Conceptualising, Measuring and Valuing the Impact of Health Technology Assessment

Eleanor Grieve

BA (Hons), MPH

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

Health Economics and Health Technology Assessment (HEHTA)

Institute of Health and Wellbeing

University of Glasgow

July 2020

Abstract

How do we assess the impact of Health Technology Assessment (HTA)? Whilst high-income countries (HIC) may have led the way, lower-income countries are increasingly beginning to develop HTA processes to assist in their healthcare decision-making. Understanding how we might quantify the costs and benefits of investing in HTA is important to policy makers and donors. Very few studies have, however, estimated the benefits of the process of HTA in terms of its value to the health system. The global expansion of HTA, its variable implementation, the lack of quantified evidence on health outcomes, along with an increasing investment in these processes at the systems level in low- and middle-income countries (LMIC) has generated greater interest from policy makers about the value and return on investment (ROI) of HTA. A lack of longer-term impact assessment (IA) may undermine its importance and value.

To fill this research gap, we have developed a methodological framework to estimate the ROI in HTA using net health benefits (NHB) as our measure of value. This is the difference between QALYs gained by an intervention and QALYs that could have been gained if the money required to deliver it had been spent on other interventions. We use a mixed-methods approach to quantify the value of HTA and to produce explanatory programme theory on the mechanisms by which HTA impact can be optimised. It is also important to consider opportunity costs when establishing HTA processes but which are often overlooked. The aim is to convey the concepts of potential and realised population NHB, and what we can attribute to the HTA process. Central to understanding this is the 'value of implementation' (VOImp). Theory-driven approaches will be used to generate and test contextual explanations for gaps between expected and actual gains in population health.

We envisage the use of this research will encourage accountability of spending decisions and help to optimise the impact of HTA in an era of investment and expansion, in particular, for LMICs, through better understanding of HTA's role in delivering health outcomes and value for money at the system level. This research will offer a forward-looking model that LMICs can point to as a reference for their own implementation.

List of presentations and papers

Oral Presentations

Grieve, E. *The Value of Health Technology Assessment: A Realist Synthesis*. HTAi. Cologne, Germany, June 2019.

Grieve, E. *The Value of Health Technology Assessment: A Mixed Methods Framework*. HTAi. Cologne, Germany, June 2019.

Grieve, E. *The Value of HTA*. NHS Healthcare Improvement Scotland 7th Annual Research Symposium: Evidence for Value in an era of realistic medicine. Edinburgh, 2019.

Grieve, E. *The Value of HTA: a mixed methods framework*. International Society for Priorities in Health. Linkoping, Sweden, 2018

Grieve, E. *Developing a theoretical framework for assessing the impact of HTA and the 'return on investment'*. Health Policy Evaluation and Technology Assessment Meetings, Xiamen, China 2017.

Grieve, E; Morris, L; Teerawattananon, Y; Newbatt, E. *From priority-setting decisions to health impact*. Panel session, International Society for Priorities in Health. Birmingham, 2016

Grieve, E. *Evaluating the impact of HTA and better decision-making on health outcomes*. Centre for Health Economics Seminar, York. 2016

Papers

Grieve E, Briggs, A, Hesselgreaves H, Wu O *et al* Conceptualising and Measuring the Value and Impact of Health Technology Assessment: A Mixed Methods Framework (under peer review Health Economics, revise and resubmit 2020)

Bouttell J, Grieve E and Hawkins N. The Role of Development-Focused Health Technology Assessment in Optimising Science, Technology and Innovation to

Achieve Sustainable Development Goal 3. Oxford University Press (in press Summer 2020)

Grieve E and Briggs A. IDSI Reference Case work stream [version 1; not peer reviewed]. *F1000Research* 2019, **8**:803 (document) (https://doi.org/10.7490/f1000research.1116834.1)

Grieve E, Hesselgreaves H, Wu O *et al*. The Value of Health Technology Assessment: a mixed methods framework [version 1; not peer reviewed]. *F1000Research* 2017, **6**:2171 (document) (https://doi.org/10.7490/f1000research.1115169.1)

Grieve E and Briggs A. A Methodological Approach for Measuring the Impact of HTA [version 1; not peer reviewed]. *F1000Research* 2017, **6**:249 (document) (https://doi.org/10.7490/f1000research.1113738.1)

Grieve, E and Briggs A. Conceptualising and Measuring the Value and Impact of Health Technology Assessment: A Mixed Methods Framework. End of Grant Report submitted to iDSI 2019.

Table of Contents

Contents

List of Tables

List of Figures

List of Equations

Acknowledgements

Thank you to my supervisors, Professors Andy Briggs, Olivia Wu and Hannah Hesselgreaves. Working with such outstanding talent is not lost on me but above all, I thank you for your friendships. Andy, for opening the door for me (the backdoor..?) to a career in HEHTA. Olivia, for being such a compassionate and supportive manager, it is truly appreciated. Hannah, for your empathy and allowing me to benefit from your invaluable support on the realist work even after your move. I would also like to thank Prof Liz Fenwick for being a fantastic line manager and teacher. And Alieda, for looking after me at work.

This research was initiated by iDSI, and my thanks especially to Kalipso Chalkidou, Francis Ruiz and Y-Ling Chi. Also, to Prof Mark Sculpher and Paul Revill for their invaluable support right from the outset. Thanks go to Itad too for being able to draw on their qualitative research in China. I would like to thank the HTA agencies which provided me with case studies for this thesis: the China National Health Development Research Center (CNHDRC), especially Wudong (Victor) Guo for allowing me to use the outputs of his economic modelling; and the Scottish Health Technologies Group (SHTG), especially to Jess Kandulu and Ed Clifton for fully supporting me in using one their HTAs. I would like to acknowledge my colleague, Janet Bouttell, for helping me out on quasi-experimental methods, and for allowing me to adapt her summary.

To dad, for all your emotional and academic support. To Bob & Bec, my sis Adama, Muhammed Hamish, Casper and Huey. I love you all. To all my friends and family who have kept me going over a tough year with your laughter and constant nonsense/messaging. Here's to love, laughter and lifelong friendships.

In memory of Tony, a deafblind child I met in Nairobi – and for all Tonys – who will always remain my motivation to strive for a fairer, more equal world and evidence-based health care for all, especially for the most vulnerable children in low-income countries.

I dedicate this to my gentle, kind and beautiful mum. I know you wanted to see me in a pink-trimmed hood one day. This is for you. Love and miss you always, mum – my best friend.

Author's Declaration

I declare that, with the exception of referencing, the research presented is original and there are no conflicts of interest.

Note that I have opted to write in the plural 'we'. Whilst this is my thesis, I wanted to recognise the contribution of others.

This research was produced as part of the International Decision Support Initiative (www.idsihealth.org), a global initiative to support decision makers in priority-setting for universal health coverage. This work received funding support from Bill & Melinda Gates Foundation, the Rockefeller Foundation and the UK Department for International Development.

Eleanor Grieve

July 2020

Abbreviations

BMGF	Bill and Melinda Gates Foundation
CGD	Center for Global Development
CHE	Centre for Heath Economics
CMO	Context-Mechanism-Outcome
CMOC	Context-Mechanism-Outcome-Configuration
CNHDRC	China National Health Development Research Center
CP	Clinical pathways
DFID	Department for International Development
EIPS	Evidence-informed Priority Setting
EUnetHTA	European Network for Health Technology Assessment
GDP	Gross Domestic Product
HITAP	Health Intervention and Technology Assessment Program
HIC	High Income Countries
HTA	Health Technology Assessment
HTAi	Health Technology Assessment international
IA	Impact assessment
IDSI	International Decision Support Initiative
iHEA	International Health Economics Association
INAHTA	International Network of Agencies for Health Technology Assessment
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
LMIC	Low- and Middle-Income Countries
M&E	Monitoring and Evaluation
MEL	Monitoring, Evaluation and Learning
MVLS	College of Medicine, Veterinary and Life Sciences
NHB	Net Health Benefits
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NLEM	National List of Essential Medicines
NMB	Net Monetary Benefits
PRICELESS	Priority Cost-effective Lessons for System Strengthening
SA	South Africa
SSA	Sub-saharan Africa
RE	Realist Evaluation
RCT	Randomised Controlled Trial
ROI	Return on Investment
RS	Realist Synthesis
SHTG	Scottish Health Technologies Group
SROI	Social Return on Investment
ToC	Theory of Change
UoG	University of Glasgow
UHC	Universal Health Coverage
VOImp	Value of Implementation
WHO	World Health Organisation

Part 1: Background and introduction

Part 1 sets out the background to this research and introduces its aims, objectives and thesis structure. The objective of this first part of the thesis is to explain how and why our HTA impact framework was developed and to put it in the context of the existing literature. It comprises chapters 1 – 3 as described below.

