups, such as oncology registrars, nurse specialists, or pharmacists, who may also prescribe chemotherapy in this context. There was no obvious means to generate a list of these individuals for the purpose of this survey.

Each person on the list of UK oncologists was sent an invitation email with a personal password protected link. The initial email was sent on 8th April 2019 and reminder emails were sent in an attempt to improve the response rate (253). The first reminder was sent on 29th April 2019 and a final reminder was sent on 21st May 2019. The survey closed on the 30th May 2019.

To disseminate the survey outside the UK, and to those within the UK who were not consultant oncologists, a generic survey link was generated for this group of respondents that could be forwarded via email. This link was sent to personal contacts, colleagues, oncology organisations who agreed to forward the link to their membership (ESMO and the Clinical Oncology Society of Australia) and the link was posted on Twitter® and on the message board of a UK medical professional organisation (Royal College of Radiologists). Some UK oncologists used the generic link to complete the survey but informed the primary researcher they had used this link rather than their personalised survey invitation. When this was the case, this was accounted for in the response rate for UK CRC oncologists. This generic link was generated and disseminated on 10th April 2019 and was closed on the 30th May 2019.

The plan was to disseminate the second survey one year after the first survey, in April 2020. Due to the COVID-19 pandemic, this was postponed to August 2020. This follow-up survey was distributed solely by email correspondence using the list of email addresses provided by those respondents who agreed to be contacted again. If respondents replied to indicate they no longer treated CRC or had taken an extended break from clinical practice, these individuals were removed from the distribution list and not used for the purposes of calculated a response rate. This survey was kept open for approximately seven weeks (31st August- 28th October 2020) and two reminder emails were sent to potential participants during that time (24th September and 12th October). Figure 5-2 outlines the timeline for survey development, piloting, and dissemination.

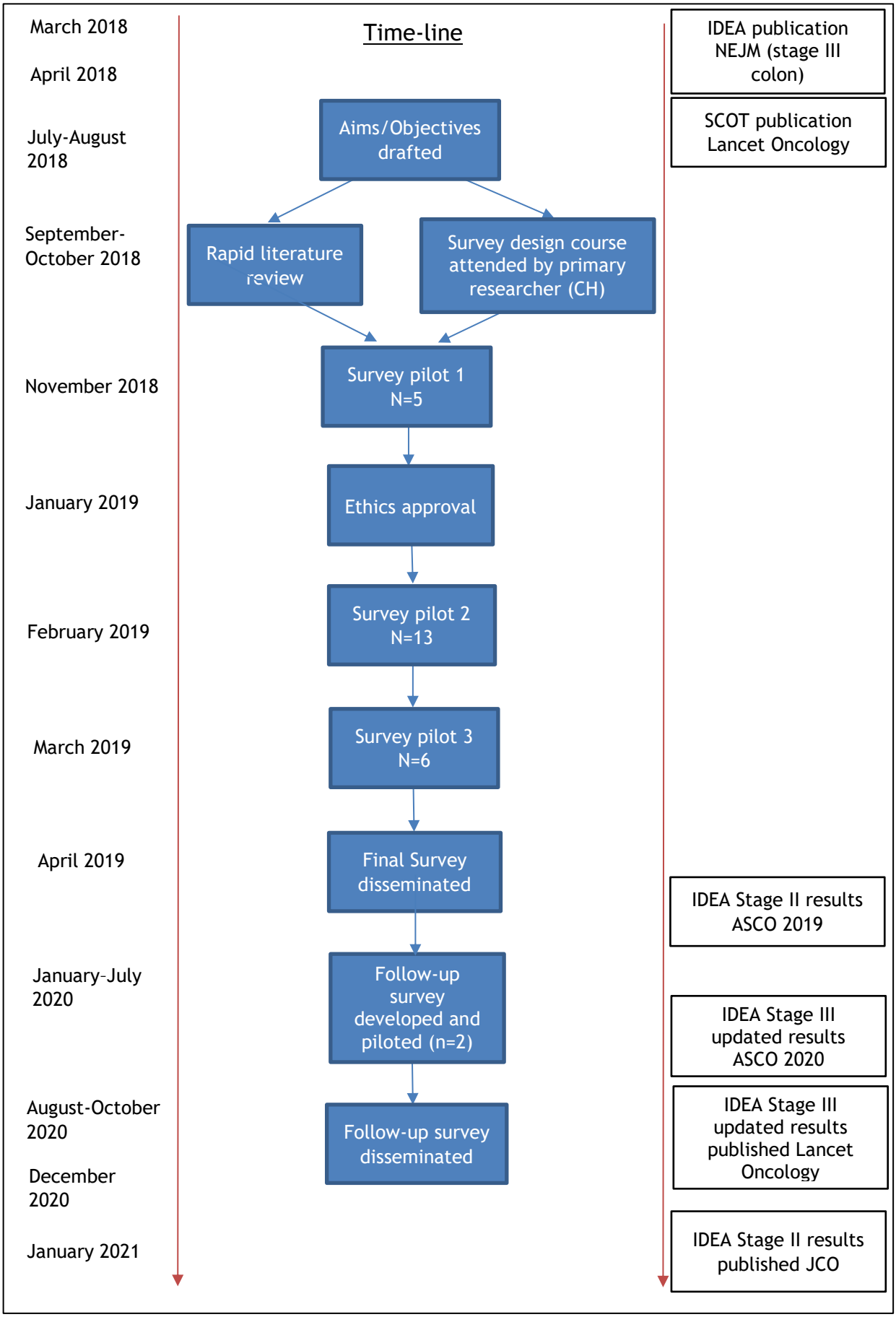


Figure 5-2 Survey development and dissemination Abbreviations: NEJM, New England Journal of Medicine; ASCO, American Society of Clinical Oncology; JCO, Journal of Clinical Oncology

The SCOT trial demonstrated 3 months of treatment was non-inferior to 6 months in the overall trial population, therefore any shortening of treatment from 6 to 3 months could be considered to align with SCOT results. However, to make it clear which changes in practice were in keeping with the subgroup results reported in the SCOT trial publication from 2018 (41), 'SCOT non-inferiority met' versus 'SCOT non-inferiority met' will be used when describing survey results based on disease risk or regimen used. Updated SCOT trial results describing 5 year OS have not yet been published, therefore these subgroup labels of 'non-inferiority met' and 'non-inferiority not met' relating to 5 year OS, rather than 3 year DFS, may change once those results are known.

5.3 Results

5.3.1 Survey response rate and respondent characteristics

In total, 265 clinicians participated in the first survey. Respondents were from Europe (180/265, 68%), USA/Canada (36/265, 14%), Asia (26/265, 10%), Australia/New Zealand (20/265, 8%), South America (2/265, 1%) and Africa (1/265, 0.4%) (Figure 5-3).

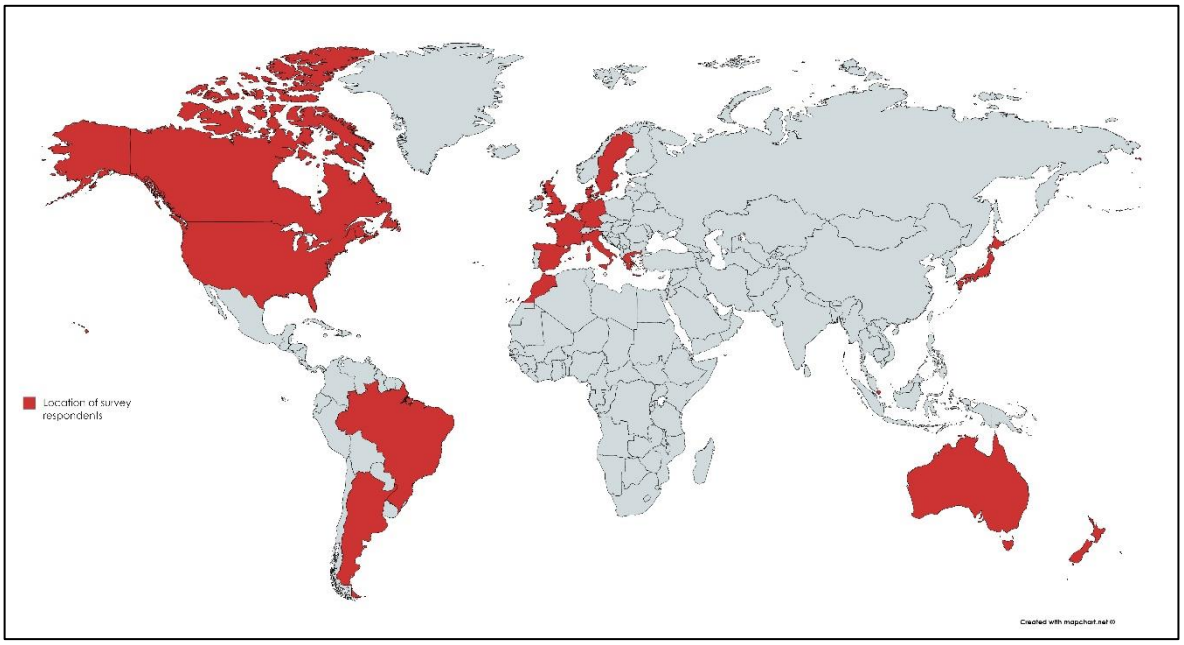


Figure 5-3 Location of survey respondents

Table 5-1 shows the location of respondents alongside an indication of which, if any, of the IDEA collaboration trials recruited in each country.

Table 5-1 Location of survey respondents according to country of recruitment to IDEA collaboration trials NA: Not applicable

Location of survey respondents	Number (%)	Patients recruited to an IDEA collaboration trial in this country?	If yes, which trial	Number of patients recruited from this location to the named trial
England	89 (34)	Yes	SCOT	5,244 (UK)
Scotland	16 (6)	Yes	SCOT	5,244 (UK)
Northern Ireland	6 (2)	Yes	SCOT	5,244 (UK)
Wales	4 (2)	Yes	SCOT	5,244 (UK)
UK (Unknown nation)	26 (10)	Yes	SCOT	5,244 (UK)
United States	35 (13)	Yes	CALGB/SWOG 80702	2,536 (USA and Canada)
Japan	25 (9)	Yes	ACHIEVE/ACHIEVE 2	1,313 (ACHIEVE) 525 (AHIEVE 2)
Australia	19 (7)	Yes	SCOT	197
Italy	11 (4)	Yes	TOSCA	3,759
Spain	6 (2)	Yes	SCOT	237
France	6 (2)	Yes	IDEA-France	2,010
Denmark	6 (2)	Yes	SCOT	311
Sweden	3 (1)	Yes	SCOT	83
Netherlands	3 (1)	No	NA	NA
Germany	3 (1)	No	NA	NA
Morocco	1 (<1)	No	NA	NA
Singapore	1 (<1)	No	NA	NA
New Zealand	1 (<1)	Yes	SCOT	16
Greece	1 (<1)	Yes	HORG	1,115
Canada	1 (<1)	Yes	CALGB/SWOG 80702	2,536 (USA and Canada)
Brazil	1 (<1)	No	NA	NA
Argentina	1 (<1)	No	NA	NA

The majority of respondents were oncologists (258/265, 97%); most had been practicing in the field of oncology for at least 10 years (196/265, 74%) and treated only or predominantly patients with CRC (215/265, 81%). The response rate from the pre-specified list of UK oncologists was 51% (126/247). In total, 106/197 (54%) of clinicians who agreed to be contacted and confirmed they still treated patients with CRC in August 2020, completed the follow up survey. They were from the UK/Europe (83/106, 78%), USA/Canada (12/106, 11%), Australia (6/106, 6%), Asia (4/106, 4%) and South America (1/106, 1%). The exact location

and characteristics of the clinicians who answered both surveys are shown in Appendix 3.

5.3.2 Clinician awareness of clinical studies

The majority of respondents (95%) were aware of clinical studies that reported results in the two years prior to April 2019, which assessed the optimal duration of doublet adjuvant chemotherapy for patients with CRC. This level of awareness was higher for UK (99%) versus international clinicians (90%) (Fisher's exact p=0.001). Almost exclusively, the studies named by respondents were the IDEA collaboration or those trials that contributed data to the IDEA collaboration. Clinicians from the UK were significantly more likely to mention the SCOT trial (51% of UK respondents versus 19% of international respondents, p<0.000 χ^2) and international respondents were more likely to name the IDEA collaboration (50% of international respondents versus 38% of UK clinicians, p=0.181 χ^2). Figure 5-4 shows the studies named by all respondents by location.

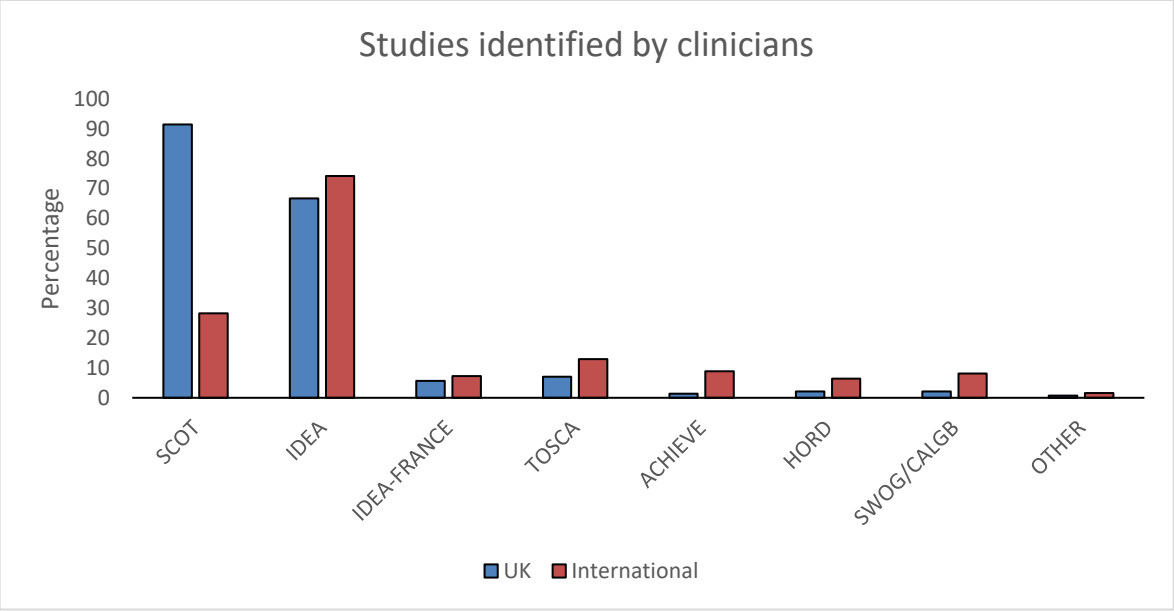


Figure 5-4 Studies identified by UK and international survey respondentsN=141 for the number of clinicians from the UK and n=124 for the number of international clinicians. Each respondent could name more than one trial. SCOT: Short Course Oncology Trial (NCT00749450); IDEA: Duration of Adjuvant Chemotherapy for Stage III Colon Cancer; TOSCA: Three or Six Adjuvant Colon (OsSC number 2007-000354-31); HORG: Hellenic Oncology Research Group (NCT01308086); CALGB/SWOG: Cancer and Leukemia Group B/South-West Oncology Group 80702 (NCT01150045); ACHIEVE: Adjuvant Chemotherapy for Colon Cancer with High Evidence (UMIN Clinical Trials Registry number UMIN 000008543). "Other" free text answers: "ACTS CC 02 trial", "SAFFA" and "Japanese trial testing 1 year of treatment but not doublet".

5.3.3 Trial impact on current clinical practice

Overall, 243 (92%) of survey respondents indicated that they had changed their practice in response to the clinical studies they had named, 3 (1%) were unsure and 19 (7%) indicated that they had not changed their practice. UK clinicians (99%) were significantly more likely to report a change in practice compared to international clinicians (83%, p<0.001 Fisher's exact). Out of the respondents who specifically named the SCOT trial (n=164), 98% (n=160) reported that they had changed their practice.

5.3.3.1 Scenarios describing patients aged under 70

For the six individual patient scenarios describing patients aged under 70, clinicians were most likely (93%) to change their practice in response to recent trials for patients with T3N1 disease (Figure 5-5a) (SCOT non-inferiority met). For the three scenarios describing patients with high-risk stage III disease (SCOT non-inferiority not met), the proportion of clinicians changing their practice was less. This proportion was similar if patients had one high-risk feature (T1-3N2, 46%; T4N1, 45%) but lower if two high-risk features were present (T4N2, 38%). For patients with stage II disease (SCOT non-inferiority not met), overall 33% of respondents indicated a practice change; this was more common for patients with MSS disease (36%) compared to MSI-H disease (30%).

Specifically looking at treatment duration chosen for each scenario, for low-risk stage III disease (SCOT non-inferiority met), using 3 months treatment (85%) was the most common choice. For patients with high-risk stage III disease (SCOT non-inferiority not met), most clinicians indicated they would chose to use over 3-6 or 6 months of treatment (84% average for three high-risk scenarios). For scenarios describing stage II disease with high-risk features (SCOT non-inferiority not met), either over 3 months (MSS disease) or active monitoring (MSI-H disease) were the most popular choices (Figure 5-5b).

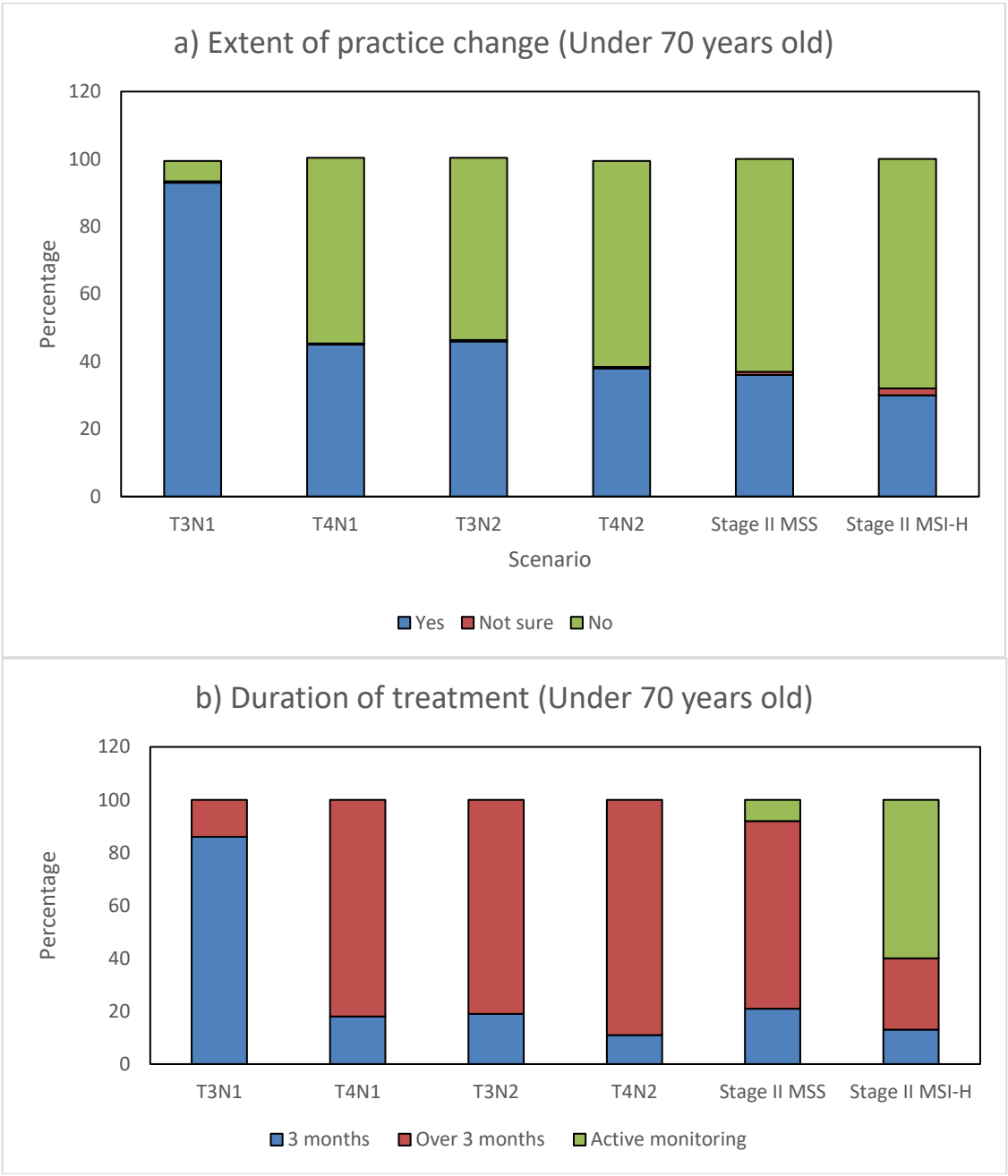


Figure 5-5 Self-reported practice change for scenarios describing patients aged under 70 years old a) Practice change b) Duration of treatment chosen.

5.3.3.2 Scenarios describing patients aged 70 years old and over

The same trends were identified for older patients but there was more heterogeneity in practice within the responses for each scenario compared to practice for younger patients. Figure 5-6a shows the percentage practice change and Figure 5-6b shows duration of treatment chosen for each scenario. Three months of treatment was the most popular duration for patients with low-risk stage III disease (SCOT non-inferiority met) and over 3 months was most popular for high-risk stage III and stage II disease (SCOT non-inferiority not met).

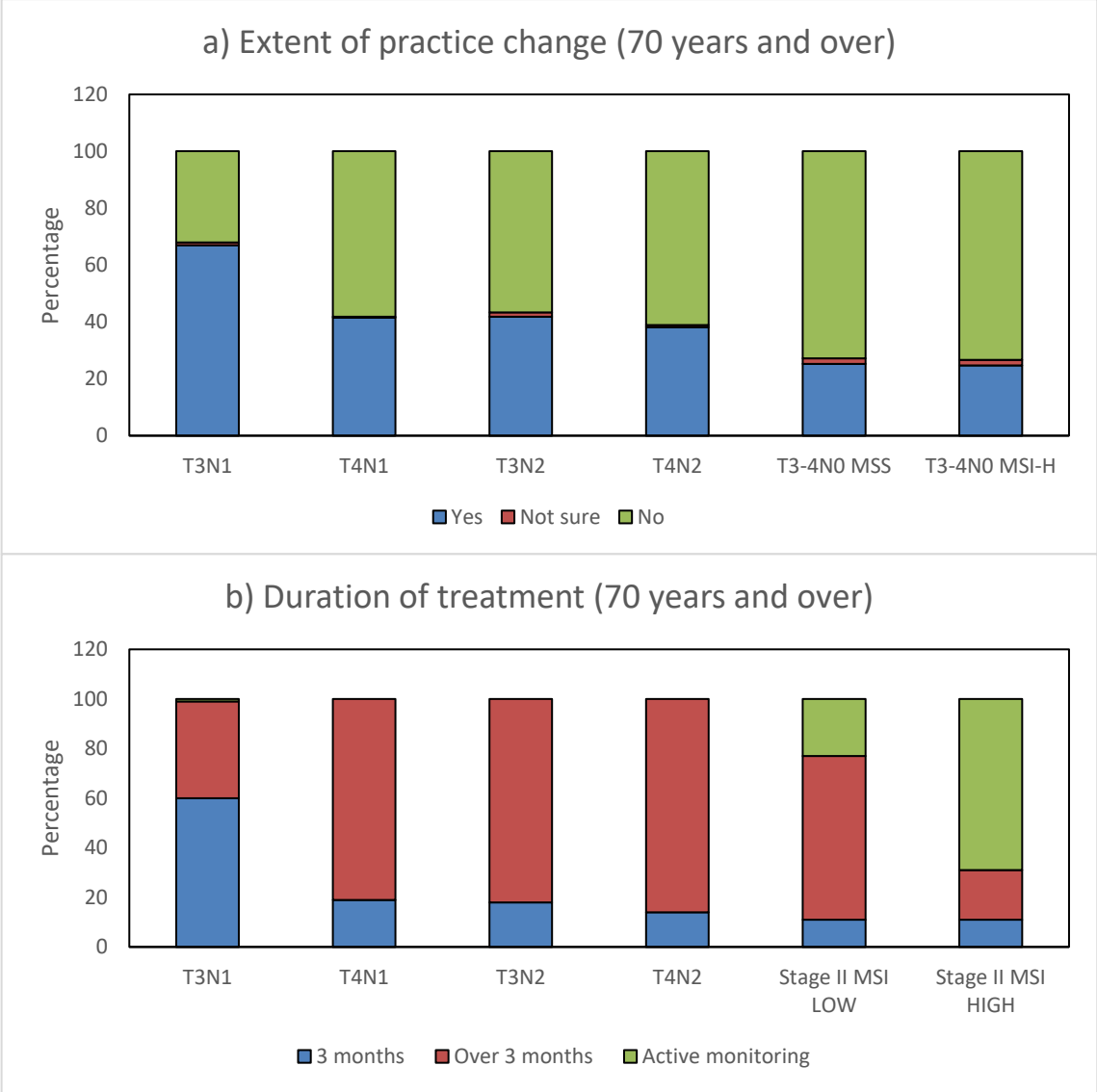


Figure 5-6 Self-reported practice change for scenarios describing patients aged 70 years old and over. a) Extent of practice change b) Duration of treatment choices

The choice of both regimen and duration of practice for all twelve scenarios are shown in Appendix 3. Summarising these results, and highlighting which practices aligned with the experimental trial arms in the IDEA study and SCOT trial, Table 5-2 shows the proportion of clinicians who specifically chose 3 months of doublet chemotherapy. For individuals who changed their approach to treatment dependent on age, the patterns of change for each scenario are shown in Appendix 3. In the second survey, the majority (92%, 98/106) of respondents reported they use biological rather than chronological age when treating patients aged 70 years and over. The definition of biological age was left to the discretion of survey respondents.

Table 5-2 Proportion of clinicians choosing 3 months of doublet chemotherapy

	Three months of doublet chemotherapy	
	<70 years	≥70 years
Low-risk stage III disease¶	86%	54%
High-risk stage III disease overall	16%	15%
High-risk stage III (T4N1)	18%	17%
High-risk stage III (T3N1)	19%	17%
High-risk stage III (T4N2)	11%	13%
Stage II* overall	16%	9%
Stage II* MSS	20%	8%
Stage II* MSI-H	12%	10%

*With high-risk features ¶ SCOT non-inferiority met for this subgroup

Lastly, clinicians were asked if they treated patients with rectal cancer in the adjuvant setting and if so, if they treated these patients using similar management strategies as they had outlined for the twelve hypothetical patient scenarios. Amongst UK respondents, 126/141 (89%) indicated they treat patients with rectal cancer in the adjuvant setting and 91/126 (72%) of these use the same treatment strategy as they outlined in the patient scenarios. For international clinicians, 117/124 (94%) indicated they treat patients with rectal cancer in the adjuvant setting and 49/117 (42%) treat those patients in a similar manner to as they outlined in the survey treatment scenarios.

5.3.4 Influences and barriers to practice change

Overall, clinicians most commonly indicated that listening to conference presentations was the mechanism of dissemination of trial results that had influenced their subsequent practice change (30%) (Figure 5-7). This was closely followed by reading a journal or specifically a high impact journal (26%), and discussion with colleagues (24%). Reading articles in the lay or social media (1%) or looking at a poster at a conference (1%) were the mechanism of trial result dissemination that were least likely to have influenced their practice.

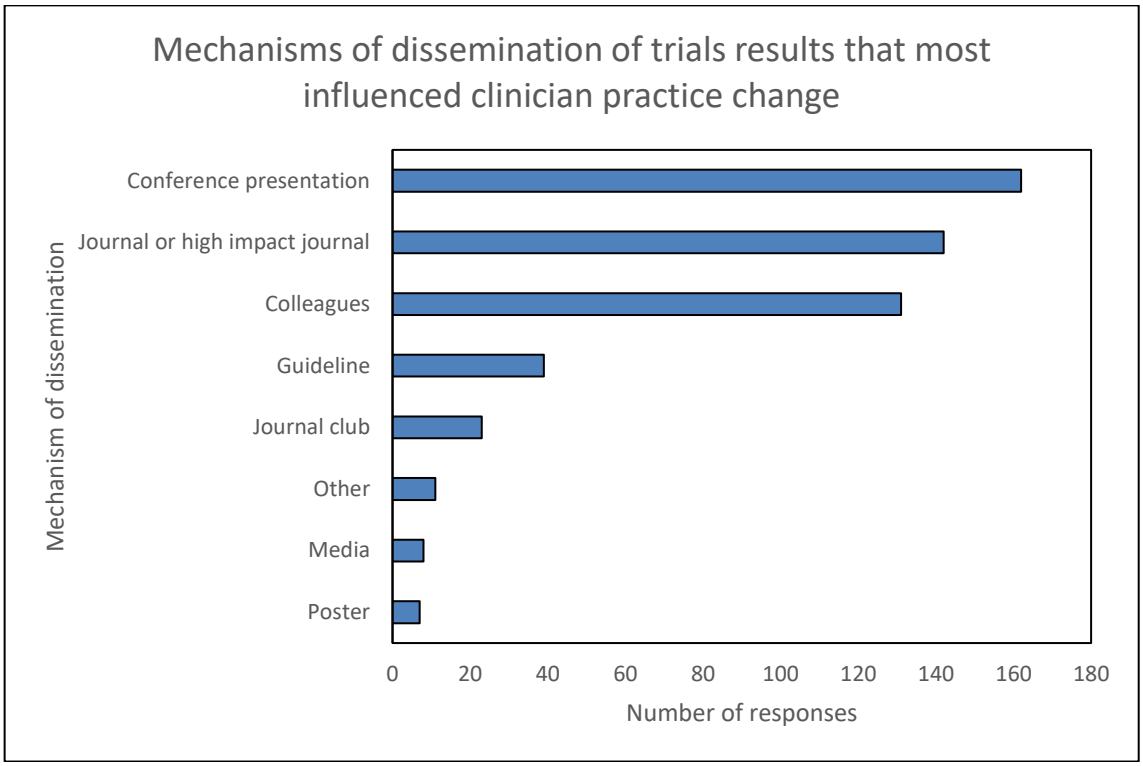


Figure 5-7 Mechanisms of dissemination of trial results that most influenced clinician practice change Overall, 243 respondents out of 265 answered this question. Respondents could chose more than one mechanism of dissemination (n=528 individual answers selected by 243 respondents indicating that on average (mean), respondents selected 2 mechanisms each).

The one UK clinician who did not change their practice explained that they were already prescribing 3 months of CAPOX chemotherapy for patients with "low-risk Duke's C disease", and 6 months of CAPOX for patients with "high-risk Duke's C disease". For them, the trial results were "confirmatory". The barriers to practice change chosen by the 18 international clinicians who reported no practice change were: a) the strength of evidence provided by recent clinical trials (72%; 13/18), b) the fact that no clinical guideline existed to support a practice change (11%; 2/18), c) their colleagues had not changed their practice (11%; 2/18) and d) they had not treated any patients yet who specifically fitted the trial criteria (6%; 1/18).

5.3.5 Timing of practice change

In the second survey, participants were asked at which time point they made their main practice change for patients with stage III colon cancer. The options were: June 2017 (IDEA and SCOT initial results disseminated), March 2018 (IDEA full publications), June 2020 (updated IDEA results). The same question was asked for patients with stage II disease but an additional time point was added:

June 2019 (IDEA stage II results presented in abstract form at ASCO conference). Overall, 82% reported their practice change for patients with stage III disease occurred either after the initial ASCO 2017 presentation or in response to the full journal publication nine months later (Figure 5-8).

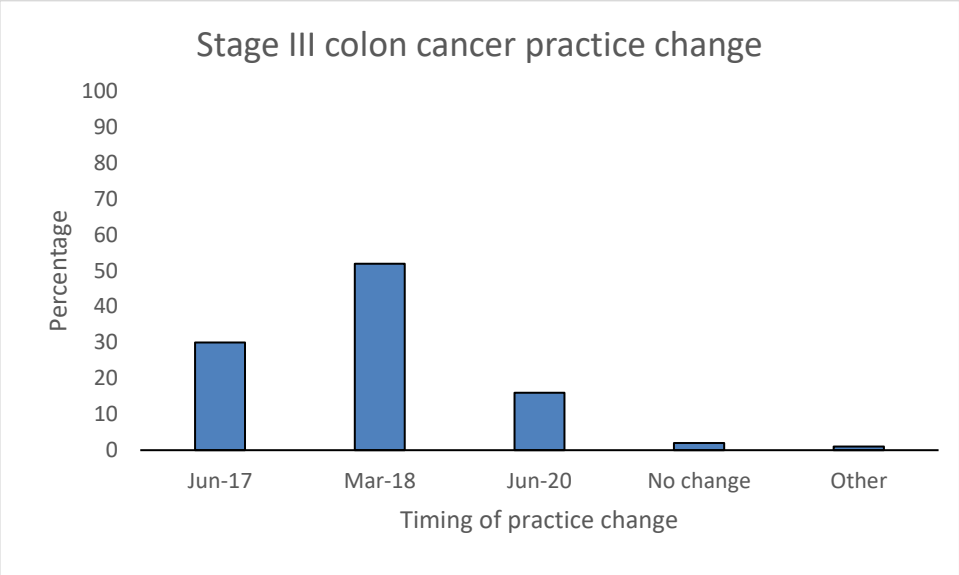


Figure 5-8 Timing of practice change (stage III disease) One clinician who chose a time-point that was not pre-specified stated that the timing of the SCOT trial results (exact time-point not indicated) had the most influence on their practice.

When describing practice change for patients with stage II disease, 57% of the second survey respondents indicated that they had changed their practice, with the largest proportion (20%) of practice change being attributed to the dissemination of stage II IDEA results at ASCO 2019 (Figure 5-9).

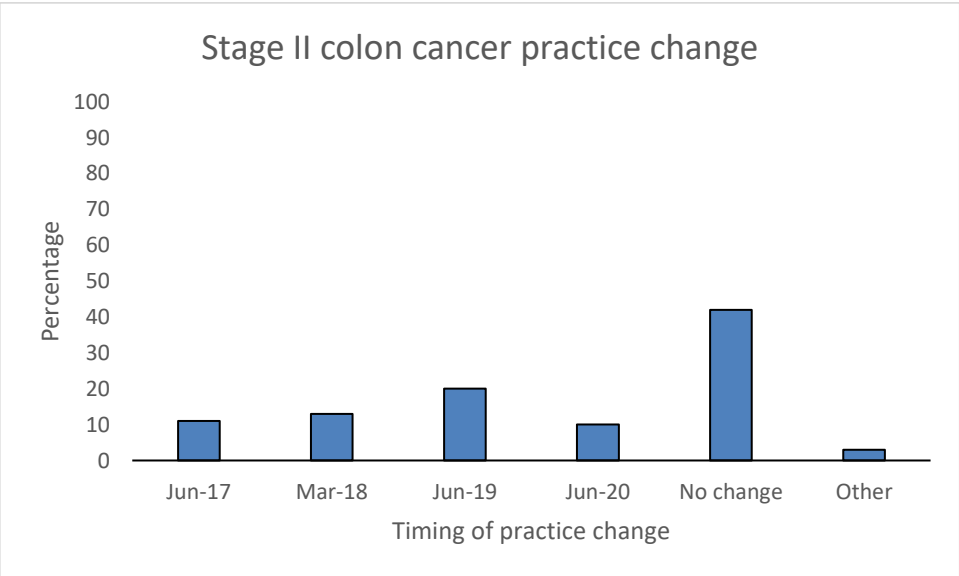


Figure 5-9 Timing of practice change (stage II) Of the three respondents who did not choose a pre-specified time-point, one stated the SCOT trial results had most influence on

their practice, one stated the ACHIEVE2 trial results and one stated: "Previous retrospective analysis of use of oxaliplatin."

5.3.6 Change in practice over time

5.3.6.1 Change in self-reported practice: Stage III

Disregarding any changes in practice driven solely by the COVID-19 pandemic in their responses, the biggest change in practice between April 2019 and August 2020 was that approximately one fifth of clinicians shortened the duration of treatment used for patients with stage III disease and one high-risk feature (T4 or N2). The change in regimen and duration of treatment chosen for patients aged under 70 by individual clinicians who answered both surveys are shown in Figure 5-10. There were no other changes affecting over 10% of clinicians in either the choice of regimen or duration of treatment. Clinicians used a similar treatment approach for low-risk stage III (SCOT non-inferiority met) and high-risk stage III disease (SCOT non-inferiority not met) with two high-risk features (T4N2) at both time-points. The results for stage III scenarios describing patients aged 70 and over are outlined in Appendix 3.

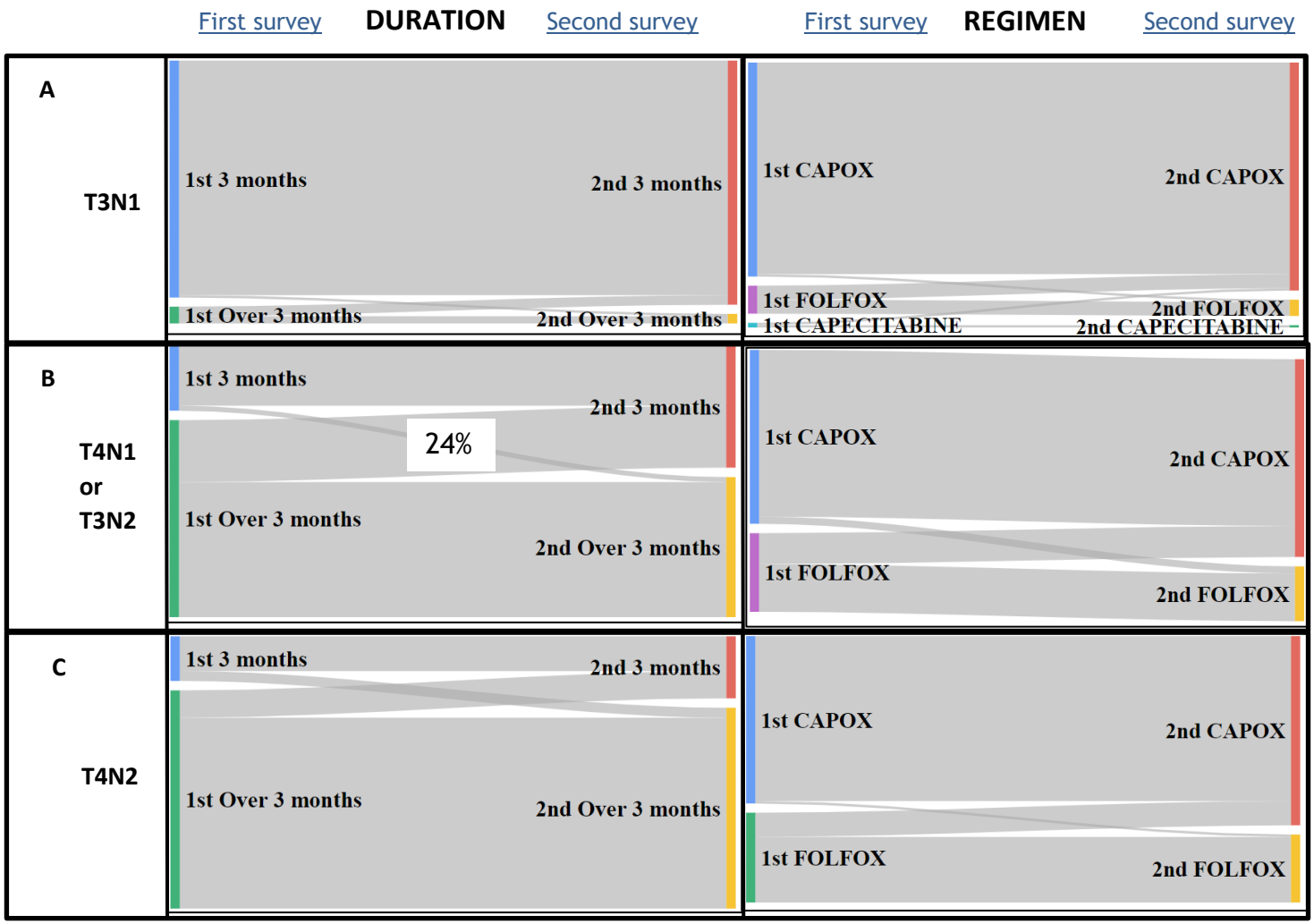

First survey
DURATION
Second survey
First survey
REGIMEN
Second survey
A
T3N1
1st 3 months
2nd 3 months
1st Over 3 months
2nd Over 3 months
1st CAPOX
2nd CAPOX
1st FOLFOX
2nd FOLFOX
1st CAPECITABINE
2nd CAPECITABINE
B
T4N1
or
T3N2
1st 3 months
2nd 3 months
24%
1st Over 3 months
2nd Over 3 months
1st CAPOX
2nd CAPOX
1st FOLFOX
2nd FOLFOX
C
T4N2
1st 3 months
2nd 3 months
1st Over 3 months
2nd Over 3 months
1st CAPOX
2nd CAPOX
1st FOLFOX
2nd FOLFOX

Figure 5-10 Change in practice between April 2019 - August 2020 for patients aged under 70 years old N=106 The first survey was carried out in April 2019 and the second survey in August 2020. Duration of treatment is on the left side of the diagram and regimens chosen is on the right side. Panel A (reading across) shows results for the scenario describing low risk stage III disease (T1-3N1) (SCOT non-inferiority met) and how choices for duration and regimen of treatment changed for the same clinicians between the first and second survey. Panel B shows the same results for two scenarios combined describing high risk stage III disease (SCOT non-inferiority not met) with one high risk feature (T4 or N2). Panel C shows the same results for the scenario describing high risk stage III disease (SCOT non-inferiority not met) with both high risk features (T4N2). The only change in opinion affecting over 20% of respondents is indicated in the diagram. The colours of the bars within this diagram are not significant.

5.3.6.2 Change in self-reported practice: Stage II

Longitudinal practice change for stage II patients was also assessed in both age groups. The scenarios were changed from describing a patient with T3/4 disease to T3 disease and T4 disease separately, meaning that a direct comparison of answers was not possible. For scenarios describing younger patients (Figure 5-11 below), 3 months of doublet chemotherapy (SCOT experimental arm treatment) was used by a minority of respondents across all scenarios at both time-points (April 2019 and August 2020). In 2020, it was clear that, in general, doublet chemotherapy was more popular for patients with T4 compared to T3 tumours. For T3 tumours, active monitoring (MSI-H) or 6 months of capecitabine (MSS) were used most frequently. A very similar pattern was seen for older patients, (Figure 12-10 in Appendix 3), although the use of doublet chemotherapy was less for patients aged 70 and over across all scenarios compared to treatment choices for younger patients. Also, active monitoring was the most popular treatment for T3 MSS tumours, compared to 6 months of capecitabine for patients aged under 70 years old.

In the second survey, clinicians were asked if they ever intentionally prescribe 3 months of fluoropyrimidine monotherapy and 22% (23/106) indicated it was a treatment strategy that they use. This question was tested indirectly in the scenario questions within both surveys, and Figure 5-11 below ('CAP3 and FU3'), and Figures 12-1 to 12-8 in Appendix 3, show that 3 months of capecitabine or 3 months of 5-fluorouracil were responses for stage II disease that were chosen by a minority of clinicians at both time-points. The directed question in the follow-up survey was complementary to this scenario information and helped to quantify the proportion of clinicians who would ever use this approach, regardless of scenario specific details.

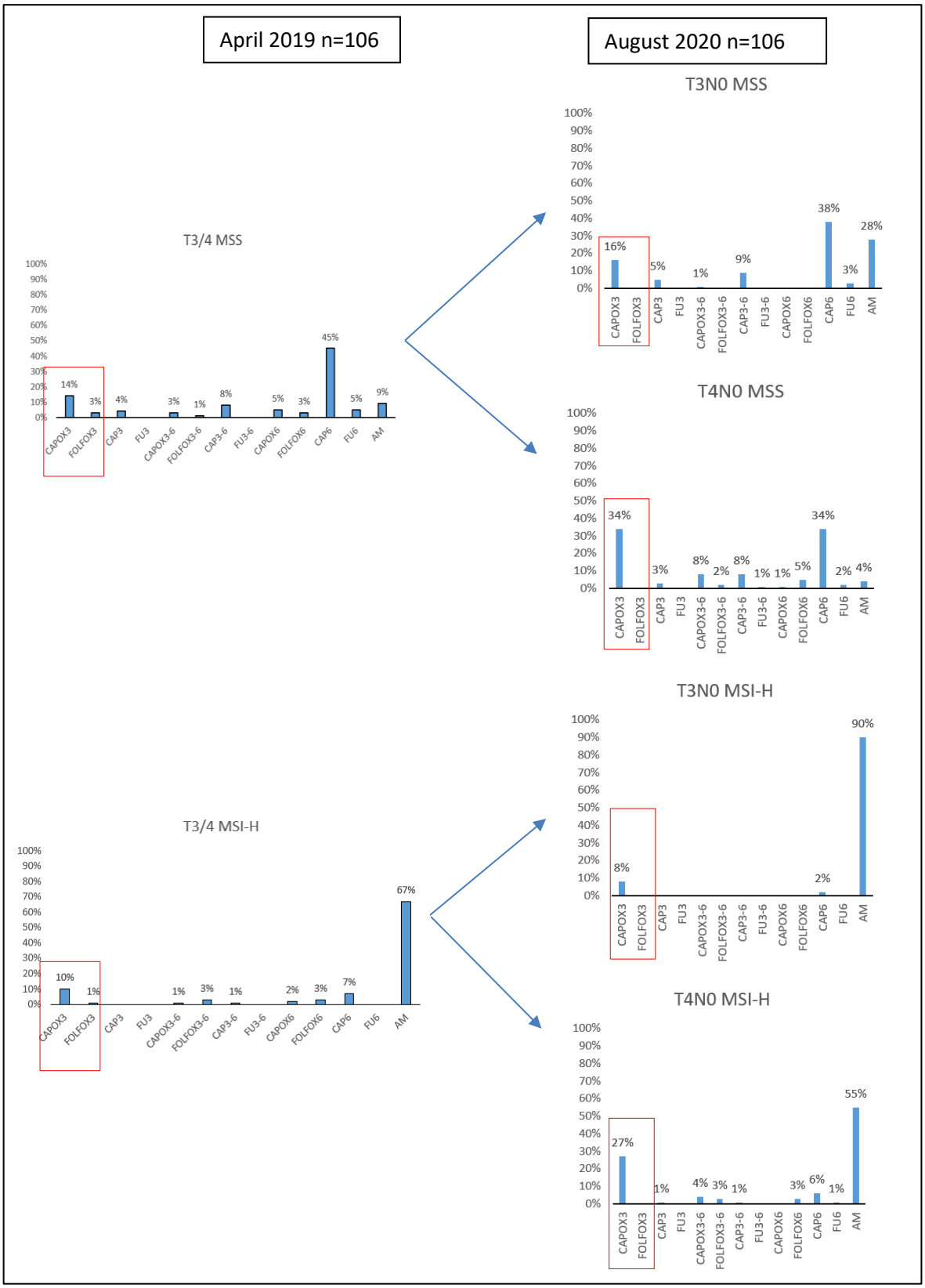


Figure 5-11 Treatment choices for patients aged under 70 with stage II disease Includes comparison of clinician choices between April 2019 and August 2020 (n=106). For 2019, this diagram includes responses only from individuals who answered both surveys. The red boxes indicate when responses align with the treatment delivered in the experimental arm of the SCOT trial/IDEA collaboration.

Specifically focusing on stage II patients, clinicians in August 2020 indicated they were more likely to use doublet chemotherapy for treating stage II disease irrespective of age after the results of the IDEA collaboration were known compared to prior to the dissemination of the IDEA findings (Figure 5-12). Despite this, there was still a minority of respondents who indicated they never used doublet chemotherapy for patients with stage II disease (16% (17/106) for patients aged under 70, 29% (31/106) for patients aged 70+).

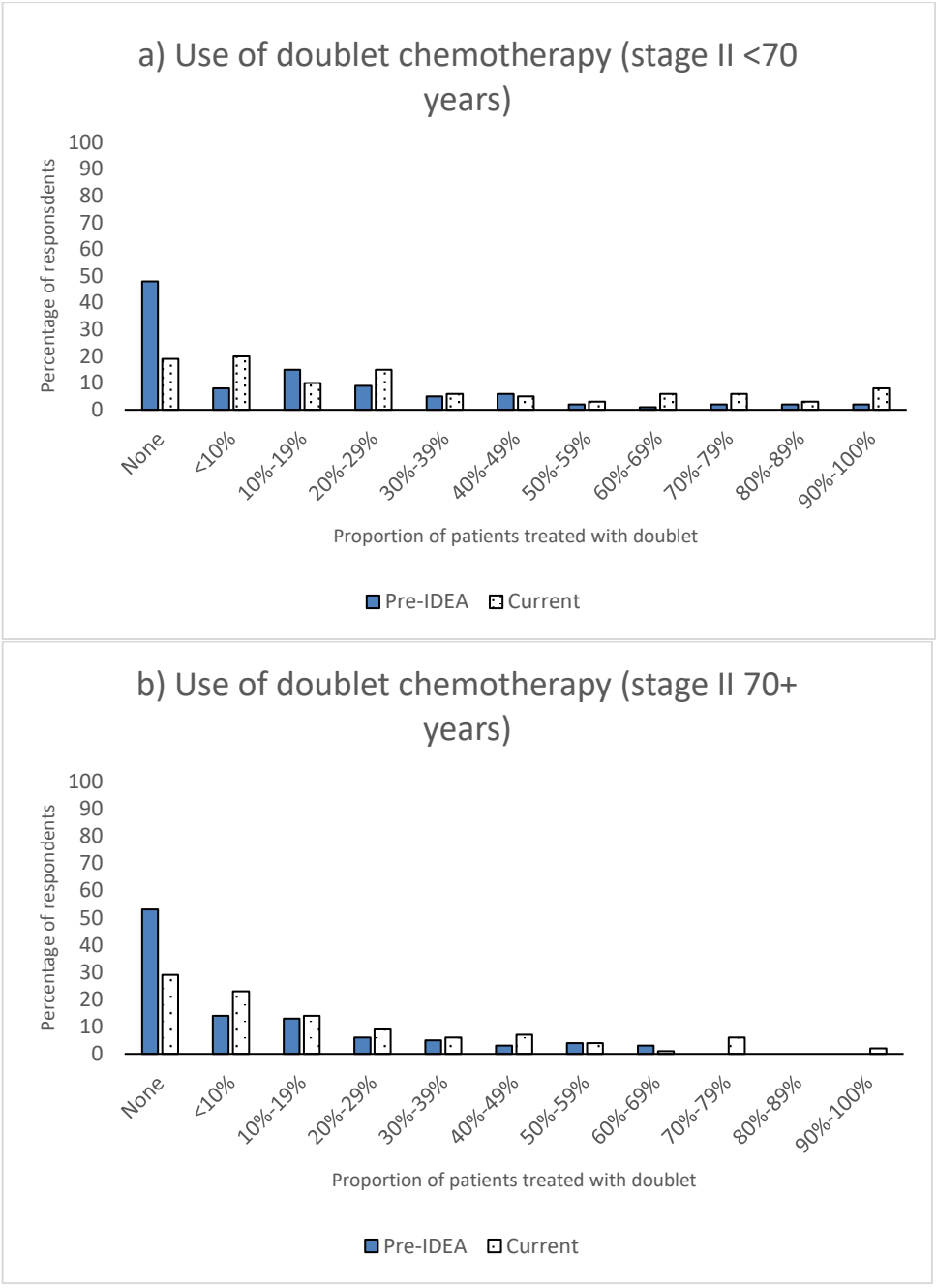


Figure 5-12 Use of doublet chemotherapy for patients with stage II disease a) aged under 70 years old b) aged 70 and over.

5.3.6.3 Change in attitudes

In keeping with their scenario related practice choices, in the first survey (Figure 5-13) clinicians most strongly agreed that CAPOX for 3 months could be considered a standard of care to treat patients with low-risk stage III disease (91%) (SCOT non-inferiority met). They most strongly disagreed that 3 months of FOLFOX could be a standard of care for high-risk stage III disease (81%) (SCOT non-inferiority not met) and there was most uncertainty (corresponding to clinicians indicating they neither agreed nor disagreed) when offering 3 months of CAPOX or FOLFOX for stage II disease (SCOT non-inferiority not met). Clinician opinions (n=106) generally remained consistent between April 2019 and August 2020 (see Figure 12-1Appendix 3). The largest changes were an increase in agreement that 3 months of FOLFOX could be an acceptable standard of care for patients with low-risk stage III disease and an increase in agreement that 3 months of CAPOX is an acceptable standard of care for stage II disease with high-risk features. There was a corresponding rise in disagreement with 3 months of FOLFOX as a standard treatment for stage II disease.

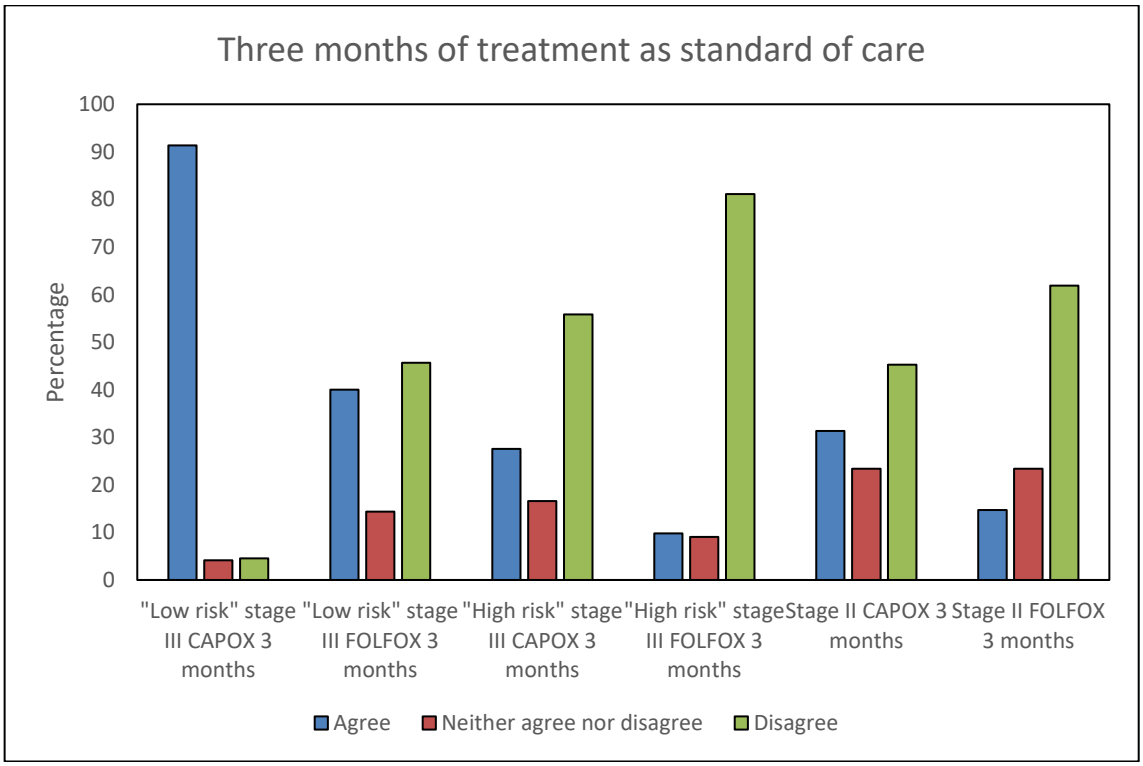


Figure 5-13 Clinician attitudes to using 3 months of treatment as standard of care for the adjuvant treatment of patients

5.3.6.4 Enduring adjuvant practice changes due to COVID-19

There were no significant differences (Fisher's exact test $p<0.05$) between management choices disregarding the influence of the COVID-19 pandemic and those made reflecting enduring changes to practice because of the COVID-19 pandemic from the August 2020 survey. This demonstrates that, even if a temporary alteration to clinician practices did occur early in the pandemic, respondents do not feel that the pandemic will have enduring changes to their practice for patients with colon cancer in the adjuvant setting. The difference in practice reported in August 2020, taking into account enduring changes from the COVID-19, are shown in Appendix 3 Table 12-5.

5.4 Discussion

This study surveyed of a large number of clinicians, all of whom confirmed they prescribed adjuvant chemotherapy to patients with CRC. The results of the IDEA collaboration were clearly well known within this group of respondents, and in particular, the SCOT trial results were well known in the UK. The increased awareness of SCOT for UK versus international clinicians may be linked to the fact that many of the UK respondents will have recruited patients to the SCOT trial (>4,000 SCOT patients from the UK). This provides a clear message to trialists that active efforts to disseminate their trial findings may be needed in countries where their trial did not recruit. It also supports the findings from Lewison et al (188) (Chapter 3) that UK guidelines, which are in part written by UK clinicians, more frequently cite UK research than would be expected based on the world literature.

The mechanisms of dissemination in the first year since the IDEA and SCOT publication that most influenced practice indicates to trialists that channels of communication and dissemination of results other than the peer-reviewed publication are important and worth focusing on in order to maximise trial impact. Guidelines were less important in this survey. Rather than indicating that guidelines do not influence practice per se, this response may reflect the time lag to widespread change in the relevant guidelines in this instance. For example, the NICE guidelines in the UK had not been updated when the first survey was distributed (254). This result prompts reflection on the high priority placed on

policy change, in particular in the REF 2014 case studies of cancer clinical trials (see Chapter 4), in particular the importance placed on policy change as a pre-requisite for practice change. Few respondents indicated that social/lay media articles or conference posters were important mechanisms of dissemination of trial results. This may reflect either that the results of these particular trials were not disseminated via those mechanisms or that these are ways of communicating the results of clinical trials that have less influence on practising clinicians.

Regardless of the location of respondents, the main message from this study was that a large proportion of clinicians have altered their practice in response to IDEA and SCOT, and that these trials can be considered practice changing. The impact of these trials on attitudes and practice occurred to a major extent by April 2019, with minimal change in the following 18 months. This demonstrates there was a very short time lag between dissemination of initial results and implementation into practice to some extent.

There was a stronger indication from UK clinicians that they regarded their practice as having changed in response to the trials they had named. This may be linked to the fact that more UK clinicians were aware of the SCOT results in April 2019 and that SCOT met its non-inferiority margin statistically, whereas this was not the case for the IDEA collaboration. In addition, it is know from IDEA (81) that UK clinicians prefer to use CAPOX whereas clinicians in locations such as France and the USA preferred FOLFOX. Given the stronger evidence for using shorter duration for CAPOX compared to the evidence for FOLFOX, this may also have influenced this difference between UK and international responses.

This study has shown that practice change has mainly occurred in line with the stratification of stage III disease into low and high-risk. This stratification was coined based on results from the IDEA collaboration, and is itself an unexpected impact from these trials. It is especially surprising that this division in treatment approach was maintained in August 2020 given that updated IDEA collaboration results showed very little clinical difference in 5-year OS between using 3 versus 6 months of CAPOX in the high-risk stage III setting. This perhaps indicates the power of the initial trial results in drawing a line along which clinicians are and are not happy to reduce treatment duration (83). An opportunity for an implementation study would be to investigate in more detail the barriers that

may exist that are preventing clinicians from changing practice for this group. These barriers may include patient actual preference or preference as perceived by clinicians, the strength of evidence, clinicians' own perception of risk and non-inferiority of shorter treatment in a higher risk setting, or prescribing practices suggested by colleagues and guidelines.

Results from a recently reported within-trial survey (255) of SCOT trial participants from Australia and New Zealand has shown that, on average, patients indicated the benefits of treatment that would warrant taking 6 months rather than 3 months of adjuvant treatment were 2-3 years extra survival beyond 5 or 15 years life expectancy, or 5-15% extra survival beyond a 5 year survival rate of 65-85%. These benefits are larger than then benefits of an extra 3 months of chemotherapy calculated in SCOT/IDEA. These results differ from a previous survey by the same group which assessed preferences for receiving 6 months of adjuvant chemotherapy for CRC versus observation, when the benefit required to accept chemotherapy was only 1 month extra survival or 1% survival gain on 5 year survival (256). These results show the importance of duration of treatment, and reflects that patients put high importance on receiving some active treatment in the adjuvant setting, but they consider carefully the balance of toxicity versus benefit of this treatment.

Clinicians also indicated they were less likely to have changed practice for patients with stage II disease. Indeed, in the August 2020 survey, some clinicians indicated they never use doublet treatment for patients with stage II disease. The stage II IDEA findings aligned closely with results for the stage III population (87), therefore the reduced uptake of study results into practice may reflect a lower use of doublet chemotherapy in this setting rather than the strength of trial evidence.

The avoidance of doublet chemotherapy in this context is highly likely to be influenced by the results of previous trials discussed in Chapter 1. For example, the subgroup analysis from the MOSAIC trial which failed to confirm an overall survival advantage from adding oxaliplatin to fluorouracil specifically for stage II patients (257). This highlights an opportunity for clinicians to improve certainty in this area by increasing dialogue around the approach to treatment of this group. It also may encourage clinician trialists responsible for the stage II IDEA

collaboration to focus on dissemination of these results specifically and to help interpretation of those results for clinicians in the context of previous trials assessing the merits of using doublet treatment specifically in this patient cohort.

This study has shown that patients with stage II MSI-H disease, regardless of age, are more likely to be offered doublet treatment or avoid adjuvant therapy altogether, compared to those with MSS stage II disease. This aligns with the results of previous clinician surveys performed before IDEA, (258) and although a minority of CRC patients have MSI-H disease (259), reinforces the case for ensuring this information is available to clinicians at the time of decision making. It also indicates that recent trial results investigating shorter duration of doublet treatment specifically are more likely to influence prescribing for MSI-H rather than MSS stage II disease for which proportionally, doublet chemotherapy is used less. Results according to MSI status are not available for the SCOT trial, however further information on molecular influences on disease outcomes dependent on treatment duration may become available with the analysis of tissue samples from SCOT trial patients in the TransSCOT programme (260).

The extent of self-reported practice change in response to recent trials was also less for older versus younger patients, concurring with results from the recently published survey of French clinicians (250). Reduced impact on practice for older patients is likely to again reflect the less frequent use of doublet chemotherapy for older versus younger patients generally. This may be due in part to previous individual and pooled trial subgroup analyses showing a lack of benefit from adding oxaliplatin to fluoropyrimidine for older patients (56, 65, 257) (see Chapter 1), although there are reports that some benefit of oxaliplatin may be maintained in older patients (261, 262), especially based on more recent trial findings (68). It also raises the bigger question of the gap in evidence for this age group and the need for randomised clinical trials specifically in this area.

Although an age cut-off was chosen for the practicalities of survey development, the vast majority of clinicians indicated they use biological rather than chronological age when making treatment decisions. It is therefore

acknowledged that clinicians are unlikely to alter their practice across such a strict age cut off in real life. This reflects the difficulty with making any treatment decisions based on age alone and explains why most national guidelines from professional bodies (NCCN, ESMO, ASCO) do not mention age. Recently updated NICE CRC guidelines do mention that age is taken into consideration but do not give direction on how this may specifically affect treatment choices (254).

As described in Chapter 1, the SAFFA trial (51) compared 3 months of protracted venous infusion 5-fluorouracil against 6 months of bolus 5-FU/leucovorin in patients with stage II/III CRC (n=801) and showed there was no OS difference between the treatment arms. This strategy has not yet been tested using modern infusional regimens, such as modified deGramont, in both arms. In the second survey, clinicians were asked if they ever intentionally used 3 months of adjuvant fluoropyrimidine monotherapy for patients with stage II disease and a minority of respondents indicated this was a treatment approach that they adopted. Ideally, it would be useful to ask this question directly regarding patients with stage III disease also, although the scenario analyses give us insight as to the likely answer. In both surveys, no clinicians chose 3 months of fluoropyrimidine monotherapy for scenarios described patients with stage III disease aged under 70. For patients with stage III disease aged 70 and over, 3 months of fluoropyrimidine was occasionally chosen by a small number of clinicians. Further investigation into clinician opinion would be helpful to understand if those who routinely use 3 months of monotherapy for stage II or stage III disease are applying the results of the SAFFA trial, or if they are extrapolating the results of the IDEA collaboration when they make this treatment choice. These results also raise another gap in existing trial evidence for clinicians making decisions in this field.

The results from this study show that, if dependent on clinician preference alone and not constrained by institutional policies, there are unlikely to be significant enduring changes due to COVID-19 going forward that differ from practice decisions based on clinical trial evidence alone. There is still a lack of published data on actual chemotherapy prescribing during the peaks of the pandemic and practice currently. More information is needed to map the effect of the

pandemic on adjuvant prescribing for CRC, especially on the choice of duration of treatment and regimens used.

The findings from these surveys support the results from previous studies performed by Iveson et al (248) and Yu et al (249, 250) in that most clinicians agreed that 3 months of adjuvant doublet chemotherapy can now be used for 'some' patients with CRC. The findings also strongly supported results from those previous surveys (248-250) and a survey of French clinicians (250) which all demonstrated the strong influence of stage III risk classification on duration of treatment prescribed. Specifically, 3 months of doublet chemotherapy was used often for low-risk stage III disease (SCOT non-inferiority met), whereas 6 months was still used for high-risk disease (SCOT non-inferiority not met). There does seem to have been a shift over time across these surveys, with more respondents indicating they would reduce treatment duration, to align with the experimental arm of the SCOT trial and IDEA collaboration, in both the low-risk and high-risk setting in this study compared to the one performed in September 2017 by Iveson et al (248). For example, in September 2017, 56% of clinicians used 3 months of treatment for low-risk disease compared to 85% for younger patients and 60% for older patients in this survey. For patients with high-risk stage III disease, in the September 2017 survey, 12% of clinicians used 3 months of treatment; this compared to 16% in younger patients and 17% in older patients in this survey. There was no distinction made between age groups in the 2017 survey for direct comparisons of these responses.

There was a strong preference for using CAPOX in the low-risk stage III setting was identified across previous surveys and those reported in this study. Iveson et al (248) demonstrated a split in using CAPOX versus FOLFOX for high-risk stage III disease, similar to the findings of this study, whereas in the survey of French clinicians (250), most used FOLFOX. These differences are likely to represent the locations of respondents within the surveys and the regimen preferences that existed pre-SCOT and IDEA. For example, most patients in IDEA-France (263) received FOLFOX, all patients in CALGB-SWOG 80702 (USA/Canada) (264)were given FOLFOX, all patients in ACHIEVE (Japan) (265) received CAPOX and the majority of patients treated in the UK within the SCOT trial (41) had CAPOX. This study was the first to report survey results for patients with stage II CRC and the

only study to include two surveys at distinct time-points, indicating a change in the approach to treating high-risk stage III patients over time, in particular the increased use of 3 months of treatment for patients with one high-risk factor (T4 or N2).

The strengths of this study include the relatively good response rate in the UK and a good spread of locations from outside the UK. Rather than providing clinicians with a list of trials and asking them if they were aware of specific trials, this survey required that clinicians wrote the name or primary author of the trial. This is a less biased test of awareness of the clinical trials of interest compared to the approach used in previous surveys (249, 250). This study offers a snapshot of clinician practice one year following the results. The surveys conducted only in France (266) and the one that mainly had respondents from Canada, USA and South America (249) can be seen as complementary, given the low numbers of respondents from France and Canada in this study. This is the also only study to look at clinician views over two time points.

This study also has limitations. A sample of clinicians were surveyed and there will be inherent response bias in any sample. Specifically, those who responded may have been more enthusiastic, more likely to be aware of trial findings and more motivated to change practice compared to survey non-responders. Response rate was captured for UK CRC oncologists but it was not possible to estimate a response rate for those clinicians located outside the UK, or who were not on the pre-specified list of UK CRC oncologists. There was also no information regarding the non-respondents amongst the UK CRC oncologists and how they differed from the respondents.

In the development of the surveys used in this study, consideration was given to the optimal way of asking clinicians about practice change in response to the results of the SCOT trial findings. Initially, the plan was to ask clinicians to recall their practice pre-SCOT trial publication and to compare this to current practice post-trial. Advice from an expert in impact evaluation (104) indicated that this approach would be open to recall bias which would undermine the validity of the responses, especially given the time period (years) from which respondents would be asked to recall their previous practice. It was identified that the

optimal way to analyse how practice has changed would have been to perform a repeated measures study with a separate survey before and after the trial publication. This was not possible due to the timeline of this study, which commenced after the SCOT trial had ended and was published. Therefore, the option of a survey shortly after the initial publication, with a follow up survey one year later, was the appropriate option to gauge initial short term and longer term changes in practice.

The impact of these trial findings specifically for patients with rectal cancer, and how this may have varied compared to colon cancer, was not explored and warrants further investigation. In addition, the impact of deficiency in MMR proteins on prescribing for stage III disease was not analysed. This may become more relevant as novel agents such as immune-checkpoint inhibitors are tested in the adjuvant setting. Also, the results of the FOXTROT trial, (267) presented at ASCO 2019, showed that six weeks of neoadjuvant chemotherapy for patients with colon cancer was well tolerated and can improve surgical morbidity. The work in this chapter did not explore how the FOXTROT results may have influenced the impact of the IDEA trials, although even if chemotherapy is moved to earlier in the disease trajectory, it is likely that the importance the IDEA results on duration of practice will still be relevant. Lastly, the patient scenarios in the surveys used in this study asked clinicians to assume patients were fit (performance status 0-1) and to indicate their choices for the majority of patients they treat. In a real world stetting these choices may be diluted based on patient, disease and other external variables. For this reason, an analysis of actual prescribing records would be a stronger indicator of current practice, although would not capture clinician perspectives and attitudes as has been achieved in this study.

5.5 Conclusion

In conclusion, the IDEA collaboration and contributory trials have impacted on clinician attitudes and self-reported practice. This study suggests that across several countries, patients with CRC are now being offered shorter adjuvant treatment. These changes will hopefully lead to the health benefits of less toxicity whilst maintaining survival gains. This study has shown that surveying research users, in this case clinicians, is a useful tool to analyse the health

sector impact of a clinical cancer trial. The results from this study will be utilised in Chapter 6 to inform estimation of the potential downstream health and economic effects from the implementation of trial findings.

6 Chapter 6: Assessing the economic value of implementation of the SCOT trial findings

6.1 Introduction

A number of the REF case studies analysed in Chapter 4 described the economic impact of clinical cancer trials; these impacts usually related to sales of drugs or describing the results of cost-effectiveness analyses conducted alongside trials, rather than predicting or evidencing economic gains from the trial findings.In this chapter, an adaptation of the methods used by Glover and colleagues (39) and Brown et al (184) were tested as an approach to evaluate the impact of the SCOT trial. Specifically, the aim of this study was to estimate the economic value of implementing the SCOT trial findings in the six countries that recruited to SCOT.

To meet this aim, the following objectives were identified:

i) Calculate the cost-effectiveness of the treatments tested within the SCOT trial from the perspective of the six countries that recruited to SCOT: Australia, Denmark, New Zealand, Spain, Sweden, and the UK.

ii) Estimate the budget impact of implementing the SCOT trial findings in these six countries using the costs calculated in part (i) of the study.

iii) Estimate the quality adjusted life year (QALY) gain from implementation of the SCOT trial findings across the six countries and calculate the monetised value of that QALY gain across a range of willingness to pay (WTP) thresholds.

iv) Use the results from parts (i)-(iii) of this study to estimate the return on investment in the SCOT trial.

6.2 Methods

The methods for this analysis align with the study objectives, and are divided into four parts: i) a cost-utility analysis ii) budget impact analysis iii) estimation of QALY gains at a population level and iv) using the results from the previous three steps to estimate the return on investment in the SCOT trial.

6.2.1 Cost-utility analysis

The main purpose of undertaking any economic evaluation is to evaluate all the changes in costs and benefits associated with a new treatment approach in comparison to the current standard of care. By allowing a comparison to alternative treatment options that are currently used in practice, it is possible to make informed decisions about what is the most cost-effective course of action.

Cost-utility analysis is a commonly used type of economic evaluation which compares the costs and outcomes of alternative interventions, where the outcomes are measured in terms of utility, which is then combined with data on survival to calculate adjusted life expectancy. The common outcome is QALYs (268), which incorporate both length of life (an assessment of the time affected by a health outcome) by the quality of that time period (the preference for being in that health state on a scale of 0 to 1) (268). Cost-utility analysis is the recommended method by many reimbursement bodies globally because survival and quality of life are important clinical trial outcomes and the generic QALY outcome from cost-utility analysis allows for comparisons across unrelated interventions and disease areas, which is particularly helpful for decision makers (112).

For the purposes of this analysis, a within trial cost-utility analysis was performed using the data from the SCOT trial. This was an update of a previously performed cost-utility analysis of the SCOT trial from a UK perspective (75). The objectives of undertaking this additional analysis were to (a) integrate updated OS and DFS data from the SCOT trial, which became available in 2019, and (b) to perform the analysis from the perspectives of each of the countries that participated in SCOT. Both of these objectives add novelty to the study and differ from the original SCOT economic analysis.

6.2.1.1 Outcomes

Survival was calculated using partitioned survival analysis and a within trial time horizon (75). As is recommended good practice in conducting a within trial cost-utility analysis for survival and costs extending beyond one year (112), discounting (269) of survival times was performed to adjust future health effects so that gains in survival occurring after year 1 were valued less than those occurring in the first

year of treatment (positive rate of time preference). Survival outcomes were available for up to 10 years for 4,592 patients, which represented longer follow up compared to the previous UK analysis that utilised survival data up to 8 years (75). Partitioned survival analysis was used to estimate the mean survival in three health states (time on treatment (ToT), DFS and recurrence).

Quality of life was estimated using patient-level responses to the EuroQol-5 dimension-3 level (EQ-5D-3L) questionnaire, (270) which were converted into utilities using preference weights. Value sets for preference weights specific to each country (271-274) were applied sequentially to the EQ-5D profiles for all patients in the SCOT trial for whom EQ-5D responses were available. EQ-5D results were only available for a sub-group of patients (n= 1 832) because collection of this data ceased after a pre-planned testing period to ensure sufficient information to inform analysis had been collected (275). In order to apply EQ-5D to the whole population, linear regression was performed to adjust for health state, regimen received, disease risk stage, age, gender, and ethnicity. This model differs from the one used in the previous UK perspective analysis by the addition of an extra co-variate representing disease stage. Specifically, rather than dividing patients into those with stage III high-risk disease (T4 or N2) and "other", in this analysis, patients were divided as either having stage II, low-risk stage III or high-risk stage III disease. In this model, 'i' refers to each patient and 't' to time from randomization:

$$Utility_{it} = \beta_0 + \beta_1 OnTreat_{it} + \beta_2 Recur_{it} + \beta_3 Arm{:}\,6M_i + \beta_4 CAPOX_i + \beta_5 LRSIII_i + \beta_6 HRSIII_i + \beta_7 MALE_i + \beta_8 AGE_i + \beta_9 AfrA_i + \beta_{10} SouA_i + \beta_{11} Chinese_i + \beta_{12} OtherEt_i + e_{it}$$

Where patient health state (on treatment and recurrence) could change over time. Standard errors were clustered at the patient level.

To obtain the final outcome measure for the cost-utility analysis (197), average QALY outcomes for each health state from the regression model and each arm were estimated by adjusting the mean utility for each health state by the survival for the same health state using an integrated quality-survival product (276). The life years for the QALY calculation were discounted (to reflect the present value

of QALYs incurred in the future (269)) at a rate of 3.5% as per recommended good practice.

6.2.1.2 Costs

The costing process in any economic evaluation involves three stages: identification of appropriate cost items, measurement of the resource use, and valuation – applying monetary unit costs related to the measured resources and multiplying them (277). In this analysis, for years 0-6, individual per patient costs were calculated by multiplying the amount of resource used (mg of drug or number of nights/days attending hospital) by the appropriate unit cost (see Chapter 2). For inpatient and day cases occurring when patients received adjuvant chemotherapy, pharmacy costs were subtracted from these unit costs to avoid counting drug costs twice. On treatment and follow up time was divided into pre-specified time points and mean costs accrued for each arm of the trial over pre-specified time intervals was estimated. Specifically, the pre-specified time-points used in this analysis aligned with the time-points at which follow up information was collected within the SCOT trial (see Table 13-5 Appendix 4). These average costs were adjusted for censoring using the Kaplan-Meier Sample Average (KMSA) method (75). Specifically, mean costs were adjusted for the probability that a patient in that arm of the trial survived to the start of the time-point. The probability of being alive was interpolated from the Kaplan-Meier survival curves (Kaplan-Meier estimator) (278, 279). The following formula outlines how the KMSA calculation was performed:

$$Expected\ cost = \sum Probability\ alive\ at\ time\ j\ x\ mean\ cost\ at\ j$$

As extended follow-up was undertaken beyond the initial trial time horizon, survival data was available for a longer duration compared to the within trial cost data, therefore a model was used to estimate annual costs for the years when actual costs data was not available (7-10 years). The same model of linear regression for estimating utilities was used, but incorporating an adjustment for year 1 costs that included chemotherapy medication and a higher use of hospitalization resources compared to all other years.

$$Yearly\ costs = \beta_0 + \beta_1 OnTreat_{it} + \beta_2 Recur_{it} + \beta_3 Arm{:}\ 6M_i + \beta_4 CAPOX_i + \beta_5 LRSIII_i + \beta_6 HRSIII_i + \beta_7 MALE_i + \beta_8 AGE_i + \beta_9 AfrA_i + \beta_{10} SouA_i + \beta_{11} Chinese_i + \beta_{12} OtherEt_i + \beta_{13} Year1_t + \beta_{14} Year1 * Arm{:}\ 6M_i + e_{it}$$

The average annual costs estimated using the model were also adjusted for the probability of a patient surviving to the start of that year using the KMSA method. Discounting was applied to all costs incurred after the first year of treatment.

6.2.1.3 Cost-utility

In order to compare the costs and effects to determine cost-utility, the incremental net monetary benefit (INMB) approach was used (280). More commonly an incremental cost-effectiveness ratio (ICER) is reported for cost-utility analyses and requires a WTP threshold (representing society's/reimbursement agencies' willingness to pay per QALY) to determine if the QALY gain from a novel treatment is cost-effective compared to the standard of care (268). For example, in England, the National Institute for Health and Care Excellence (NICE) use a WTP threshold of £20,000-£30,000 per QALY for making their recommendations (197). In an ICER calculation, a novel treatment is assessed as cost-effective compared to the comparator treatment if the difference in costs divided by the difference in QALYs between the treatments is less than or equal to the WTP threshold. The INMB incorporates the WTP threshold into the calculation:

$$INMB = (WTP\ x\ difference\ in\ QALY) - Difference\ in\ cost$$

An INMB over zero indicates that an intervention is cost-effective compared to usual care at the specified WTP threshold. In some countries, a WTP may not be used, or it may not be explicitly stated in public documents even if one is used (281).

For this analysis, the monetary value of one annual gross domestic product (GDP) per capita for each country was used as a WTP threshold. This was based on the WHO guidelines for appropriate cost-effectiveness thresholds, which suggest using a value of 1-3 times the GDP per capita (282) and was used as the best available

benchmark that would be consistent across countries. Using country specific WTP thresholds would have been inconsistent and often they are not explicitly expressed by agencies for many countries. The lower end of the WHO range was used because it is recognised that the WTP per QALY of individuals or populations are often below this GDP per capita. Also, although proponents of using the QALY as an outcome measure in economic evaluations suggest that the value of a QALY is the same under all conditions, there is some evidence that the WTP of individuals for a QALY may vary based on if a treatment is life extending or life improving. Lower value may be placed on life improving treatments, such as the shorter duration of chemotherapy under evaluation in this study (283).

Six separate fully pooled, one country costing cost-utility analyses were conducted, one for each country, followed by a fully pooled, multi-country costing approach for the purposes of subgroup analysis. For the one country costing approach, resource use and outcome results for all patients in the trial were used, and single country specific unit costs and utility values were applied to all patients. For the fully pooled, multi-country approach, unit costs specific to the country patients were recruited from were used. A healthcare system perspective was adopted and calculations were performed in USD unless otherwise specified.

6.2.1.4 Sensitivity analysis

Probabilistic sensitivity analysis (284) was undertaken to address and quantify potential sampling uncertainty within the cost-utility analysis (285). Bootstrapping (110, 286) with 1000 iterations was the method utilised to perform this sensitivity analysis. Bootstrapping is the method of re-sampling with random replacement of variables from the original sample to create an empirical distribution to act as an estimate of the true population distribution from which the sample was drawn. This bootstrapped distribution was used to calculate 95% confidence intervals (CIs) around the primary results. Deterministic sensitivity analysis (284), used to account for variations in a specific input parameter or set of parameters, was used in this study to calculate the incremental NMB over a range of WTP thresholds. Finally, cost-effective acceptability curves (CEACs) (287) and cost-effectiveness planes were used to illustrate the probability of cost-effectiveness of the two treatment durations over a range of WTP thresholds

(288). A main CEAC for the overall trial results was produced, as well as CEACs for important patient, disease and treatment related subgroups.

6.2.1.5 Subgroup analysis

Subgroup analysis was performed using a fully pooled analysis in terms of outcomes but by using a multi-country costing approach. This meant that unit costs specific to each country were applied to patients in the SCOT trial that were recruited from that country. Given that most patients were recruited from the UK, unit costs from this location were given the most weight in the analysis. The subgroups chosen for further investigation aligned with the important pre-planned and post-hoc subgroup analyses from the SCOT trial (41) and IDEA collaboration (81): extended risk stage and treatment regimen. Differences in cost-effectiveness between trial arms by gender and age group were also investigated. Age was categorised into under 70 versus 70 and over, as this was a clinically meaningful split.

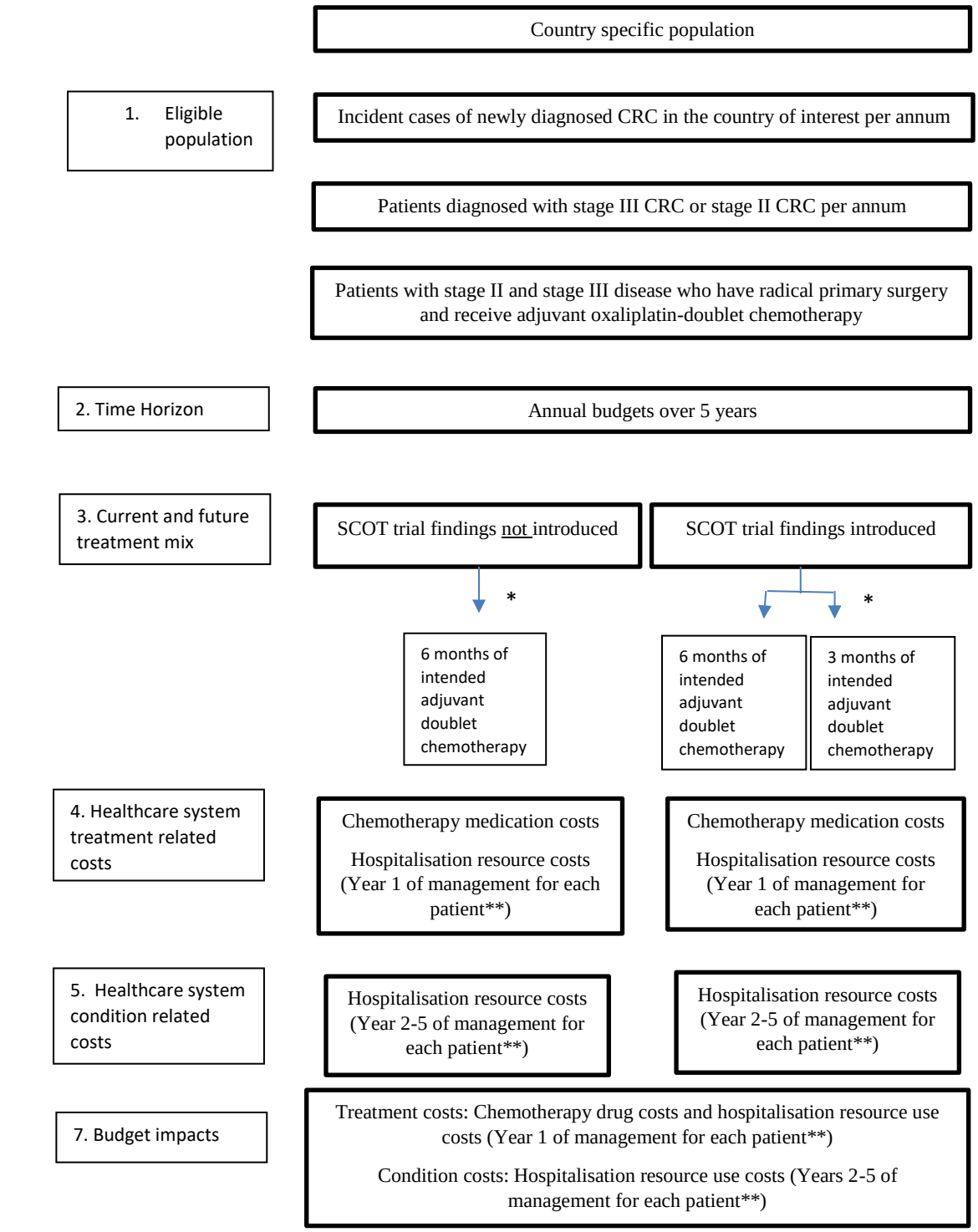
6.2.2 Budget impact analysis

A budget impact analysis was the method used for the second part of this study. The purpose of a budget impact is the 'forecast of rates of use (or changes in rates of use) with their consequent short and medium-term effects on budgets and other resources to help health service managers plan changes that result from the introduction of a new technology.' (Dictionary of Health Economics from (289)) Budget impact differs from the types of economic analyses described above in that the effectiveness of treatments is generally not considered and costs are considered over a short-term time horizon that would be relevant to budget planning (1-5 years, rather than potential lifetime costs for an economic evaluation). Budget impact analyses are most often used by reimbursement bodies after an initial economic evaluation, to help estimate the likely financial consequences for a healthcare system of a decision to fund a novel therapy.

6.2.2.1 Budget impact framework

Figure 6-1 outlines the analytical framework used for the budget impact analysis. The eligible population for this budget impact analysis consisted of patients diagnosed with stage II or III CRC who receive adjuvant oxaliplatin-

based doublet chemotherapy. The budget impact of SCOT trial implementation was calculated by estimating the change in healthcare costs if a specified proportion of patients were treated with an intended 3 months, rather than 6 months of chemotherapy after the publication of the SCOT trial results. An assumption was made that any patient receiving doublet chemotherapy in the adjuvant setting pre-SCOT would receive 6 months of treatment. A summary of the values used in the budget impact analysis and a list of the sources is provided in Appendix 4.



* This framework specifically focuses on the use of 6 versus 3 months of doublet chemotherapy. Other treatments that may be used, e.g. 6 months of fluoropyrimidine monotherapy. The doublet regimen assessed in this study is CAPOX or FOLFOX. In real life, occasionally patients have a contra-indication to fluoropyrimidine-based chemotherapy and may receive an alternative, such as raltitrexed, S1, UFT/LV or TAS-102 (depending on the country of treatment).

**All patients enter the model at year 1 of their treatment. For those entering in year 1 of the model, costs for years 1-5 will be included. For those entering in year 2 of the model, costs for years 1-4 will be included. For those entering in year 3 of the model, costs for year 1-3 will be included etc.

Figure 6-1 Budget impact analysis analytical framework

6.2.2.2 Data sources for budget impact framework

The population of interest included in the budget impact analysis was estimated for each country using an epidemiological approach, based on the incidence of disease (118), identified from published literature and country specific reports. Incident rather than prevalent cases were utilised because the treatment of interest is only prescribed once for newly diagnosed stage II or III CRC. Data on undiscounted chemotherapy costs were taken from the cost-utility analysis (part (i)) of this study). Non-drug costs of resource utilisation for both the standard treatment and the intervention from the trial were also utilised. Those hospitalisation costs incurred in the first year were assumed to be treatment related and those in the following years (1-4) were assumed to be condition related. Including non-drug costs was in keeping with the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) guidelines but it recognised that this is not advocated by some country-specific budget impact analysis guidelines (Australia)(117).

Rather than making the assumption that all patients would receive shorter treatment after the SCOT trial findings were disseminated, the results from the clinician survey (Chapter 5) were used to estimate the proportion of patients who would receive shorter treatment post-SCOT. The survey estimates accounted for differences in practice change for patients aged under 70 versus 70 and over. The details of how the survey results were applied are included in the footnotes of Table 13-6 in Appendix 4.

Figure 6-2 demonstrates how the eligible population was calculated and the extent of practice change applied. The time horizon chosen for the budget impact analysis was 5 years and no discounting was used, both in keeping with ISPOR guidance (118).