Chapter 1 introduces the rationale for, and background to, this research. It also outlines the thesis chapters which are structured in 4 Parts.

Chapter 2 is a review of the literature on evaluating the impact of HTA. We are specifically interested in the impact HTA has on net health benefits (NHB) and its value to the health system.

Chapter 3 discusses philosophical and methodological issues regarding causality, the different ways of thinking about this, the nature of its measurement and the challenge it presents in the design of impact evaluations, especially of complex interventions such as HTA. A comparison group (a counterfactual) enables us to estimate changes in outcome that can we can attribute to an intervention, here the HTA process. A comparator is crucial to impact evaluations as without this, it can lead to erroneous measures and conclusions of impact with attribution which could be wrongly assigned or interpreted. This chapter provides a discussion of, on the one hand, the counterfactual approach to cause and effect and different means by which we might measure and apply this to HTA; and, on the other, theory driven approaches to understand what underpins outcomes leading to impact.

1 Introduction

1.1 Background

Health care resources are finite in every setting and, irrespective of the financing and organisation of a country's healthcare system, decisions on what interventions to cover, and under what circumstances, have to be made in an evidence-based and fair way(1). Health technology assessment (HTA) is one of the tools for priority setting. The World Health Organisation (WHO) defines HTA as: 'the systematic evaluation of properties, effects, and/or impacts of health technology. It is a multidisciplinary process to evaluate the social, economic, organisational and ethical issues of a health intervention or health technology. The main purpose of conducting an assessment is to inform a policy decision making'(2). This is just one of several definitions of HTA in current usage. We expand on its definition and terminology later in this chapter and make a distinction between HTA at the systems level, and its discrete application to one or several health technologies. In essence, HTA helps decision-makers to understand the consequences of one (or several) alternative course of actions and to select the options that produce the best outcomes at the lowest cost.

Globally, a growing commitment to universal health coverage (UHC) is promoting the role and institutionalisation of HTA (3-6) including in low- and middle-income countries (LMIC). Institutionalisation of HTA is seen as pivotal to supporting UHC as a means of supporting better allocation of finite resources, cost containment and the maximisation of health. All countries face complex financing challenges on the journey to sustainable development. As countries' Gross Domestic Product (GDP) per capita increases, they are expected to transition from aid and to take responsibility for strategic planning of their investments to progress towards UHC. For those countries transitioning from donor support to a greater reliance on domestic resources and consequently, with an increased risk of financing gaps in the social sector in particular(7), priority setting in health is a necessary response to inform decision-makers on which interventions are best able to achieve UHC goals in a budget constrained environment(8).

Thus, whilst high-income countries (HIC) may have led the way, LMIC are increasingly beginning to develop HTA processes to assist in their healthcare

decision-making. For example, in India, HTA has been used to inform national clinical guidelines and quality standards to improve the quality-of-care delivery; and in Sub-Saharan Africa, countries such as Ghana, Tanzania and Zambia are seeking to use HTA for their health benefit package design and other purchasing decisions(9-11). In Thailand, a country at the forefront of UHC, HTA has been used to inform decision-making for over a decade, notably to define the benefits package and the National List of Essential Medicines (NLEM)(12).

The benefits concerning the link between HTA at the systems level and outcomes in terms of health improvements have, however, rarely been quantified. A lack of longer-term impact assessment (IA) may undermine its importance and value. Furthermore, it is evident that significant amounts of scarce resources are invested in HTA. It has been estimated that the National Institute for Health and Care Excellence (NICE) for England and Wales depends on about 2,000 external experts and spends on average £150,000 for conducting an hta on each new drug, as based on an invitation to tender issued by the Department of Health in 2009(13). Although countries where HTA has been institutionalised spend relatively little as a percentage of total health spending (estimates range from 0.01–1 percent of total public spending) on these processes(14, 15), HTA competes with other health priorities in the context of scarce resources. Budget constraints and a lack of funding was identified as a key reason by 77 countries out of 111 country respondents for not investing in HTA according to a survey undertaken by the WHO in 2015(16).

The global expansion of HTA, the lack of quantified evidence on health outcomes, its variable implementation resulting in sub-optimal impact, along with an increasing investment in these processes at the systems level, has generated greater interest from policy makers and donors about the value and return on investment (ROI) in HTA. Indeed, this research was initiated by the international Decision Support Initiative (iDSI), a partnership-based initiative working towards achieving UHC and which is funded by the Bill and Melinda Gates Foundation (BMGF), the Rockefeller Foundation and the UK Department for International Development (DFID). The next two sub-sections describe the work of iDSI in relation to this research.

1.1.1 The international Decision Support Initiative

Established in November 2013, iDSI is a global network of health, policy and economic expertise, working to increase the value and impact of health spending(11). The initiative was launched by the Global Health and Development Group at Imperial College London, the Health Intervention and Technology Assessment Program, Thailand (HITAP), the Center for Global Development, Priority Cost Effective Lessons for Systems Strengthening South Africa (PRICELESS SA), and the China National Development and Research Center (CHNDRC). iDSI supports countries to make better decisions about how much public money to spend on healthcare and how to make that money go further. Their strategy is 'to work with LMIC governments and global development funders to create lasting, country-owned institutional capacity for evidence-informed priority setting – leading to more cost-effective, equitable and sustainable resource allocation which will translate into higher quality healthcare coverage, reduced financial impoverishment for households, and ultimately better health and more lives saved'(11). iDSI offers skills and expertise on priority setting in health, ranging from the production of research and cost-effectiveness tools to policy support and technical assistance. It relies on national institutions and invests in building local capacity to conduct priority setting exercises. To this end, iDSI supports countries to build a sustainable and locally relevant HTA mechanism for priority setting to support progress towards UHC(17).

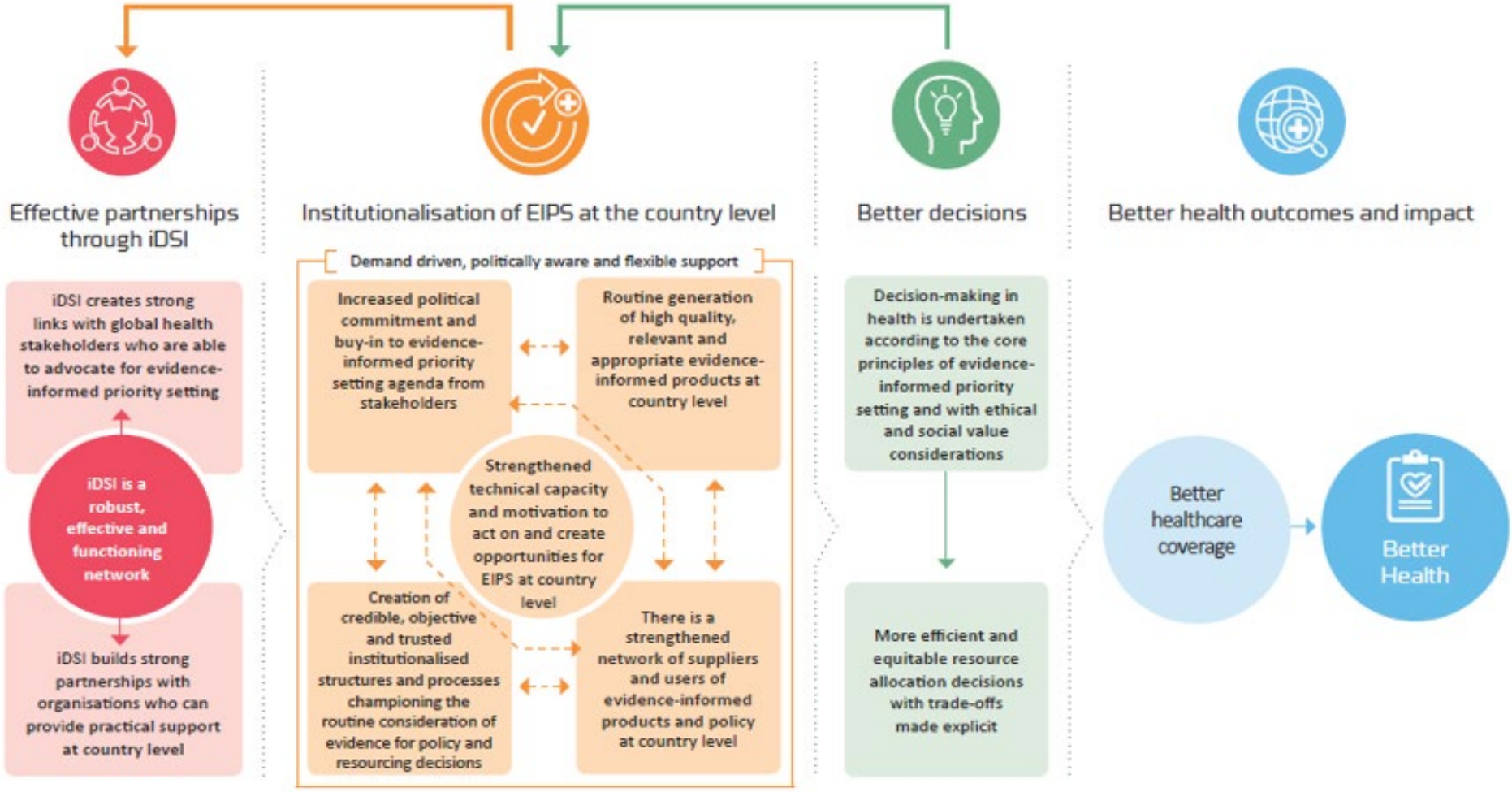
1.1.2 Measuring the impact of 'Better Decisions'

Integral to their work is iDSI's commitment to measuring the impact of 'better decisions' to share learning with partners around the world. In order to do this, a specialist monitoring and evaluation (M&E) agency, itad[1], partnered with iDSI to develop a monitoring, evaluation and learning (MEL) framework based on a Theory of Change (ToC). This ToC [Figure 1.1] states that combining demand-driven support and policy-informed knowledge products with institutional and procedural support would encourage better decisions about the use of resources for health(18). It encompasses the three main components of 'effective partnerships', 'stronger institutions' and 'better decisions' to help elevate the

[1] Note: itad is not an abbreviation

value of priority-setting. As resources are used more consistently to prevent, detect and treat the major burdens of disease across societies, the theory stipulates that population health improves.

Figure 1-1 itad/iDSI Theory of Change

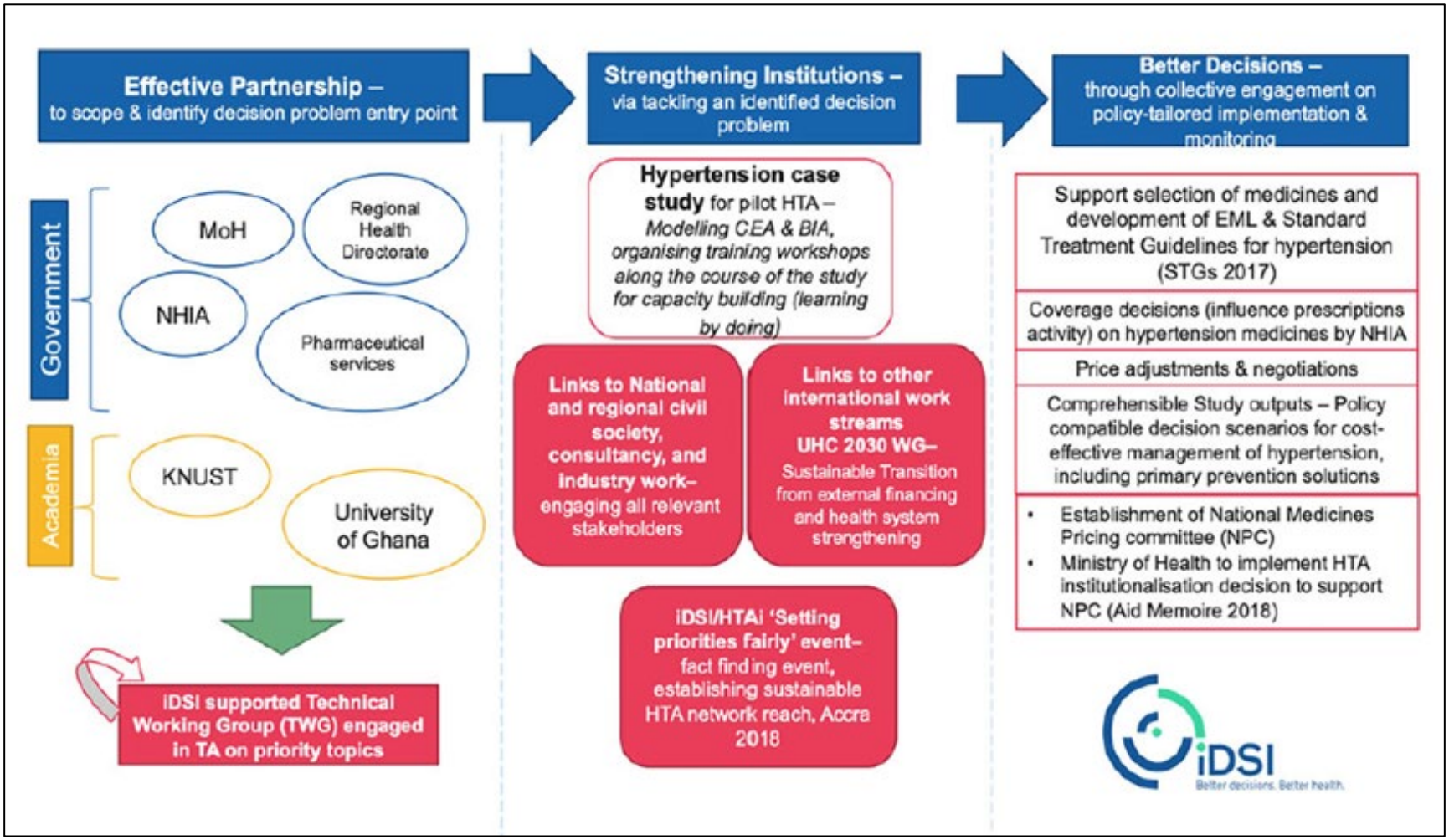


Source: Reproduced from itad Theory of Change Review Report with permission from Itad Ltd(18)

The ToC recognises that there is a complex translation process between better decisions and better health dependent on many assumptions about local factors and systems, including linkage between decisions and budgets, delivery, implementation and data accuracy. The key assumptions in this ToC are: decisions are implemented, health practitioner behaviour follows evidence and policy, beneficiaries choose to access healthcare when it is available and, lack of healthcare, or poor quality of healthcare, is a key driver for poor health outcomes(18). Only when those decisions result in implementation and practice change, can better health be achieved. Yet, we know that implementation of HTA recommendations and decisions are variable(19-23). The data for studies of the implementation of reimbursement decisions are scarce but available evidence indicates that there is a gap between decisions and implementation, resulting in an inefficient use of new drugs and a consequent loss of value(24).

Figure 1.2 shows the ToC being applied in practice by iDSI in introducing HTA to Ghana. This reflects the positioning of HTA in Ghana but stops at better decisions(9). It is this last part of the ToC regarding implementation of those decisions and their translation into health outcomes which is currently under-evaluated and under-theorised so that everything we are focusing on here is a step on from HTA dissemination. Thus, it is on this final part of the ToC that our research concentrates, to both quantify impact as well as theorise the mechanisms by which the impact of HTA can be optimised.