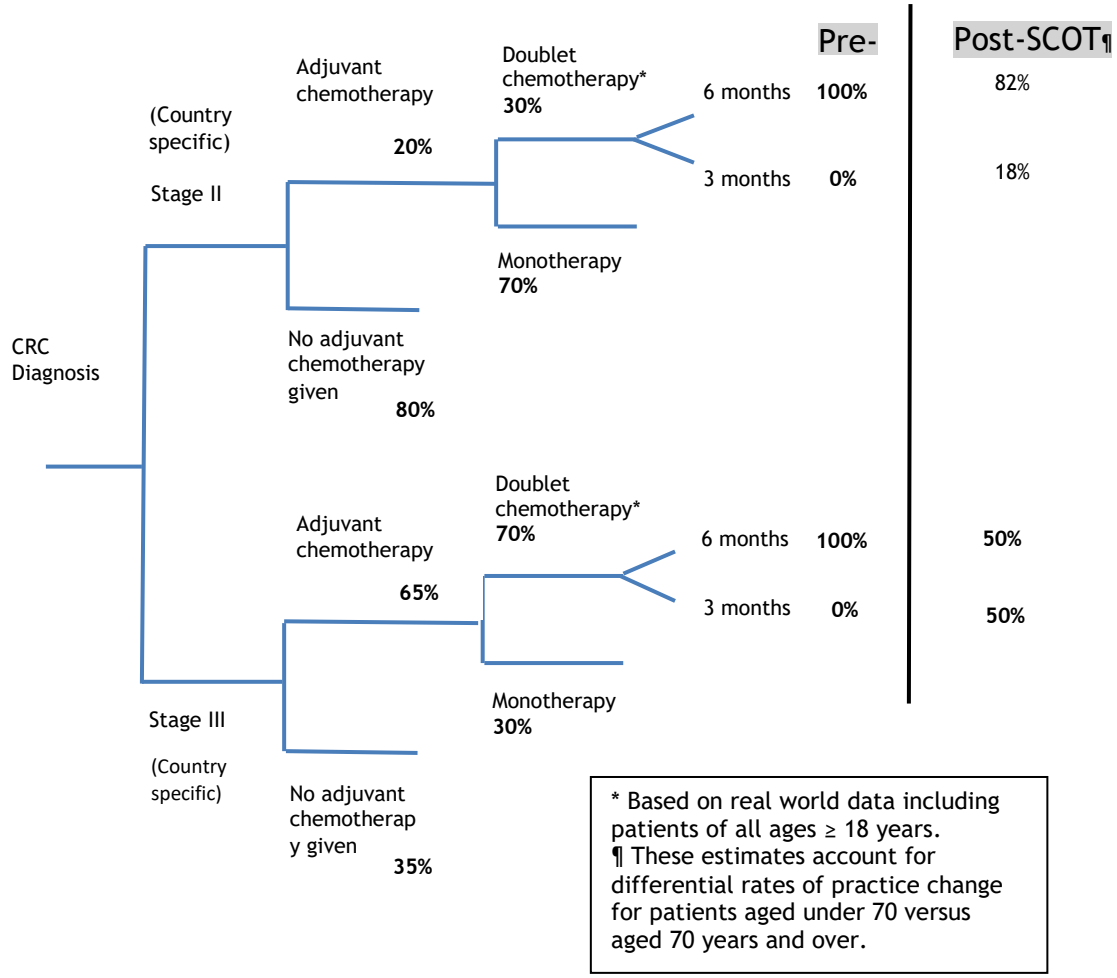


Figure 6-2 Eligible population and extent of practice change

6.2.2.3 Sensitivity analysis

Uncertainty around key parameters in the budget impact analysis was addressed in a one-way sensitivity analysis. A set of base case parameter values were chosen to act as the main comparator around which sensitivity analyses for each country were performed by changing the values of these main cost drivers. The parameters that were altered were: i) the proportion of patients diagnosed with CRC that received adjuvant chemotherapy, ii) the proportion of those patients who received the intervention of a shorter duration of chemotherapy, and iii) the time horizon of the analysis. As is recognised good practice, inputs for the sensitivity analyses were taken from published literature and reports where possible rather than altering values by random amounts (118).

Three scenario analyses were also performed. First, patients with a diagnosis of colon cancer only were considered by excluding patients with rectal cancer. Next, the budget impact of patients with stage III disease only was estimated. These scenario analyses were conducted due to the increased uncertainty regarding the use of doublet chemotherapy and practice change in response to SCOT for patients with rectal cancer and stage II disease (see Chapters 1 and 5).

A third scenario analysis was undertaken to calculate the budget impact from a societal, rather than a healthcare system perspective. Specifically, the loss of productivity from absenteeism at work during chemotherapy treatment, and the costs to patients to travel to hospital in the first year of treatment were included. There are three main approaches used in the literature to value productivity loss. These are shown in Table 13-4 Appendix 4 (based on (290, 291)). The human capital approach and the time patients were absent from employment during chemotherapy treatment were the focus of this evaluation. To calculate productivity loss, average country specific annual wages were used to calculate the earnings lost if 50% of patients aged 65 or under did not work during their time on treatment. The value of 50% was chosen based on the range of values provided by clinicians in response to a survey question about the proportion of patients that continue full time employment during adjuvant chemotherapy treatment (Appendix 3 Figure 12-12). The average time on treatment aligned with the time on treatment for each trial arm within the SCOT trial, calculated in the cost-utility analysis. The proportion of individuals who work part time in each country was also taken into account. Given the median age of patients receiving chemotherapy in the SCOT trial was 65 years, to calculate the proportion of patients aged 65 and under, the number receiving chemotherapy was divided in half. The productivity loss per patient using this approach was calculated using the following equation:

$$Productivity\ loss = 0.5x(nC) * ((T * Y) * (R * pt))$$

nC: Number of patients with CRC treated with adjuvant chemotherapy annually. T: Time on treatment. Y: Average yearly salary before tax. R: Country specific full time employment rate. Pt: country specific proportion of patients in part-time work.

The unit costs to calculate productivity loss are detailed in Table 6-1. One source was used for the average annual wage for each country (OECD Economic Outlook, Volume 2019, Issue 2) (292).

Table 6-1 Average yearly wage

Cost category	Unit cost (USD)	Source
Productivity loss estimates: average yearly wage (USD)		
Australia	54 501	OECD Economic Outlook Volume 2019, Issue 2
Denmark	57 150	
New Zealand	44 031	
Spain	38 758	
Sweden	46 695	
United Kingdom	47 226	

To calculate travel costs, an assumption that patients had a 30 mile round trip to attend hospital for chemotherapy or outpatient appointments was taken from the previous CRC literature (293). Travel for inpatient admissions was not included because although the number of nights spent in hospital for each patient was collected, it was unclear how many separate inpatient admissions this represented. It was also unclear what proportion of patients would have paid for their own travel to hospital for the purposes of these admissions rather been admitted to hospital from an outpatient clinic or being brought by ambulance. A mileage cost of 0.66 USD/mile was used to calculate travel costs (294). Travel costs were calculated using individual patient level data from the SCOT trial and were adjusted for the probability of survival (as per costs for main budget impact analysis). Patients of all ages were included for calculation of travel costs. The average travel cost incurred per patient in year 1 of treatment was calculated as follows:

$$Travel\ costs = (nC) * (Dc * 30(m))$$

nC: Number of patients with CRC treated with adjuvant chemotherapy annually. Dc: Average total number of daycase or outpatient attendances. m: Average cost of mileage per mile.

For the purposes of calculating both productivity loss and travel costs from the perspective of each country, purchasing power parity for private consumption

(295), rather than health specific purchasing power parity (used for chemotherapy and hospitalisation costs), was used to convert country specific currencies to USD.

6.2.3 Calculation of QALYs gained from implementation

Part iii) of the analysis was a calculation of the average QALY gain if the SCOT trial was implemented in these six countries over a 5-year time horizon. The average per patient QALY outcomes for the first 5 years for 3 versus 6 months of treatment were used from the within trial cost-utility analysis and applied to the eligible population of patients affected by SCOT trial implementation using the same BIA framework described above.

6.2.4 The value of implementation of the SCOT trial

An estimation of the cost of running the SCOT trial was obtained from a representative from Glasgow CRUK Clinical Trials Unit that ran the SCOT trial in collaboration with the Oncology Clinical Trials Unit in Oxford. The economic returns on the investment in the SCOT trial were calculated by subtracting the sum of the budget impact and monetised QALYs from the total cost of performing the SCOT trial.

6.3 Results

6.3.1 Cost-effectiveness of 3 months versus 6 months of treatment

6.3.1.1 Description of SCOT trial patients

Key patient, disease and treatment related factors for patients recruited from each country in the SCOT trial (n=6,065) are outlined in Table 6-2. Across all locations, average age was similar, there were more males than females, and most patients were fit, with a WHO performance status of zero. Notable differences included that Australian clinicians preferred to use FOLFOX (78%), whereas those from other locations used more CAPOX. There were no patients with stage II disease recruited from New Zealand, although the overall number from this country was small. There were more patients with low versus high-risk stage III disease in all countries, except New Zealand and Sweden. Disease compliance was higher for the 3-month versus the 6-month arm and higher for

fluoropyrimidine compared to oxaliplatin, regardless of treatment arm. Oxaliplatin compliance was lowest in Denmark and Sweden and in particular oxaliplatin compliance dipped to under 50% for patients in the 6-month arm from Denmark.

Table 6-2 SCOT trial patient characteristics dependent on location

Number of patients (percentage) Total N=6,065 (100%)		Australia N=196 (3.2%)	Denmark N=310 (5.1%)	New Zealand N=16 (0.3%)	Spain N=233 (3.8%)	Sweden N=82 (1.4%)	United Kingdom N=5,228 (86.3%)
Patient characteristic							
Mean age (years, min-max) Median (IQR)		64 (39-83) 65 (58-71)	64 (23-81) 65 (59-70)	65 (48-78) 66 (61-71)	63 (38-80) 64 (57-69)	63 (32-78) 65 (59-69)	63 (20-85) 65 (58-70)
Gender (Number, percentage)	Male Female	112 (57%) 84 (43%)	178 (57%) 132 (43%)	11 (69%) 5 (31%)	143 (61%) 90 (39%)	51 (62%) 31 (38%)	3,177 (61%) 2,051 (39%)
Performance status (Number, percentage)	0 1	160 (82%) 36 (18%)	248 (80%) 62 (20%)	12 (75%) 4 (25%)	180 (77%) 53 (23%)	75 (91%) 7 (9%)	3,642 (70%) 1,586 (30%)
Disease characteristic							
Extended risk stage	II Low III High III	4 (2%) 117 (60%) 75 (38%)	85 (27%) 129 (42%) 96 (31%)	0 (0%) 8 (50%) 8 (50%)	62 (27%) 99 (42%) 72 (31%)	7 (9%) 34 (41%) 41 (50%)	956 (18%) 2281 (44%) 1991 (38%)
Treatment							
Drug regimen (number, percentage)	FOLFOX CAPOX	152 (78%) 44 (22%)	45 (15%) 265 (85%)	4 (25%) 12(75%)	110 (47%) 123 (53%)	40 (49%) 42 (51%)	1620 (31%) 3608 (69%)
Treatment compliance (Median; IQR)							
Fluoropyrimidine	3 month 6 month	97% (89-99%) 90% (58-97%)	94% (86-98%) 79% (49-92%)	91% (87-96%) 79% (65-92%)	97% (91-99%) 92% (80-98%)	93% (88-98%) 85% (59-96%)	94% (81-99%) 82% (56-94%)
Oxaliplatin	3 month 6 month	97% (89-99%) 70% (53-87%)	87% (63-97%) 43% (25-66%)	99% (90-100%) 78% (74-79%)	98% (93-99%) 85% (60-97%)	95% (87-98%) 62% (37-72%)	95% (81-99%) 70% (45-86%)

6.3.1.2 Outcomes

Survival

Figure 6-3 shows separate Kaplan-Meier curves for each partitioned survival outcome across the whole study period (maximum survival in either arm was 10.6 years), using data from all evaluable patients (n=6,065).

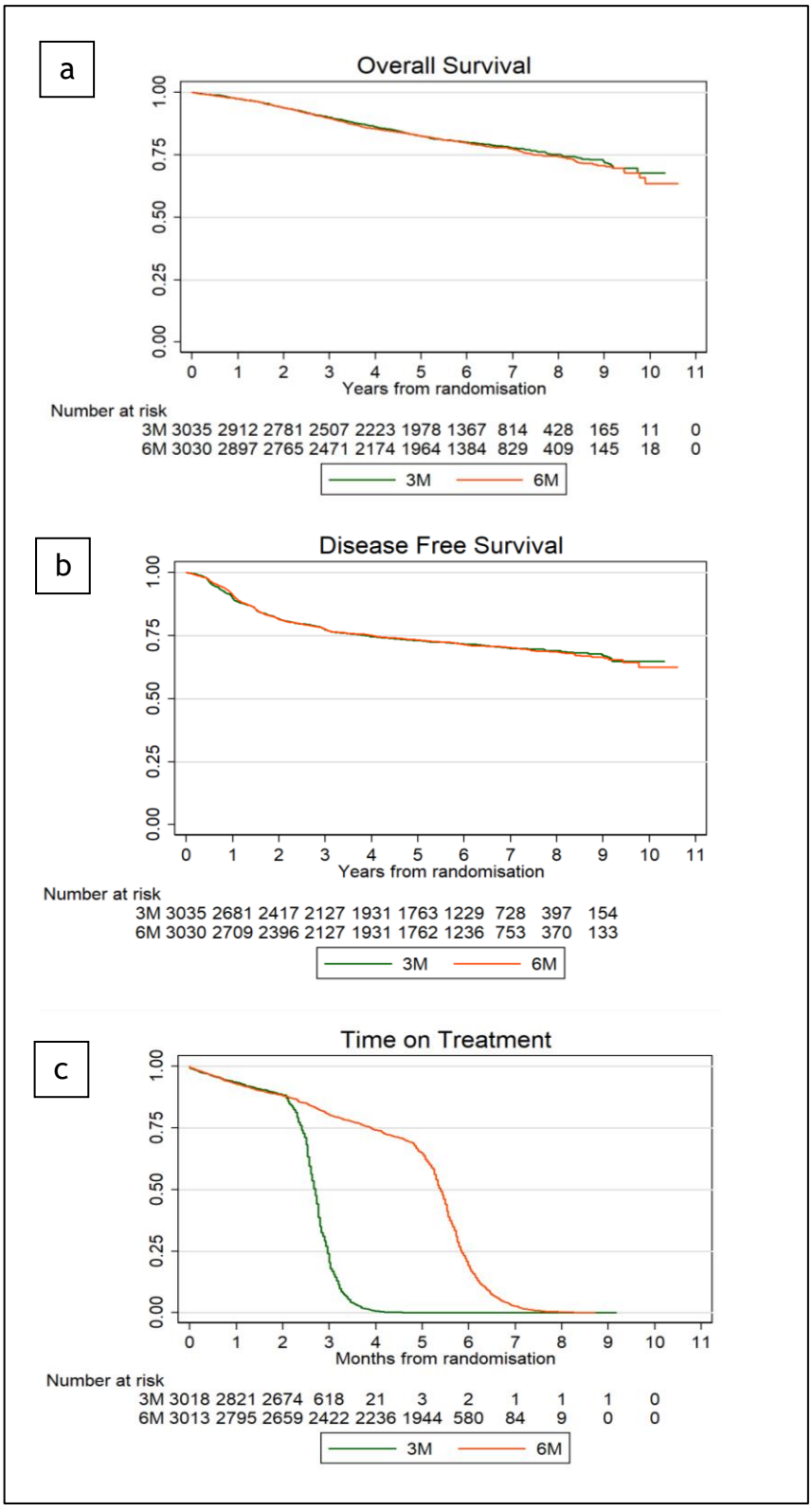


Figure 6-3 Kaplan-Meier curves a: Overall survival b: Disease free survival c: Time on treatment.

Figure 6-4 shows the Kaplan-Meier curves (max duration 10 years) showing the differences in the area under the curves for recurrence, DFS, and time on treatment for the 3-month and 6-month trial arms.

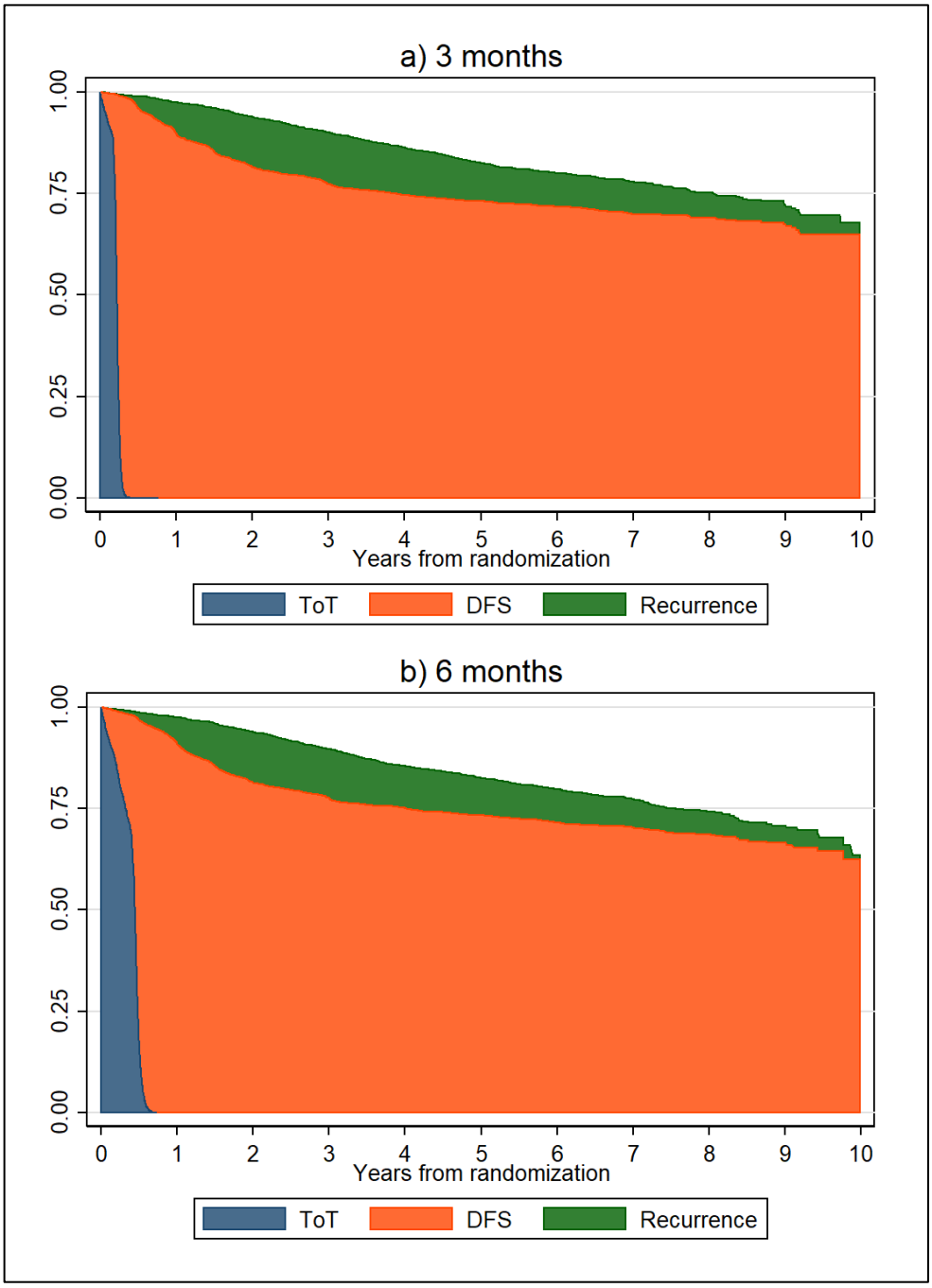


Figure 6-4 Survival curves partitioned by health state a: 3 month arm b: 6 month arm.

The average time spent in these three health states for patients in each trial arm is shown in Table 6-3. DFS was longer but recurrence was also higher for the 3-month arm (neither statistically significant). Longer average time in the DFS state was partly influenced by the shorter time on treatment. Time on treatment was significantly longer in the 6-month versus 3-month treatment arm, however, it was not double, reflecting the poorer treatment compliance with longer treatment (see Table 6-2). Mean life expectancy was higher for patients in the 3 versus 6-month arm (not statistically significant).

Table 6-3 Survival times for each trial arm

Health state	3 month arm N=3035 (50.04%)	6 month arm N=3030 (49.96%)	Incremental (CI)	p-value
ToT	0.21 (0.21, 0.22)	0.39 (0.39, 0.40)	-0.18 (-0.19, -0.17)	<0.001
DFS	7.35 (7.21, 7.49)	7.17 (7.02, 7.31)	0.19 (-0.01, 0.38)	0.063
Recurrence	0.82 (0.74, 0.90)	0.78 (0.69, 0.85)	0.05 (-0.06, 0.17)	0.351
Total (OS)	8.39 (8.27, 8.50)	8.33 (8.21, 8.44)	0.06 (-0.10, 0.22)	0.443
Kaplan-Meier estimate used for computation of expected time in each health state. Survival time estimated up to 10 years post randomisation for n=4,592 and up to 8 years for 1,473 patients. Missing values for ToT means that the number of patients used for this calculation was lower (n=3018 for 3 month arm and n=3013 for 6 month arm). Undiscounted times used.				

Quality of life

The linear regression output for UK utilities is provided below and in included in Appendix 4 for other countries. As expected, utilities were reduced when patients were receiving chemotherapy treatment, patients who had longer treatment (6-month arm) and when they experienced disease recurrence compared to a disease free state. Patients who received CAPOX had a minimally higher average utility compared to patients prescribed FOLFOX.

Table 6-4 Regression output for UK utilities

UK utilities		
Variable	Co-efficient	S.E.
Health states (ref: disease free)		
On treatment	-0.0396*	0.004
Recurrence	-0.0694*	0.015
Arm: 6 months (ref: 3 months)	-0.0154*	0.007
Characteristics		
CAPOX	0.0042	0.008
Low-risk stage III	0.0002	0.012
High-risk stage III	-0.0062	0.011
Male	0.0163*	0.007
Age	0.0016*	0.001

Ethnic (ref: Caucasian)		
African/Caribbean	-0.0817*	0.039
South Asian	-0.1453*	0.536
Chinese	-0.0477	0.076
Other	0.0186*	0.022
Constant		
Comparison is a 65 year old, Caucasian female patient on 3-month trial arm in a DFS health state and stage II disease treated with FOLFOX. *p<0.05		

Figure 6-5 demonstrates the average utility values at baseline and up to eight years for patients in each arm of the SCOT trial for each country. Once country specific utility weights were applied, health-related quality of life estimates were lowest for New Zealand and highest for Sweden. Given that fully-pooled outcome data was used for this calculation, any differences solely reflect differences in the estimates of how individuals from those specific countries value quality of life, rather than any observed difference between patients from different countries within the SCOT trial.

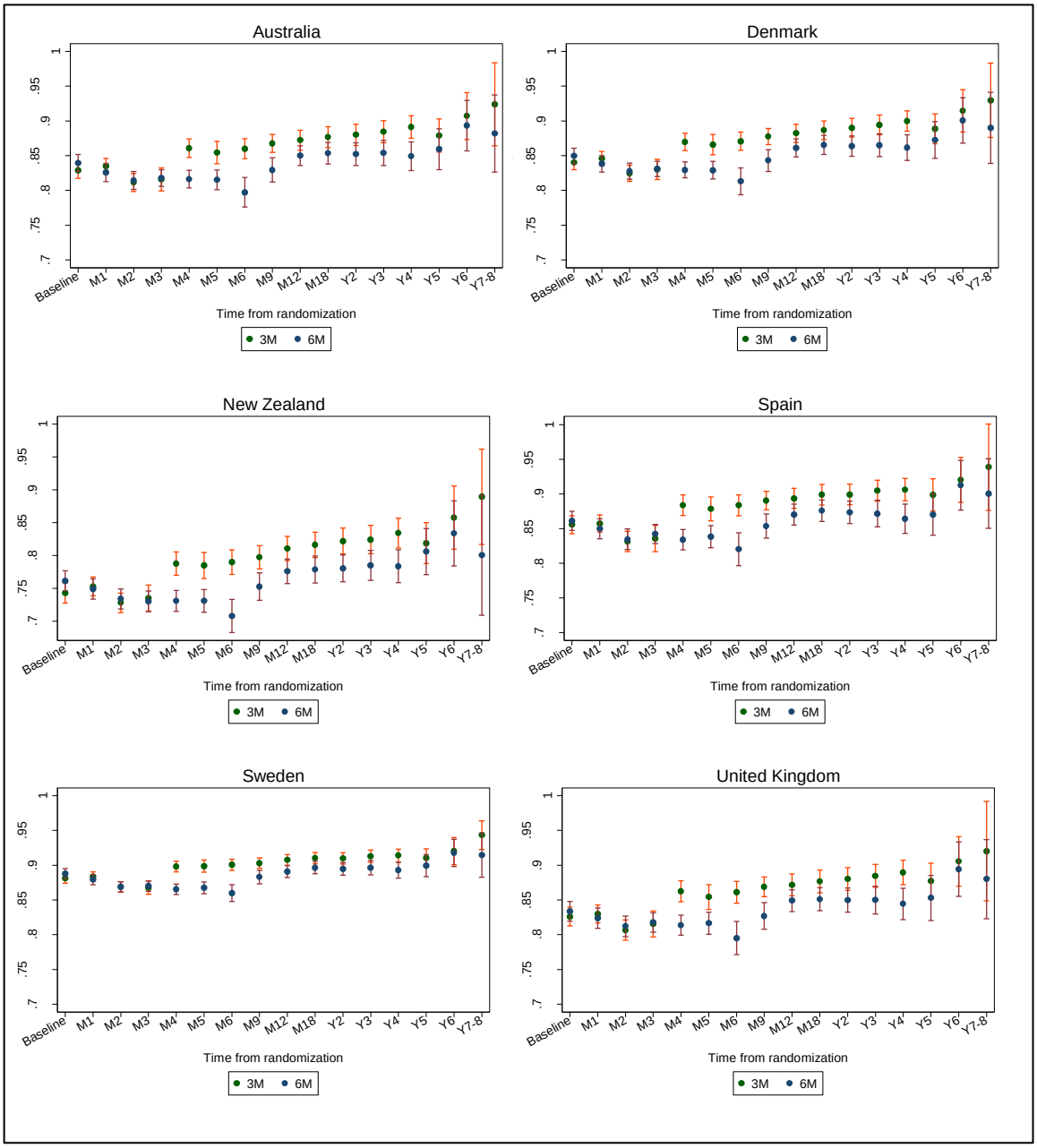


Figure 6-5 Fully pooled results for quality of life calculated by applying country specific utilities to EQ-5D responses Relevant 95% confidence intervals are represented by error bars either side of each point estimate. "M": month, "Y", year.

The final outcome of interest was QALY gain. The QALY gain from using 3 versus 6 months of treatment is shown in Table 6-6.

6.3.1.3 Costs

The output from the linear regression model used to extrapolate costs for years 7, 8 and 9 from a UK perspective are shown below (Table 6-5) and in Appendix 4 for other countries.

Table 6-5 Regression output for yearly costs from UK perspective

UK costs (Regression used to model costs for years 7-10)		
Variable	Co-efficient	S.E.
Health states (ref: disease free)		
On treatment	13,974*	314
Recurrence	10,706*	800
Arm: 6 months (ref: 3 months)	282	153
Arm (6 months)*Year 1	672	382
Year 1	6114*	349
Characteristics		
CAPOX	-4915*	291
Low-risk III	49	328
High-risk III	217	304
Male	-394	269
Age	23	13
Ethnic (ref: Caucasian)		
African/Caribbean	-1550	844
South Asian	127	1048
Chinese	453	1679
Other	-231	428
Constant	4508*	939
Comparison is a Caucasian female patient on 3-month trial arm aged 65 years with stage II disease, treated with FOLFOX after year 1 in a disease free state.		

The average undiscounted chemotherapy and hospitalisation cost per patient for each country of using 3 versus 6 months of treatment and the difference between treatment durations are shown in Figure 6-6 and mean discounted costs for each country are outlined in Table 6-6. These costs represent the combined costs over a period of up to 10 years. Spain and the UK had the lowest chemotherapy unit costs, whereas chemotherapy unit costs were very similar for Australia, New Zealand, Sweden, and Denmark. The differences in hospitalisation costs aligned with the ratios between UK and other country costs used in the calculation (see Chapter 2). Costs calculated using a multi-country approach aligned closely with those from the UK because most patients in the SCOT trial were from this location.

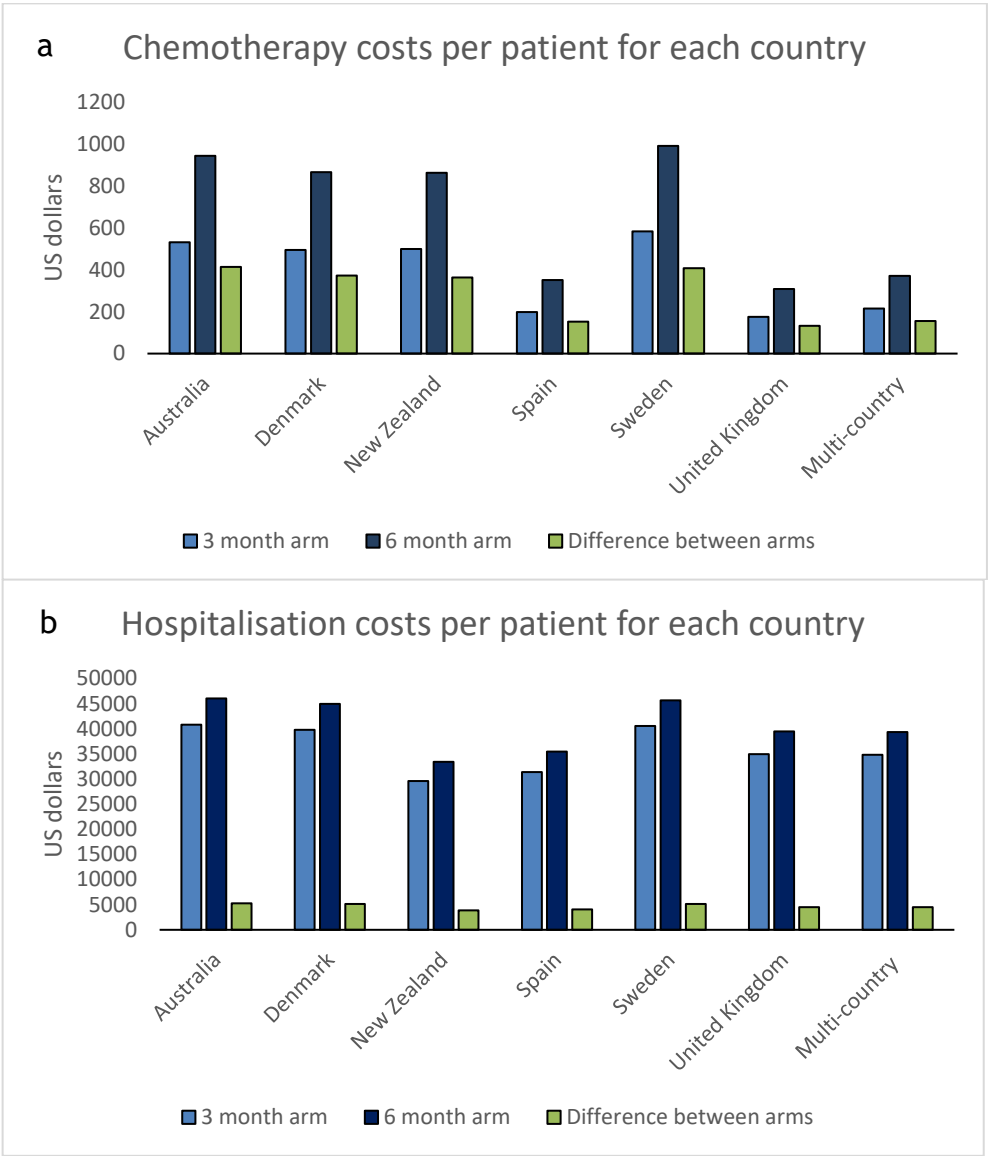


Figure 6-6 Mean undiscounted costs per patient Costs given per arm and difference between trial arms for each country a) chemotherapy medication costs b) hospitalisation costs.

6.3.1.4 Cost utility, sensitivity and sub-group analyses

The cost-effectiveness of 3 versus 6 months of treatment was estimated by comparing the costs and outcomes for each trial arm. Overall, 3 months of treatment was a dominant strategy, being more effective and cost-saving compared to 6 months across all locations (Table 6-6). The differences between EQ-5D responses (Figure 6-3) are also reflected in the difference in QALY gains between countries, which ranged from 0.11 for Sweden to 0.17 for New Zealand (only statistically significant for New Zealand).

The INMB was greater than $8,000 across all locations, with over 99% probability that the 3-month arm was cost-effective for all countries at a WTP threshold of one GDP/capita.

Table 6-6 Country specific cost-utility analysis results

Arm	Mean (discounted) Costs (95% CI)	QALYs (95% CI)	NMB using 1 x GDP per capita (95% CI)	Probability of being CE at WTP of 1 x GDP per capita (%)
Australia			GDP: 53, 000 USD	
3 months	37,289 (35,520; 39,226)	6.28 (6.17; 6.40)	295, 494 (288,886; 302,170)	99.6%
6 months	42,830 (40,691; 44,999)	6.13 (6.00; 6.26)	282, 158 (275, 037; 239,783)	0.4%
Incremental	-5,541 (-8,383; -2,624)	0.14 (-0.01; 0.30)	13,337 (4,265; 22,533)	3M dominates
Denmark			GDP: 62,000 USD	
3 months	36,357 (34,639; 38,242)	6.35 (6.25;6.47)	357,653 (350,583; 365,053)	99.6%
6 months	41,744 (39,660; 43,856)	6.22 (6.10; 6.34)	343, 768 (335,926;352,194)	0.4%
Incremental	-5,386 (-8156; -2544)	0.13 (-0.01; 0.28)	13,884 (4,011;24,119)	3M dominates
New Zealand			GDP: 42,000	
3 months	27,133 (25,889; 28,514)	5.80 (5.67; 5.93)	216,261 (210,639;222,158)	99.6%
6 months	31,264 (29,736; 32,793)	5.63 (5.49; 5.77)	204,983 (198,714; 211,365)	0.4%
Incremental	-4,131 (-6148; - 2013)	0.17 (0.002; 0.34)	11,278 (3,886; 19,311)	3M dominates
Spain			GDP: 31,000	
3 months	28,443 (27,119; 29,909)	6.44 (6.33; 6.56)	171,245 (167,196; 175,464)	99.8%
6 months	32,583 (30,961; 34,217)	6.29 (6.16; 6.41)	162,273 (158,018; 166,987)	0.2%
Incremental	-4,140 (-6286; -1893)	0.15 (-0.00; 0.31)	8,972 (3,409; 14,602)	3M dominates
Sweden			GDP: 52,000	
3 months	37,104 (35,244; 39,079)	6.56 (6.46; 6.65)	303,778 (298,231; 309, 245)	99.5%
6 months	42,515 (40,353; 44,770)	6.44 (6.34; 6.54)	292,493 (286,508; 298,774)	0.5%
Incremental	-5411 (-8383; -2418)	0.11 (-0.02; 0.24)	11,285 (3557; 18,942)	3M dominates
UK			GDP: 40,000 USD	
3 months	31,629 (30,144; 33, 269)	6.27 (6.16; 6.39)	219,347 (214,184; 224, 628)	99.6%
6 months	36,182 (34,368; 38,023)	6.12 (5.99; 6.25)	208,673 (203,051; 214,704)	0.4%

Incremental	-4553 (-6955; -2056)	0.15 (-0.01; 0.31)	10,674 (3,683; 17,928)	3M dominates
Multi-country			NICE threshold £30,000=$42,000 USD	
3 months	31,594 (30,092; 33,227)	6.27 (6.16; 6.39)	231, 932 (226,577;237,410)	99.6%
6 months	36,150 (34,333; 38,008)	6.12 (6.00; 6.25)	220,949 (215,104;227,210)	0.4%
Incremental	-4,557 (-6932; -2097)	0.15 (-0.01; 0.31)	10,983 (3,684;18,664)	3M dominates

Figure 6-7 demonstrates the INMB across a range of WTP thresholds. Sweden had one of the highest cost-savings, therefore at lower WTP provided the highest INMB. As the WTP for a QALY gain increases, the INMB was highest from a New Zealand perspective because of the highest QALY gain when New Zealand specific EQ-5D weights were applied in the calculation of utilities.

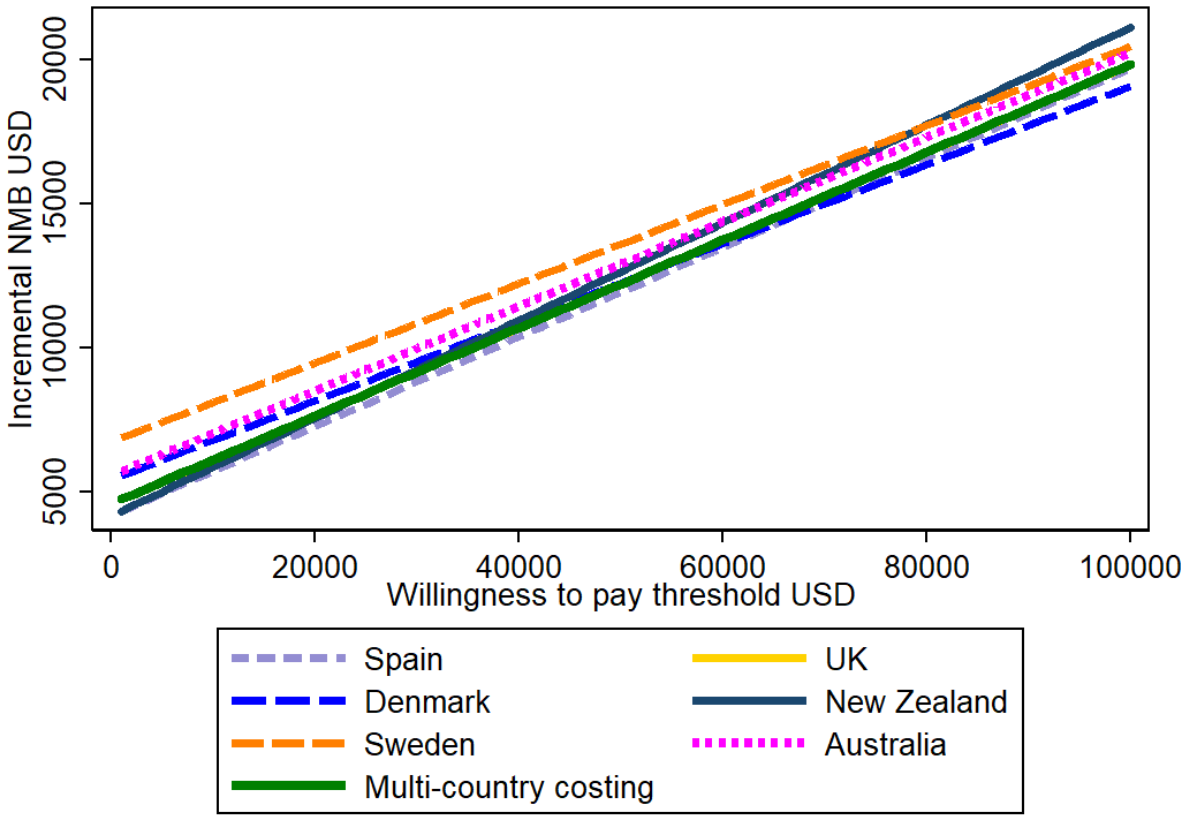


Figure 6-7 INMB over a range of WTP thresholds for each country INMB: Incremental net monetary benefit. WTP: willingness to pay. NB. UK and multi-country results are overlapping.

The CEACs for each location demonstrated that 3 months of treatment is cost-effective compared to 6 months over a large range of thresholds up to $200,000/QALY. Focusing on the cost-effectiveness planes, for all countries the point estimates for cost-effectiveness of 3 versus 6 months of adjuvant treatment for CRC were all within the South East quadrant, that is, the shorter treatment was the dominant strategy. Also, the majority of the bootstrapped estimations were also in the lower right hand quadrant of the cost-effectiveness plane, indicating that three months of treatment produced both a cost saving and a quality of life gain. A minority of data points sat in the lower left quadrant, indicating that shorter treatment is still cost saving but leads to reduced QALYs. The country specific WTP thresholds of one GDP per capita are displayed on each cost-effectiveness plane as a dashed line. Point estimates and bootstrapped results lying below the line show that 3 months of treatment is cost-effective compared to 6 months using the specific threshold indicated. It is clear that most bootstrapped estimations were under these thresholds for all six countries.

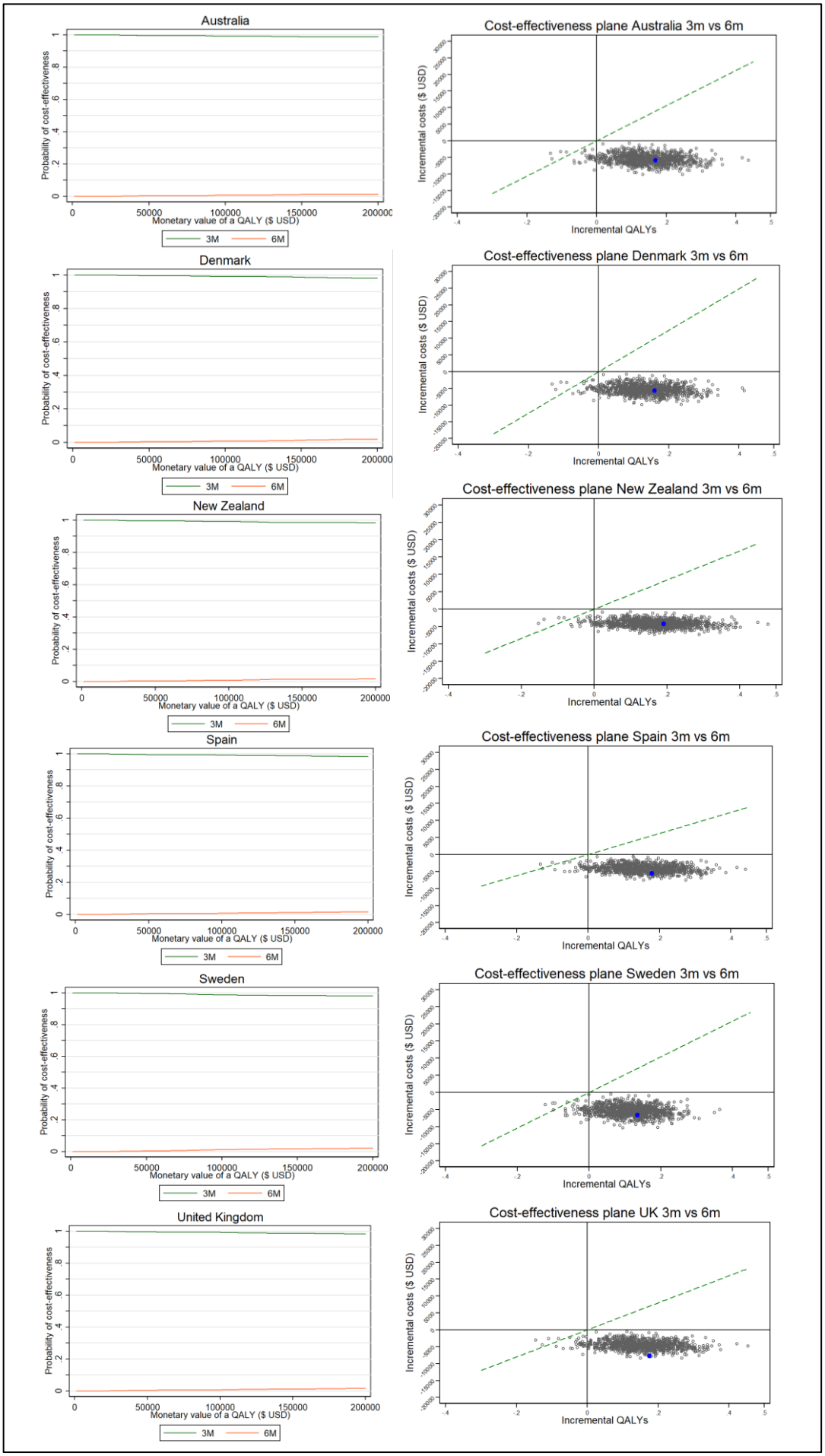

Australia
Probability of cost-effectiveness
Monetary value of a QALY ($ USD)
3M
6M
Cost-effectiveness plane Australia 3m vs 6m
Incremental costs ($ USD)
Incremental QALYs
Denmark
Cost-effectiveness plane Denmark 3m vs 6m
New Zealand
Cost-effectiveness plane New Zealand 3m vs 6m
Spain
Cost-effectiveness plane Spain 3m vs 6m
Sweden
Cost-effectiveness plane Sweden 3m vs 6m
United Kingdom
Cost-effectiveness plane UK 3m vs 6m

Figure 6-8 Country specific CEACs and cost-effectiveness planes Displaying 1000 bootstrapped iterations of the incremental cost/QALY for 3 months versus 6 months of treatment from the perspective of all six healthcare. The blue circle on the cost-effectiveness planes indicates the relevant incremental cost/QALY point estimate and the green dashed line indicates a WTP threshold (1 GDP/capita) for each country.

The results of subgroup analyses are provided in Figure 13-2 Appendix 4. There was most uncertainty for patients receiving FOLFOX. At a WTP threshold of $42,000 USD, the probability of 3 months of FOLFOX being cost-effective was 77%, compared to 99% for CAPOX. Comparing patients aged under 70 versus 70 and over, the cost-effectiveness planes show us that, for the older group, a number of bootstrapped estimations lie in the upper right quadrant, indicating that 3 months of treatment is more costly than 6 months of treatment, although at the same time providing a QALY gain. Interestingly, for male patients, the CEAC shows cost-effectiveness of 3 months compared to 6 months of treatment dips at higher WTP thresholds and the corresponding cost-effectiveness plane shows that this is driven by a reduction in QALY gain, rather than an increase in costs for the 3-month arm. Focusing on extended stage of disease, there was the greatest spread of bootstrapped results for patients with stage II disease which was also the smallest of these subgroups (n=1,114). At a WTP threshold of $42,000, the probability of 3 months of treatment being cost-effective compared to 6 months was 90.3%, 96.3%, and 87.5% for patients with stage II, low-risk stage III and high-risk stage III respectively.

6.3.2 Analysis of budget impact

Table 6-7 shows the number of patients for each country estimated to receive 3 months rather than 6 months of adjuvant doublet chemotherapy after implementation of the SCOT trial results. This estimate was highest for the UK and lowest for New Zealand. Chemotherapy medication costs, treatment related and condition related hospitalisation costs for each country over 5 years show that most cost savings from SCOT trial implementation were mainly driven by a reduction in the treatment related hospitalisations (year 1) for the shorter treatment approach of 3 months versus 6 months.

Combining all relevant healthcare costs, the estimated healthcare system savings of implementation of SCOT trial findings over 5 years ranged from $3.6 million USD (New Zealand) to over $61.4 million USD (UK). The combined base

case budget impact was $152 million USD. The values are provided in country specific currencies in Appendix 4 Table 13-17. The impact for New Zealand was lowest due to the smallest eligible population and cost difference per patient. Although the cost-utility analysis showed that the cost savings per patient from using shorter treatment were highest from the perspective of Australia, the UK had the largest budget impact due to the biggest eligible population.

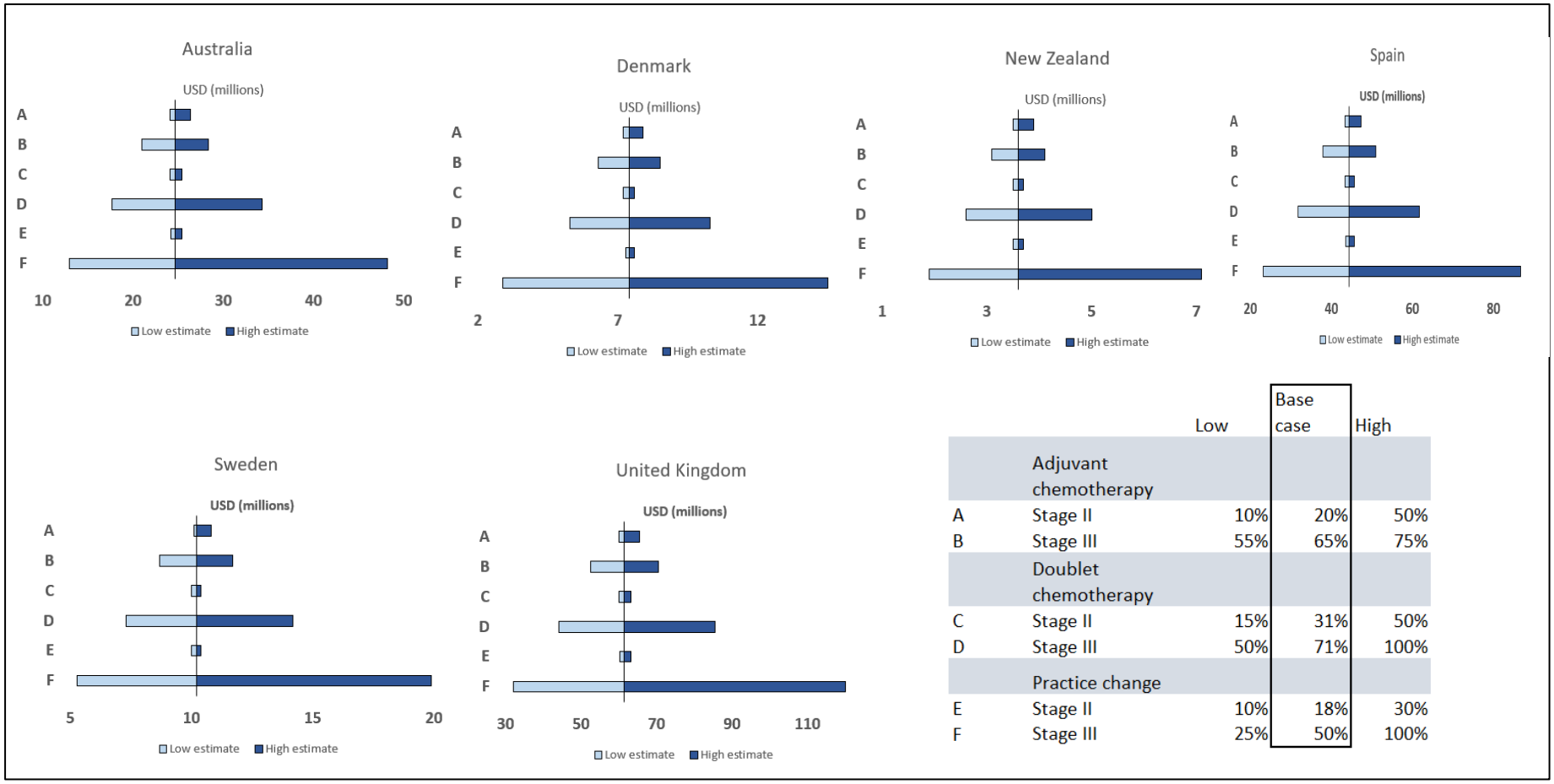
Table 6-7 Budget impact analysis base case and scenario analysis in country specific currency All values for budget impact included in this table indicate cost savings USD (millions).

	Australia	Denmark	New Zealand	Spain	Sweden	United Kingdom
Annual number of patients receiving 3 months of treatment who would have received 6 months pre-SCOT	957	495	187	2,304	406	2,906
Chemotherapy medication costs over 5 years	2.0	0.5	0.3	1.8	0.8	1.9
Treatment related hospitalisations in (year 1 for each individual patient) over 5 years	21.4	6.5	3.1	40.2	8.8	56.1
Condition related hospitalisations (years 2-5 for each individual patient) over 5 years	1.3	0.4	0.2	2.4	0.6	3.4
Total budget impact = Medication cost + Cost of treatment and condition related hospitalisations	24.7	7.4	3.6	44.4	10.2	61.4

6.3.2.1 Budget impact sensitivity analysis

A one-way sensitivity analysis was performed in which key parameters used in the budget impact calculation were increased or decreased in turn to investigate the effect on the extent of budget impact. These parameters and the values used in the sensitivity analysis are outlined in Figure 6-9, alongside a Tornado

diagram for each country showing the corresponding variation in 5-year cost savings. For example, if it was assumed that all patients with stage III CRC receive 3 rather than 6 months of adjuvant doublet chemotherapy, the overall potential budget savings amounted to $297 million USD.



		Low	Base case	High
	Adjuvant chemotherapy			
A	Stage II	10%	20%	50%
B	Stage III	55%	65%	75%
	Doublet chemotherapy			
C	Stage II	15%	31%	50%
D	Stage III	50%	71%	100%
	Practice change			
E	Stage II	10%	18%	30%
F	Stage III	25%	50%	100%

Figure 6-9 Budget impact analysis one-way sensitivity analysis Parameters used to perform the sensitivity analysis are shown in the table embedded within this figure. There are six tornado diagrams, one for each country displaying 6 one-way sensitivity analyses. The letter on the right hand side of each diagram indicates the parameters changed, according to the values displayed in the table. For example, for "A", the proportion of patients with stage II CRC receiving adjuvant chemotherapy has been altered from 10% (low estimate) to 50% (high estimate) around the baseline case of 20%. The line in the middle of each tornado diagram indicates the baseline budget impact estimated for that country (see Table 6-7 for those results listed under "Total budget impact".

Removing patients with stage II disease from the analysis (Table 6-8) had little effect (total savings $145 million USD) because of the small population deemed eligible to be affected by a practice change in the base scenario for this study. Excluding patients with rectal cancer in a scenario analysis led to a decrease in the budget impact by less than half (total $102 million USD) because incidence of rectal cancer is generally less than colon cancer for stage II/III disease.

Productivity loss had a significant monetary impact over five years, in excess of the healthcare budget impact for the same time-period. The combined travel cost impact over 5 years across all countries amounted to $4.5 million USD. The total estimated budget impact across all countries over 5 years from a societal perspective (adding base case budget impact, productivity loss and travel costs) amounted to $340 million across all six countries.

Table 6-8 Budget impact scenario analysis results

	Australia	Denmark	New Zealand	Spain	Sweden	United Kingdom
Base case total budget impact = Medication cost + Cost of treatment and condition related hospitalisations	24.7	7.4	3.6	44.4	10.2	61.4
Scenario analysis (otherwise as per base scenario)						
Rectal Cancer Excluded	17.1	5.0	2.4	29.2	6.9	41.3
Stage II excluded	23.5	7.1	3.4	42.4	9.8	58.7
Productivity loss (50% aged 65 years and under stop work)	29.2	10.1	5.1	48.8	11.8	78.8
Travel costs	0.6	0.2	0.1	1.5	0.3	1.8

6.3.3 QALY gain from SCOT trial implementation

In order to undertake an assessment of economic impact in the SCOT trial following the approach used by Glover et al (39), it was necessary to calculate the country specific QALY gain from implementation. This was calculated using the same incident population over 5 years as in the budget impact model. Average QALY outcomes from the 3 and 6-month arms of the trial were applied

to the estimated proportion of patients receiving an intended 6 versus 3 months of treatment pre and post-SCOT. Table 6-9 shows the average QALY gain and the monetised values of these gains.

Table 6-9 QALY gain over 5 years from a country specific perspective and the monetised value of that QALY gain *WTP threshold: $42,000 USD

	Australia	Denmark	New Zealand	Spain	Sweden	United Kingdom
Average QALY gain from implementation	384	104	84	921	102	1163
Monetised QALY gain* (USD millions)	16.1	4.4	3.5	38.7	4.3	48.8

Figure 6-10 demonstrates the change in monetised QALY gains for each country at different WTP thresholds using the main point estimate for QALY outcome from the two arms of the SCOT trial calculated in the cost-utility analysis (part i) of this study).

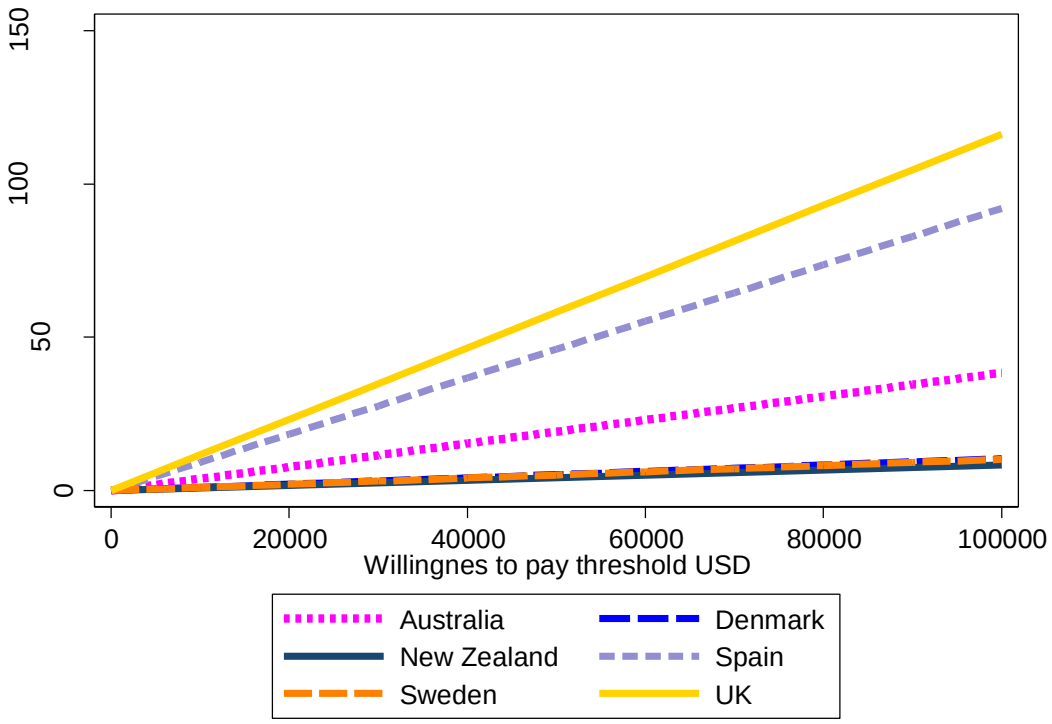


Figure 6-10 Monetised QALY gain from implementation of SCOT findings

6.3.4 Estimating the return on the investment in the SCOT trial

The overall value of conducting and implementing the SCOT trial was calculated by comparing the amount spent on developing and running the trial across six countries (see Appendix 4), to the cost savings and monetised health benefits from implementation. Table 6-10 shows the components of this calculation. Overall, subtracting the costs to run the trial, the economic value of conducting and implementing the SCOT trial was $259 million USD. Adopting a societal perspective, by adding productivity loss and travel calculations (188.3), brought this figure to $447 million USD.

Table 6-10 Calculation of economic value of investment in the SCOT trial Return on investment (USD): (-8.8) – (151.7 + 115.8) = 258.7 million USD gain.

	Australia	Denmark	New Zealand	Spain	Sweden	United Kingdom	Total
Input							
Project specific and core CTU funding to develop and run SCOT	8.8 million USD (deficit)						-8.8
Output							
Base case total budget impact from implementation (USD millions)	24.7	7.4	3.6	44.4	10.2	61.4	+151.7
Monetised QALY gain* (USD millions)	16.1	4.4	3.5	38.7	4.3	48.8	+115.8
Total							+258.7 USD

6.4 Discussion

Three months of adjuvant, oxaliplatin-based chemotherapy is cost-effective and cost saving compared to 6 months from the perspective of all countries that recruited to the SCOT trial. Using updated survival data, mean life expectancy was better in the 3 month arm (not statistically significant), whereas in the previous cost-utility analysis, life expectancy was non-significantly higher in the 6 month arm. In addition, although there was a QALY gain from shorter treatment in the previous cost-utility analysis, it was higher in this study (only

statistically significant for New Zealand), driven both by increased life expectancy and quality of life improvement. Any cost-effectiveness differences between countries was driven by the unit costs for hospitalisations and chemotherapy medication, as well as the country specific weights attached to quality of life measures.

Regarding unit costs, publicly available list prices were used for every country. Although it is unlikely any within-country variation in these list prices would change the cost-effectiveness outcomes, it is important for decision makers from each country to read these results in the context of country specific cost used. For example, the UK costs calculated from this analysis are lower than estimated in a previous UK perspective cost-utility analysis (75). This is because a difference source of unit costs was used in this analysis (eMIT) compared to the previous analysis (BNF). eMIT costs were felt to be more appropriate for this analysis for two reasons. Firstly, eMIT is the source of unit costs for generic drugs recommended by NICE guidance (197). Secondly, when BNF unit costs were investigated for use in this analysis, these costs were in some instances ten times higher compared to both eMIT costs and medication unit costs for the other countries in this analysis. This supports the concept that even within country sources can vary and interpretation of any results should be mindful of this variation. If BNF costs had been used in this analysis, the cost savings from SCOT trial implementation may have been several times the current estimate.

In this study, the cost-effectiveness of shorter treatment was most certain for patients with low-risk stage III disease. The previous cost-utility analysis from a UK perspective (75) split patients into high-risk stage III and “other” and therefore it was not possible to explore the difference for patients between stage II and low-risk stage III. The cost-effectiveness for patients with low-risk stage III disease was driven by non-significant improvements in life expectancy, QALY gain, and cost saving associated with shorter treatment.

Subgroup analysis also revealed that the cost-effectiveness of using 3 versus 6 months of treatment was less certain for patients receiving FOLFOX, especially at higher WTP thresholds. This was because of the small mean benefit in life expectancy from using longer treatment with this regimen (not statistically significant). However, FOLFOX for 3 months was still seen to be a cost-effective

treatment compared to 6 months because of the costs savings and improved quality of life associated with shorter treatment. This contrasts with a recent model based cost-utility analysis specifically relevant to patients with high-risk stage II disease (296), which used information from the SCOT trial (utilities), the TOSCA trial (adverse events) and the pooled high-risk stage II IDEA analysis (survival outcomes). The authors of that model-based study concluded that, although the cost-effectiveness of 3 months of CAPOX was demonstrated, 3 months of FOLFOX was not cost-effective compared to 6 months of FOLFOX. This result was driven by the fact that, as in this study, FOLFOX was a more expensive treatment compared to CAPOX and 3 months of FOLFOX was less effective than 6 months.

Approximately 1.8 million people (297) are diagnosed with CRC per annum globally, with the annual cost of managing this disease projected at over $39 billion USD (298). Around half of these patients (299) present with stage II-III disease, therefore making savings relevant to this patient cohort has the significant cost consequences. This study shows that implementing SCOT trial findings in six high-income countries translates to savings of hundreds of millions of dollars, in the order of a 17-fold increase on the original SCOT investment, and the total global impact is likely to be several times this estimation. Adding in the QALY gain related to this implementation, for the base-case analysis from a healthcare perspective meant the return on the initial investment was 30-fold; if societal perspective gains were included, this increased to over 50-fold. This is in line with previous estimates of the return on medical research in the USA (twenty fold return) (300) and compares favourably to a previous analysis of non-cancer clinical trial impact performed by US authors who identified a four-fold return on investment (301). That US study analysed the impact of a programme of trials but was limited by the fact that if cost-utility and implementation information was not available for certain trials, only the costs of conducting the trial, and none of the benefits, were included in the estimation. Also, the implementation of new treatment strategies tested within many of the individual trials assessed in that study occurred at a cost to the health-service rather than being cost-saving.

The main part of this budget impact analysis considered direct healthcare service costs only. A previous estimate of the burden of all cancers on the European Union economy was calculated at $126.3 billion for the year 2009 (302), with 40% related to direct healthcare costs and 40% related to productivity costs. The analysis from the report Medical Research: What's it Worth (149) (Table 3-1 Chapter 3), suggested that valuing either health gain or work costs will often outweigh direct health care savings. The findings from this study supports the results from both of those studies, and taking a societal perspective significantly increased the cost savings related to SCOT trial finding implementation. The incidence of CRC is increasing in younger patients (303), therefore it could be argued that this provides even greater rationale to include the impact of treatment, morbidity and mortality from CRC on patient employment when evaluating the downstream effects of research findings.

The strengths of this study are the use of within trial data to calculate costs, utilities, and survival gains associated with the treatments being assessed. The analysis was also strengthened by using the implementation estimate from a survey of clinicians. This method applied evidenced practice change rather than making assumptions on potential practice change, which have inherent limitations. Using clinician self-reported change was a more realistic approach to evidencing impact than previous cancer trial impact studies have utilised. Using these assessments of implementation add weight to this type of analysis, because without these results it would be impossible to know the extent of practice change or if change in practice would have occurred anyway, without the clinical trial and its associated cost.

This study has some important limitations. Firstly, using a fully pooled approach, varying unit costs alone are likely to underestimate between-country differences compared to fully splitting the analysis (112). A partially split approach was avoided because two of the countries that recruited to SCOT (Sweden and Denmark) did not collect hospitalisation resource use information. A fully split approach was not used for the same reason. Also small numbers of patients from some locations and varying follow up times for patients from different countries would have led to increased uncertainty about the validity of outcome measures based on patients from each country alone.

For the purposes of the budget impact analysis, an assumption was made that the clinicians who reported they prescribed 3 months of doublet chemotherapy post-SCOT, used 6 months of doublet chemotherapy pre-SCOT. It would have been optimal to have an estimation of actual prescribing practices pre-SCOT. There is most uncertainty surrounding how this assumption applies to patients with stage II disease because improvement in OS from adding oxaliplatin to fluoropyrimidine has not been demonstrated and, (63) (304) for this reason, clinicians may give fluoropyrimidine monotherapy. This uncertainty was addressed by only including the small proportion of stage II patients that are likely to receive doublet chemotherapy, by varying this parameter in the budget impact sensitivity analysis, and by excluding patients with stage II disease as one of the budget impact scenario analyses. Similarly, despite being included in several international guidelines, the use of doublet chemotherapy in the adjuvant setting for rectal cancer is more controversial compared to colon cancer, with fewer randomised trials to support its use. The budget changes relevant to rectal cancer treatment were therefore more uncertain compared to those for colon cancer and this is the reason why patients with rectal cancer were also excluded within a scenario analysis. Lastly, the budget impact analysis assumed that the proportion of patients receiving FOLFOX (32%) versus CAPOX (68%) pre- and post-SCOT was the same as in the SCOT trial (41). Given the higher uncertainty regarding the non-inferiority of using 3 months of FOLFOX, clinicians may switch to CAPOX. This is most relevant to clinicians from Australia, who showed a preference for FOLFOX within SCOT.

A societal perspective was included in a sensitivity analysis, recognising the merits of taking a broader approach to evaluating the impact from medical research. Specifically, the value of lost earnings was calculated, in line with the previous analysis by Brown et al (184). However, the approach to assessing societal gains in this way are not without controversy. In particular, using the human capital approach has been criticised because it only recognises the productivity loss associated with those of working age, and does not incorporate the value of the loss of leisure time or informal care giving for those of any age in the population. The approach is also limited in that it may overestimate the loss accrued when patients do not attend work due to treatment or illness, because in a real-world setting, their workload may be picked up by a colleague

or their post may be filled with another worker during that time. Although there is disagreement in the literature regarding whether productivity costs relating to treatment should be considered (versus those relating to morbidity and/or mortality only), previous analyses relevant to adjuvant CRC treatment have included the cost to society of patients not being able to attend working during the period they are receiving chemotherapy (296).

In the study by Glover and colleagues (39), the authors calculated an internal rate of return on investment into cancer research as a percentage return on the investment every year for perpetuity. This was not calculated in this study because the QALY gain assessed in the cost-utility analysis was not for the lifetime of patients enrolled in the trial. Not using the internal rate of return approach means that the results from this study cannot be compared with others in the literature that are reported in that manner.

For the SCOT trial, where the QALY gain was minimal and did not meet statistical significance for 3 months versus 6 months of treatment, it is not clear that adding this extra step of valuing QALYs, beyond on the budget impact analysis alone, was necessarily worthwhile. Clearly, the main value from implementing SCOT findings will be on cost savings (the first of Buxton's categories (100) as described in Chapter 3) and in reducing toxicity from treatment, rather than on extra years of life gained. Nevertheless, it was useful to explore how this calculation of QALY gain from implementation is performed because it will be important if applying this approach to other clinical trials, in particular to trials with a superiority end-point. The economic value of implementing superiority trials will be a balance between the QALY gains from using a new treatment (the fourth of Buxton's categories), minus the extra cost of using that treatment in a real world healthcare setting.

6.5 Conclusion

This study has widened the transferability of cost-utility analysis results from the SCOT trial. This type of analysis alongside multi-national clinical trials ensures all countries that recruited to the trial have results relevant to their decision makers. This study has also estimated positive impact on 5-year healthcare budgets from implementing SCOT trial findings of over $150 million USD across

six countries, with potential additional value being added from QALY gain from shorter treatment. These vast savings could fully justify the investment in conducting the SCOT trial.

7 Chapter 7: Impact of the SCOT trial on local practice

7.1 Introduction

Assessing practice change is one important indicator of how clinicians and patients are interpreting and discussing cancer trial results. Evaluating clinical practice will indicate if clinicians consider that a novel treatment that has been tested within a clinical trial is suitable for use within a real world patient population and if the institutional environment is favourable for practice change in line with trial findings.

As shown in Chapter 5, a survey of research users such as clinicians is one tool that can be used to explore the impact of cancer trials on practice. One of the shortcomings of using a survey includes the inherent selection bias, specifically, only a sample of the population of practising clinicians will respond to a survey and it is not always clear how representative the sample is of the population of interest. Another limitation of the survey in Chapter 5 is that clinicians' self-reported practice in the years following publication of SCOT and the IDEA collaboration have been investigated, however there is no documentation of prescribing practices from before the results of these trials were known. Lastly, it was not clear from the survey alone if what clinicians have reported they do aligns with their actions.

To address these issues, and to assess the impact of the SCOT trial on clinical practice in a different way, the aim of the work in this chapter was to evaluate clinical practice change in response to the SCOT trial using administrative healthcare data. Specifically, the chemotherapy prescribing records for patients with a diagnosis of CRC within one health board in Scotland (Greater Glasgow and Clyde (GG&C)) before and after publication of the SCOT trial findings were used. GG&C is the largest health board in Scotland and covers a population of 1.2 million individuals. Within this study, five different approaches to evaluating the impact of the SCOT trial on prescribing practices were tested to explore the optimal way of assessing trial impact.

The objectives of this chapter were to:

i) Describe patient, disease, treatment information and outcomes relevant to the cohort of patients receiving adjuvant chemotherapy for CRC in GGC

ii) Analyse the impact of the SCOT trial findings on clinical practice within this health board

iii) Test a number of quantitative methods to undertake this analysis

7.2 Methods

The process of data acquisition and linkage of datasets is outlined in Chapter 2. The variables used for this analysis and the datasets from which they were derived are outlined in Table 14-2 Appendix 5.

7.2.1 Cohort derivation

The inclusion criteria used to define the relevant cohort of patients for the purpose of this analysis were:

- Patients of any age diagnosed with stage II or stage III colon or rectal cancer who had received adjuvant chemotherapy for CRC with 5-fluorouracil or capecitabine alone or in combination with oxaliplatin. These patients were identified using a regimen descriptor variable in ChemoCare.

- Patients who received adjuvant chemotherapy between January 2010 and March 2018.

The incident, rather than prevalent, population was used. This meant that only those patients who started adjuvant chemotherapy within these dates were included. Follow up was until the end of January 2019.

Exclusion criteria were:

- Patients receiving chemotherapy regimens other than those described

- Patients who received adjuvant chemotherapy as part of a clinical trial

- Patients with known metastatic disease (even if this metastatic disease was resected)

- Patients who received only capecitabine monotherapy in combination with radiotherapy. (Patients were included if they received capecitabine monotherapy concurrently with radiotherapy but also received adjuvant chemotherapy. The concomitant capecitabine prescription was not included for the purposes of this analysis.) Patients who received short course radiotherapy at any stage were included as long as all other criteria were met.

- Patients who received neo adjuvant chemotherapy (other than capecitabine in combination with radiotherapy).

No national data dictionary existed for ChemoCare datasets across Scotland, therefore it was not known if the ChemoCare systems in other locations would have regimen names that could be used for cohort derivation as was the case for ChemoCare WoSCAN. Identifiable clinical portal records were also not available at a national level. Acknowledging these issues, to provide information on how best to define the cohort when using an anonymised national dataset, the study cohort for GG&C was also defined using i) Cancer Audit/QPI and ii) ACaDMe SMR06 datasets, and the three methods of cohort definition were compared.

The method of derivation for the final cohort of patients used for analysis was based on the ChemoCare system. An updated search for the GGC cohort was performed in January 2020 to identify patients who commenced adjuvant chemotherapy between April 2018 and March 2019.

Patients were divided into two groups based on the regimen of adjuvant chemotherapy they received at cycle one of treatment. Although it was recognised that clinicians may switch chemotherapy regimens during treatment due to tolerability, for the purposes of this study, the choice of therapy at cycle one was assumed to indicate a clinician's preferred therapy for each individual patient. The two patients groups were: a) those who received oxaliplatin-fluoropyrimidine

doublet chemotherapy at cycle one and b) those who received fluoropyrimidine monotherapy at cycle one. This split was based on the hypothesis that SCOT trial findings were most likely to impact patients who received the treatment that was tested within the SCOT trial (doublet chemotherapy). The cohort of patients receiving monotherapy at initiation of adjuvant therapy were used as a control/comparison group.

In order to analyse the impact of the SCOT trial findings on practice, a measure of the duration of treatment received by each patient was required. This was calculated by using the number of cycles of chemotherapy delivered, taking account of the regimen of chemotherapy delivered at each cycle, for example, FOLFOX or CAPOX, which are delivered fortnightly or three weekly respectively. This assessment did not utilise the calendar time and therefore between-cycle delays to treatment were intentionally not accounted for within this calculation.

A binary indicator to demonstrate if the duration of chemotherapy received was over 3 months or less than or equal to 3 months was also used. Three months was chosen in line with the experimental arm of the SCOT trial (41). The assumption was made that if a patient received over 3 months of treatment, it was unlikely that the clinician and patient intended to align practice to the experimental arm of the SCOT trial. It was recognised in advance that a proportion of patients would not reach 3 months of treatment due to tolerability, toxicity, or patient choice to stop treatment, rather than a conscious decision from the clinician at the time of treatment initiation to give this duration. However, it was also assumed that the proportion of patients not reaching 3 months of treatment due to tolerability, toxicity or patient choice, would be stable pre and post-SCOT.

The influence of the SCOT trial findings on clinical practice were represented using a dichotomous, time dependent independent variable with 0 representing the time-period prior to 1st June 2017, when the SCOT trial results were first publicised at ASCO, and 1 representing the time-period after June 2017.

7.2.2 Statistical methods used to analyse chemotherapy datasets

7.2.2.1 Descriptive analysis and comparison of proportions

Descriptive statistics (percentages/counts) were used as the first method to analyse the chemotherapy prescribing data. Average duration of chemotherapy delivered was compared at different time-points using the median duration of treatment and the non-parametric Mann-Whitney U test (110). A binary outcome variable representing the duration of treatment received was also used and the proportion of patients receiving longer treatment across two time-points was compared using the Chi-squared test (110).

7.2.2.2 Regression analysis

Univariate linear regression (110) was used to investigate the change in average duration of chemotherapy received (continuous variable), dependent on the timing of the dissemination of the SCOT trial results. Univariate analysis was also performed to explore the effect of patient, disease, and treatment related variables on the duration of chemotherapy prescribed. Multi-variate analysis was used to explore the impact of the SCOT trial on chemotherapy duration whilst accounting for these other co-variates. Co-variates were added to the multivariate model in a step-wise manner (110).

For the purposes of the regression analyses using individual patient level data, the following patient, disease, and treatment related factors were identified in advance as important co-variates to be used within any regression analysis: gender, age (≤ 70 years old and >70 years old), social deprivation category, stage of disease (stage II, low-risk stage III and high-risk stage III) and the chemotherapy regimen used at the first cycle of treatment. This age cut-off was chosen as a clinically relevant, as identified in the piloting stage for the survey (Chapter 5 Results).

The equation used for the multivariate linear regression analysis was as follows:

$$Treatment\ duration\ (weeks) = \beta_0 + \beta_1 SCOT + \beta_2 Gender + \beta_3 Agegroup + \beta_4 SIMD + \beta_5 riskstage + \beta_6 Regimen + e$$

This process was repeated but using logistic regression (110) with a binary outcome (dependent) variable, signifying if a patient received over three months of treatment. The following equation was used for the multivariate logistic regression:

$$Log\ odds\ of\ receiving\ over\ 3\ months\ of\ treatment = \beta_0 + \beta_1 SCOT + \beta_2 Gender + \beta_3 Agegroup + \beta_4 SIMD + \beta_5 riskstage + \beta_6 Regimen + e$$

7.2.2.3 Segmented regression with individual patient level data

Segmented regression, also known as piecewise regression (305), was the third approach tested. This method was used to account for whether a gradual change in practice may have occurred over the whole time-period being studied, which could have explained any change around the time of dissemination of SCOT trial results. Time (in years) was included as a continuous variable from the time the first patient commenced treatment. Two additional variables were also created for each patient to allow calculation of the change in intercept of the fitted regression slopes pre versus post-June 2017. The first variable ("time1") represented time prior to June 2017 and the second represented time after June 2017 ("time 2"). These were created as mutually exclusive variables with time1 taking a value of zero if time2 took any value over zero and time2 taking a value of zero if time1 took any value over zero. Both of these variables were then included in a linear regression model to understand the trend in outcome (treatment duration in weeks) over time in two distinct periods, pre-June and post-June 2017. Two further variables were created to represent the intercepts of the predicted trend lines pre-June 2-17 and post-June 2017. This allowed visualisation of the fitted regression line pre-SCOT and post-SCOT and calculation of the difference in the slopes and intercepts of both lines as an indication of the impact of the SCOT trial findings on clinical practice.

The equations used to apply piecewise regression using linear and logistic regression methods are shown below:

$$Treatment\ duration\ (weeks) = \beta_0 + \beta_1 time1 + \beta_2 time2 + \beta_3 int1 + \beta_4 int2 + e$$

$$Log\ odds\ of\ receiving\ over\ 3\ months\ of\ treatment = \beta_0 + \beta_1 time1 + \beta_2 time2 + \beta_3 int1 + \beta_4 int2 + e$$

An important assumption for the purposes of these analyses was that variables other than time were unlikely to be confounding the effect of the SCOT trial findings given that there were minimal changes in demographics, disease or treatment characteristics in the cohort before and after the time of dissemination.

7.2.2.4 Interrupted time series analysis

The dataset was converted into a time series by calculating the mean duration of treatment received by patients within consecutive, monthly time-periods. These monthly mean durations were plotted graphically to illustrate the change in practice across the whole study period. The time series was interrupted at the time of the initial dissemination of the SCOT trial findings and separate linear regression lines were fitted for the pre versus post-SCOT period. This analysis was repeated by creating a time series of the proportion of patients receiving over 3 months of treatment per month and using logistic regression to test the difference between pre and post-SCOT prescribing.

The three main variables used for the purposes of this analysis were:

T: the time elapsed since the start of the study. This was measured in months.

X_t: a dummy variable indicating the pre-intervention period (coded 0) and post-intervention period (coded 1).

Y_t: the outcome (duration of treatment in weeks when using linear regression or log odds of the proportion of patients receiving over 3 months of treatment when using logistic regression) at time t.

The following regression model was used:

$$Yt = B0 + B1T + B2Xt + B3TXt$$

This model intentionally accounted for the effect of an underlying time trend on the intervention of interest within the time series dataset. *B*0 represented the intercept of the regression curve and the baseline level of the outcome (proportion of patients receiving over 3 months of treatment or treatment duration in weeks) at time zero. *B*1 indicated the underlying time trend across the whole period. *B*2 represented any level change following the intervention and *B*3 indicated the slope change by using the interaction between the underlying time and intervention (TXt). *B* coefficients were derived from the regression output to understand if there was a significant underlying time trend, if there was a significant level change and or if there was a slope change. The counterfactual situation in which the pre-SCOT patterns of prescribing continued unchanged was modelled using linear regression. This allowed the predicted mean treatment duration based on actual data versus the counterfactual situation to be compared. For logistic regression, exp(*B*) represented the corresponding log odds.

This model was checked for seasonality by using a visual inspection of the time series plot (306). To check for autocorrelation, visual inspection of residuals and partial residuals from the model were used (306). This was supplemented by using a Durbin-Watson test (307), with a d-statistic of 2.0 indicating no serial correlation, a result closer to 0 indicating evidence of positive correlation and a result closer to 4 indicating negative correlation.

7.2.2.5 Median percentage dose delivered

The median percentage dose of chemotherapy mediations received by individual patients was calculated as an alternative way to investigate practice change.
The median percentage dose received was compared pre versus post-SCOT using a Mann-Whitney U test (110).

7.3 Results

7.3.1 Cohort derivation

Figure 7-1 shows how the final study cohort using GG&C data was identified using ChemoCare data. Appendix 5 Figure 14-1 shows the results for the comparison of methods for cohort derivation.

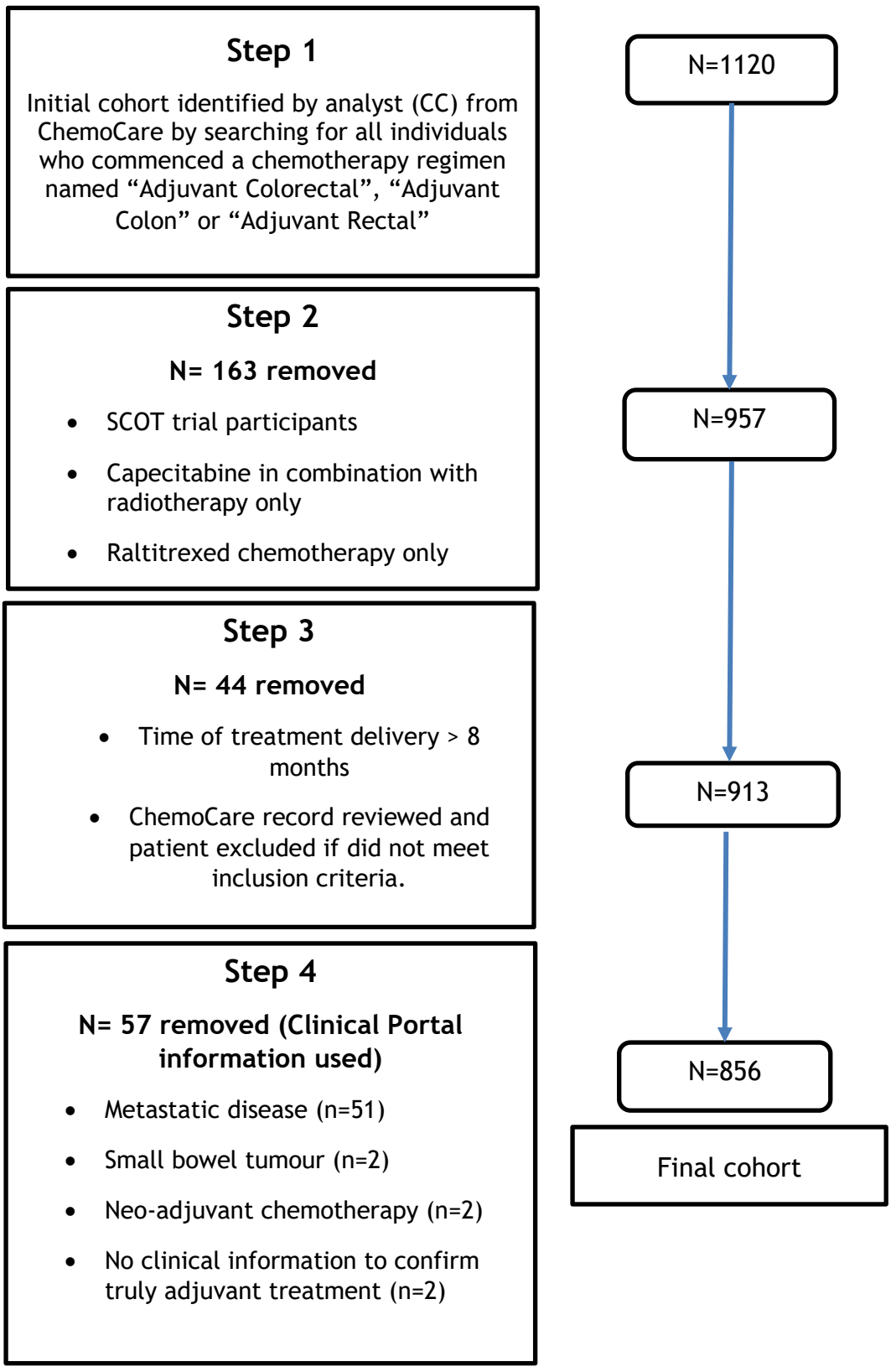


Figure 7-1 GG&C cohort definition using ChemoCare dataset

7.3.2 Descriptive analysis of cohort

The final cohort of patients totalled 998 individuals; 856 from the first data acquisition in March 2019 and 142 from the update in January 2020. Table 7-1 outlines demographic, disease and treatment related variables for the cohort of patients treated with adjuvant chemotherapy in GGC between 2010-2019 and any differences in the cohorts of patients starting chemotherapy pre versus post June 2017.

Table 7-1 Description of whole cohort receiving adjuvant chemotherapy in GGC "Pre-SCOT": before June 2017. "Post-SCOT": After June 2017.

Demographic, disease or treatment related variable	*Pre-SCOT (percentage) n= 755*	*Post-SCOT (percentage) n= 243*	*Total Number (percentage) n=998*
Sex			
Male	394 (52%)	128 (53%)	522 (52%)
Female	361 (48%)	115 (47%)	476 (48%)
Age			
Median age (IQR)	64 (56-71)	65 (56-71)	64 (56-71)
Age group			
≤70 years	555 (74%)	183 (75%)	738 (74%)
>70 years	199 (26%)	61 (25%)	260 (26%)
Deprivation category			
1	257 (34%)	72 (30%)	329 (33%)
2	118 (16%)	48 (20%)	166 (17%)
3	88 (12%)	36 (15%)	124 (12%)
4	111 (15%)	29 (12%)	140 (1%4)
5	176 (23%)	56 (23%)	232 (23%)
Unknown	5 (1%)	2 (1%)	7 (1%)
Location of disease			
Colon	551 (73%)	194 (80%)	745 (75%)
Rectosigmoid or Rectum	204 (27%)	49 (20%)	253 (25%)
Duke's stage			
B	230 (30%)	77 (32%)	307 (31%)
C	521 (69%)	166 (68%)	687 (69%)
Unknown	4 (1%)	0 (0%)	4 (0%)
T stage			
X	1 (0%)	0 (0%)	1 (0%)
0	0 (0%)	1 (0%)	1 (0%)
1	19 (3%)	8 (3%)	27 (3%)
2	46 (6%)	13 (5%)	59 (6%)
3	410 (54%)	144 (59%)	554 (56%)
4	273 (36%)	77 (32%)	350 (35%)
Unknown	6 (1%)	0 (0%)	6 (1%)
N stage			
0	229 (30%)	78 (32%)	307 (31%)
1	339 (45%)	117 (48%)	456 (46%)
2	181 (24%)	48 (20%)	229 (23%)
Unknown	6 (1%)	0 (0%)	6 (1%)
Stage III risk groups			
Low-risk	240 (46%)	83 (50%)	323 (47%)

High-risk	280 (54%)	82 (50%)	362 (53%)
Treatment regimen			
CAPOX	385 (51%)	114 (47%)	499 (50%)
FOLFOX	99 (13%)	40 (16%)	139 (14%)
Capecitabine	255 (34%)	87 (36%)	342 (34%)
IV 5-fluorouracil	15 (2%)	3 (1%)	18 (2%)

The average age of patients receiving adjuvant treatment was 64 years (IQR 56-71), with a higher proportion of men (52%) compared to women (48%). A large proportion (33%) of patients in this group were in the most deprived deprivation category (SIMD 1). This corresponds to the areas of red shading in Figure 7-2 which outlines the distribution of SIMD deciles in the City of Glasgow and surrounding areas. The majority of patients GGC who received adjuvant chemotherapy had a diagnosis of colon cancer (75%) and there was a higher proportion of patients with stage III (69%) compared to stage II (31%) disease.

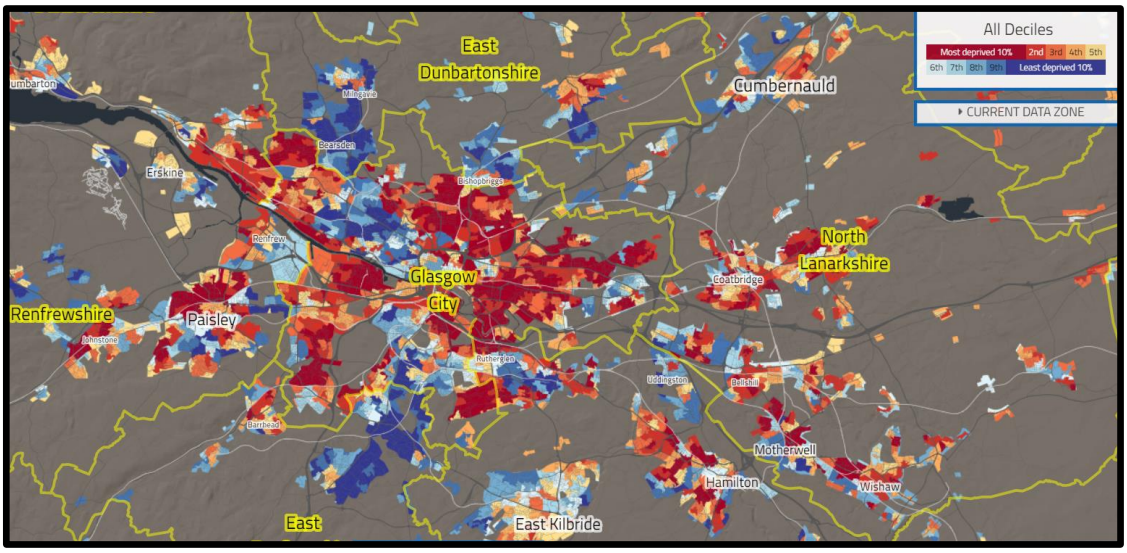


Figure 7-2 Scottish Index of Multiple Deprivation categories (deciles) for Glasgow City and surrounding areas Dark red indicates the most deprived decile. Adapted from source: https://simd.scot/#/simd2020/BTTTFTT/11/-4.2530/55.8579/. Accessed 16th June 2020.