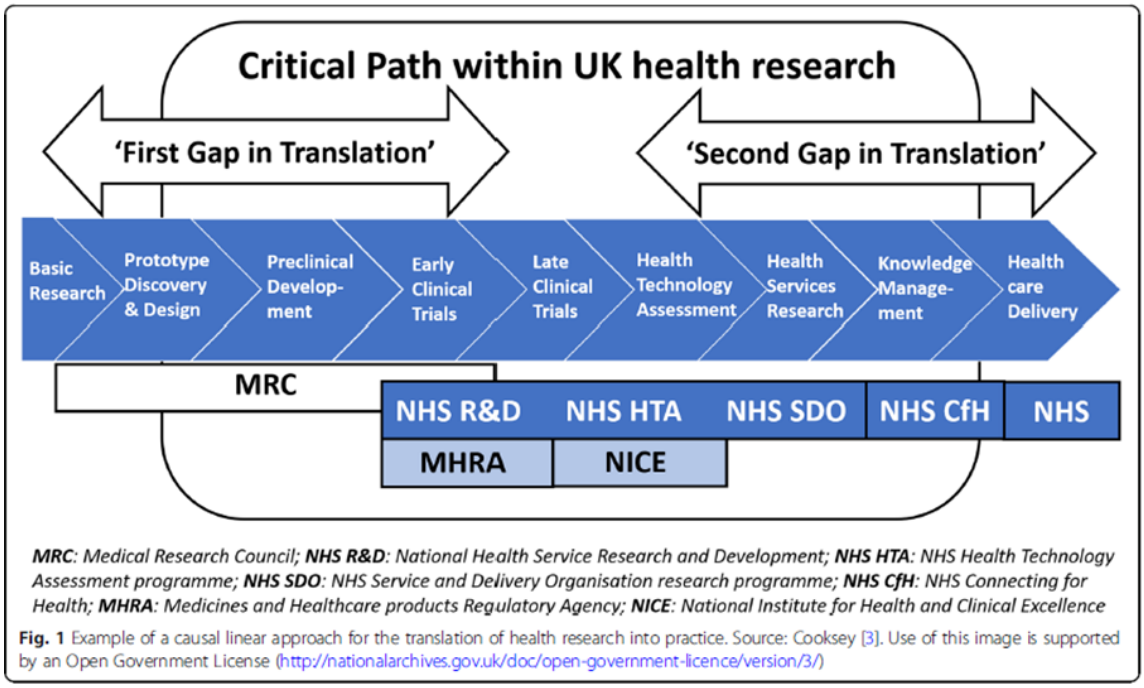
Figure 1-2 Example of ToC in practice



Source: Hollingworth et al. Implementing health technology assessment in Ghana to support universal health coverage: building relationships that focus on people, policy, and process. International Journal of Technology Assessment in Health Care. 2020;36(1):8-11(9). Reproduced with permission from Cambridge University Press and Copyright Clearance Center.

Finally, this linear schematic [Figure 1.3] potentially masks a lot of the complexity of getting evidence into practice (in a UK context) with HTA placed about half way down the chain between research at one end and healthcare delivery at the other(25). We can, though, usefully draw on this to develop a realist lens [see Chapter 3] by which to highlight the complexity in the implementation of HTA findings and to address the interplay between complexity-informed policy making - without any expectation that variance is taken out of decision-making practices by HTA. We use a realist approach to embrace the established understanding that health outcomes are complex, context-bound and produced by systems, not by singular interventions.

Figure 1-3 From HTA to health care delivery



Source: Cooksey D, 2006(26) cited in Braithwaite et al, When complexity science meets implementation science: a theoretical and empirical analysis of systems change. BMC Medicine. 2018;16(1):63(25). Use of this image is supported by the Creative Commons CC BYhttps://creativecommons.org/licenses/.

1.2 Terminology

1.2.1 Health Technology Assessment

We consider here some definitions and terminology surrounding HTA, and introduce how we are defining HTA in this thesis. We presented the WHO definition of HTA above. Until very recently, the Professional Society for Health

Economics and Outcomes Research (ISPOR) defined HTA as 'an evidence-based, multidisciplinary process intended to support healthcare decision making by assessing properties and effects of one or more new or existing health technologies in comparison with a current standard. Aiming at determining added value, HTA uses explicit analytical frameworks based on research and the scientific method in a systematic, transparent, unbiased way'(27). Typically, HTA refers to individual studies on a specific technology, and both WHO and ISPOR's original definitions reflected this by referring solely to the systematic evaluation of a health technology (or technologies). A health technology can include evidence-based interventions, practices or policies.

More recently, an international joint task group co-led and convened by the International Network of Agencies for Health Technology Assessment (INAHTA) and Health Technology Assessment international (HTAi), have proposed an updated definition of HTA which has replaced other's definitions. The new definition is as follows: 'A multidisciplinary process that uses explicit methods to determine the value of a health technology at different points in its lifecycle. The purpose is to inform decision-making in order to promote an equitable, efficient and high-quality health system'(28). Whilst this definition, similarly, refers to the evaluation of a health technology, it also brings in the value to the health system as a whole – with perhaps more potential for traction at the policy level. Indeed, increasingly HTA is used to refer to a more systematic 'process at the systems levels to inform priority setting and decision making ie as a tool for priority setting with its explicit consideration of costs and benefits'(17, 29).

This is a key distinction we make throughout this thesis. We are concerned with HTA's value to the health system, with our focus being the evaluation of HTA at the systems (ie a country or jurisdiction) level. We thus distinguish between 'HTA' (in capitals) at the systems level as a tool or process for priority setting, and 'hta' (in small letters) as an assessment and/or appraisal of an individual technology (or technologies). However, in order to get to the value of investing in HTA at the systems levels (hereafter, referred to as HTA), we need to look at what the process is delivering. In other words, we need to quantify and aggregate the value of individual htas. Additionally, whilst we recognise HTA may include several activities, for example, horizon scanning, assessment and

appraisal, pricing, we consider its overall purpose is to inform decision-making(28). We also apply the definition of HTA as 'locally relevant, fair and evidence-informed processes'(30) used to inform health benefit plans or essential medicines lists recognising that existing priority setting mechanisms, in LMICs in particular, are not always (or, at all) referred to as HTA but to which this framework, we believe, can still usefully apply.

Where we refer to the implementation of hta recommendations, we mean this to be about enacting upon those decisions (rather than the undertaking of hta as a process of itself). We may use the terms recommendations, findings and decisions interchangeably acknowledging that these are all the outcomes of an hta, be it the scientific evidence from an assessment or the actual evidence-informed decision following its appraisal. Either way, we are focusing on the step beyond decision-making and dissemination to the implementation and realisation of those decisions. We also use 'implementation' and 'uptake' interchangeably as used in the implementation science literature when referring to the adoption of evidence-based interventions(31).

Finally, others refer to HTA at different levels by categorising it into 3 types(32):

- Micro: appraisal of individual technologies (or related groups thereof)
- Meso: clinical practice guidelines to manage patient care pathways within a healthcare system
- Macro: efficiency, organisation and strengthening of the healthcare system

We see this distinction relating to the type of health technology being assessed. This terminology would, therefore, come under our usage of the term hta as a discrete application of assessment and/or appraisal of one (or several) micro, meso or macro level interventions.

1.2.2 Value

The concept of value can take many different forms - normative and subjective, and indeed, in health economics, has many different usages. It can refer to a utility theory of value representing individual preferences, value for money in terms of cost-effectiveness, the valuation of health states and outcomes, and value frameworks incorporating many different elements of value. In this thesis, we refer to value (whether it be that of a technology or hta/HTA) in terms of net health benefits (NHB) [see Chapter 4], reflecting that a health economist's concern with value will always mean looking for improvement in final (health) outcomes(33). We define the value of HTA to be achieved through increasing the uptake of net beneficial technologies and decreasing the uptake of non-net beneficial technologies. We refer to the value of an hta in terms of how its attributable NHB compared with the cost of undertaking the hta process. We also refer to the value of HTA at a systems level as an aggregate of the value of all individual htas.

1.2.3 Outcomes and impact

Realist inquiry talks about 'outcomes' whereas the objective of this research is to measure 'impact'. The former usually refers to the goal of an intervention (for example, we refer to the uptake of a technology following an hta as our outcome of interest), the latter to longer term consequences (for example, on health). It is helpful to use this conceptualisation which comes from classic logic models (and which are influential in public health) and is useful even within the realist school(34). However, the terms are often used in diametrically opposed ways in different sectors. For example, in health evaluation, 'impact' is short term and 'outcome' is long term. In international development, 'outcome' is short term and usually specific to participants in the program, and impact is both long term and (often) population wide. In realist evaluation, the term outcome is used to include all kinds of outcomes and impacts, depending on things like the scale and time period of the evaluation. (35)

In this thesis, we refer to an impact framework - the term 'impact' fitting the current rhetoric of research bodies better which are concerned with the changes and benefits attributable to research(36). We keep the usage of outcomes in

the realist synthesis though we do make a distinction between intermediate outcomes and final impact [Chapter 6, Table 6.2].

1.3 The aims and objectives of this research

1.3.1 Aim

This research aims to conceptualise, value and measure the impact of HTA at a systems level.