The only significant difference between patients treated prior to June 2017 compared to post June 2017 was a higher proportion of patients with rectal cancer (25% post versus 20% pre) with a corresponding lower proportion of patients with colon cancer treated after June 2017 (75% versus 80%) (χ^2 p=0.034).

In total, (n=638, 64%) of patients initiated treatment with oxaliplatin doublet chemotherapy, whereas (n=360, 36%) of patients received fluoropyrimidine monotherapy at cycle one of treatment. Across both groups, there was a clear

preference (84% of patients) in prescription of capecitabine based regimens (CAPOX or capecitabine alone) compared to regimens using intravenous fluoropyrimidine (FOLFOX or 5-fluorouracil alone, 16% of patients). Table 14-3 and Table 14-4 in Appendix 5 outline the patient, disease, and treatment characteristics for the cohort of patients commencing treatment with doublet chemotherapy versus those receiving fluoropyrimidine monotherapy separately. Figure 14-2 in Appendix 5 shows the treatment received for the 35% (n=349) of patients who switched regimens during their treatment.

7.3.3 Description of practice change

7.3.3.1 Overall cohort

The distribution of treatment duration based on the number of cycles of chemotherapy delivered for the overall cohort is shown below.

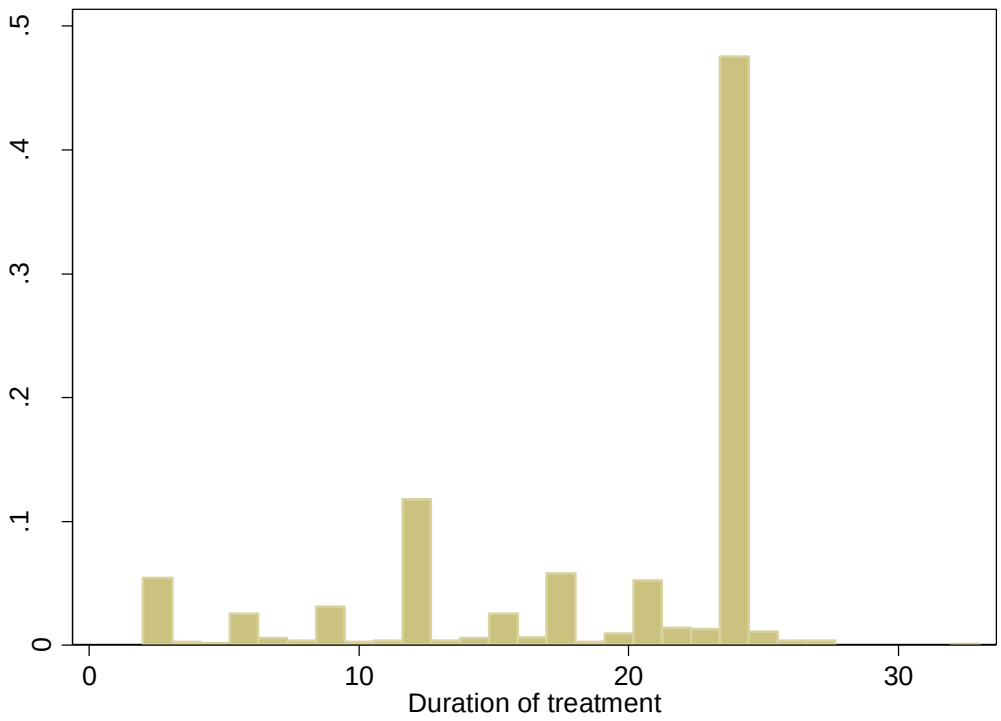


Figure 7-3 Histogram demonstrating the distribution of treatment duration (weeks) for GG&C cohort.

The average duration of treatment delivered based on number of cycles was 24 weeks (median, IQR 12-24 weeks) and overall the majority (73%; 732/998) of patients received over 3 months of treatment. Figure 7-4 shows the proportion of all patients prescribed adjuvant treatment who received over 3 months of

treatment by year of treatment. The yearly median duration of treatment, calculated using number of cycles is also displayed. Overall, the median duration of treatment prior to June 2017 was 24 weeks (IQR 18-24 weeks). This decreased to 12 weeks (IQR 12-24 weeks) for patients commencing chemotherapy after this time point until the end of the study (Mann-Whitney U test p<0.001).

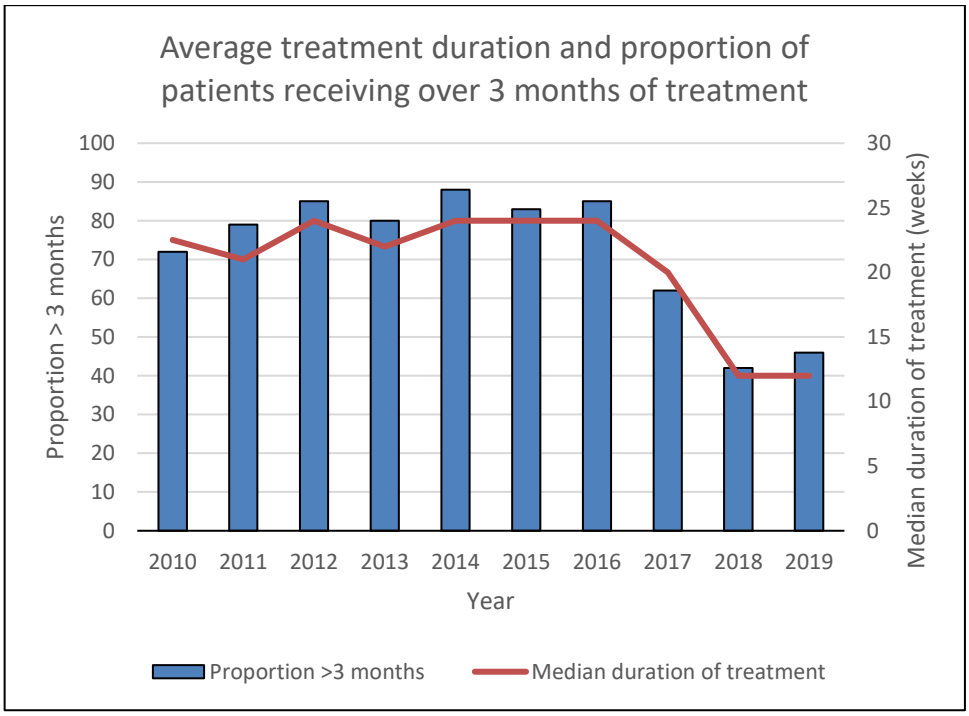


Figure 7-4 Average treatment duration by year Proportion of patients receiving over three months of adjuvant chemotherapy (primary y axis (left)) and median duration of treatment in weeks (secondary axis (right) from 2010-2019.

Figure 7-5 shows there was a decrease from 82% to 47% in the proportion of patients receiving over three months of treatment post versus pre-June 2017 (χ^2 p<0.001).

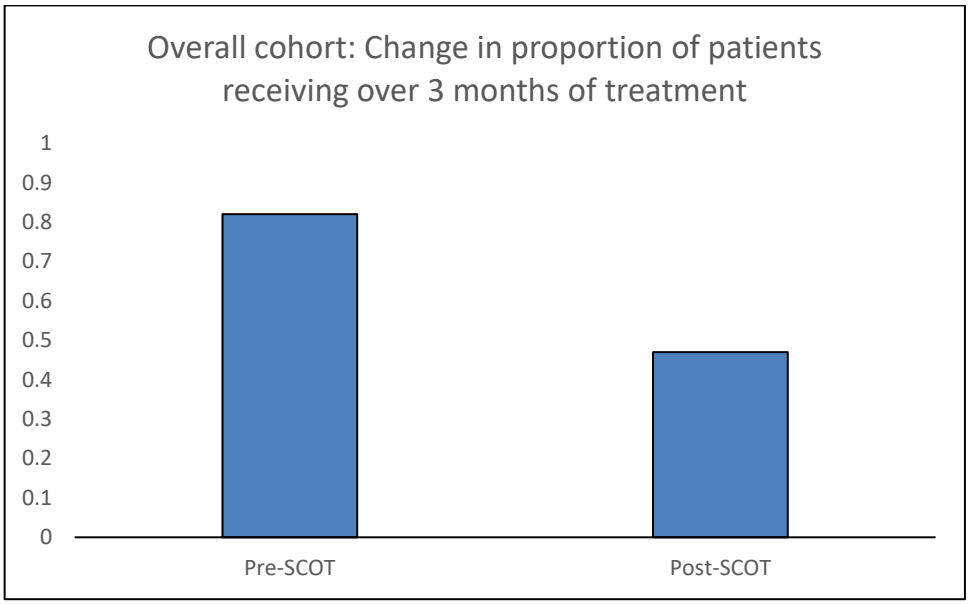


Figure 7-5 Change in proportion of patients receiving over 3 months of treatment.

7.3.3.2 Practice change according to type of regimen

The median duration of treatment for patients specifically receiving doublet chemotherapy at their first cycle of treatment (n=638) was also 24 weeks (IQR 12-24) across the whole time-period. Figure 7-6a shows the proportion of patients who started treatment with oxaliplatin-fluoropyrimidine chemotherapy that received over 3 months of treatment by year, and the median duration of treatment across the whole period (2010-2019). The proportion of patients getting over 3 months of treatment was approximately the same or higher in the years 2010-2016 compared to in the overall cohort. However, the opposite was true for years 2017-2019, with this proportion dropping compared to the overall cohort. The median duration of treatment pre-June 2017 was 24 weeks (IQR 18-24) and after June 2017 was 12 weeks (IQR 12-20 weeks), (Mann-Whitney U test $p<0.001$).

Figure 7-6b displays the same information for patients commencing treatment with single agent fluoropyrimidine chemotherapy. There was minimal difference in the proportion receiving over 3 months of chemotherapy or the median duration of treatment in 2017-2019 compared to the previous years. The median duration of treatment pre-June 2017 was 24 weeks (IQR 15-24) and after June 2017 was 24 weeks (IQR 14-20 weeks) (Mann-Whitney U test $p=0.865$).

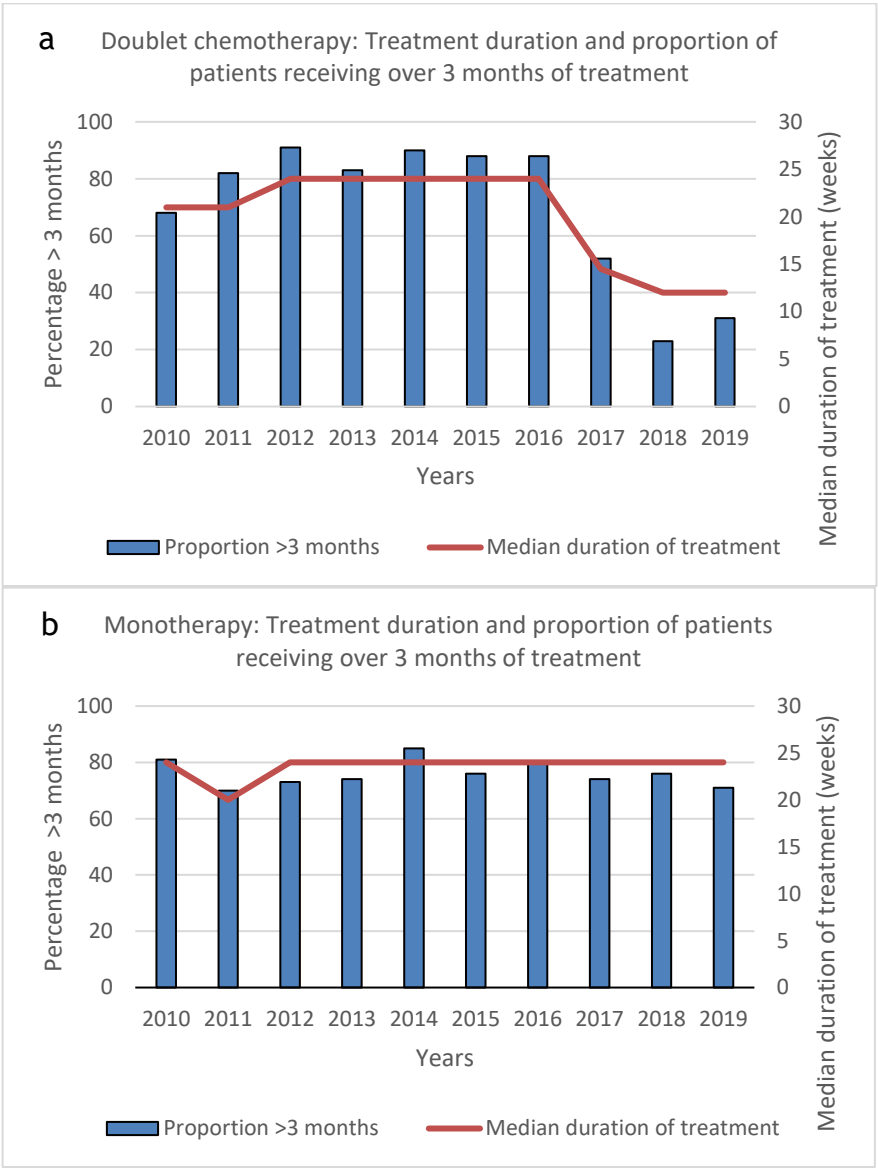


Figure 7-6 Duration of treatment by year dependent on regimen prescribed Proportion of patients treated receiving over 3 months of adjuvant chemotherapy (primary y-axis (left)) and median duration of treatment in weeks (secondary y-axis (right) from 2010-2019. a: Patients receiving doublet chemotherapy at first treatment cycle. b: Patients receiving monotherapy at first treatment cycle.

Dividing time into a dichotomous variable to investigate the impact of the SCOT trial findings, prior to June 2017, 85% of patients who were given oxaliplatin-fluoropyrimidine doublet chemotherapy received over 3 months of treatment. This significantly declined to 31% of patients in the post-SCOT period (χ^2 $p<0.001$) (Figure 7-7). There was no significant change in the proportion of patients receiving over 3 months of treatment after June 2017 (76%) compared to before June 2017 (77%, χ^2 $p=0.774$) for those patients treated with fluoropyrimidine monotherapy.

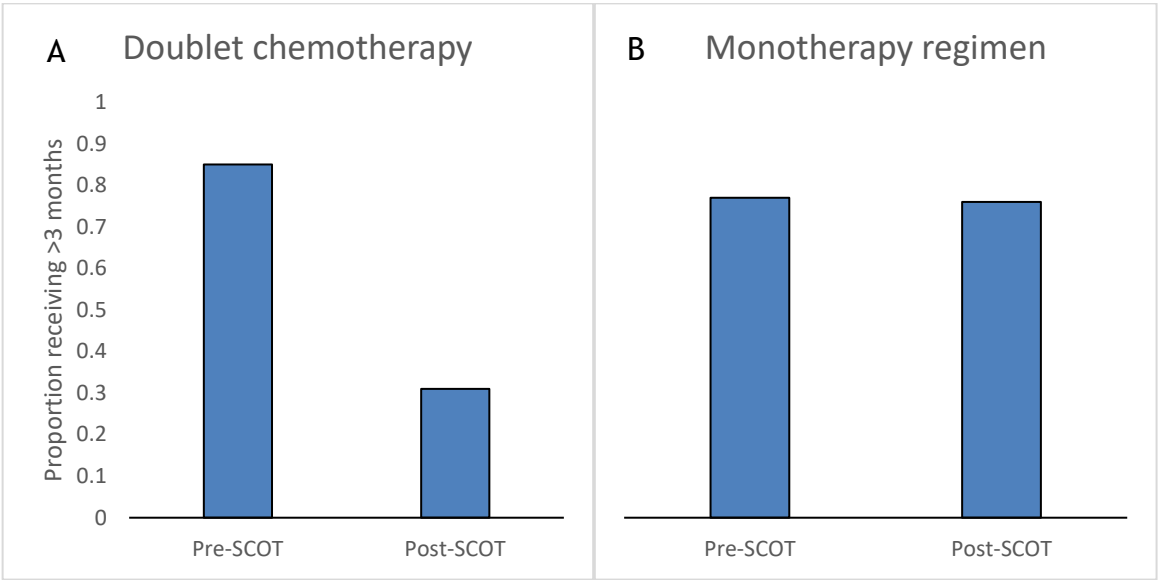


Figure 7-7 Change in proportion of patients who received over 3 months of treatment pre June 2017 versus post June 2017 A: Patients receiving doublet chemotherapy at cycle one. B: Patients receiving fluoropyrimidine monotherapy at cycle one.

7.3.4 Regression analysis

7.3.4.1 Linear regression analysis

There was a significant decrease (6.30 weeks reduction, p<0.001) in the average duration of treatment received after June 2017 compared to before this time-point for patients receiving doublet chemotherapy at their initial treatment cycle (Table 7-2). For patients who received monotherapy at their initial treatment cycle, there was no significant change in treatment duration after, compared to prior to June 2017 (0.32 week decrease, p=0.718), even adjusting for other patient and disease related factors (0.24 week decrease, p=0.794).

For patients treated with doublet chemotherapy, looking at other patient and disease characteristics across the whole study period, those with high-risk stage III disease received 2.7 weeks longer treatment compared to those with stage II disease (p=<0.001) (Table 7-2). Adjusting for patient and disease characteristics, there was still a significant decrease in average treatment duration after June 2017 (“post SCOT”) compared to before this time-point (6.34 weeks reduction, p<0.001). For patients receiving monotherapy, there were no co-variates that significantly affected the average treatment duration on univariate analysis or multivariate analysis.

Table 7-2 Results of univariate and multivariate linear regression

	Doublet		Monotherapy	
	Unadjusted co-efficient (95% CI)	Adjusted co-efficient (95% CI)	Unadjusted co-efficient (95% CI)	Adjusted co-efficient (95% CI)
SCOT				
Post-SCOT	-6.30 (-7.4, -5.21)	-6.34 (-7.45, -5.23)	-0.32 (-2.08, 1.43)	-0.24 (-2.06, 1.58)
Age group				
≥ 70	-1.55 (-2.98, 0.16)	-1.51 (-2.83, -0.19)	-0.76 (-2.29, 0.76)	-0.47 (-2.18, 1.24)
Sex				
Female	-0.37 (-1.40, 0.66)	-0.18 (-1.13, 0.76)	-0.46 (-1.98, 1.06)	-0.53 (-2.08, 1.01)
Risk stage				
Low-risk III	0.99 (-0.55, 2.53)	1.91 (0.47, 3.34)	-1.65 (-3.62, 0.33)	-1.75 (-3.90, 0.40)
High-risk III	2.71 (1.20, 4.23)	3.07 (1.66, 4.48)	-0.29 (-2.21, 1.63)	-0.16 (-2.26, 1.93)
Regimen				
FOLFOX	0.03 (-1.22, 1.28)	0.19 (-0.98, 1.36)	0.92 (-2.57, 4.40)	0.75 (-2.86, 4.37)
Disease site				
Rectal	-0.41 (-1.56-0.75)	-0.92 (-1.99, 0.15)	0.53 (-1.31, 2.36)	0.57 (-1.34, 2.49)
SIMD				
2	-0.44 (-2.00, 1.12)	-0.22 (-1.64, 1.19)	-2.17 (-4.43, 0.01)	-2.28 (-4.59, 0.03)
3	-0.24 (-2.00, 1.53)	0.34 (-1.27, 1.95)	-1.92 (-4.44, 0.51)	-2.06 (-4.52, 0.40)
4	0.63 (-1.00, 2.26)	0.54 (-0.94, 2.03)	0.17 (-2.27, 2.62)	0.12 (-2.35, 2.59)
5	-0.10 (-1.48, 1.28)	0.15 (-1.12, 1.41)	-0.99 (-3.08, 1.10)	-1.11 (-3.25, 1.03)

Focusing on patients who received doublet chemotherapy, on multivariate analysis there was a significant interaction between the dichotomous time variable (pre versus post June 2017) and regimen delivered at cycle one ($p<0.001$), and between the same time variable and disease risk stage ($p=0.0059$). Patients receiving CAPOX were more likely have shorter treatment after June 2017 compared to before this time-point whereas the same change was not seen for patients receiving FOLFOX (Table 7-3). Regarding disease risk stage, there was a significant decrease in treatment duration after June 2017 across all risk stages but the change was most marked for patients with low-risk stage III disease (8.6 weeks decrease (95% CI 9.97, -7.23, $p<0.001$).

Table 7-3 Change in treatment duration post-SCOT for patients prescribed doublet treatment dependent on regimen and risk stage

	Average change in treatment duration post-SCOT (SCOT co-efficient) in weeks	*p value*
First regimen		
CAPOX	-8.05 (-9.25, -6.84)	<0.001
FOLFOX	-1.02 (-3.44, 1.39)	0.404
Risk stage		
Stage II	-5.64 (-9.65, -1.62)	0.006
Low-risk III	-8.60 (-9.97, -7.23)	<0.001
High-risk III	-4.04 (-5.75, -2.34)	<0.001

Regarding patients who received fluoropyrimidine monotherapy, there was a significant interaction with time as a dichotomous variable and disease risk stage (p=0.0206) and age group (p=0.0067). Looking at these subgroups in more detail demonstrated that there was no significant change in treatment duration regardless of age or disease risk stage (Table 7-4).

Table 7-4 Change in treatment duration post-SCOT for patients prescribed monotherapy dependent on age and risk stage

	Average change in treatment duration post-SCOT in weeks	*p value*
Age		
<70	0.80 (-4.86, 3.27)	0.696
>70	-0.34 (-2.42, 1.73)	0.744
Risk stage		
Stage II	0.89 (-1.32, 3.10)	0.429
Low-risk III	-2.15 (-8.24, 3.95)	0.484
High-risk III	-3.29 (-6.85, 0.27)	0.069

7.3.4.2 Logistic regression analysis

Focusing on patients who received doublet chemotherapy, the odds of a patient receiving over 3 months of treatment decreased by 92% after June 2017 compared to before this time-point (OR 0.08, 95%CI 0.05-0.12, p<0.001). There was minimal change in these odds (OR 0.07, 95% CI 0.04-0.11, p<0.001) in a multivariate analysis adjusting for patient and disease related characteristics (Table 7-5).

For patients who received monotherapy, there was no significant change in the proportion of patients receiving over 3 months treatment (OR 0.92, 95%CI 0.53-1.61, p=0.774) post-SCOT. Patients in SIMD category 2 were less likely (OR 0.47,

95%CI 0.23-0.96, p=0.038) than those in category 1 (most deprived) to receive over 3 months of treatment but the overall effect of SIMD on treatment duration was not significant (p=0.1367). There were no other patient or disease characteristics that significantly affected the proportion of patients receiving over 3 months of treatment.

Table 7-5 Results of univariate and multivariate logistic regression.

	Doublet		Monotherapy	
	Unadjusted odds ratio (95% CI)	Adjusted odds ratio (95% CI)	Unadjusted odds ratio (95% CI)	Adjusted odds ratio (95% CI)
SCOT				
Post-SCOT	0.08 (0.05-0.12)	0.07 (0.04-0.11)	0.92 (0.53-1.61)	0.92 (0.51-1.66)
Age group				
≥ 70	0.63 (0.40-0.99)	0.54 (0.31-0.94)	0.65 (0.40-1.06)	0.64 (0.37-1.13)
Sex				
Female	0.90 (0.64-1.27)	0.95 (0.63-1.44)	0.78 (0.48-1.28)	0.74 (0.45-1.23)
Risk stage				
Low-risk III	0.93 (0.57-1.52)	1.47 (0.82-2.62)	0.60 (0.33-1.10)	0.66 (0.34-1.30)
High-risk III	2.05 (1.23-3.42)	3.23 (1.75-6.00)	0.96 (0.51-1.80)	1.14 (0.56-2.29)
Regimen				
FOLFOX	1.08 (0.71-1.64)	1.20 (0.71-2.02)	1.07 (0.34-3.34)	0.94 (0.28-3.16)
Disease site				
Rectal	1.03 (0.70-1.51)	0.84 (0.53-1.34)	1.04 (0.57-1.88)	1.02 (0.54-1.92)
SIMD				
2	0.76 (0.45-1.26)	0.74 (0.40-1.37)	0.47 (0.23-0.96)	0.43 (0.21-0.91)
3	0.88 (0.49-1.59)	1.15 (0.56-2.37)	0.52 (0.24-1.13)	0.47 (0.22-1.04)
4	0.88 (0.51-1.52)	0.78 (0.41-1.50)	1.14 (0.47-2.80)	1.14 (0.46-2.83)
5	0.86 (0.54-1.36)	0.90 (0.52-1.57)	0.63 (0.32-1.26)	0.61 (0.30-1.25)

As found on linear regression, using logistic regression, there was a significant interaction for patients receiving doublet chemotherapy between both the regimen received (p<0.001) and disease risk stage (p=0.031) with the timing of the SCOT trial results. Subgroup analysis showed that for patients prescribed CAPOX treatment at cycle one, the odds of receiving over 3 months of treatment decreased by 97% post-SCOT (p<0.001) whereas for those prescribed FOLFOX the odds decreased by 51% and this drop was non-significant (p=0.1130). There was a significant decrease in the odds of receiving over 3 months of treatment across all disease stages, but the change was most marked for patients with low-risk stage III disease (OR 0.03, 95%CI 0.01-0.06, p<0.001). Table 7-6 shows the odds ratio for receiving over 3 months of treatment post-SCOT compared to pre-SCOT by regimen and risk stage.

Table 7-6 Change in odds of receiving over 3 months of treatment post-SCOT for patients prescribed doublet treatment dependent on regimen and risk stage

	Odds ratio for receiving over 3 months of treatment post-SCOT compared to before this time-point	*p value*
First regimen		
CAPOX	0.03 (0.02, 0.06)	<0.001
FOLFOX	0.49 (0.21, 1.18)	0.113
Risk stage		
Stage II	0.06 (0.01, 0.29)	<0.001
Low-risk III	0.03 (0.01, 0.06)	<0.001
High-risk III	0.16 (0.08, 0.30)	<0.001

For patients receiving monotherapy, disease risk stage (p=0.0040) and age group (p=0.0080) interacted significantly with the timing of publication of SCOT trial results. There was no significant change in the proportion of patients receiving over 3 months of treatment post-SCOT for either age group. Regarding patients with different disease stages, there was no significant change in prescribing for patients with stage II or low-risk stage III disease but there was a significant decrease in the proportion of patients receiving over 3 months for individuals who had high-risk stage III disease (OR 0.27, 95%CI 0.07, 0.98, p=0.047). Table 7-7 shows the odds ratio for receiving over 3 months of treatment post-SCOT compared to pre-SCOT by age group and risk stage.

Table 7-7 Change in odds of receiving over 3 months of treatment post-SCOT for patients prescribed monotherapy dependent on regimen and risk stage

	Change in odds of receiving over 3 months of treatment post-SCOT	*p value*
Age		
<70	0.62 (0.28, 1.35)	0.227
>70	1.33 (0.51, 3.45)	0.563
Risk stage		
Stage II	1.69 (0.74, 3.88)	0.213
Low-risk III	0.33 (0.66, 1.92)	0.217
High-risk III	0.27 (0.07, 0.98)	0.047

7.3.5 Segmented regression

7.3.5.1 Segmented linear regression

The third approach used to investigate clinical practice change was segmented logistic and linear regression. Assessing the impact of the SCOT trial findings by using the date when the trial results were disseminated, whilst also accounting for any time related trend, there was a significant change in the number of weeks of doublet treatment received after versus before June 2017. The average duration of doublet chemotherapy delivered just before June 2017 was 22.0 weeks of chemotherapy (95% CI 21.0-23.0) and just after June 2017 was 16.0 weeks (95% CI 14.0-18.0), which was a significant decrease (Figure 7-8, -6.5 weeks, 95% CI-8.3 to -3.8, p<0.001). In addition to this level change, there was also significant change in treatment duration over time in the post SCOT period (graph slope post June 2017 in Figure 7-8) compared to the pre-SCOT period (slope change -2.2 weeks/year, 95%CI -3.8 to -0.57, p=0.008).

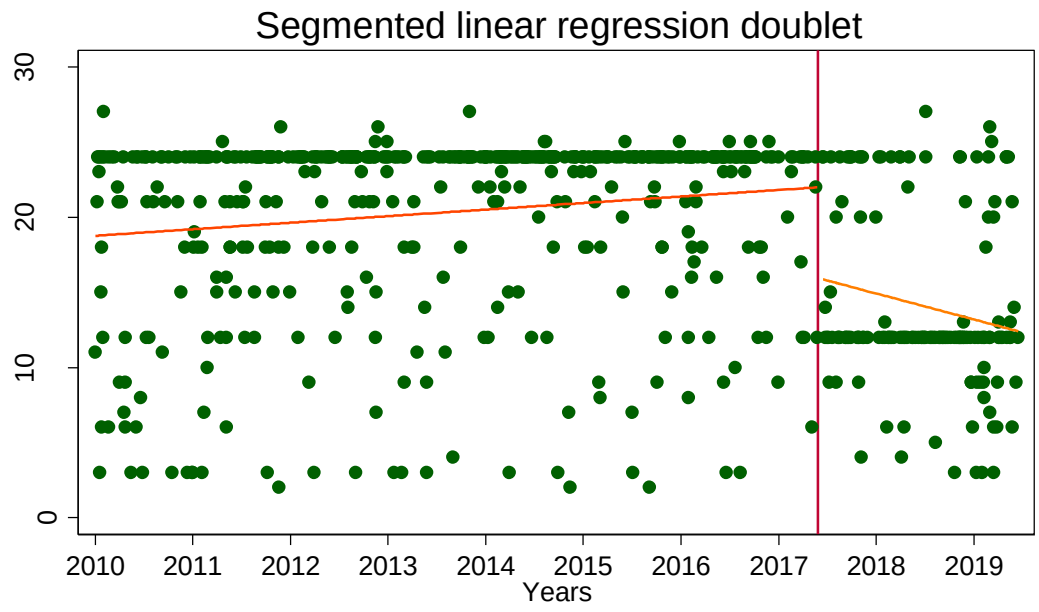


Figure 7-8 Segmented linear regression (doublet)

Patients receiving monotherapy in the adjuvant setting were used as a comparator group. There was no significant change in prescribing identified pre versus post June 2017 in either the treatment duration just before versus just after June 2017 (0.56 week decrease, 95%CI -4.06, 2.94, p=0.753) or in the time trend pre versus post-SCOT (Figure 7-9) (slope change 0.05 weeks/year, 95%CI -

2.56, 2.66, p=0.971). There was one obvious outlier in this plot (2014 receiving over 30 weeks of treatment). The individual patient level data was reviewed for this patient and there was no reason to indicate that this was not adjuvant therapy being delivered. For this reason, this individual was kept within the analysis.

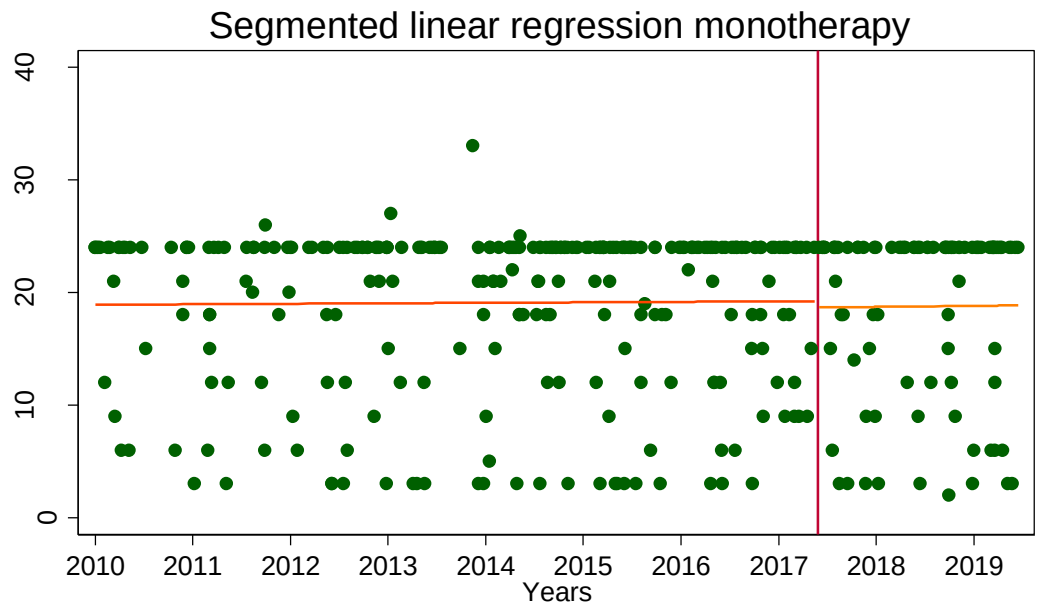


Figure 7-9 Segmented linear regression (monotherapy)

7.3.5.2 Segmented logistic regression

Segmented logistic regression (Figure 7-10) showed a significant change in the proportion of patients receiving over 3 months of treatment post June 2017 compared to pre June 2017 when accounting for this underlying time trend for individuals who received doublet chemotherapy at cycle one of treatment. There was a significant change in the level of the regression lines (OR 0.07 (0.03-0.16 p<0.001) corresponding with a 93% decrease in the odds of a patients receiving over 3 months of treatment post versus pre-SCOT. Despite this acute jump in prescribing, there was no significant difference in the slope (trend) of the fitted regression lines in the pre versus post June 2017 trend (OR 0.60 (0.34-1.08, p=0.091). This indicates that there was no difference in the trend over time pre-SCOT versus post-SCOT in the proportion of patients receiving over 3 months of treatment.

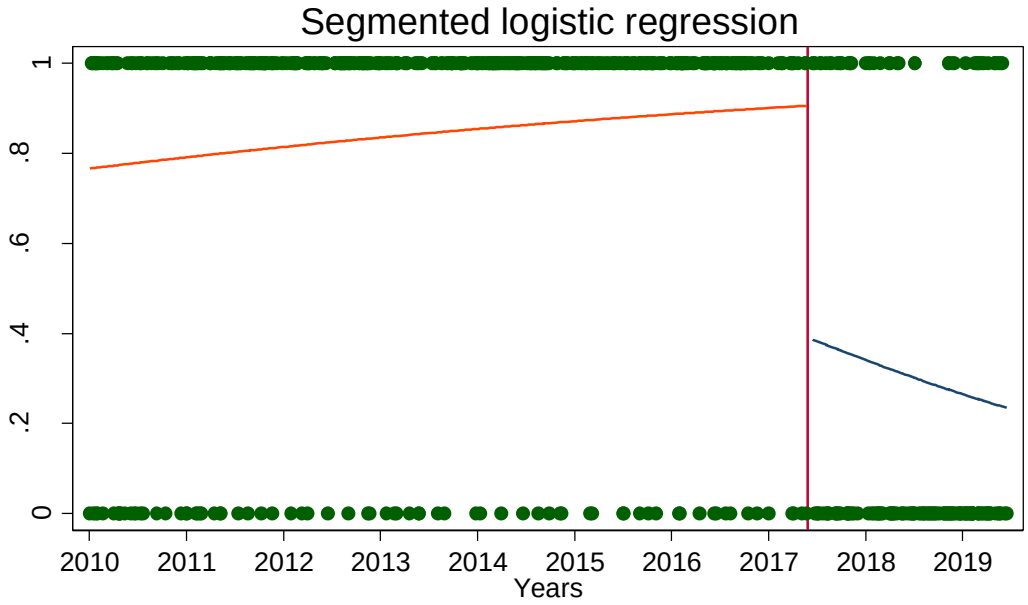


Figure 7-10 Segmented logistic regression (doublet)

In comparison, the same analysis applied to patients initiating treatment with monotherapy demonstrated minimal change in the fitted trend (Figure 7-11) post versus pre June 2017. Specifically, there was no significant change in the level/intercept (OR 1.14 (95% CI 0.36-3.55) p=0.827) or slope (OR 0.77 (95% CI 0.35-1.69) p=0.512) of the pre versus post SCOT prescribing practices.

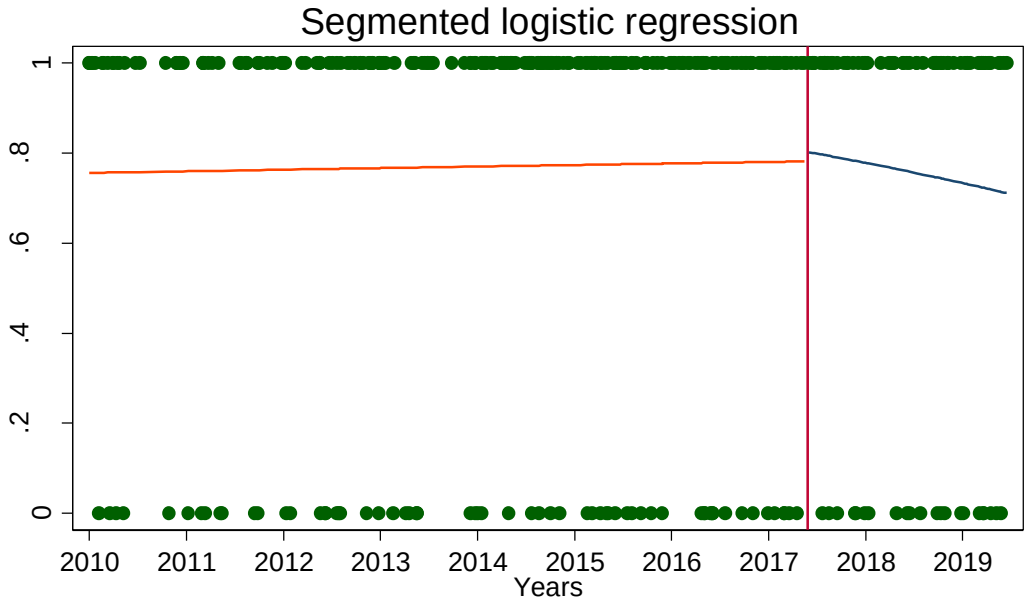


Figure 7-11 Segmented logistic regression (monotherapy)

7.3.6 Interrupted time series analysis

7.3.6.1 ITSA using a continuous outcome variable

The fourth approach used to analyse the impact of the SCOT trial findings on practice was interrupted time series analysis. First, a scatter plot of the time series data was constructed to identify any underlying trends in the time series data. On visual inspection of the time series plot (Figure 7-12) there were no obvious seasonal trends in chemotherapy prescribing and there was no obvious autocorrelation in the time series (see Appendix 5 for results of these tests).

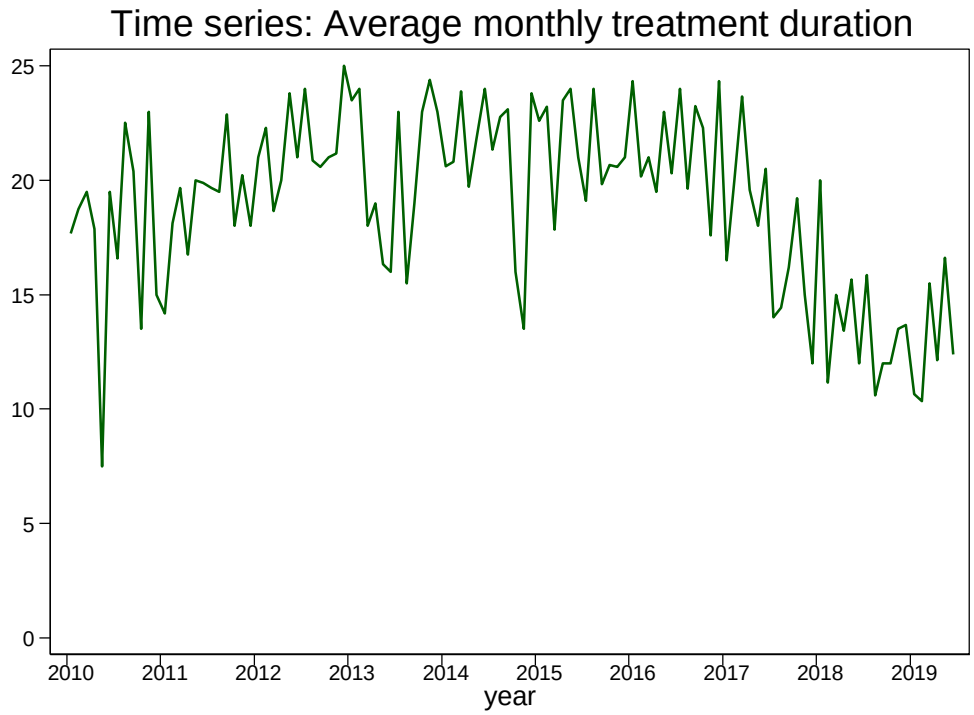


Figure 7-12 Time series plot for patients receiving doublet chemotherapy (average monthly duration of treatment)

This time series was intentionally interrupted at the date 1st June 2017, and trends pre and post June 2017 were assessed and compared. There was an upward trend in the average duration of treatment prior to this time point (slope: increase in 0.04 weeks/month (95%CI 0.01, 0.06) p<0.001) (Figure 7-13) and a downward trend in treatment duration post-SCOT (slope: decrease by 0.18weeks/month (95%CI -0.32, -0.04) p=0.012). There was a significant change in average treatment duration post versus pre June 2017 (decrease by 5.72 weeks (95% CI -8.20,-3.24) p<0.001) with a statistically significant change in the slopes (slope change -0.22 (95%CI -0.37, -0.07) p=0.005). The counterfactual

situation was modelled in which the SCOT trial results were not disseminated, and the pre-June 2017 trend continued unchanged. This counterfactual situation is represented by the blue dashed line in Figure 7-13.

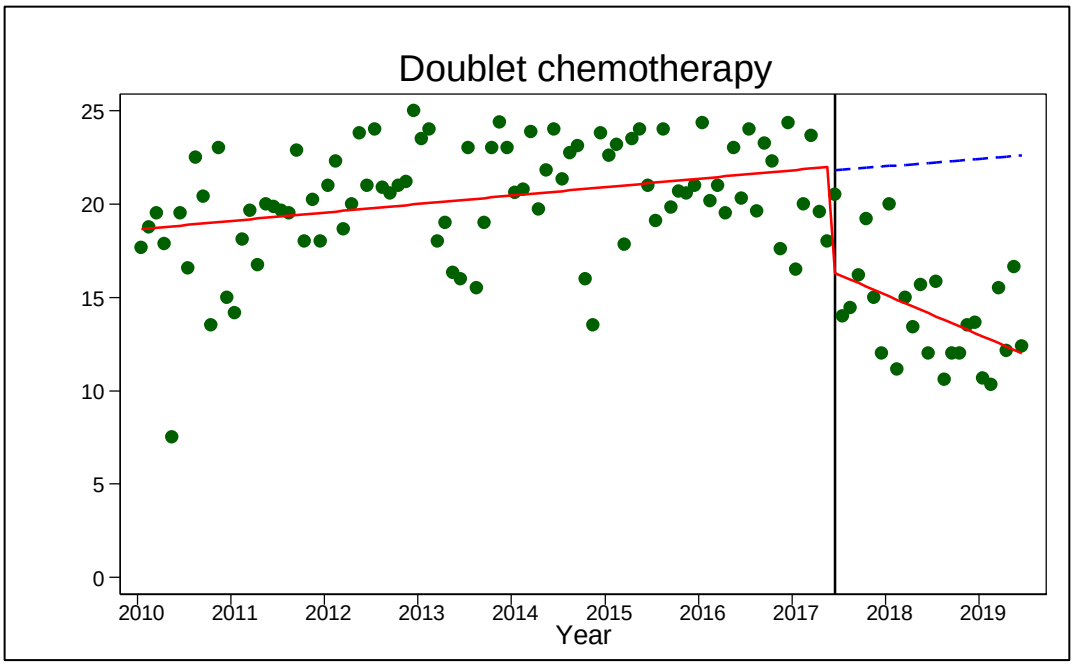


Figure 7-13 ITSA doublet (linear regression)

Patients receiving monotherapy at cycle 1 were used a comparator group. This cohort achieved the pre-specified criteria for a suitable control group using ITSA because there was no difference in the average proportion of patients receiving over 3 months of treatment prior to June 2017 (p=0.241) or any difference in the trend over time (line slope) pre-SCOT (p=0.818) compared to patients receiving doublet chemotherapy. The time series for this patient cohort showed no obvious seasonal trends in prescribing (Figure 7-14) and there was no autocorrelation (see Appendix 5 for test results).

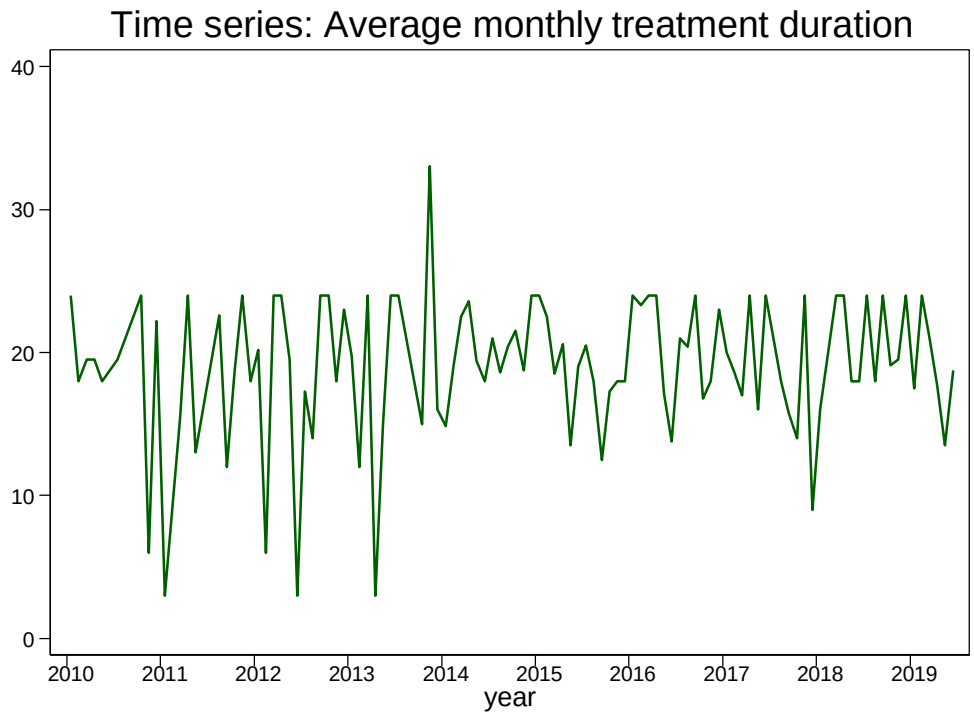


Figure 7-14 Time series plot for patients receiving monotherapy (average monthly duration of treatment)

Interrupting the time series at June 2017 showed there was no significant difference in post SCOT versus pre-SCOT prescribing practices for this patient cohort (Figure 7-15). There was a non-significant trend over time in average treatment duration pre-SCOT (0.03 weeks/month (95%CI -0.01, 0.07) p=0.162) which was unchanged post-SCOT (0.03 weeks/month (95%CI -0.20, 0.25) p=0.826), difference -0.01 (95%CI -0.29, 0.28) p=0.970). There was no change in the average duration of treatment received around the June 2017 time-point (decrease by 1.12 weeks (95%CI -5.62, 3.37) p=0.625).

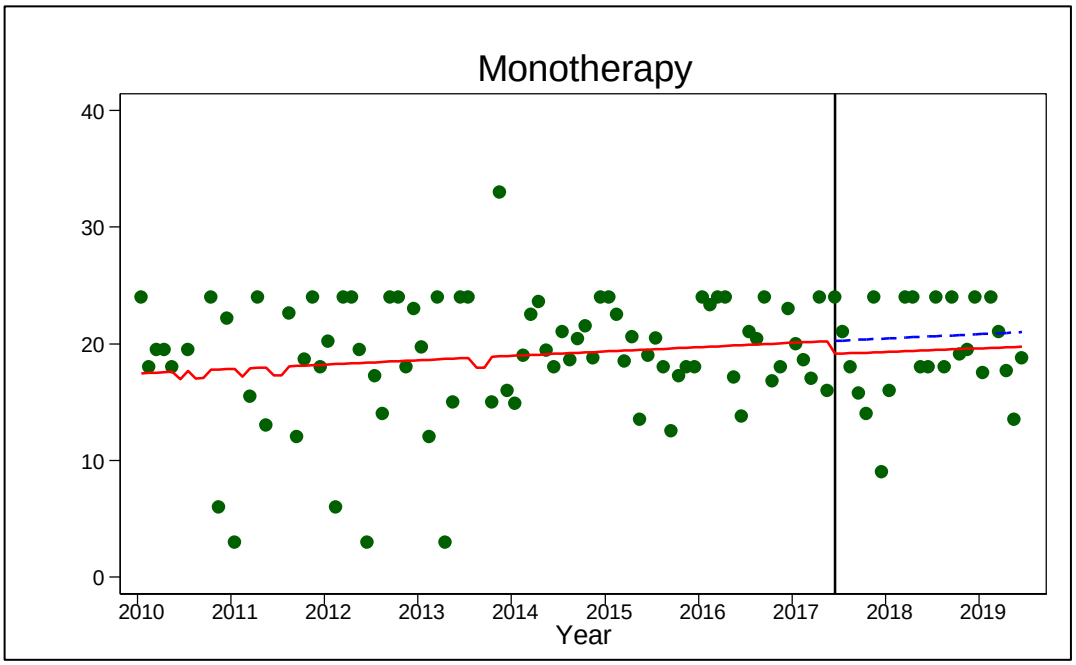


Figure 7-15 ITSA monotherapy (linear regression)

7.3.6.2 ITSA using a binary outcome variable

This analysis was repeated by plotting the proportion of patients who received over 3 months of treatment per month in GG&C from 2010-2019 and investigating the difference in prescribing pre versus post June 2017. There was no obvious seasonality and no autocorrelation detected for either of these time-series (see Appendix 5).

For patients receiving doublet treatment, interrupting the time series at June 2017 showed a significant change in the odds of receiving over 3 months of treatment post-SCOT (decrease in odds by 94% (OR 0.06, 95%CI 0.02, 0.17), p<0.001). There was no significant change in the slope of the pre versus post SCOT for the trend in patients receiving over 3 months of treatment over time (difference in slopes 1.06 (95%CI 0.90, 1.00) p=0.068) (Figure 7-16).

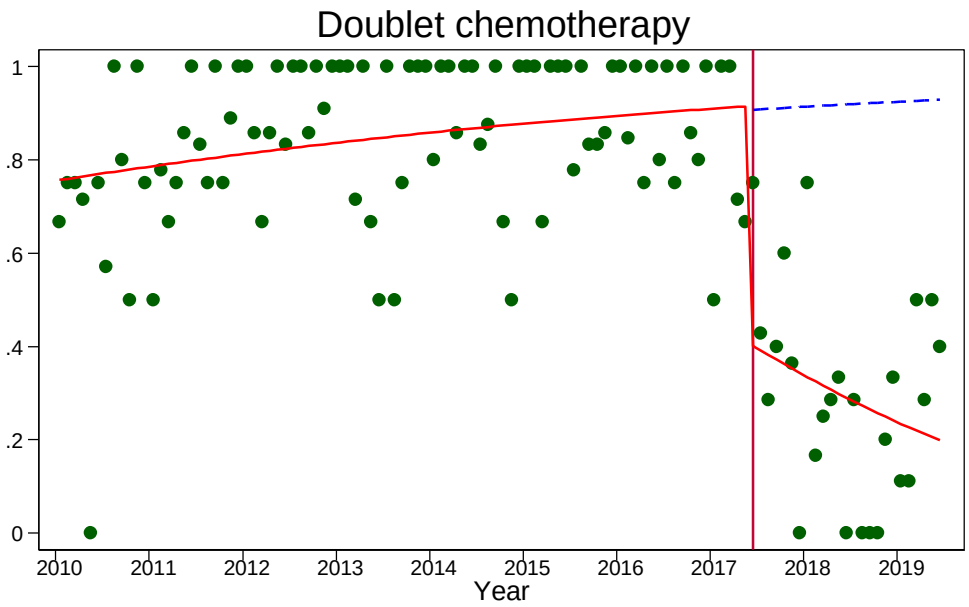


Figure 7-16 ITSA doublet (logistic regression)

For patients receiving monotherapy, there was no significant change in prescribing identified (Figure 7-17). The level change in the odds of a patient receiving over 3 months of treatment was 1.12 ((95%CI 0.32, 3.82) p=0.863) and the slope change was 0.96 ((95%CI 0.90, 1.02) p=0.174) (Figure 7-17).

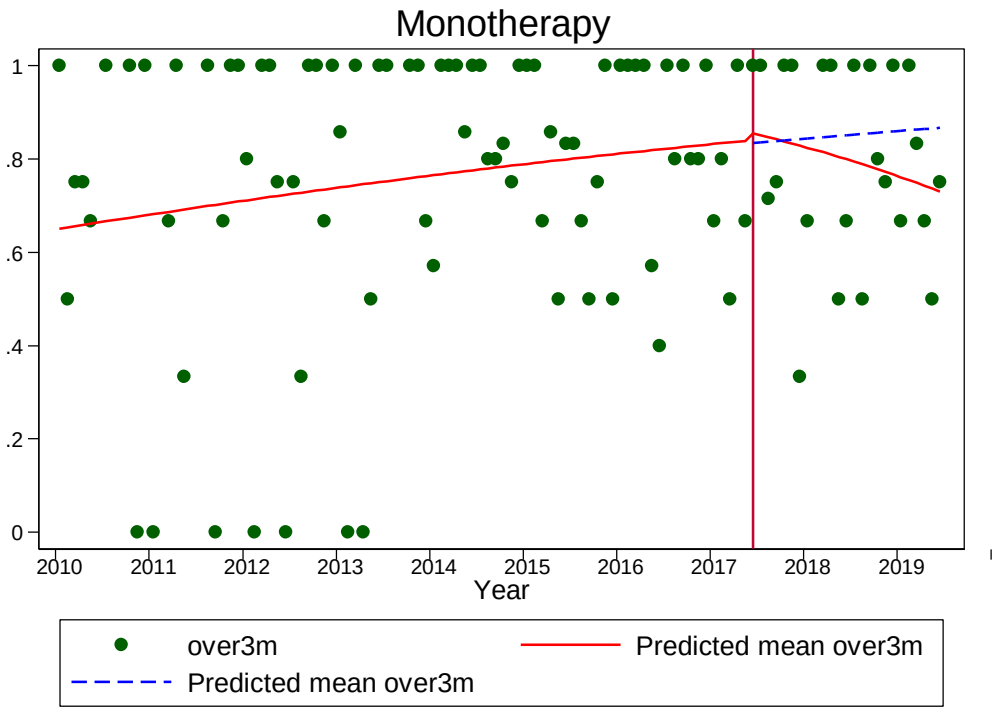


Figure 7-17 ITSA monotherapy (logistic regression)

Table 7-8 summarises results from the four ITSAs, showing the average duration of chemotherapy delivered and the estimated proportion of patients receiving over 3 months of treatment at the last time series interval (June 2019), both for patients receiving doublet chemotherapy at and those receiving monotherapy at cycle one of treatment. The same outcomes estimated from the counterfactual plots are included for comparison. These give an estimation of the prescribing practices if no change at occurred in June 2017 and the trend line from pre-June 2017 had continued unchanged.

For patients initiating treatment with doublet chemotherapy, the proportion of patients getting over 3 months of treatment decreased from 93% (95% CI 84-97) in the modelled counterfactual situation to 20% (95% CI 10-37) using the real world data. Similarly, the average treatment duration of treatment fell from 23 weeks (95% CI 22-23) to 12 weeks (95% CI 11-13) for this patient cohort. For those receiving monotherapy at cycle one, the decrease was less marked and non-significant with 73% (95% CI 60-84) of patients receiving over 3 months of treatment in June 2019 versus 87% (95% CI 74-94) in the counterfactual situation. The change in average treatment duration from 21 weeks in real life to 20 weeks in the predicted model.

Table 7-8 Comparison of outcomes from time series versus counterfactual models

	Prescribing practice post-SCOT (June 2019)	Doublet (95% CI)	Monotherapy (95% CI)
Time series model	Proportion of patients receiving >3 months of treatment (%)	20 (10-37)	73 (58-84)
	Mean duration of treatment (weeks)	12 (11-13)	20 (18-22)
Counterfactual	Proportion of patients receiving >3 months of treatment (%)	93 (84-97)	87 (74-94)
	Mean duration of treatment (weeks)	23 (22-23)	21 (19-23)

Taking the cohort of patients and the outcome that showed the biggest change across June 2017, an exploratory analysis was done for patients receiving doublet chemotherapy. ITSA was repeated using the binary outcome of the proportion of patients receiving over three months of treatment, but interrupting the time series at two arbitrary time-points: June 2014 and January

2016. The results are shown in Appendix 5. There was no significant change in the trend or level of practice change at these time points.

7.3.7 Median dose delivered

The last approach used to investigate practice change post-SCOT trial was calculation of the average dose of treatment delivered before and after trial dissemination. The median dose of fluoropyrimidine and oxaliplatin delivered to patients in GG&C before versus after June 2017 is shown in Table 7-9 according to the regimen that patients were prescribed at cycle one of their treatment. There was a clear decrease in the median dose delivered for patients receiving CAPOX, but this was not replicated for other regimens.

Table 7-9 Median chemotherapy doses received

	Fluoropyrimidine		
First Regimen	Median dose delivered (IQR)		
	Pre-SCOT	Post-SCOT	Difference (Mann-Whitney U)
CAPOX	86% (72-96%)	49% (41-51%)	p<0.001
FOLFOX	81% (50-93%)	82% (49-98%)	p=0.6028
Capecitabine	77% (51-92%)	81% (48-91%)	p=0.7799
5-fluorouracil	89% (63-97%)	90% (8-99%)	p=0.8590
	Oxaliplatin		
First Regimen	Median dose delivered (IQR)		
	Pre-SCOT	Post-SCOT	Difference (Mann-Whitney U)
CAPOX	63% (42-79%)	46% (37-50%)	p<0.001
FOLFOX	61% (44-72%)	53% (45-75%)	p=0.7097
Capecitabine	0% (0-0%)	0% (0-0%)	p=0.3103
5-fluorouracil	0% (0-0%)	0% (0-65%)	p=0.3159

7.4 Discussion

A significant change in the average duration of adjuvant chemotherapy received and in the proportion of patients receiving over 3 months of chemotherapy occurred after the dissemination of the SCOT trial findings in June 2017 for patients treated in GG&C. This impact was driven by a change in adjuvant treatment with a doublet regimen.

Comparing patients who were initiated on doublet treatment in GG&C health board between 2010-2019 (see Table 14-3 Appendix 5) compared to patients

enrolled in the SCOT trial, GG&C patients were, on average, younger (median age GGC 62 years versus 65 years in SCOT trial). Also, there were more patients with rectal cancer in the GGC cohort (22% vs 18%) however, this may have been related to the fact that rectosigmoid tumours were included in this categorisation in the GGC cohort whereas they may have been included with colon cancers in the SCOT trial. In the SCOT trial, 33% of patients received FOLFOX chemotherapy, and in particular 31% of UK patients in SCOT received FOLFOX (see results in Chapter 6). This compared to 22% in the GG&C cohort of patients. There was a similar proportion of patients with stage II cancer in both the SCOT trial (18%) and this real world cohort (15%).

Using five different approaches, alternative ways of quantitatively assessing the extent of the SCOT trial impact were explored. The first approach was a descriptive analysis and comparison of proportions. These were a powerful way of identifying the main message from the data, this is, that practice did change across the time-point of interest, and this approach provided in basic terms, the extent of that change.

An advantage of the second approach used was that several factors other than the SCOT trial which could potentially impact treatment duration, or affect the likelihood of a patient receiving over 3 months of treatment, were assessed and accounted for in the analysis of practice change. On reflection, the main research question was to investigate the timing of the SCOT trial on treatment duration. Unless there was a strong pre-existing reason to expect that patient, disease, or regimen related factors were significantly different in the pre-versus post-SCOT period, it is unlikely these variables would be confounding the effect of the SCOT trial on the outcome of interest. This was supported by the results of this analysis, which showed there was minimal difference in the change in prescribing post June 2017 when accounting for these factors; indeed the extent of practice change was larger when these factors were incorporated into a multivariate analysis. Although an important analysis to have performed, this gave confidence to the decision not to include these factors within the next two analysis approaches.

The third approach used was segmented regression and the main advantage over the regression already performed was that the underlying effect of time trends

were accounted for whilst also using individual patient level data. This approach addressed the question of whether any changes in prescribing identified were a reflection of an underlying trend over time that would have occurred regardless of the SCOT trial, rather than the trial itself. Although using individual patient level data was an advantage as it used every data point available, plotting every individual's duration of treatment or binary outcome as to whether they received over 3 months of treatment meant the plots were crowded and susceptible to the influence of outlier values.

Using ITSA provided a better way compared to segmented regression with individual patient data to visualise changes in prescribing over time. This was because each estimate on the scatterplot represented the mean duration or proportion of patients receiving over 3 months of treatment for that month. As for segmented regression, using ITSA it was possible to adjust for time specific trends that may have overestimated the effect of time on practice. Additional advantages to ITSA were the ability to plot the counterfactual situation in which the prescribing trends were unaffected by any change occurring around June 2017, which again provided a powerful visual presentation of the impact on prescribing that had occurred. The main disadvantage of using ITSA was that the data was aggregated into monthly groups, meaning that the granularity from the raw data was lost.

For the first four approaches, the use of two outcomes in the analysis were compared: treatment duration in weeks and the proportion of patients receiving over 3 months of treatment. On reflection, using a binary indicator gave a stronger indication of potential impact of the SCOT trial, given that if patients received over 3 months of chemotherapy, it was unlikely they were being prescribed therapy with the experimental arm of the SCOT trial in mind. Using a continuous variable of treatment duration was less relevant to the research objective, which was to assess SCOT impact.

The final approach to evaluate practice change was to calculate average doses of chemotherapy delivered. An advantage of this approach was that it allowed comparison with the results of the SCOT trial. The median dose of fluoropyrimidine chemotherapy delivered in the six-month arm of the SCOT trial was 83.2% (56.7-95.7%), and this was comparable to the pre-SCOT dose received

across all four regimens in this cohort. The median dose of oxaliplatin chemotherapy delivered in the six-month arm of the SCOT trial was 70.2% (44.3-87.1%), with the average dose of oxaliplatin delivered for patients commencing treatment with CAPOX or FOLFOX in GG&C being slightly less than this trial estimate. The results using median dose delivered pre versus post-SCOT supported the findings from the earlier approaches that there was the clearest change in practice for patients prescribed CAPOX chemotherapy at cycle one of treatment.

An important finding from the literature review in Chapter 3 was that attributing impact to a specific research study might be challenging. In an attempt to address this issue, the counterfactual situation, in which the SCOT trial did not occur, was assessed in this study in two ways. Firstly, a comparator group of patients who received monotherapy at cycle one were analysed to test if the timing of the dissemination of the SCOT trial findings had the same effect on practice compared to the cohort receiving doublet chemotherapy. There was no significant change in practice using any of the methodologies in this study for patients receiving monotherapy at cycle one of treatment. Secondly, within the ITSAs which modelled prescribing trends for patients, the pre-SCOT regression lines were extended to predict the average treatment duration or proportion of patients receiving over 3 months of treatment for each month post-SCOT if the pre-SCOT regression line had continued unchanged.

There are limitations to this study. Chemotherapy prescribing practice was only assessed in one health board. It is likely that clinicians will be influenced by the prescribing practices of their colleagues (see Chapter 5 survey results) and that different locations may have separate local prescribing guidelines. GG&C is also the health board where the original chief investigator of the SCOT trial was based and where the CTU that developed the trial is housed. It is possible that for these reasons, clinicians in GG&C were more likely than other locations to adopt the SCOT trial findings; the adoption of evidence and guidelines has been shown in previous studies to occur to a greater extent in the location where the research or guidelines were developed (308). It would therefore be useful to understand if the same prescribing practices changes identified in this study

occurred on a wider scale, for example at a regional, national (Scotland and/or UK) or international level.

Performing the ITSA analysis in particular would benefit from having a larger group of patients. When using ITSA, the number of data points available either side of the intervention and the number of observations at each time interval are both factors which affect the power of the analysis (306). Ideally, a minimum of nine data points pre and post intervention would be available and at least 100 observations per time interval is encouraged (309). Although this study met the first criteria, the second was not reached. Having a larger cohort to perform ITSA would also mean that separate ITSA for patient cohorts grouped by important co-variates could be explored.

The attribution of the practice change observed in this study to the SCOT trial was assessed using the timing of the SCOT trial abstract publication and an investigation of the counterfactual as described above. Despite the merits of this approach, it must be acknowledged that other events may have occurred at the time point chosen that are not described or that are unknown (for example a change in clinician staffing at GGC, individual clinician experience, change in management) which may be contributing to practice change. The IDEA collaboration, to which SCOT contributed, published their results at the same ASCO conference in June 2017. It is therefore impossible to distinguish the effect of these abstracts individually using quantitative methods, and a more qualitative approach would be required. Considering the results from the survey (Chapter 5) in which many UK clinicians indicated they were aware of both the SCOT trial and the IDEA collaboration, it is highly likely that clinicians were applying the findings from both of these studies when making clinical decisions after June 2017.

Finally, this cohort was defined by those receiving chemotherapy. There was therefore no information provided on patients who did not received adjuvant chemotherapy and no estimation of the proportion of patients diagnosed or undergoing major surgery that receive adjuvant treatment. This information would be useful to understand the reach of any practice change and if the impact of the SCOT trial could be increased by ensuring all patients who are eligible for adjuvant chemotherapy are getting access to treatment.

7.5 Conclusion

This chapter has analysed the impact of the SCOT trial findings on clinical practice and explored a variety of different methodologies for assessing this impact. A selection of these approaches will be used and developed in Chapter 8 for analysis of national level chemotherapy prescribing data. The lessons learned from this analysis would be equally applicable to assessing the impact of other clinical trials in future. Even if the outcome variable of interest (treatment duration or proportion of patients receiving over 3 months of treatment in this study) is not the same, the approach of using the time the results were published as a dichotomous variable, will be relevant.

8 Chapter 8: Impact on national practice

8.1 Introduction

The analysis in Chapter 7 gave important insights into the extent of chemotherapy prescribing practices in GG&C both before and after the SCOT trial results were initially disseminated. As described in the discussion section of Chapter 7, these results are not necessarily generalisable to a wider patient group of patients. In order to address this limitation, an analysis of prescribing practices at a national level was performed. This required a completely different approach to data acquisition due to the governance issues and logistics involved in accessing chemotherapy prescribing data on a national scale.

Out of the approaches to analysis tested in Chapter 7, three were used in this study: descriptive statistics, regression analysis, and ITSA. The regression analysis and ITSA were performed using a binary outcome variable only, rather than comparing the use of a binary and continuous outcome variable as was done in Chapter 7. The national dataset was also analysed without dividing patients into those who initially received doublet versus single agent chemotherapy. Instead, the regimen prescribed at cycle one was used as a co-variate in the regression analysis.

The aim of this study was to evaluate the implementation of SCOT trial results nationally in Scotland to provide a better idea of the scale and reach of trial impact. The objectives of this chapter were to:

i) Explore the feasibility of using administrative healthcare data to identify a cohort of patients who received adjuvant chemotherapy in Scotland and reflect on how this compares to using locally acquired data

ii) Describe the cohort of patients who receive adjuvant treatment, including their survival outcomes

iii) Utilise a selection of methods tested in Chapter 7 to analyse the impact of the SCOT trial findings on adjuvant clinical practice at a national level.

For the purposes of reporting the results of this analysis, the time-period prior to June 2017 is termed ‘pre-SCOT’ and after June 2017 is ‘post-SCOT’.

8.2 Methods

The process of gaining access to national level datasets is described in Chapter 2 and variables used in the final analysis on a national level are outlined in Appendix 5 Table 14-6. Additional information on patient co-morbidity (Charlson index) and the location where treatment was delivered (one of three cancer networks in Scotland) was available in the national compared to the local GG&C dataset, and these were added as co-variates in the regression analysis.

The descriptive analysis performed in this study was extended (compared to Chapter 7) to include an assessment of survival. This was to allow a more in depth comparison of this cohort compared to patients enrolled in the SCOT and IDEA trials. Three and 5-year overall and CRC cause specific survival was calculated using the Kaplan-Meier method and plotted graphically using a Kaplan-Meier curve. Survival was defined from the date of first chemotherapy cycle delivered and cause specific survival was defined by using death from CRC. The non-parametric log-rank test was used to test the difference in survival between different patient cohorts that had been divided into groups using categorical variables. A more in depth description of the regimens and duration of treatment delivered for this cohort was also performed. Patients were divided into groups dependent on age and disease stage to mimic the patient scenarios used in the first survey described in Chapter 5.

In this national study, logistic regression analysis was extended to include a subgroup analysis to understand the change in treatment duration after June 2017 for groups of patients with specific characteristics. In addition, patients who started chemotherapy after June 2017 only were analysed to identify if any patient characteristics were associated with a higher likelihood of a patient still receiving over 3 months of treatment after the results of the SCOT trial were disseminated.

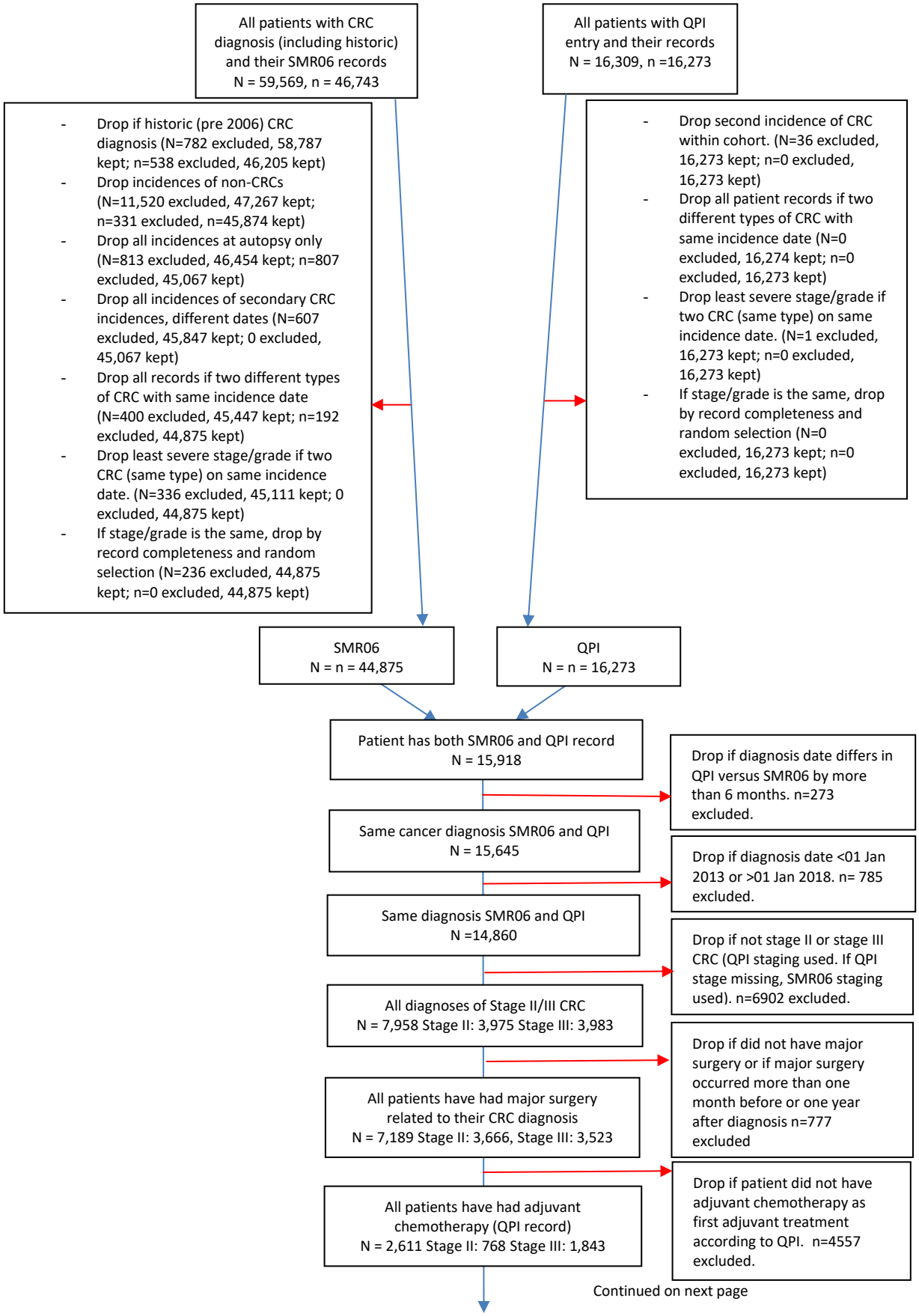
Finally, ITSA was performed for the overall patient cohort and separately for the cohort of patients considered SCOT eligible versus ineligible. Patients were

deemed “SCOT eligible” if they had stage II disease with high-risk features or stage III CRC and commenced treatment with fluoropyrimidine-oxaliplatin doublet chemotherapy within 11 weeks of major CRC surgery. High-risk features for stage II disease were defined as: T4 disease, extramural venous invasion, poor differentiation, less than 10 lymph nodes retrieved at operation, obstruction or perforation. There was no information available on perineural or lymphatic vascular invasion in this dataset, therefore these were not used for patient selection. In this national analysis, due to the increased number of patients for inclusion in the analysis compared to the local data, any months within the time series that contained data for less than five patients were excluded.

8.3 Results

8.3.1 Feasibility of using administrative healthcare data to define a national cohort

Figure 8-1 outlines the steps taken to create the final patient cohort for this analysis. The aim was to identify patients with CRC who received adjuvant chemotherapy after surgery in Scotland between 2013 and 2018.



Continued on next page

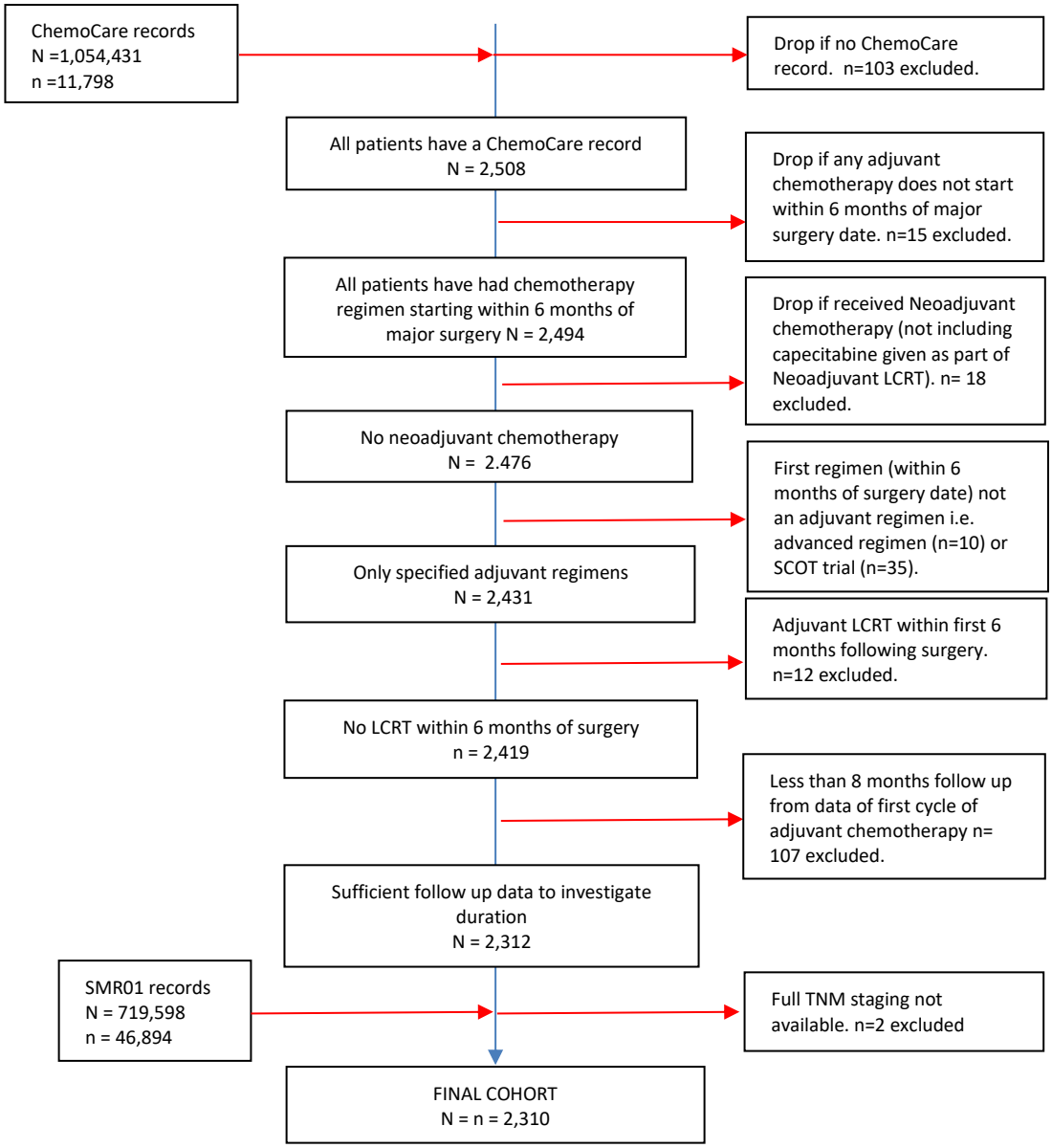


Figure 8-1 National cohort derivation using national, linked datasets

8.3.2 Description of overall cohort

Using this method of cohort derivation, between January 2013 and January 2018, 7,958 patients in Scotland were diagnosed with stage II/III CRC (stage II 3,975 (50%) stage III 3,983 (50%)). In total, 7,189 (90%) of these patients underwent major CRC surgery within a year of their diagnosis (stage II 3,666 (51%), stage III 3,523 (49%)). Of those patients diagnosed with CRC, 2,611 (33%) received adjuvant chemotherapy (19% (n=768) of stage II patients and 46% (n=1,843) of patients diagnosed with stage III). This represented 36% of those who had

undergone major surgery (stage II: 21%, stage III: 52%). In total, 2,101 (57%) of patients with stage II disease who had major surgery within a year of diagnosis were identified as fitting into the high-risk category and 667 (87%) of stage II patients who received adjuvant chemotherapy fitted into the high-risk category. Table 8-1 shows the proportion of patients with stage II disease that had high-risk features for each of these cohorts. As described above, this table does not include information on lymphovascular invasion or obstruction/perforation because this information was not available within this dataset.

Table 8-1 High-risk disease features for patients with stage II disease

	Number	Percentage
Patients with stage II CRC who had major surgery	3,666	100%
Any high-risk feature	2101	57%
Of patients with any high-risk feature (n=2101):		
Extramural venous invasion	1355	64%
Lymph node yield <10	176	8%
Poor differentiation	507	24%
T4 disease	882	42%
Patients with stage II CRC who had major surgery and chemotherapy	768	100%
Any high-risk feature	667	87%
Of patients with any high-risk feature (n=667)		
Extramural venous invasion	474	71%
Lymph node yield <10	29	4%
Poor differentiation	135	20%
T4 disease	637	96%

8.3.3 Practice change cohort

The final cohort for the purposes of analysing the impact of the SCOT trial on prescribing consisted of 2,310 patients (Figure 8-1). Table 8-2 describes the characteristics of this cohort and compares patients treated with adjuvant chemotherapy pre-June 2017 and post-June 2017.

Three quarters of patients were aged 70 and under and there was a slight male preponderance. A majority of patients had stage III CRC and patients were more likely to have colon rather than rectal cancer. As was the case for local/GG&C prescribing (Chapter 7), there was a clear preference for using capecitabine-based regimens compared to those based on using intravenous 5-fluorouracil.

Almost half of patients were from the West of Scotland (WoSCAN), with the remainder split almost evenly between the South-East (SCAN) and North (NoSCAN). Of note, the proportion of patients included from SCAN after June 2017 was less compared to pre-June 2017 because granular chemotherapy prescribing records post-SCOT were available for a shorter time-period compared to in NoSCAN and WoSCAN. There were 452 patients treated in the GG&C health board during this period (first chemotherapy dose from June 2013 to May 2018). The cohort of patients list in the local analysis identified from ChemoCare (Chapter 7) as having treatment during the same period was 528.

Table 8-2 Cohort characteristics

	Pre-SCOT (Number (Percentage))	Post-SCOT (Number (Percentage))	Total (Number (Percentage))
Number	1957 (85%)	353 (15%)	2310 (100%)
Age			
Median age (IQR)	65 (57-71)	65 (57-71)	64 (55-71)
Age groups			
70 and under	1466 (75%)	261 (74%)	1727 (75%)
Over 70	491 (25%)	92 (26%)	583 (25%)
Gender			
Male	1034 (53%)	203 (58%)	1237 (54%)
Female	923 (47%)	150 (42%)	1073 (46%)
Location			
SCAN	489 (25%)	37 (10%)	526 (23%)
WoSCAN	966 (49%)	197 (56%)	1163 (50%)
NoSCAN	502 (26%)	119 (34%)	621 (27%)
Stage			
II	593 (30%)	88 (25%)	681 (29%)
III	1364 (70%)	265 (75%)	1629 (71%)
Risk stage			
II	593 (30%)	88 (25%)	681 (29%)
Low-risk III	626 (32%)	122 (35%)	748 (32%)
High-risk III	738 (38%)	143 (41%)	881 (38%)
Regimen			
CAPOX/RALOX	1023 (52%)	177 (50%)	1200 (52%)
FOLFOX	164 (8%)	54 (15%)	218 (9%)
Cap alone	709 (36%)	110 (31%)	819 (35%)
FU alone	61 (3%)	12 (3%)	73 (3%)
Site			
Colon	1530 (78%)	276 (78%)	1806 (78%)
Rectum	427 (22%)	77 (22%)	504 (22%)
SIMD			
1	343 (18%)	59 (17%)	402 (17%)
2	368 (19%)	70 (20%)	438 (29%)
3	388 (20%)	90 (25%)	478 (21%)

4 5	399 (20%) 459 (23%)	64 (18%) 70 (20%)	463 (20%) 529 (23%)
Charlson score			
Mean	0.54 (SD 0.92)	0.38 (0.77)	0.1 (0.9)

Figure 8-2 shows the Kaplan-Meier curve for OS from date of first chemotherapy for the whole cohort. There were 370 deaths from any cause with a median follow up for 3.68 years (95%CI 3.59-3.77 years). Separate Kaplan-Meier curves showing the difference in OS for subgroups are included in Appendix 5.

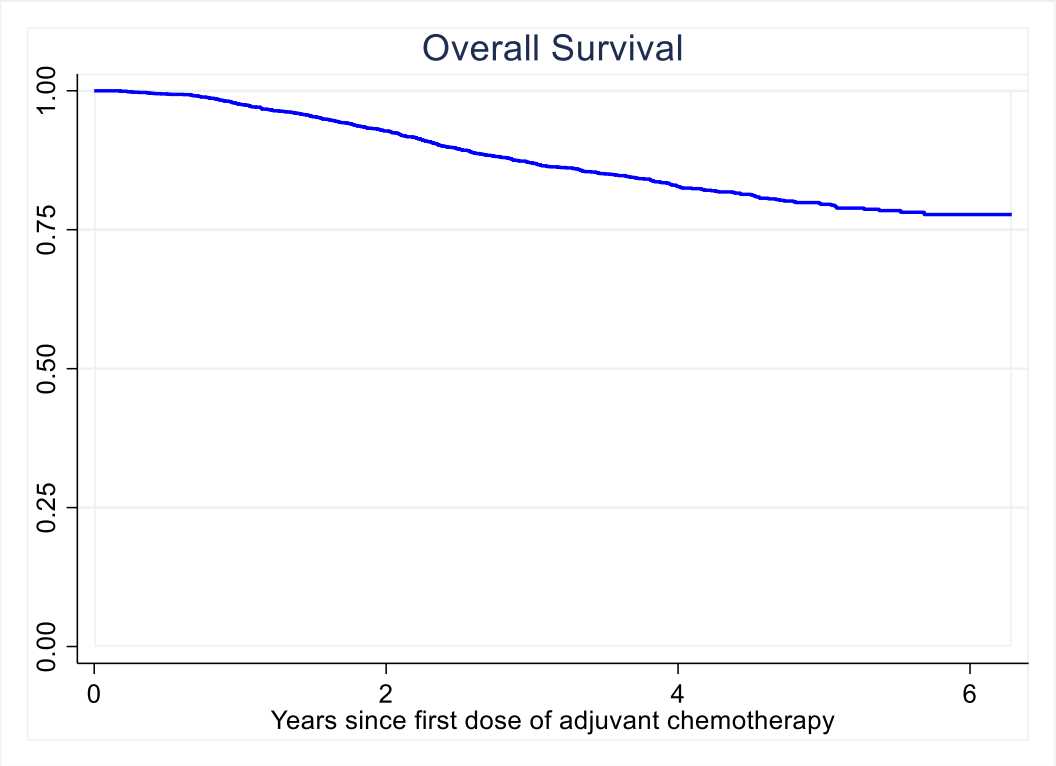


Figure 8-2 Kaplan-Meier curve showing OS from date of first chemotherapy cycle

Figure 8-3 shows the Kaplan-Meier curve for CRC survival. There were 307 deaths from CRC in this cohort with a median follow up time of 3.62 years (95%CI 3.50-3.70 years). Separate Kaplan-Meier curves showing the difference in CRC specific survival for subgroups are included in Appendix 5. Median survival was not reached for either OS or CRC outcomes. Table 14-7 in Appendix 5 lists 3 and 5-year OS and CRC survival for the whole cohort and for patient subgroups divided by patient, disease, and treatment characteristics.

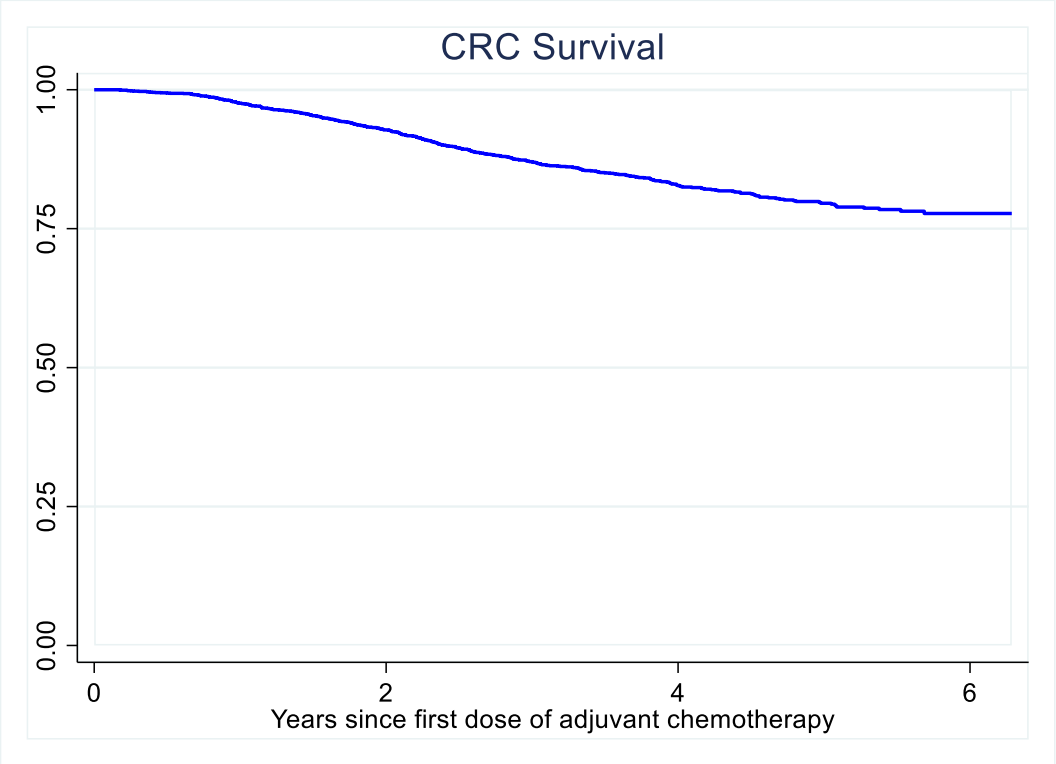


Figure 8-3 Kaplan-Meier curve showing cancer specific survival (CRC) from date of first chemotherapy

8.3.4 A description of practice change

There was a significant decrease in the proportion of patients receiving over 3 months of treatment post-SCOT (75%) versus pre-SCOT (42%, χ^2 p<0.001) (Figure 8-4).

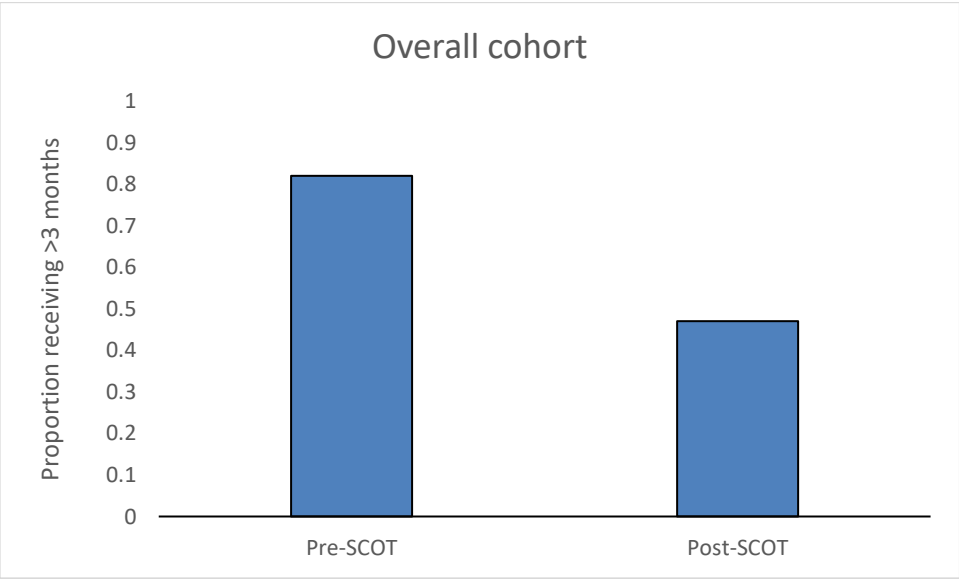


Figure 8-4 Change in proportion of patients receiving over 3 months of treatment pre- versus post-SCOT.

The median duration of treatment for the whole cohort was 18 weeks (IQR 12-24). Prior to June 2017 median treatment duration was 21 weeks (IQR 14-24 weeks) and this decreased to 12 weeks (IQR 12-21weeks) for patients starting adjuvant chemotherapy after June 2017 (Mann-Whitney U test p<0.001). Figure 8-5 demonstrates the median duration of treatment received per year.

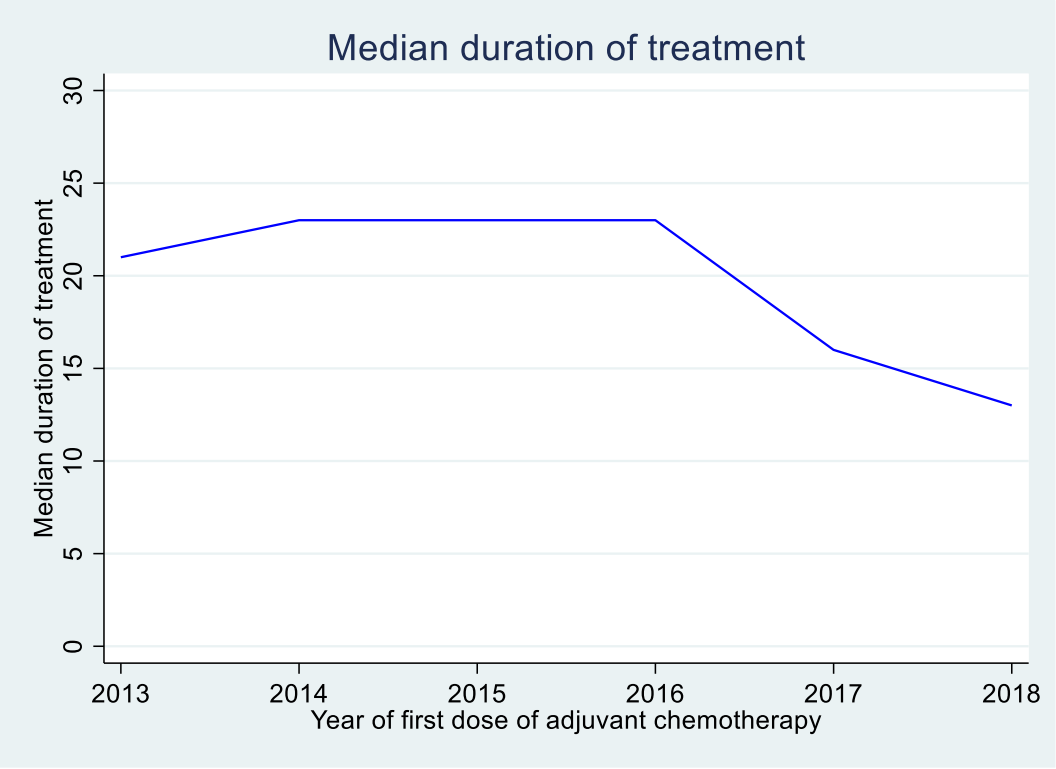


Figure 8-5 Median duration of treatment per year from 2013-2018 Treatment duration calculated in weeks using cycles of treatment. Y-axis label (weeks).

For patients aged under 70 years old, prior to June 2017, over 3 months of CAPOX was a clear preference for patients with stage III disease, and using capecitabine for over 3 months was utilised most often for patients with stage II disease (Figure 8-6). After June 2017, the use of 3 months or less of CAPOX increased across all scenarios, in particular for patients with low-risk stage III disease. Patients with high-risk stage III disease and with N2 disease as the only high-risk factor were more likely to receive the shorter duration of CAPOX (25%) compared to those with T4 disease as the only high-risk factor (T4N1 14%).

Regarding patients aged 70 and over, there was a higher use of capecitabine monotherapy across all scenarios compared to treatment of patients in the younger age group (Figure 8-6). Post-SCOT, again, as for younger patients, a rise in the use of CAPOX for 3 months or less was observed, in particular for those

with stage III disease, although numbers in this sub-group were small (<20) so these results should be interpreted with caution.

Table 8-3 shows how these results compare with the findings from the first survey reported in Chapter 5.

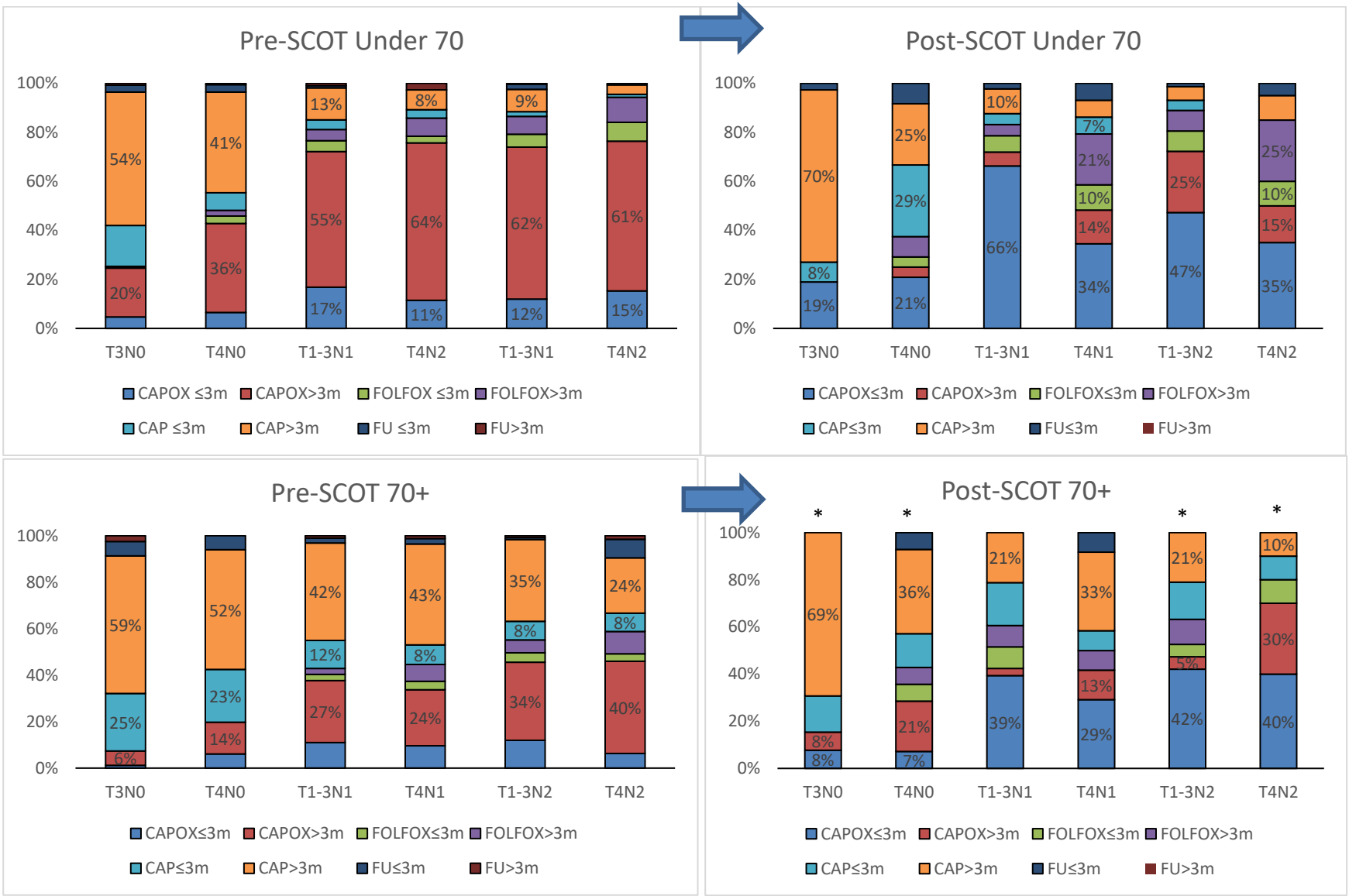

Pre-SCOT Under 70
100%
80%
60%
40%
20%
0%
54%
20%
41%
36%
13%
55%
8%
64%
9%
62%
61%
17%
11%
12%
15%
T3N0
T4N0
T1-3N1
T4N2
T1-3N1
T4N2
CAPOX ≤3m
CAPOX>3m
FOLFOX ≤3m
FOLFOX>3m
CAP ≤3m
CAP>3m
FU ≤3m
FU>3m
Post-SCOT Under 70
100%
80%
60%
40%
20%
0%
70%
8%
19%
25%
29%
21%
10%
66%
7%
21%
10%
14%
34%
25%
47%
25%
10%
15%
35%
T3N0
T4N0
T1-3N1
T4N1
T1-3N2
T4N2
CAPOX≤3m
CAPOX>3m
FOLFOX≤3m
FOLFOX>3m
CAP≤3m
CAP>3m
FU≤3m
FU>3m
Pre-SCOT 70+
100%
80%
60%
40%
20%
0%
59%
25%
6%
52%
23%
14%
42%
12%
27%
43%
8%
24%
35%
8%
34%
24%
8%
40%
T3N0
T4N0
T1-3N1
T4N1
T1-3N2
T4N2
CAPOX≤3m
CAPOX>3m
FOLFOX≤3m
FOLFOX>3m
CAP≤3m
CAP>3m
FU≤3m
FU>3m
Post-SCOT 70+
*
*
*
*
100%
80%
60%
40%
20%
0%
69%
8%
8%
36%
21%
7%
21%
39%
33%
13%
29%
21%
5%
42%
10%
30%
40%
T3N0
T4N0
T1-3N1
T4N1
T1-3N2
T4N2
CAPOX≤3m
CAPOX>3m
FOLFOX≤3m
FOLFOX>3m
CAP≤3m
CAP>3m
FU≤3m
FU>3m

Figure 8-6 Treatment regimen and duration of chemotherapy delivered pre- versus post-SCOT

Table 8-3 Proportion of patients receiving 3 months or less of doublet chemotherapy post-SCOT Results from national administrative dataset compared to the first survey results (Chapter 5).

	Three months of doublet chemotherapy			
	<70 years		≥70 years	
	Survey	National data	Survey	National data
Low-risk stage III disease	86%	73%	54%	48%
High-risk stage III disease	16%	48%	15%	42%
T4N1	18%	45%	17%	29%
T1-3N2	19%	56%	17%	47%
T4N2	11%	45%	13%	50%
Stage II* overall	16%	22%	9%	11%
Stage II* MSS	20%		8%	
Stage II* MSI-H	12%		10%	
T4N0		19%		8%
T3N0		25%		14%

8.3.5 Logistic regression: overall cohort

On univariate analysis, there was a decrease of 76% in the odds of receiving over 3 months of treatment in the post-SCOT period compared to before the results of this trial were known (unadjusted OR 0.24, 95% CI 0.19-0.30, p<0.001).

Looking across the whole study period, the chemotherapy regimen used at cycle one, socio-economic status, location of treatment and risk stage significantly affected the likelihood of a patients receiving more than 3 months of treatment, regardless of when the SCOT trial results were disseminated (Table 8-4). Specifically, patients with low-risk stage III disease were less likely overall to receive over 3 months of treatment compared to those with stage II disease (OR 0.70 (0.56-0.88)). Individuals receiving FOLFOX (OR 0.52, 95%CI 0.39-0.70) and those receiving 5-fluorouarcil (OR 0.12 (0.07-0.21)) were less likely to receive over 3 months of treatment compared to patients treated with CAPOX, whereas patients treated with capecitabine monotherapy were more likely to receive over 3 months (OR 1.27, 95% CI 1.03-1.56). Patients treated in the North of Scotland were less likely to receive over 3 months of treatment compared to the South East (OR 0.45, 95% CI 0.35-0.58). Lastly, patients in some of the less socio-economically deprived areas (SIMD 3 (OR 0.64, 95%CI 0.48-0.86) and 4 (OR 0.73, 95% CI 0.54-0.98) were less likely to receive over 3 months of treatment compared to SIMD group 1 (most deprived).

As was observed in the local analysis in Chapter 7, adjustment for all of the included variables on multivariate logistic regression did not markedly affect the odds of a patient receiving longer treatment post-SCOT, which was significant (Table 8-4). Assessment for interactions between co-variates and the timing of the SCOT trial showed there was a significant interaction with the SCOT trial variable and regimen, risk stage and co-morbidity. These were investigated further as part of a subgroup analysis.

Table 8-4 Univariate and multivariate analysis of factors affecting the likelihood of patients receiving over 3 months of treatment Abbreviations: OR, odds ratio; Ref, reference category; SCOT, Short Course Oncology Treatment; WoS, West of Scotland; NoS, North of Scotland; SCAN, South East Scotland; CAPOX, capecitabine and oxaliplatin chemotherapy; FOLFOX, 5-fluorouracil and oxaliplatin chemotherapy; SIMD, Scottish Index of Multiple Deprivation.

		Unadjusted OR	Lower 95% CI	Upper 95% CI	Adjusted OR	Lower 95% CI	Upper 95% CI
SCOT	Pre-SCOT	Ref					
	Post-SCOT	0.240135	0.189879	0.303692	0.22594	0.175847	0.290304
Age group	<70	Ref					
	70+	0.871744	0.711503	1.068073	0.815386	0.645342	1.030236
Sex	Male	Ref					
	Female	0.884206	0.739303	1.057511	0.8007	0.65867	0.973357
Risk stage	Stage II	Ref					
	LRIII	0.698989	0.557327	0.876659	0.800292	0.61309	1.044656
	HRIII	0.988799	0.789606	1.238243	1.289375	0.98007	1.696295
Location	SCAN	Ref					
	WoS	1.120699	0.883893	1.420947	1.46312	1.126871	1.899704
	NoS	0.450374	0.35023	0.579153	0.635802	0.481086	0.840274
Regimen	CAPOX						
	FOLFOX	0.519386	0.386298	0.698327	0.557514	0.400948	0.775218
	Capecitabine	1.268454	1.03406	1.55598	1.371677	1.068775	1.760424
	5-Fluorouracil	0.119525	0.068472	0.208642	0.143764	0.079551	0.259807
Disease site	Colon						
	Rectum	0.77052	0.624019	0.951414	0.795	0.629358	1.004237
SIMD	1						
	2	0.840213	0.618554	1.141302	1.026409	0.734774	1.433796
	3	0.642452	0.47879	0.862058	0.930591	0.668743	1.294965
	4	0.728553	0.540424	0.982174	1.033795	0.738428	1.447308
	5	0.871969	0.649252	1.171085	1.083815	0.780519	1.504966
Charlson score	0						
	1	1.051083	0.845609	1.306486	0.97436	0.768825	1.23484
	>1	1.04336	0.779296	1.3969	0.896923	0.649357	1.238871

8.3.5.1 Subgroup analysis

The percentage of patients receiving over 3 months of treatment prior to and after June 2017 for each subgroup is outlined in Table 8-5.

Table 8-5 The proportion of patients receiving over 3 months of treatment pre versus post-SCOT dependent on patient, disease, and treatment related factors Abbreviations: LR, low-risk; HR, high-risk; SCAN, South East Scotland; WoS, West of Scotland; NoS, North of Scotland; CAPOX, capecitabine and oxaliplatin; FOLFOX, 5-fluorouracil and oxaliplatin; CAP, capecitabine; FU, 5-fluorouracil; SIMD, Scottish Index of Multiple Deprivation.

	Pre-SCOT	Post-SCOT
	>3 months of chemotherapy	>3 months of chemotherapy
Age group		
Under 70	77%	40%
70 and over	72%	49%
Gender		
Male	77%	43%
Female	73%	42%
Risk stage		
II	75%	61%
LR3	73%	24%
HR3	78%	47%
Location		
SCAN	78%	27%
WoS	82%	47%
Nos	60%	39%
Regimen		
CAPOX	80%	23%
FOLFOX	57%	56%
CAP	77%	72%
FU	28%	0%
Disease Site		
Colon	76%	46%
Rectum	73%	30%
SIMD		
1	80%	44%
2	77%	44%
3	71%	41%
4	74%	36%
5	76%	47%
Charlson index groups		
0	77%	38%
1	73%	56%
>1	73%	50%

Multi-variate logistic regression analysis was performed for each sub-group separately to understand the effect of the SCOT trial whilst accounting for other variables. Table 8-6 demonstrates the adjusted odds of getting over 3 months of treatment after June 2017 compared to pre-June 2017 for patients in each subgroup.

Table 8-6 Subgroup analysis Abbreviations: LR, low-risk; HR, high-risk; SCAN, South East Scotland; WoS, West of Scotland; NoS, North of Scotland; CAPOX, capecitabine and oxaliplatin; FOLFOX, 5-fluorouracil and oxaliplatin; SIMD, Scottish Index of Multiple Deprivation.

		Unadjusted SCOT OR	Lower 95% CI	Upper 95% CI	Adjusted SCOT OR	Lower 95% CI	Upper 95% CI
Risk stage	Stage II	0.5378	0.3370	0.8582	0.5405	0.3262	0.8956
	LR III	0.1125	0.0715	0.1770	0.0883	0.0540	0.1444
	HR III	0.2519	0.1737	0.3652	0.2540	0.1703	0.3788
Age group	Under 70	0.2048	0.1555	0.2697	0.1853	0.1377	0.2494
	70+	0.3743	0.2378	0.5892	0.3542	0.2173	0.5772
Sex	Male	0.2206	0.1612	0.3019	0.2022	0.1437	0.2845
	Female	0.2617	0.1834	0.3734	0.2541	0.1745	0.3701
Regimen	CAPOX	0.0746	0.0510	0.1093	0.0656	0.0438	0.0984
	FOLFOX	0.9309	0.5010	1.7296	0.9435	0.4186	2.1269
	Capecitabine	0.7608	0.4847	1.1940	0.7009	0.4397	1.1174
	5-Fluorouracil	1.0000					
Location	SCAN	0.1062	0.0499	0.2263	0.0822	0.0365	0.1851
	WoS	0.1937	0.1400	0.2680	0.1903	0.1362	0.2660
	NoS	0.4287	0.2849	0.6452	0.3531	0.2174	0.5734
Disease site	Colon	0.2651	0.2035	0.3454	0.2554	0.1932	0.3376
	Rectum	0.1608	0.0944	0.2737	0.1060	0.0568	0.1981
SIMD	1.0000	0.1948	0.1092	0.3474	0.2097	0.1131	0.3888
	2.0000	0.2424	0.1427	0.4118	0.2093	0.1177	0.3721
	3.0000	0.2797	0.1741	0.4494	0.2702	0.1608	0.4540
	4.0000	0.2003	0.1148	0.3497	0.1586	0.0863	0.2914
	5.0000	0.2811	0.1678	0.4709	0.2808	0.1587	0.4969
Co-morbidity	Charlson 0	0.1891	0.1427	0.2506	0.1768	0.1308	0.2390
	Charlson 1	0.4653	0.2745	0.7888	0.4416	0.2501	0.7796
	Charlson >1	0.3631	0.1595	0.8266	0.2696	0.1048	0.6934

For those patients who received CAPOX, on unadjusted and adjusted regression analysis there was a large and significant change in the proportion of patients receiving over 3 months of treatment post-SCOT versus pre-SCOT whereas no significant change was seen for patients receiving FOLFOX or capecitabine (unadjusted). On adjusted analysis, the change in capecitabine prescribing just

met significance at the 0.05 level. No patients received over 3 months of 5-fluorouarcil in the post-SCOT period.

Practice change post-SCOT was significant regardless of disease type (colon versus rectum), sex, socio-economic status and co-morbidity. It was also significant for all locations and both age groups, although it was less marked for patients living in the North of Scotland and for patients aged over 70 years. In both instances, this was because pre-SCOT, there was already a large proportion of patients treated with 3 months or less of adjuvant chemotherapy (North of Scotland (40%), aged over 70 (28%)). Regarding risk stage, on adjusted analysis, there was a significant change for patients with low-risk stage III and high-risk stage III.

8.3.5.2 Adjuvant treatment post-SCOT only

Focusing only on patients starting chemotherapy after June 2017 (n=353), results from a multivariate model demonstrated that patients receiving FOLFOX (OR 3.82 95% CI 1.90-7.70) or capecitabine (OR 8.33 95% CI 4.31-16.09) were more likely to still receive over 3 months of treatment compared to those receiving CAPOX post-SCOT. This aligns with the post-SCOT treatment choices shown in Figure 8-6. Patients with low-risk stage III disease treated after June 2017 were significantly less likely to receive the longer duration of treatment compared to those with stage II disease (OR 0.37 95% CI 0.18-0.78). Patients with a Charlson Index of 1 (higher level of comorbidity) were more likely to have over 3 months of treatment compared to those with an index of zero (OR 2.26 95% CI 1.15-4.44).

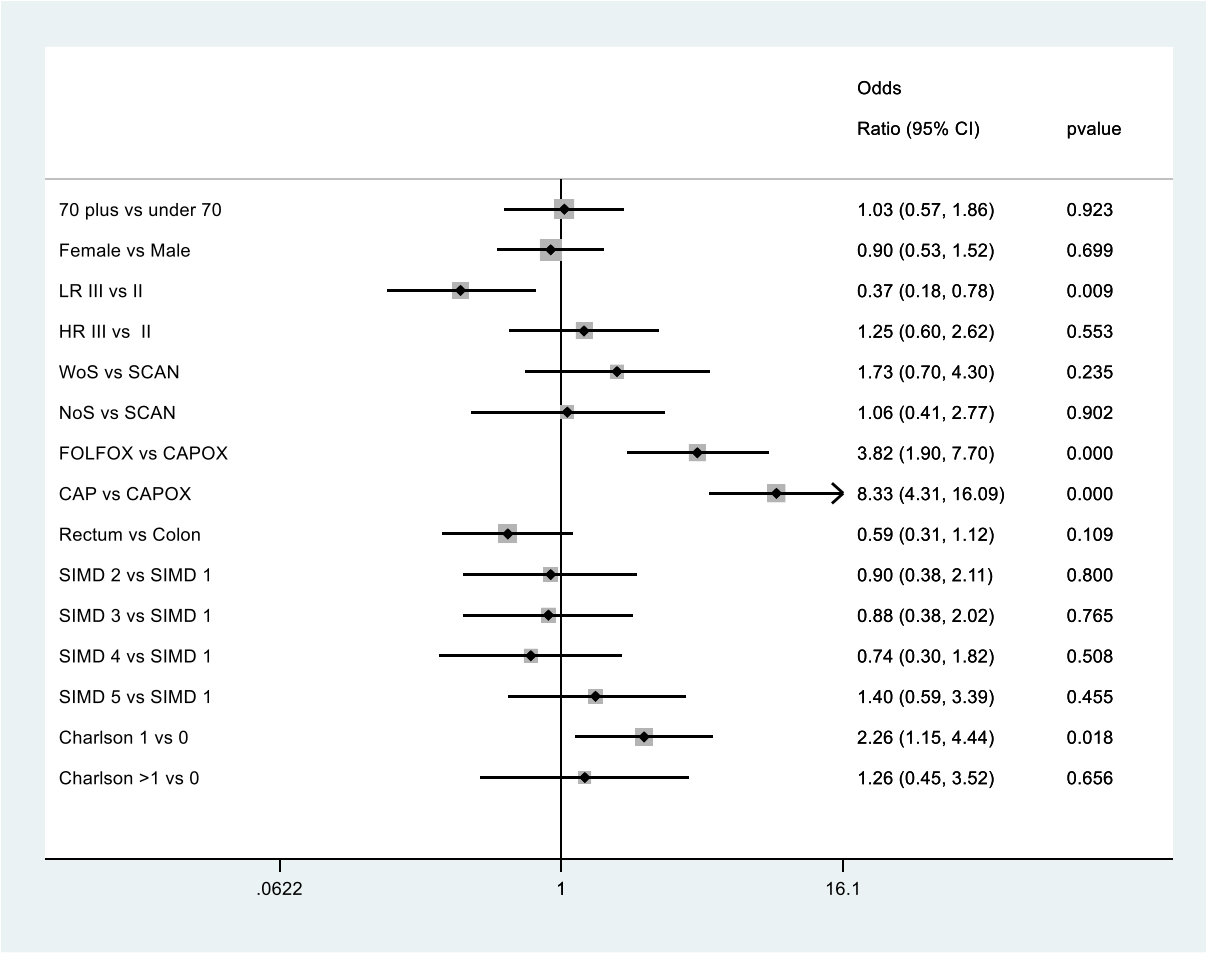


Figure 8-7 Adjusted odds of receiving over 3 months of treatment for patients starting adjuvant treatment after June 2017 Abbreviations: LR, low-risk; HR, high-risk; WoS, West of Scotland; NoS, North of Scotland; SCAN, South-East Scotland; FOLFOX, 5-fluorouracil and oxaliplatin; CAPOX, capecitabine and oxaliplatin; CAP, capecitabine; SIMD, Scottish Index of Multiple Deprivation; vs, versus.

8.3.6 Interrupted time-series analysis

There was a clear decrease in the monthly proportion of patients receiving over 3 months of treatment after June 2017 compared to prior to this time-point (Figure 8-8). The estimated proportion of patients receiving over 3 months of treatment at the end of the study period (chemotherapy start date in April 2018) was 34% (95% CI 28-40%) based on actual prescribing data, compared to 67% (95% CI 57-76%) based on extrapolation of the trend in the treatment duration pre-SCOT (absolute reduction in percentage by 33%).

There was a significant decrease in the proportion of patients receiving over 3 months of treatment immediately after June 2017, compared to immediately before this time point (a level difference of 0.41 (0.27-0.61, $p<0.001$)). There was also a significant change in the trends in prescribing over time, pre versus

post-SCOT (difference in slopes of graphs pre versus post-SCOT: 0.95 (0.91-0.99, p=0.024).

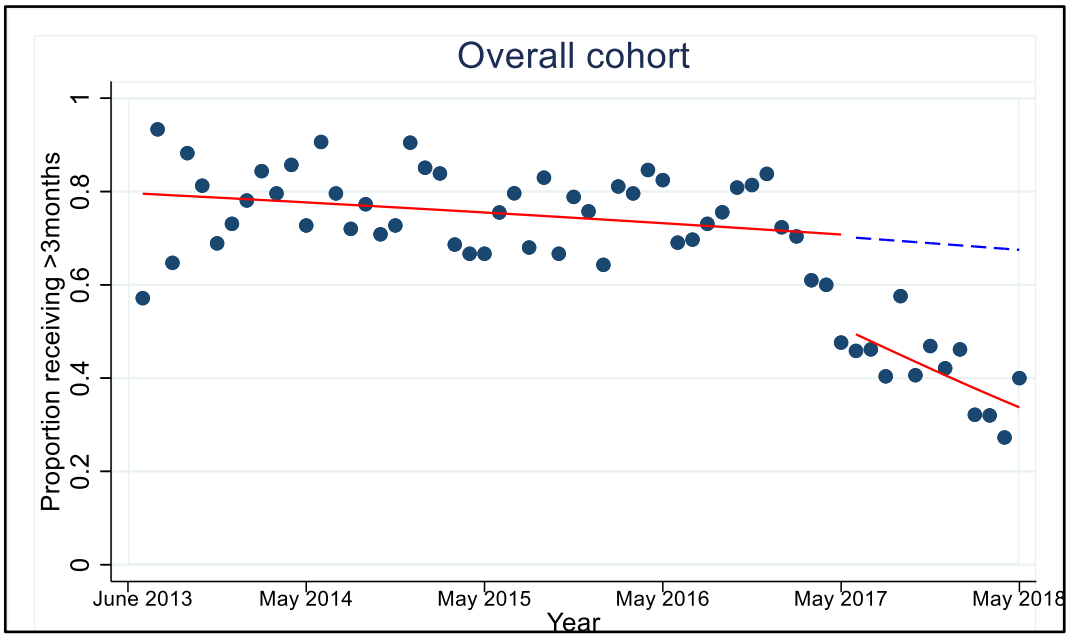


Figure 8-8 Change in monthly proportion of patients receiving over 3 months of treatment (overall cohort) The blue dots in the diagram indicate the monthly proportion of patients receiving over 3 months of chemotherapy. The red lines show the predicted trend in the proportion of patients receiving over 3 months of chemotherapy based on actual prescribing data. The dashed blue line indicates the predicted proportion of patients receiving over 3 months of chemotherapy if the pre-SCOT trend line (red) is continued uninterrupted (counterfactual situation). The month signifies the date which a patient started chemotherapy. Overall, there was an absolute decrease in percentage of 33% for the proportion of patients receiving over 3 months of treatment at the end of the study period from the actual versus counterfactual trend line.

Practice change was driven by changes in treatment for patients who met the SCOT trial criteria. For these SCOT eligible patients (Figure 8-9), the difference between the proportion of patients receiving over 3 months of treatment using actual data (24%, 95%CI 17-31%) versus the counterfactual situation (75%, 95% CI 63-84%) was 51%**.** There was a large and significant difference in the proportion of patients receiving over 3 months of treatment immediately after SCOT compared to before June 2017 (OR 0.20 (0.12-0.35, p<0.001)). There was also a significant change in the trend in prescribing over time post-SCOT versus pre-SCOT (change in the slope of the post-SCOT versus pre-SCOT graphs: 0.93 (0.87-1.00, p=0.0038)).

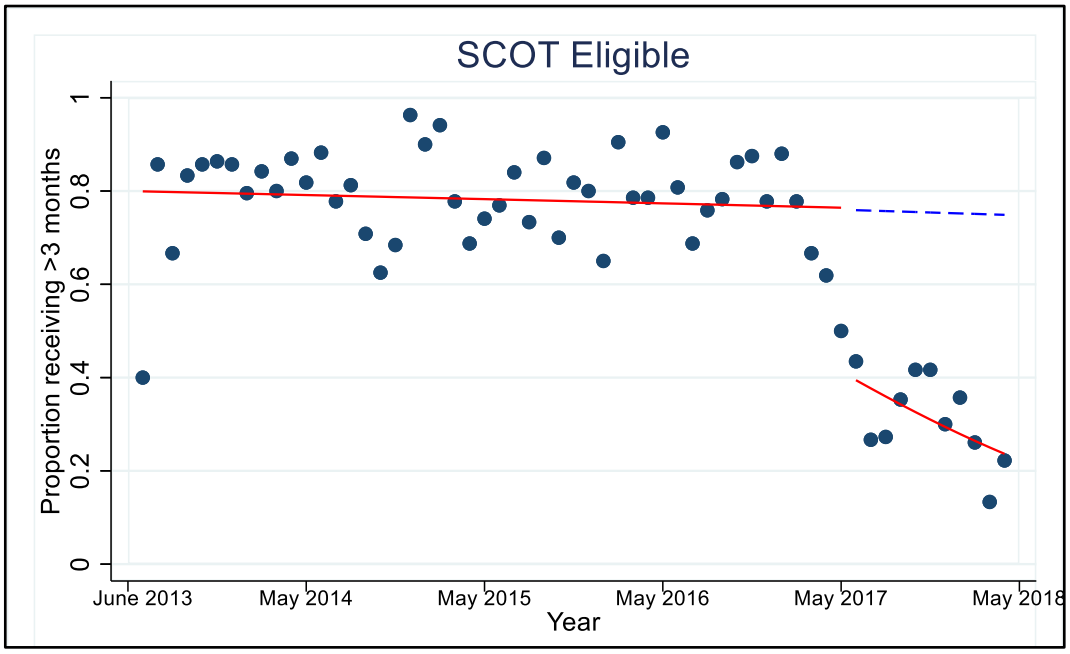


Figure 8-9 Change in monthly proportion of patients receiving over 3 months of treatment (SCOT eligible) The blue dots in the diagram indicate the monthly proportion of patients receiving over 3 months of chemotherapy. The red lines show the predicted trend in the proportion of patients receiving over 3 months of chemotherapy based on actual prescribing data. The dashed blue line indicates the predicted proportion of patients receiving over 3 months of chemotherapy if the pre-SCOT trend line (red) is continued uninterrupted (counterfactual situation). The month signifies the date which a patient started chemotherapy. Overall, absolute decrease in percentage of 51%for the proportion of patients receiving over 3 months of treatment at the end of the study period for the trend line based on actual data versus the counterfactual. The pre-SCOT time trend was 1.00 (0.98-1.01, p=0.615), post SCOT was 0.93 (.87-0.99, p=0.615).

For SCOT ineligible patients (Figure 8-10), the predicted proportion of patients receiving over 3 months of treatment at the end of the study period (June 2018) using actual prescribing data was 58% (95% CI 50-65%) compared to a counterfactual estimate of 61% (95% CI 49-71%) (3% difference). There was no significant difference in the slope or level change of the prescribing trends across the June 2017 time point (difference in slope 0.99 (0.91-1.08, p=0.871); level difference of 0.91 (0.44-1.86, p=0.787)).

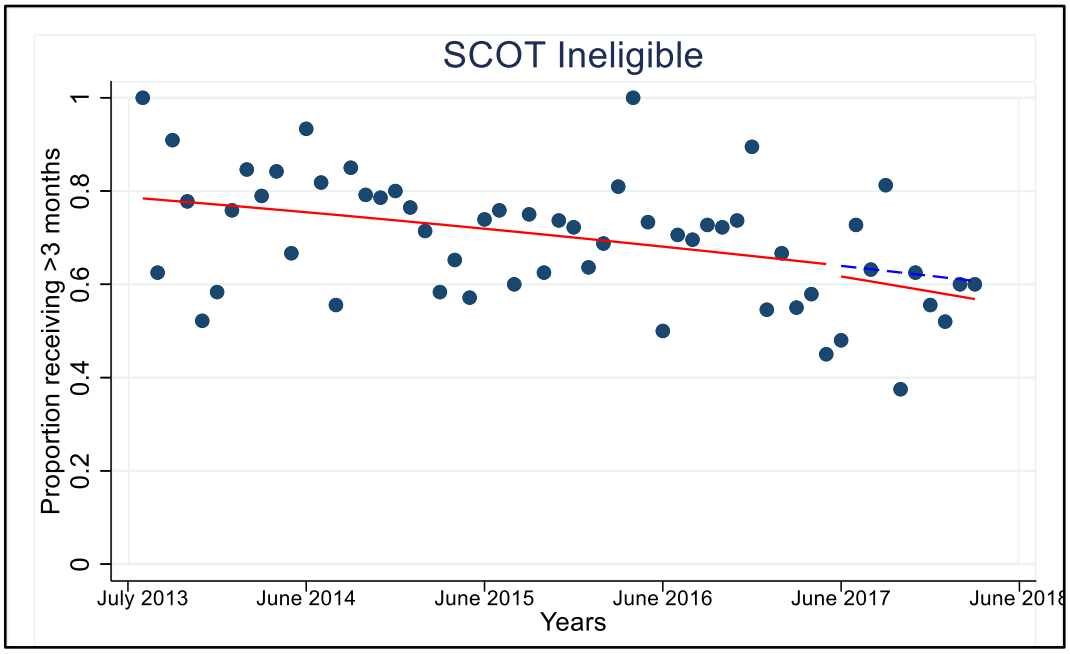


Figure 8-10 Change in monthly proportion of patients receiving over 3 months of treatment (SCOT ineligible) The blue dots in the diagram indicate the monthly proportion of patients receiving over 3 months of chemotherapy. The red lines show the predicted trend in the proportion of patients receiving over 3 months of chemotherapy based on actual prescribing data. The dashed blue line indicates the predicted proportion of patients receiving over 3 months of chemotherapy if the pre-SCOT trend line (red) is continued uninterrupted (counterfactual situation). The month signifies the date which a patient started chemotherapy. 3% change (increase) in proportion of patients receiving over 3 months of treatment at the end of the study period for the actual versus counterfactual trend line. The pre-SCOT time trend was 0.98 (0.97-1.00, p=0.044), post SCOT was 0.98 (.89-1.07, p=0.621)

8.4 Discussion

This study shows that it was feasible to identify a cohort of patients with CRC who were treated with adjuvant chemotherapy in Scotland using administrative healthcare data. The demographics, disease, and treatment related factors were broadly similar to the characteristics of the GG&C cohort (2010-2019). CAPOX was the preferred doublet regimen used by clinician for this Scottish population, as was the case in the GG&C study and in the SCOT trial. The main difference between the patients in the GG&C (Chapter 7) versus this national level study was that there was a higher percentage (33%) of patients living in the most deprived locations (SIMD group 1) in GG&C compared to the national level (17%).

In this study 17% of patients had a diagnosis of rectal cancer; this compared to 18% in the SCOT trial (41) and 25% of patients in the GG&C analysis. The discrepancy in GG&C results is likely to be in a large part due to the fact that rectosigmoid tumours were classified as rectal rather than colon tumours in the GG&C analysis, rather than a reflection of a truly higher incidence of rectal

cancer in that location. SCOT was the only trial that contributed to the IDEA collaboration that recruited patients with rectal cancer and there is less consensus regarding the use of adjuvant chemotherapy for rectal cancer generally (Chapter 1). Therefore, although this study has demonstrated a significant reduction in the proportion of patients receiving over 3 months of chemotherapy post-SCOT, the underlying assumption that 6 months is an accepted standard of care in this setting is less robust and any interpretation of practice or practice change associated with implementation of trial findings must be interpreted in this context.

The 5-year OS for this cohort (Appendix 5) compared favourably to the survival reported in the IDEA collaboration (82.4% for 3 month arm and 82.8% for 6 month arm). The limitation of this comparison is that a minority of patients in this study had sufficient follow up to provide a reliable estimate of 5-year OS. Unfortunately, the result from this study could not be compared to 5-year OS specifically from the SCOT trial because the longer-term results from SCOT alone have not yet been published.

This analysis demonstrated that the duration of adjuvant chemotherapy delivered in Scotland changed significantly and rapidly within months of the dissemination of the SCOT trial findings. The extent of change was similar to results for patients from GG&C, although in GG&C the proportion of patients receiving over 3 months of chemotherapy in the pre-SCOT period was higher (82% versus 75%). This study has shown that population level data can provide important insights into real world treatment choices and allow analysis of how, and when, clinical trial evidence is implemented into practice. The population level analysis widens the generalisability of these results compared to the local study of prescribing practices that was performed (Chapter 7). This was especially important to explore given that the trial in question was developed by clinicians working in the local GG&C location.

When practice change did occur, the biggest change in practice was relevant to patients who met the SCOT trial inclusion criteria. This implies that for the SCOT trial (and IDEA collaboration), the findings of these trials were probably not being applied too far beyond patients resembling the original trial populations, justifying the approach to budget impact analysis used in Chapter 6 which only

considered practice change for patients receiving doublet chemotherapy. The change in practice was seen most clearly for patients receiving CAPOX, whereas in contrast, patients receiving FOLFOX were more likely to still receive over 3 months of treatment post-SCOT. A change in practice particularly affecting patients prescribed CAPOX aligns with the results from the pre-planned subgroup analyses from SCOT and IDEA, which both showed that non-inferiority was met when CAPOX was used but not for FOLFOX. Practice change was also highest for patients with low-risk stage III disease compared to those with stage II or high-risk stage III CRC, again in line with the strength of evidence from SCOT and IDEA subgroup analyses. It is interesting to note that this distinction in application of trial results by subgroup occurred in Scotland even before the full publication of SCOT trial results was disseminated (April 2018).

It was clear (Figure 8-9) that the monthly proportion of SCOT eligible patients receiving over 3 months of chemotherapy started to decline in February 2017, approximately four months prior to the dissemination of SCOT results. If patients started adjuvant treatment at the end of February, they would be reaching 3 months of treatment at the end of May/start of June. This phenomenon could therefore be explained if clinicians not only changed their practice for patients commencing treatment in June 2017, but if they used the results to terminate at 3 months the treatment that was ongoing at that date.

Looking at duration of treatment and the regimens used in more detail pre-versus post-SCOT allowed direct comparison with the survey results from Chapter 5. The results of the survey reported in Chapter 5 highlighted that age may affect the extent of practice change, with older patients being less affected by the SCOT trial findings due to the lower use of doublet chemotherapy in this cohort. Although there was a bigger change observed for patients aged under 70 in this study, there was still a significant change in practice for those aged 70 and over, and adjusting for age did not significantly alter the change the reduction in odds of patients receiving over 3 months of treatment post-SCOT. In keeping with the survey results (Chapter 5), there was a lower use of doublet chemotherapy for patients in the older age group in general.

For patients aged under 70, there was a higher use of 3 months of doublet chemotherapy across all stage III scenarios in this Scottish cohort compared to

the survey results. This may be explained by the influence of international clinicians, who were less likely to change practice and more likely to use FOLFOX, compared to UK oncologists. It also suggests that clinicians are not over reporting their practice change in the survey, and if anything, the survey estimates may have under-estimated SCOT and IDEA impact. For both age groups in the national cohort, patients with high-risk stage III disease were less likely to receive the shorter duration of doublet chemotherapy if they had T4 as their only high-risk feature compared to N2. This does not fit with the survey results that showed minimal difference in practice between patients with either one of the high-risk stage III features, but does fit with the practice change that was predicted by an independent clinician reviewing the IDEA results at the ESMO plenary session described in Chapter 1 (78).

Overall, the proportion of patients receiving 3 months or less of doublet chemotherapy in this real world dataset, and the proportion of clinicians choosing this treatment option in the survey, were very similar for stage II and low-risk stage III disease, for both age groups (<10% difference between survey and national data). The main difference between these studies related to treatment decisions for high-risk stage III CRC. In the survey, less than 20% of clinicians chose 3 months of treatment for each of the high-risk scenarios, whereas in the national dataset, around 40-50% of patients on average were receiving 3 months or less of this treatment option. One explanation for this could be that in the real world setting, patients were receiving shorter treatment because of toxicity, rather than an intentional choice on the part of their treating physician. However, if this was true, the same logic would apply to low-risk stage III disease, where the same magnitude of difference was not observed. The national dataset looked at prescribing up to mid-2018, whereas the survey was performed in 2019, therefore an argument regarding increased adoption of trial results nationally would also not stand. The explanation for this difference may lie in clinicians underreporting a change in practice if they perceive the change to be controversial, or it may reflect differences in practice change between Scottish versus other UK or international clinicians, especially those using FOLFOX.

Only half of patients diagnosed with stage III disease during this period in Scotland received adjuvant chemotherapy. This is in line with recent estimates from England (258) and raises the question of why so few patients are receiving this treatment. The QPI target for adjuvant chemotherapy delivery for patients aged 50-74 with stage III CRC in Scotland is 70%; those who refuse treatment or receive neoadjuvant chemotherapy are excluded from the calculation (310). The proportion of patients with stage II disease receiving adjuvant chemotherapy was even less, but this was expected, and is partly dictated by the proportion of patients with stage II disease with high-risk features. This finding shows the utility of using real world data to estimate trial impact. Specifically, these results demonstrate that predicting practice change associated with the SCOT trial based on incidence of stage II/III CRC alone would have over-estimated the health service impact.

One limitation of this study (and the GG&C study in Chapter 7) is that it is not possible to claim causality of the SCOT trial on the practice change identified, only to suggest that the SCOT trial contributed to or was associated with the difference in prescribing practices seen after June 2017, compared to before June 2017. Despite this limitation, in combination with the survey results (Chapter 5), these two methods present a strong picture of practice change in response to SCOT, showing, as suggested in the literature review in Chapter 3, a mixed methods approach to evaluation can be worthwhile.

This study used four years of data from before the SCOT trial findings were disseminated but only one year of data after June 2017. It is therefore not clear from this study if this change in practice was maintained over a longer period or if any changes to the treatment of the SCOT ineligible cohort would have emerged with longer follow up. It would be interesting to have data for additional years to know if the effects of disease stage and regimen remain, now that stage II results (87) and updated OS (83) results have been published. The results from the GG&C (Chapter 7) give reassurance that the trends seen in this study in 2018 are likely to be maintained in 2019. In future work it would also be interesting to see the effect of the COVID-19 pandemic on prescribing in this patient cohort. It would not be unreasonable to expect that clinicians may have had more reason to use a shorter treatment duration during the pandemic in

order to reduce healthcare service use. It would be interesting to explore if any change in prescribing practices during the pandemic revert to pre-COVID trends after the main viral peaks have subsided, as predicted by the results in Chapter 5.

There was also insufficient follow up post-SCOT to compare health outcomes such as survival pre versus post June 2017. Even if longer follow up had been available, no routine collection of patient reported outcomes currently exist nationally for CRC patients to evaluate the potential decrease in toxicity (for example peripheral neuropathy and diarrhoea) that may accompany the practice change observed. The likely impact of SCOT on these health outcomes can only be estimated at this stage, using the numbers of patients in this cohort. The SCOT trial findings demonstrated that the percentage of patients experiencing at least grade 2 neuropathy was 58% in the 6-month trial arm and 25% with an intended 3-months of treatment. Applying this to the number of patients receiving oxaliplatin-fluoropyrimidine doublet treatment post-SCOT in this cohort (n=231), if no practice change had occurred and 75% of patients had still received over 3 months of chemotherapy in June 2018, approximately 100 patients would be expected to suffer from neuropathy that was severe enough to affect their activities of daily living. In contrast, the prediction of the proportion of patients having over 3 months of chemotherapy using actual data (24%) means that this number will drop to 32. In addition, Chapter 6 indicated that the discounted cost savings per patient over 10 years from using 3 rather than 6 months of chemotherapy are approximately $4,553 USD. Applying these savings to patients receiving doublet chemotherapy in this real world cohort translates to a healthcare system saving in Scotland of $523,595 (approximately £382,224) for patients commencing chemotherapy within the one-year period post-SCOT. Also, if it is assumed that patients who receive an intended 6 months rather than 3 months of treatment have at least 6 additional consultant outpatient visits and chemotherapy day bed visits, for patients in the SCOT eligible cohort, this would translate to a reduction in requirement of almost 700 outpatient and 700 day case trips per year in Scotland.

In addition to the results from the individual patient analysis, there are other important learning points from this work regarding accessibility to real world

data. If datasets such as this one used in this study are to be available to investigate cancer trial impact in future, it is important to be transparent about the process of data acquisition. Table 8-7 explains the main challenges encountered to achieve data access, and outlines recommendations for conducting the same process in future.

Many of these recommendations in Table 8-7 are aimed at policy makers and specifically the Scottish government. One of the major suggestions is to invest in staff and infrastructure to support data acquisition and linkage in particular to employ more people to support data extraction from ChemoCare sites in the North and East of Scotland. Looking to the future, a more permanent solution to the difficulty around accessing data from the five different ChemoCare locations in Scotland would be to hold this data centrally, as is currently done in England (311).

Another major recommendation is to replace the current model for access to administrative health care data in Scotland of 'link and destroy' for individual research projects (312). Currently, multiple researchers might apply for the same or similar data to be linked, but this is all done in silos. Although it does help to address many of the privacy concerns from data controllers, it is not an efficient use of data providers' resources. The current model, in which this dataset will be destroyed in 5-10 years, poses a threat to research integrity and transparency, in the sense that it will difficult to reproduce results that have stemmed from a bespoke linked dataset that has subsequently been destroyed.

Table 8-7 Challenges of building a linked, administrative healthcare dataset for CRC on a national scale and recommendations for repeating this process for other tumour types in future.

Issue	Barrier	Recommendation*
Data specification	No national data dictionary existed to describe information held in each ChemoCare system. This made dialogue with ChemoCare analysts to plan the project specification more difficult, ultimately leading to the requirement for a substantial PBPP amendment to reflect the actual data received versus the variables originally approved.	Data dictionaries for datasets being linked are a requirement to know in advance which data variables will be accessed and linked.
Capacity	The length of time required to obtain ChemoCare datasets was partly attributable to a lack of capacity for staff within regional cancer networks to engage with the process. The responsibility for physically downloading reports from the ChemoCare system was often performed by an individual whose major responsibility was service provision. In addition, one ChemoCare site had specific difficulties with the software required to store large datasets.	Investment is required to ensure sufficient staff capacity at regional sites so that resource is not being diverted from service provision without proper recognition of this effort. Staff capacity at central sites needs to be sufficiently robust so that there is no slowing of data transfer and linkage set-up due to external pressures such as annual leave/sickness/other projects. There should be continuity in the staff managing data transfer and linkage.
Data transfer	Each transfer of data from data providers external to PHS required careful communication between the sender and recipient because data deposited in the secure transfer environments was automatically deleted if not picked up within 72 hours. Launch of a new secure file transfer system coincided with the data transfer and indexing process and at times, it was necessary to utilise a separate platform because of problems with the new system.	All parties involved in a data transfer need to prioritise communicating effectively within the same time window regarding a data transfer if the data transfer is to be successful. A secure data transfer platform is required. It should be straightforward to use by central and regional data analysts, with ready access to information technology support if any technical issues arise.
Data linkage	Data linkage took longer than anticipated because the first attempt experienced technical problems. This unexpected difficulty was partly due to the number of datasets being linked, as well as the impact of the Covid-19 pandemic when resource reprioritisation was required within PHS.	Easier data linkage would be possible if the data was held by a central data controller, such is the case for SMR datasets in Scotland and as occurs with the Systemic Anti-Cancer Therapy database in England (311).

Data Access	The datasets outlined in Chapter 2 Figure 2-4 represent most of the full dataset that was approved, however, unscheduled care data is still not available.	A secure research environment to store and analyse data is required to meet data governance and privacy requirements. The NSH is one example of this type of research environment. Others exist and some are industry-led, for example, AIMES Management Services Ltd. Preserving linked data, such as the datasets described in this project, as a repository should be a priority. This will facilitate data access for future researchers and reduce wastage of resources.
Resources	A substantial portion of the time-line stipulated by the funder for this project was dedicated to data access. The timeliness of data access has previously been documented as a barrier in several other UK projects (313, 314) and raises a broader issue around the ability of early career researchers to use nationally linked cancer datasets that include chemotherapy data in Scotland within the current landscape. The cost for data access correlates with the number of datasets external to PHS being linked, which also makes it infeasible for an early career researcher without links to another group or significant funding.	Training of staff at regional sites is required to ensure they have the skills required for efficient extraction, analysis, and transfer of large datasets. These staff also need access to proper information technology infrastructure that can deal with large datasets.

* Regional datasets = the same information for different locations within the same country are held by individual data controllers at a regional level, for example ChemoCare datasets. Regional sites = the organisations holding regional datasets. Central datasets = datasets which are stored and maintained at a national level, for example SMR datasets in Scotland. Central sites = the organisations holding central datasets.

Accessing administrative data is a hot topic in the light of the COVID-19 pandemic, when the importance of having access to real world data in a timely manner to know how patients are being treated and the corresponding outcomes became clear. A group led by Dr Ben Goldacre has been commissioned by the UK government to investigate how healthcare data can be used both efficiently and safely for research purposes in light of the COVID-19 pandemic. This review, the results of which are expected in mid-2021, is being performed in parallel with the Data Strategy for Health and Social care which aims to improve data use within and by the UK government (315).

There are efforts in Scotland to streamline and improve the use of cancer specific administrative datasets. For example, the Cancer Medicines Outcomes Project (CMOP) was commissioned by the Scottish government in response to the 2016 Beating Cancer: Ambition and Action report (316). The overarching aim of the CMOP programme is to maximise the use of electronic records to understand outcomes for patients treated with cancer medicines in Scotland, and one of the key objectives is to test the scalability of linking cancer medicines datasets at a national level (317). A separate programme of work, the Scottish Cancer Registry and Intelligence Service (SCRIS) (318), was established in 2017 with the aim of creating a national Cancer Intelligence Platform, with national reporting of cancer outcomes and treatments available to approved users via a dashboard. In late 2020, chemotherapy prescribing data (ChemoCare) covering 100% of the population was added to this platform. Due to data privacy concerns, this system is not intended to grant access to researchers to analyse individual patient level data, rather its primary function is for use by service providers in their delivery of cancer services in Scotland.

More recently, stakeholders within the Scottish Government have recognised the potential benefits of preserving the linkage between datasets and storing data indefinitely for use by multiple research projects, whilst maintaining appropriate information governance protocols. Specifically, Research Data Scotland (RDS) was launched in 2019 as part of the Programme for Government (21). This is a not for profit organisation which aims to improve the economic, social and environmental wellbeing of Scottish residents by enabling access to linked data, not limited to healthcare datasets, for research in the public good.

Allowing researchers to use this type of administrative data to assess the impact of clinical trials in a real world population could be beneficial for government organisations, such as the Scottish Medicine Consortium in Scotland or NICE in England to inform future decision making on drug approvals. These regulatory bodies often have to rely on trial data to extrapolate the effectiveness and cost of treatments in a real world setting. In addition, knowing how trial evidence impacts on practice within a population will allow the government to understand practice across different locations and build national treatment guidelines that are appropriate and relevant across the whole country. Evaluating trial impact in

this way will also show policy makers if there are any issues with equity of provision of care across communities, for example, based on socio-demographic characteristics of the population. In this way, trial impact assessment could be used to co-ordinate access to medicines and other new treatments in a way that improves the health of the population as a whole.

8.5 Conclusion

The duration of adjuvant CRC treatment changed at a population level in Scotland after June 2017. This suggests a rapid translation of clinical trial evidence into practice and it is predicted this will lead to important health and health service impact at a national level. This study has shown that it is feasible to use patient level prescribing data at a national level in Scotland to investigate clinical trial impact, although significant investment of time and resources was required to make this successful.

9 Chapter 9: Discussion

This thesis has argued that cancer trial impact is important and that understanding how to approach its evaluation is worthwhile. This refutes a more traditional dogma that the act of performing research itself and the pursuit of knowledge alone are intrinsically valuable. In 2021, we live in a society of rising pressures on government, charity, and industry budgets. Increasingly, these organisations want to know if the cancer research they chose to financially support, and the cancer trials they decide to develop, are actually making a difference. This study has investigated how these impacts can be evaluated.

9.1 Summary of this research

The work in Chapter 3 identified that there is not one accepted approach to evaluate research impact and that literature discussing the assessment of cancer research impact is scarce. This gap was addressed by identifying key examples of cancer research impact evaluation, thereby allowing the categories, methods and frameworks used for impact evaluation within these cancer studies to be highlighted and discussed.

An analysis of the REF 2014 case studies in Chapter 4 demonstrated that higher education institutions mainly relied on documentary analysis, expert testimony, economic evaluations and occasionally, pre-existing audits of practice, to describe cancer trial impact. There were no approaches to impact assessment used within these case studies that had not already been identified in the overview of the literature (Chapter 3), nor were there any examples of in-depth impact analyses undertaken specifically for the purpose of the REF 2014. Nevertheless, these case studies did show the types of impacts most relevant and most often assessed for clinical cancer trials specifically. These categories were policy impact, impact on new knowledge, and benefits to health or the health sector.

The case study in this thesis showed that clinicians were using evidence from a large, phase III clinical trial (SCOT) in practice. This impact was successfully demonstrated using surveys and analysis of prescribing data. As highlighted in the previous literature (Chapter 3), this mixed method approach to evaluation

was useful. The limitation of the survey was that it was not clear if self-reported practice reflected real world prescribing, and the drawback of administrative data analysis was the ability to assess the attribution of any changes specifically to clinicians' application of trial results; these approaches were therefore complementary.

The dataset used for the national analysis represented the first time that systemic anti-cancer data has been collated on a Scotland-wide basis, and linked with other administrative datasets for the purposes of research. This allowed exploration of practice change on a national scale and overcame many of the issues with generalisability from analysis of local data only. This dataset also provided better transparency around how the patient cohort was derived, and gave information on patients that received chemotherapy, as well as those not treated with adjuvant chemotherapy across all locations. Although the barriers to accessing data were ultimately overcome, the process of data acquisition has provided lessons for how this issue could be tackled in future.

Lastly, this thesis has shown that health economic approaches can be applied to assess the impact of a clinical cancer trial, and that survey data can be used to strengthen this approach by providing an indication of implementation of trial results. In this instance, the value of the health sector and health impacts from trial implementation far outweighed the initial research investment to conduct the SCOT trial.

9.2 Finding a framework for SCOT trial impact evaluation

Six frameworks for impact evaluation were identified and discussed in Chapter 3 as being potentially relevant to cancer research. Appendix 6 includes a description of how these frameworks can be applied to the SCOT trial, using the results from the analyses conducted in Chapters 5-8, as well as additional documentary analysis. A modified version of the Payback framework has been included below (Figure 9-1) to summarise the impacts of the SCOT trial. An accompanying list of indicators of impact that are relevant to SCOT can be found at the following link: http://researchdata.gla.ac.uk/1135/, based on the list of indicators developed for the purposes of the content analysis described in Chapter 4.

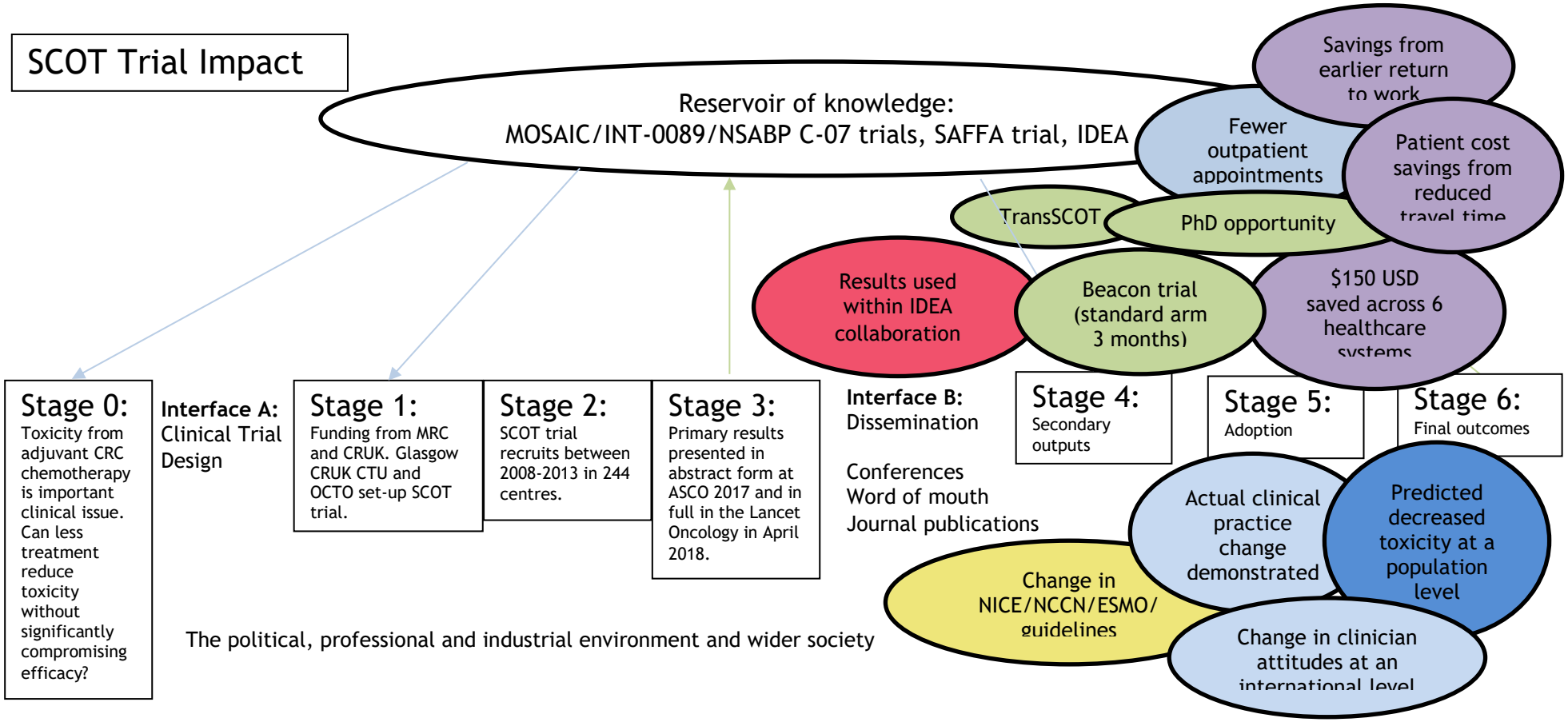


Figure 9-1 The Payback Framework modified to consider how the impact of the SCOT trial Adapted from the original Payback Framework (15)

The modified Payback framework is drawn above as a linear pathway but it is unlikely that pathways to impact will be unidirectional, predictable or consistent between different cancer trials. Figure 9-2 below shows a more realistic depiction of the pathways to impact applied to the SCOT trial. If using the modified Payback framework, or any of the other frameworks outlined in Appendix 6, it is important to be flexible in considering the processes through which impacts occur and the order in which they may occur over time.

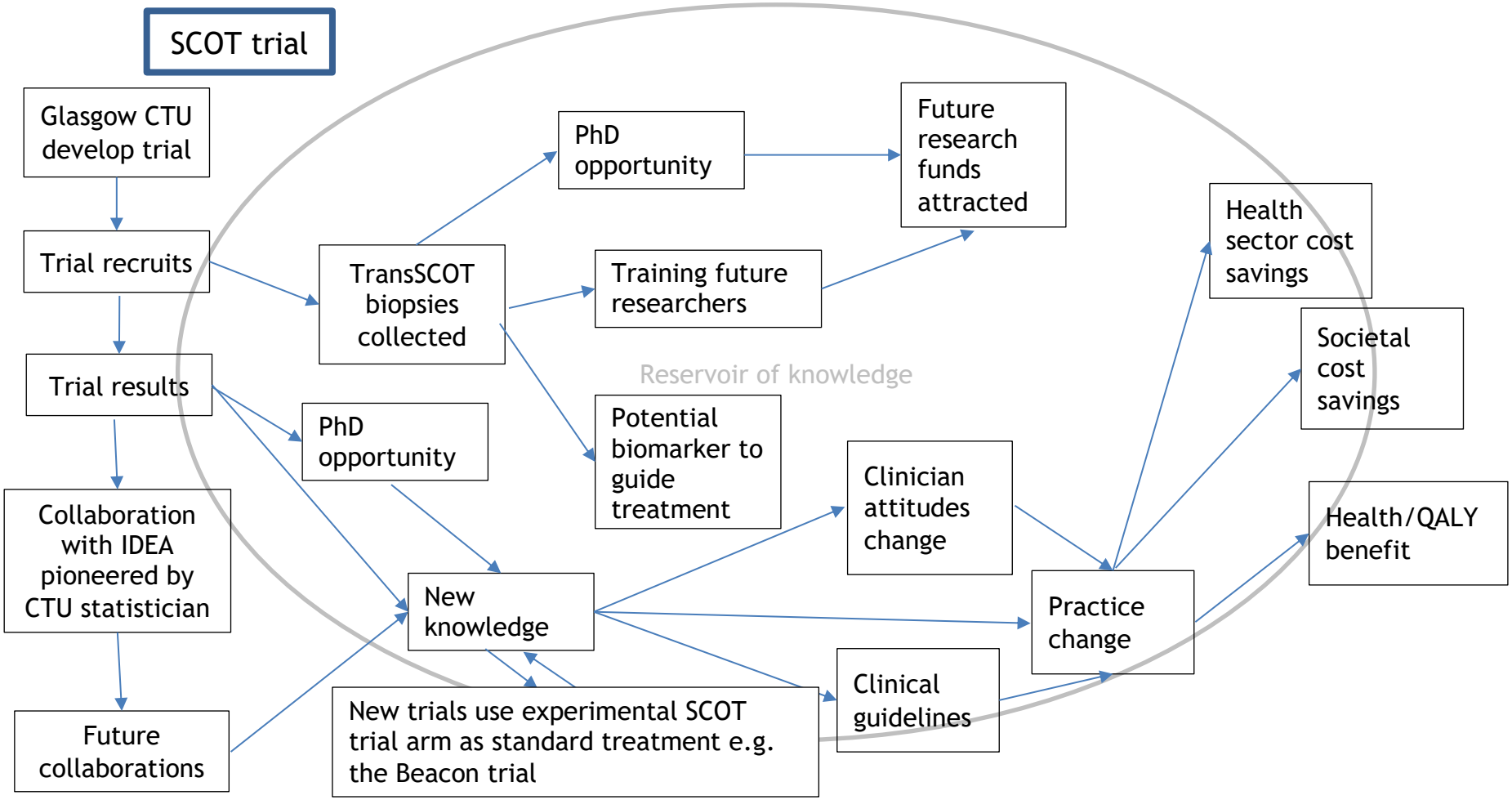


Figure 9-2 Potential pathways to impact for the SCOT trial Impacts in grey are those not identified for SCOT but examples of other types of impacts that could be investigated for other clinical cancer trials.

9.3 Lessons learned for future trial design from the SCOT impact case study

Using the survey results from Chapter 5, it is clear that the post-hoc subgroup analysis, which stratified stage III patients into low and high-risk, had an important influence on the implementation of SCOT trial results. This raises the question of whether the difference in outcome according to tumour and nodal stage could have been predicted, and if the analysis could have been planned rather than post-hoc. If designing a similar non-inferiority trial in the future, it would be important to consider that clinicians treating any type of cancer may be reluctant to reduce standard treatment for higher risk patients and it would be useful to try to define high-risk patients *a priori*. Reflecting on the previous surveys performed by Blinman and colleagues (255, 256), it would also be useful to ask patients in advance of developing a trial, what decrease in benefit they consider to be not unacceptably worse than the standard of care, and to plan the statistical trial analysis accounting for these preferences.

The analysis of administrative healthcare data reported in Chapters 7 and 8 demonstrated that the treatment regimen used (CAPOX versus FOLFOX) also impacted on implementation of trial findings. The subgroup analysis based on regimen type was pre-planned, but the findings from the impact evaluation raise the question as to whether more could have been done to account for potential differences in outcome based on regimen, for example using random allocation of patients between FOLFOX and CAPOX treatment. Prior to SCOT, the X-ACT trial had shown that capecitabine monotherapy was not only non-inferior to fluorouracil, but there was a trend to better survival (293) and the SAFFA trial indicated that the method of administration of intravenous 5-fluorouracil may have had an effect on efficacy of treatment (51). In defence of the SCOT trialists, not randomising by the regimen prescribed allowed for a pragmatic trial design and reflected the freedom of choice oncologists would have had in the clinic. It appears however, in this instance, by not using randomisation, it is now impossible to know from the SCOT trial results if there is a true difference in efficacy between the regimens and this has made the trial findings more difficult to interpret and implement.

The results of the survey in Chapter 5 indicate there was more uncertainty around the use of shorter treatment for patients with stage II disease and that clinicians were less comfortable commenting on their practice for rectal cancer in the adjuvant setting. Patients with stage II disease and/or rectal cancer were included in the SCOT trial but not across the whole IDEA collaboration. If planning the IDEA collaboration today, a better approach would be to agree in advance the population of patients to be recruited across IDEA trials. The survey results in Chapter 5 also showed that clinicians approached the treatment of patients aged over 70 differently to younger patients, with the results from Chapters 7 and 8 indicating that patients receiving doublet chemotherapy in a real world setting are, on average, younger that those recruited to SCOT. Overall, these findings highlight to trialists to be mindful of patient age and to think of how clinicians can be reassured in future regarding how to apply trial results to older patients. This may involve doing a separate trial for older patients, which has been done in the FOCUS2 trial (233) (discussed in Chapter 4), or ensuring sufficient numbers of patients are recruited to older cohorts to allow robust conclusions to be drawn from subgroup analyses.

Regardless of these drawbacks, what is clear from the impact assessment is the huge benefit from planned international collaboration at the outset of performing a trial. In addition, SCOT was the only IDEA collaboration trial that collected health economic data, which allowed thorough comparison of the cost-effectiveness of 3 versus 6 months of treatment; this should be encouraged for other phase III RCTs going forward. SCOT also collected pathological specimens for patients recruited to the trial (260) which has allowed a programme of translational work to be subsequently developed (TransSCOT). Impact evaluation (Appendix 6) has highlighted that analysis of these specimens has already led to new research and the generation of new knowledge. In this way, impact assessment demonstrates the wider impact that the SCOT trial is already having on future research, which would have been the case regardless of whether non-inferiority of shorter treatment had been demonstrated.

9.4 Do the methods used in this thesis work for assessing clinical cancer trial impact?

9.4.1 Surveys

The advantages of using surveys for cancer trial impact assessment (Chapter 5) were that a large number of clinicians from several locations could be reached, this approach allowed an element of quantification of impact on practice, and compared to administrative data, had the benefit that clinician awareness of trials and attitudes towards trial results could be assessed.

One of the major drawbacks, which will apply to any survey being conducted in this context, was that there was no recognised list of CRC oncologists that existed and was easily accessible, for the UK or internationally. Time and effort was required to build a UK list for the purposes of this project, but despite this effort, there was no way of checking the final list was correct, and the list will now be out of date. There was also no information available on non-responders, making it impossible to explore potential response bias. These issues could be potentially be addressed by the creation of a live list of all such oncologists, maintained by research institutions or funders, such as the NCRI, CRUK or Royal Colleges in the UK, or may be something that the clinical trials community may want to build in future, as a collaborative, nationally held resource. The use of any list would still require buy-in from clinicians to garner survey responses, and the potential for response bias, whereby more motivated researchers are more likely to reply, would still exist. A further issue with surveys is the burden on respondents, who are often busy clinicians. The fact that the surveys used in this study provided mainly quantitative results can also be seen as a limitation as well as a strength.

In future, surveys could be used to question researchers and trialists themselves regarding all impacts from their trials. This would be a useful approach to cancer trial impact evaluation in particular if the person undertaking the assessment was not familiar with the trial or topic area, or if a large number of trials were being assessed simultaneously. The Payback Framework could be used to structure these

surveys, and one of the empirical examples of cancer research impact identified in Chapter 3 (Donovan et al) shows how this can be done (195).

9.4.2 Health economic approaches

The method used in this study to assess the economic value of cancer research, adapted from Glover et al (39) and Brown et al (184), was successfully applied to one cancer clinical trial. This analysis relied heavily on health economic data collected as part of the SCOT trial. This type of analysis would still have been possible if health economic trial data did not exist but would have relied on cost, resource, and utility data either collected as part of previous trials or linked with administrative datasets.

An interesting piece of work would be to apply this type of assessment for a number of trials, for example, all trials conducted by one CTU or supported by a specific funder. This approach to assessing a programme of trials has been performed previously (2006) in a non-cancer setting by Johnston and colleagues, who assessed the return on the investment into phase III RCTs (n=28) funded by the US National Institute of Neurological disorders (301). These types of evaluations at the programme level can be limited by selective advocacy of the most impressive outputs, and any evaluation of a programme of cancer trial work would need to include the investment into all trials, not only those that met their primary objective. If assessing the impact of other trials in future, in particular trials with a superiority end-point, it would be useful to understand if, when new interventions are applied to routine clinical practice, they are achieving the same level of benefit in a real world setting compared to the select groups of patients recruited to phase III clinical trials (319).

The analysis in Chapter 6 used implementation of SCOT trial results at the core of the assessment. Earlier phase trials are less likely to impact on practice and will have other important consequences for drug sales and future research, collaborations, or stopping a futile treatment from being developed further. If assessing a programme of trials that includes a mixture of early and late phase studies, the approach used in Chapter 6 would need to be adapted to allow

assessment of these different types of impacts. Alternatively, it could be recognised in advance that investment into early phase trials are critical to the pathway of developing and running larger trials that will ultimately change practice and health, and this investment is simply added to the cost-deficit side of the equation.

9.4.3 Administrative data analysis

The approach of using prescribing data for the analyses conducted in Chapters 7 and 8 provided robust evidence of the impact of the SCOT trial. The strength of using administrative data was that it provided a clear message that practice change actually occurred, and that this change could be measured. Different methods to analyse prescribing data were tested. In this instance, the change in practice was obvious even using simple descriptive statistics. Out of the other methods tested, ITSA was the additional method that added the most value; the strengths of the ITSA were the ability to visualise the trends in prescribing pre and post-SCOT, as well as the counterfactual situation. ITSA may be even more useful as a method to apply to the assessment of other clinical cancer trials where practice change is less obvious using descriptive analysis alone.

9.4.4 Documentary analysis

Although not the focus of any of the results chapters, documentary analysis was utilised when applying impact frameworks to describe the impact of the SCOT trial (Appendix 6). Analysis of secondary sources of information in this way is likely to be a core part of any research impact assessment and in this instance complemented the results of specific assessments (Chapters 5-8) by allowing a narrative of clinical trial impacts to be generated.

9.4.5 Case studies

Within this study, the methods outlined above were used to populate an impact case study for the SCOT trial. Case studies are a standalone method for impact evaluation, and put simply, involve the process of taking a specific research project or programme of work and building a narrative around the impact of that

research. Case studies are a good approach to evaluate the impact of clinical trials because each trial is a research entity that can be the focus of a separate case study. Impact case studies are now a well-known approach to impact assessment in the UK because of the REF 2014, but their utility should be viewed as extending beyond the REF.

9.4.6 Methods not used in this thesis

There were several methods for research impact assessment identified in Chapter 3 that were not employed in this study, but that could be relevant to cancer trial assessment in future. The most obvious is an in-depth assessment of policy impact. Although this was done to a limited extent using documentary analysis in Appendix 6, Lewison et al (188, 189, 191) (Chapter 3) have shown that this could be performed on a larger scale using bibliometric software. This may be an approach to cancer trial impact that could be used by funders or research organisations to understand how a collection of cancer trials influences policy, and how soon after trial investment this impact occurs.

All of the approaches described above may be broadly applicable to clinical trials covering disease entities other than cancer. The differences for trials assessing other diseases may include the time lines expected for changes in health from trial implementation. For example, cancer trials often assess improvements in survival over many years, whereas other disease sites such as cardiology or infectious diseases, may focus on much shorter time lines such as an improvement in blood pressure, or, as we have seen in recent COVID-19 trials, 28-day mortality (320). Also, the treatments tested in trials may differ for other fields, for example, there may be a stronger focus on surgical techniques or device interventions, rather than on investigational medicinal products. This study has focused on a trial assessing drug treatments, but even within cancer trials, different approaches to impact evaluation may be required for trials focusing on a different part of the cancer management pathway, for example, screening or diagnosis.

9.5 Reflection on cancer trial impact evaluation from the perspective of different cancer trial stakeholders

In an ideal world, the impact of every cancer trial that is performed would be assessed. In addition, this assessment would be considered and planned, both before a trial is carried out and after the results have been disseminated. This is an aspirational statement but it is a useful exercise to consider how this could be done. Specific suggestions for how cancer trial stakeholders can engage with and support the process of cancer trial impact evaluation going forward are discussed below.

9.5.1 Funders

Cancer research funders have indicated that achieving wider impacts are a key driver for their work. For example, CRUK state the aim of their research is 'to beat cancer sooner', with the objective to see three-quarters of patients diagnosed with cancer surviving the disease within the next twenty years (321). The NCI, the largest cancer research organisation in the world, have a mission to help 'all people live longer, healthier lives' (322). Finally, the Institute of Cancer Research in the UK has a vision of 'a world where people can live their lives free of cancer as a life-threatening disease' (323). It is not clear how funders will be able to show their donors they are achieving the aims set out in their mission statements, unless they actively evaluate and reflect on the impact of research performed as a results of their funding decisions.

Cancer charities such as CRUK have faced major pressures on their budgets due to the COVID-19 pandemic and they face tough decisions surrounding funding allocation going forward. Using a framework, such as the one outlined in Section 9.2.7, to create impact narratives, could help funders decide how to allocate their increasingly limited funds. An example of good practice was identified in Chapter 3 in the article from Donovan et al (195) which described an evaluation of the impact of investments by the National Breast Cancer Charity in Australia. Cancer funders could perform such an assessment focusing specifically on cancer trials,

whilst keeping their mission statement at the core of any reflection on the impacts identified.

Whereas using impact narratives can help funders with allocation of resources, using economic approaches to impact assessment could be used to advocate for increased funding from donors. In particular, adopting an economic approach to assessment can show members of the public that their investment in charity funded research will have health and societal gains which far outweigh the initial charity contribution, and at times may even be cost-saving, especially in countries where the healthcare systems are run using public money.

Lastly, despite the rapid uptake of SCOT trial results into practice, this study has shown that it still took over ten years from the SCOT study being granted research funding to practice impact. Using impact evaluation to document trial timelines and to reflect on if current trial timelines are appropriate will be a useful exercise for funders. Events during the COVID-19 pandemic have shown the speed with which impactful trials can be performed when required. A better understanding of trial development and set-up timelines, through impact evaluation, could help funders outline timelines in which they expect their trials to be completed and potentially increase the impact from their investments.

9.5.2 Research institutions

We know that research institutions are being asked to evaluate the impact of their work, and that in the UK, core government research funding is dependent upon these evaluations. Chapter 4 showed that a major focus for cancer impact narratives from higher education institutions in the REF 2014 was on policy. There was an apparent gap in assessing less obvious impacts such as improvements in research capacity and innovation, and on longer-term impacts on health and the economy. Now this gap has been highlighted, research institutions can be aware of how their approaches to impact assessment can be improved going forward. The results from Chapter 6 show that the economic value of research can often far outweigh the investment. This type of impact evaluation approach could be used

by research institutions to lobby the UK government to increase the overall pot of money being distributed in the REF.

We know that in 2014, preparation of submissions to the REF cost UK higher education institutions £246 million, with £55million spent on impact evaluation. This was a 133% increase from the Research Assessment Exercise in 2008 which did not include evaluation of research impact (241). If the assessment of real-world cancer trial impact is to be sustainable, and if it is to happen for reasons other than the REF, provision of resources to perform such evaluations will need to be addressed either through core funding or specifically within project specific research grants.

9.5.3 The pharmaceutical industry

Pharmaceutical companies appear to be increasingly interested in sharing their clinical trial data and using real world data to explore the effectiveness and impact of their products on patient health outside of clinical trials. Pharmaceutical companies will often be considering the impact of their drug, rather than the impact of a specific clinical trial, but the principal is similar. This interest is reflected in the recent acquisition of Flatiron, an oncology specific electronic health records company, by a large, cancer-focused pharmaceutical business (Roche) for $1.9 billion USD (324). Flatiron have partnerships with a large network of cancer clinics and research facilities. In addition, companies such as Lily, GSK and Johnston and Johnston are also now routinely sharing their clinical trial data for the purposes of research, on platforms such as Vivli (325), the Yoda Project (326), and Clinical Study Data Request.com (327). This is a clear recognition that trial data can be used for more than the purpose of the original trial proposal. Also, as we have seen with the development of the Oxford-Astrazeneca COVID-19 vaccine, there was a commitment by the pharmaceutical company to provide the vaccine at a not-for-profit cost for a specific time-period. Clearly, these companies are partly driven by financial gain, but there are indications that the industry as a whole recognises the importance of working closely with researchers and looking at real world data to make better investment decisions and achieve their institutional missions.

9.5.4 Cancer researchers

Cancer researchers are increasingly being asked on research funding applications to think about the impact of their proposed trials, and this study will provide these researchers with a framework for how to approach these assessments. It is important for researchers to recognise the burden of doing impact assessment early and to ask for proper investment from their institutions and funders to support these activities. There are also implications from the SCOT impact case study for CRC clinicians and trialists designing their own trials. Firstly, this study has shown that real world impact can often be diluted if not as many patients as expected actually receive cancer treatment tested within the trial. It has also shown that it can be difficult to predict when subgroup analyses from clinical trials will be used in practice and how they will be applied.

9.5.5 Clinical Trials Units and the wider trials community

In the UK, CTUs and the broader trial community, including NCRI groups and trial networks, make decisions on which cancer trials are developed. In this way, just like government policy makers and funders, the trial community have control of allocating investment of research time and money. A clear way for the trials community to engage with the impact agenda could start with each CTU or trials group having a clear mission statement, co-developed with patients, around the types of impact they would like to see arising from their trials. It is likely that CTUs would benefit from coming together to discuss and share their thoughts and to build a culture of evaluation in which approaches to cancer trial impact assessment are routinely shared between CTUs, peers, and stakeholders. If not being done already, evaluation of trial impact at the stage of trial development could improve decision making around which trials to support, by looking further than the merits of the scientific question being asked. This could be done in a straightforward manner by using the modified Payback Framework described above, and producing a narrative for each potential trial around the expected impacts that could be achieved.

Whereas *ex-ante* evaluation could help with funding allocation, assessing the impact of trials that have already been performed could help groups such as CTUs advocate for more investment overall from their funders, and allow analysis of what types of trials are aligning with their overall goals. These *ex-post* impact narratives should be stored in an accessible way to provide a rich resource for early career trialists, including statisticians, clinicians and others involved in trial development. Such a resource would allow trialists to learn from the experience of those individuals who have designed trials before them, and would also help to compare the wider impacts of newer trials with complex designs. If the act of reflecting on trial impact is not considered or prioritised, this opportunity for learning how to improve on trial design is lost. The trials community could also disseminate patient-facing cancer trial impact narratives, for example, via their CTU websites. The COVID-19 vaccine trials have provided an unprecedented example of trial impact occurring on an international scale with significant media coverage. In most other circumstances, patients who participate in clinical trials are not always aware of the trial results, or how their participation may have benefited other individuals.

9.5.6 Patients

Patients often help the trial community to decide which trial ideas to develop. The recent impact summit, described in Chapter 1 (35), identified that designing and building research programmes with central partnership between patients and researchers in mind helps to maximise the impact from those research projects. Specific to cancer trials, CTUs could ask their patient advocates which impacts the CTU should focus on, the optimal trade-off between the burden of the trial schedule for patients recruited to the trial versus the potential impact of that trial, and which research questions are most important when considering the downstream impacts they would like to see come to fruition. They could also ask patient representatives to help disseminate their trial results on patient-facing platforms and to lobby for trial results to be cited in policy and implemented in practice.

A patient representative was involved in publishing a description of the process to access health care data described in Chapter 2. They were surprised at the length of time required to achieve data access. In their own words (included with permission), “From a patient perspective it’s hard to comprehend that the data isn’t linked and that this type of work is done time and time again on the same data – seems wasteful and blinkered.” Hopefully the account included in this study will demonstrate the need for both researchers and patients to lobby policy makers for better infrastructure to be in place to streamline this process in future and to avoid wastage of public investment.

9.5.7 Unanswered questions across stakeholder groups

There are unanswered questions and criticisms of impact evaluation that need to be addressed. First, it is not yet clear whose responsibility it will be to perform these evaluations. If impact evaluation is to be taken seriously, it requires infrastructure and investment in people who have the skill set to plan and conduct impact assessments. Given the investment, critics will ask if research impact assessment is actually a cost-effective exercise in its own right. Unless the results of impact assessments are communicated, and provide lessons that are used, the value of performing impact evaluation and the enthusiasm to invest resources to do so will wane.

Assessments also need to be performed in a timely manner so that the results can actually be fed back to develop more impactful cancer trials. The reality is that trialists are often busy clinicians and CTU directors and project managers are often stretched in their capacity. In order to incorporate reflection on impact and its evaluation into their current responsibilities, these individuals will need to have the time and resources available in order to plan an impact strategy. In addition, there will need to be funds to pay for these activities, either from core funding or incorporated into trial specific grants.

It is important to safeguard individual researchers when it comes to cancer trial impact evaluation. As described in Chapter 2, when bibliometrics were first developed, they were used for decision making around researcher performance

and promotion. The mistake was being made of counting what could be counted rather than what was a meaningful reflection of research performance and quality. There is a similar risk when it comes to broader impact evaluation. It will be important not to place undue pressure on an individual researcher to produce impactful research. As highlighted in one of the reviews in Chapter 3 (169), research by definition is an investigation of the unknown, which is dictated by uncertainty. Research success can be due to luck and serendipity and can sometimes take years to achieve. Pressuring individuals to ensure their research produces wider impacts may be counterintuitive because it may limit risk taking and the expansive, creative thinking that is an intrinsic part of the research process. Even when applied at the programme or institute level, care needs to be taken regarding how impact evaluation is used by institutions to make short-term investment decisions and promotions. This differs from encouraging researchers to maximise the impact of nay work they do, which can be achievable. Table 9-1 offers suggestions of how better to evaluate, communicate and maximise cancer trial impact in the future.

Table 9-1 Recommendations for cancer trial impact assessment Abbreviations: HEI, higher education institution; CTU, Clinical Trials Unit.

	Recommendations:	Target group
Evaluating the impact of cancer trials	- Educate trialists to anticipate the types of data required to evaluate impact and the collection methods to acquire this data.	HEIs Funders
	- Use indicators of cancer trial impact (for example, those in the content analysis coding manual) to identify the wider impacts of future trials and to describe the impact of completed trials.	Trialists CTUs HEIs
	- Assess how cancer trial results are used by decision makers. This will create a narrative of the pathways through which impact occurs (direct and indirect). This process may uncover unexpected and less clearly defined impacts.	Trialists CTUs Researchers HEIs
	- Identify examples of researchers or patients actively contributing to maximising trial impact.	Trialists CTUs
	- Evaluate the impact of negative trials. Demonstrating impacts that do not rely on positive trial results will encourage funders and researchers to adopt a broader approach to trial assessment.	Trialists CTUs
	- Provide funding and support for robust cancer trial impact evaluation.	Funders HEIs
	- Give researchers access to real world data, in particular healthcare administrative datasets and real world cost data.	Policy makers
Communicating the impact of cancer trials	- Publicise cancer trial impact evaluations. Platforms for publicising evaluations could include patient facing charity websites, CTU websites and clinical trial registries as well as more formal channels such as open access publications.	Trialists CTUs HEIs Funders
Maximising the impact of cancer trials	- Incorporate impact assessment into the trial design process. This will generate ideas for researchers and CTUs of how they can take a more active role in maximising impact.	Trialists CTUs

	- Continue to provide opportunities for trialists to engage with stakeholders, including patients, in the planning stages of clinical trial design to specifically explore the types of wider trial impacts that are important to stakeholders.	Funders CTUs
	- Ask patients what type of impacts are important from their perspective and develop trials that aim to achieve these impacts.	Trialists CTUs

9.6 Strengths and limitations

A substantial strength of this study is the novelty of the question addressed. This study has brought clarity to a broad topic and distilled this into a more tangible discussion specific to cancer trials. Specifically, identification and critique of key examples of cancer research impact assessment and analysis of the REF case studies with a focus on cancer trials, are both novel additions to the impact evaluation literature.

An easier option for the focus of this study would have been to look solely at the implementation of trial results. However, after becoming familiar with the topic of research impact evaluation, it was evident that focusing on implementation alone would have ignored other types of impact relevant to cancer trials. The decision to adopt a broader approach to looking at trial impact was supported by the results from the REF case study analysis, which showed that cancer trials that do not meet their primary endpoint can still be impactful. Another strength of this work was the achievement of analysing SCOT trial impact at a national level. Prior to this study, no resource existed for this analysis to be carried out, and it was also not clear how the data could be used to investigate trial impact.

There are also several limitations to consider. First, a decision was made to perform an in-depth evaluation of one trial rather than to perform superficial assessments across several trials. The disadvantage of this approach is that it did not demonstrate the generalisability of the methods chosen and raises the question as to whether the methods used to evaluate SCOT could be applied to other trials. Secondly, the methodology used to assess SCOT trial impact was mainly quantitative, with no in-depth exploration of why impact occurred or if barriers to impact existed. It was also difficult to pick apart the impact of the combined

analysis (IDEA) from the impact of a single trial (SCOT) and this would have needed to be asked specifically within the surveys used or through using interviews with clinicians. Taking a qualitative approach to understand the process of trial impact would have been more relevant if there was a strong signal from a quantitative analysis that the impact expected had not occurred.