1.3.2 Objectives

- To develop a conceptual and methodological framework to quantify the impact of HTA.

- To explore the use of mixed methods in order to produce data-driven theory that considers individual, interpersonal, institutional and systems-level components and their interactions on the mechanisms by which HTA impact can be optimised.

1.4 Ethics

As all data are secondary, no ethics permission was required. Indeed, our framework aims to utilise routine administrative data and HTA outputs which are available in the public domain.

1.5 Thesis structure and chapter outline

The thesis is structured into 4 parts and contains 10 chapters. Each of the 4 parts is summarised below along with chapter outlines.

1.5.1 Part 1: Background and introduction

Part 1 sets out the background to this research and introduces its aims, objectives and thesis structure. The objective of this first part of the thesis is to explain how and why our HTA impact framework was developed and to put it in the context of the existing literature. It comprises chapters 1 – 3 as described below.

Chapter 1: Introduction

Chapter 1 introduces the rationale for, and background to, this research. It also outlines the thesis chapters which are structured in 4 Parts.

Chapter 2: Literature Review

Chapter 2 is a review of the literature on evaluating the impact of HTA. We are specifically interested in the impact HTA has on net health gains and its value to the health system.

Chapter 3: Philosophical and methodological issues regarding causation

Chapter 3 discusses philosophical and methodological issues regarding causality, the different ways of thinking about this, the nature of its measurement and the challenge it presents in the design of impact evaluations, especially of complex interventions such as HTA. A comparison group (a counterfactual) enables us to estimate changes in outcome that can we can attribute to an intervention, here the HTA process. A comparator is crucial to impact evaluations as without this, it can lead to erroneous measures and conclusions of impact with attribution which could be wrongly assigned or interpreted. This chapter provides a discussion of, on the one hand, the counterfactual approach to cause and effect and different means by which we might measure and apply this to HTA; and, on the other, theory driven approaches to understand what underpins outcomes leading to impact.

1.5.2 Part 2: A framework for conceptualising, measuring and valuing the impact of HTA

Part 2 presents our methodological framework for measuring the impact of HTA. This is a key output of this research. It describes both the quantitative framework and the realist theory driven evaluation. It comprises of chapters 4 - 6 as described below.

Chapter 4: NHB-ROI Framework

Chapter 4 presents our Net-Health Benefit Return on Investment (NHB-ROI) Framework for quantifying the impact HTA and how it addresses limitations

identified in Chapters 2 and 3. We believe this is important as a lack of longer-term impact assessment may undermine the importance and value of HTA. We describe our mixed methods approach to quantify the value of investing in HTA as well as to explain how outcomes and impact are achieved. This framework addresses one of the most overlooked yet, one of the most critical aspects of evaluation, namely the opportunity costs associated with the capital investment and ongoing running costs required to sustain an HTA infrastructure at the systems level(37). If we think priority setting should be better informed by evidence, and use tools such as HTA, this necessarily implies institutional change to establish a sustainable system. Consequently, we have to consider the opportunity costs associated with its establishment and ongoing running(38).

Chapter 5: Realist protocol

Chapter 5 outlines our plans for how we will review and analyse the literature as a realist synthesis to develop theory as to what underpins HTA outcomes. By outcomes, we refer here to the uptake or implementation of hta decisions. Whilst impact on NHB is our final outcome of interest, this is usually only observable over a much longer timeframe than research follow-up allows and thus, we use modelled health outcomes. We theorise contextual factors and mechanisms to provide greater insight into how 'better decisions' translate into 'health impact' through the uptake of hta findings.

Chapter 6: Realist synthesis

Chapter 6 presents the realist synthesis and its results. We draw on formal theories in the literature to help diagnosis of the implementing context and to understand readiness to implement or enact upon those decisions following an hta. We extract context-mechanism-outcomes (CMO) from the literature retrieved from our original literature search (see Chapter 2) as well as additional articles reviewed as part of the realist synthesis. We then draw lines between these CMO and our programme theory to demonstrate the links and hypotheses on capability and willingness to implement.

1.5.3 Part 3: Case studies

Part 3 presents two case studies as illustrative examples of how the framework can be put into practice. We use case study design to illustrate the value of introducing a single hta. We draw on routine administrative data and hta reports to populate our impact framework from which we estimate the NHB-ROI of these processes. It comprises of Chapters 7 – 9 as described below.

Chapter 7: China National Health Development Research Center case study

Chapter 7 draws on an existing evaluation of stroke clinical pathways (CP) as undertaken by the China National Health Development Research Center (CNHDRC) to populate our NHB-ROI framework. Uptake of the CP was based on an evaluation using routinely collected longitudinal data and their longer-term impact was modelled as part of the hta. We draw on these results to estimate realised population NHB and what we can attribute to the hta process. We offset this against the estimated costs of investing in HTA. By offsetting the attributable benefits against the costs associated with the process, we show high returns from simply introducing a single hta.

Chapter 8: Scottish Health Technologies Group case study

Chapter 8 uses the economic outputs of an existing HTA Evidence Note produced by the Scottish Health Technologies Group (SHTG) on a glucose monitoring sensor for diabetics to populate our framework. We analyse routine administrative data on its prescribing using quasi-experimental methods to establish a counterfactual and a credible measure of uptake of the technology, and what can be attributed to the hta. Again, by offsetting the attributable benefits against the costs associated with the process, we show high returns from simply introducing a single hta.

Chapter 9

Chapter 9 brings together the programme theory with the two case studies. This shows how the ROI-NHB and realist framework looks when brought together.

1.5.4 Part 4: Synthesis

Part 4 synthesises the research, discusses implications and makes recommendations. It comprises Chapter 10 as described below.

Chapter 10: Recommendations and Further Research

Chapter 10 provides a synthesis of the research to make practical recommendations for decision-makers and researchers. We outline what our framework adds to the existing body of work on HTA impact and, suggest potential next steps for future research.

1.6 Description of my role and contribution of others

As stated in my Declaration, I have opted to write throughout the thesis in the plural 'we' to recognise the contribution of others who I have formally thanked in my Acknowledgements. Here, I explicitly describe my role and the invaluable support I received from my supervisors and other colleagues as part of an iterative process, enabling me to develop the research into a coherent body of work.

UoG was sub-contracted by the CHE, University of York and as part of iDSI, during 2014-16 to develop a conceptual framework to measure the impact of HTA. Based on this work, a larger grant to further develop this research was awarded directly to HEHTA, UoG from 2016 -18. This grant supported my time to undertake this research and forming part of my PhD (enrolled 2016 – 2020).

I independently undertook the background literature review (Chapter 2) and continued to update this throughout the duration of the research. I am very grateful for the guidance and input from UoG MVLS librarian, Paul Cannon, on search terms and strategy.

Initial work was co-developed by CHE, York and HEHTA, UoG. Conveying the concepts of potential and realised population NHB, dependent upon the extent to which clinical practice is changed ie the VOImp, was the starting point to this work. At HTAsiaLink conference 2015, CHE/HEHTA jointly presented on this work, albeit at an early stage of the framework. There, I used a case study from

HITAP to illustrate these concepts empirically, expressing impact in terms of full and current implementation and equating this to potential and realised population NHB. I am indebted to Profs Mark Sculpher and Paul Revill, CHE for their invaluable contribution of coming up with these initial concepts which have underpinned the research from the outset, in particular, Chapter 4.

Explaining the gap between expected and actual gain in health was the next step. As iDSI had developed a theory-based approach to their monitoring and evaluation, it seemed appropriate to be able to tie into this work by theorising the mechanisms by which impact could be optimised (as well as quantifying impact). Having independently reviewed theory-based approaches to evaluation, I opted to draw upon realist inquiry to highlight the complexity in the implementation of HTA findings. Prof Hannah Hesselgreaves, who subsequently joined HEHTA, had expertise in applying this approach in evaluations. I developed the protocol and carried out the realist synthesis (Chapters 5 and 6) under the supervision of Hannah, drawing on her extensive expertise in this area. Again, I acknowledge Paul Cannon's help in developing search terms.

The VOImp was further developed at a later stage with the support of my main supervisor, Prof Andy Briggs, with the addition of drawing in a counterfactual to obtain a measure of attributable NHBs (Figures 4-2 and 4-3). I independently explored methods by how we might practically undertake such an analysis, drawing on the most appropriate quasi-experimental methods given what routine health data may be available. I acknowledge my colleague, Janet Bouttell, for helpfully and kindly allowing me to adapt her summary review of methods in this field (Table 3-5). Given the different paradigms to causality which I had drawn upon to develop the framework (counterfactual and theory-based), I brought these together in Chapter 3 to illustrate the challenge of the counterfactual in the design of impact evaluations of complex interventions.