Real world change in health outcomes are often cited by researchers and funders as the ultimate mission from cancer research. An important expected impact from SCOT was a reduction in treatment toxicity in a real world setting, in particular, a decrease in peripheral neuropathy from using shorter treatment. Nevertheless, in this study, a detailed analysis of health outcomes attributable to SCOT was not undertaken. This type of analysis would have required access to administrative health records that included a measurement of treatment toxicity. The ChemoCare system in the South-East of Scotland does record this type of toxicity information but these variables did not form part of the national data set used in this project.

Associated with the work in Chapter 6, is the limitation of the concept of payback on the initial investment in the SCOT trial. Specifically, the financial investment and the subsequent cost savings in this assessment come from different sources; the financial investment is mostly provided by charities and research institutions, whereas the cost savings are enjoyed by the public health service, or if taking a wider perspective, by employers and patients. An implicit assumption in this study, and in many previous studies on this topic (39, 187), was that the cost of running a cancer trial can be subtracted from the cost savings to a healthcare system or patients. However, these costs come from separate budgets and the opportunity cost of investing in a cancer trial is not the same as the opportunity cost of investing in a new medicine or health technology. The only way this can be considered payback, is if a broad, societal view is taken of research investment. Specifically, one must assume that charities and research institutions are investing with the aim of societal benefit and their investment source has been provided by the same society it is looking to improve. This assumption may not hold when thinking about trials purely financed by industry.

There is overlap with some of the economic approaches to impact assessment with more traditional health economic 'return on investment' (328) calculations. One of the well-publicised criticisms of return of investment assessments is that, in the majority of cases, there is a positive return. Thinking about the similarities with impact assessment poses the question of whether it is useful to perform an evaluation that is ultimately always going to indicate that doing research has had some form of return compared to the input to the research process itself.

Lastly, an important limitation of this work was the minimal patient involvement. Initial conversations with a patient advocate were carried out in the first year of this study but it was difficult to have fruitful conversations because the concept of research impact and its assessment was still being understood and explored by the primary researcher (CH).

9.7 Future directions: Research questions arising and gaps identified from this work

In future it would be interesting to interview stakeholders such as patients, funders, researchers, research institutions and CTUs to investigate which categories of impact they value. In one of the articles identified in Chapter 3, the authors used the results of their literature review to ask patients which types of impacts from health research were most important (163). The development of patient-centred metrics of impact would also be worthwhile. As we move to embed patient reported outcomes into routine datasets, this will present a new opportunity to analyse trial impact from this more patient centred perspective. If what is important to patients is established at the outset, the impact assessment could be focused to explore to what extent this outcome was impacted by clinical trial implementation in a real world setting.

In Chapter 4, there was reflection on the burden for patients in participating in cancer trials, for example, the extra clinic visits, tests, and side effects from potentially futile treatment. In SCOT, patients recruited to the experimental arm of the trial experienced less toxicity from treatment and received less treatment, therefore the burden to patients was not the focus of this case study. However, if

considering the overall 'investment', financial or otherwise, into cancer trials in general, this concept of the burden on patients as part of the investment into a trial, could be explored further.

As described in Chapter 3, the concept of evaluating the returns from research initiated in the USA and Australia in the early 2000s. At that time, this approach was not identified as assessing the 'impact' of research, rather it was known as the concept of showing 'exceptional returns' (300) from research by looking at its downstream socioeconomic effects. An editorial '*Putting a value on medical research'*, published in the Lancet in 2006 (319), after Buxton's WHO review in 2004 (100), argued the importance of developing this under researched area of evaluating returns from research. It began with a quotation from the renowned American health activist and philanthropist, Mary Lasker, *"If you think research is expensive, try disease."* Despite this call to action, two decades on, and this field of research is still relatively under developed.

The approach to assess the economic value of conducting the SCOT trial used in this study focused on the cost-savings and health gains attributable to SCOT implementation. As outlined above, no assessment was made of the spillover effect of performing SCOT to the private sector or UK GDP. Examples of spillover may include private companies using 3 months of doublet chemotherapy in the standard of care arm in their next clinical trials. Researchers who received training as part of the SCOT trial effort may work on industry-funded trials in future, thus contributing the expertise learnt through their time working on a publicly funded trial, to future research efforts funded by the private sector. Also, the knowledge learnt through studies performed as part of the translational programme of work associated with the SCOT trial is likely to produce knowledge that can be used by pharmaceutical companies in their efforts to find new treatments for CRC. Although there have been previous attempts by US (329, 330) and UK (179) authors to estimate the extent to which publicly funded research can affect privately funded research and development, and the processes through which this may occur, the literature on how to approach this issue specifically is not well developed and is generally not disease specific. There is a need for a better understanding of this process and how this process could be realistically applied to

the assessment of the economic value of investment in one or a number of clinical cancer trials.

Chapter 8 demonstrated the extent of practice change after June 2017 relating to adjuvant chemotherapy prescribing in Scotland, but it would be interesting to calculate the real world cost savings related to this change. This would be possible because an update to the national dataset (December 2020) now contains detailed information on the doses of hospital and community medication prescriptions, as well as patient level information costing linked to inpatient and outpatient hospital attendances. More work is required to assess the best way to use this cost information within an analysis of the cost implications from trial implementation. Other research questions that have arisen from the work specifically described in Chapter 8 include why the proportion of patients receiving adjuvant chemotherapy was lower than expected, in particular for patients with stage III disease. Also, it is unknown why regional variation in the duration of adjuvant treatment delivered existed within Scotland prior to the dissemination of the SCOT trial findings, even accounting for co-morbidity and the regimen used.

This study has considered the impact of the SCOT trial in detail. This raises the question regarding the next trial that is needed for patients in this area. The next trial in this space is likely to include patients with colon cancer only. This is because in the decade since the SCOT trial was developed and performed, there has been a divergence in approach to managing colon versus rectal cancer and in particular, there has been a shift in relation to rectal cancer to giving more treatments in the neoadjuvant setting (331-333). The results from Chapters 5-8 showed that many patients with stage II disease or those aged over 70 receive single agent chemotherapy in the adjuvant setting. Although SCOT provided evidence for using a shorter duration of doublet treatment, these results do not provide evidence to support using a shorter duration of monotherapy. If a trial comparing treatment duration for monotherapy was being designed, it would be another opportunity (as was performed for SCOT), to collect patient tumour samples and to perform detailed molecular profiling of tumours. It may also be an opportunity to use a cross-sectional trial design, as was performed for one of the IDEA collaboration trials (264) to test the addition of other treatments in this

context. Lastly, Chapter 5 indicated that the molecular characteristics of CRC tumours, in particular mismatch repair protein deficiency, are important determinants of treatment decisions. Recent evidence has shown that MSI-high tumours of all types respond well to immune checkpoint blockade (334) (335). There is a current UK trial that is assessing the addition of a checkpoint inhibitor immunotherapy to adjuvant chemotherapy for patients with MSI-high stage III disease (336). This trial is already using 3 months duration of treatment in the standard of care treatment schedule.

9.8 Final reflections

Performing cancer trials is a major industry that requires huge financial investment, as well as input from cancer patients. Going forward, research funders are going to be even more judicious in their investment decisions. There cannot be a randomised clinical trial to answer every uncertainty in clinical practice, therefore if cancer trials are being performed, it is imperative that we learn as much as possible from the process of conducting each trial, as well as from the trial results. Ultimately, a major reason for doing research impact evaluation is to make better decisions in future; decisions around what research to do, how it should be performed and prioritised, how the results can be interpreted and used, and what we can learn from completed trials that will improve decisions for trials in development.

In order for the findings of this study to be impactful, the messages outlined in this discussion need to reach policy makers, researchers, patients, and the cancer trial community. These stakeholders need to agree that impact assessment is worthwhile and invest resources at an institutional level to adopt cancer trial impact evaluation. This study demonstrated there was a lack of impact assessment tools relevant to cancer clinical trials. Instrumental impacts from this study will arise through the utilisation by stakeholders of the framework relevant to cancer clinical trials provided above (modified Payback Framework) and/or the list of indicators of cancer trial impact provided in the content analysis coding manual. Further value will be realised if impact evaluation undertaken using these approaches affects future trial design and/or if impact assessments highlight when

completed trials are not as impactful as expected, and subsequent actions are taken to maximise impact. Another important direct impact from this study is the access to a novel administrative dataset in Scotland that can now be used to answer other important research questions on a national scale.

Conceptual impacts from the findings of this study may include opening the dialogue around how impact evaluation is relevant to cancer clinical trials and laying the foundation to develop this concept specific to cancer research. This study will demonstrate to UK researchers that impact evaluation does not need to be only relevant to the REF, and there is the potential to take greater control of the impact agenda going forward. Bringing wider impact further up the agenda for CTUs, researchers and funders may ultimately refocus the lens through which trial success is viewed and the Payback Framework modified to increase relevance to cancer trials provides an effective tool through which to carry out such impact evaluation.

Appendices

10 Appendix 1 CRC chemotherapy regimens

Table 10-1 Chemotherapy regimens used to treat CRC in the adjuvant setting Abbreviations: FU, fluoropyrimidine; LV, leucovorin; OX, oxaliplatin; m, modified; mg, milligram; m2 meter squared; BD, twice a day; PO, per oral; CRC, colorectal cancer.

Name	Type	FU	LV	Ox	Timing
CAPOX	Oral FU (SCOT trial)	Capecitabine 1000mg/m2 BD PO 14 days every 21 days	NA	130mg/m^2	3 weekly
Mayo Clinic or Bossett	Bolus FU	Bolus FU 425mg/m2	20mg/m2 (Low dose)	Nil	Monthly (5 days every month)
Roswell Park	Bolus FU	500mg/m2 bolus	500mg/m2 (High dose)	Nil	Weekly (For 6 out of 8 weeks)
FLOX	Bolus FU	Roswell Park (weekly)	Roswell Park (High dose)	85mg/m^2	Weekly (For 6 out of 8 weeks)
FU infusion	INFUSIONAL FU (SAFFA TRIAL 12 week arm)	5-FU 300mg/m^2/day	Not specified	Nil	Over 12 weeks
FU5LV2 (DeGramont)	Infusional FU	400mg/m^2 bolus + 600mg/m^2 22 hour continuous infusion	LV 200mg/m^2 over 2hours	Nil	2 weekly
FOLFOX4	Infusional FU	400mg/m^2 bolus + 600mg/m^2 22 hour continuous infusion	LV 100mg/m^2	85mg/m^2 oxaliplatin	2 weekly
mFOLFOX6	Infusional and bolus FU	400mg/m^2 bolus + 2400mg/m^2 infusion via 46 hours continuously	200mg/m^2 IV infusion over 2 hours	85mg/m^2 oxaliplatin	2 weekly
mFOLFOX6 (SCOT trial)	Infusional and oral FU	400mg/m^2 bolus + 2400mg/m^2 infusion via 46 hours continuously	L-folinic acid 175mg or Folinic acid 350mg	85mg/m^2	2 weekly

11 Appendix 2 Review Search Terms

11.1 Search Terms for systematic review

Keywords:
Impact (title)
AND (framework* OR pathway* OR tool* OR measur* OR categor* OR demonstr* OR evaluat* OR method* OR model* OR metric* OR assess*)
Research (title)

MeSH terms:
Research, Biomedical Research, Health Services Research.

11.2 Search strategy Ovid Medline

Table 11-1 Search terms Ovid Medline

1	impact.ti.
2	(framework* or pathway* or tool* or measur* or categor* or demonstr* or evaluat* or model* or method* or metric* or assess*).ti.
3	1 and 2
4	exp Research/
5	exp Biomedical Research/
6	exp Health Services Research/
7	research.ti.
8	4 or 5 or 6 or 7
9	3 and 8
10	limit 9 to (english language and yr="1998 -Current")

12 Appendix 3 Supplementary survey material

Table 12-1 Clinical scenarios included in both surveys

	Survey 1 April 2019	Survey 2 August 2020
Patients aged under 70	Colorectal cancer	Colon Cancer
1	T3N1	T3N1
2	T4N1	T4N1
3	T3N2	T3N2
4	T4N2	T4N2
5	T3/4 MSS	T3MSS
6	T3/4 MSI-H	T4MSS
7	-	T3 MSI-H
8	-	T4 MSI-H
Patients aged 70 years and over	Colorectal cancer	Colon Cancer
	Same six scenarios as above	Same eight scenarios as above
Post COVID enduring changes	Colorectal cancer	Colon Cancer
	-	Same 16 scenarios as above (8 patients aged under 70, 8 patients aged 70 and over).

Table 12-2 Locations of respondents to first survey

Location	Number (%)
England	89 (34)
Scotland	16 (6)
Northern Ireland	6 (2)
Wales	4 (2)
UK (Unknown nation)	26 (10)
United States	35 (13)
Japan	25 (9)
Australia	19 (7)
Italy	11 (4)
Spain	6 (2)
France	6 (2)
Denmark	6 (2)
Sweden	3 (1)
Netherlands	3 (1)
Germany	3 (1)
Morocco	1 (<1)
Singapore	1 (<1)
New Zealand	1 (<1)
Greece	1 (<1)
Canada	1 (<1)
Brazil	1 (<1)
Argentina	1 (<1)

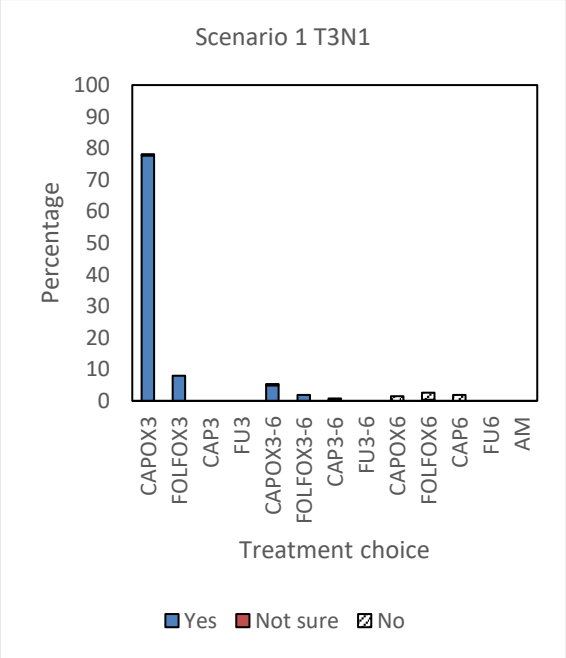
Table 12-3 Practice details for respondents to first survey

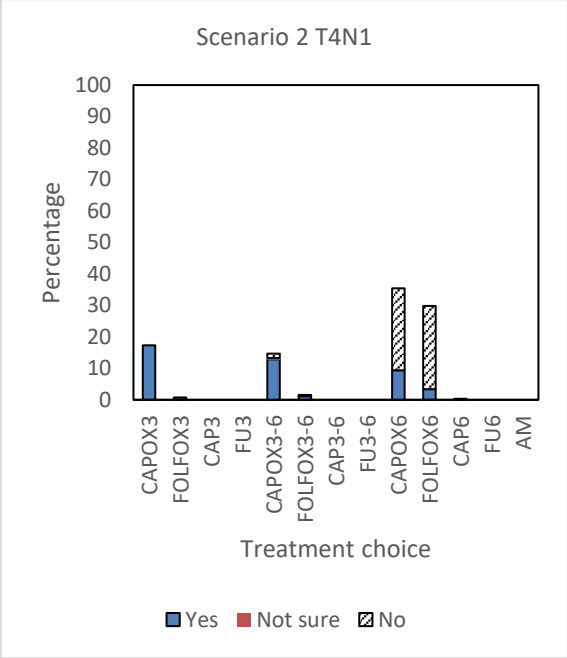
Practice details	UK n=141 Number (%)	International n=124 Number (%)	Overall n=265 Number (%)
Clinical oncology	75 (53%)	11 (9%)	86 (32%)
Medical oncology	65 (46%)	106 (85%)	171 (65%)
Radiation oncology	1 (1%)		1 (0%)
Surgeon		6 (5%)	6 (2%)
Nurse practitioner		1 (1%)	1 (0%)
Type of practice			
Clinical academic (public and private)	4 (3%)	23 (19%)	27 (10%)
Clinical academic (public only)	12 (9%)	48 (39%)	60 (23%)
Clinical academic (private only)		9 (7%)	9 (3%)
Health service (public and private)	46 (33%)	8 (6%)	54 (20%)
Health service (public only)	79 (56%)	28 (22%)	107 (40%)
Health service (private only)		8 (6%)	8 (3%)
Duration of practice			
<2 years	3 (2%)	4 (3%)	7 (3%)
2-5 years	11 (8%)	12 (10%)	23 (9%)
6-9 years	27 (19%)	12 (10%)	39 (15%)
10-20 years	66 (47%)	46 (37%)	112 (42%)
>20 years	34 (24%)	50 (40%)	84 (32%)
Time spent treating patients with CRC			
Only CRC	10 (7%)	15 (12%)	25 (9%)
Predominantly CRC	109 (77%)	81 (65%)	190 (72%)
Minority CRC	22 (16%)	28 (23%)	50 (19%)

Table 12-4 Location and practice details for respondents of the follow- up survey

Characteristics	Overall n=106 Number (%)
Location	
Australia	6 (6%)
Brazil	1 (1%)
Canada	1 (1%)
Denmark	3 (3%)
France	3 (3%)
Germany	2 (2%)
Greece	1 (1%)
Italy	3 (3%)
Japan	3 (3%)
Netherlands	2(2%)
Singapore	1(1%)
Spain	3 (3%)
Sweden	2 (2%)
UK	64 (60%)

USA	11 (10%)
Speciality	
Clinical oncology	38 (36%)
Medical oncology	67 (63%)
Surgeon	1 (1%)
Type of practice	
Clinical academic (public and private)	9 (8%)
Clinical academic (public only)	27 (25%)
Clinical academic (private only)	4 (4%)
Health service (public and private)	16 (15%)
Health service (public only)	48 (45%)
Health service (private only)	2 (2%)
Duration of practice	
<2 years	2 (2%)
2-5 years	10 (9%)
6-9 years	12 (11%)
10-20 years	45 (42%)
>20 years	37 (35%)
Time spent treating patients with CRC	
Only CRC	10 (9%)
Predominantly CRC	83 (78%)
Minority CRC	13 (12%)





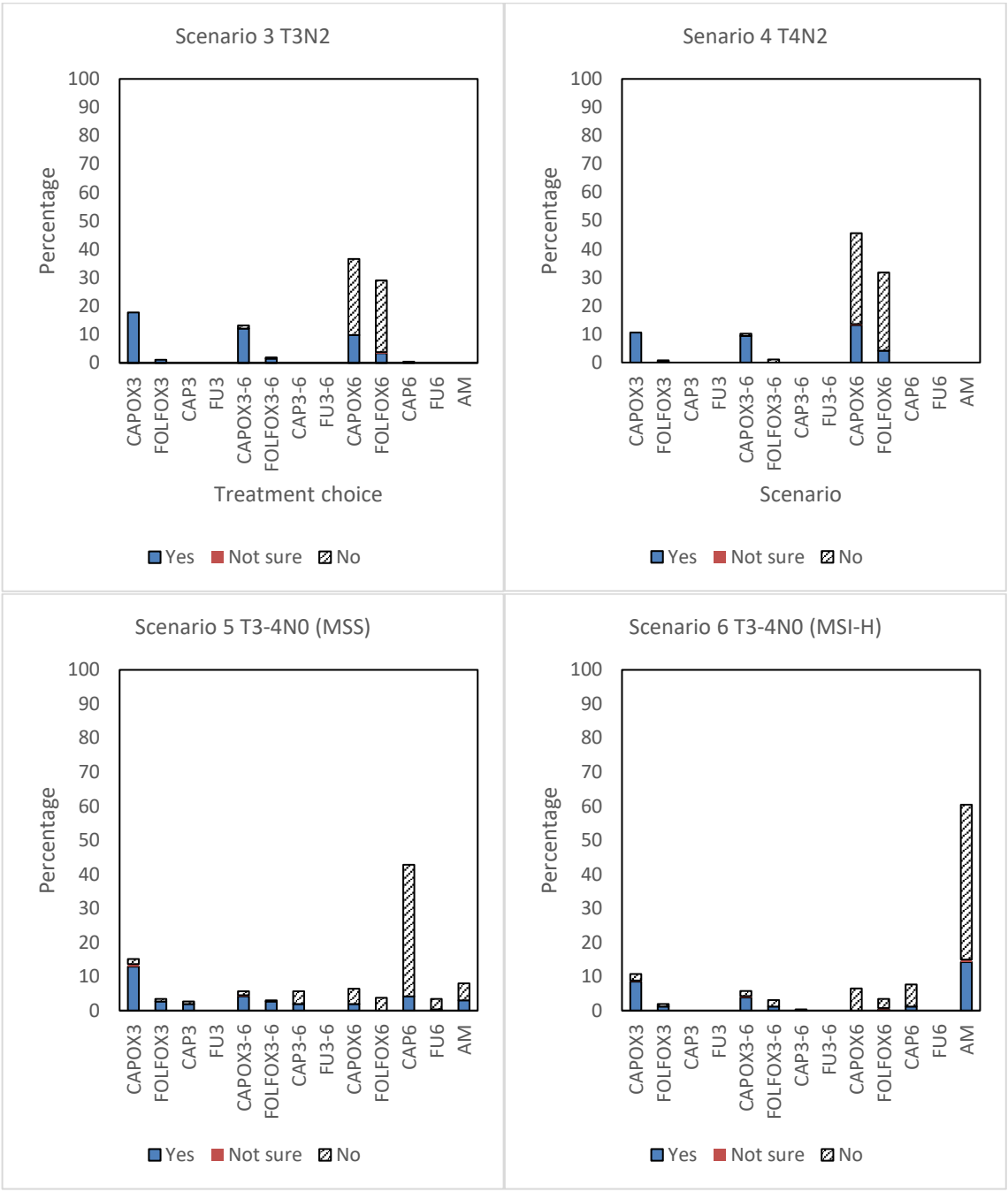
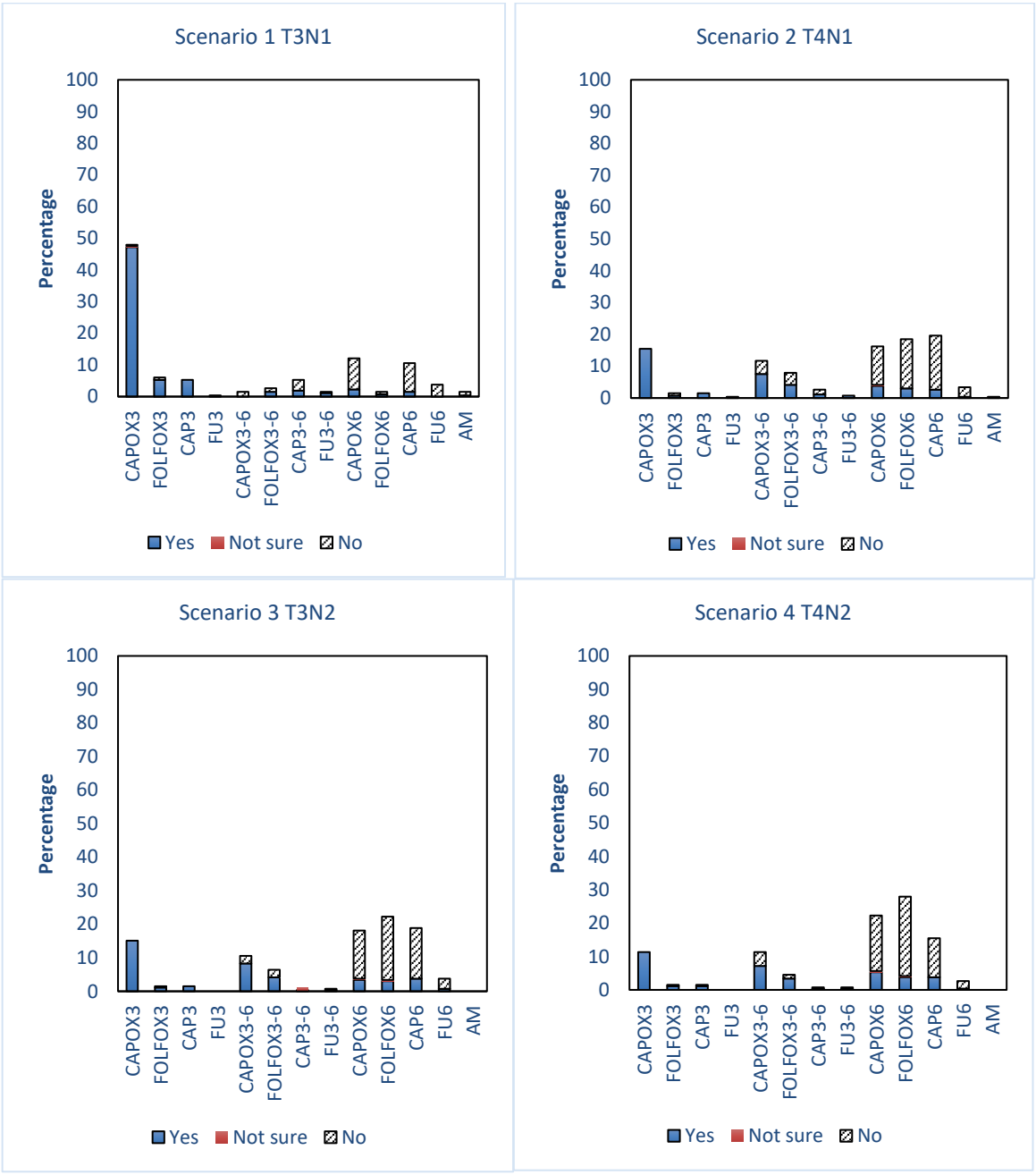


Figure 12-1 Regimen and duration of treatment chosen for six scenarios describing patients aged under 70 years old The colour of the bars indicates the proportion of respondents who indicated their treatment choice represented a change in practice in response to clinical trials (blue), no change in practice (grey) or if they were unsure (orange).


Scenario 1 T3N1
Percentage
100
90
80
70
60
50
40
30
20
10
0
CAPOX3
FOLFOX3
CAP3
FU3
CAPOX3-6
FOLFOX3-6
CAP3-6
FU3-6
CAPOX6
FOLFOX6
CAP6
FU6
AM
Yes
Not sure
No
Scenario 2 T4N1
Percentage
100
90
80
70
60
50
40
30
20
10
0
CAPOX3
FOLFOX3
CAP3
FU3
CAPOX3-6
FOLFOX3-6
CAP3-6
FU3-6
CAPOX6
FOLFOX6
CAP6
FU6
AM
Yes
Not sure
No
Scenario 3 T3N2
Percentage
100
90
80
70
60
50
40
30
20
10
0
CAPOX3
FOLFOX3
CAP3
FU3
CAPOX3-6
FOLFOX3-6
CAP3-6
FU3-6
CAPOX6
FOLFOX6
CAP6
FU6
AM
Yes
Not sure
No
Scenario 4 T4N2
Percentage
100
90
80
70
60
50
40
30
20
10
0
CAPOX3
FOLFOX3
CAP3
FU3
CAPOX3-6
FOLFOX3-6
CAP3-6
FU3-6
CAPOX6
FOLFOX6
CAP6
FU6
AM
Yes
Not sure
No

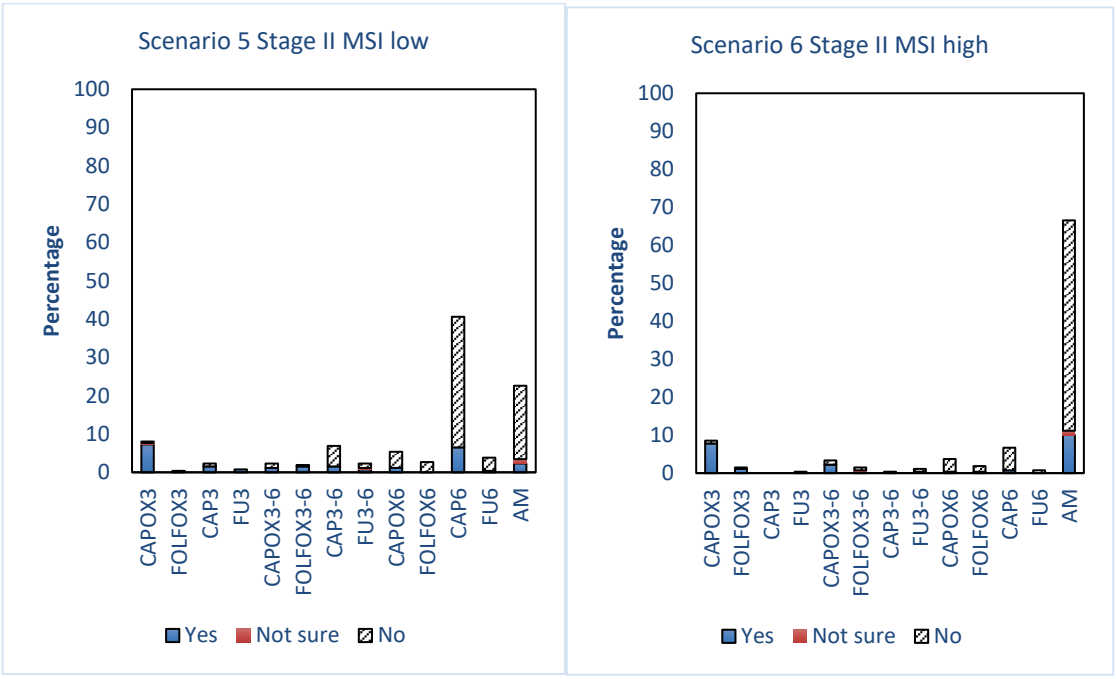


Figure 12-2 Regimen and duration of treatment chosen for six scenarios describing patients aged 70 and over The colour of the bars indicates the proportion of respondents who indicated their treatment choice represented a change in practice in response to clinical trials (blue), no change in practice (dashed) or if they were unsure (orange).

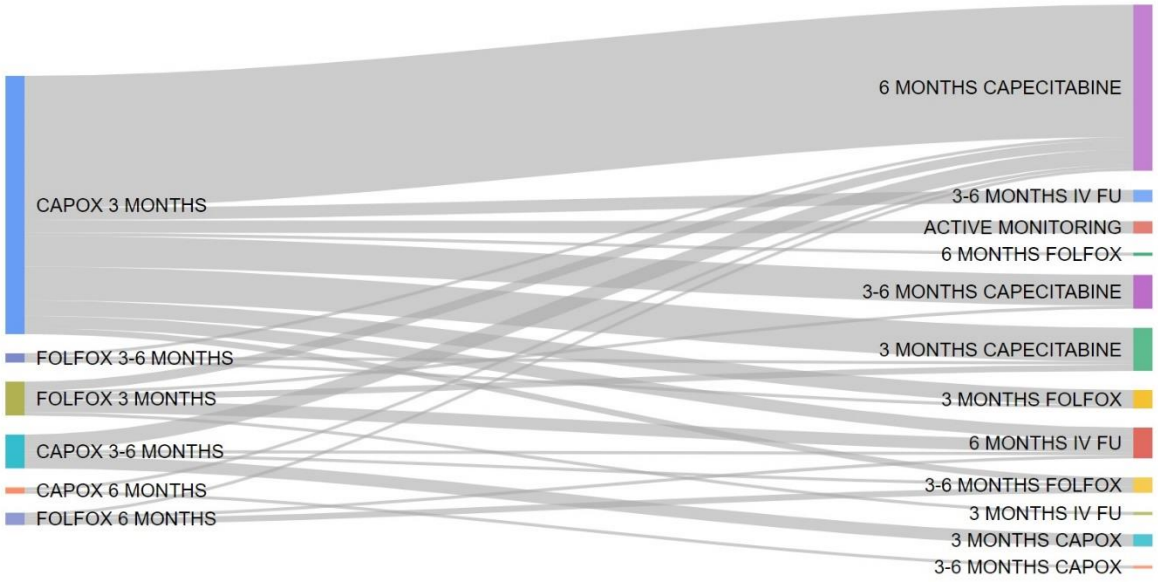


Figure 12-3 Scenario 1: T3N1 disease. Details of regimens used by the clinicians (43% of total) who selected different treatment regimens for patients dependent on age Responses for those clinicians who did not change their mind are not shown. Patients under 70 years old (left) versus over 70 years old (right). The most common change in practice for low risk stage III patients was from doublet chemotherapy for 3 months in younger patients to capecitabine monotherapy for 6 months in older patients. The colours of the bars within this diagram are not significant.

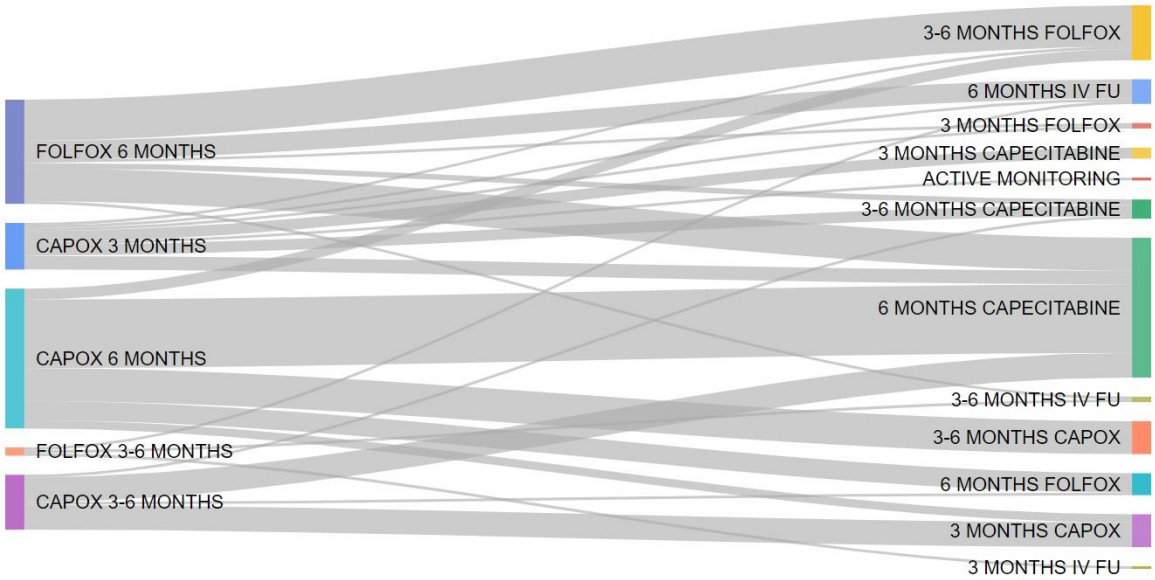


Figure 12-4 Scenario 2: T4N1 disease. Details of regimens used by the clinicians (49% of total) who selected different treatment regimens for patients dependent on age. Responses for those clinicians who did not change their mind are not shown. Patients under 70 years old (left) versus over 70 years old (right). For high-risk stage III disease, the biggest change was from 6 months of doublet chemotherapy to 6 months of capecitabine. The colours of the bars within this diagram are not significant.

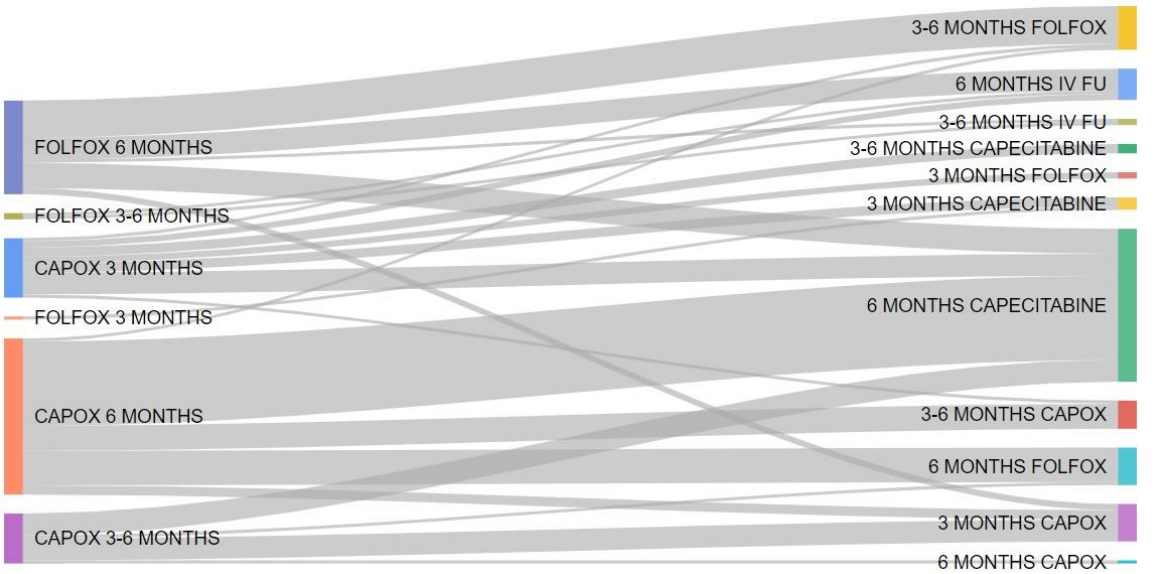


Figure 12-5 Scenario 3: T4N1 disease. Details of regimens used by the clinicians (44% of total) who selected different treatment regimens for patients dependent on age. Responses for those clinicians who did not change their mind are not shown. Patients under 70 years old (left) versus over 70 years old (right). For high-risk stage III disease, the biggest change was from 6 months of doublet chemotherapy to 6 months of capecitabine. The colours of the bars within this diagram are not significant.

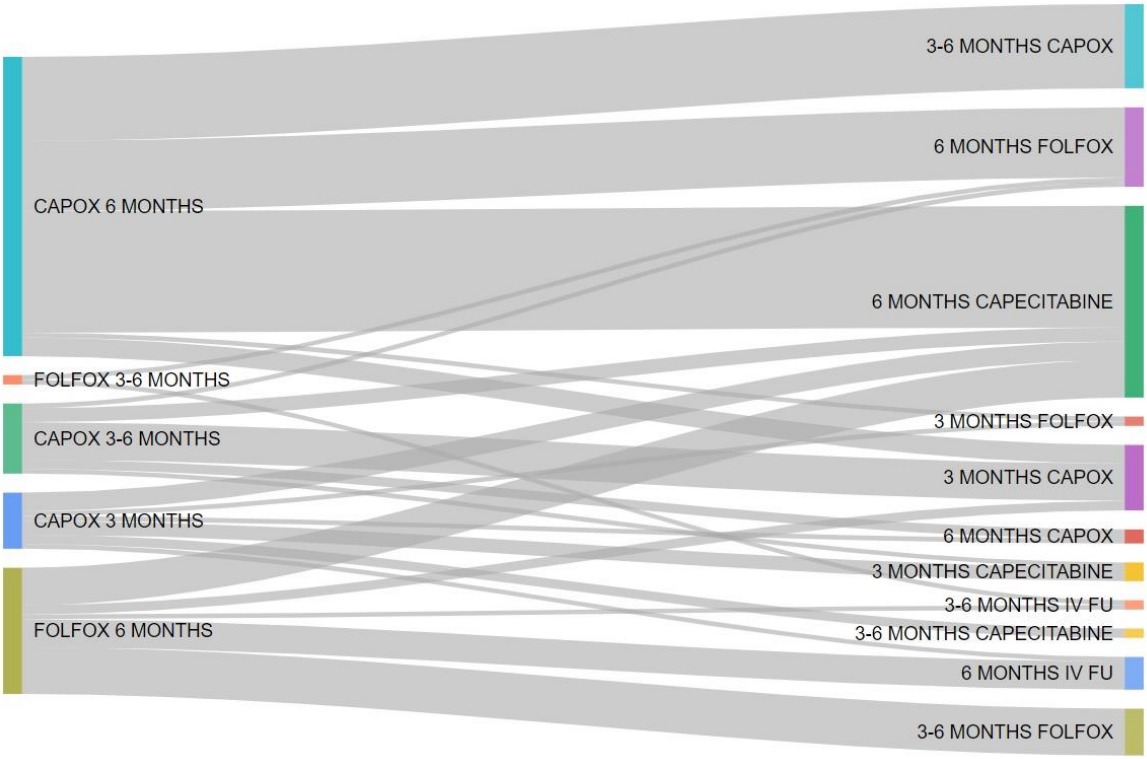


Figure 12-6 Scenario 4: T4N2 disease. Details of regimens used by the clinicians (45% of total) who selected different treatment regimens for patients dependent on age. Responses for those clinicians who did not change their mind are not shown. Patients under 70 years old (left) versus over 70 years old (right). For high-risk stage III disease, the biggest change was from 6 months of doublet chemotherapy to 6 months of capecitabine. The colours of the bars within this diagram are not significant.

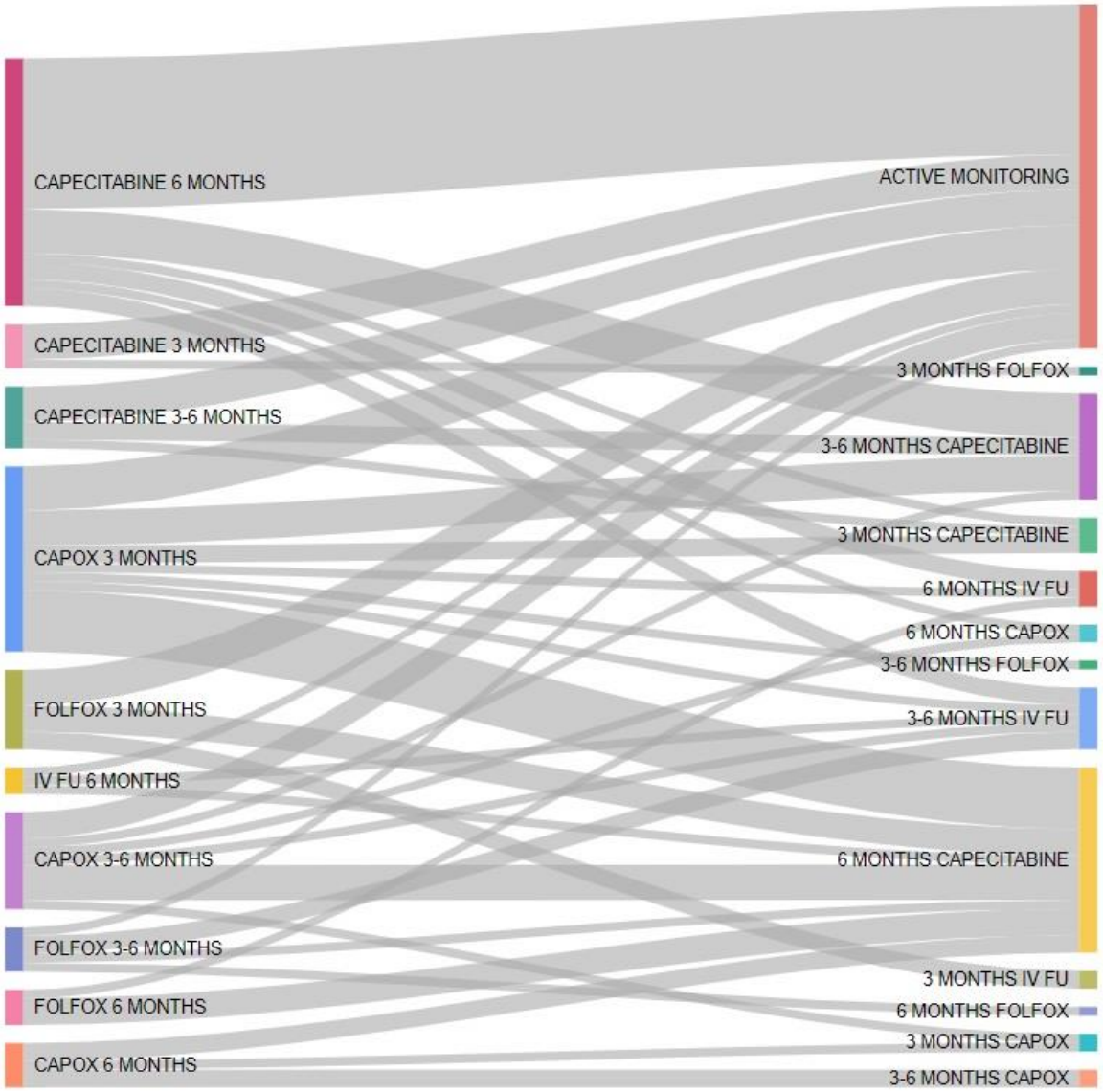


Figure 12-7 Scenario 5: T3-4N0 MSI Low disease. Details of regimens used by the clinicians (38% of total) who selected different treatment regimens for patients dependent on age. Responses for those clinicians who did not change their mind are not shown. Patients under 70 years old (left) versus over 70 years old (right). For stage II disease, the switch was most often from 6 months of capecitabine for younger patients to active monitoring for patients aged 70 and over. The colours of the bars within this diagram are not significant.

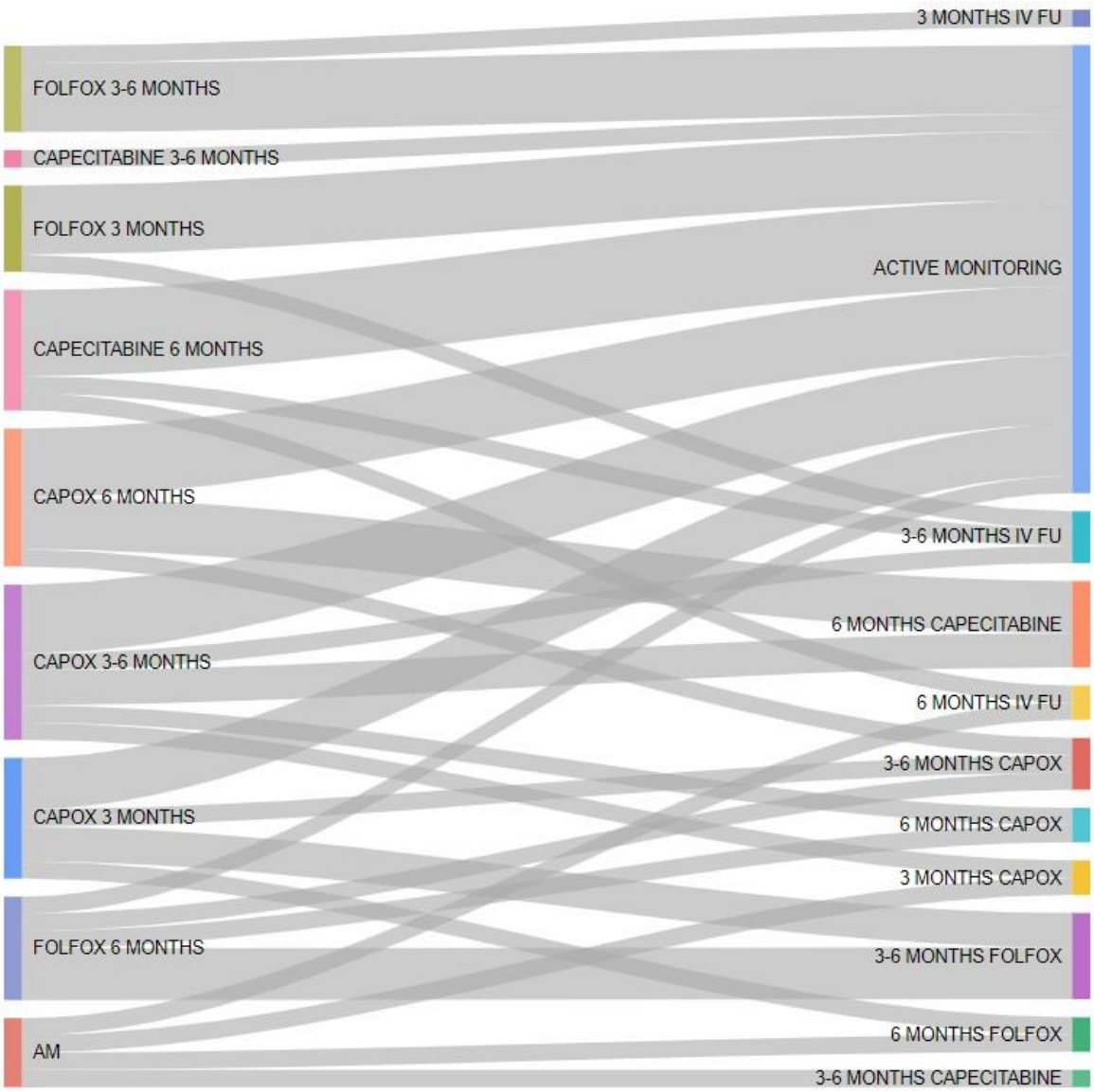
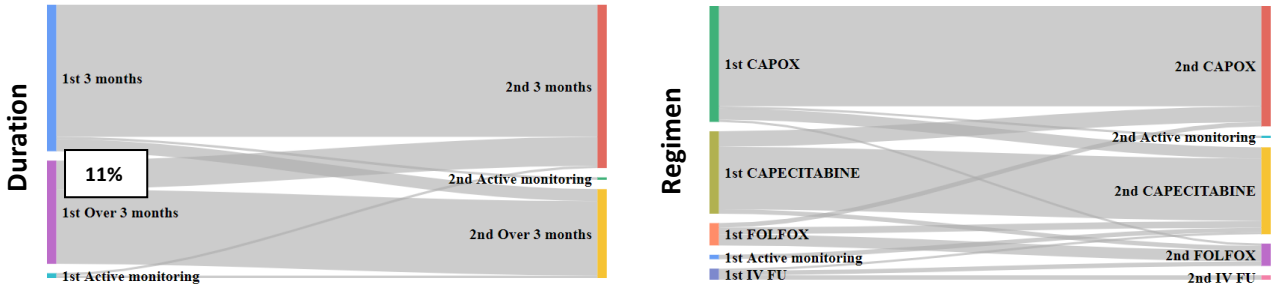
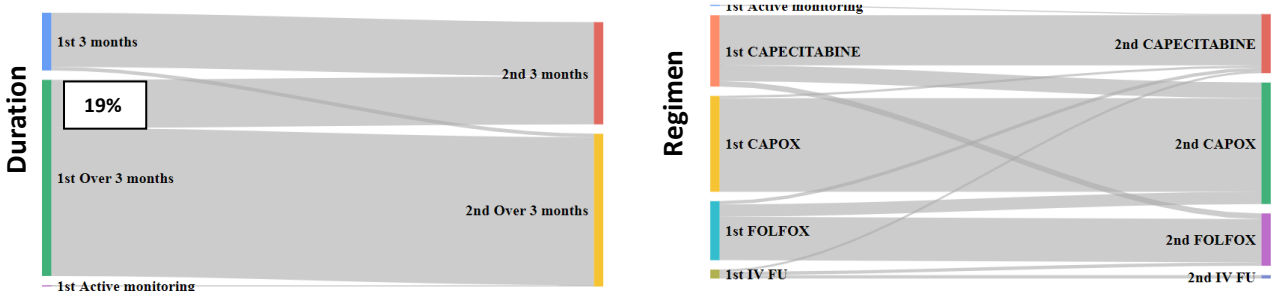


Figure 12-8 Scenario 6: T3-4N0 MSI high disease. Details of regimens used by the clinicians (28% of total) who selected different treatment regimens for patients dependent on age. Responses for those clinicians who did not change their mind are not shown. Patients under 70 years old (left) versus over 70 years old (right). For stage II disease, the switch was most often from 6 months of capecitabine for younger patients to active monitoring for patients aged 70 and over. The colours of the bars within this diagram are not significant.

Low risk stage III (≥ 70 years)



High risk stage III T4 or N2 (≥ 70 years)



High risk stage III T4 and N2 (≥ 70 years)

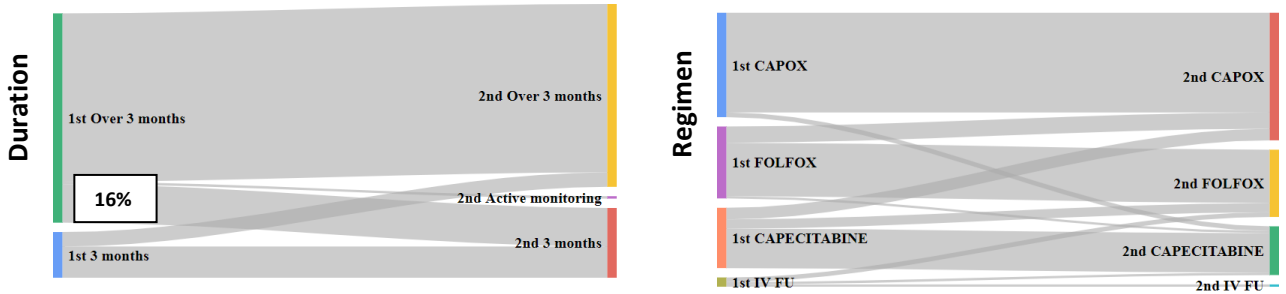


Figure 12-9 Switch in individual treatment choices for scenarios describing patients with stage III disease aged 70 and over The left side of each graph displays the treatment choices in April 2019. The right side of the graphs depict the treatment choices for the same group of clinicians (n=106) in August 2020. If an individual has changed their choice this is shown by a diagonal, rather than straight connecting grey line. The colours of the bars within this diagram are not significant.

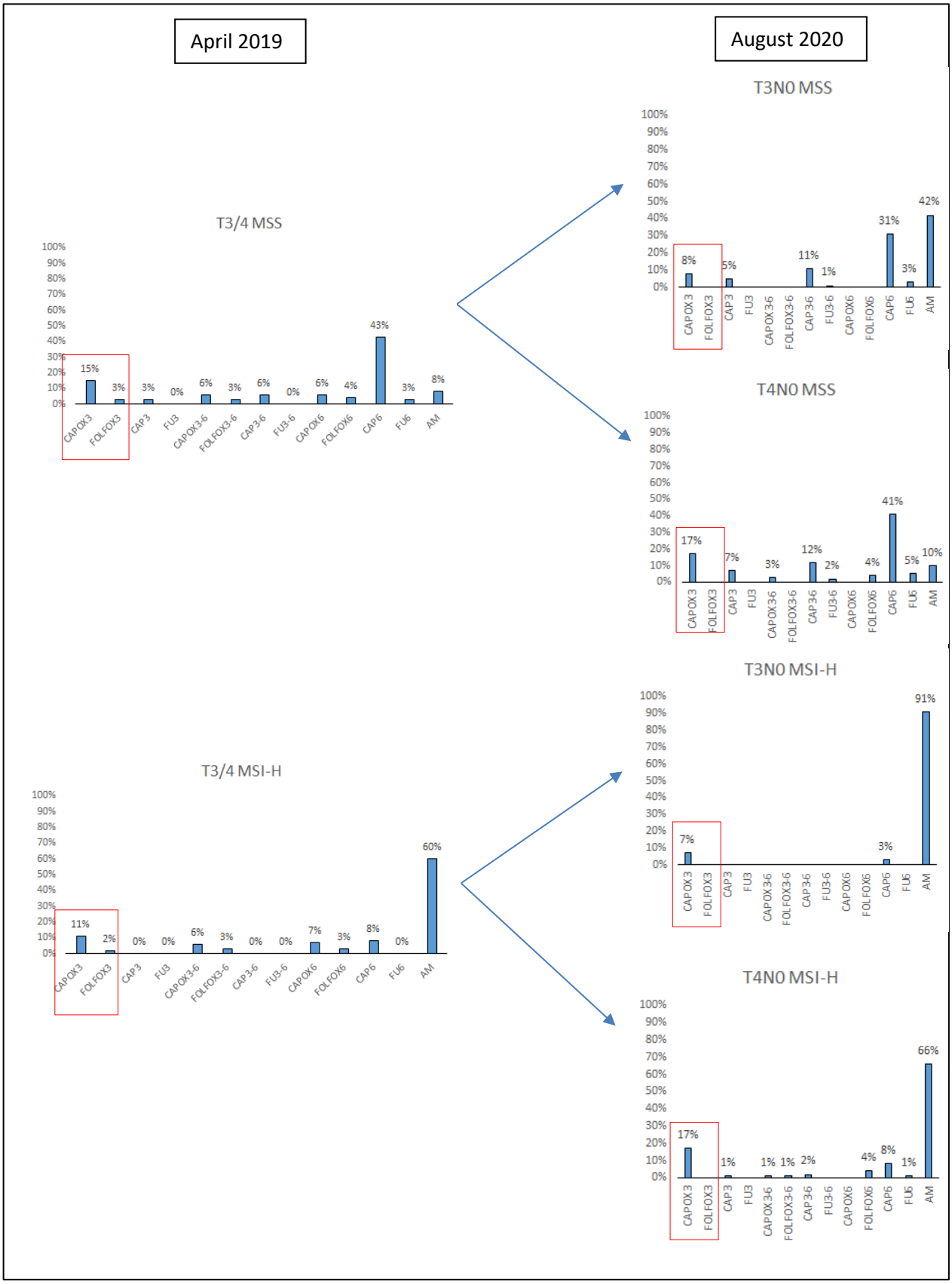


Figure 12-10 Treatment choices for patients aged 70 and over with stage II disease with a comparison of clinician choices between April 2019 and August 2020 (n=106) For patients with MSS disease, in August 2020 most clinicians still chose 6 months of capecitabine (31%) or active monitoring (42%) for T3 disease. However, the use of active monitoring increased (42% versus 8%) and the use of capecitabine decreased (43% vs 31%) compared to responses from the same clinicians to the T3/4 MSS scenario in April 2019. Overall, 8% of clinicians chose 3 months of CAPOX and none chose over 3 months of CAPOX to treat T3 MSS disease. Also no clinicians chose to use FOLFOX of any duration in August 2020.

Regarding patients with T4 MSS disease, the commonest treatment choice in August 2020 was 6 months of capecitabine (41%). There was a slight increase in the use of CAPOX (15% to 17%) and a slight increase in active monitoring (8% to 10%) compared to the responses for T3/4 MSS disease in April 2019. Only 7% of respondents indicated they use over 3 months of doublet chemotherapy in this situation (3% CAPOX and 4% FOLFOX). Regarding MSI-H disease, the most common choice both in April 2019 (60%) and August 2020 (T3: (91%) T4 66%), was active monitoring. At both time-points, there was a higher use of doublet chemotherapy compared to fluoropyrimidine monotherapy for both T3 and T4 disease. The red boxes indicate when responses align with the treatment delivered in the experimental arm of the SCOT trial/IDEA collaboration.

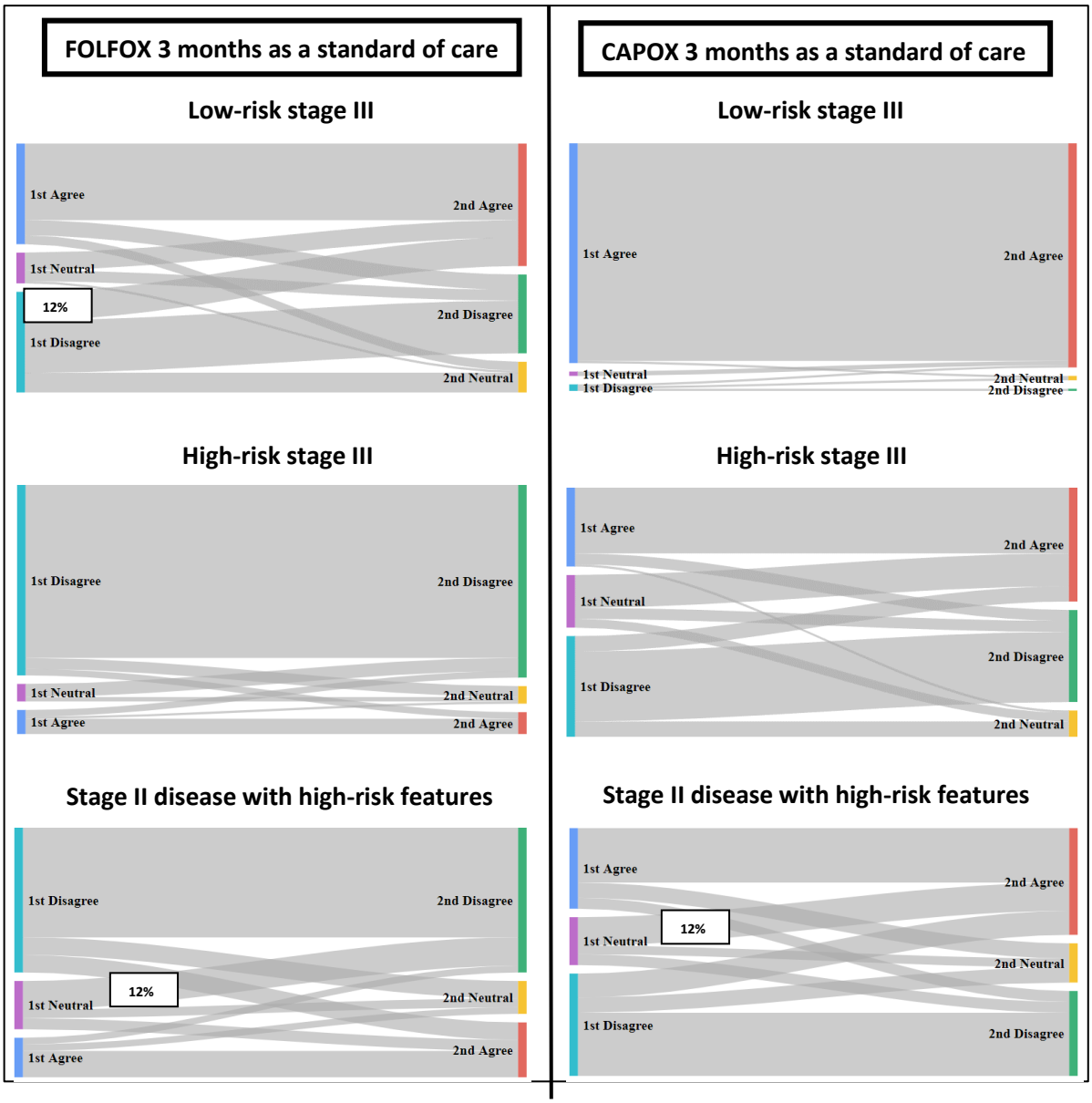


Figure 12-11 Change in individual clinician opinions between April 2019-August 2020 for the group of clinicians who answered both surveys (n=106). If an individual has changed their choice this is shown by a diagonal, rather than straight connecting grey line. Any changes >10% are highlighted. Low-risk stage III (SCOT non-inferiority met), high-risk stage III (SCOT non-inferiority not met), stage II disease with high-risk features (SCOT non-inferiority not met). The colours of the bars within this diagram are not significant.

Table 12-5 Change in clinician (n=106) preferences in August 2020. Disregarding (pre) and taking into account (post) effects from the COVID-19 pandemic.

	Under 70 Pre-COVID	Post-COVID	Fisher's		70 and over Pre-COVID	Post-COVID	Fisher's
T1-3N1				T1-3N1			
Capox 3m	96	96	1	Capox 3m	54	53	1
Capox 3-6m	2	2	1	Capox 3-6m	0	0	NA
Capox 6m	0	0	NA	Capox 6m	0	0	NA
Folfox 3m	5	4	1	Folfox 3m	9	8	1
Folfox 3-6m	1	0	1	Folfox 3-6m	0	0	NA
Folfox 6m	1	1	1	Folfox 6m	1	0	NA
Capecitabine 3m	0	1	1	Capecitabine 3m	2	2	1
Capecitabine 3-6m	0	1	1	Capecitabine 3-6m	1	6	0.119
Capecitabine 6m	1	1	1	Capecitabine 6m	26	23	1
FU 3m	0	0	NA	FU 3m	0	0	NA
FU 3-6m	0	0	NA	FU 3-6m	0	0	NA
FU 6m	0	0	NA	FU 6m	2	2	1
AM	0	0	NA	AM	1	6	0.119
T4N1				T4N1			
Capox 3m	51	55	0.583	Capox 3m	36	37	1
Capox 3-6m	17	20	0.718	Capox 3-6m	9	11	0.815
Capox 6m	14	13	1	Capox 6m	7	3	0.332
Folfox 3m	3	2	1	Folfox 3m	3	2	1
Folfox 3-6m	2	3	1	Folfox 3-6m	8	10	0.806
Folfox 6m	19	13	0.338	Folfox 6m	11	8	0.632
Capecitabine 3m	0	0	NA	Capecitabine 3m	4	6	0.748
Capecitabine 3-6m	0	0	NA	Capecitabine 3-6m	6	9	0.594
Capecitabine 6m	0	0	NA	Capecitabine 6m	20	18	0.858
FU 3m	0	0	NA	FU 3m	0	0	NA
FU 3-6m	0	0	NA	FU 3-6m	0	0	NA
FU 6m	0	0	NA	FU 6m	2	1	1
AM	0	0	NA	AM	0	1	1
T3N2				T3N2			
Capox 3m	45	54	0.271	Capox 3m	39	37	0.886
Capox 3-6m	18	18	1	Capox 3-6m	10	11	1
Capox 6m	21	13	0.19	Capox 6m	8	5	0.569
Folfox 3m	0	0	NA	Folfox 3m	1	1	1
Folfox 3-6m	2	3	1	Folfox 3-6m	7	6	1
Folfox 6m	20	18	0.858	Folfox 6m	17	14	0.698

Capecitabine 3m	0	0	NA	Capecitabine 3m	2	7	0.17
Capecitabine 3-6m	0	0	NA	Capecitabine 3-6m	4	7	0.538
Capecitabine 6m	0	0	NA	Capecitabine 6m	17	17	1
FU 3m	0	0	NA	FU 3m	0	0	NA
FU 3-6m	0	0	NA	FU 3-6m	0	0	NA
FU 6m	0	0	NA	FU 6m	1	1	1
AM	0	0	NA	AM	0	0	NA
T4N2				T4N2			
Capox 3m	25	35	0.129	Capox 3m	26	29	0.754
Capox 3-6m	22	20	0.863	Capox 3-6m	5	11	0.192
Capox 6m	31	23	0.27	Capox 6m	14	11	0.671
Folfox 3m	0	0	NA	Folfox 3m	1	1	1
Folfox 3-6m	2	4	0.683	Folfox 3-6m	8	8	1
Folfox 6m	26	24	0.749	Folfox 6m	19	18	1
Capecitabine 3m	0	0	NA	Capecitabine 3m	2	5	0.445
Capecitabine 3-6m	0	0	NA	Capecitabine 3-6m	4	5	1
Capecitabine 6m	0	0	NA	Capecitabine 6m	15	17	0.848
FU 3m	0	0	NA	FU 3m	0	0	NA
FU 3-6m	0	0	NA	FU 3-6m	0	0	NA
FU 6m	0	0	NA	FU 6m	1	1	1
AM	0	0	NA	AM	0	0	NA
T3N0 MSS				T3N0 MSS			
Capox 3m	17	18	1	Capox 3m	8	8	1
Capox 3-6m	1	0	1	Capox 3-6m	0	0	NA
Capox 6m	0	1	1	Capox 6m	0	0	NA
Folfox 3m	0	0	NA	Folfox 3m	0	0	NA
Folfox 3-6m	0	0	NA	Folfox 3-6m	0	0	NA
Folfox 6m	0	0	NA	Folfox 6m	0	0	NA
Capecitabine 3m	5	7	0.768	Capecitabine 3m	5	4	1
Capecitabine 3-6m	10	12	0.822	Capecitabine 3-6m	12	12	1
Capecitabine 6m	40	33	0.386	Capecitabine 6m	33	29	0.651
FU 3m	0	0	NA	FU 3m	0	0	NA
FU 3-6m	0	0	NA	FU 3-6m	1	0	1
FU 6m	3	1	0.621	FU 6m	3	2	0.683
AM	30	34	0.654	AM	44	51	0.407
T3N0 MSI-H				T3N0 MSI-H			
Capox 3m	8	8	1	Capox 3m	7	6	1
Capox 3-6m	0	0	NA	Capox 3-6m	0	0	NA
Capox 6m	0	0	NA	Capox 6m	0	0	NA
Folfox 3m	0	0	NA	Folfox 3m	0	0	NA
Folfox 3-6m	0	0	NA	Folfox 3-6m	0	0	NA
Folfox 6m	0	0	NA	Folfox 6m	0	0	NA

Capecitabine 3m	0	0	NA	Capecitabine 3m	0	0	NA
Capecitabine 3-6m	0	2	0.498	Capecitabine 3-6m	0	0	NA
Capecitabine 6m	2	1	1	Capecitabine 6m	3	1	0.621
FU 3m	0	0	NA	FU 3m	0	0	NA
FU 3-6m	0	0	NA	FU 3-6m	0	0	NA
FU 6m	0	0	NA	FU 6m	0	0	NA
AM	95	93	0.829	AM	96	98	0.806
T4N0 MSS				T4N0 MSS			
Capox 3m	36	33	0.77	Capox 3m	18	20	0.724
Capox 3-6m	8	2	0.101	Capox 3-6m	3	2	1
Capox 6m	1	2	1	Capox 6m	0	0	NA
Folfox 3m	0	0	NA	Folfox 3m	0	0	NA
Folfox 3-6m	2	2	1	Folfox 3-6m	0	1	1
Folfox 6m	5	4	1	Folfox 6m	4	2	0.683
Capecitabine 3m	3	9	0.134	Capecitabine 3m	7	7	1
Capecitabine 3-6m	8	9	1	Capecitabine 3-6m	13	10	0.66
Capecitabine 6m	36	35	1	Capecitabine 6m	43	40	0.778
FU 3m	0	0	NA	FU 3m	0	0	NA
FU 3-6m	1	0	1	FU 3-6m	2	0	1
FU 6m	2	2	1	FU 6m	5	4	1
AM	4	7	0.538	AM	11	20	0.119
T4N0 MSI-H				T4N0 MSI-H			
Capox 3m	29	22	0.353	Capox 3m	18	16	0.852
Capox 3-6m	4	6	0.748	Capox 3-6m	1	1	1
Capox 6m	0	0	NA	Capox 6m	0	0	NA
Folfox 3m	0	0	NA	Folfox 3m	0	0	NA
Folfox 3-6m	3	1	0.621	Folfox 3-6m	1	1	1
Folfox 6m	3	4	1	Folfox 6m	4	2	0.683
Capecitabine 3m	1	1	1	Capecitabine 3m	1	0	1
Capecitabine 3-6m	1	2	1	Capecitabine 3-6m	2	2	1
Capecitabine 6m	6	8	0.783	Capecitabine 6m	8	6	0.783
FU 3m	0	0	NA	FU 3m	0	0	NA
FU 3-6m	0	0	NA	FU 3-6m	0	0	NA
FU 6m	1	1	1	FU 6m	1	1	1
AM	58	59	1	AM	70	77	0.372

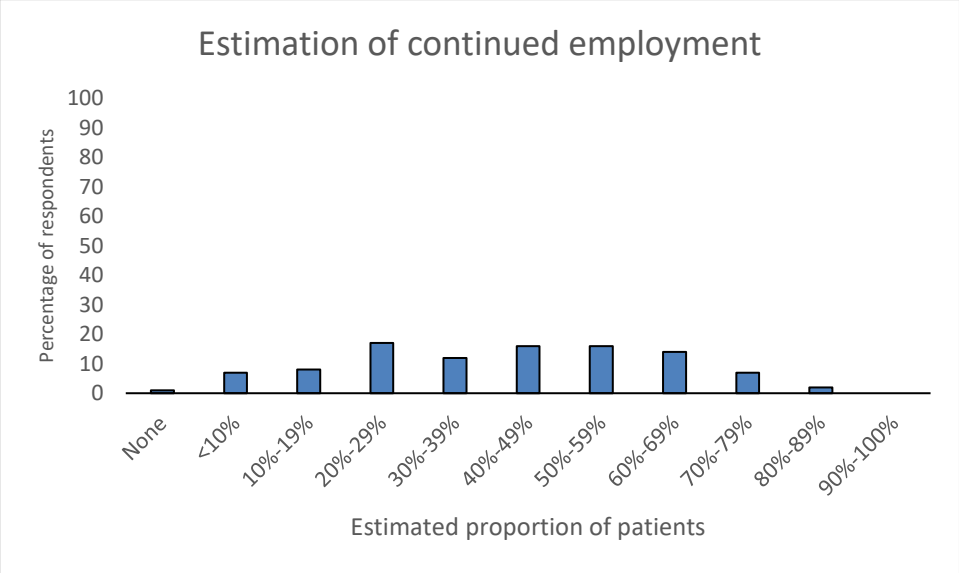


Figure 12-12 Clinicians' estimation of the proportion of patients of working age who continue to work full time during their adjuvant treatment for CRC

13 Appendix 4 Supplementary material for economic evaluation

Table 13-1 Types of economic evaluation. Information adapted from text in (15)

Type of economic evaluation	Description
Cost-effectiveness analysis	Costs are measured in monetary terms whereas consequences are measured in natural units of effectiveness, for example, units of blood pressure reduction, live-years gained. The effect of interest must be common to the alternative treatments being compared.
Cost-utility analysis	Costs are measured in monetary terms but outcomes are measured in utilities, and typically reported as quality-adjusted life years (QALYs). As QALYs are a generic measure of health gain, this allows the cost-effectiveness of treatments within different areas of health care to be compared to assess the opportunity cost of spending on those treatments within a fixed budget. Other generic measures of health gain that can be used include the disability adjusted life year (DALY) or health years equivalent (HYE).
Cost-benefit analysis	Both costs and effects are measured in monetary units. Outcome measures must be converted to monetary values and this may be done using individuals' willingness to pay for those outcomes, for

	example, health gain. CBA allows comparison of the cost-effectiveness of interventions in different disciplines.
Cost-minimisation analysis	Comparison of the costs of alternative interventions. The consequences of the interventions are assumed to be equal.

Table 13-2 Types of multi-country perspective economic evaluations (Adapted from ((112)))

	Source of effectiveness data			
Source of resource data		All participating countries	Sample of countries	Single country
	All participating countries	Fully pooled analysis	Partially pooled analysis	Partially pooled/split analysis
	Sample of countries	Partially pooled analysis	Partially pooled/split analysis	Partially split analysis
	Single country	Partially pooled/split analysis	Partially split analysis	Fully split analysis

Table 13-3 Pros and cons of different types of multi-county economic evaluations (Adapted from ((112)))

Advantage	Fully pooled, multi-country costing	Fully pooled, one-country costing	Partially split, one-country costing	Fully split, one-country costing
Maintains patient-level relationship between resource use and clinical benefits	√	√		√
Maintains patient-level relationship between resource use and costs	√		√	√
Maximises statistical power for treatment effect	√	√	√	
Minimises collection of unit cost data		√	√	√
Allows consistent reporting of treatment effects in economic and clinical manuscripts	√	√	√	

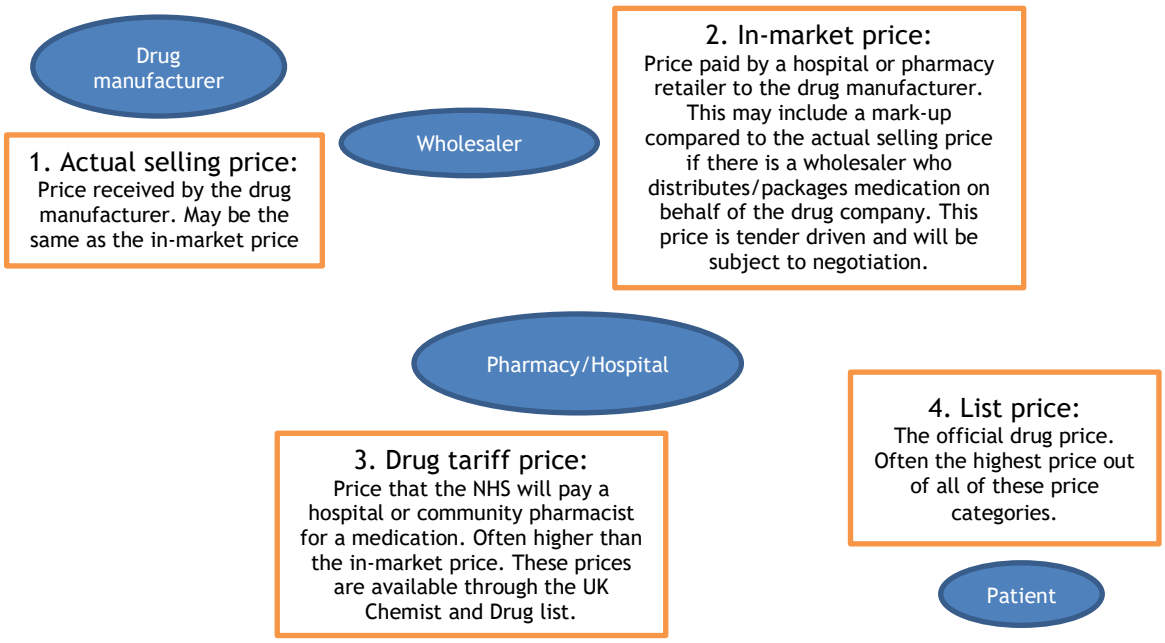


Figure 13-1 An explanation of types of unit costs. This figure was made based on a discussion with a representative from the pharmaceutical industry.

Table 13-4 Types of productivity loss calculation

Productivity costs estimation	Description
Human capital approach	The value of absenteeism at work for patients who are unwell is measured using gross (pre-tax) wage. This can include paid or unpaid work. It is contentious whether this estimation should include time away from work when receiving treatment. The criticism of the human capital approach is that it estimates potential lost production rather than the loss experience in real life which will be mitigated by compensating factors, such as an employee replacing that individual in the workplace (the basis of the friction cost approach).
Friction cost approach	Considers productivity costs from an employer's perspective. The friction period is the time that employers take to recover from the productivity loss from an absent employee. For example, the time to fill this vacancy. Often short term vacancies produce similar estimate to using the human capital approach but for longer term absence, the productivity loss calculated using the friction cost approach will be less than the human capital approach.
Incorporation of productivity costs into health effects	Productivity costs are included in the measure of health effects when, for example, the QALY is used, because it is assumed that individuals will incorporate the impact of health conditions on their ability to work when they value the impact of a health condition on their quality of life. Using

	this approach, it would therefore be double counting if estimating health outcomes using QALYs and also including an estimation of productivity loss separately.

Table 13-5 Time-point of data collection These time-points following the approach taken for the SCOT trial economic analysis (75).

Time-point (Survival to start of the time period)	Cost period (Costs incurred during this time period)
1 (Survival at baseline/time zero)	1 (0-1.5 months)
2 (Up to 1.5 months)	2 (1.5-3 months)
3 (Up to 3 months)	3 (3-4.5 months)
4 (Up to 4 months)	4 (4-5 months)
5 (Up to 5 months)	5 (5-6 months)
6 (Up to 6 months)	6 (6-9 months)
7 (Up to 9 months)	7 (9-12 months)
8 (Start of year 12)	8 (12-18 months)
9 (Up to 18 months)	9 (18-24 months)
10 (Start of year 2)	10 (2-3 years)
11 (Start of year 3)	11 (3-4 years)
12 (Start of year 4)	12 (4-5 years)
13 (Start of year 5)	13 (5-6 years)
14 (Start of year 6)	14 (6-7 years)
15 (Start of year 7)	15 (7-8 years)

16 (Start of year 8)	16 (8-9 years) No actual cost data
17 (Start of year 9)	17 (9-10 years) No actual cost data

Table 13-6 Variables and values used to populate the budget impact framework

Parameter	Value	Source of information
Incident population diagnosed with stage III disease per year	Colon; Rectum	All incidence estimates from Globoan (2018) (337).
Australia	2740; 1215	Stage proportion from Australian Institute of Health and Welfare (https://www.aihw.gov.au/reports/cancer/cancer-data-in-australia/data).(338)
Denmark	817; 407	Stage proportion from Danish report (339).
New Zealand	523; 250	Total number from NZ Ministry of Health. Stage proportions based on Australian source.
Spain	6271; 3268	Stage proportion based on UK source (340).
Sweden	1145; 546	Stage proportion based on Swedish report provided via personal communication with
UK	8100; 3934	Stage proportion from Cancer Research UK Bowel Statistics. (299)
Incident population diagnosed with stage II disease per year	Colon/Rectum	
Australia	2740; 1215	All incidence estimates from Globoan (2018). Stage proportions from same source used for stage III.
Denmark	888; 292	
New Zealand	523; 250	
Spain	6030; 3143	
Sweden	976; 465	
UK	7788; 3784	
Proportion of patients with stage III CRC that receive adjuvant chemotherapy		
Base case	65%	Mean of low (56%) and high (75%) estimate. Also, aligns approximately with: Swedish report (341) 60% colon 2018 Van Steenbergen et al (342) 2011 Netherlands 68% for patients aged 65-74 years. Lima et al 2011 (343) Canada 61%. Upadhyag et al 2015 (344) USA 65%.
Low estimate	55%	Taylor et al 2020 (258) 56% England
High estimate	75%	Boland et al 2013 (345) 74% USA (excluded aged 80+) Ortiz et al 2019 (346) 75% Puerto Rico
Proportion of patients with stage II CRC that receive adjuvant chemotherapy		
Base case	20%	Yang et al 2018 (304, 347) USA 21% Taylor et al 2020 (258) England 16% Swedish report 2018 (341) 20% colon (10% rectum)
Low estimate	10%	Chosen by primary researcher as a reasonable low estimate for the purposes of this budget impact analysis.
High estimate	50%	Fotheringham et al 2019 (348) 50-60% UK
Out of those patients with stage III CRC that receive adjuvant chemotherapy, the proportion who receive doublet treatment.		

Base case	71%	Taylor et al (Table 1) (258).This estimate includes patients of all ages ≥18 years old.
Low estimate	50%	Chosen by primary researcher as a reasonable low estimate for the purposes of this budget impact analysis.
High estimate	100%	Chosen by primary researcher as a reasonable high estimate for the purposes of this budget impact analysis.
Out of those patients with stage II CRC that receive adjuvant chemotherapy, the proportion who receive doublet treatment.		
Base case	31%	Taylor et al (Table 1) (258).This estimate includes patients of all ages ≥18 years old.
Low estimate	15%	Chosen by primary researcher as a reasonable low estimate for the purposes of this budget impact analysis.
High estimate	50%	Chosen by primary researcher as a reasonable high estimate for the purposes of this budget impact analysis.
Percentage of clinicians using 3 months of doublet chemotherapy to treat patients under 70 with stage III CRC post-SCOT trial (i.e. practice change)		
Base case	*54%*	*International survey**
Percentage of clinicians using 3 months of doublet chemotherapy to treat patients aged 70+ with stage III CRC post-SCOT trial (i.e. practice change)		
Base case	*37%*	*International survey***
Overall practice change estimate Stage III		
Base scenario	50%	Data from a real world data set showed that of those patients receiving doublet chemotherapy, 76% were aged under 70 years old and 24% were aged 70 years and over (258). Combining these real world estimates with practice change calculated using an international survey: 0.76 x 54% = 41% 0.24 x 37% = 9% Sum=50%
Low estimate	25%	Chosen by primary researcher as a reasonable low estimate for the purposes of this budget impact analysis.
High estimate	100%	Chosen by primary researcher as a reasonable low estimate for the purposes of this budget impact analysis.
Percentage of clinicians using 3 months of doublet chemotherapy to treat patients under 70 with stage II CRC post-SCOT trial (i.e. practice change if assume that 100% of patients prescribed doublet chemotherapy pre-SCOT receive an intended 6 months of treatment)		
Base case	*20%*	*International survey. ***

Percentage of clinicians using 3 months of doublet chemotherapy to treat patients aged 70+ with stage II CRC post-SCOT trial.		
Base case	*8%*	*International survey. *****
Overall practice change estimate Stage II		
Base scenario	18%	Data from a real world data set showed that of those patients receiving doublet chemotherapy, 76% were aged under 70 years old and 24% were aged 70 years and over (258). Combining these real world estimates with practice change calculated using an international survey: 0.76 x 20% = 15% 0.24 x 8% = 2% Sum=18%
Low estimate	10%	Chosen by primary researcher as a reasonable low estimate for the purposes of this budget impact analysis.
High estimate	30%	Chosen by primary researcher as a reasonable low estimate for the purposes of this budget impact analysis.

*Based on hypothetical patient scenarios for patients aged <70 years old. This calculation accounts for differences in clinician practice for low risk stage III (T1-3N1) and high-risk stage III (T4 or N2) disease. In the survey, 86% clinicians reported they use 3 months of doublet chemotherapy for low-risk stage III disease and 16% of clinicians reported they use 3 months of doublet chemotherapy for high-risk stage III disease. In the SCOT trial (Lancet Oncology publication April 2018 Figure 3), approximately 54% of stage III patients in the trial cohort had low risk disease and 46% had high-risk disease. Therefore, the overall proportion of patients with stage III disease receiving doublet chemotherapy for 3 months post-SCOT = (86% 3 months doublet chemotherapy x 54% patients with low risk stage III CRC) + (16% 3 months doublet chemotherapy x 46% patients with high risk stage III CRC) = 54%.

** Based on hypothetical patient scenarios for patients aged 70+ years old. In the survey, 53% of clinicians reported that they use doublet chemotherapy for 3 months for low-risk stage III disease and 15% of clinicians reported they use 3 months of doublet chemotherapy for high-risk stage III disease. Overall proportion of patients treated with stage III disease receiving doublet chemotherapy for 3 months therefore = (53% 3 months doublet chemotherapy x 56% patients with low risk stage III CRC) + (15% 3 months doublet chemotherapy x 46% patients with high risk stage III CRC) = 37%.

***Based on hypothetical patient scenarios for patients aged <70 years old. Proportion of clinicians using 3 months of doublet chemotherapy post-SCOT was 20% for microsatellite stable disease and 12% for MSI-H disease. Assuming approximately 20% of patients with stage II CRC will have MSI-H disease, overall proportion receiving 3 months of doublet chemotherapy post-SCOT: (0.8 x 20%) + (0.2x12%) = 18%.

**** Based on hypothetical patient scenarios for patients aged <70 years old. Proportion of clinicians using 3 months of doublet chemotherapy post-SCOT was 8% for microsatellite stable disease and 10% for MSI-H disease. Assuming approximately 20% of patients with stage II CRC will have MSI-H disease, overall proportion receiving 3 months of doublet chemotherapy post-SCOT: (0.8 x 8%) + (0.2 x 10%) = 8%.

Table 13-7 Utility regression (Australia)

Australia utilities		
Variable	Co-efficient	S.E.
Health states (ref: disease free)		
On treatment	-0.0378*	0.004
Recurrence	-0.0644*	0.014
Arm: 6 months (ref: 3 months)	-0.0144*	0.007

Characteristics		
CAPOX	0.0040	0.007
Low risk stage III	-0.0023	0.010
High risk stage III	-0.0086	0.010
Male	0.0213*	0.007
Age	0.0014*	0.000
Ethnic (ref: Caucasian)		
African/Caribbean	-0.0719*	0.035
South Asian	-0.1229*	0.047
Chinese	-0.0302	0.064
Other	0.0166	0.0210
Constant	0.7701	0.027
Explanation: Comparison is a 65-year old, Caucasian female patient on the 3 month trial arm in a disease free health state and stage II disease treated with FOLFOX. *p<0.05		

Table 13-8 Utility regression (Denmark)

Denmark utilities		
Variable	Co-efficient	S.E.
Health states (ref: disease free)		
On treatment	-0.0363*	0.003
Recurrence	-0.0572*	0.013
Arm: 6 months (ref: 3 months)	-0.0129*	0.006
Characteristics		
CAPOX	0.0037	0.006
Low risk stage III	-0.0017	0.009
High risk stage III	0.0072	0.009
Male	0.0193*	0.006
Age	0.0013*	0.000
Ethnic (ref: Caucasian)		
African/Caribbean	-0.0630*	0.031
South Asian	-0.1123*	0.043
Chinese	-0.0235	0.052
Other	0.0150	0.019
Constant	0.7909	0.024
Explanation: Comparison is a 65-year old, Caucasian female patient on the 3 month trial arm in a disease free health state and stage II disease treated with FOLFOX. *p<0.05		

Table 13-9 Utility regression (New Zealand)

New Zealand utilities		
Variable	Co-efficient	S.E.
Health states (ref: disease free)		
On treatment	-0.0526*	(0.005)
Recurrence	-0.0777*	(0.017)

Arm: 6 months (ref: 3 months)	-0.0178*	(0.009)
Characteristics		
CAPOX	0.0051	(0.009)
Low risk stage III	-0.0060	(0.013)
High risk stage III	-0.0129	(0.013)
Male	0.0288*	(0.009)
Age	0.0016*	(0.000)
Ethnic (ref: Caucasian)		
African/Caribbean	-0.0859*	(0.039)
South Asian	-0.1161*	(0.048)
Chinese	-0.0467	(0.072)
Other	0.0253	(0.026)
Constant	0.6931	(0.034)
Explanation: Comparison is a 65-year old, Caucasian female patient on the 3 month trial arm in a disease free health state and stage II disease treated with FOLFOX. *p<0.05		

Table 13-10 Utility regression (Spain)

Spain utilities		
Variable	Co-efficient	S.E.
Health states (ref: disease free)	-0.0360*	0.004
On treatment	-0.0632*	0.015
Recurrence	-0.0157*	0.007
Arm: 6 months (ref: 3 months)	0.0052	0.008
Characteristics		
CAPOX	0.0052	0.008
Low risk stage III	0.0002	0.010
High risk stage III	-0.0074	0.010
Male	0.0129	0.007
Age	0.0014*	0.000
Ethnic (ref: Caucasian)		
African/Caribbean	-0.0725*	0.036
South Asian	-0.1529*	0.056
Chinese	-0.0473	0.091
Other	0.0243	0.020
Constant	0.7957	0.029
Explanation: Comparison is a 65-year old, Caucasian female patient on the 3 month trial arm in a disease free health state and stage II disease treated with FOLFOX. *p<0.05		

Table 13-11 Utility regression (Sweden)

Sweden utilities		
Variable	Co-efficient	S.E.
Health states (ref: disease free)		
On treatment	-0.0242*	(0.002)

Recurrence	-0.0355*	(0.008)
Arm: 6 months (ref: 3 months)	-0.0091*	(0.004)
Characteristics		
CAPOX	0.0026	(0.004)
Low risk stage III	-0.0008	(0.006)
High risk stage III	-0.0057	(0.006)
Male	0.0115*	(0.004)
Age	0.0008*	(0.000)
Ethnic (ref: Caucasian)		
African/Caribbean	-0.0317	(0.020)
South Asian	-0.0668*	(0.027)
Chinese	-0.005	(0.038)
Other	0.0148	(0.012)
Constant	0.8466	(0.016)
Explanation: Comparison is a 65-year old, Caucasian female patient on the 3 month trial arm in a disease free health state and stage II disease treated with FOLFOX. *p<0.05		

Table 13-12 Yearly cost regression (Australia)

Australia costs (Regression used to model costs for years 7-10)		
Variable	Co-efficient	S.E.
Health states (ref: disease free)		
On treatment	16,127*	375
Recurrence	12,674*	961
Arm: 6 months (ref: 3 months)	334	180
Year 1	7,178*	420
Year 1*Arm (6 months)	756	455
Characteristics		
CAPOX	-5,683*	346
Low risk III	70	390
High risk III	260	360
Male	-476	321
Age		16
Ethnic (ref: Caucasian)		
African/Caribbean	-1,866	978
South Asian	137	1,220
Chinese	486	1,964
Other	-297	505
Constant	5,104	1,119
Explanation: Comparison is a Caucasian female patient on the 3 month trial arm aged 65 years with stage II disease, treated with FOLFOX after year 1 in a disease free state.		

Table 13-13 Yearly cost regression (Denmark)

Denmark costs (Regression used to model costs for years 7-10)		
Variable	Co-efficient	S.E.

Health states (ref: disease free)		
On treatment	15,795*	365
Recurrence	12,335*	934
Arm: 6 months (ref: 3 months)	325	176
Year 1	6,997*	408
Year 1*Arm (6 months)	744	423
Characteristics		
CAPOX	-5,559*	337
Low risk III	65	380
High risk III	251	351
Male	-462	312
Age	78	15
Ethnic (ref: Caucasian)		
African/Caribbean	-1811	957
South Asian	156	1195
Chinese	471	1,915
Other	-290	492
Constant	5,007	1,089
Explanation: Comparison is a Caucasian female patient on the 3 month trial arm aged 65 years with stage II disease, treated with FOLFOX after year 1 in a disease free state.		

Table 13-14 Yearly cost regression (New Zealand)

New Zealand costs (Regression used to model costs for years 7-10)		
Variable	Co-efficient	S.E.
Health states (ref: disease free)		
On treatment	11,960*	261
Recurrence	8,948*	659
Arm: 6 months (ref: 3 months)	237	129
Year 1	5,153*	286
Year 1*Arm (6 months)	589	317
Characteristics		
CAPOX	-4204*	243
Low risk III	32	272
High risk III	180	255
Male	-322	223
Age	18	11
Ethnic (ref: Caucasian)		
African/Caribbean	-1273	720
South Asian	98	887
Chinese	425	1422
Other	-171	360
Constant	3939	779

Explanation: Comparison is a Caucasian female patient on the 3 month trial arm aged 65 years with stage II disease, treated with FOLFOX after year 1 in a disease free state.

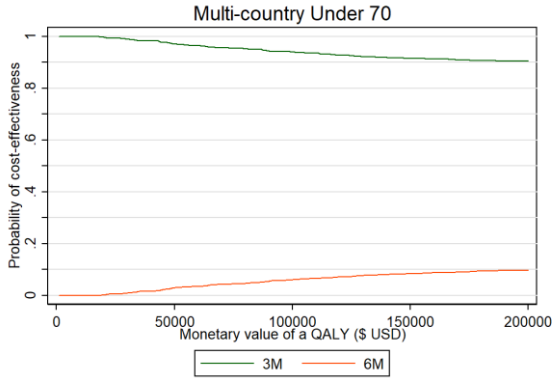
Table 13-15 Yearly cost regression (Spain)

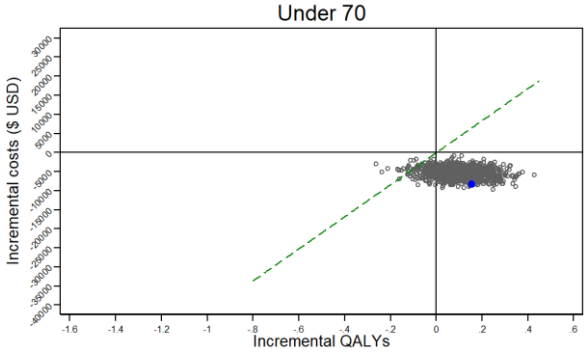
Spain costs (Regression used to model costs for years 7-10)		
Variable	Co-efficient	S.E.
Health states (ref: disease free)		
On treatment	12,636*	(279)
Recurrence	9,533*	(706)
Arm: 6 months (ref: 3 months)	252	137
Year 1	5,474*	307
Year 1*Arm (6 months)	617	338
Characteristics		
CAPOX	-4443*	259
Low risk III	38	290
High risk III	129	271
Male	-346	239
Age	20	12
Ethnic (ref: Caucasian)		
African/Caribbean	-1365	761
South Asian	107	940
Chinese	435	1,508
Other	-191	382
Constant	4,132	832
Explanation: Comparison is a Caucasian female patient on the 3 month trial arm aged 65 years with stage II disease, treated with FOLFOX after year 1 in a disease free state.		

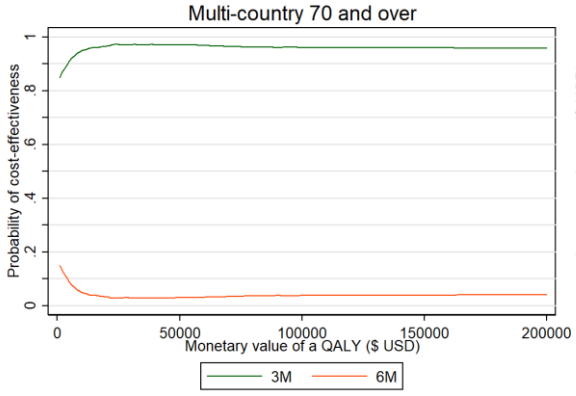
Table 13-16 Yearly cost regression (Sweden)

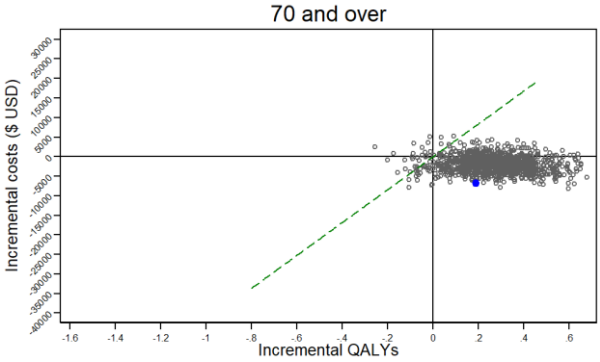
Sweden costs (Regression used to model costs for years 7-10)		
Variable	Co-efficient	S.E.
Health states (ref: disease free)		
On treatment	12,636*	1,034
Recurrence	15,548*	397
Arm: 6 months (ref: 3 months)	342	183
Year 1	7,225*	455
Year 1*Arm (6 months)	673	482
Characteristics		
CAPOX	-5,484*	361
Low risk III	103	412
High risk III	269	373
Male	-522	340
Age	34	17
Ethnic (ref: Caucasian)		

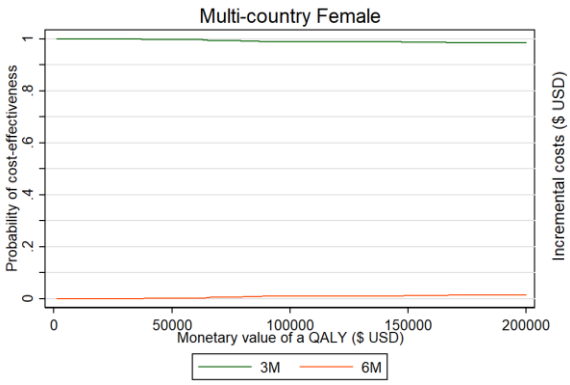
African/Caribbean	-2,011	957
South Asian	201	1,223
Chinese	305	1,952
Other	-400	517
Constant	4,584	1,183
Explanation: Comparison is a Caucasian female patient on the 3 month trial arm aged 65 years with stage II disease, treated with FOLFOX after year 1 in a disease free state.		

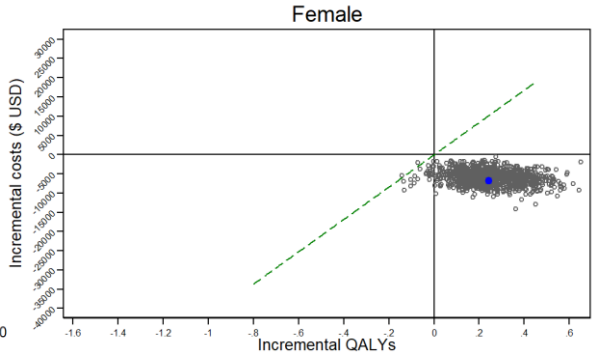


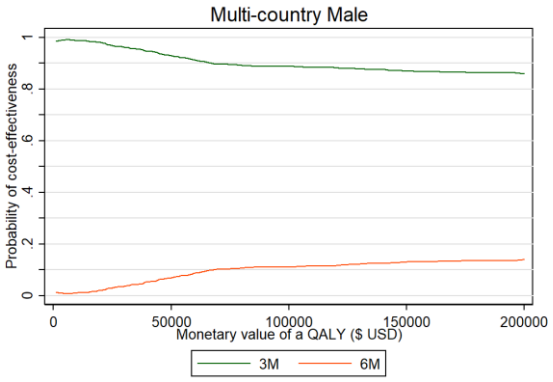


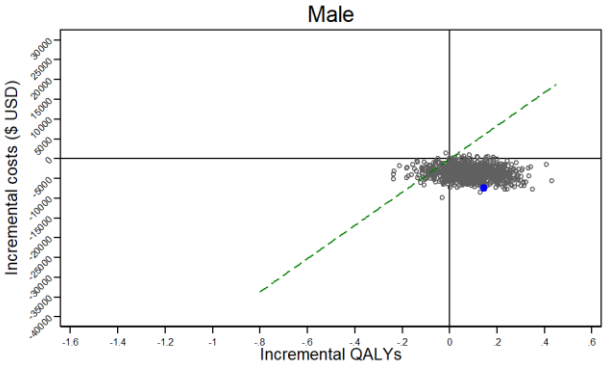


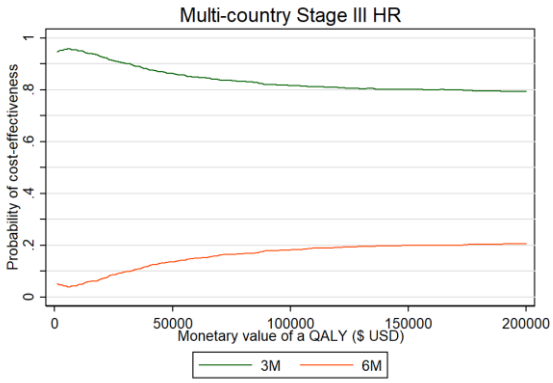


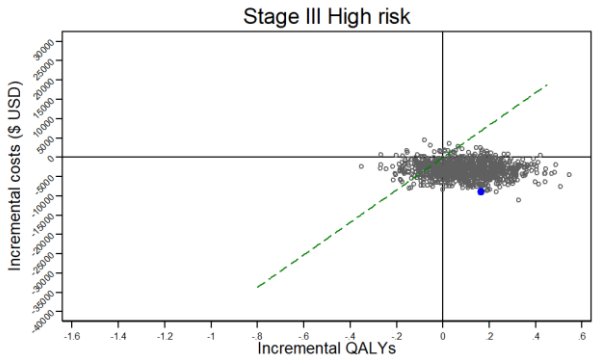


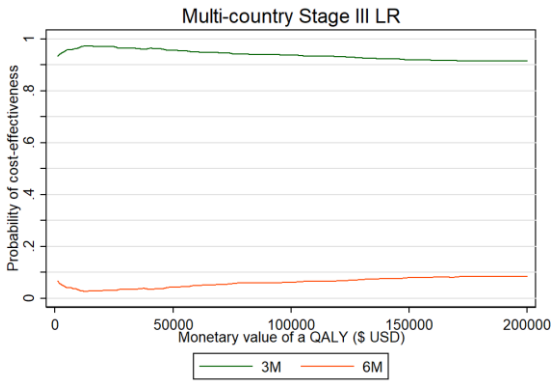


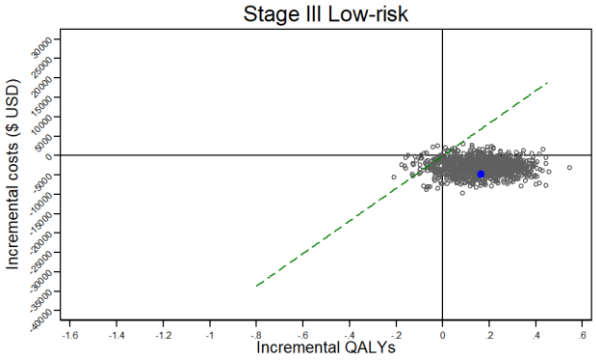

Multi-country Male
Probability of cost-effectiveness
0
.2
.4
.6
.8
1
0
50000
100000
150000
200000
Monetary value of a QALY ($ USD)
3M
6M

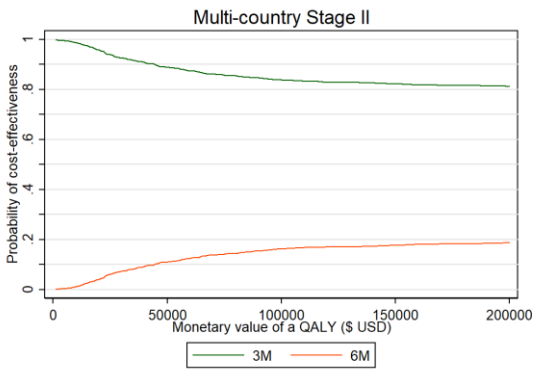

Male
Incremental costs ($ USD)
Incremental QALYs

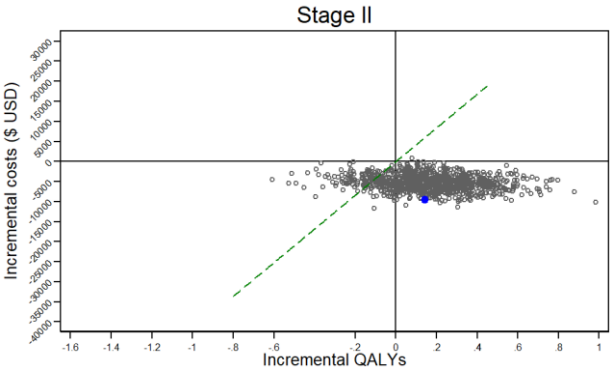

Multi-country Stage III HR
Probability of cost-effectiveness
0
.2
.4
.6
.8
1
0
50000
100000
150000
200000
Monetary value of a QALY ($ USD)
3M
6M


Stage III High risk
Incremental costs ($ USD)
Incremental QALYs


Multi-country Stage III LR
Probability of cost-effectiveness
0
.2
.4
.6
.8
1
0
50000
100000
150000
200000
Monetary value of a QALY ($ USD)
3M
6M


Stage III Low-risk
Incremental costs ($ USD)
Incremental QALYs


Multi-country Stage II
Probability of cost-effectiveness
0
.2
.4
.6
.8
1
0
50000
100000
150000
200000
Monetary value of a QALY ($ USD)
3M
6M


Stage II
Incremental costs ($ USD)
Incremental QALYs

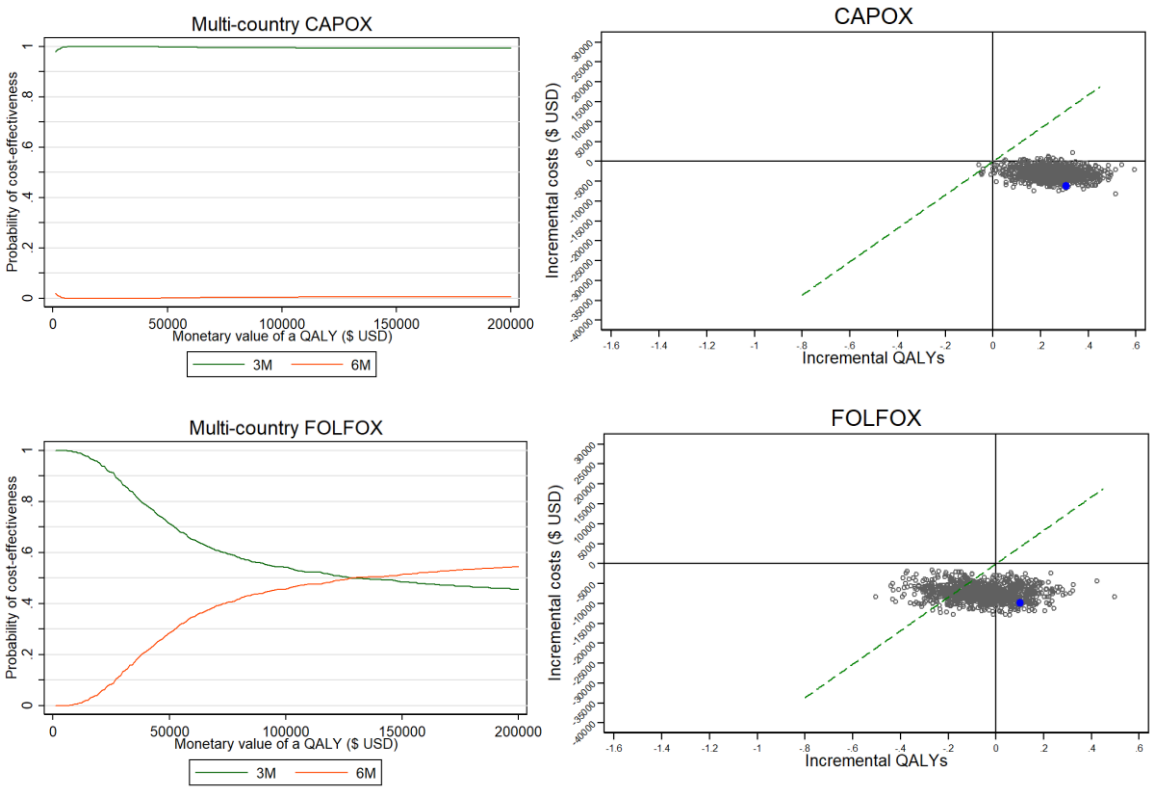


Figure 13-2 Subgroup analyses CEACs and cost-effectiveness planes The WTP threshold plotted for the subgroup cost-effectiveness planes was $42,000, in line with the UK NICE guidance of £30,000 (197).

Table 13-17 Budget impact analysis results in country-specific currency

Country currency (rounded to nearest 100,000)	Australia	Denmark	New Zealand	Spain	Sweden	United Kingdom
Base scenario budget impact						
Chemotherapy medication costs over 5 years	3,000,000 AUD	3,800,000 Kr	$400,000	€1,300,000	9,400,000 kr	£1,400,000
Treatment related hospitalisations in (year 1 for each individual patient) over 5 years	32,100,000 AUD	44,300,000 Kr	$3,600,000	€30,500,000	100,100,000 kr	£41,800,000
Condition related hospitalisations (years 2-5 for each individual patient) over 5 years	2,000,000 AUD	2,700,000 Kr	$200,000	€1,800,000	6,400,000 kr	£2,500,000
Total budget impact = Medication cost + Cost of treatment and condition related hospitalisations	37,000,000 AUD	50,700,000 Kr	$4,300,000	€33,600,000	116,000,000 kr	£45,800,000
Scenario analysis (otherwise as per base scenario)						
Rectal Cancer Excluded	25,600,000 AUD	34,000,000 Kr	$2,900,000	€22,100,000	78,600,000 kr	£30,800,000
Stage II Excluded	35,300,000 AUD	48,500,000 Kr	$4,100,000	€32,100,000	111,400,000 kr	£43,700,000

The SCOT study received the following project specific funding:
UK
SCOT Original Grant:
MRC (subsequently transferred to NIHR NETSCC)
Grant Ref: G0601706
Duration: 96 months, 1st Dec 2006 – 30th November 2014
Total awarded: MRC contribution £2,449,391 (FEC £3,061,732)

CRUK CTU Core Funding (Estimated):
Used in the 12 month gap between the MRC and HTA funding.
2 X 1.0 Full Time Equivalent (FTE) Clinical Trial Coordinators
0.25 FTE Project Manager
0.5 FTE Clinical Trial Administrator
0.3 FTE Statistician

Total cost = £129,000 (Grant: C6716/A9894)

Extension to Follow-up:
NIHR HTA
Grant Ref: 14/140/84
Duration: 30 months, 1st Dec 2015 – 31st May 2018
Total awarded: £274,695

Within the UK, the study was also supported with NCRN/SCRN/NIHR infrastructure funding within the participating sites.

(UK total £2,853,086, appox $3.8 million USD)

Sweden
Swedish Cancer Society (as part of a larger (2 million SEK; approx. $175,000 USD) annual grant for CRC research; proportion for SCOT trial not quantifiable.)

Denmark
Interreg grant. Approximately $700,000 USD (5 million Danish Kr).

These funders had no role in the identification, design, conduct, and reporting of the analysis.

Total project specific funding approx. 8.8 million USD

14 Appendix 5 Supplementary material relevant to analysis of administrative healthcare data

Table 14-1 Datasets request as part of PBPP application 1718-0026

Dataset	*Data controller*	*Description*	*Years requested for the purposes of this project*
NRS Deaths	NRS	This dataset is collected by National Records Scotland (NRS), which is a Scottish government institution. It contains information on date, cause and place of death for all deaths registered in Scotland since 1974. PHS is granted access to extracts from this dataset for research/linkage purposes.	2006-2018
SMR00	PHS	This dataset contains patient level episode data on outpatient appointments across all specialities (except A & E and Genito-urinary medicine). Data collection began in Scotland in the 1990s. Data collection within 6 weeks of outpatient attendance.	1997-2018
SMR01	PHS	SMR01 comprises patient level episode data on hospital inpatient and day case discharges from acute specialities in Scotland. Data is available in computerised format from 1968.	1997-2018
SMR04	PHS	SMR04 contains data for patients receiving care in Mental Health facilities (inpatient and day cases).	2006-2018
SMR06	PHS	SMR06 is also known as the Scottish Cancer Registry and established 1954. Collects information relevant to the diagnosis and management of malignant neoplasms, as well as carcinoma in situ and some benign tumours. Data is collected annually. CORECT-R Scotland has requested information on patients with a diagnosis of CRC only.	2006-2018
ChemoCare WoSCAN	WoS Cancer Network	Regional chemotherapy prescribing dataset.	2006-2018 but reliable data from 2012 onwards
ChemoCare SCAN	SCAN Cancer Network	Regional chemotherapy prescribing dataset.	2012-2018
ChemoCare Grampian	Grampian Cancer Network	Regional chemotherapy prescribing dataset.	2006-2018 but reliable data from 2012 onwards
ChemoCare Tayside	Tayside Cancer Network	Regional chemotherapy prescribing dataset.	2006-2018 but reliable data from 2012 onwards
ChemoCare Highlands	Highlands Cancer Network	Regional chemotherapy prescribing dataset.	2006-2018 but reliable data from 2012 onwards
QPI WoS	NHS Greater Glasgow and Clyde	National prospective audit dataset collected and stored regionally on an annual basis (April each year). NHS boards are required to report their activity against QPIs as part of a mandatory national cancer quality programme. Healthcare Improvement Scotland is responsible for the external quality assurance of cancer services against tumour specific QPIs.	2013-2018
QPI SCAN	NHS Lothian	National prospective audit dataset collected and stored regionally on an annual basis (April each year).	2013-2018

QPI NoSCAN	NHS Tayside, NHS Grampian and NHS Highland	National prospective audit dataset collected and stored regionally on an annual basis (April each year).	2013-2018
PLICS	PHS	PLICS is the patient level information costing system and contains cost variables derived from SMR00, SMR01 and SMR04. This top down costing system was developed to allow hospital costs to be attributed to patient activity in a detailed way to reflect key cost drivers such as length of stay and apportions hospital site and speciality specific direct costs to individual patient records on admission, per day, for theatre time and specific high cost items. PLICS is not yet available for SMR06.	Financial year 2014/15-2017 (SMR01 2012 start)
SICSAG	Scottish Intensive Care Society Audit Group	SICSAG is the Scottish Intensive Care Society Audit Group dataset and contains both episode level and daily information provided for each patient.	2006-2018
Radiotherapy	NA	Detailed radiotherapy data is currently not available on a national basis in Scotland and held loco-regionally by radiotherapy centres. Key information on radiation treatment delivered (for example if radiotherapy was delivered and date of treatment) is currently available within the Scottish Cancer Registry (SMR06). However, granular radiotherapy data (for example dose, technique and modality) is currently held by individual hospital institutions which deliver radiotherapy. A process is in development to make radiotherapy data available nationally – this involves Scottish radiotherapy centres sending data extract to Public Health England, who curate the data to a common standard prior to returning to Public Health Scotland.	Not currently available
Prescribing Information System	PHS	The Prescribing Information System (PIS) is a data source for all prescribing of medicines (and their costs) that are prescribed and dispensed in the community in Scotland. Includes medications prescribed in hospital but dispensed in the community but not those dispensed in hospital. Information for this dataset is supplied by the Practitioner and Counter Fraud Services Division.	2010-2018
Accident and Emergency	PHS	Originally established in 2007 to monitor compliance of each NHS board with the maximum four hour waiting time target. Departments may submit individual episode level data (detailed information on each patient attendance) or aggregate level data (often smaller minor injury units). Sites that submit episode level data account for 94% of national A and E attendances.	2011-2018
GP Out of Hours	PHS	A Scottish government commissioned (2014) dataset to improve understanding of activity, demand and capacity at a national level for primary care out of hours services.	2014-2018
Scottish Ambulance Service (SAS)	PHS	The SAS dataset contains individual level records of all patient contact with the service.	2011-2018
NHS 24	PHS	The NHS 24 dataset contains individual level records of all patient contact with the service.	2011-2018

14.1 Comparison of methods of cohort derivation for the GG&C cohort

The primary researcher (CH) did not have access to patient level information from QPI and ACaDMe raw datasets and therefore this step of cohort derivation was reliant on the ChemoCare (CC) and QPI (FC) data providers. Within the ACaDMe dataset, patients with a diagnosis CRC (ICD-10 codes 18-20) and with Duke's B or Duke's C disease in SMR06, who had undergone CRC surgery and who received chemotherapy within six months of this surgery were identified. Duke's stage was used because it had a superior completion rate within SMR06 compared to TNM staging. When using the QPI dataset, patients with either Duke's B or C CRC who had received adjuvant chemotherapy were identified using a variable called "ADJONC" and criteria "2" which indicated that chemotherapy was given to that patient with adjuvant intent. This variable (ADJONC) was available from 2010 (pre-QPI) and continued to be collected as part of the QPI dataset.

The accuracy of the SMR06 diagnostic ICD-10 codes and the staging and pathological data provided by the Cancer Audit data was assessed on a small subset of patients within the final ChemoCare cohort using patient identifiable data. For the purposes of this analysis, if there were disease staging or pathological details missing from these data sources, these were retrieved by looking at the individual patient electronic records. If this information was still not available, these fields were designated as missing.

Figure 14-1 shows the cohorts defined using three different approaches. The largest cohort was obtained using Cancer Audit/QPI data (n=1096) and 334 patients were identified in the Cancer Audit/QPI cohort that were not present in the final ChemoCare cohort. The majority (88%, 164/186) of patients who met the inclusion criteria but who were not the ChemoCare dataset resided outside GG&C; they were included in the QPI GG&C data extraction because their CRC diagnosis occurred in the GG&C health board. Cancer Audit data identifies all patients diagnosed with CRC in GG&C whereas the ChemoCare dataset defines patient location using the health board of residence. Patients who reside outside the GG&C health board can be diagnosed in GG&C if they are referred to a

tertiary treatment centre for imaging, diagnostic biopsies, or surgery and this explains why the Cancer Audit cohort is larger. The approximate specificity and sensitivity of using Cancer Audit data was estimated. Overall, using Cancer Audit/QPI data to identify patients with stage II/III disease who received adjuvant chemotherapy had a sensitivity of approximately 90% and a specificity of 95% compared to what was considered the gold standard for this cohort, which was ChemoCare data which was cleaned and prepared using identifiable electronic portal records.

The ACaDMe (SMR06) cohort was the smallest of the three cohorts used and the majority of these patients were found in the Cancer Audit or ChemoCare databases. Thirty-nine patients were identified using ACaDME that were not in the other two databases but only two of these patients met the inclusion criteria. Because the primary researcher did not have direct access to the ACaDME datasets, it was not possible to interrogate this method of cohort derivation further. For example, it was unclear which codes had been used to signify that the patient had undergone CRC surgery.

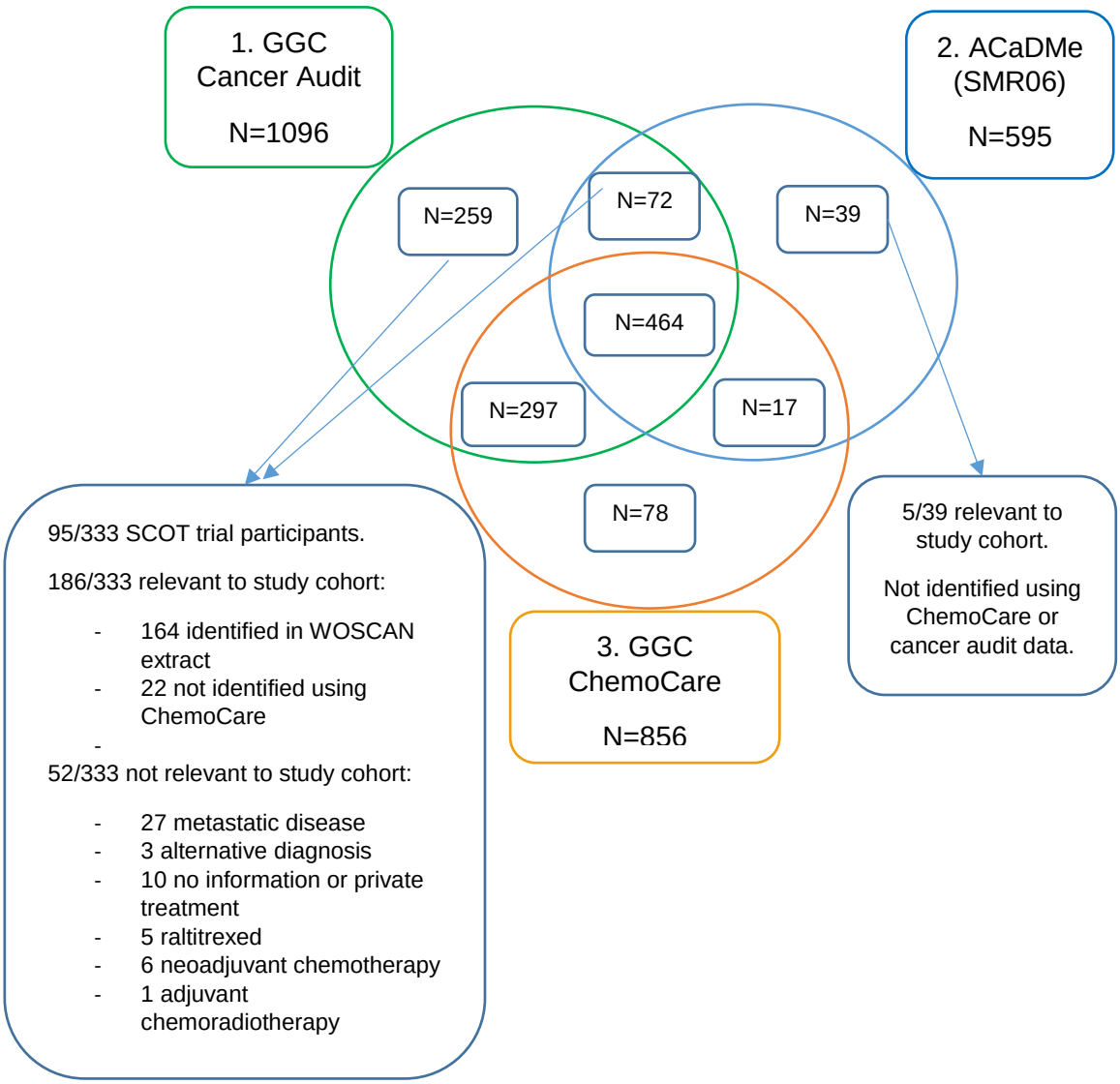


Figure 14-1 Comparison of GG&C cohort definition using three approaches

The cohort chosen for further analysis was the original in ChemoCare cohort. ICD-10 codes were available for 99% (n=851/856) patients and twenty patients were randomly picked to check the accuracy of the ICD-10 code against patient identifiable data and these were correct in 19/20 cases. One case had recorded the diagnosis as colon cancer, whereas a code of rectosigmoid cancer would have been more specific. Disease staging information and pathological details were provided by the QPI dataset. TNM staging was available for 849/858 (99%) records and Duke's staging was available for 841/856 (98%) records. A different set of twenty electronic records were reviewed to check the accuracy of these entries. There was 100% concordance between the Tumour and Nodal staging (TNM codes), Duke's staging and pathological details (for example, number of nodes sampled, margin positivity, degree of differentiation) for these patients from the QPI dataset and the patient identifiable records.

Table 14-2 Variables used for GG&C analysis

Category of variable	Variable name/description	Name in dataset	Variable type	Dataset	Formation of derived variables
Disease characteristics	Tumour stage	tstage	Raw	AcaDME (SMR06) and Clinical Portal	NA
	Nodal stage	nstage	Raw	AcaDME (SMR06) and Clinical Portal	NA
	Extended risk stage	riskstage	Derived	NA	Derived from T-stage and N-stage
	Duke's stage	dukes	Raw	AcaDME (SMR06) and Clinical Portal	NA
	Site of disease	Site	Raw	AcaDME (SMR06) and Clinical Portal	NA
Patient characteristics	Gender	Gender	Raw	AcaDME (SMR06) and Clinical Portal	NA
	Age	age	Derived	NA	Derived by analyst (CC) from ChemoCare record using date of birth. Age at first chemotherapy treatment
	Age grouping	ageg2	Derived	NA	Derived from age by primary researcher (CH). Groupings were under 70 and 70 and over.
	Scottish Index of Multiple Deprivation category	Depcat	Derived	NA	Derived by analyst (CC) from ChemoCare record using post-code and SIMD codes (quintiles).
	Height	HEIGHT	Raw		NA
	Weight	WEIGHT	Raw		NA
	Body surface area	BSA	Derived	NA	Derived by primary researcher (CH) using HEIGHT and WEIGHT
	Cohort identifier	ID	Derived	NA	Derived by primary researcher (CH) to anonymise data
Treatment characteristics	Drug regimen*	REGIMEN	Raw	ChemoCare GG&C	NA
	Drug name	DRUGNAME	Raw	ChemoCare GG&C	NA
	Date of chemotherapy delivery	APT_DATE	Raw	ChemoCare GG&C	NA
	Drug dose	DRUGDOSE	Raw	ChemoCare GG&C	NA
	Drug dose per m2	DOSEM2	Derived	NA	Derived by primary researcher (CH) using DRUGDOSE and BSA
	First adjuvant chemotherapy regimen prescribed	first_regimen	Derived	NA	Derived by primary researcher (CH) from ChemoCare record using drug name, drug dose and date of chemotherapy delivery

	Duration of treatment (calculated using cycles)	total_weeks	Derived	NA	Derived by primary researcher (CH) from ChemoCare record using drug name, drug dose and date of chemotherapy delivery
	Date of first adjuvant chemotherapy treatment	first_Apt	Derived	NA	Derived by primary researcher (CH) from ChemoCare record using drug name, drug dose and date of chemotherapy delivery
	Year of first adjuvant chemotherapy treatment	year	Derived	NA	Derived by primary researcher (CH) from ChemoCare record using date of chemotherapy delivery
	Month of first adjuvant chemotherapy treatment	month	Derived	NA	Derived by primary researcher (CH) from ChemoCare record using date of chemotherapy delivery
	Indication of if patient received over 3 months of treatment	over12w	Derived	NA	Derived by primary researcher (CH) using derived variable "total_weeks"
	Time since first patient in cohort started adjuvant treatment	time	Derived	NA	Derived by primary researcher (CH) using date of first adjuvant chemotherapy delivered
	Indication if adjuvant chemotherapy delivered before or after ASCO 2017	ASCO	Derived	NA	Derived by primary researcher (CH) using date of first adjuvant chemotherapy delivered
	Indication of if patient's first adjuvant chemotherapy regimen was monotherapy or doublet treatment	doublet	Derived	NA	Derived by primary researcher (CH) from ChemoCare record using drug name and drug dose

Table 14-3 Demographic, disease and treatment related variables for patients receiving doublet chemotherapy at cycle one (GG&C analysis)

Demographic, disease or treatment related variable	Pre-SCOT (percentage) n= 484 (76%)	Post-SCOT (percentage) n= 154 (24%)	Total Number (percentage) n=638
Sex			
Male	262 (54%)	76 (49%)	338 (53%)
Female	222 (46%)	78 (51%)	300 (47%)
Age			
Median age (IQR)	62 (54-67)	62 (53-68)	62 (54-68)
Age group			
≤70 years	414 (86%)	127 (82%)	541 (85%)
>70 years	70 (14%)	27 (18%)	97 (15%)
Deprivation category			
1	167 (35%)	50 (32%)	217 (34%)
2	79 (16%)	24 (16%)	103 (16%)
3	50 (10%)	23 (15%)	73 (11%)
4	70 (14%)	20 (13%)	90 (14%)
5	115 (24%)	35 (23%)	150 (24%)
Unknown	3 (1%)	2 (1%)	2 (0%)
Location of disease			
Colon	346 (71%)	118 (77%)	464 (73%)
Rectosigmoid or Rectum	138 (29%)	36 (23%)	174 (27%)
Duke's stage			
B	81 (17%)	16 (10%)	97 (15%)
C	400 (83%)	138 (90%)	538 (84%)
Unknown	3 (1%)	0 (0%)	3 (0%)
T stage			
X	1 (0%)	0 (0%)	1 (0%)
0	0 (0%)	1 (1%)	1 (0%)
1	15 (3%)	8 (5%)	23 (4%)
2	35 (7%)	11 (7%)	46 (7%)
3	240 (50%)	81 (53%)	321 (50%)
4	188 (39%)	53 (34%)	241 (38%)
Unknown	5 (1%)	0 (0%)	5 (1%)
N stage			
0	80 (17%)	16 (10%)	96 (15%)
1	252 (52%)	100 (65%)	352 (55%)
2	147 (30%)	38 (25%)	185 (29%)
Unknown	5 (1%)	0 (0%)	5 (1%)

Stage III risk groups			
Low risk	177 (44%)	75 (54%)	252 (47%)
High risk	222 (56%)	63 (46%)	285 (53%)
Treatment regimen			
CAPOX	383 (80%)	114 (74%)	499 (78%)
FOLFOX	99 (20%)	40 (26%)	139 (22%)

Table 14-4 Demographic, disease and treatment related variables for patients receiving single agent chemotherapy at cycle one (GG&C analysis)

Demographic, disease or treatment related variable	Pre-SCOT (percentage) n= 270 (75%)	Post-SCOT (percentage) n= 90 (25%)	Total Number (percentage) n= 360
Sex			
Male	132 (49%)	52 (58%)	184 (51%)
Female	138 (51%)	38 (42%)	176 (49%)
Age			
Median age (IQR)	70 (61-74)	69 (61-73)	69 (61-74)
Age group			
≤70 years	141 (52%)	56 (63%)	197 (55%)
>70 years	129 (48%)	34 (38%)	163 (45%)
Deprivation category			
1	90 (33%)	22 (24%)	112 (31%)
2	39 (14%)	24 (27%)	63 (18%)
3	37 (14%)	14 (16%)	51 (14%)
4	41 (15%)	9 (10%)	50 (14%)
5	61 (23%)	21 (23%)	82 (23%)
Unknown	2 (1%)	0 (0%)	2 (1%)
Location of disease			
Colon	204 (76%)	77 (86%)	281 (78%)
Rectosigmoid or Rectum	66 (24%)	13 (14%)	79 (22%)
Duke's stage			
B	149 (55%)	61 (68%)	210 (58%)
C	120 (44%)	29 (32%)	149 (41%)
Unknown	1 (0%)	0 (0%)	1 (0%)
T stage			
X	0 (0%)	0 (0%)	0 (0%)
0	0 (0%)	0 (0%)	0 (0%)
1	4 (2%)	0 (0%)	4 (1%)
2	11 (4%)	2 (2%)	13 (4%)
3	170 (63%)	63 (70%)	233 (65%)
4	84 (31%)	25 (28%)	109 (30%)
Unknown	1 (0%)	0 (0%)	1 (0%)
N stage			

0	149 (55%)	62 (69%)	211 (59%)
1	86 (32%)	18 (20%)	104 (29%)
2	34 (13%)	10 (11%)	44 (12%)
Unknown	1 (0%)	0 (0%)	1 (0%)
Stage III risk groups			
Low risk	63 (53%)	8 (29%)	71 (48%)
High risk	57 (48%)	20 (71%)	77 (52%)
Treatment regimen			
Capecitabine	255 (94%)	87 (97%)	342 (95%)
IV 5-fluorouracil	13 (6%)	3 (3%)	18 (5%)

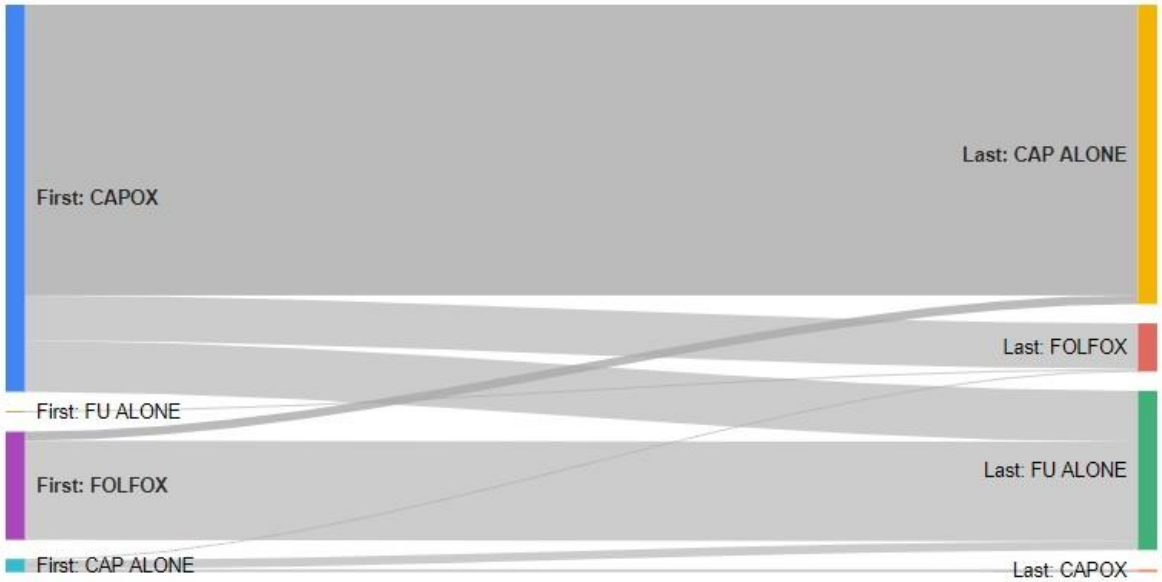


Figure 14-2 Treatment switch First regimen: CAPOX (n=265, 76%), FOLFOX (n=74, 21%), CAP alone (n=15, 4%), FU alone (n=1, (0.3%)). The largest proportion of treatment switching was a change from using CAPOX in cycle one to capecitabine monotherapy at the last cycle of treatment delivered (n=199 of 349, 57%).

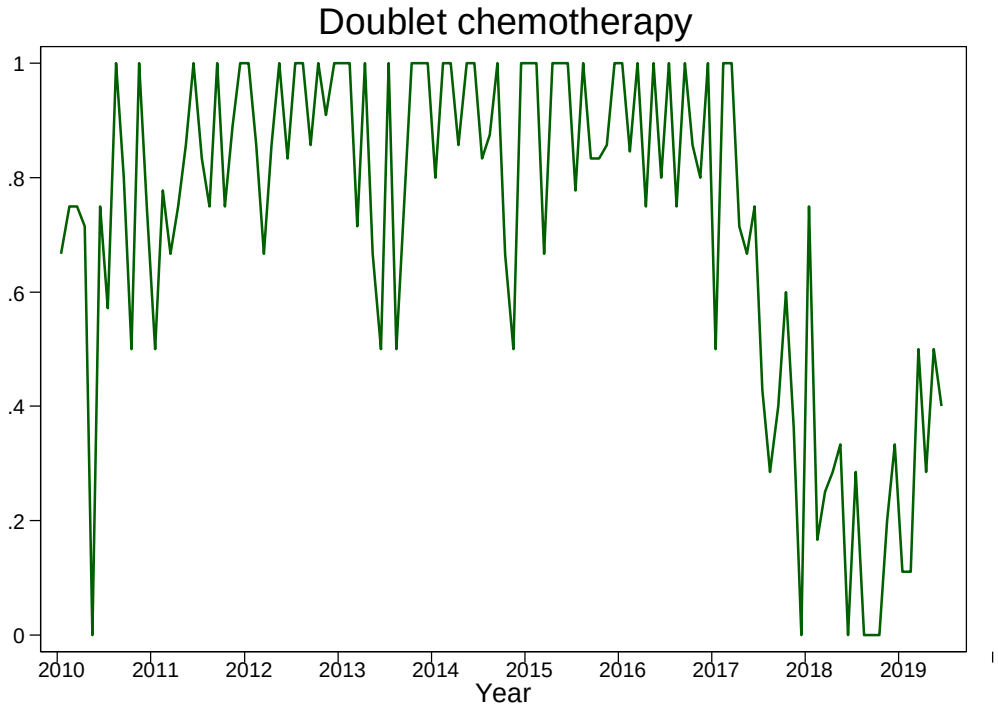


Figure 14-3 Time series to check for seasonality (logistic regression for GG&C doublet chemotherapy analysis)

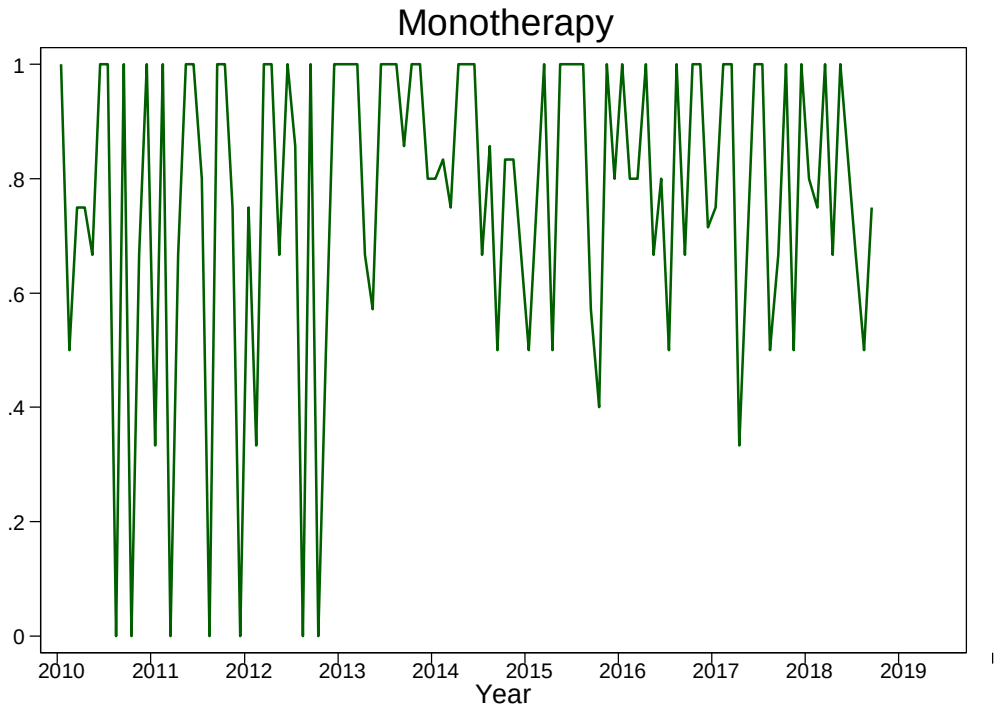


Figure 14-4 Time series to check for seasonality (logistic regression for GG&C single agent chemotherapy analysis)

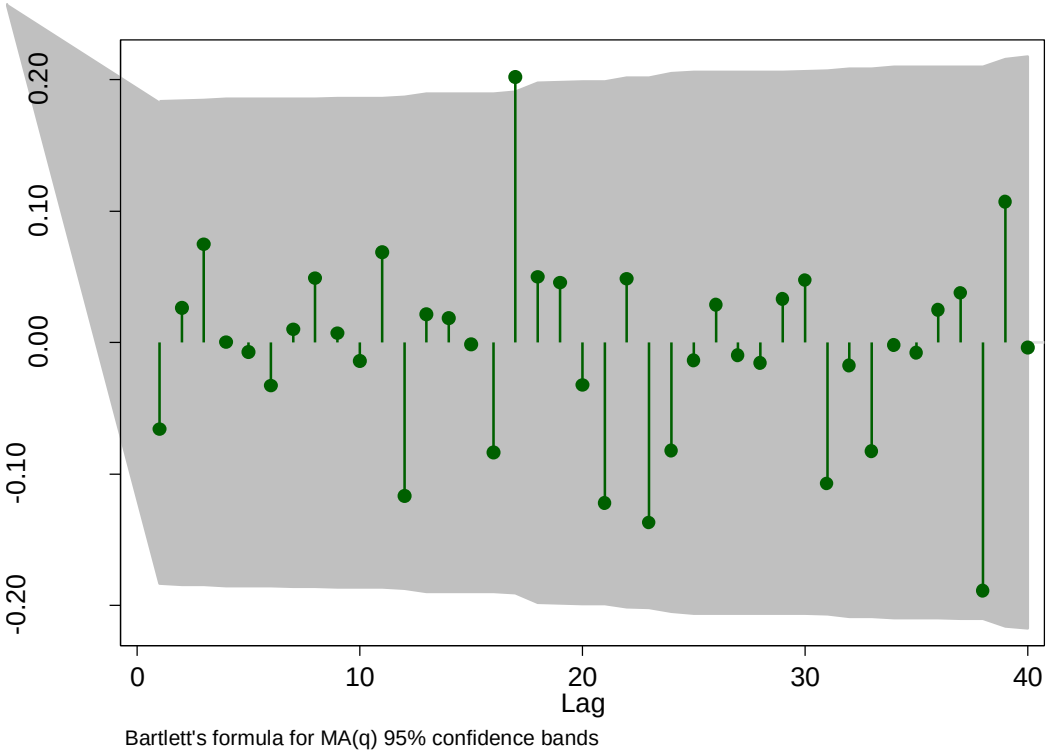


Figure 14-5 Plot of residuals (linear regression doublet chemotherapy)

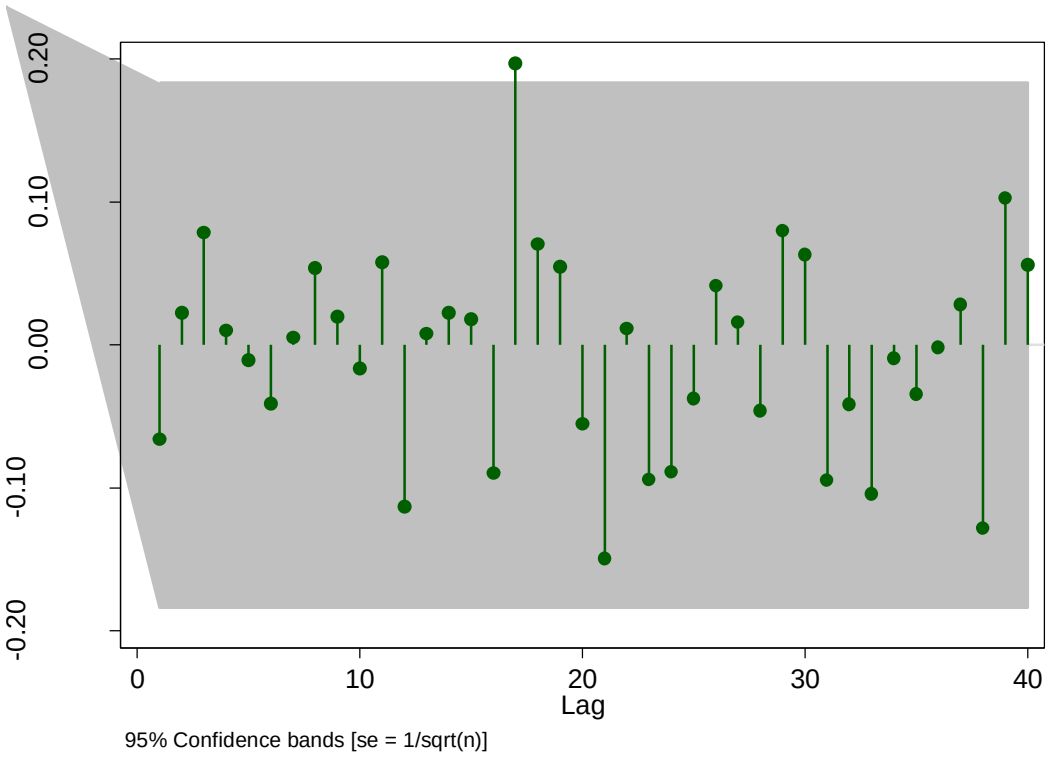


Figure 14-6 Plot of partial residuals (linear regression doublet chemotherapy)

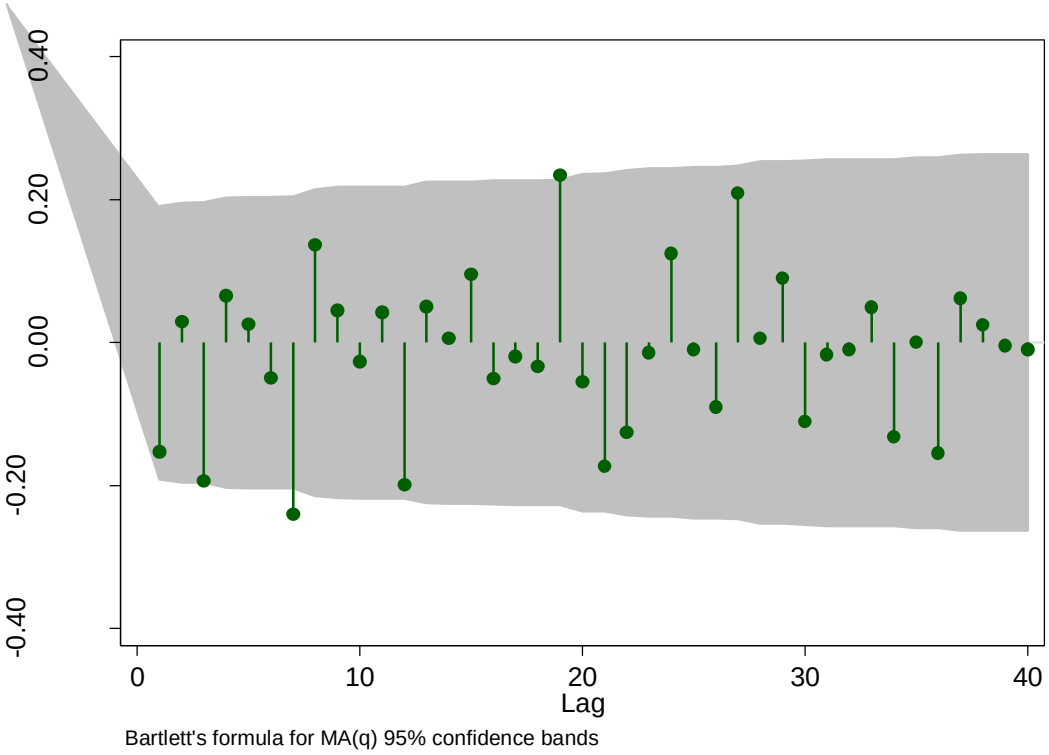


Figure 14-7 Plot of residuals (linear regression monotherapy)

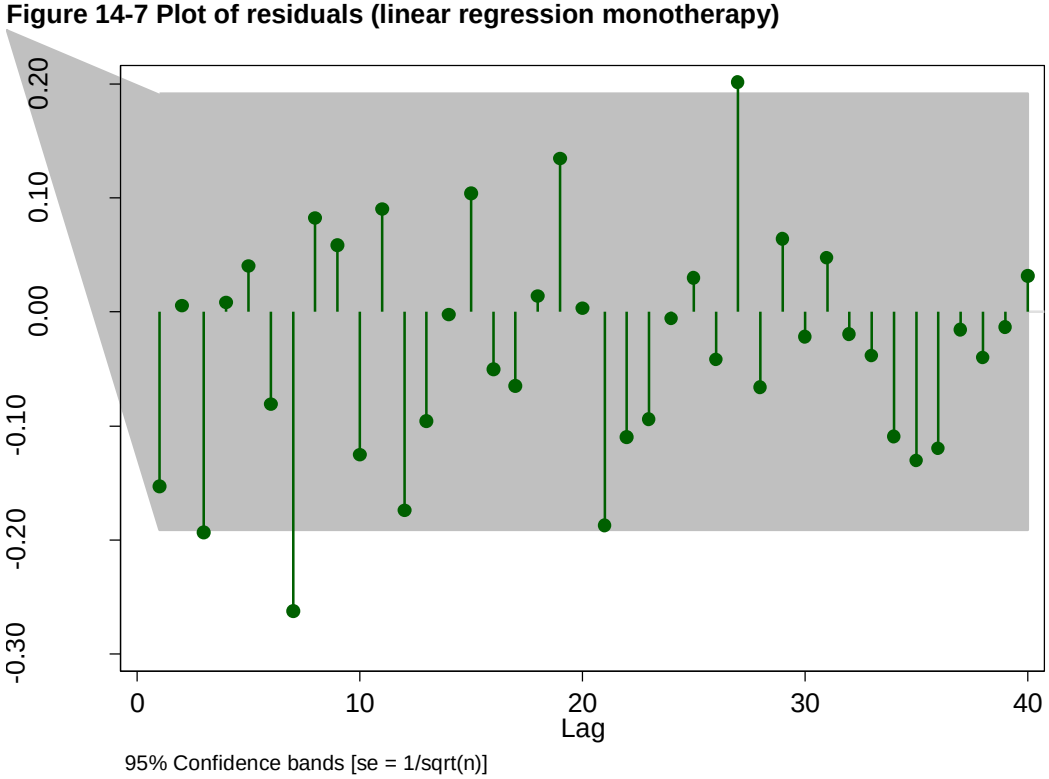


Figure 14-8 Plot of partial residuals (linear regression monotherapy)

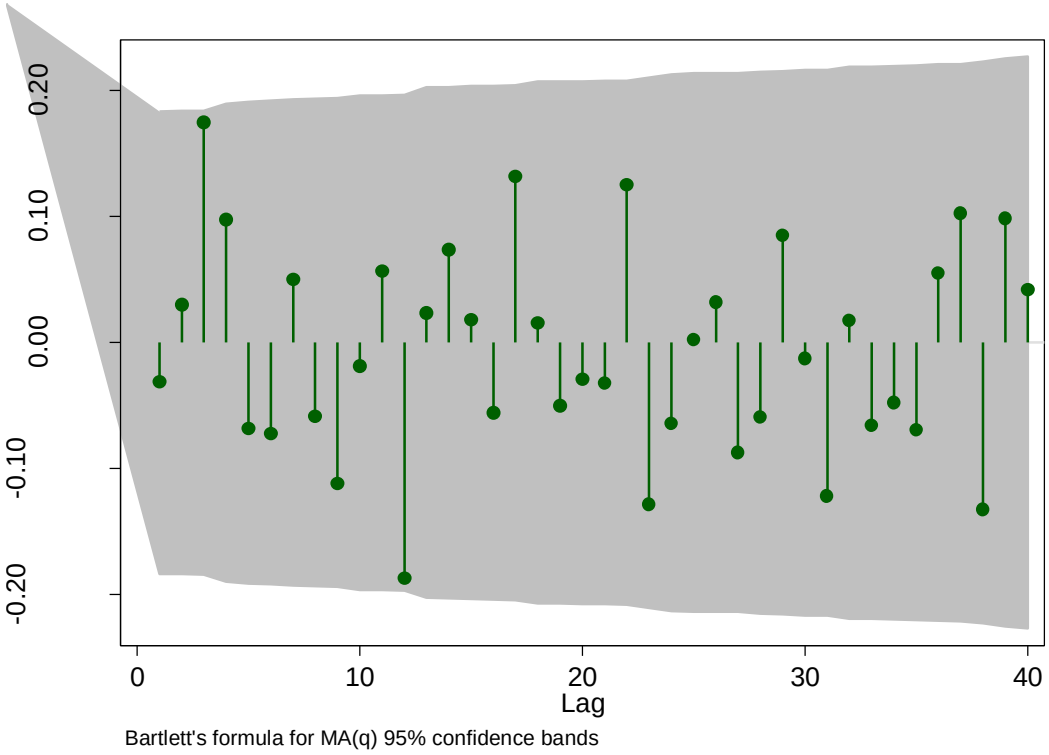


Figure 14-9 Plot of residuals (logistic regression doublet chemotherapy)

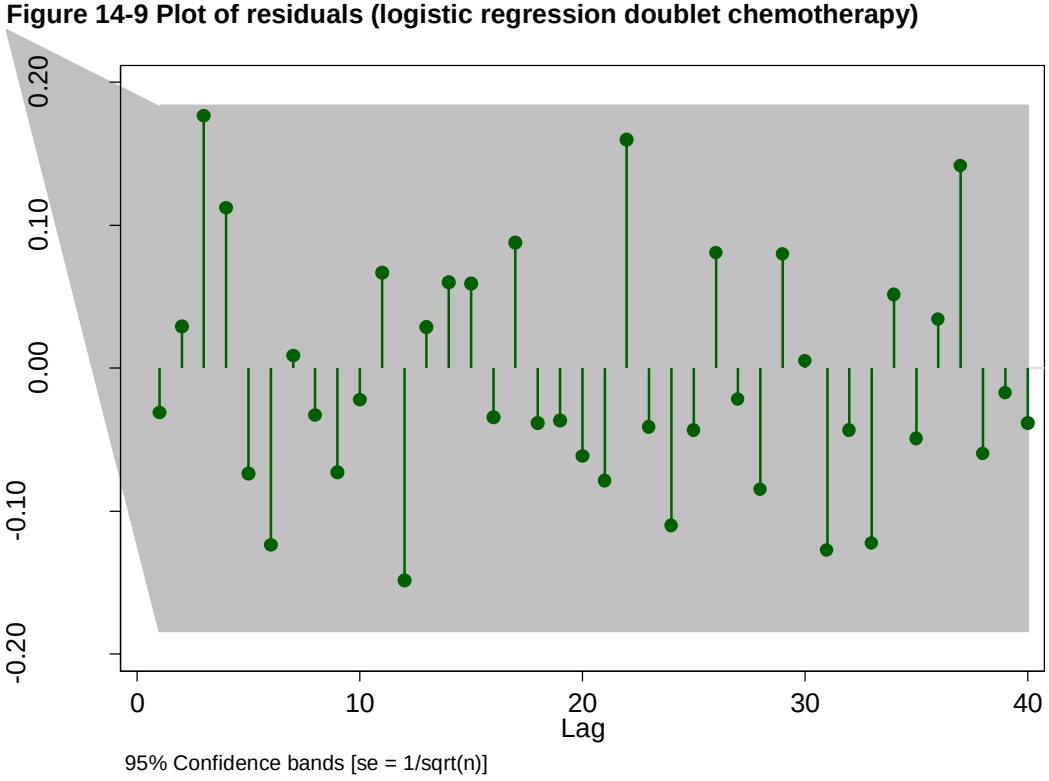


Figure 14-10 Plot of partial residuals (logistic regression doublet chemotherapy)

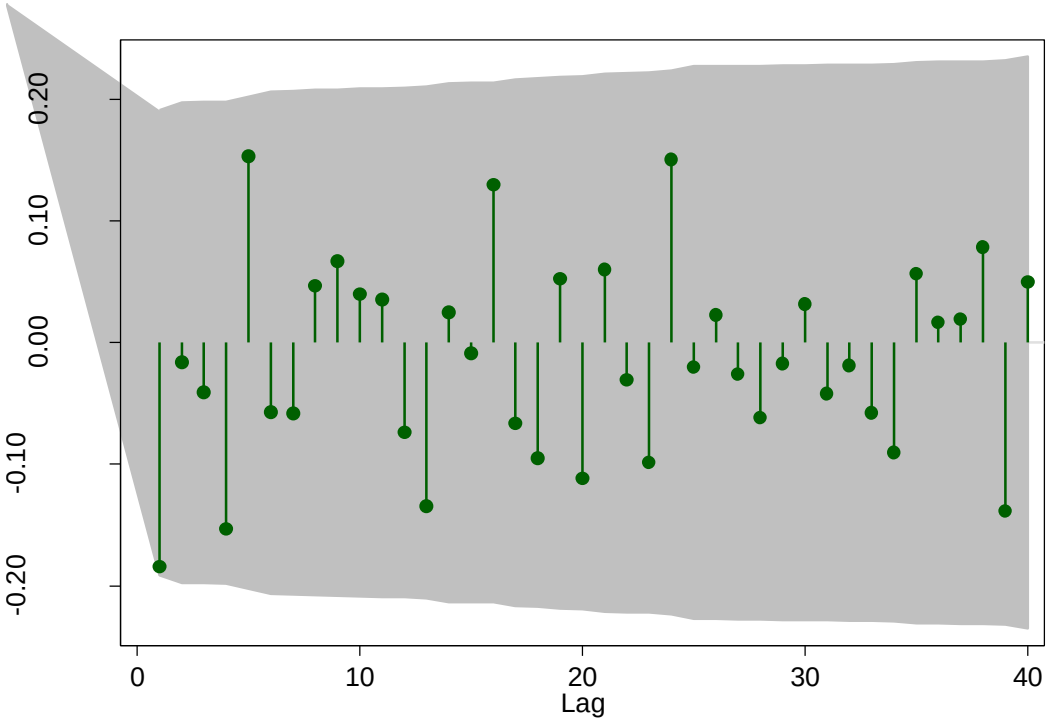


Figure 14-11 Plot of residuals (logistic regression monotherapy)

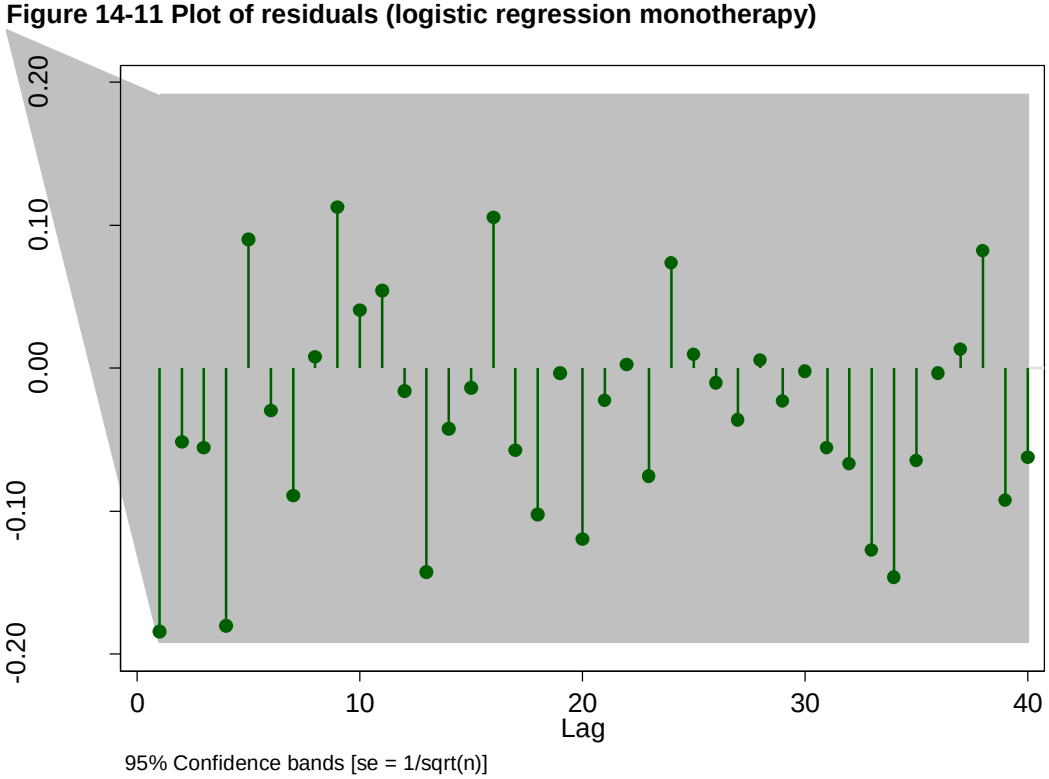


Figure 14-12 Plot of partial residuals (logistic regression monotherapy)

Table 14-5 Results of Durbin-Watson tests for autocorrelation

Analysis	Durbin-Watson test statistic
Linear regression doublet	2.13
Linear regression monotherapy	2.14
Logistic regression doublet	2.05

Logistic regression monotherapy	2.35

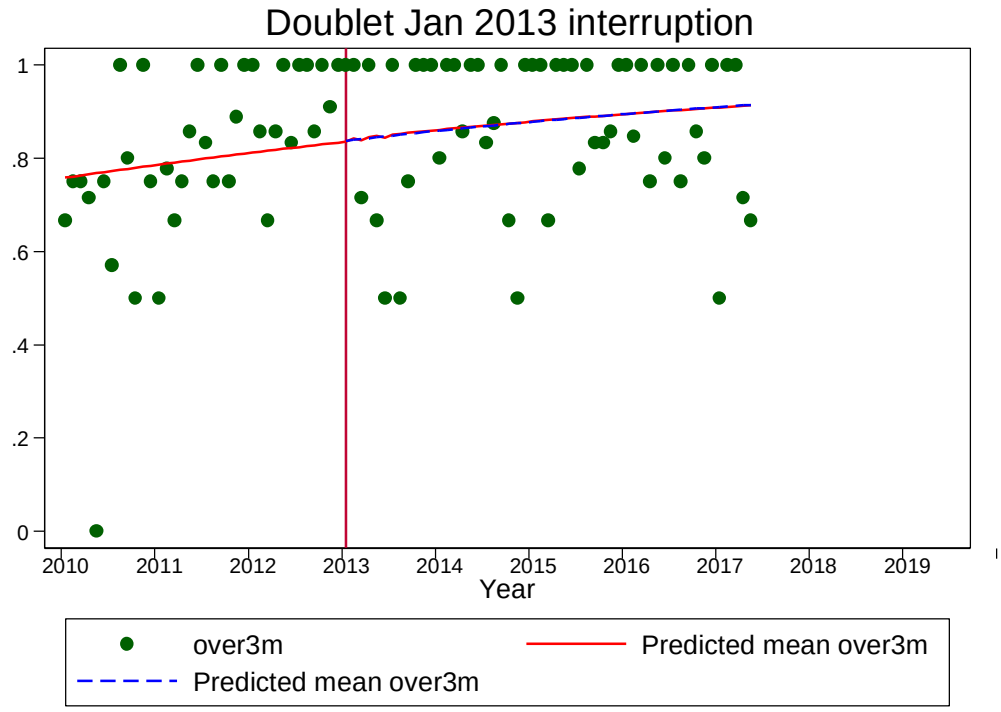


Figure 14-13 ITSA with interruption at January 2013

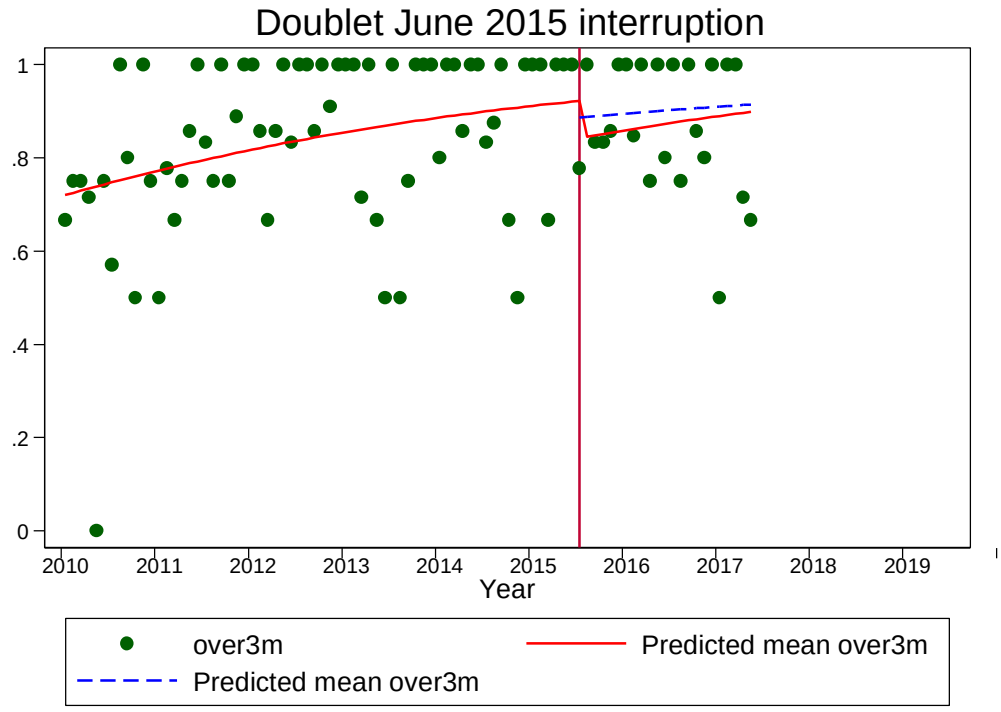


Figure 14-14 ITSA with interruption at June 2015

Table 14-6 Variables used for national analysis

Category of variable	Variable name/description	Name in dataset	Variable type for purposes of analysis	Codes to identify cohort	Variable type	Dataset	Formation of derived variables
Disease information	ICD-10		Cohort Definition	C18, C19, C20	Raw	SMR06	NA
Diagnosis	Previous CRC	previous_crc_cancer	Cohort Definition	No previous diagnosis	Derived by eDRIS	SMR06	Derived by eDRIS. Patient has diagnosis of ICD10 18-20 cancer prior to 2006.
Diagnosis	Diagnosis	incidence_date_fmt	Cohort Definition	2013-2018	Raw	SMR06	NA
Diagnosis	Diagnosis	diagdate_fmt	Cohort Definition	Date of diagnosis between 1st January 2013 and 1st January 2018	Raw	QPI	NA
Disease type	Site		Working variable	C18, C19, C20	Derived	SMR06	C18 and C19=colon (Code=1). C20=Rectum (Code=2).
Disease information	Site	disease	Working variable	1,2	Derived	QPI	site is a raw variable in QPI which is descriptive e.g. Caecum, ascending colon. There are 12 options. These were converted first to numbers. 1 and 3-10 are colon, 2 is "currently unknown await update", 11 is rectum and 12 is not recorded. These numbers are used to create a new, derived variable called "disease". 1 and 3-10 are "1" (Colon), 11 is "2" (rectum) and 2 and 12 are "." (unknown). If the QPI raw variable was 2 or 12, disease type information was taken from SMR06 ICD10 codes when available.
Treatment	Operation code	opcode2 and opcode2b	Working variable	OPCS4 codes	Raw	QPI	NA
Treatment	Operation code	opcode_derived and opcode2b_derived	Cohort Definition	1,2,3,4,9	Derived	QPI	OPCS4 codes grouped according to major operation/palliative operation/bypass/stoma formation/missing
Patient demographic	Age		Working variable	NA	Derived	QPI	Date of diagnosis minus date of birth
Patient demographic	Sex	Sex	Working variable	NA	Raw	SMR06	NA
Disease information	Extended risk stage		Working variable	NA	Derived	QPI	T3/4N0 = risk stage 1; T1-3N1=risk stage 2; T4N1, T1-3N2, T4N2 = risk stage 3.
Patient demographic	Social deprivation	simd2016_sc_quintile	Working variable	NA	Raw	SMR06	

Patient demographic	Charlson score	quan_morbidity	Working variable	NA	Derived	SMR01	Index created by giving patients a score of 1 for each specific co-morbidity related to an episode of inpatient care from the year before diagnosis until death or last follow up (not including malignancy). These conditions are: acute myocardial infarction, chronic heart failure, peripheral vascular disease, cerebrovascular disease, dementia, COPD, rheumatoid disease, peptic ulcer disease, mild liver disease, diabetes with no complications, diabetes with complications, hemi or paraplegia, renal disease, moderate to severe liver disease, HIV. These scores were adjusted using quan weightings (x 2 for dementia and mod/sever liver disease, x4 for HIV).
Patient demographic	Charlson group	charlson_group	Working variable	NA	Derived	SMR01	Patients with no significant conditions given a grouping of zero, those with one are in group 1 and those with more than one are in group 2.
Treatment	Regimen	regimen	Cohort Definition	Patients receiving advanced/metastatic regimens at their first cycle of chemotherapy after major surgery removed.	Raw	ChemoCare	NA
Regimen calculated	Regimen calculated from individual drugs	regimen_calc	Cohort Definition	CAPOX, FOLFOX, capecitabine, 5-fluorouarcil	Derived	ChemoCare	This is derived from the individual medications prescribed for an individual patient. There is a variable "regimen" in the raw data but regimens which include the same drugs may have different names within different ChemoCare systems. Regimen calculated names therefore derived using the actual drugs received.
First regimen		first_regimen	Working variable	NA	Derived	ChemoCare	This is the first regimen (calculated) received after major surgery for patients with stage II/III CRC.
Patient demographic	Location	location	Working variable	NA	Derived	ChemoCare	The location of the ChemoCare prescription by Cancer Network. The North of Scotland covers Tayside, Grampian and the Highlands.
Treatment	Treatment time in weeks	total_weeks	Working variable	NA	Derived	ChemoCare	Based on the number of cycles of treatment received in the adjuvant setting. Cycles are either 2 weekly or 3 weekly depending on the medications used in each cycle. Two or three weeks are added to this variable to account for the time taken for the last cycle.

Treatment	Indicator of if patient received over 3 months of treatment	over3m	Working variable	NA	Derived	ChemoCare	Derived from treatment time in weeks. If treatment time is 13 weeks or over the patient is categorised as having over 3 months of treatment (code=1). If a patient has less than 13 weeks i.e. up to 12 weeks and 6 days, the patient is coded as having less than 3 months of treatment (code=0).
Patient demographic	Age group (under 70 versus 70+)	age_group	Working variable	NA	Derived		Age_group=0 if age at diagnosis is 70 years or less and Age group =1 if age at diagnosis is over 70
Patient demographic	SCOT eligibility	SCOTelig	Working variable	NA	Derived	QPI and Chemocare	Patients were deemed “SCOT eligible” if they had stage II disease with high risk features or stage III CRC and commenced treatment with fluoropyrimidine-oxaliplatin doublet chemotherapy within 11 weeks of major CRC surgery. SCOT ineligible=0, SCOT eligible=1.
Disease information	Duke's stage (QPI raw)	dukes_original (in raw database called dukes)	Working variable	NA	raw	QPI	
Disease information	Duke's stage (QPI raw)	dukes	Working variable	NA	derived	QPI	Original variable renamed dukes_original. Dukes convered to a byte variable with the following labels: 1 ="Dukes A" 2 ="Dukes B" 3= "Dukes C1" 4 ="Dukes C2" 5= "Dukes D" 96= "Not applicable" 99 ="Not recorded"
Disease information	Colorectal stage derived from QPI TNM	dukes_derived_qpi	Cohort Definition	2 and 3	Derived	QPI	Colorectal stage derived from TNM staging. 1 ="Dukes A" 2 ="Dukes B" =3 "Dukes C" 4 ="Dukes D". When this information is missing, it is derived directly from the "dukes" variable (TNM assumed to be more reliable than raw dukes).
Disease information	T-stage (raw)	finalt_original (called finalt in raw database)	Working variable	NA	raw	QPI	NA
Disease information	T-stage (derived)	finalt	Working variable	NA	derived	QPI	Original variable renamed finalt_original. Finalt converted to a byte variable with the following labels: 1=T0, 2=T1, 3=T2, 4=T3, 5=T4, 6=T4a, 7=T4b, 8=Tx, 9=Tis, 96=not applicable, 99=not recorded (including missing)
Disease information	N-stage (raw)	finaln_original (called finaln in raw database)	Working variable	NA	raw	QPI	NA
Disease information	N-stage (derived)	finaln	Working variable	NA	Derived	QPI	Original variable renamed finaln_original. Finaln converted to a byte variable with the following labels: 1=N0, 2=N1, 3=N1a, 4=N1b, 5=N1c, 6=N2, 7=N2a, 8=N2b, 9=NX, 96=not applicable, 99=not recorded (including missing)
Disease information	M-stage (raw)	finalm_original	Working variable	NA	raw	QPI	NA

Disease information	M-stage (derived)	finalm	Working variable	NA	Derived	QPI	Original variable renamed finalm_original. Finalm converted to a byte variable with the following labels: 1=M0, 2=M1, 3=M1a, 4=M1b, 5=M1c, 6=Mx, 96=not applicable, 99=not recorded (including missing)
Disease information			Working variable	NA			
Disease information	Dukes (SMR06 raw)	stage_colorectal	Working variable	NA	raw	SMR06	NA
Disease information	Dukes (SMR06 derived)	dukes_stage	Working variable	NA	Derived	SMR06	Original variable (stage_colorectal) converted to a numerical variable with the following values: 1=A, 2=B, 3=C, C1, C2, 4=D, 9=9.
Disease information	Clinical T-stage (raw)	stage_clinical_t	Working variable	NA	raw	SMR06	NA
Disease information	Clinical T-stage (derived)	clinical_stage_t	Working variable	NA	Derived	SMR06	Derived from stage_clinical_t. 1=1/1a/1b/1b/1m/is. 2=2/2a/2b/2c. 3=3/3a/3b/3c. 4=4/4a/4b/4c/4d. 5=X. 9=missing
Disease information	Clinical N-stage (raw)	stage_clinical_n	Working variable	NA	raw	SMR06	NA
Disease information	Clinical N-stage (derived)	clinical_stage_n	Working variable	NA	Derived	SMR06	Derived from stage_clinical_n. 0=0, 1=1/1a/1b/1b/1m/is. 2=2/2a/2b/2c. 3=3/3a/3b/3c. 5=X. 9=missing
Disease information	Clinical M-stage (raw)	stage_clinical_m	Working variable	NA	raw	SMR06	NA
Disease information	Clinical M-stage (derived)	clinical_stage_m	Working variable	NA	Derived	SMR06	Derived from stage_clinical_m. 0=0, 1=1/1a/1b/1c. 5=X. 9=missing.
Disease information	Pathological T-stage (raw)	stage_pathologic_t	Working variable	NA	raw	SMR06	NA
Disease information	Pathological T-stage (derived)	path_stage_t	Working variable	NA	Derived	SMR06	Derived from stage_pathologic_t. 1=1/1a/1b/1b/1m/is. 2=2/2a/2b/2c. 3=3/3a/3b/3c. 4=4/4a/4b/4c/4d. 5=X. 9=missing
Disease information	Pathological N-stage (raw)	stage_pathologic_n	Working variable	NA	raw	SMR06	
Disease information	Pathological N-stage (derived)	path_stage_n	Working variable	NA	Derived	SMR06	Derived from stage_pathologic_n. 0=0, 1=1/1a/1b/1b/1m/is. 2=2/2a/2b/2c. 3=3/3a/3b/3c. 5=X. 9=missing
Disease information	Pathological M-stage (raw)	stage_pathologic_m	Working variable	NA	raw	SMR06	
Disease information	Pathological M-stage (derived)	path_stage_m	Working variable	NA	Derived	SMR06	Derived from stage_pathologic_m. 0=0, 1=1/1a/1b/1c. 5=X. 9=missing.

Disease information	Final T-stage	finalt_SMR	Working variable	NA	Derived	SMR06	Clinical t stage (clinical_stage_t) used. When this is missing, path_stage_t used. Also, clinical_stage_t_replaced with path_stage_t if path_stage_t upgrades/worse grade than clinical staging.
Disease information	Final N-stage	finaln_SMR	Working variable	NA	Derived	SMR06	Clinical n stage (clinical_stage_n) used. When this is missing, path_stage_n used. Also, clinical_stage_t_replaced with path_stage_n if path_stage_n upgrades/worse grade than clinical staging.
Disease information	Final M-stage	finalm_SMR	Working variable	NA	Derived	SMR06	Clinical m stage (clinical_stage_m) used. When this is missing, path_stage_m used. Also, clinical_stage_m_replaced with path_stage_m if path_stage_m upgrades/worse grade than clinical staging.
Disease information	Colorectal stage (derived)	stage_derived	Working variable	NA	Derived	SMR06	Derived from finalt/n/m_SMR according to AJCC staging (stage I/II/III/IV/mising)
Disease information	Colorectal stage (derived)	dukes_derived_SMR	Cohort Definition	2 and 3	Derived	SMR06	Overall stage derived from SMR. Derived TNM stage (stage_derived) used as baseline. If this information is missing, dukes_derived fills in the missing variables.
Disease information	TNM T stage	finalt_cohort	Working variable	NA	Derived	SMR06 & QPI	Used to specify extended risk stage. TNM staging from QPI used. 1=finalt1/2/9, 2=finalt3, 3=finalt4, 4=finalt5/6/7, 9=finalt5/96/99. Any missing variables in QPI replaced if SMR06 not missing.
Disease information	TNM N stage	finaln_cohort	Working variable	NA	Derived	SMR06 & QPI	Used to specify extended risk stage. TNM staging from QPI used. Any missing variables in QPI replaced if SMR06 not missing.