In terms of aggregating from an individual hta to impact at the systems level in order to quantify a NHB-ROI (Equation 4-5 and Figure 4-4), I am indebted to Prof Briggs for drawing this out. I independently undertook Chapters 7 and 8 in applying the quantitative framework to two case studies as illustrative of the proof-of-method. I am immensely grateful to CNHDRC and SHTG for their

generosity in allowing me to use their technology assessments. This also enabled me to test the theory in Chapter 9 and to integrate the mixed methods. In Chapter 10, I synthesised the research to make practical recommendations for decision-makers and researchers. This last chapter makes reference to a book chapter co-authored by myself and colleagues in HEHTA, Janet Bouttell and Prof Neil Hawkins.

Finally, I am indebted to all those who have provided constructive feedback on drafts throughout, including my supervisors, Olivia Wu, Hannah Hesselgreaves and Andy Briggs and many iDSI colleagues. Below is a summary of these contributions by chapter.

Chapter 1
Chapter 2 – with acknowledgements to Paul Cannon
Chapter 3 – with acknowledgements to Janet Bouttell
Chapter 4 – informed by early work with York and co-developed with Andy Briggs
Chapter 5 – under supervision of Hannah Hesselgreaves/ librarian input
Chapter 6 – under supervision of Hannah Hesselgreaves
Chapter 7 – with thanks to CNHDRC
Chapter 8 – with thanks to SHTG
Chapter 9
Chapter 10 – reference to Bouttell, Grieve, Hawkins book chapter

2 Impact of HTA: a literature review

2.1 Introduction

The benefits concerning the link between HTA and impact in terms of net health improvements have rarely been quantified; 'the literature on assessment of HTA influence is still quite limited and there is little on longer term effects on clinical practice and health outcomes'(39). By net health improvements, we mean the difference between health gained by an intervention and health that could have been gained if the money required to deliver it had been spent on other interventions. Capturing the impact of HTA on clinical practice and health is difficult given the potential requirements in terms of data. Even in countries where HTA is well established, evidence which identifies such impact in terms of net health gains is limited. It is especially challenging in LMIC where routine administrative monitoring and evaluation data are often lacking, leading to poor data capture and reporting. This has led some to claim that 'there is currently insufficient evidence that the use of priority setting tools improves health outcomes and reverses existing inequities...we have ample evidence that the lack of a rational and transparent process generates inequity and stagnation in mortality levels'(40).

Whilst much research has been undertaken on establishing what factors influence improved decision-making(41-44), how those influences on decision-making interact with local context and health systems, leading to impact on clinical practice and health, is less understood. For countries with greater capacity constraints, how decision-making interacts with 'context' leading to impact on health is even less explored and arguably, of critical importance(45). Straus et al(46), as cited in(47), state 'a review of the existing literature on HTA reveals a startling lack of depth, particularly on the impact HTA has had on health-care budgets, efficiency and on societal health outcomes….whereas the previous 10 years have been well-spent on building the HTA/evidence-based medicine infrastructure and evidence base, the next 10 should focus on the outcomes'.

2.2 Search terms and strategy

A literature review on the evaluation of HTA was undertaken, focusing on impact in terms of health gains, and the methods and frameworks used. The search terms [Table 2.1] were run on Ovid MEDLINE(R) without Revisions 1996 to May Week 4 2020, and Embase 1996 to 2020 Week 22, last updated 1 June 2020.

Table 2-1 Search terms

1	exp Technology Assessment, Biomedical/
2	((health technolog$ adj3 assessment$) or HTA).mp.
3	1 or 2
4	(("evaluat$" or "method$" or "framework$" or "model$" or "concept$" or "empirical research" or "theor$" or "cost benefit analys$") adj5 (health technolog$ or HTA)).tw.
5	(("impact$" or "value" or "gain$" or "benefit$" or "influence$" or "return$ on investment" or "ROI" or "social adj3 return$" or "economic adj3 return$") adj5 (health technolog$ or HTA)).tw.
6	4 or 5
7	3 and 6

Inclusion criteria: framework, model or evaluation of HTA; implementation or uptake of a technology following an hta.

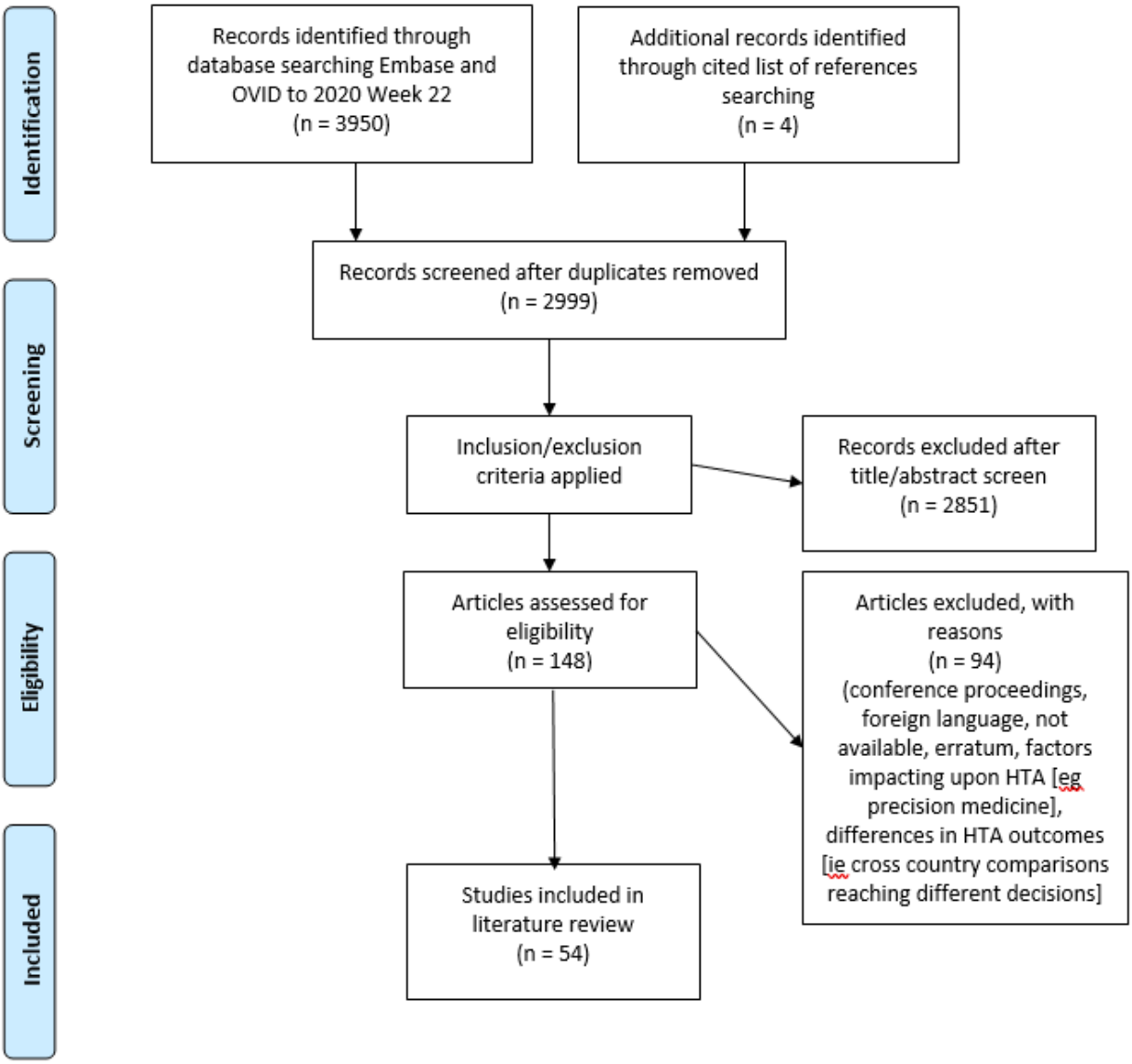
Exclusion criteria: methodology on measuring outcomes and/or value in economic evaluations; uptake of a technology without going through an hta; foreign language.