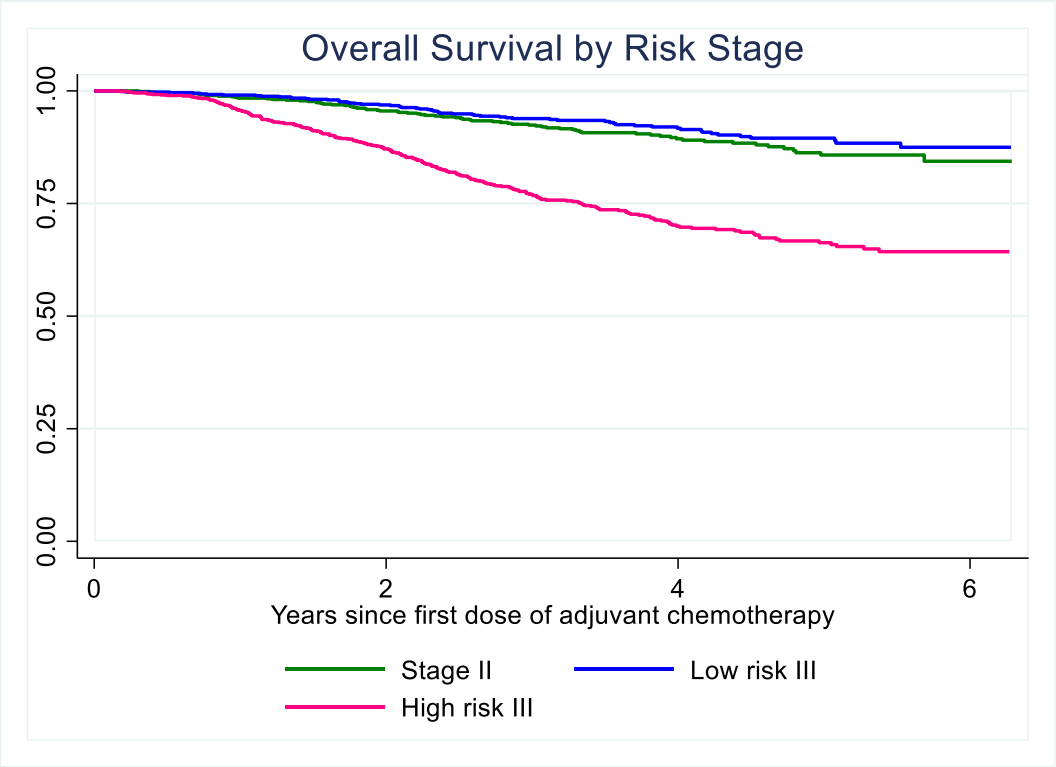


Figure 14-15 OS Kaplan-Meier curve by risk stage

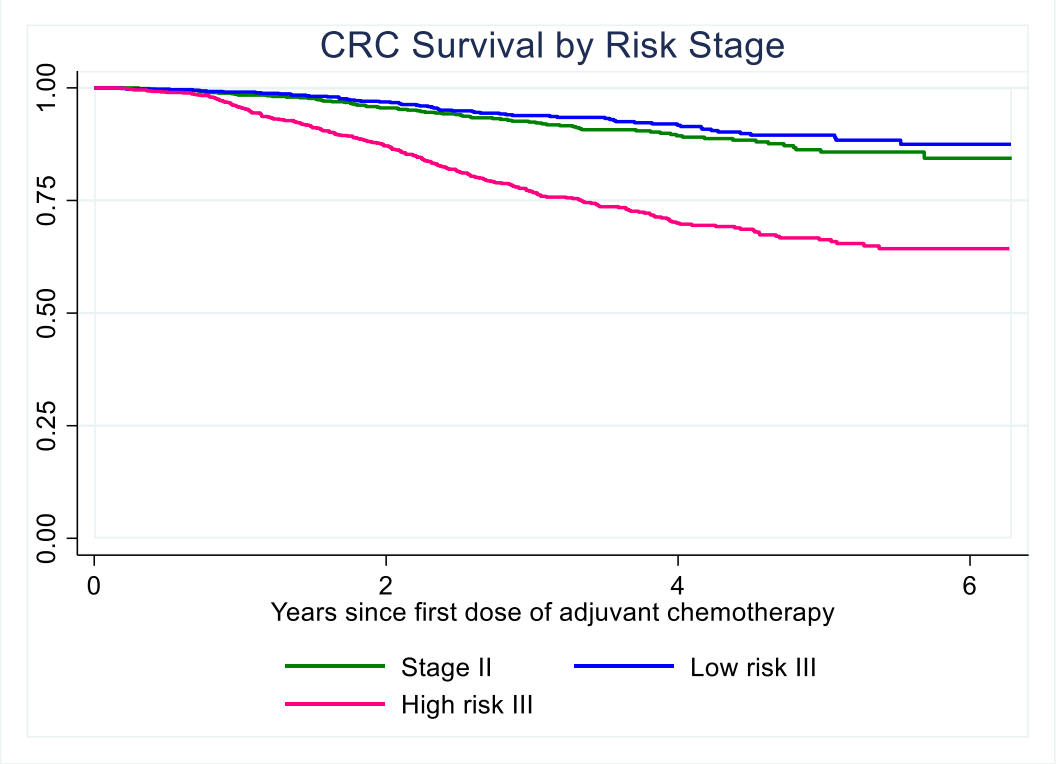


Figure 14-16 CRC Kaplan-Meier curve by risk stage

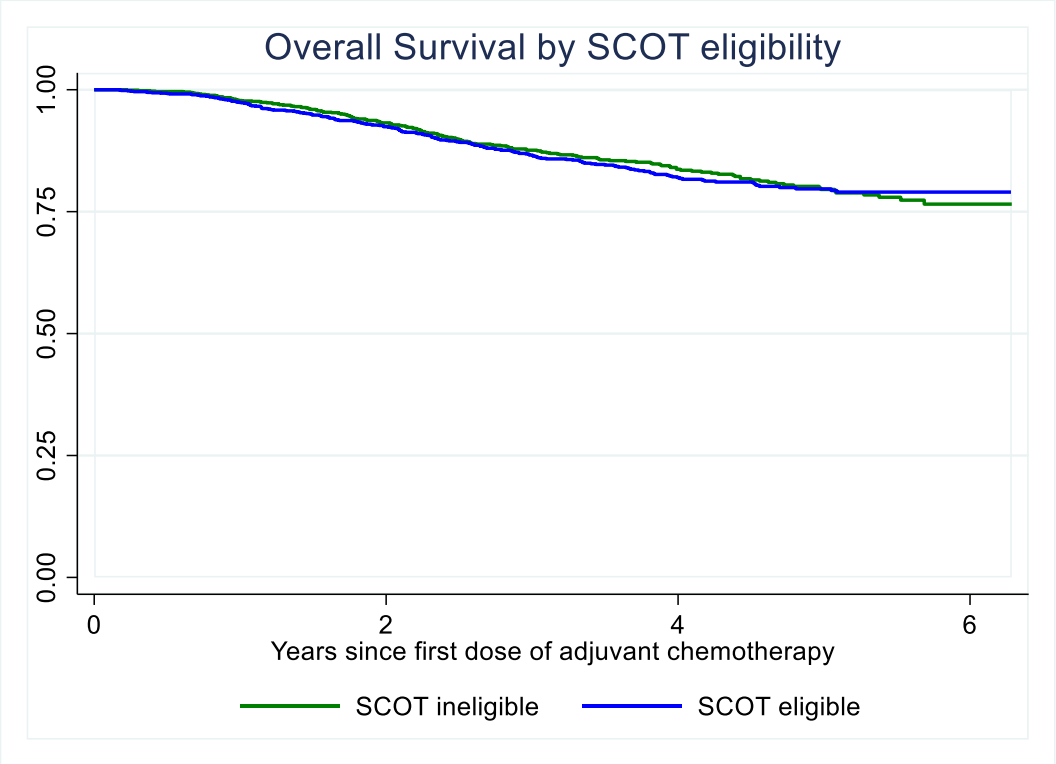


Figure 14-17 OS Kaplan-Meier curve by SCOT eligibility

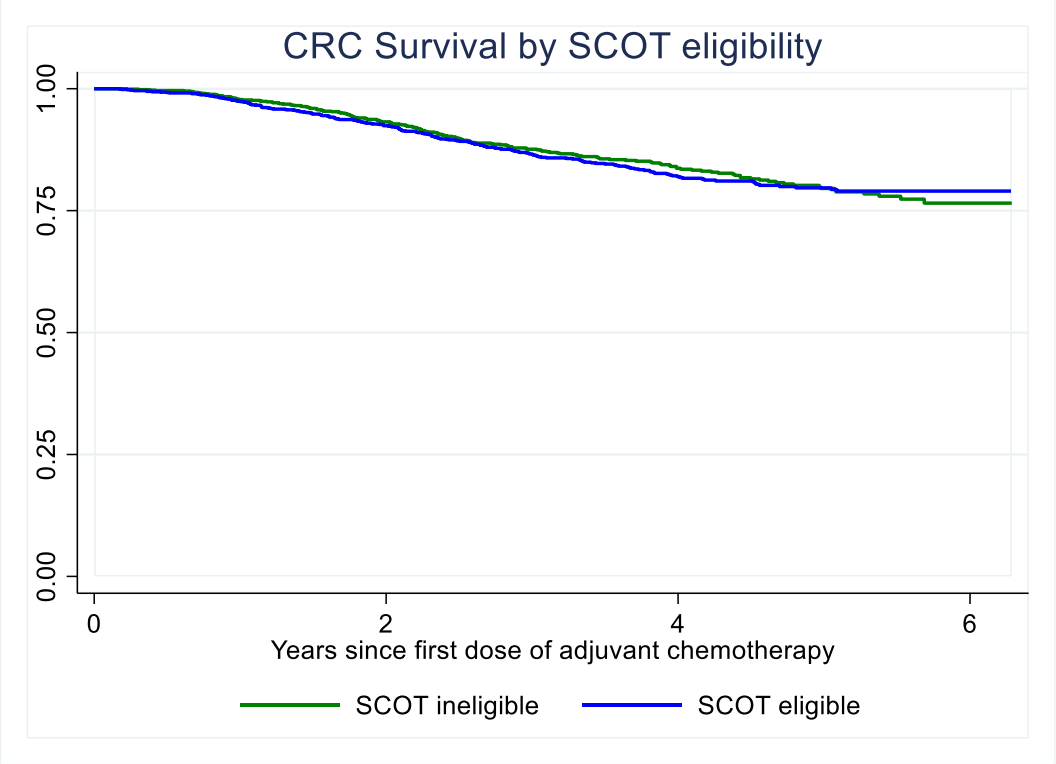


Figure 14-18 CRC Kaplan-Meier curve by SCOT eligibility

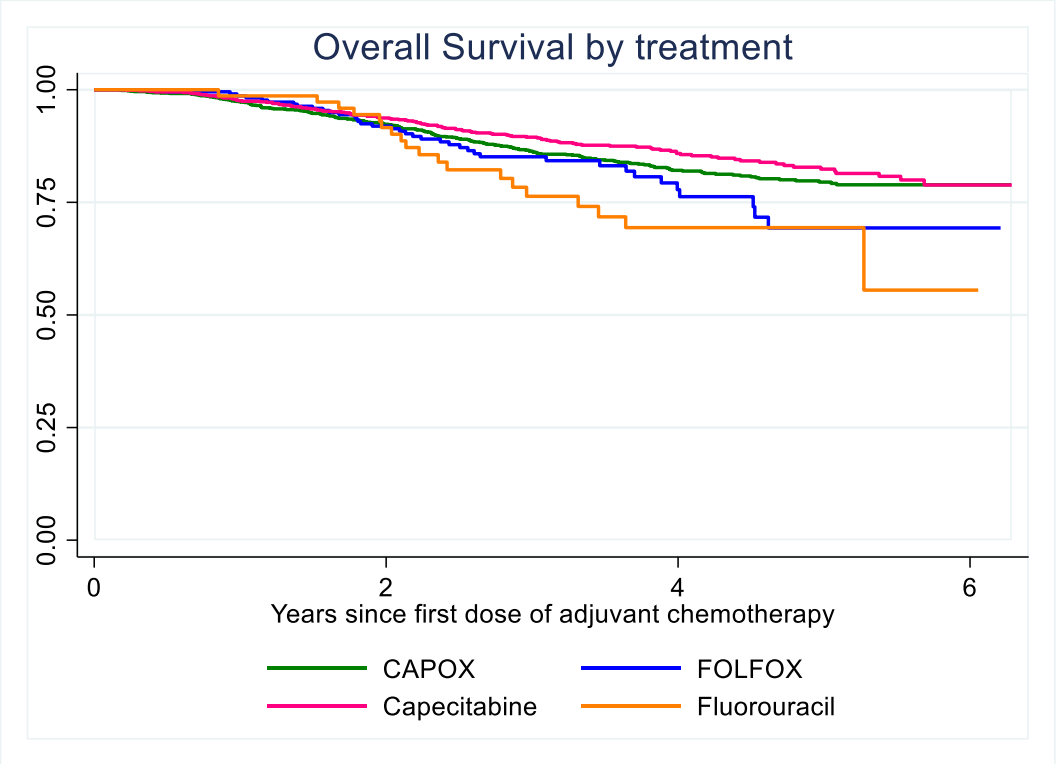


Figure 14-19 OS Kaplan-Meier curve by regimen

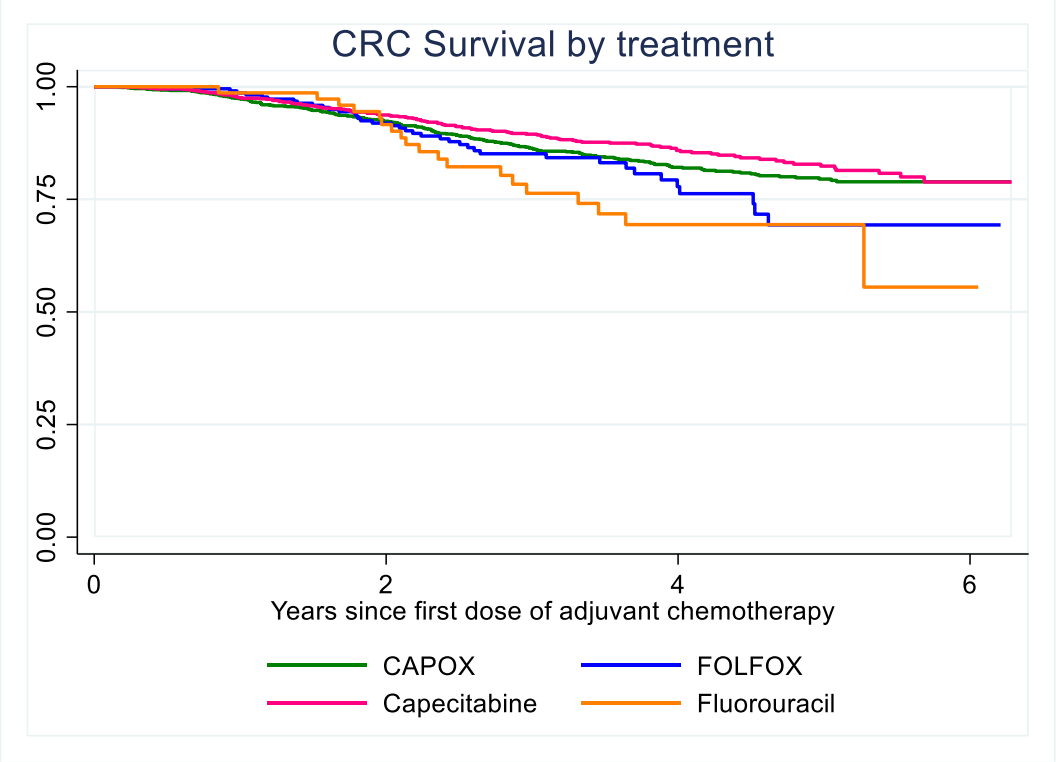


Figure 14-20 CRC Kaplan-Meier curve by regimen

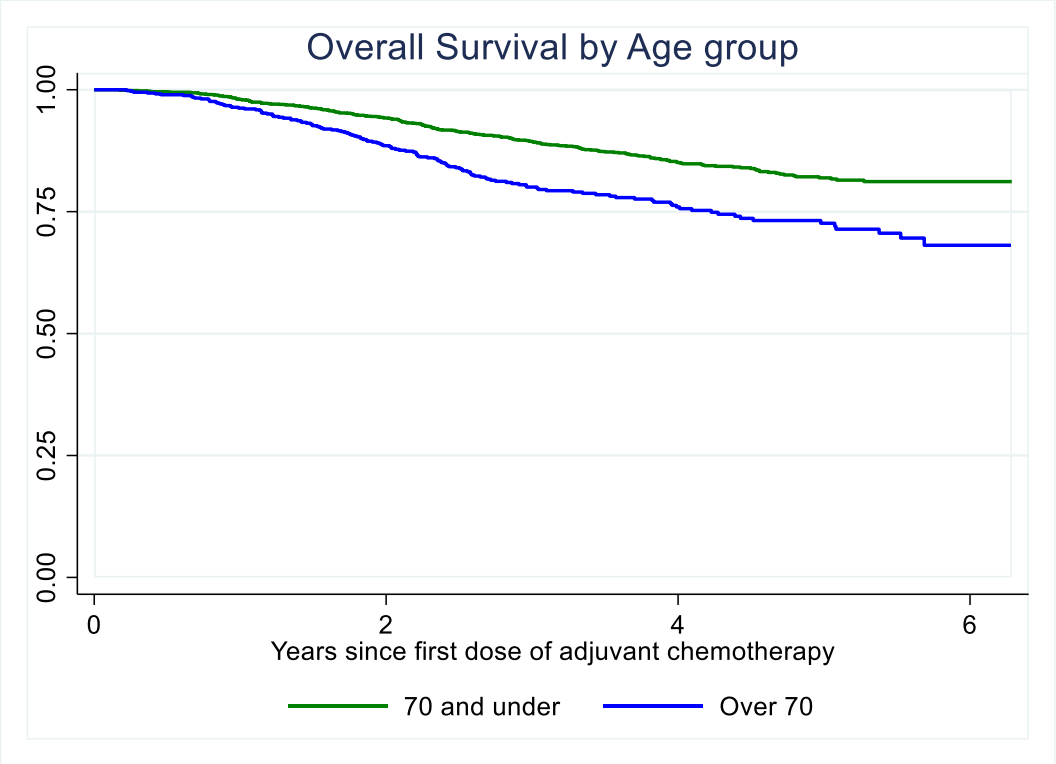


Figure 14-21 OS Kaplan-Meier curve by age group

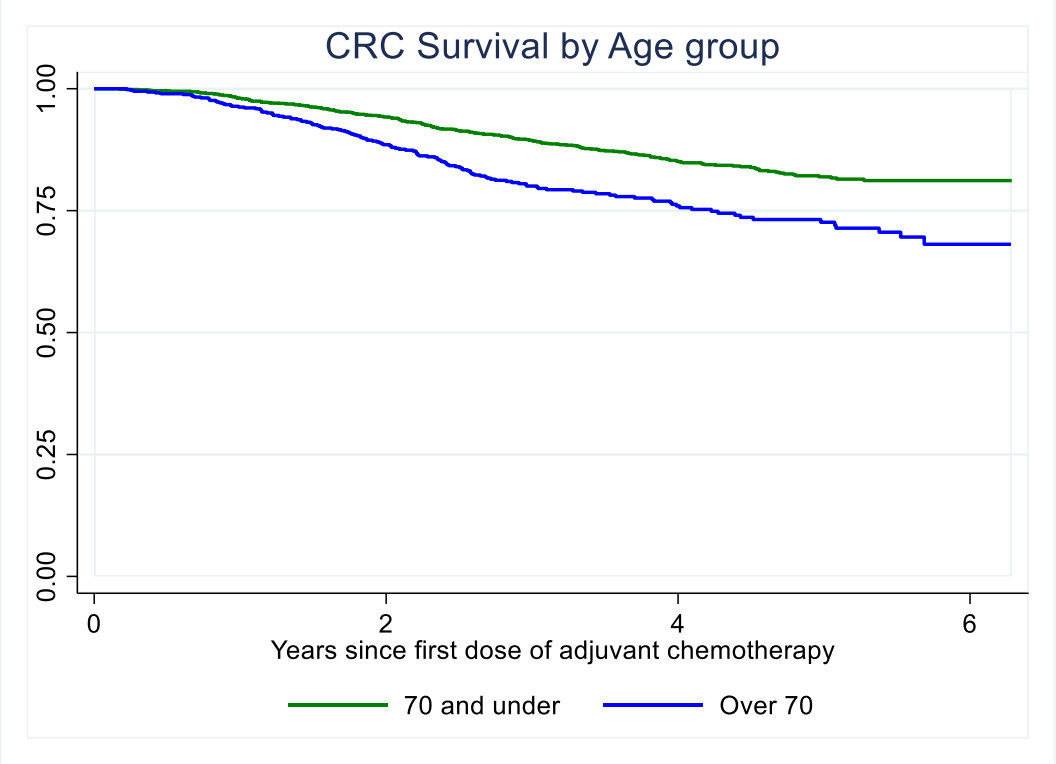


Figure 14-22 CRC Kaplan-Meier curve by age group

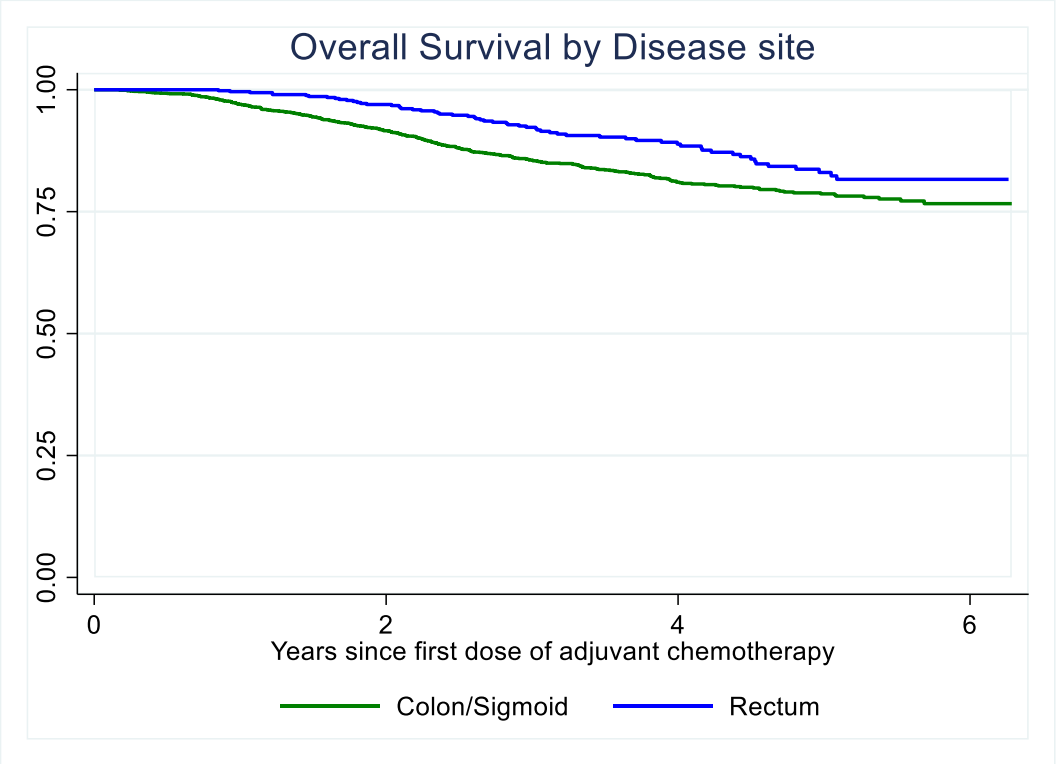


Figure 14-23 OS Kaplan-Meier curve by disease site

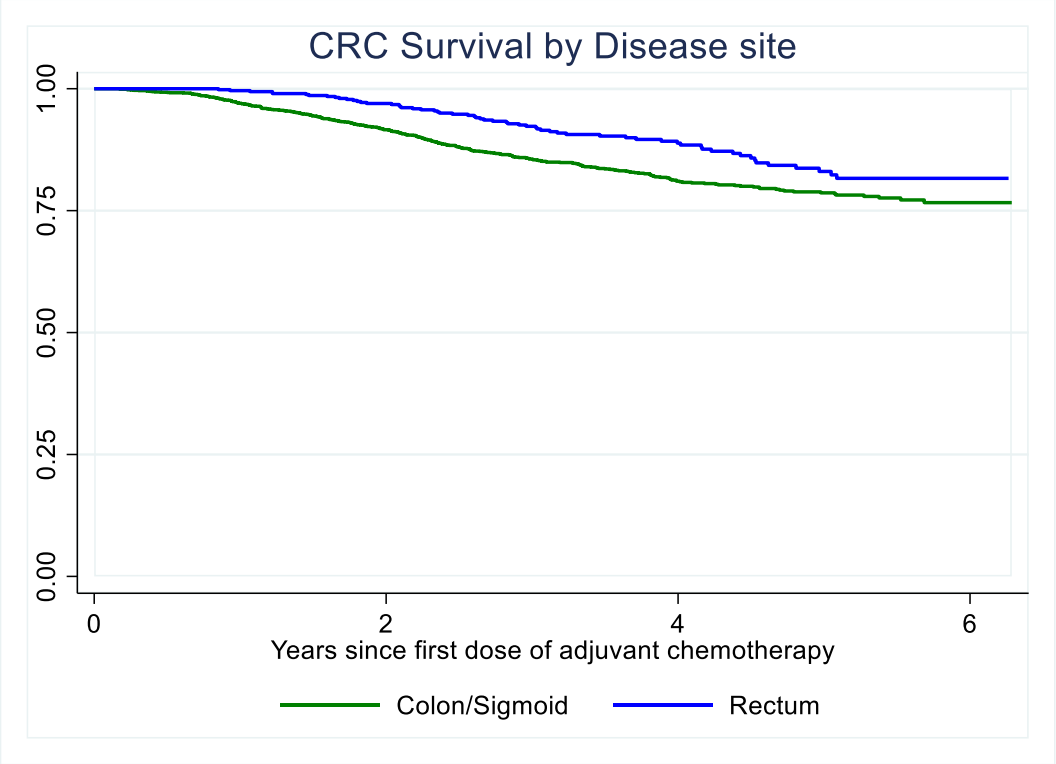


Figure 14-24 CRC Kaplan-Meier curve by disease site

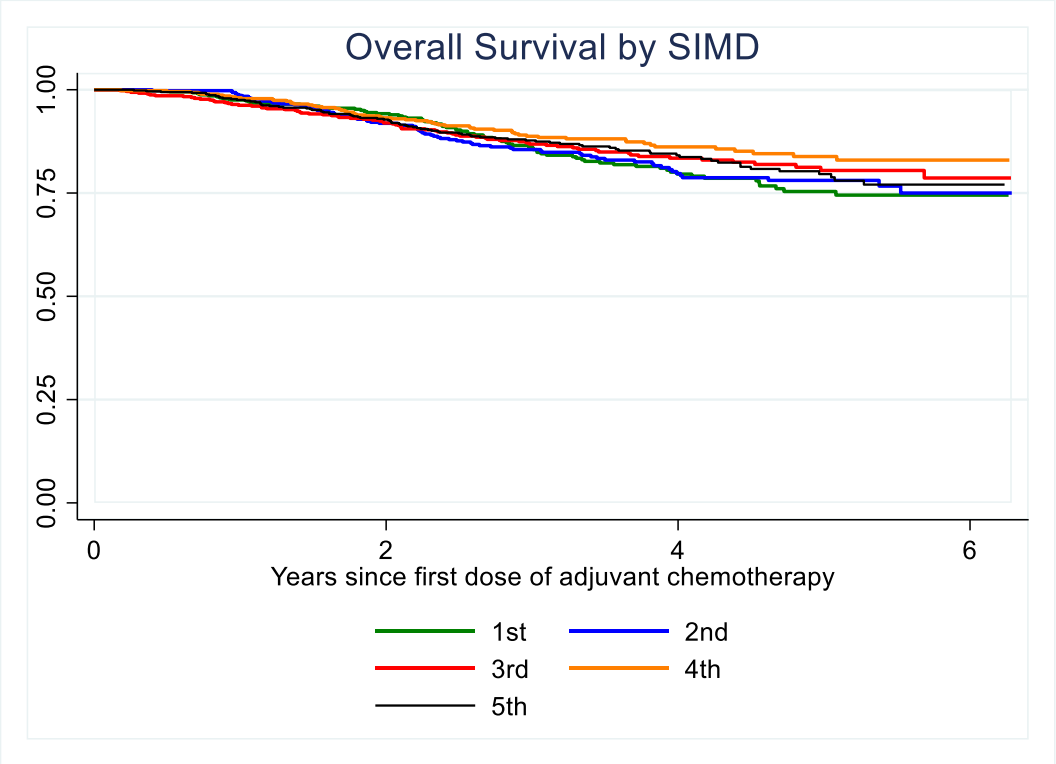


Figure 14-25 OS Kaplan-Meier curve by SIMD

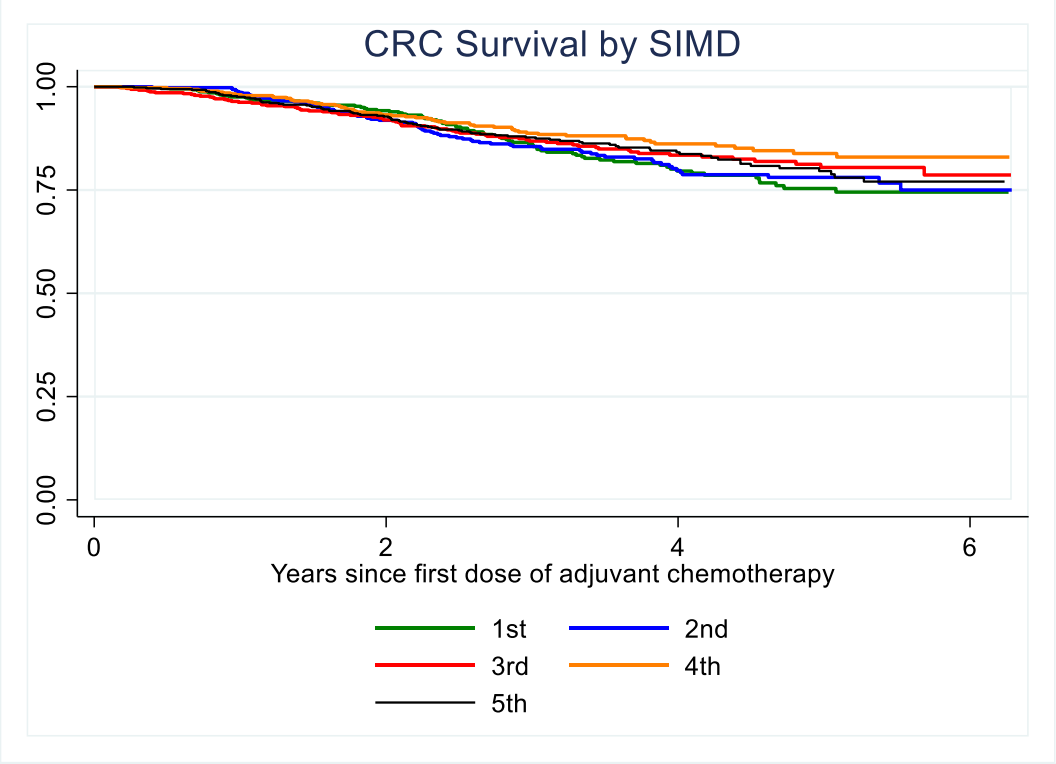


Figure 14-26 CRC Kaplan-Meier curve by SIMD

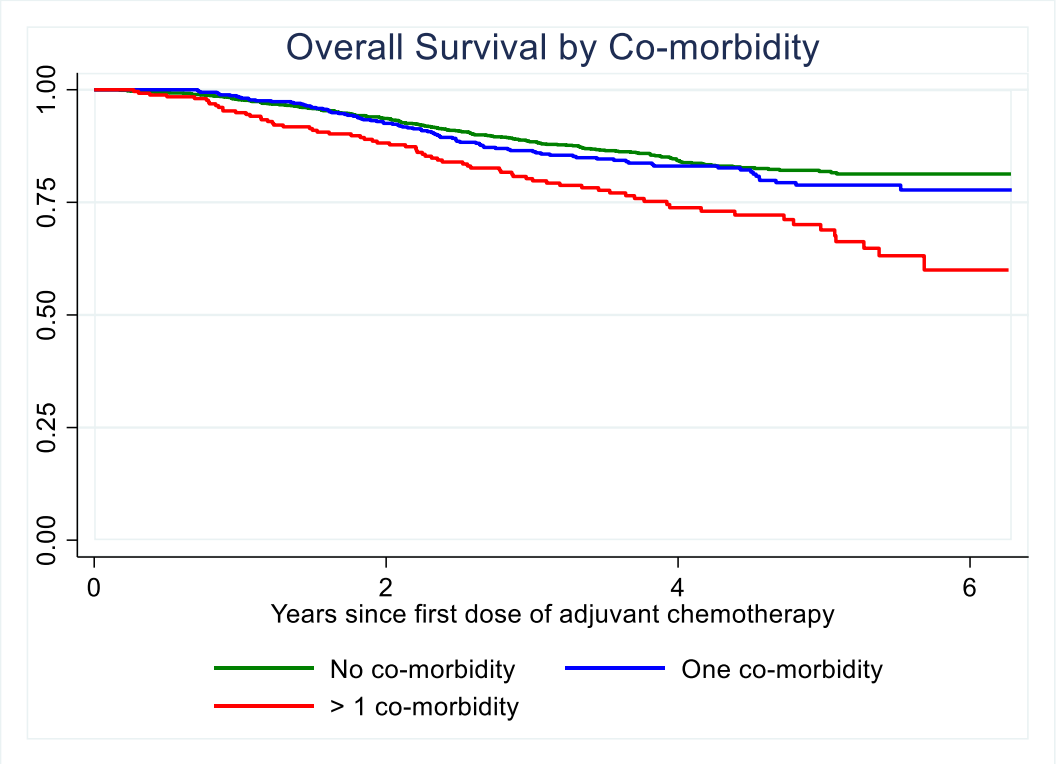


Figure 14-27 OS Kaplan-Meier curve by Charlson group

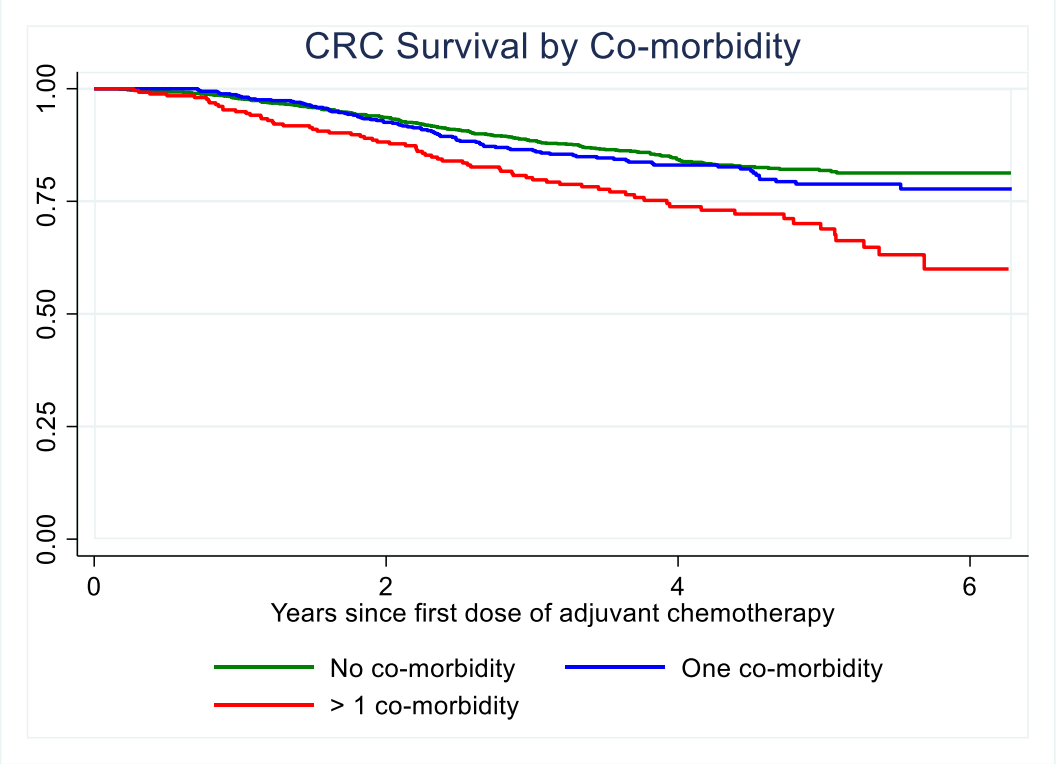


Figure 14-28 CRC Kaplan-Meier curve by Charlson group

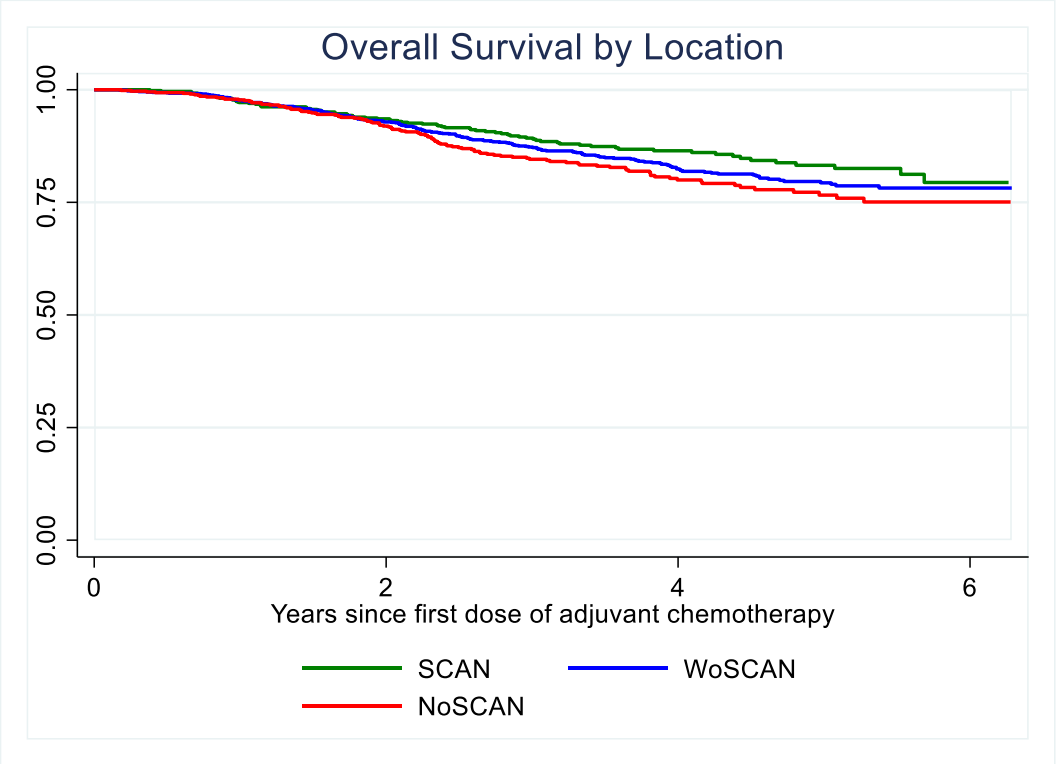


Figure 14-29 OS Kaplan-Meier curve by location

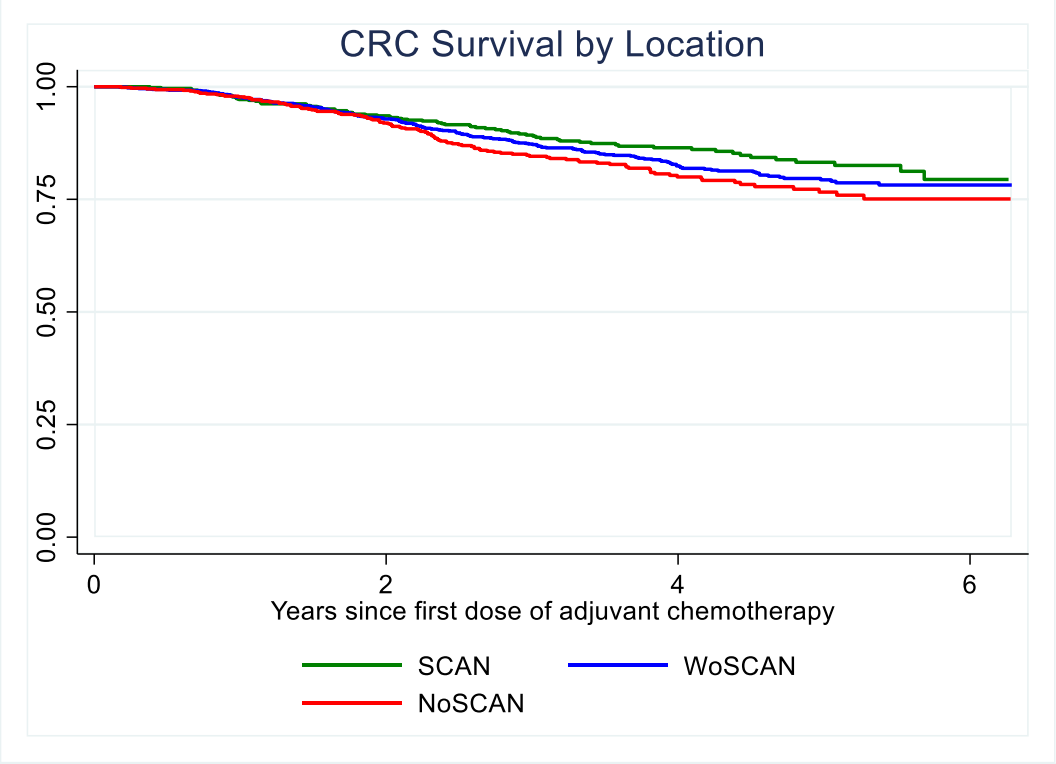


Figure 14-30 CRC Kaplan-Meier curve by location

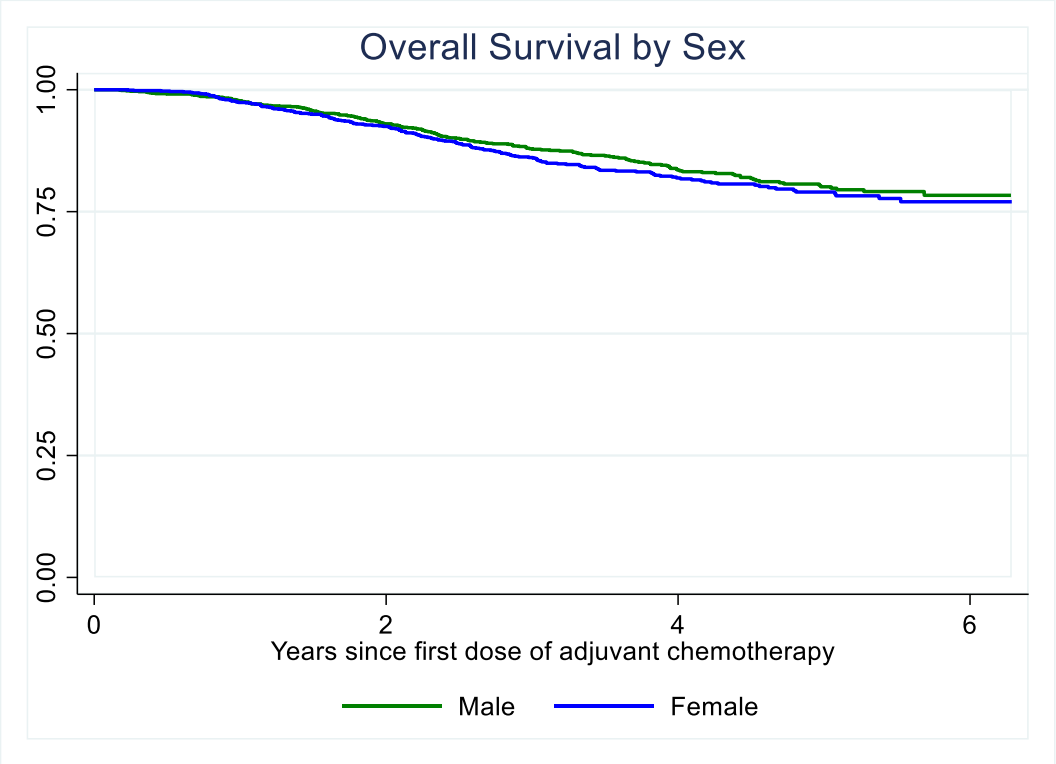


Figure 14-31 OS Kaplan-Meier curve by sex

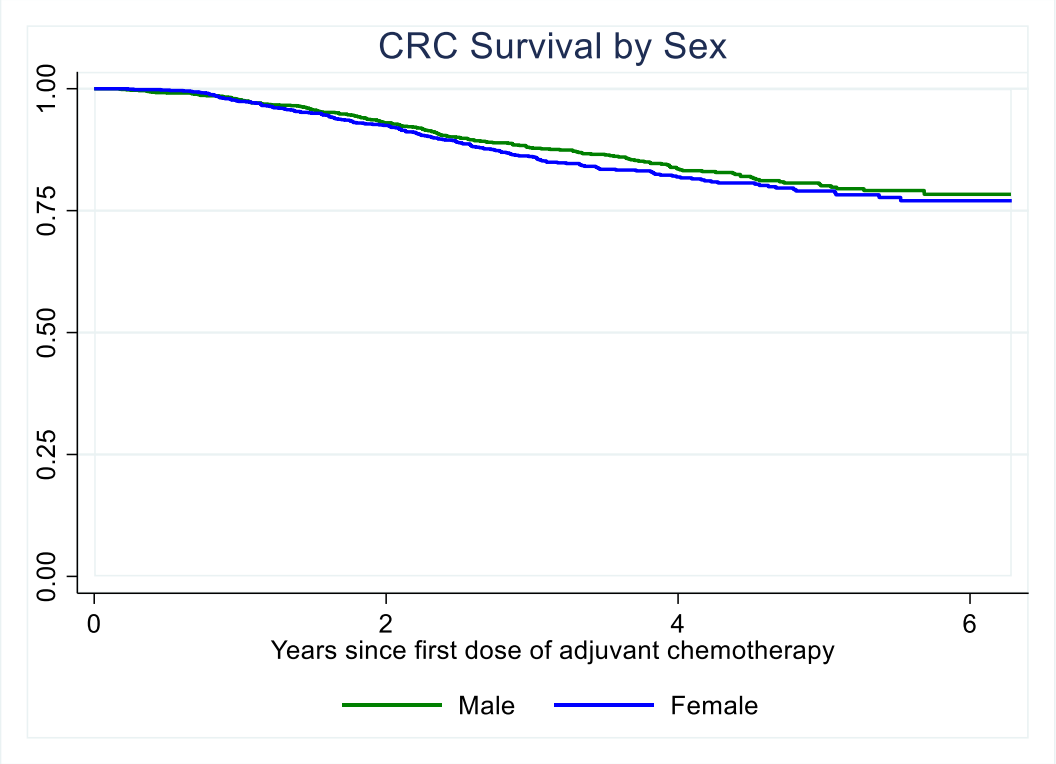


Figure 14-32 CRC Kaplan-Meier curve by sex

Table 14-7 OS and CRC survival estimates

	Overall 3 year survival	CRC specific 3 year survival		Overall 5 year survival	CRC specific 5 year survival	Logrank OS	Logrank CRC
Overall	86 (84-87)	88 (87-90)	Overall	79 (77-81)	82 (80-84)	NA	NA

Risk stage			Risk stage			<0.001	<0.001
II	92 (89-94)	94 (92-95)	II	86 (82-89)	89 (86-92)		
LR3	93 (91-95)	95 (93-96)	LR3	89 (85-91)	91 (88-93)		
HR3	75 (72-78)	78 (75-81)	HR3	65 (61-69)	69 (65-72)		
SCOT eligible			SCOT eligible			0.8946	0.3927
No	86 (84-89)	89 (87-91)	No	79 (75-82)	83 (80-85)		
Yes	86 (83-88)	88 (85-89)	Yes	79 (76-82)	82 (79-84)		
Regimen			Regimen			0.0074	0.0018
CAPOX	86 (83-87)	87 (85-89)	CAPOX	79 (76-82)	81 (78-84)		
FOLFOX	84 (78-89)	87 (81-91)	FOLFOX	70 (59-79)	74 (63-83)		
CAP	88 (86-90)	81 (77-85)	CAP	91 (89-93)	86 (82-88)		
FU	76 (64-85)	79 (67-88)	FU	69 (55-80)	75 (61-84)		
Age group			Age group			<0.001	<0.001
<70	88 (87-90)	90 (89-92)	<70	81 (79-84)	84 (82-86)		
>70	79 (75-83)	82 (78-85)	>70	71 (66-76)	76 (71-80)		
Disease site			Disease site			0.004	0.0094
Colon	85 (83-87)	87 (85-89)	Colon	78 (76-80)	81 (79-83)		
Rectum	90 (87-93)	92 (88-94)	Rectum	82 (76-86)	85 (80-88)		
SIMD			SIMD			0.2061	0.1567
1	84 (79-87)	86 (82-89)	1	74 (68-79)	78 (72-83)		
2	86 (83-89)	86 (82-89)	2	80 (75-84)	80 (75-84)		
3	87 (84-90)	88 (85-91)	3	78 (73-83)	84 (79-87)		
4	85 (81-88)	90 (87-93)	4	78 (73-82)	86 (81-90)		
5	88 (85-91)	90 (87-92)	5	83 (78-87)	82 (77-86)		
Charlson			Charlson			<0.001	0.07
0	88 (86-89)	89 (87-91)	0	81 (79-84)	84 (81-86)		
1	85 (82-88)	87 (84-90)	1	79 (74-83)	81 (76-84)		
>1	79 (73-83)	85 (80-89)	>1	66 (58-73)	76 (69-82)		

Location			Location			0.0675	0.0304
SCAN	88 (85-91)	89 (86-92)	SCAN	83 (78-86)	84 (80-87)		
WoS	86 (84-88)	89 (87-91)	WoS	79 (75-81)	83 (80-86)		
NoS	84 (80-87)	85 (82-88)	NoS	76 (71-80)	79 (74-82)		
Sex			Sex			0.302	0.3524
Male	87 (85-89)	79 (76-81)	Male	90 (88-91)	82 (80-85)		
Female	84 (82-87)	78 (75-81)	Female	87 (84-89)	82 (78-84)		

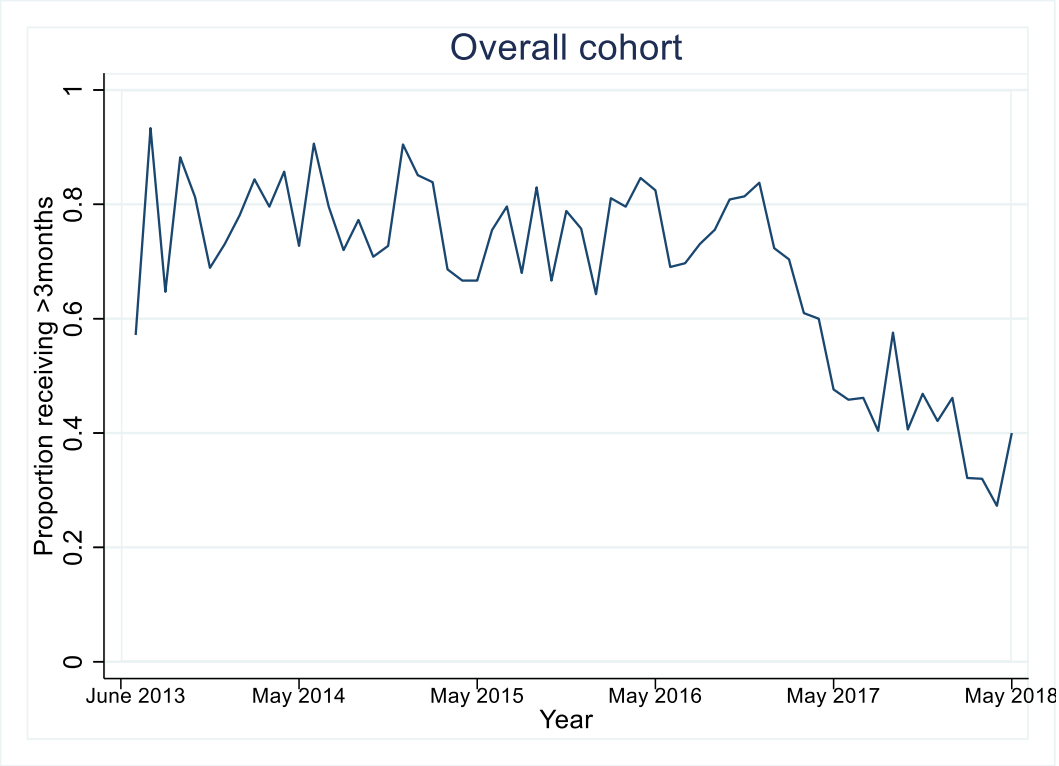


Figure 14-33 Time series for national cohort to check for seasonality

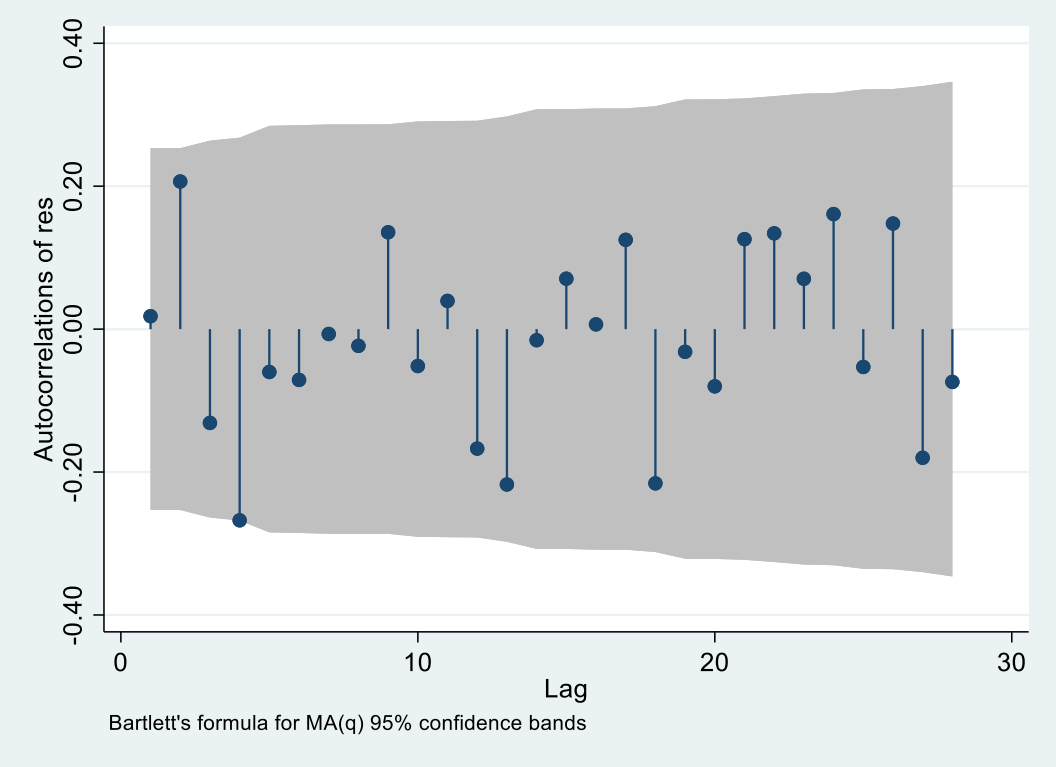


Figure 14-34 Plot of residuals for time series (national overall cohort)

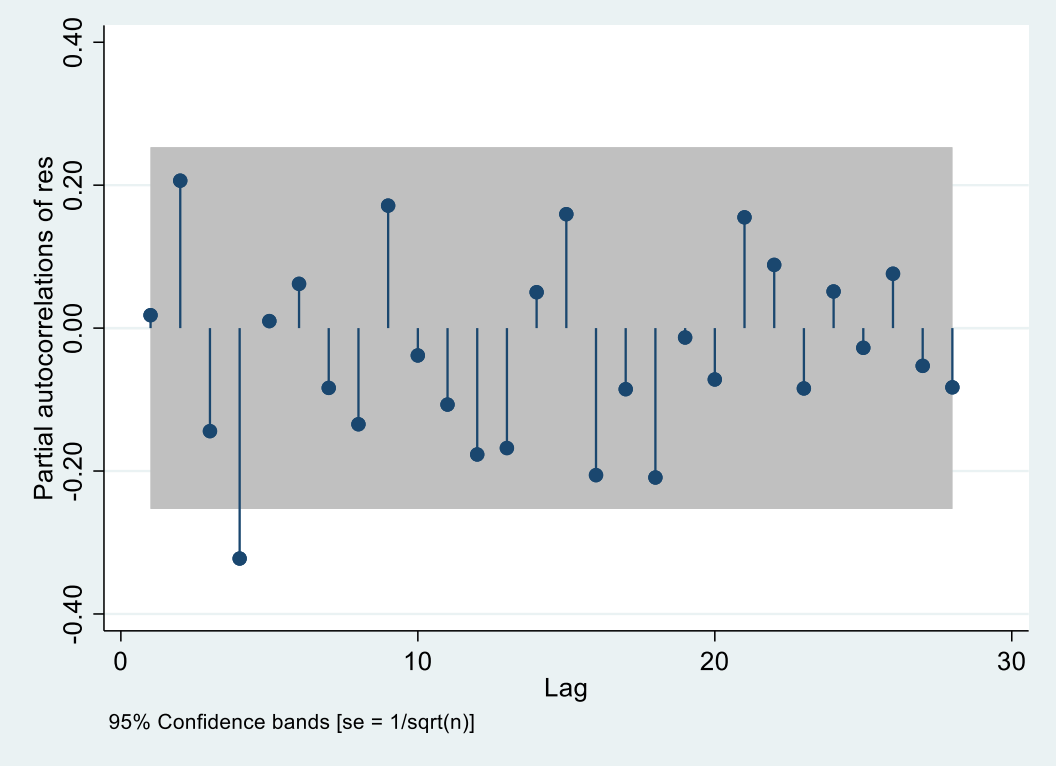


Figure 14-35 Plot of partial residuals for time series (national overall cohort)

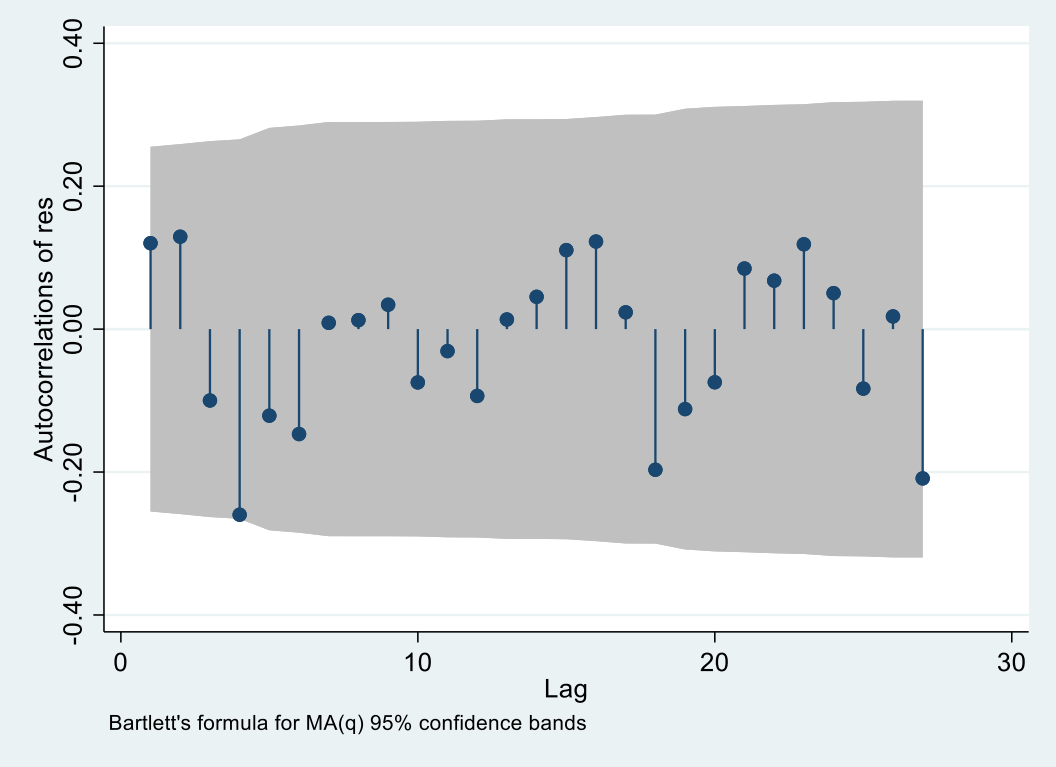


Figure 14-36 Plot of residuals for time series (national SCOT eligible cohort)

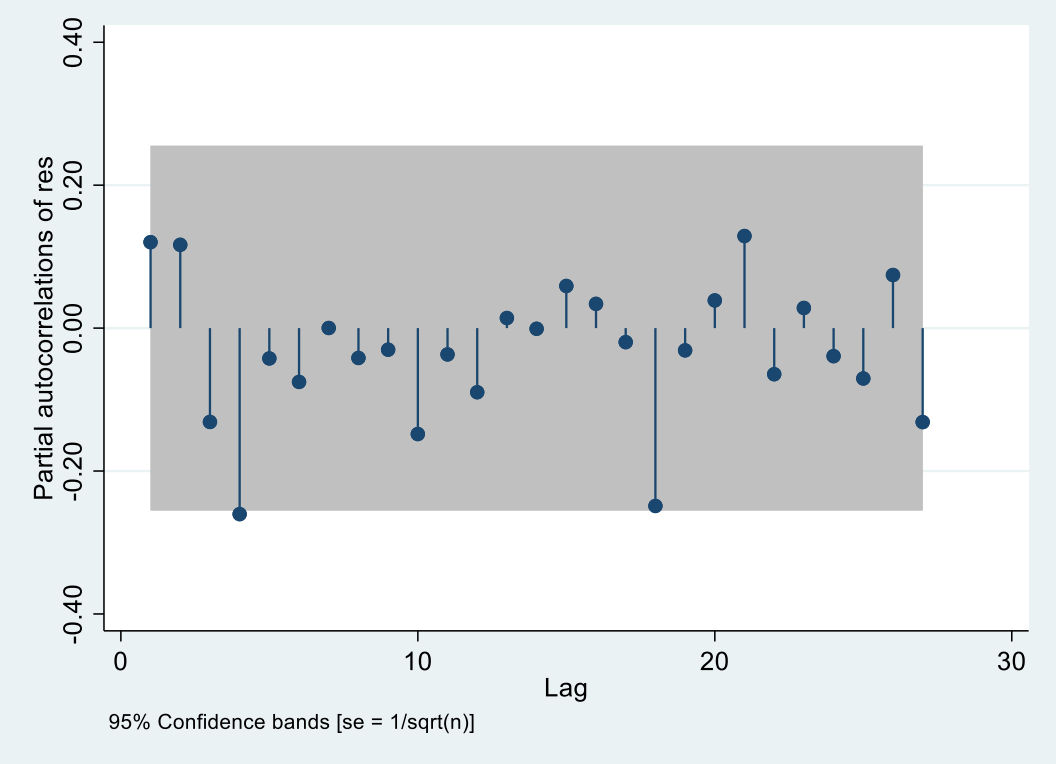


Figure 14-37 Plot of partial residuals for time series (national SCOT eligible cohort)

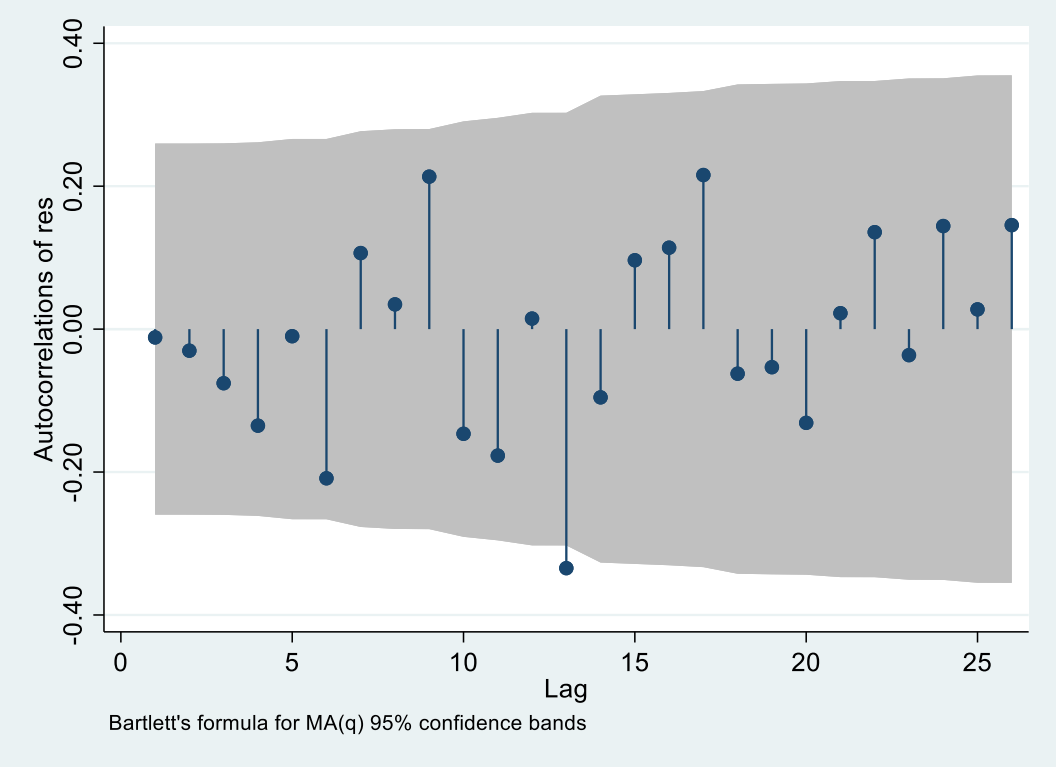


Figure 14-38 Plot of residuals for time series (national SCOT ineligible cohort)

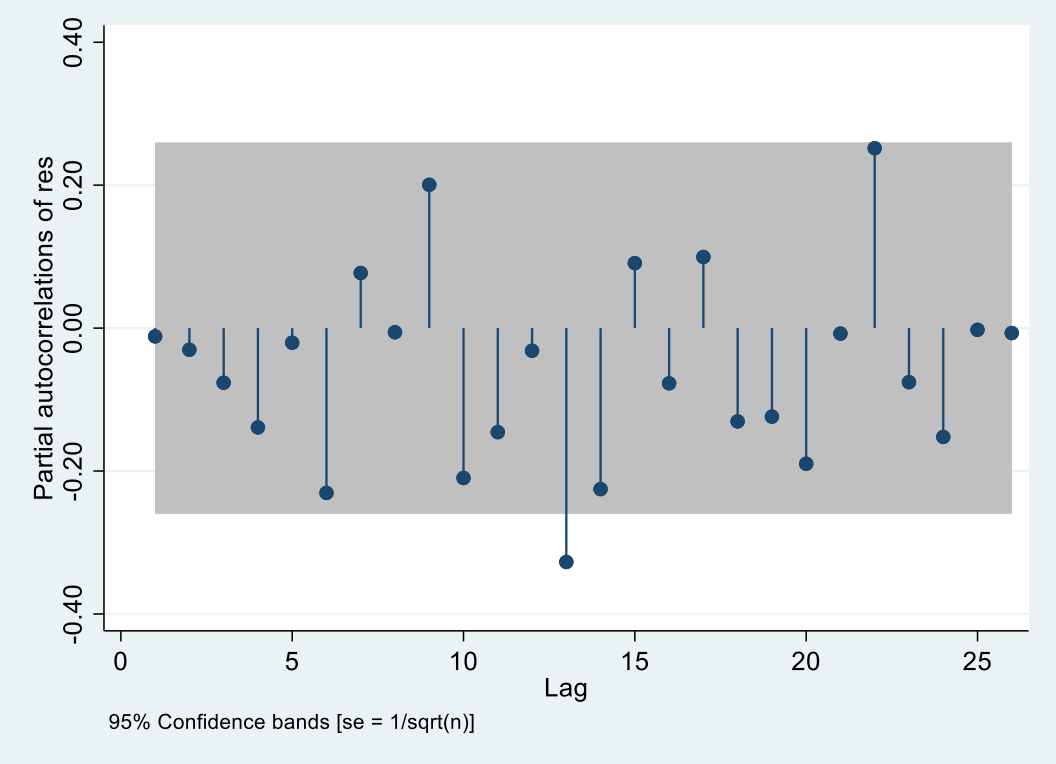


Figure 14-39 Plot of partial residuals for time series (national SCOT ineligible cohort)

Table 14-8 Durbin-Watson test statistics for national level ITSAs

Analysis	Durbin-Watson test statistic
Logistic regression overall	1.84
Logistic regression SCOT eligible	1.50
Logistic regression SCOT ineligible	1.96

15 Appendix 6 Comparison of impact frameworks

As discussed in Chapters 1 and 3, research impact frameworks can be used to structure data collection for impact evaluation, as well as communicating the results of an evaluation. Below, six different frameworks that were identified from a review of the literature (Chapter 3) are used to describe the impact of the SCOT trial.

15.1 Payback framework

A description of the payback framework was first published in in the 1990s (15). The framework consists of a modified logic model (see Figure 3-3) alongside a list of five impact categories (Figure 15-1). This was initially developed to evaluate health services research.

Category	Definition
1. Knowledge	Journal articles; conference presentations; books; book chapters; research reports
2. Benefits to future research and research use	• Better targeting of future research
	• Development of research skills, personnel and overall research capacity
	• A critical capacity to absorb and utilise appropriately existing research including that from overseas
	• Staff development and educational benefits
3. Benefits from informing policy and product development	• Improved information bases for political and executive decisions
	• Other political benefits from undertaking research
	• Development of pharmaceutical products and therapeutic techniques
4. Health and health sector benefits	• Improved health
	• Cost reduction in delivery of existing services
	• Qualitative improvements in the process of delivery
	• Improved equity in service delivery
5. Broader economic benefits	• Wider economic benefits from commercial exploitation of innovations arising from R&D
	• Economic benefits from a healthy workforce and reduction in working days lost

Figure 15-1 Payback framework categories of impact. Taken from (15)

Stage 0: Topic/Issue Identification

In the early 2000s it was known that adjuvant fluoropyrimidine-oxaliplatin doublet chemotherapy provided a survival benefit for patients with stage II and stage III colorectal cancer. The standard treatment was six months duration which had been shown to be as effective as longer treatment. The clinical

problem with this treatment was the high rate of toxicity, in particular peripheral neuropathy that accompanied this duration of therapy

Interface A: Project specification and selection

Professor Jim Cassidy (oncologist) and Mr Jim Paul (CTU statistician) at the University of Glasgow identified this was an issue and developed a clinical trial idea to explore this concept further. In the meantime, other clinicians and trial groups in a number of countries worldwide were also developing clinical trials to address this concern.

Stage 1: Inputs to research

The SCOT trial was developed by the University of Glasgow CTU. A successful funding application was made to the Medical Research Council and the trial was awarded £2.4 million in investment in 2006. A decision was made to open the SCOT trial in multiple centres and both in the UK and internationally. Other funding was invested from CRUK, NIHR, the Swedish Cancer Society and Interreg for the purposes of this trial.

Stage 3: Research process

The SCOT trial opened in 2008 and recruited patients until 2013. In total, 6,144 patients registered for the trial, 6,088 were randomised and 6,065 provided consent for their data to be used.

Stage 3: Primary outputs from research

The SCOT trial findings met the pre-specified end-point showing that 3 months of treatment was non-inferior to 6 months of treatment for the overall trial population. Pre-planned subgroup analysis showed that there was an unexpected differences in outcome depending on the regimen selected by the treating physician. A post-hoc analysis showed differences in outcome for stage III patients dependent on the size of the tumour and nodal involvement.

Interface B: Dissemination

The dissemination of the SCOT trial findings provide examples of the first payback in the framework, “knowledge”. The findings were initially disseminated in June 2017 at the ASCO conference by Professor Tim Iveson, the trial principal investigator. The results were also discussed at the ESMO conference in September 2017 and published in full in the Lancet Oncology in April 2018. Results to patients specifically with high-risk stage II disease were presented separately at ESMO 2018. The SCOT trial was the largest contributor to the International Duration of Adjuvant treatment (IDEA) collaboration and these collaboration results were disseminated initially at ASCO 2017 and ESMO 2017. The full publication of primary results was published in the New England Journal of Medicine in March 2018. Results for high-risk stage II patients were presented at ASCO 2019, ESMO 2019 and published in full in the Journal of Clinical Oncology in December 2020. Finally, overall survival results for IDEA were disseminated at ASCO 2020 and published in full in the Lancet Oncology in November 2020.

Stage 4: Secondary outputs: policy making/product development

This stage overlaps with category three, benefits from informing policy and product development. The main impact on policy from the SCOT trial has been on clinical guidelines from medical professional bodies. The survey undertaken for this thesis (Chapter 5) outlined which guidelines clinicians use in practice. Out of this list of guidelines, the NICE guidelines, Japanese and NCCN guidelines have been updated since the dissemination of the SCOT trial and IDEA collaboration findings and have specifically cited one or both of these trial findings as the basis for their recommendations.

Stage 5: Adoption by practitioners and public

The results of the survey (Chapter 5) and real world prescribing data analysis (Chapters 7 and 8) have demonstrated that the SCOT trial findings have been adopted into practice by clinicians.

Stage 6: Final outcomes

It is predicted that the practice change outlined above will lead to health benefits for patients who will experience less toxicity from a shorter duration of treatment. There are also cost benefits from the perspective of the healthcare services providing this treatment.

Stock or reservoir of knowledge

The impact of this trial must be evaluated in the context of other knowledge. For example, the results of the SCOT trial will be adopted and used within the context of the finding of the IDEA collaboration and each of the other five trials that contributed. The SCOT trial findings will build on previous knowledge from clinical trials conducted prior to SCOT (Chapter 1).

Categories of impact

Using the list of categories of payback, those not picked up by using the logic model include:

Better targeting of future research: The results of the SCOT trial will impact on future clinical trials investigating chemotherapy in the neoadjuvant or adjuvant setting. For example, the UK based POLEM trial has used 3 months of CAPOX as its standard adjuvant arm (336). Also, there is a major parallel work stream of translational work (TransSCOT) currently in progress using translational samples collected from patients in the SCOT trial (260).

Staff development and educational benefits: This PhD is using patient level data from the SCOT trial and has provided the opportunity for further research in this context. Individual patient level data is being used for the purposes of contributing to a separate higher degree (OCTOPUS project, University of Manchester (349)). Specifically related to health economics, anonymised patient level data is being used as a teaching resource on an educational course at the University of Glasgow.

Economic benefits from a healthy workforce and reduction in working days lost: As described in Chapter 6, if the SCOT trial findings are implemented and shorter adjuvant treatment is used, it is highly likely that a proportion of

patients will return to work sooner, providing significant monetary benefits at a societal level.

15.2 Research utilisation ladder

This framework focuses on the utilisation of research results by decision makers and others in society and aligns with “user” focused definitions of research impact (see Chapter 1). The utilisation of research is considered a process consisting of several stages. In their article describing the research utilisation ladder, Landry et al (204) created a modified version of Knott and Wildavsky’s scale (1980) of research utilisation (image below) and used this scale to assess the impact of social science research. In their study, the scale is used to survey over 2,000 social science researchers regarding how they self-assess their research is used in practice. In the survey, the authors also collected details regarding the research in question and the action of the researchers to promote research use. A Likert scale was used to ask researchers of how well they rate the utilisation of their research and these results were converted into a binary, quantifiable result. Specifically, each rung on the ladder is assigned a zero score if the survey participant answers negatively on the Likert scale or a one score if the participant answers in a positive manner. The details collected regarding research and its use are used as co-variates in a regression model to predict which researchers/what type of research is likely to climb highest on the knowledge utilisation ladder and which stages/rungs of the ladder may present barriers to research impact. Below, this framework is applied to the SCOT trial in both a narrative and quantitative fashion.

Stages of the Ladder of Knowledge Utilization

Stage 1	Transmission: I transmitted my research results to the practitioners and professionals concerned.
Stage 2	Cognition: My research reports were read and understood by the practitioners and professionals concerned.
Stage 3	Reference: My work has been cited as a reference in the reports, studies, and strategies of action elaborated by practitioners and professionals.
Stage 4	Effort: Efforts were made to adopt the results of my research by practitioners and professionals.
Stage 5	Influence: My research results influenced the choice and decision of practitioners and professionals.
Stage 6	Application: My research results gave rise to applications and extension by the practitioners and professionals concerned.

Figure 15-2 Landry's ladder of knowledge utilisation Taken from (204)

Stage 1: Evidenced via conference presentations and publications of the SCOT trial results. (This would score 1).

Stage 2: The survey reported in Chapter 5 shows that clinicians are aware of the SCOT trial findings. (This would score 1).

Stage 3: The SCOT trial findings have been cited in the NICE guidelines and other professional guidelines internationally. (This would score 1).

Stage 4: The survey results showed that practitioners were making efforts to use the SCOT trial results in their own practice. (This would score 1).

Stage 5: Again, the survey in this thesis (Chapter 3) shows that the SCOT trial and IDEA collaboration did influence the practice change described by survey participants. (This would score 1).

Stage 6: Both the survey (Chapter 3) and administrative database analysis (Chapter 6) demonstrate that the SCOT trial findings have changed practice. (This would score 1).

15.3 Research Impact Framework

The Research Impact Framework (RIF) was developed by researchers (2006) at the London School of Hygiene and Tropical Medicine (LSHTM). These authors recognised that there was an increasing requirement for academics to describe the impact of the research they carry out, but that this usually performed on an ad hoc basis, making it difficult to learn from and to make comparisons between these evaluations. Their aim was to provide a framework specifically for health researchers who have had no training in impact evaluation, to enable them to evaluate the impact of their work.

The authors developed their framework firstly by drawing on previous literature to identify four main areas of research impact that would form the basis of their framework. This information was used to develop an interview guide, which was subsequently used to question researchers at LSHTM regarding the impact of their work. Through these interviews the authors developed, alongside the primary researchers themselves, impact narratives around seven health services and policy research projects and around four more basic research projects from the performed by researchers at the departments of Epidemiological and Public Health and Infectious and Tropical Disease. They also conducted four case studies using secondary data such as reports and previous impact evaluations. Thematic analysis of interview transcripts was performed and the results were used to build the RIF. Specifically, 27 categories of impact were identified from the interview transcripts and other sources, grouped under four themes. In some instances, these categories were divided further into sub-categories, which represented more specific examples of impact or potential impact that may arise from health research. The four main themes were: research impact, policy impact, service impact and societal impact. This framework went through further iteration and validation through discussion with LSHTM researchers and by testing the framework using LSHTM researchers and research projects. A summary of the framework is provided below. The impact areas, categories and sub-categories were reviewed to identify those most relevant to the SCOT trial.

Narrative areas	RESEARCH PROJECT/PROGRAMME DESCRIPTION	Key dates (mm/yy)
Description of research project/programme	Topics/research area: Geopolitical contexts: Funders and budget: Research management, influencing events and challenges:	
1. Research-related impact	1.1 Type of problem/knowledge 1.2 Research methods used 1.3 Publications and papers 1.4 Products, patents and translatability potential 1.5 Research networks 1.6 Leadership and awards 1.7 Research management 1.8 Communication	
2. Policy impact	2.1 Level of policy-making 2.2 Type of policy 2.3 Nature of policy impact 2.4. Policy networks 2.5 Political capital	
3. Service impact	3.1 Type of services: health/intersectoral 3.2 Evidence-based practice 3.3 Quality of care 3.4 Information systems 3.5 Services management 3.6 Cost-containment and cost-effectiveness	
4. Societal impact	4.1 Knowledge, attitudes and behaviour 4.2 Health literacy 4.3 Health status 4.4 Equity and human rights 4.5 Macroeconomic/related to the economy 4.6 Social capital and empowerment 4.7 Culture and art 4.8 Sustainable development outcomes	

Figure 15-3 Five main impact areas included in the Research Impact Framework Taken from (99)

Description of research project/programme

The topic area is research into adjuvant treatment for patients with colorectal cancer. The research under evaluation is a phase III multi-centre, international trial conducted in six countries. The budget to perform this research is outlined in **Error! Reference source not found.** Appendix 4.

Research-related impact

The problem that the SCOT trial addressed was the high level of toxicity, in particular neuropathy, from the previous standard duration of adjuvant chemotherapy given to patients with high risk stage II and stage III colorectal cancer. Due to the high incidence of CRC globally, thousands of patients were being treated with six months of treatment and experiencing toxicity each year. The research method used to investigate the merits of a shorter treatment duration was a large phase III clinical trial and the findings from this trial were published at international conferences and in peer-reviewed journals (details in

Chapter 1). The SCOT trial contributed to a global research network (IDEA) and provided the opportunity for collaboration and pooling of results. Professor Jim Cassidy and laterally, Professor Tim Iveson were the chief investigators of this trial and both worked closely alongside the lead statistician Mr Jim Paul. The trial was developed and managed by the CRUK Glasgow Clinical Trials Unit.

Policy impact

The level of policy impact of the SCOT trial was local, national, and international and the type of policy was mainly clinical guidelines. Specifically, the SCOT trial has been cited in the 2020 NICE guidelines. The IDEA collaboration, to which it was the largest contributor, has been cited in Japanese and NCCN guidelines.

Service impact

The SCOT trial has impacted on the health service. As shown in Chapter 6, there has been a change in practice corresponding with the SCOT trial findings, leading to a reduction in the duration of adjuvant chemotherapy delivered. Chapter 6 has demonstrated that shorter duration of treatment is a cost-effective strategy and that it will be cost-saving.

Societal impact

Although difficult to measure, it is expected that using shorter treatment in the adjuvant setting for CRC will benefit the health of patients by reducing the toxicity related to longer treatment, without significantly compromising survival outcomes. Macroeconomic impacts are likely to be seen at a societal level if patients given shorter treatment return to work, leisure and caring duties sooner.

Specific categories/sub-categories relevant to the SCOT trial:

Area	Category	Sub-category
Research related impacts	Type of problem/knowledge addressed	Evidence of effectiveness of interventions

	Publications and papers Research networks and user involvement Research leadership Communication	Addressing research gaps and testing new hypotheses Publications in scientific journals Technical reports Citation of research publications by other researchers Formal academic talks/presentations Guidelines
Policy Impact	Level of policy making Type of policy Nature of policy influence	Local/National/International Guidelines Instrumental
Service Impacts	Evidence based practice Cost-containment and cost-effectiveness	
Societal impact	Health status Macroeconomic/related to the economy	 Healthy workforce outcomes Value of health gain

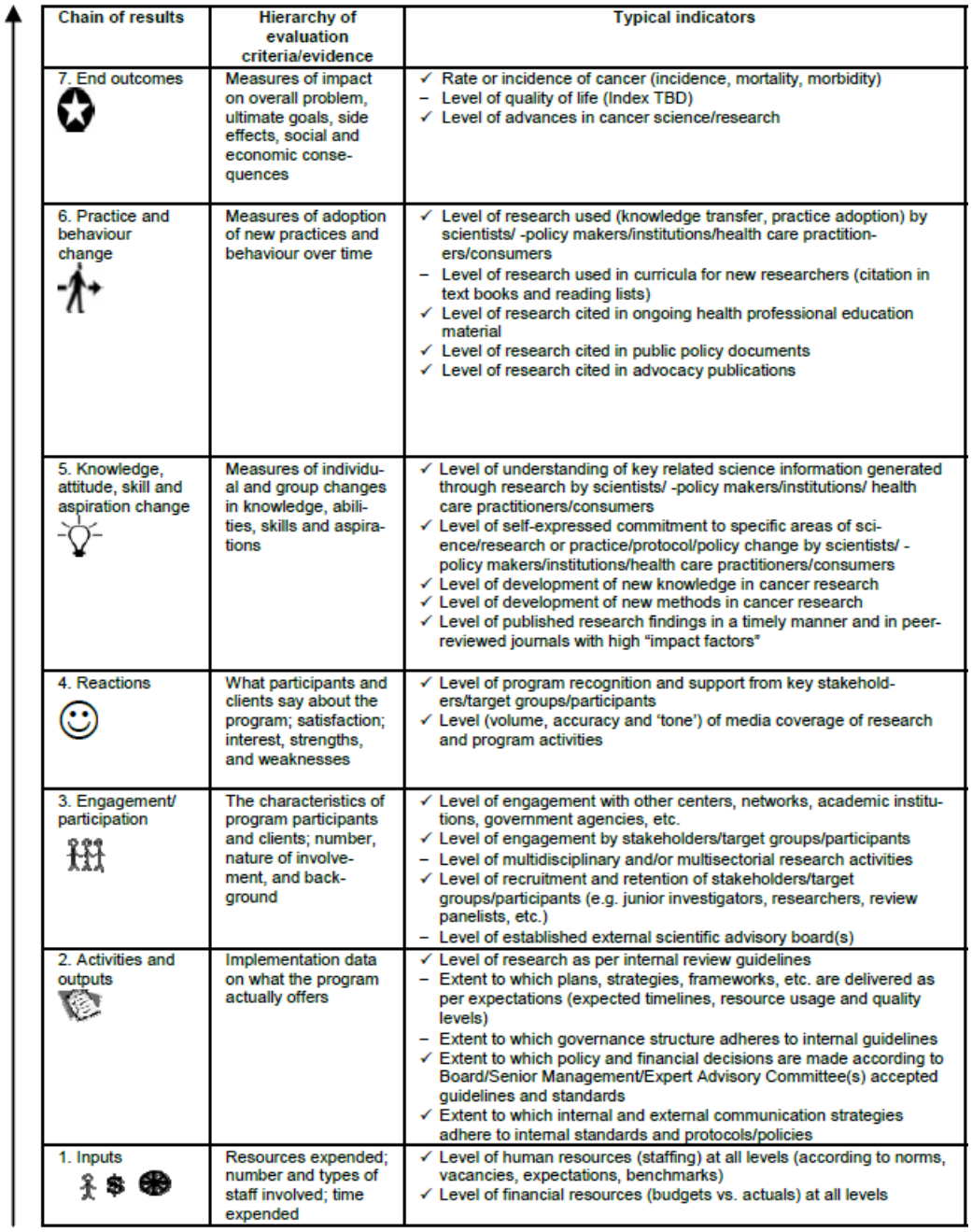
15.4 Montague and Valentim (modified CAHS)

Montaague and Valentim have described the development of a novel impact framework based on a combination of the Canadian Academy of Health Sciences (CAHS) framework and the Bennett and Rockwell theory of action. The CAHS framework was developed using the Payback framework and the five main impact categories of the CAHS closely align with those from the Payback model (Advancing knowledge, building capacity, informing decision-making, health impacts, and broad socio-economic impacts). The CAHS framework consists of a logic model and the five impact categories are populated with 60 impact indicators. A full report (2007) of who and how the CAHS framework was developed, alongside a full list of the indicators included is publicly available (21).

Montague and Valentim suggest that there are challenges to using the CAHS framework in practice and that it is not straightforward to link impacts and tell a story of how impact has occurred. Recognising this challenge, they call on the Bennett and Rockwell theory of action, which describes how impact can occur

from a chain of events. The authors apply the 60 indicators of impact from the CAHS framework to the seven stages in the Bennett hierarchy to produce a new framework to evaluate the impact of health research. Montague and Valentim apply their framework to evaluate the impact of a cancer trial. This work was commissioned by the Canadian Cancer Society, Canada's largest health charity, which wanted to better evaluate the impact of their investments.

The image below shows the chain of results (left hand column) along with typical indicators of impact corresponding to each stage in the theory of change (right hand column). A summary of the main indicators of SCOT trial impact according to these stages is provided in the table below. In addition, a diagram of these indicators of impact occurring over time is provided in line with the diagram used in Montague and Valentim's evaluation of a Canadian cancer trial (MA-17 trial).



Chain of results	Hierarchy of evaluation criteria/evidence	Typical indicators
7. End outcomes	Measures of impact on overall problem, ultimate goals, side effects, social and economic consequences	✓ Rate or incidence of cancer (incidence, mortality, morbidity) – Level of quality of life (Index TBD) ✓ Level of advances in cancer science/research
6. Practice and behaviour change	Measures of adoption of new practices and behaviour over time	✓ Level of research used (knowledge transfer, practice adoption) by scientists/ -policy makers/institutions/health care practitioners/consumers – Level of research used in curricula for new researchers (citation in text books and reading lists) ✓ Level of research cited in ongoing health professional education material ✓ Level of research cited in public policy documents ✓ Level of research cited in advocacy publications
5. Knowledge, attitude, skill and aspiration change	Measures of individual and group changes in knowledge, abilities, skills and aspirations	✓ Level of understanding of key related science information generated through research by scientists/ -policy makers/institutions/ health care practitioners/consumers ✓ Level of self-expressed commitment to specific areas of science/research or practice/protocol/policy change by scientists/ -policy makers/institutions/health care practitioners/consumers ✓ Level of development of new knowledge in cancer research ✓ Level of development of new methods in cancer research ✓ Level of published research findings in a timely manner and in peer-reviewed journals with high "impact factors"
4. Reactions	What participants and clients say about the program; satisfaction; interest, strengths, and weaknesses	✓ Level of program recognition and support from key stakeholders/target groups/participants ✓ Level (volume, accuracy and 'tone') of media coverage of research and program activities
3. Engagement/ participation	The characteristics of program participants and clients; number, nature of involvement, and background	✓ Level of engagement with other centers, networks, academic institutions, government agencies, etc. ✓ Level of engagement by stakeholders/target groups/participants – Level of multidisciplinary and/or multisectorial research activities ✓ Level of recruitment and retention of stakeholders/target groups/participants (e.g. junior investigators, researchers, review panelists, etc.) – Level of established external scientific advisory board(s)
2. Activities and outputs	Implementation data on what the program actually offers	✓ Level of research as per internal review guidelines – Extent to which plans, strategies, frameworks, etc. are delivered as per expectations (expected timelines, resource usage and quality levels) – Extent to which governance structure adheres to internal guidelines ✓ Extent to which policy and financial decisions are made according to Board/Senior Management/Expert Advisory Committee(s) accepted guidelines and standards ✓ Extent to which internal and external communication strategies adhere to internal standards and protocols/policies
1. Inputs	Resources expended; number and types of staff involved; time expended	✓ Level of human resources (staffing) at all levels (according to norms, vacancies, expectations, benchmarks) ✓ Level of financial resources (budgets vs. actuals) at all levels

Figure 15-4 Framework for research impact assessment. Taken from (193)

Table 15-1 Montague and Valentim's framework applied to the SCOT trial

Chain of results	Hierarchy of evaluation criteria/evidence	Indicator of impact
7. End outcomes	Measures of impact on overall problem, ultimate goals, side effects, social and economic consequences	Chapter 6 shows that the trial findings have confirmed shorter treatment to be cost-effective in most situations. The budget impact analysis predicts the likely cost savings at the health service budget level if implementation occurs. Wider, society level economic impacts include the reduction in productivity loss due to an earlier return to work for patients who are treated with the shorter treatment. Finally, health outcomes

		are likely to be affected at a population level, specifically a reduction in toxicity and longer-term side effects related to longer treatment.
6. Practice and behaviour change	Measures of adoption of new practices and behaviour over time	Review of real world prescriptions provide a robust indication of uptake of trial results (Chapter 7 and 8 results).
5. Knowledge, attitude, skill and aspiration change	Measures of individual and group changes in knowledge, abilities, skills and aspirations	Again, the survey in Chapter 5 has captured this to some extent.
4. Reactions	What participants and clients say about the program, satisfaction, interest, strengths and weaknesses	Clinician, patient, and policy maker reactions to the trial results. This can be demonstrated through formal publication of conference discussions and the survey such as the one described in Chapter 5 of this thesis.
3. Engagement and participation	The characteristics of program participants and clients, number, nature of involvement and background	Dissemination of trial findings in the academic and lay literature.
2. Activities and outputs	Implementation data on what the program actually offers	Running of the trial and the trial results.
1. Inputs	Resources expended, number and types of staff involved, time expended	The financial and time investment by funders and researchers to perform the SCOT trial.