2.3 Results

2.3.1 Articles retrieved

The Prisma flow diagram is shown below [Figure 2.1]. Annex A lists the final articles included.

Figure 2-1 Prisma flow diagram



2.3.2 HTA impact frameworks

Several HTA evaluation frameworks were identified in the literature. We describe those acknowledging the need to capture impact on health in chronological order below(19, 48-54).

Jacob & McGregor framework

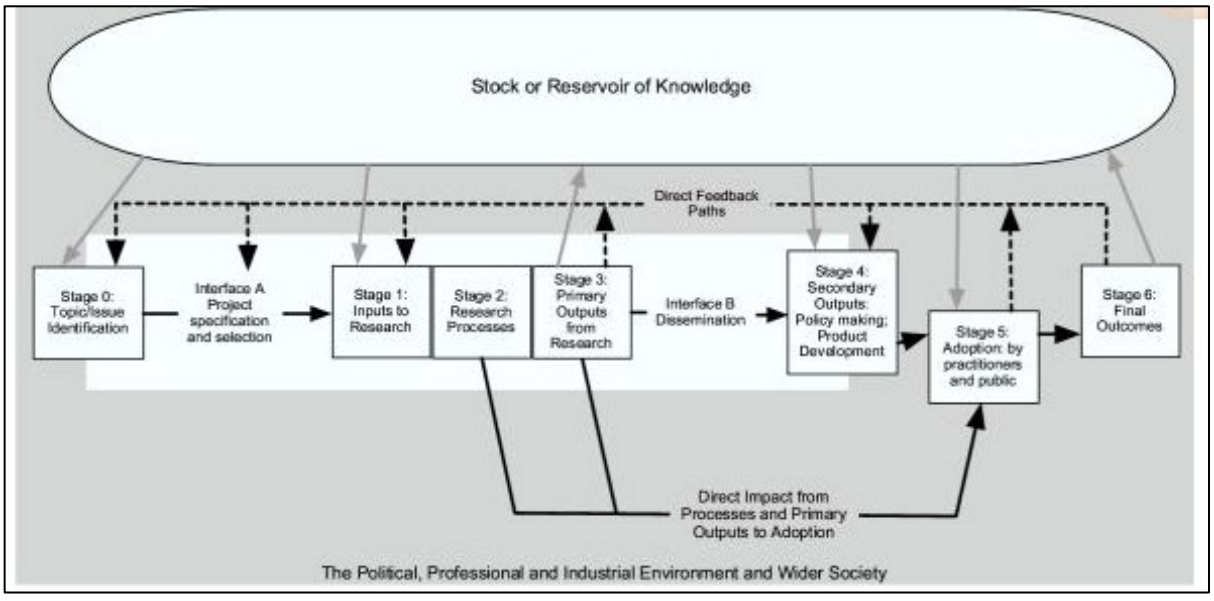
Jacob and McGregor, in one of the earliest HTA impact frameworks(49), define the impact of htas so as to 'influence the diffusion and use of health technology in such a way as to increase the efficiency of the health care system (by increasing its effectiveness or reducing its costs)'. Using interviews,

questionnaires and data banks, they estimated the impact of 21 hta reports on policies, technology diffusion and utilisation. They found cost-savings into the millions (Canadian dollars) through a systematic documentation of its effects.

Payback framework

Buxton and Hanney's Payback Framework(50), previously commissioned to assess the UK's National Institute for Health Research (NIHR) HTA programme's first ten years (1993–2003)(55), is recognised as the most dominant methodological framework used to assess the impact of healthcare research(56). This framework categorises benefits of health research ranging from traditional knowledge production, and research training and targeting, to impacts on policy, product development, health and economic gains [Figure 2.2]. Measurable impacts primarily focus on knowledge generation, perceived policy impact and, to some extent, on practice – with the conclusion that impact on knowledge generation is more easily quantified than that on policy, behaviour or especially, health gain. They note that, in spite of the dedication of HTA agencies to evaluation, the impact of their technology assessments on health policy is rarely assessed.

Figure 2-2 Payback model



Source: Reprinted by permission from Springer Nature and Copyright Clearance Center. Springer Nature, Health Research Policy and Systems, Proposed methods for reviewing the outcomes of health research: the impact of funding by the UK's 'Arthritis Research Campaign', Hanney S, SPRINGER NATURE LICENSE, 2004 cited in Oortwijn et al (57).

Davies et al

Davies et al(51) used decision analytic modelling to conduct ex ante assessments of the relative value for money of htas in terms of their expected net costs and health benefits to society in order to assist in their prioritisation for assessment [Figures 2.3 – 2.5]. It is assumed that the principal objective of HTA is to provide information and evidence to influence health care practice and improve the efficiency of health care provision. The perspective of the model includes consideration of the costs and benefits to the research funding body, the providers of health and social care services and, the patients who are likely to receive the health care interventions(51). They acknowledge that priorities could be set according to a range of criteria but focus on a quantitative measure of 'value' or 'payback', noting that most previous analyses were ex post assessments. Broader costs and benefits to society (value of knowledge, political and administrative benefits, skills etc) were excluded partly because they are difficult to assess but also as they would accrue irrespective of the technology in question. Figure 2.3 illustrates the processes through which HTA can change the efficiency of health care provision. Figure 2.4 illustrates the range of factors which may modify the impact of HTA on the provision of health care. Figure 2.5 illustrates the conceptual structure of the prioritisation model. Utilisation rates were assumed to increase (decrease) with the addition of evidence on the effectiveness (ineffectiveness) of a technology from the hta.