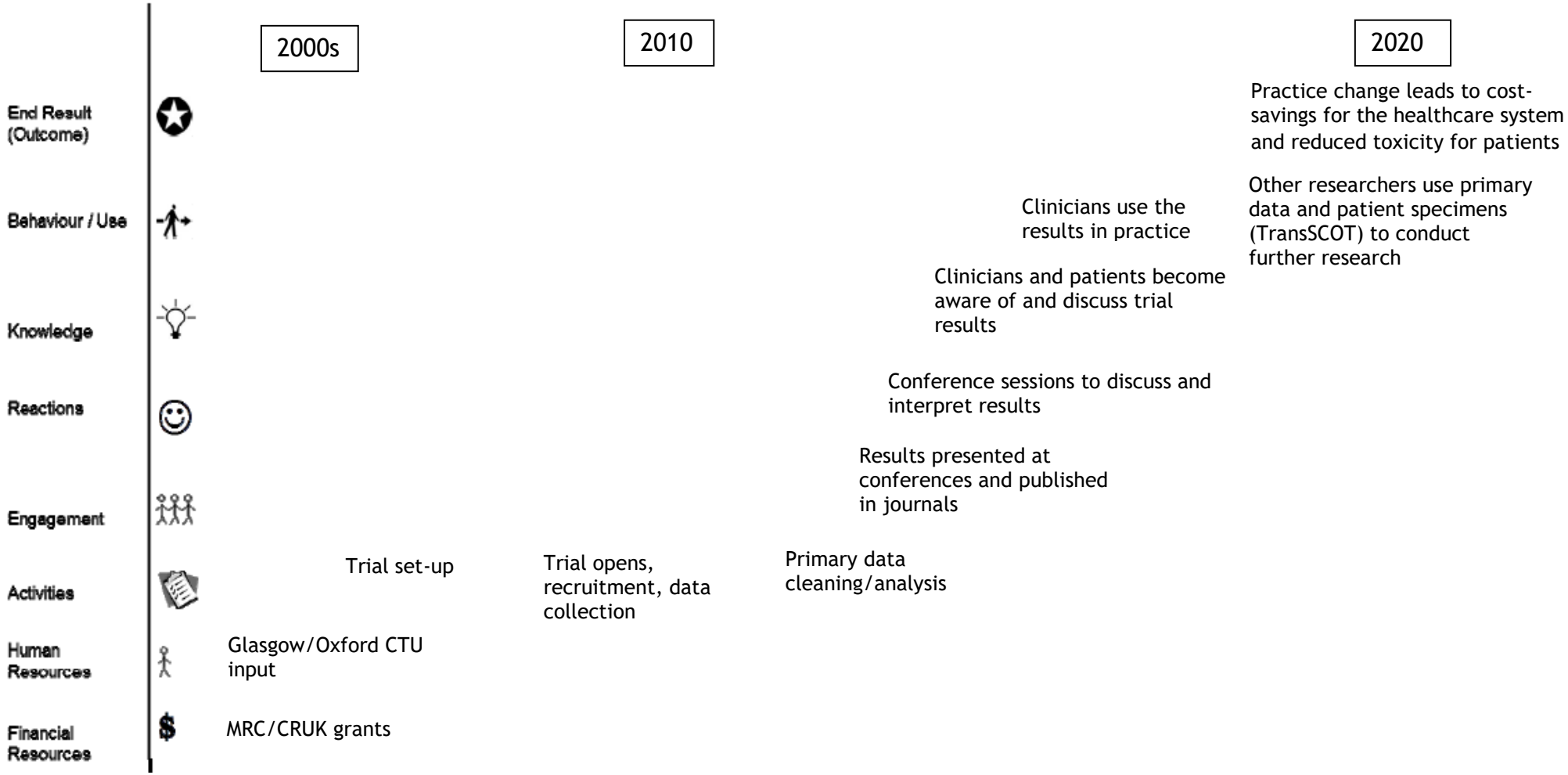


Figure 15-5 Montague and Valentim's framework applied to the SCOT trial. Adapted from (193)

15.5 Weiss's logic model (modified United way framework)

In his 2005 essay, Anthony Weiss raises a call to arms to medical researchers and funders to change the way the impact of medical research is evaluated. He argues for a shift away from an outputs based approach that focuses on publication counts and journal impact factors and a move to an outcomes based approach that evaluates how research brings us one step closer to the real goal for any medical research – to improve disease. He adapts a logic model described in a report by not for profit organisation United Way, which was developed in 1996 to better evaluate the impact of programs of work performed by human service, research and charitable institutions. Weiss adapts this logic model and applies it to medical research (image below) and describes ways in which impact at each level of the logic model could be evaluated and the challenges to impact occurring and its evaluation at each stage. In particular, Weiss highlights that a barrier to research impact often occurs between the publication of research results and the application of those results into practice, the "efficacy-effectiveness gap". Although a simple model, this is still considered a type of impact framework. How this framework could be applied to investigate the impact of the SCOT trial is described below.

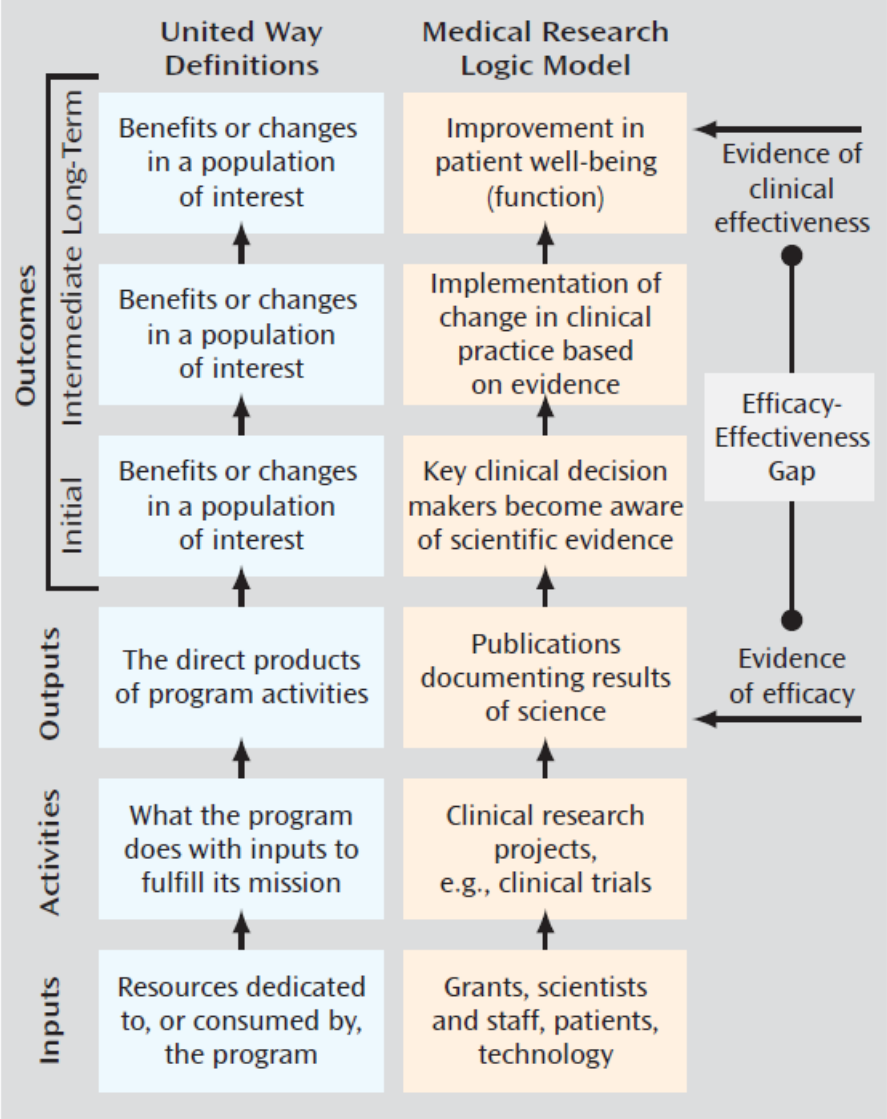


Figure 15-6 Anthony Weiss's logic model Taken from (202)

Inputs Project specific financial investment from funders, CTU investment, and input of time and effort from clinicians and patients enrolled in SCOT.

Clinical research projects The SCOT trial is performed in six countries over 5 years.

Publications documenting results of science The results of the SCOT trial are first disseminated at ASCO 2017 and the full publication is in April 2018. Updated results were communicated at ASCO 2020 and the SCOT trial has contributed to the IDEA collaboration and the corresponding publications.

Key clinical decision makers become aware of scientific evidence The survey results reported in Chapter 5 demonstrate that a large number of clinicians were aware of the SCOT trial results. It is impossible to know how representative this sample is of clinicians generally but the large number and resounding result regarding SCOT trial and IDEA collaboration awareness indicate that it is likely most clinicians working in this field will have some idea of the SCOT trial findings.

Implementation of change in clinical practice based on evidence In his essay, Anthony Weiss suggests that one of the most robust ways to understand if research results are used in practice is to analyse real world prescribing records. The results in Chapter 8 demonstrate that it is highly likely the SCOT trial results influenced the duration of adjuvant chemotherapy given to patients in Scotland.

Improvement in patient well-being (function) It can be hypothesised that with a proven reduction in treatment duration that the incidence of toxicity and side effects from treatment seen in the SCOT trial are likely to also be demonstrated in the real world setting. Also, the SCOT trial showed there was no clinically significant difference in survival outcomes between using six versus three months of treatment. Longer follow up data would be required to confirm if this was also the case in a real world setting. Also, better collection of toxicity and patient reported outcomes would be necessary to demonstrate if a corresponding reduction in harm from treatment occurred.

15.6 Healthcare research impact matrix

The final framework considered was that from Cruz-Riveria and colleagues (2018) from the Centre for Patient Reported Outcomes at the University of Birmingham. This matrix was developed by reviewing the existing literature to find examples of methodological frameworks and summarising the synthesising these frameworks and the impact metrics included within each framework. The structure of the matrix was based on frameworks including Payback (15), the RIF (203)and Anthony Weiss's logic model (202). The final matrix is displayed below. The authors group their impact categories into temporal groups, short, mid and long-term and the arrow implies a linear pathway from research to impact occurring.

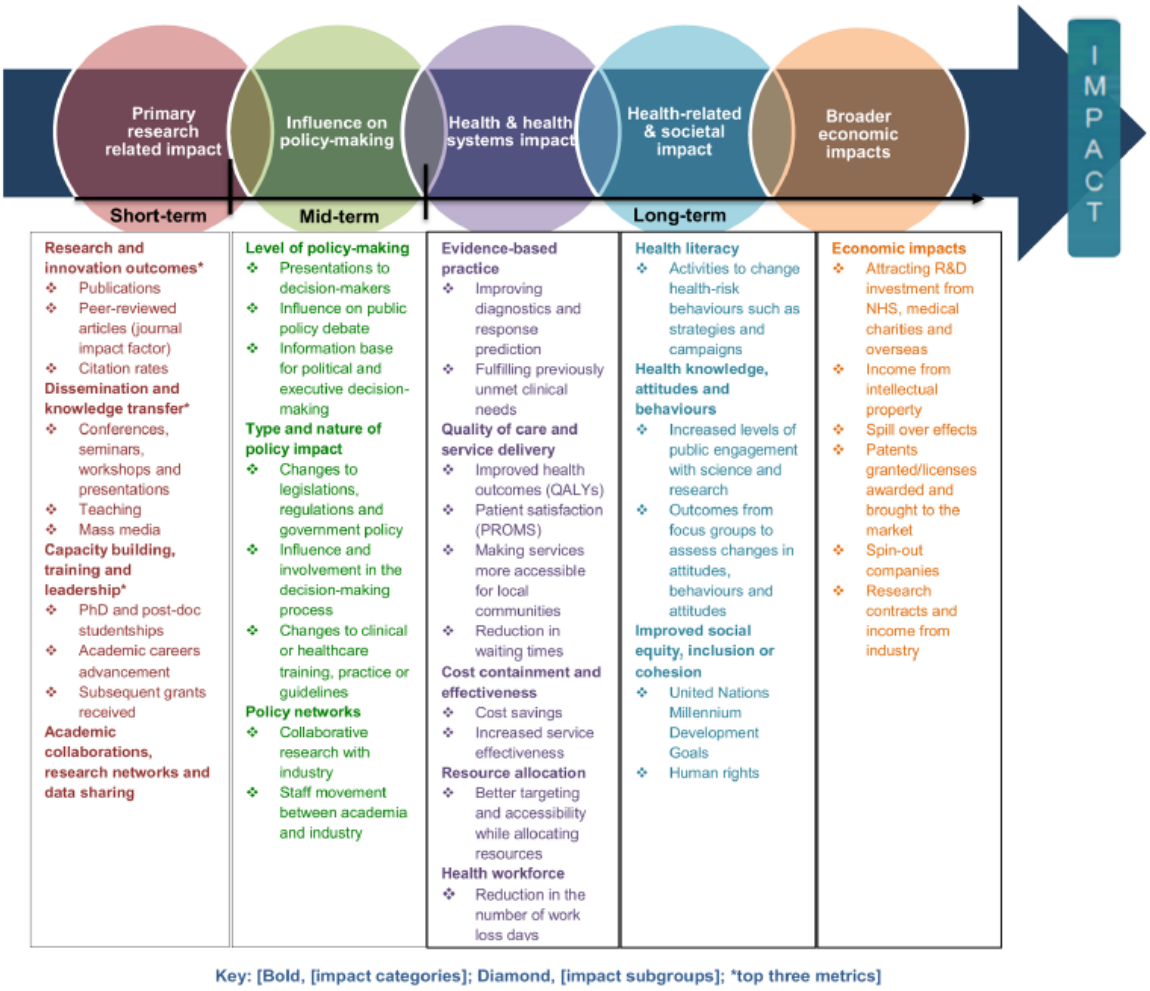


Figure 15-7 Cruz-Rivera's impact matrix Taken from (170)

16 Appendix 7 Funding and information governance approvals

Beatson cancer charity
DETERMINED TO BEAT CANCER

Dr Catherine Hanna
CRUK Research Fellow
Clinical Trials Unit
The Beatson West of Scotland Cancer Centre
1053 Great Western Road
Glasgow
G12 0YN

19th March 2018

Dear Catherine,

Funding Application Number:
Funding Application Title: Research project – To assess the "real-life" impact of a trial on chemotherapy prescribing in Scotland

I am delighted to confirm that we have approved your application for funding the above project, subject to your agreement to the following terms and conditions:

1. The Beatson Cancer Charity Funding Evaluation Form, attached, must be completed and submitted to us by 30th June 2019 and annually thereafter.

2. Beatson Cancer Charity should be acknowledged and our logo displayed on equipment where possible and all related materials and we should receive copy of such materials in advance for approval.

3. Agreement to ongoing marketing and PR activity including digital and media activity to support fundraising and to demonstrate the success and the impact of the studies.

4. We may publicise this award in internal and external communications, including digital and media activity and we may request that you provide appropriate information in relation to the funding award including case study materials in support of this activity.

5. All payments should be claimed no later than 30th June 2021.

6. Written acceptance of the above terms and conditions should be submitted to me by 23rd March 2018 prior to any funds being released or reimbursed in connection with this application.

Beatson Cancer Charity
Beatson West of Scotland
Cancer Centre
1053 Great Western Road
Glasgow, G12 0YN

Office: 0141 212 0505
Email: info@beatsoncancercharity.org
Online: www.beatsoncancercharity.org
Twitter: @beatson_charity
Facebook.com/beatsoncancercharity

Catherine Hanna
catherine.hanna@glasgow.ac.uk

Data Protection Officer
Information Governance Department
NHS Greater Glasgow & Clyde
2nd Floor, 1 Smithhills Street
Paisley
PA1 1EB

Date: 18/07/2019

Enquiries to: Isobel Brown
Tel: 0141 355 2020
Email: Isobel.Brown@ggc.scot.nhs.uk

Dear Catherine

- **Re: Assessing the "real life" impact of the SCOT trial results in Scotland: Have prescribing practices changed after publication of this trial and if so at what rate have they changed?**

Thank you for your Caldicott application received on 17/07/2019 regarding your proposed Research Project.

I have reviewed this application and can confirm that I am happy to approve this application on behalf of the Caldicott Guardian.

Please note that this approval only covers access to NHSGGC patients.

Please find attached a signed copy of your application for your records.

Yours sincerely

Isobel Brown
Data Protection Officer
Information Governance

Public Benefit and Privacy Panel for Health and Social Care
nss.PBPP@nhs.net
www.informationgovernance.scot.nhs.uk

Dr Catherine Hanna,
Beatson West of Scotland Cancer Centre,
Great Western Road,
Glasgow,
G12 OYN

Date: 22nd May 2018
Your Ref:
Our Ref:

Dear Dr Hanna

Re Application: Assessing the "impact" of clinical trial results: Proposal to assess chemotherapy prescribing practices in Scotland prior to and following the publication of the Short Course Oncology Trial (SCOT).
Version: 20180514 PBPP CHv2

Thank you for your application for consideration by the Public Benefit and Privacy Panel for Health and Social Care. Your application has undergone proportionate governance review and has been approved.

This approval is given to process data as specified in the approved application form, and is limited to this. Approval is valid for the period specified in your application until 31st January 2021. You are required to notify the Panel Manager of any proposed change to any aspect of your proposal, including purpose or method of processing, data or data variables being processed, study cohorts, individuals accessing and processing data, timescales, technology/infrastructure, or any other relevant change.

On conclusion of your proposal, as part of NHS Scotland Governance and monitoring we will require you to complete an End of Project reporting form to demonstrate that you have complied with the obligations outlined such as data destruction or submission of references for publications of findings.

I would take this opportunity to remind you of the declaration you have made in your application form committing you to undertakings in respect of information governance, confidentiality and data protection. . It is the responsibility of the applicant and their organisation to ensure that their study complies with current and new legislation.

Requests for access to NHS Scotland data as part of this approved application should be supported by evidencing a copy of your approval letter and application form to the relevant local board contacts/data providers.

Please note that summary information about your application and its approval, including the title and nature of your proposal, will be published on the panel website (www.informationgovernance.scot.nhs.uk).

I hope that your proposal progresses well,

Yours sincerely

Dr Marian Aldhous

Panel Manager
NHS Scotland Public Benefit and Privacy Panel for Health and Social Care
Email: nss.PBPP@nhs.net

List of References

1. UK CR. CRUK: "Cancer Statistics for the UK." [cited 2019 20th May 2019]. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics-for-the-uk.
2. Cancer Research UK. Cancer risk statistics 2019 [Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/risk.
3. Institute NC. [National Cancer Institute "About NCI" webpage]. Available from: https://www.cancer.gov/about-nci/budget/fact-book/data/research-funding.
4. UK CR. CRUK: "Current clinical trial research" [cited 2019 20th May 2019]. Available from: https://www.cancerresearchuk.org/our-research-by-cancer-topic/our-clinical-trial-research/current-clinical-trial-research.
5. Mattina J, Carlisle B, Hachem Y, Fergusson D, Kimmelman J. Inefficiencies and Patient Burdens in the Development of the Targeted Cancer Drug Sorafenib: A Systematic Review. PLoS Biol. 2017;15(2):e2000487.
6. Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gulmezoglu AM, et al. How to increase value and reduce waste when research priorities are set. Lancet. 2014;383(9912):156-65.
7. Ioannidis JPA, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. Lancet. 2014;383(9912):166-75.
8. Richard S. BMJ blogs [Internet]. Godlee F, editor2018. [31st July 2019]. Available from: https://blogs.bmj.com/bmj/2018/07/30/richard-smith-measuring-research-impact-rage-hard-get-right/.
9. Higher Education Funding Council for England. REF2014 Impact Case Studies 2014 [Available from: https://impact.ref.ac.uk/casestudies/FAQ.aspx.
10. Oxford English Dictionary 2020 [Available from: https://www.oed.com/view/Entry/163432?rskey=PSlUSC&result=1&isAdvanced=false#eid.
11. Expertisecentrum Onderzoek en Ontwikkelingsmonitoring. The history of bibliometrics 2021 [Available from: https://www.ecoom.be/nodes/degeschiedenisvanbibliometrie/en.
12. Stevens H, Dean A, Wykes M. DESCRIBE project: Definitions, Evidence, and Structures to Capture Research Impacts and Benefits. 2013.
13. Adam P, Ovseiko PV, Grant J, Graham KEA, Boukhris OF, Dowd AM, et al. ISRIA statement: ten-point guidelines for an effective process of research impact assessment. Health research policy and systems. 2018;16(1):8.
14. Williams K, Grant J. A Comparative Review of How the Policy and Procedures to Assess Research Impact Evolved in Australia and the UK. Res Evaluat. 2018;27(2):93-105;; 13.
15. Donovan C, Hanney S. The 'Payback Framework' explained. Res Evaluat. 2011;20(3):181-3.
16. McKenna HP. Research assessment: the impact of impact. International journal of nursing studies. 2015;52(1):1-3.
17. Chowdhury G, Koya K, Philipson P. Measuring the Impact of Research: Lessons from the UK's Research Excellence Framework 2014. PloS one. 2016;11(6):e0156978.
18. Terama E, Smallman M, Lock SJ, Johnson C, Austwick MZ. Beyond Academia - Interrogating Research Impact in the Research Excellence Framework. PloS one. 2016;11(12):e0168533.

19. Greenhalgh T, Fahy N. Research impact in the community-based health sciences: an analysis of 162 case studies from the 2014 UK Research Excellence Framework. BMC medicine. 2015;13:232.
20. Government A. Engagement and Impact Assessment 2018/9 2019 [Available from: https://dataportal.arc.gov.au/EI/NationalReport/2018/pages/introduction/index.html?id=background.
21. Canadian Academy of Health Sciences. Canadian Academy of Health Sciences ASSESSMENT REPORT. MAKING AN IMPACT A Preferred Framework and Indicators to Measure Returns
on Investment in Health Research. 2009.
22. Alliance CHSaPR. Making an Impact: A shared framework for assessing the impact of health services and policy research on decision-making. 2018.
23. AstraZeneca. Annual Reports 2021 [Available from: https://www.astrazeneca.com/investor-relations/annual-reports.html.
24. NPC. Impact measurement, evaluation and data 2021 [Available from: https://www.thinknpc.org/themes/measure-and-manage-impact/impact-measurement-evaluation-and-data/.
25. Organisations TNCfV. Impact 2021 [Available from: https://www.ncvo.org.uk/practical-support/information/impact.
26. Gomes DS, ; C. The; impact; generated; by; publicly; and; charity-funded; research; in; the; United; Kingdom:; a; systematic; literature; review. Health; research; policy; and; systems. 2019;17(1):22-019-0425-2.
27. Cope D, Carter IM, Morgan Jones M, Grant J, Husbands Fealing K, Horton A, et al. 7 Essays on Impact. University of Exeter; 2013. Contract No.: Report.
28. Penfield T, Baker M, Scoble R, Wykes M. Assessment, Evaluations and Definitions of Research Impact: A Review. Res Evaluat. 2014;23(1):21--32.
29. Alla K, Hall WD, Whiteford HA, Head BW, Meurk CS. How do we define the policy impact of public health research? A systematic review. Health research policy and systems. 2017;15(1):84.
30. UK Research and Innovation. Pathways to impact 2019 [Available from: https://webarchive.nationalarchives.gov.uk/20200203113953/https://www.ukri.org/innovation/excellence-with-impact/pathways-to-impact/.
31. REF2014. Search REF Impact Case Studies 2014 [Available from: https://impact.ref.ac.uk/casestudies/.
32. Lavis J, Ross S, McLeod C, Gildiner A. Measuring the impact of health research. Journal of Health Services & Research Policy. 2003;8(3):165-70.
33. Eccles MP, Mittman BS. Welcome to Implementation Science. Implementation Science. 2006;1(1):1.
34. Bauer MS, Damschroder L, Hagedorn H, Smith J, Kilbourne AM. An introduction to implementation science for the non-specialist. BMC Psychology. 2015;3(1):32.
35. Hanney SR, Ovseiko PV, Graham KER, Chorzempa H, Miciak M. A systems approach for optimizing implementation to impact: meeting report and proceedings of the 2019 In the Trenches: Implementation to Impact International Summit BMC Proceedings. 2020;14(6):10.
36. Paoletti X, Postel-Vinay S. Phase I–II trial designs: how early should efficacy guide the dose recommendation process? Annals of Oncology. 2018;29(3):540-1.
37. Head SJ, Kaul S, Bogers AJ, Kappetein AP. Non-inferiority study design: lessons to be learned from cardiovascular trials. Eur Heart J. 2012;33(11):1318-24.

38. National Cancer Institute. Types of Trials 2020 [
39. Glover M, Buxton M, Guthrie S, Hanney S, Pollitt A, Grant J. Estimating the returns to UK publicly funded cancer-related research in terms of the net value of improved health outcomes. BMC medicine. 2014;12:99.
40. UK CR. Cancer research – benefiting patients AND the economy 2014 [Available from: https://scienceblog.cancerresearchuk.org/2014/06/16/cancer-research-benefiting-patients-and-the-economy/.
41. Iveson TJ, Kerr RS, Saunders MP, Cassidy J, Hollander NH, Tabernero J, et al. 3 versus 6 months of adjuvant oxaliplatin-fluoropyrimidine combination therapy for colorectal cancer (SCOT): an international, randomised, phase 3, non-inferiority trial. The Lancet Oncology. 2018;19(4):562-78.
42. Obrand DI, Gordon PH. Incidence and patterns of recurrence following curative resection for colorectal carcinoma. Diseases of the colon and rectum. 1997;40(1):15-24.
43. Andre T, Boni C, Mounedji-Boudiaf L, Navarro M, Tabernero J, Hickish T, et al. Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. The New England journal of medicine. 2004;350(23):2343-51.
44. Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. CA: a cancer journal for clinicians. 2017;67(2):93-9.
45. Wolmark N, Fisher B, Rockette H, Redmond C, Wickerham DL, Fisher ER, et al. Postoperative adjuvant chemotherapy or BCG for colon cancer: results from NSABP protocol C-01. Journal of the National Cancer Institute. 1988;80(1):30-6.
46. Laurie JA, Moertel CG, Fleming TR, Wieand HS, Leigh JE, Rubin J, et al. Surgical adjuvant therapy of large-bowel carcinoma: an evaluation of levamisole and the combination of levamisole and fluorouracil. The North Central Cancer Treatment Group and the Mayo Clinic. Journal of Clinical Oncology. 1989;7(10):1447-56.
47. Moertel CG, Fleming TR, Macdonald JS, Haller DG, Laurie JA, Goodman PJ, et al. Levamisole and fluorouracil for adjuvant therapy of resected colon carcinoma. The New England journal of medicine. 1990;322(6):352-8.
48. Efficacy of adjuvant fluorouracil and folinic acid in colon cancer. International Multicentre Pooled Analysis of Colon Cancer Trials (IMPACT) investigators. Lancet. 1995;345(8955):939-44.
49. O'Connell MJ, Laurie JA, Kahn M, Fitzgibbons RJ, Jr., Erlichman C, Shepherd L, et al. Prospectively randomized trial of postoperative adjuvant chemotherapy in patients with high-risk colon cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1998;16(1):295-300.
50. Wolmark N, Rockette H, Mamounas E, Jones J, Wieand S, Wickerham DL, et al. Clinical trial to assess the relative efficacy of fluorouracil and leucovorin, fluorouracil and levamisole, and fluorouracil, leucovorin, and levamisole in patients with Dukes' B and C carcinoma of the colon: results from National Surgical Adjuvant Breast and Bowel Project C-04. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1999;17(11):3553-9.
51. Chau I, Norman AR, Cunningham D, Tait D, Ross PJ, Iveson T, et al. A randomised comparison between 6 months of bolus fluorouracil/leucovorin and 12 weeks of protracted venous infusion fluorouracil as adjuvant treatment in

colorectal cancer. Annals of oncology : official journal of the European Society for Medical Oncology. 2005;16(4):549-57.
52. Haller DG, Catalano PJ, Macdonald JS, O'Rourke MA, Frontiera MS, Jackson DV, et al. Phase III Study of Fluorouracil, Leucovorin, and Levamisole in High-Risk Stage II and III Colon Cancer: Final Report of Intergroup 0089. Journal of Clinical Oncology. 2005;23(34):8671-8.
53. Twelves C, Wong A, Nowacki MP, Abt M, Burris H, Carrato A, et al. Capecitabine as Adjuvant Treatment for Stage III Colon Cancer. New England Journal of Medicine. 2005;352(26):2696-704.
54. Kuebler JP, Wieand HS, O'Connell MJ, Smith RE, Colangelo LH, Yothers G, et al. Oxaliplatin Combined With Weekly Bolus Fluorouracil and Leucovorin As Surgical Adjuvant Chemotherapy for Stage II and III Colon Cancer: Results From NSABP C-07. Journal of Clinical Oncology. 2007;25(16):2198-204.
55. Quasar Collaborative G, Gray R, Barnwell J, McConkey C, Hills RK, Williams NS, et al. Adjuvant chemotherapy versus observation in patients with colorectal cancer: a randomised study. Lancet. 2007;370(9604):2020-9.
56. Yothers G, O'Connell MJ, Allegra CJ, Kuebler JP, Colangelo LH, Petrelli NJ, et al. Oxaliplatin As Adjuvant Therapy for Colon Cancer: Updated Results of NSABP C-07 Trial, Including Survival and Subset Analyses. Journal of Clinical Oncology. 2011;29(28):3768-74.
57. Andre T, Quinaux E, Louvet C, Colin P, Gamelin E, Bouche O, et al. Phase III study comparing a semimonthly with a monthly regimen of fluorouracil and leucovorin as adjuvant treatment for stage II and III colon cancer patients: final results of GERCOR C96.1. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007;25(24):3732-8.
58. Schmoll HJ, Tabernero J, Maroun J, de Braud F, Price T, Van Cutsem E, et al. Capecitabine Plus Oxaliplatin Compared With Fluorouracil/Folinic Acid As Adjuvant Therapy for Stage III Colon Cancer: Final Results of the NO16968 Randomized Controlled Phase III Trial. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2015;33(32):3733-40.
59. Chau I, Cunningham D. Adjuvant therapy in colon cancer: current status and future directions. Cancer treatment reviews. 2002;28(5):223-36.
60. Köhne C-H, Bedenne L, Carrato A, Bouché O, Popov I, Gaspà L, et al. A randomised phase III intergroup trial comparing high-dose infusional 5-fluorouracil with or without folinic acid with standard bolus 5-fluorouracil/folinic acid in the adjuvant treatment of stage III colon cancer: The Pan-European Trial in Adjuvant Colon Cancer 2 study. European journal of cancer. 2013;49(8):1868-75.
61. de Gramont A, Larsen AK, Tournigand C, Louvet C, Andre T, Gercor. Update on targeted agents for adjuvant treatment of colon cancer in 2006. Gastrointestinal cancer research : GCR. 2007;1(4 Suppl 2):S47-9.
62. Haller DG, Tabernero J, Maroun J, de Braud F, Price T, Van Cutsem E, et al. Capecitabine plus oxaliplatin compared with fluorouracil and folinic acid as adjuvant therapy for stage III colon cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2011;29(11):1465-71.
63. André T, de Gramont A, Vernerey D, Chibaudel B, Bonnetain F, Tijeras-Raballand A, et al. Adjuvant Fluorouracil, Leucovorin, and Oxaliplatin in Stage II to III Colon Cancer: Updated 10-Year Survival and Outcomes According to BRAF Mutation and Mismatch Repair Status of the MOSAIC Study. Journal of Clinical Oncology. 2015;33(35):4176-87.

64. Petersen SH, Harling H, Kirkeby LT, Wille-Jørgensen P, Mocellin S. Postoperative adjuvant chemotherapy in rectal cancer operated for cure. The Cochrane database of systematic reviews. 2012;2012(3):CD004078-CD.
65. Meyers BM, Cosby R, Quereshy F, Jonker D. Adjuvant Chemotherapy for Stage II and III Colon Cancer Following Complete Resection: A Cancer Care Ontario Systematic Review. Clinical oncology. 2017;29(7):459-65.
66. Tournigand C, Andre T, Bachet j, Teixeira L, Boni C, Clingan PR, et al. FOLFOX4 as adjuvant therapy in elderly patients (pts) with colon cancer (CC): Subgroup analysis of the MOSAIC trial. Journal of Clinical Oncology. 2010;28(15_suppl):3522-.
67. Haller DG, O'Connell MJ, Cartwright TH, Twelves CJ, McKenna EF, Sun W, et al. Impact of age and medical comorbidity on adjuvant treatment outcomes for stage III colon cancer: a pooled analysis of individual patient data from four randomized, controlled trials. Annals of oncology : official journal of the European Society for Medical Oncology. 2015;26(4):715-24.
68. Salem ME, Yin J, Goldberg RM, Pederson LD, Wolmark N, Alberts SR, et al. Evaluation of the change of outcomes over a 10-year period in patients with stage III colon cancer: pooled analysis of 6501 patients treated with fluorouracil, leucovorin, and oxaliplatin in the ACCENT database. Annals of oncology : official journal of the European Society for Medical Oncology. 2020;31(4):480-6.
69. McCleary NJ, Meyerhardt JA, Green E, Yothers G, de Gramont A, Van Cutsem E, et al. Impact of age on the efficacy of newer adjuvant therapies in patients with stage II/III colon cancer: findings from the ACCENT database. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2013;31(20):2600-6.
70. Glimelius B, Osterman E. Adjuvant Chemotherapy in Elderly Colorectal Cancer Patients. Cancers (Basel). 2020;12(8):2289.
71. Grothey A. Oxaliplatin-safety profile: neurotoxicity. Seminars in oncology. 2003;30(4 Suppl 15):5-13.
72. Land SR, Kopec JA, Cecchini RS, Ganz PA, Wieand HS, Colangelo LH, et al. Neurotoxicity from oxaliplatin combined with weekly bolus fluorouracil and leucovorin as surgical adjuvant chemotherapy for stage II and III colon cancer: NSABP C-07. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007;25(16):2205-11.
73. Pachman DR, Qin R, Seisler DK, Smith EM, Beutler AS, Ta LE, et al. Clinical Course of Oxaliplatin-Induced Neuropathy: Results From the Randomized Phase III Trial N08CB (Alliance). Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2015;33(30):3416-22.
74. Andre T, Boni C, Navarro M, Tabernero J, Hickish T, Topham C, et al. Improved overall survival with oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment in stage II or III colon cancer in the MOSAIC trial. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2009;27(19):3109-16.
75. Robles-Zurita J, Boyd KA, Briggs AH, Iveson T, Kerr RS, Saunders MP, et al. SCOT: a comparison of cost-effectiveness from a large randomised phase III trial of two durations of adjuvant Oxaliplatin combination chemotherapy for colorectal cancer. British journal of cancer. 2018;119(11):1332-8.
76. Iveson T, Boyd KA, Kerr RS, Robles-Zurita J, Saunders MP, Briggs AH, et al. 3-month versus 6-month adjuvant chemotherapy for patients with high-risk stage II and III colorectal cancer: 3-year follow-up of the SCOT non-inferiority RCT. 2019;23:64.

77. Pectasides D, Karavasilis V, Papaxoinis G, Gourgioti G, Makatsoris T, Raptou G, et al. Randomized phase III clinical trial comparing the combination of capecitabine and oxaliplatin (CAPOX) with the combination of 5-fluorouracil, leucovorin and oxaliplatin (modified FOLFOX6) as adjuvant therapy in patients with operated high-risk stage II or stage III colorectal cancer. BMC Cancer. 2015;15:384.
78. Sobrero A, Grothey A, Iveson T, Labianca R, Yoshino T, Taieb J, et al. The hard road to data interpretation: 3 or 6 months of adjuvant chemotherapy for patients with stage III colon cancer? Annals of oncology : official journal of the European Society for Medical Oncology. 2018;29(5):1099-107.
79. Roseweir AK, Park JH, Hoorn St, Powell AGMT, Aherne S, Roxburgh CSD, et al. Histological phenotypic subtypes predict recurrence risk and response to adjuvant chemotherapy in patients with stage III colorectal cancer. The Journal of Pathology: Clinical Research. 2020;6(4):283-96.
80. Alexander PG, Roseweir AK, Pennel KAF, van Wyk HC, Powell AGMT, McMillan DC, et al. The Glasgow Microenvironment Score associates with prognosis and adjuvant chemotherapy response in colorectal cancer. British journal of cancer. 2021;124(4):786-96.
81. Grothey A, Sobrero AF, Shields AF, Yoshino T, Paul J, Taieb J, et al. Duration of Adjuvant Chemotherapy for Stage III Colon Cancer. The New England journal of medicine. 2018;378(13):1177-88.
82. National Comprehensive Cancer Network Guidelines (Colon Cancer) Version 4.2018. NCCN; 2018 Accessed 10th March 2019.
83. Sobrero AF, Andre T, Meyerhardt JA, Grothey A, Iveson T, Yoshino T, et al. Overall survival (OS) and long-term disease-free survival (DFS) of three versus six months of adjuvant (adj) oxaliplatin and fluoropyrimidine-based therapy for patients (pts) with stage III colon cancer (CC): Final results from the IDEA (International Duration Evaluation of Adj chemotherapy) collaboration. Journal of Clinical Oncology. 2020;38(15_suppl):4004-.
84. André T, Meyerhardt J, Iveson T, Sobrero A, Yoshino T, Souglakos I, et al. Effect of duration of adjuvant chemotherapy for patients with stage III colon cancer (IDEA collaboration): final results from a prospective, pooled analysis of six randomised, phase 3 trials. The Lancet Oncology. 2020;21(12):1620-9.
85. American Society of Clinical Oncology. Final IDEA Collaboration Results Support Just 3 Months of Adjuvant CAPOX for Most Patients With Stage III Colon Cancer 2020 [Available from: https://dailynews.ascopubs.org/do/10.1200/ADN.20.200177/full/.
86. Lieu CH, editor Adjuvant therapy: When did it get so complex? American Society of Clinical Oncology; 2020; Virtual.
87. Iveson T, Sobrero AF, Yoshino T, Sougklakos I, Ou F-S, Meyers JP, et al. Prospective pooled analysis of four randomized trials investigating duration of adjuvant (adj) oxaliplatin-based therapy (3 vs 6 months {m}) for patients (pts) with high-risk stage II colorectal cancer (CC). Journal of Clinical Oncology. 2019;37(15_suppl):3501-.
88. Iveson TJ, Sobrero AF, Yoshino T, Souglakos I, Ou F-S, Meyers JP, et al. Duration of Adjuvant Doublet Chemotherapy (3 or 6 months) in Patients With High-Risk Stage II Colorectal Cancer. Journal of Clinical Oncology. 2021;39(6):631-41.
89. Iveson T, Kerr R, Saunders MP, Hollander NH, Tabernero J, Haydon AM, et al. Final DFS results of the SCOT study: An international phase III randomised (1:1) non-inferiority trial comparing 3 versus 6 months of oxaliplatin based

adjuvant chemotherapy for colorectal cancer. Journal of Clinical Oncology. 2017;35(15_suppl):3502-.
90. Hunt H, Pollock A, Campbell P, Estcourt L, Brunton G. An introduction to overviews of reviews: planning a relevant research question and objective for an overview. Syst Rev. 2018;7(1):39.
91. Hunt H, Pollock A, Campbell P, Estcourt L, Brunton G. An introduction to overviews of reviews: planning a relevant research question and objective for an overview. Systematic Reviews. 2018;7.
92. Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. International journal of evidence-based healthcare. 2015;13(3):132-40.
93. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. Open medicine : a peer-reviewed, independent, open-access journal. 2009;3(3):e123-30.
94. Greenhalgh T, Thorne S, Malterud K. Time to challenge the spurious hierarchy of systematic over narrative reviews? European journal of clinical investigation. 2018;48(6).
95. Bryman A. Social Research Methods 4th Edition. Fourth ed. Oxford: Oxford University Press; 2012 2012. 809 p.
96. Krippendorff K. Content Analysis. An Introduction to its Methodology. 2nd Edition ed: SAGE Publications; 2004.
97. Berelson B. Content analysis in communication research. New York, NY, US: Free Press; 1952. 220- p.
98. Guthrie S, Krapels J, Lichten CA, Wooding S. 100 Metrics to Assess and Communicate the Value of Biomedical Research: An Ideas Book. . 2016.
99. Kuruvilla S, Mays N, Pleasant A, Walt G. Describing the impact of health research: a Research Impact Framework. BMC health services research. 2006;6:134.
100. Buxton M, Hanney S, Jones T. Estimating the economic value to societies of the impact of health research: a critical review. Bulletin of the World Health Organization. 2004;82(10):733-9.
101. O'Connor C, Joffe H. Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. International Journal of Qualitative Methods. 2020;19:1609406919899220.
102. Oppenheim AN. Questionnaire Design, Interviewing and Attitude Measurement. 2nd ed. New York, USA: Basic Books Inc; 1999.
103. Graham B. Developing a Questionnaire: Cointinuum; 2000.
104. Jensen EL, C. Doing Real Research: Sage Publications Limited; 2016.
105. Boynton PM, Greenhalgh T. Selecting, designing, and developing your questionnaire. Bmj. 2004;328(7451):1312.
106. Evans J. Desinging a Questionnaire: A short course for health professionals. In: University of Stirling, editor. 2018. p. 10.
107. Thwaites Bee D, Murdoch-Eaton D. Questionnaire design: the good, the bad and the pitfalls. Archives of disease in childhood - Education & practice edition. 2016;101(4):210.
108. Taber KS. The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. Research in Science Education. 2018;48(6):1273-96.
109. Boynton PM. Administering, analysing, and reporting your questionnaire. Bmj. 2004;328(7452):1372.

110. B.R. K, J.A.C. S. Essential Medical Statistics. 2nd ed: Blackwell Science Ltd; 2006.
111. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. PharmacoEconomics. 2013;31(5):361-7.
112. Glick HA, Doshi, J.A., Sonnad, S.S., Polsky, D. Economic Evaluation in Clinical Trials. Second ed. Briggs A, Gray, A., editor. Oxford: Oxford University Press; 2015.
113. World Health Organisation. CHOosing Interventions that are Cost Effective (WHO-CHOICE) 2019 [Available from: https://www.who.int/choice/country/country_specific/en/.
114. Stenberg K, Lauer JA, Gkountouras G, Fitzpatrick C, Stanciole A. Econometric estimation of WHO-CHOICE country-specific costs for inpatient and outpatient health service delivery. Cost effectiveness and resource allocation : C/E. 2018;16:11.
115. Scotland PH. Acute Medical Speciality Costs 2020 [Available from: https://www.isdscotland.org/Health-Topics/Finance/Costs/Detailed-Tables/Speciality-Costs/Acute-Medical.asp.
116. Organisation for Economic Co-operation and Development. Purchasing power parities (PPP) 2019 [Available from: https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm.
117. Foroutan N, Tarride J-E, Xie F, Levine M. A methodological review of national and transnational pharmaceutical budget impact analysis guidelines for new drug submissions. Clinicoecon Outcomes Res. 2018;10:821-54.
118. Sullivan SD, Mauskopf JA, Augustovski F, Jaime Caro J, Lee KM, Minchin M, et al. Budget impact analysis-principles of good practice: report of the ISPOR 2012 Budget Impact Analysis Good Practice II Task Force. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2014;17(1):5-14.
119. Oxford Uo. UK Colorectal Cancer Intelligence Hub CORECT-R 2020 [Available from: https://www.ndph.ox.ac.uk/corectr/corect-r.
120. NHS Health Research Authority. Establishing a UK Colorectal Cancer Intelligence Hub 2018 [Available from: https://www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/establishing-a-uk-colorectal-cancer-intelligence-hub/.
121. PUBLIC BENEFIT AND PRIVACY PANEL FOR HEALTH AND SOCIAL CARE-HSC-PBPP 2021 [Available from: https://www.informationgovernance.scot.nhs.uk/pbpphsc/.
122. Healthcare Improvement Scotland. Cancer Quality Performance Indicators (QPIs) 2020 [Available from: http://www.healthcareimprovementscotland.org/our_work/cancer_care_improvement/cancer_qpis/quality_performance_indicators.aspx.
123. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. Journal of Chronic Diseases. 1987;40(5):373-83.
124. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Medical care. 2005;43(11):1130-9.
125. International cQ. NVivo 2021 [Available from: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/support-services/nvivo-downloads.

126. 2021 cM. Microsoft Excel 2016 [Available from: https://www.microsoft.com/en-gb/microsoft-365/excel.
127. 2021 cM. Microsoft Word 2016 [Available from: https://www.microsoft.com/en-gb/microsoft-365/word.
128. LLC S. New in STATA 14 release 2021 [Available from: https://www.stata.com/stata14/.
129. Developers G. Google Charts 2020 [Available from: https://developers.google.com/chart.
130. Ltd J. Online surveys (formerly BOS) 2021 [Available from: https://www.onlinesurveys.ac.uk/.
131. Gunn A, Mintrom M. Social Science Space Network Blogger [Internet]2017. [cited 2019]. Available from: https://www.socialsciencespace.com/2017/01/five-considerations-policy-measure-research-impact/.
132. Callaway E. Beat it, impact factor! Publishing elite turns against controversial metric. Nature. 2016;535(7611):210-1.
133. Cancer Research UK. Measuring the impact of research 2017 [Available from: https://www.cancerresearchuk.org/funding-for-researchers/research-features/2017-06-20-measuring-the-impact-of-research.
134. Gordon LG, Bartley N. Views from senior Australian cancer researchers on evaluating the impact of their research: results from a brief survey. Health research policy and systems. 2016;14:2.
135. Research Councils UK. Excellence with impact 2018 [Available from: https://webarchive.nationalarchives.gov.uk/20180322123208/http://www.rcuk.ac.uk/innovation/impact/.
136. Thompson MK, Poortmans P, Chalmers AJ, Faivre-Finn C, Hall E, Huddart RA, et al. Practice-changing radiation therapy trials for the treatment of cancer: where are we 150 years after the birth of Marie Curie? British journal of cancer. 2018;119(4):389-407.
137. Srinivasaiah N, Joseph B, Mackey P, Monson JR. How do we manage early rectal cancer? A national questionnaire survey among members of the ACPGBI after the preliminary results of the MRC CR07/NCIC CO16 randomized trial. Colorectal disease : the official journal of the Association of Coloproctology of Great Britain and Ireland. 2008;10(4):357-62.
138. Downing A, Morris EJ, Aravani A, Finan PJ, Lawton S, Thomas JD, et al. The Effect of the UK Coordinating Centre for Cancer Research Anal Cancer Trial (ACT1) on Population-based Treatment and Survival for Squamous Cell Cancer of the Anus. Clinical oncology. 2015;27(12):708-12.
139. Tsang Y, Ciurlionis L, Kirby AM, Locke I, Venables K, Yarnold JR, et al. Clinical impact of IMPORT HIGH trial (CRUK/06/003) on breast radiotherapy practices in the United Kingdom. The British journal of radiology. 2015;88(1056):20150453.
140. South A, Parulekar WR, Sydes MR, Chen BE, Parmar MK, Clarke N, et al. Estimating the Impact of Randomised Control Trial Results on Clinical Practice: Results from a Survey and Modelling Study of Androgen Deprivation Therapy plus Radiotherapy for Locally Advanced Prostate Cancer. European urology focus. 2016;2(3):276-83.
141. Horsley T DO, Sampson M. . Checking reference lists to find additional studies for systematic reviews. . Cochrane Database of Systematic Reviews 2011(8).

142. Green BN, Johnson CD, Adams A. Writing narrative literature reviews for peer-reviewed journals: secrets of the trade. Journal of chiropractic medicine. 2006;5(3):101-17.
143. Raftery JH, S. Greenhalgh, T. Glover, M. Blatch-Jones, A. Models and applications for measuring the impact of health research: update of a systematic review for the Health Technology Assessment programme. Health technology assessment. 2016;20(76):1-254.
144. Hanney SB, M.; Green, C.; Coulson, D.; Raftery, J. An assessment of the impact of the NHS Health Technology Assessment Programme. Health technology assessment. 2007;11(53):82.
145. Greenhalgh T, Raftery J, Hanney S, Glover M. Research impact: a narrative review. BMC medicine. 2016;14:78.
146. Hanney SR, Gonzalez-Block MA, Buxton MJ, Kogan M. The utilisation of health research in policy-making: concepts, examples and methods of assessment. Health research policy and systems. 2003;1(1):2.
147. Coryn CLS, Hattie JA, Scriven M, Hartmann DJ. Models and mechanisms for evaluating government-funded research - An international comparison. American Journal of Evaluation. 2007;28(4):437-57.
148. Philipp-Bastian Brutscher SW, Grant J. Health Research Evaluation Frameworks: An International Comparison. RAND Europe; 2008.
149. Health Economics Research Group OoHE, RAND Europe. Medical Research: What's it worth? Estimating the economic benefits from medical research in the UK. London: London: UK Evaluation Forum; 2008.
150. Boaz A, Fitzpatrick S, Shaw B. Assessing the impact of research on policy: A literature review. Science and Public Policy. 2009;36(4):255-70.
151. Marjanovic S HS, Wooding Steven. A Historical Reflection on Research Evaluation Studies, Their Recurrent Themes and Challenges. Technical Report.; 2009.
152. Yazdizadeh B, Majdzadeh R, Salmasian H. Systematic review of methods for evaluating healthcare research economic impact. Health Research Policy & Systems. 2010;8:6.
153. Banzi R, Moja L, Pistotti V, Facchini A, Liberati A. Conceptual frameworks and empirical approaches used to assess the impact of health research: an overview of reviews. Health Research Policy & Systems. 2011;9:26.
154. Hanney S. Ways of assessing the economic value or impact of research: is it a step too far for nursing research? Journal of Research in Nursing, vol 16, no 2, Mar 2 2011, p 151-166. 2011.
155. Patel VM, Ashrafian H, Ahmed K, Arora S, Jiwan S, Nicholson JK, et al. How has healthcare research performance been assessed? A systematic review. Journal of the Royal Society of Medicine. 2011;104(6):251-61.
156. Ruscio J, Seaman F, D'Oriano C, Stremlo E, Mahalchik K. Measuring Scholarly Impact Using Modern Citation-Based Indices. Measurement: Interdisciplinary Research and Perspectives. 2012;10(3):123-46; 24.
157. Bornmann L. What is societal impact of research and how can it be assessed? a literature survey. Journal of the American Society for Information Science & Technology. 2013;64(2):217-33.
158. Susan Guthrie, Watu Wamae, Stephanie Diepeveen, Wooding S, Grant J. Measuring Research. A guide to research evaluation frameworks and tools.: RAND Europe Cooperation 2013.
159. Smith KM, Crookes E, Crookes PA. Measuring Research "Impact" for Academic Promotion: Issues from the Literature. Journal of Higher Education Policy and Management. 2013;35(4):410-20; 11.

160. Carpenter CR, Cone DC, Sarli CC. Using publication metrics to highlight academic productivity and research impact. Academic Emergency Medicine. 2014;21(10):1160-72.
161. Milat AJ, Bauman AE, Redman S. A narrative review of research impact assessment models and methods. Health Research Policy & Systems. 2015;13:18.
162. Moed HF, Halevi G. Multidimensional assessment of scholarly research impact. Journal of the Association for Information Science & Technology. 2015;66(10):1988-2002.
163. Pollitt A, Potoglou D, Patil S, Burge P, Guthrie S, King S, et al. Understanding the relative valuation of research impact: a best-worst scaling experiment of the general public and biomedical and health researchers. BMJ open. 2016;6(8).
164. Thonon F, Boulkedid R, Delory T, Rousseau S, Saghatchian M, van Harten W, et al. Measuring the outcome of biomedical research: a systematic literature review. PloS one. 2015;10(4):e0122239.
165. Paul Wouters MT, Kayvan Kousha, Ludo Waltman, Sarah de Rijcke, Alex, Rushforth aTF. The Metric Tide Literature Review (Supplementary Report I to the Independent Review of the Role of Metrics in Research Assessment and Management). 2015.
166. Agarwal A, Durairajanayagam D, Tatagari S, Esteves SC, Harlev A, Henkel R, et al. Bibliometrics: tracking research impact by selecting the appropriate metrics. Asian Journal of Andrology. 2016;18(2):296-309.
167. Chikoore L. Perceptions, motivations and behaviours towards 'research impact': a cross-disciplinary perspective: University of Loughborough; 2016.
168. Wimmer EN, Rethlefsen ML, Jarvis C, Shipman JP. Understanding Research Impact: A Review of Existing and Emerging Tools for Nursing. Journal of Professional Nursing. 2016;32(6):401-11.
169. Bornmann L. Measuring Impact in Research Evaluations: A Thorough Discussion of Methods For, Effects of and Problems with Impact Measurements. Higher Education: The International Journal of Higher Education Research. 2017;73(5):775-87; 13.
170. Cruz Rivera S, Kyte DG, Aiyegbusi OL, Keeley TJ, Calvert MJ. Assessing the impact of healthcare research: A systematic review of methodological frameworks. PLoS Medicine / Public Library of Science. 2017;14(8):e1002370.
171. Deeming S, Searles A, Reeves P, Nilsson M. Measuring research impact in Australia's medical research institutes: a scoping literature review of the objectives for and an assessment of the capabilities of research impact assessment frameworks. Health Research Policy & Systems. 2017;15(1):22.
172. Peter N, Kothari A, Masood S. Identifying and understanding research impact: A review for occupational scientists. J Occup Sci. 2017;24(3):377-92.
173. Reale E, Avramov D, Canhial K, Donovan C, Flecha R, Holm P, et al. A review of literature on evaluating the scientific, social and political impact of social sciences and humanities research. Res Evaluat. 2017;27(4):298-308.
174. Newson RK, ; L.;; Rychetnik,; L.;; Milat,; A.;; Bauman,; A. Looking; both; ways: a; review; of; methods; for assessing; research impacts; on; policy; and; the; policy; utilisation of research. Health; research; policy; and; systems. 2018;16(1):54-018-0310-4.
175. Pedrini ML, Valentina; Battaglia, Mario Alberto; Zaratin, Paola. Assessing the health research's social impact: a systematic review. Scientometrics. 2018;114(3):1227-50.

176. Weishuhn PH, ; Katharina;; Ferretti,; Johanna. Research; Impact; Assessment; in; Agriculture--A; Review; of; Approaches; and; Impact; Areas. Research; Evaluation. 2018;27(1):36-42;; 7.
177. Braithwaite J, Herkes J, Churruca K, Long J, C Pomare C, Boyling C, et al. Comprehensive Researcher Achievement Model (CRAM): a framework for measuring researcher achievement, impact and influence derived from a systematic literature review of metrics andmodels. BMJ open. 2019;9(3).
178. Heyeres MT, ; K.;; Yang,; Y.;; Yan,; L.;; Jiang,; H. The; characteristics; and; reporting; quality; of; research; impact; case; studies:; A; systematic; review. Evaluation; and; program; planning. 2019;73:10-23.
179. Sussex J, Feng Y, Mestre-Ferrandiz J, Pistollato M, Hafner M, Burridge P, et al. Quantifying the economic impact of government and charity funding of medical research on private research and development funding in the United Kingdom. BMC medicine. 2016;14(1):32.
180. Raftery J HS, Greenhalgh T, Glover M & Blatch-Jones A. Models and applications for measuring the impact of health research: Update of a systematic review for the Health Technology Assessment Programme(. Health technology assessment. 2016;20(76).
181. Buxton M, Hanney S. How can payback from health services research be assessed? Journal of health services research & policy. 1996;1(1):35-43.
182. England PH. Introduction to logic models 2018 [Available from: https://www.gov.uk/government/publications/evaluation-in-health-and-well-being-overview/introduction-to-logic-models.
183. Klautzer L, Hanney S, Nason E, Rubin J, Grant J, Wooding S. Assessing policy and practice impacts of social science research: the application of the Payback Framework to assess the Future of Work programme. Res Evaluat. 2011;20(3):201-9.
184. Brown MLN, S. G.; Shibley, L. M. Adjuvant therapy for stage III colon cancer: economics returns to research and cost-effectiveness of treatment. Journal of the National Cancer Institute. 1994;86(6):424-30.
185. Ugolini DB, A.; Parodi, S.; Casilli, C.; Santi, L. Assessing research productivity in an oncology research institute: the role of the documentation center. Bulletin of the Medical Library Association. 1997;85(1):33-8.
186. Ugolini DC, C.; Mela, G. S. Assessing oncological productivity. is one method sufficient? European journal of cancer. 2002;38(8):1121-5.
187. Coyle DG, E.; Wells, G. The assessment of the economic return from controlled clinical trials. A framework applied to clinical trials of colorectal cancer follow-up. The European journal of health economics : HEPAC : health economics in prevention and care. 2003;4(1):6-11.
188. Lewison GS, R. The impact of cancer research: how publications influence UK cancer clinical guidelines. British journal of cancer. 2008;98(12):1944-50.
189. Lewison G, Tootell S, Roe P, Sullivan R. How do the media report cancer research? A study of the UK's BBC website. British journal of cancer. 2008;99(4):569-76.
190. Saad EDM, A.; Masson, A. L.; Prisco, F. E. The geography of clinical cancer research: analysis of abstracts presented at the American Society of Clinical Oncology Annual Meetings. Annals of oncology : official journal of the European Society for Medical Oncology. 2010;21(3):627-32.
191. Lewison G, Markusova V. The evaluation of Russian cancer research. Res Evaluat. 2010;19(2):129-44.

192. Lakdawalla DNS, E. C.; Jena, A. B.; Reyes, C. M.; Goldman, D. P.; Philipson, T. J. An economic evaluation of the war on cancer. Journal of health economics. 2010;29(3):333-46.
193. S M, R V. Evaluation of RT&D: from 'prescriptions for justifying' to 'user-oriented guidance for learning'. Res Evaluat. 2010;19(4):251-61.
194. Sullivan R, Lewison G, Purushotham AD. An analysis of research activity in major UK cancer centres. European journal of cancer. 2011;47(4):536-44.
195. Donovan C, Butler L, Butt AJ, Jones TH, Hanney SR. Evaluation of the impact of National Breast Cancer Foundation-funded research. The Medical journal of Australia. 2014;200(4):214-8.
196. Guthrie S, Pollitt A, Hanney S, Grant J. Investigating Time Lags and Attribution in the Translation of Cancer Research: A Case Study Approach. Rand health quarterly. 2014;4(2):16.
197. Excellence NIfHaC. Guide to the Methods of Technology Appraisal: 2013. 2013.
198. Government of Canada. Statistics Canada 2020 [Available from: https://www.statcan.gc.ca/eng/start.
199. Glover M, Montague E, Pollitt A, Guthrie S, Hanney S, Buxton M, et al. Estimating the returns to United Kingdom publicly funded musculoskeletal disease research in terms of net value of improved health outcomes. Health research policy and systems. 2018;16(1):1.
200. Wooding S, Pollitt A, Castle-Clarke S, Cochrane G, Diepeveen S, Guthrie S, et al. Mental Health Retrosight: Understanding the Returns From Research (Lessons From Schizophrenia): Policy Report. Rand health quarterly. 2014;4(1):8-.
201. Guthrie S, Hafner M, Bienkowska-Gibbs T, Wooding S. Returns on research funded under the NIHR Health Technology Assessment (HTA) Programme. . Health technology assessment. 2015;19(67).
202. Weiss AP. Measuring the impact of medical research: moving from outputs to outcomes. The American journal of psychiatry. 2007;164(2):206-14.
203. Kuruvilla S, Mays N, Pleasant A, Walt G. Describing the impact of health research: a Research Impact Framework. BMC health services research. 2006;6:134.
204. Landry RA, N. Lamari, M. Climbing the Ladder of Research Utilization. Evidence from Social Science Research. Science Communication. 2001;22(4):396-422.
205. Dissemination UoYCfRa. PROSPERO International prospective register of systematic reviews 2021 [Available from: https://www.crd.york.ac.uk/prospero/.
206. 2014REF. Decisions on assessing research impact. Northavon House, Coldharbour Lane, Bristol, BS16 1QD; 2011. Contract No.: 8th August 2019.
207. Terämä E, Smallman M, Lock SJ, Johnson C, Austwick MZ. Beyond Academia – Interrogating Research Impact in the Research Excellence Framework. PloS one. 2016;11(12):e0168533.
208. King's College London and Digital Science. The nature, scale and beneficiaries of research impact: An initial analysis of Research Excellence Framework (REF) 2014 impact case studies. HEFCE; 2015.
209. Greenhalgh T, Fahy N. Research impact in the community-based health sciences: an analysis of 162 case studies from the 2014 UK Research Excellence Framework. BMC medicine. 2015;13.
210. Russell Group. About 2019 [Available from: https://www.russellgroup.ac.uk/.

211. Russell Group. Research Excellence Framework 2014 [Available from: https://russellgroup.ac.uk/news/research-excellence-framework/.
212. Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer. The Lancet. 2005;365(9453):60-2.
213. University College London. Introduction of aromatase inhibitors for the treatment of breast cancer. 2014 [Available from: https://impact.ref.ac.uk/casestudies/CaseStudy.aspx?Id=34811.
214. Queen Mary University of London. Anastrozole for oestogen receptor positive breast cancer. 2014 [Available from: https://impact.ref.ac.uk/casestudies/CaseStudy.aspx?Id=18352
215. University of Manchester. Improving outcomes of women diagnosed with and at increased risk of breast cancer: the results of translational research and national and international clinical trials.: Higher Education Funding Council,; 2014 [Available from: https://impact.ref.ac.uk/casestudies/CaseStudy.aspx?Id=28102.
216. University of Manchester. Preventing bone loss in patients treated for breast cancer. : Higher Education and Funding Coucil,; 2014 [Available from: https://impact.ref.ac.uk/casestudies/CaseStudy.aspx?Id=28032.
217. Institute of Cancer Research. Aromatase inhibitors in breast cancer treatment: Higher Education and Funding Council,; 2014 [Available from: https://impact.ref.ac.uk/casestudies/CaseStudy.aspx?Id=41981.
218. Hamdy FC, Donovan JL, Lane JA, Mason M, Metcalfe C, Holding P, et al. 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. New England Journal of Medicine. 2016;375(15):1415-24.
219. University of Bristol. Avoiding harm and evaluating benefit: establishing and implementing an evidence-based policy for prostate cancer screening in the UK: Higher Education Funding Council for England; 2014 [Available from: https://impact.ref.ac.uk/casestudies/CaseStudy.aspx?Id=40165.
220. University of Cambridge. The ProtecT Trial and Associated Translational Research – Management of Localised Prostate Cancer -Neal 2014 [Available from: https://impact.ref.ac.uk/casestudies/CaseStudy.aspx?Id=29986.
221. University of Oxford. UOA02-05: Hormone Replacement Therapy and Cancer Risk: The Million Women Study 2014 [Available from: https://impact.ref.ac.uk/casestudies/CaseStudy.aspx?Id=9539.
222. University College London. Inclusion of flexible sigmoidoscopy in the UK Bowel Cancer Screening Programme 2014 [Available from: https://impact.ref.ac.uk/casestudies/CaseStudy.aspx?Id=23145.
223. Nottingham Uo. Saving lives through faecal occult blood screening for bowel cancer 2014 [Available from: https://impact.ref.ac.uk/casestudies/CaseStudy.aspx?Id=40998.
224. University of Leeds. Case Study 4. Improving chemotherapy, radiotherapy and patient outcomes for colorectal cancer through patient-focused integrated clinical trials 2014 [cited 2019. Available from: https://impact.ref.ac.uk/casestudies/CaseStudy.aspx?Id=27417.
225. University of Edinburgh. C: Detailed analysis of trial of lapatinib in combination with capecitabine in advanced, HER2+ breast cancer leads to marketing authorisation worldwide. 2014.
226. Higher Education Funding Council. Decisions on assessing research impact. Northavon House, Coldharbour Lane, Bristol, BS16 1QD; 2011 March 2011.
227. Hinrichs S, Grant J. A new resource for identifying and assessing the impacts of research. BMC medicine. 2015;13(1):148.

228. World Health Organisation. Health Statistics and Information systems Disease burden and mortality estimates 2019 [
229. Cancer Research UK. Cancer incidence for common cancers 2020 [Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence/common-cancers-compared#heading-Zero.
230. Cancer Research UK. Cancer mortality for common cancers 2020 [Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/mortality/common-cancers-compared#heading-Zero.
231. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, et al. Improved Survival with Ipilimumab in Patients with Metastatic Melanoma. New England Journal of Medicine. 2010;363(8):711-23.
232. Bundred NJ, Kenemans P, Yip CH, Beckmann MW, Foidart J-M, Sismondi P, et al. Tibolone increases bone mineral density but also relapse in breast cancer survivors: LIBERATE trial bone substudy. Breast Cancer Research. 2012;14(1):R13.
233. Seymour MT, Thompson LC, Wasan HS, Middleton G, Brewster AE, Shepherd SF, et al. Chemotherapy options in elderly and frail patients with metastatic colorectal cancer (MRC FOCUS2): an open-label, randomised factorial trial. Lancet. 2011;377(9779):1749-59.
234. Wong CH, Siah KW, Lo AW. Estimation of clinical trial success rates and related parameters. Biostatistics. 2018;20(2):273-86.
235. Strzebonska K, Waligora M. Umbrella and basket trials in oncology: ethical challenges. BMC Medical Ethics. 2019;20(1):58.
236. Imperial College London. Clinical Development of Temozolomide: An Anticancer Drug that Improves Survival of Patients with Brain Cancer (Glioma). 2014.
237. Department for the Economy, Higher Education Funding Council for Wales, Research England, Scottish Funding Council. REF 2018/02 Consultation on the draft panel criteria and working methods. 2018.
238. Department for the Economy, Higher Education Funding Council for Wales, Research England, Scottish funding Council. REF 2019/01 Guidance on Submissions. 2019.
239. Clarivate Analytics. InCites Journal Citation Reports 2019 [Available from: https://jcr.clarivate.com/JCRJournalHomeAction.action.
240. Queen Mary University of London. Improved sensitivity of breast cancer screening with two-view mammography 2014 [Available from: https://impact.ref.ac.uk/casestudies/CaseStudy.aspx?Id=18319.
241. Stern N. Research Excellence Framework (REF) review: Building on success and learning from experience. 2016.
242. HEFCE. Index of revisions to the 'Guidance on submissions' (2019/01). 2020.
243. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics. 1977;33(1):159-74.
244. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb). 2012;22(3):276-82.
245. Bobak CA, Barr PJ, O'Malley AJ. Estimation of an inter-rater intra-class correlation coefficient that overcomes common assumption violations in the assessment of health measurement scales. BMC Medical Research Methodology. 2018;18(1):93.
246. Carlisle BG, Doussau A, Kimmelman J. Patient burden and clinical advances associated with postapproval monotherapy cancer drug trials: a retrospective cohort study. BMJ open. 2020;10(2):e034306.

247. Iveson TJ, Sobrero AF, Yoshino T, Souglakos I, Ou F-S, Meyers JP, et al. Duration of Adjuvant Doublet Chemotherapy (3 or 6 months) in Patients With High-Risk Stage II Colorectal Cancer. Journal of Clinical Oncology. 2021:JCO.20.01330.
248. Iveson T, Hanna C, Iveson P, Zhang S, Levasseur A, Meyerhardt J. The Early Impact of the IDEA Collaboration Results: How the Results Changed Prescribing Practice. JNCI Cancer Spectrum. 2021;5(4).
249. Yu IS, Pereira AAL, Lee M, Korphaisarn K, Marshall J, Segelov E, et al. Medical Oncologists' Perspectives on How the Results of the IDEA Collaboration Impact the Adjuvant Treatment of Stage III Colon Cancer. The oncologist. 2019.
250. Ouali K, Cohen R, Turpin A, Neuzillet C, Rousseau B, Garcia-Larnicol ML, et al. Impact of the IDEA study on clinical practice for stage III colon cancer patients: A French GERCOR - FFCD - GI UNICANCER national survey. Annals of Oncology. 2019;30:v210.
251. Muzny DM, Bainbridge MN, Chang K, Dinh HH, Drummond JA, Fowler G, et al. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012;487(7407):330-7.
252. Kawakami H, Zaanan A, Sinicrope FA. Microsatellite instability testing and its role in the management of colorectal cancer. Current treatment options in oncology. 2015;16(7):30-.
253. Edwards PJ, Roberts I, Clarke MJ, Diguiseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. The Cochrane database of systematic reviews. 2009(3):MR000008.
254. Excellence NIfHaC. Colorectal cancer NICE guideline [NG151] Published date: 29 January 2020 2020 [Available from: https://www.nice.org.uk/guidance/ng151/chapter/Recommendations#management-of-local-disease.
255. Blinman P, Martin A, Jefford M, Goldstein D, Boadle D, Morris M, et al. Patients' Preferences for 3 Months vs 6 Months of Adjuvant Chemotherapy for Colon Cancer. JNCI Cancer Spectrum. 2021;5(1).
256. Blinman P, Duric V, Nowak AK, Beale P, Clarke S, Briscoe K, et al. Adjuvant chemotherapy for early colon cancer: what survival benefits make it worthwhile? European journal of cancer. 2010;46(10):1800-7.
257. Tournigand C, Andre T, Bonnetain F, Chibaudel B, Lledo G, Hickish T, et al. Adjuvant therapy with fluorouracil and oxaliplatin in stage II and elderly patients (between ages 70 and 75 years) with colon cancer: subgroup analyses of the Multicenter International Study of Oxaliplatin, Fluorouracil, and Leucovorin in the Adjuvant Treatment of Colon Cancer trial. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2012;30(27):3353-60.
258. Taylor JC, Swinson D, Seligmann JF, Birch RJ, Dewdney A, Brown V, et al. Addressing the variation in adjuvant chemotherapy treatment for colorectal cancer: Can a regional intervention promote national change? International journal of cancer. 2020;n/a(n/a).
259. Battaglin F, Naseem M, Lenz HJ, Salem ME. Microsatellite instability in colorectal cancer: overview of its clinical significance and novel perspectives. Clinical advances in hematology & oncology : H&O. 2018;16(11):735-45.
260. B.E. Engelmann TI, E. Høgdall, N.H. Holländer. 4060 - TransSCOT – translational research based on the Short Course Oncology Therapy (SCOT) cohort. European Society for Medical Oncology2016.
261. Haller DG, O'Connell MJ, Cartwright TH, Twelves CJ, McKenna EF, Sun W, et al. Impact of age and medical comorbidity on adjuvant treatment outcomes

for stage III colon cancer: a pooled analysis of individual patient data from four randomized, controlled trials. Annals of oncology : official journal of the European Society for Medical Oncology. 2015;26(4):715-24.
262. McCleary NJ, Meyerhardt JA, Green E, Yothers G, de Gramont A, Van Cutsem E, et al. Impact of Age on the Efficacy of Newer Adjuvant Therapies in Patients With Stage II/III Colon Cancer: Findings From the ACCENT Database. Journal of Clinical Oncology. 2013;31(20):2600-6.
263. Andre T, Vernerey D, Mineur L, Bennouna J, Desrame J, Faroux R, et al. Three Versus 6 Months of Oxaliplatin-Based Adjuvant Chemotherapy for Patients With Stage III Colon Cancer: Disease-Free Survival Results From a Randomized, Open-Label, International Duration Evaluation of Adjuvant (IDEA) France, Phase III Trial. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2018;36(15):1469-77.
264. Meyerhardt JA, Shi Q, Fuchs CS, Niedzwiecki D, Zemla TJ, Kumthekar P, et al. Celecoxib in addition to standard adjuvant therapy with 5-fluorouracil, leucovorin, oxaliplatin (FOLFOX) in stage III colon cancer: Results from CALGB/SWOG 80702. Journal of Clinical Oncology. 2020;38(15_suppl):4003-.
265. Yoshino T, Yamanaka T, Oki E, Kotaka M, Manaka D, Eto T, et al. Efficacy and Long-term Peripheral Sensory Neuropathy of 3 vs 6 Months of Oxaliplatin-Based Adjuvant Chemotherapy for Colon Cancer: The ACHIEVE Phase 3 Randomized Clinical Trial. Jama Oncol. 2019.
266. Ouali K, Turpin A, Neuzillet C, Rousseau B, Garcia-Larnicol M-L, Tournigand C, et al. Impact of the IDEA Collaboration Study Results on Clinical Practice in France for Patients With Stage III Colon Cancer: A National GERCOR - PRODIGE Survey. Clinical Colorectal Cancer. 2020.
267. Seymour MT, Morton D. FOxTROT: an international randomised controlled trial in 1052 patients (pts) evaluating neoadjuvant chemotherapy (NAC) for colon cancer. Journal of Clinical Oncology. 2019;37(15_suppl):3504-.
268. Drummond M, Sculpher M, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. Fourth ed: Oxford; 2015.
269. Attema AE, Brouwer WBF, Claxton K. Discounting in Economic Evaluations. PharmacoEconomics. 2018;36(7):745-58.
270. Brooks R. EuroQol: the current state of play. Health policy. 1996;37(1):53-72.
271. Ramos-Goni JM, Rivero-Arias O. eq5d: A command to calculate index values for the EQ-5D quality-of-life instrument. Stata J. 2011;11(1):120-5.
272. Devlin NJ, Hansen P, Kind P, Williams A. Logical inconsistencies in survey respondents' health state valuations -- a methodological challenge for estimating social tariffs. Health economics. 2003;12(7):529-44.
273. Viney R, Norman R, King MT, Cronin P, Street DJ, Knox S, et al. Time trade-off derived EQ-5D weights for Australia. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2011;14(6):928-36.
274. Burstrom K, Sun S, Gerdtham UG, Henriksson M, Johannesson M, Levin LA, et al. Swedish experience-based value sets for EQ-5D health states. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2014;23(2):431-42.
275. Boyd KA, Briggs AH, Paul J, Iveson T, Midgely R, Harkin A, et al. Analysis of adverse events and quality of life data for an economic evaluation of adjuvant chemotherapy in colorectal cancer: when can we stop collecting? Trials. 2011;12(Suppl 1):A41-A.

276. Billingham LJ, Abrams KR. Simultaneous analysis of quality of life and survival data. Statistical methods in medical research. 2002;11(1):25-48.
277. Raftery J. Costing in economic evaluation. BMJ (Clinical research ed). 2000;320(7249):1597-.
278. Lin DY, Feuer EJ, Etzioni R, Wax Y. Estimating medical costs from incomplete follow-up data. Biometrics. 1997;53(2):419-34.
279. Etzioni R, Urban N, Baker M. Estimating the costs attributable to a disease with application to ovarian cancer. Journal of clinical epidemiology. 1996;49(1):95-103.
280. Paulden M. Calculating and Interpreting ICERs and Net Benefit. PharmacoEconomics. 2020;38(8):785-807.
281. Woods B, Revill P, Sculpher M, Claxton K. Country-Level Cost-Effectiveness Thresholds: Initial Estimates and the Need for Further Research. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2016;19(8):929-35.
282. Cameron D, Ubels J, Norström F. On what basis are medical cost-effectiveness thresholds set? Clashing opinions and an absence of data: a systematic review. Glob Health Action. 2018;11(1):1447828-.
283. Nimdet K, Chaiyakunapruk N, Vichansavakul K, Ngorsuraches S. A Systematic Review of Studies Eliciting Willingness-to-Pay per Quality-Adjusted Life Year: Does It Justify CE Threshold? PloS one. 2015;10(4):e0122760.
284. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD. Model Parameter Estimation and Uncertainty: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. Value in Health. 2012;15(6):835-42.
285. Reed SD. Statistical considerations in economic evaluations: a guide for cardiologists. Eur Heart J. 2014;35(25):1652-6.
286. Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. Health economics. 1997;6(4):327-40.
287. Fenwick E, Marshall DA, Levy AR, Nichol G. Using and interpreting cost-effectiveness acceptability curves: an example using data from a trial of management strategies for atrial fibrillation. BMC health services research. 2006;6(1):52.
288. Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Medical decision making : an international journal of the Society for Medical Decision Making. 1998;18(2 Suppl):S68-80.
289. Mauskopf J, Earnshaw SR, Brogan A, Wolowacz S. Budget-Impact Analysis of Health Care Interventions: Adis, Cham; 2017.
290. Gol-Montserrat J, del Burgo MLM, Quecedo L, del Llano JE. Analysis of Productivity Costs in Cancer: A Systematic Review. Global & Regional Health Technology Assessment. 2017;4(1):grhta.5000262.
291. Pritchard C, Sculpher M. Productivity Costs: Principles and Practice in Economic Evaluation. Office of Health Economics, editor: Office of Health Economics,; 2000.
292. Organisation for Economic Co-operation and Development. OECD Economic Outlook, Volume 2019 Issue 2 2019 [Available from: https://www.oecd-ilibrary.org/economics/oecd-economic-outlook-volume-2019-issue-2_9b89401b-en.
293. Cassidy J, Douillard JY, Twelves C, McKendrick JJ, Scheithauer W, Bustova I, et al. Pharmacoeconomic analysis of adjuvant oral capecitabine vs intravenous

5-FU/LV in Dukes' C colon cancer: the X-ACT trial. British journal of cancer. 2006;94(8):1122-9.
294. GOV.UK. Travel - mileage and fuel rates and allowances 2019 [Available from: https://www.gov.uk/government/publications/rates-and-allowances-travel-mileage-and-fuel-allowances/travel-mileage-and-fuel-rates-and-allowances.
295. OECD. Purchasing Power Parities - Frequently Asked Questions (FAQs) 2021 [Available from: http://www.oecd.org/sdd/purchasingpowerparities-frequentlyaskedquestionsfaqs.htm.
296. Jongeneel G, Greuter MJE, van Erning FN, Koopman M, Vink GR, Punt CJA, et al. Model-based evaluation of the cost effectiveness of 3 versus 6 months' adjuvant chemotherapy in high-risk stage II colon cancer patients. Therapeutic Advances in Gastroenterology. 2020;13:1756284820954114.
297. World Health Organisation. Fact sheet: Cancer 2018 [Available from: https://www.who.int/news-room/fact-sheets/detail/cancer.
298. Bradley CJ, Lansdorp-Vogelaar I, Yabroff KR, Dahman B, Mariotto A, Feuer EJ, et al. Productivity savings from colorectal cancer prevention and control strategies. Am J Prev Med. 2011;41(2):e5-e14.
299. Cancer Research UK. Bowel Cancer Statistics 2016 [
300. Rosenberg LE. Exceptional economic returns on investments in medical research. The Medical journal of Australia. 2002;177(7):368-71.
301. Johnston SC, Rootenberg JD, Katrak S, Smith WS, Elkins JS. Effect of a US National Institutes of Health programme of clinical trials on public health and costs. Lancet. 2006;367(9519):1319-27.
302. Luengo-Fernandez R, Leal J, Gray A, Sullivan R. Economic burden of cancer across the European Union: a population-based cost analysis. The Lancet Oncology. 2013;14(12):1165-74.
303. Vuik FE, Nieuwenburg SA, Bardou M, Lansdorp-Vogelaar I, Dinis-Ribeiro M, Bento MJ, et al. Increasing incidence of colorectal cancer in young adults in Europe over the last 25 years. Gut. 2019;68(10):1820-6.
304. Esnaola NF, Mauldin P, Ebeling M, Hulsey T, Liu B, Thomas MB, et al. Adjuvant chemotherapy for stage II colon cancer: Who (really) needs it? Journal of Clinical Oncology. 2011;29(15).
305. McGee VE, Carleton WT. Piecewise Regression. Journal of the American Statistical Association. 1970;65(331):1109-24.
306. Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. International journal of epidemiology. 2017;46(1):348-55.
307. Gujarati. DN, Porter. DC. Basic Econometrics. Fifth ed: McGraw-Hill Education; 2009.
308. Dixon S, Coleman P, Nicholl J, Brennan A, Touch S. Evaluation of the impact of a technology appraisal process in England: the South and West Development and Evaluation Committee. Journal of health services research & policy. 2003;8(1):18-24.
309. Jandoc R, Burden AM, Mamdani M, Levesque LE, Cadarette SM. Interrupted time series analysis in drug utilization research is increasing: systematic review and recommendations. Journal of clinical epidemiology. 2015;68(8):950-6.
310. Network WoSCNCCMC. Audit Report Colorectal Cancer Quality Performance Indicators Clinical Audit Data: 1st April 2018 and 31st March 2019. 2019.

311. Public Health England. SACT Systemic Anti-cancer Therapy Chemotherapy Dataset 2011 [Available from: http://www.chemodataset.nhs.uk/home.
312. Harron K, Dibben C, Boyd J, Hjern A, Azimaee M, Barreto ML, et al. Challenges in administrative data linkage for research. Big Data Soc. 2017;4(2):2053951717745678-.
313. Lugg-Widger FV, Angel L, Cannings-John R, Hood K, Hughes K, Moody G, et al. Challenges in accessing routinely collected data from multiple providers in the UK for primary studies: Managing the morass. International Journal of Population Data Science. 2018;3(3).
314. Dattani N, Hardelid P, Davey J, Gilbert R. Accessing electronic administrative health data for research takes time. Archives of disease in childhood. 2013;98(5):391-2.
315. GOV.UK. Policy paper National Data Strategy 2020 [Available from: https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy.
316. Scottish Government. Beating Cancer: Ambition and Action (2016) update: achievements, new action and testing change. 2016.
317. Cancer Medicines Outcome Programme Team. Cancer Medicines Outcomes Programme (CMOP) Phase 1 Report (October 2016 – March 2020). 2020 August 2020.
318. Public Health Scotland. Scottish Cancer Registry 2020 [Available from: https://www.isdscotland.org/Health-Topics/Cancer/Scottish-Cancer-Registry/.
319. Blakemore C, Davidson J. Putting a value on medical research. The Lancet. 2006;367(9519):1293-5.
320. Dexamethasone in Hospitalized Patients with Covid-19. New England Journal of Medicine. 2020;384(8):693-704.
321. UK CR. Our strategy to beat cancer sooner 2020 [Available from: https://www.cancerresearchuk.org/about-us/our-organisation/our-strategy-to-beat-cancer-sooner#:~:text=Our%20vision%20is%20to%20bring,doubled%20%E2%80%93%20today%20half%20will%20survive.
322. Health NIo. National Cancer Insitute 2020 [Available from: https://www.nih.gov/about-nih/what-we-do/nih-almanac/national-cancer-institute-nci.
323. The Institute of Cancer Research. Our mission 2020 [Available from: https://www.icr.ac.uk/about-us/our-mission.
324. Flatiron. Cancer is smart. Together, we can be smarter. 2021 [Available from: https://flatiron.com/.
325. Data VCfGCR. A global clinical research data sharing platform 2021 [Available from: https://vivli.org/.
326. University Y. 2021 [Available from: https://yoda.yale.edu/.
327. ideaPoint. Welcome to ClinicalStudyDataRequest.com. 2021.
328. Consortium YHE. Return on Investment [online] York2016 [Available from: https://yhec.co.uk/glossary/return-on-investment/.
329. Toole Andrew A. Does Public Scientific Research Complement Private Investment in Research and Development in the Pharmaceutical Industry? The Journal of Law and Economics. 2007;50(1):81-104.
330. Ward MR, Dranove D. The Vertical Chain of Research and Development in the Pharmaceutical Industry. Economic Inquiry. 1995;33(1):70-87.
331. Conroy T, Lamfichekh N, Etienne P-L, Rio E, Francois E, Mesgouez-Nebout N, et al. Total neoadjuvant therapy with mFOLFIRINOX versus preoperative chemoradiation in patients with locally advanced rectal cancer: Final results of

PRODIGE 23 phase III trial, a UNICANCER GI trial. Journal of Clinical Oncology. 2020;38(15_suppl):4007-.
332. Garcia-Aguilar J, Patil S, Kim JK, Yuval JB, Thompson H, Verheij F, et al. Preliminary results of the organ preservation of rectal adenocarcinoma (OPRA) trial. Journal of Clinical Oncology. 2020;38(15_suppl):4008-.
333. Hospers G, Bahadoer RR, Dijkstra EA, van Etten B, Marijnen C, Putter H, et al. Short-course radiotherapy followed by chemotherapy before TME in locally advanced rectal cancer: The randomized RAPIDO trial. Journal of Clinical Oncology. 2020;38(15_suppl):4006-.
334. U.S. Food and Drug Administration. FDA grants accelerated approval to pembrolizumab for first tissue/site agnostic indication 2017 [Available from: https://www.fda.gov/drugs/resources-information-approved-drugs/fda-grants-accelerated-approval-pembrolizumab-first-tissuesite-agnostic-indication.
335. Chalabi M, Fanchi LF, Dijkstra KK, Van den Berg JG, Aalbers AG, Sikorska K, et al. Neoadjuvant immunotherapy leads to pathological responses in MMR-proficient and MMR-deficient early-stage colon cancers. Nat Med. 2020;26(4):566-76.
336. Royal Marsden NHS Foundation Trust. Avelumab Plus 5-FU Based Chemotherapy as Adjuvant Treatment for Stage 3 MSI-High or POLE Mutant Colon Cancer (POLEM): clinicaltrials.gov; 2020 [Available from: https://clinicaltrials.gov/ct2/show/NCT03827044.
337. International Agency for Research on Cancer: World Health Organisation. Denmark Source: Globocan 2018
2018 [Available from: https://gco.iarc.fr/today/data/factsheets/populations/208-denmark-fact-sheets.pdf.
338. Australian Institute of Health and Welfare (AIHW) Canberra. 2018 Cancer Data in Australia. Canberra, Australia; 2018.
339. Danish Colorectal Cancer Group. Nationwide database for cancer in colon and rectum (DCCG.dk) Clinical Basic Report 2018. 2018.
340. Red Espanola de Registros de Cancer (REDECAN). ESTIMACIONES DE LA INCIDENCIA DEL CÁNCER EN ESPAÑA, 2019. 2019.
341. Swedish Colorectal Cancer Registry. Onkologirapport 2018 från Svenska kolorektalcancerregistret
2020.
342. van Steenbergen LN, Rutten HJT, Creemers GJ, Pruijt JFM, Coebergh JWW, Lemmens VEPP. Large age and hospital-dependent variation in administration of adjuvant chemotherapy for stage III colon cancer in southern Netherlands. Annals of Oncology. 2010;21(6):1273-8.
343. Lima ISF, Yasui Y, Scarfe A, Winget M. Association between receipt and timing of adjuvant chemotherapy and survival for patients with stage III colon cancer in Alberta, Canada. Cancer. 2011;117(16):3833-40.
344. Upadhyay S, Dahal S, Bhatt VR, Khanal N, Silberstein PT. Chemotherapy use in stage III colon cancer: a National Cancer Database analysis. Ther Adv Med Oncol. 2015;7(5):244-51.
345. Boland GM, Chang GJ, Haynes AB, Chiang YJ, Chagpar R, Xing Y, et al. Association between adherence to National Comprehensive Cancer Network treatment guidelines and improved survival in patients with colon cancer. Cancer. 2013;119(8):1593-601.
346. Ortiz-Ortiz KJ, Tortolero-Luna G, Rios-Motta R, Veintidos-Feliu A, Hunter-Mellado R, Torres-Cintron CR, et al. Use of adjuvant chemotherapy in patients

with stage III colon cancer in Puerto Rico: A population-based study. PloS one. 2018;13(3):e0194415.
347. Yang Y, Yang Y, Yang H, Wang F, Wang H, Chen Q, et al. Adjuvant chemotherapy for stage II colon cancer: who really needs it. Cancer management and research. 2018;10:2509-20.
348. Fotheringham S, Mozolowski GA, Murray EMA, Kerr DJ. Challenges and solutions in patient treatment strategies for stage II colon cancer. Gastroenterology report. 2019;7(3):151-61.
349. Centre He-R. OCTOPUS: Obesity and Cancer TOgether imPact Upon Survival 2020 [Available from: https://www.herc.ac.uk/research_project/octopus/.