Figure 2-3 Davies et al – Process of HTA impact on health care provision

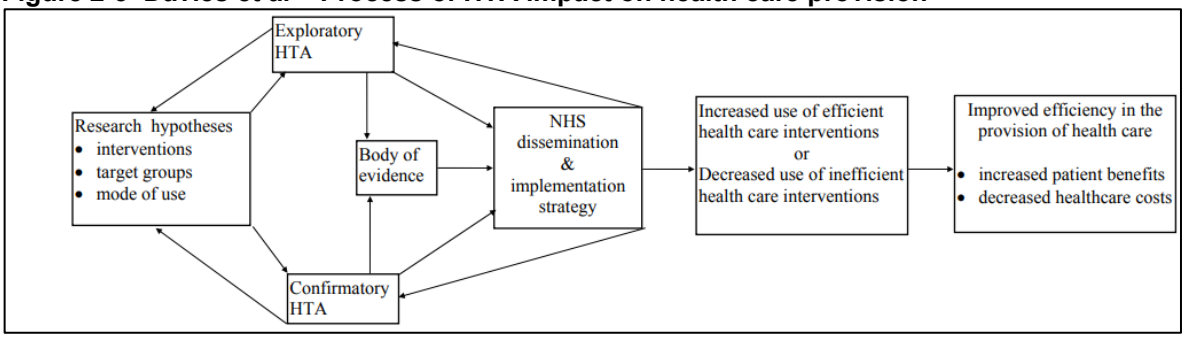


Figure 2-4 Davies et al - Adoption and utilisation of HTA results

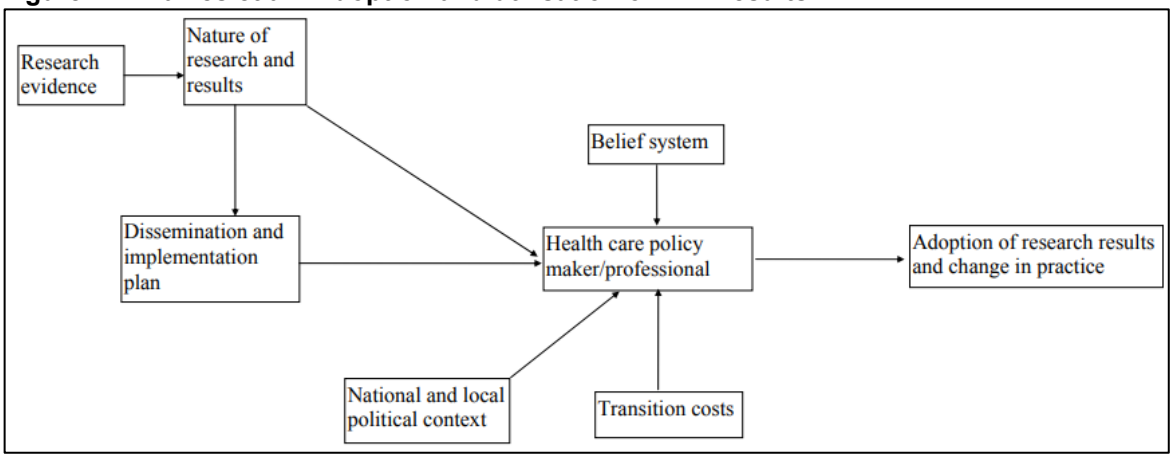
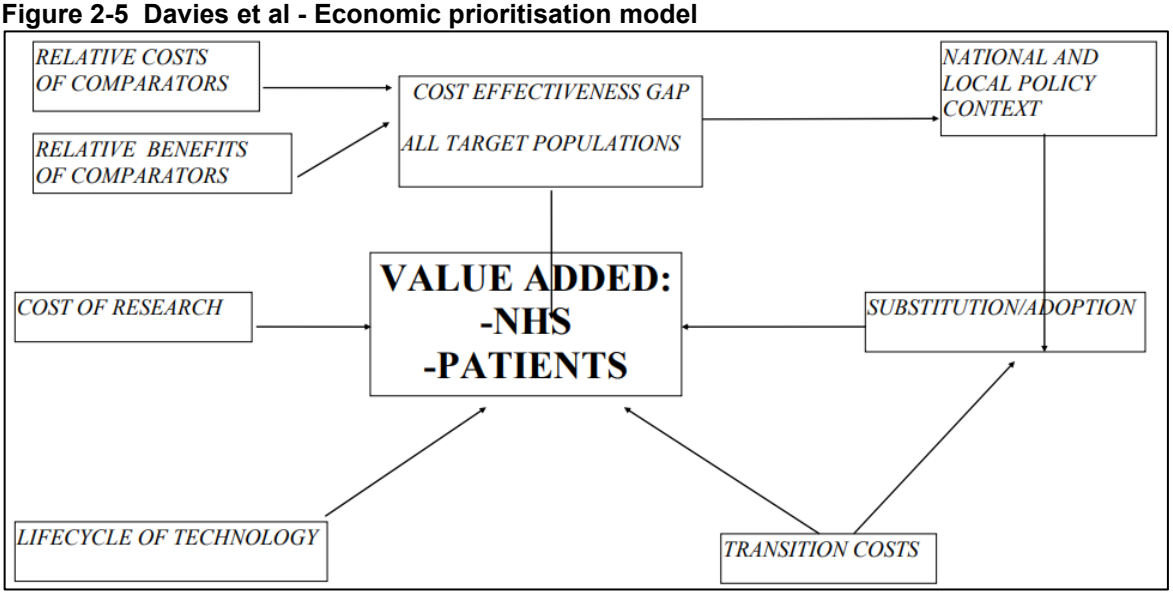


Figure 2-5 Davies et al - Economic prioritisation model



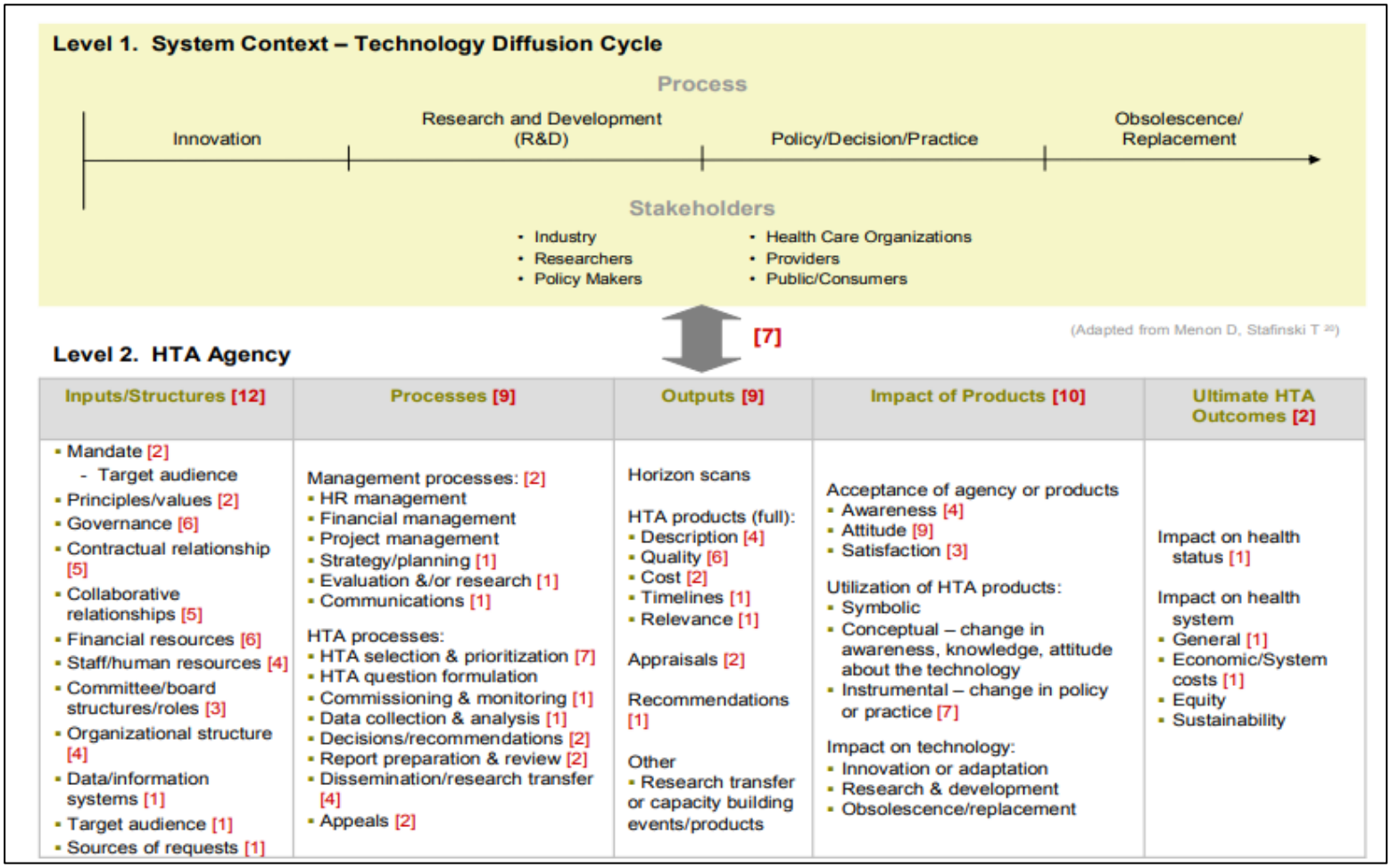
Source: Figs 2.3 – 2.5 Reproduced from Davies et al. Prioritizing investments in health technology assessment. Can we assess potential value for money? Article in International Journal of Technology Assessment in Health Care · February 2000(51) with permission from Cambridge University Press and Copyright Clearance Center

Wanke et al

Wanke et al(53) propose a generic evaluation framework for HTA agencies. They conducted a review of agencies to understand what aspects of HTA have been evaluated, and how. They took a logic model approach to their conceptual model which was informed by Stufflebeam's classification of 22 evaluation approaches(58), Hailey's exploration of HTA effectiveness(59), existing literature at that time on the impact of HTA(49, 60-63) and a survey to members of the International Network of Agencies for Health Technology Assessment (INAHTA)(54). They also identified four older HTA evaluations noting their limitations of accounting for attribution and upstream outcomes(49, 64-66).

Wanke et al's framework [Figure 2.6](53) depicts the range of dimensions conceived for an evaluation of an HTA agency. Level 1 depicts the context in which HTA agencies operate. In Level 2, the HTA which operates within the Level 1 technology diffusion process, is depicted as a logic model with inputs/structures, processes, outputs, outcomes and ultimate impact. The items listed under these headings represent the items upon which an evaluation may be focused. The authors state that only when an HTA agency has matured and is stable, should an impact evaluation be undertaken with a view to determining the programme's merit or worth. Most of the reviews or evaluations included in their review (14 out of 16) were identified by a direct request to INAHTA members.

Figure 2-6 Wanke et al framework



Source: [Figure 4 Dimensions evaluated] in Margaret Wanke, Don Juzwishin, Richard Thornley, Liza Chan HTA Initiative #16. An Exploratory Review of Evaluations of Health Technology Assessment Agencies. February 2006(53). Permission sought from the Alberta Heritage Foundation for Medical Research (AHFMR) 17 June 2020 by email.

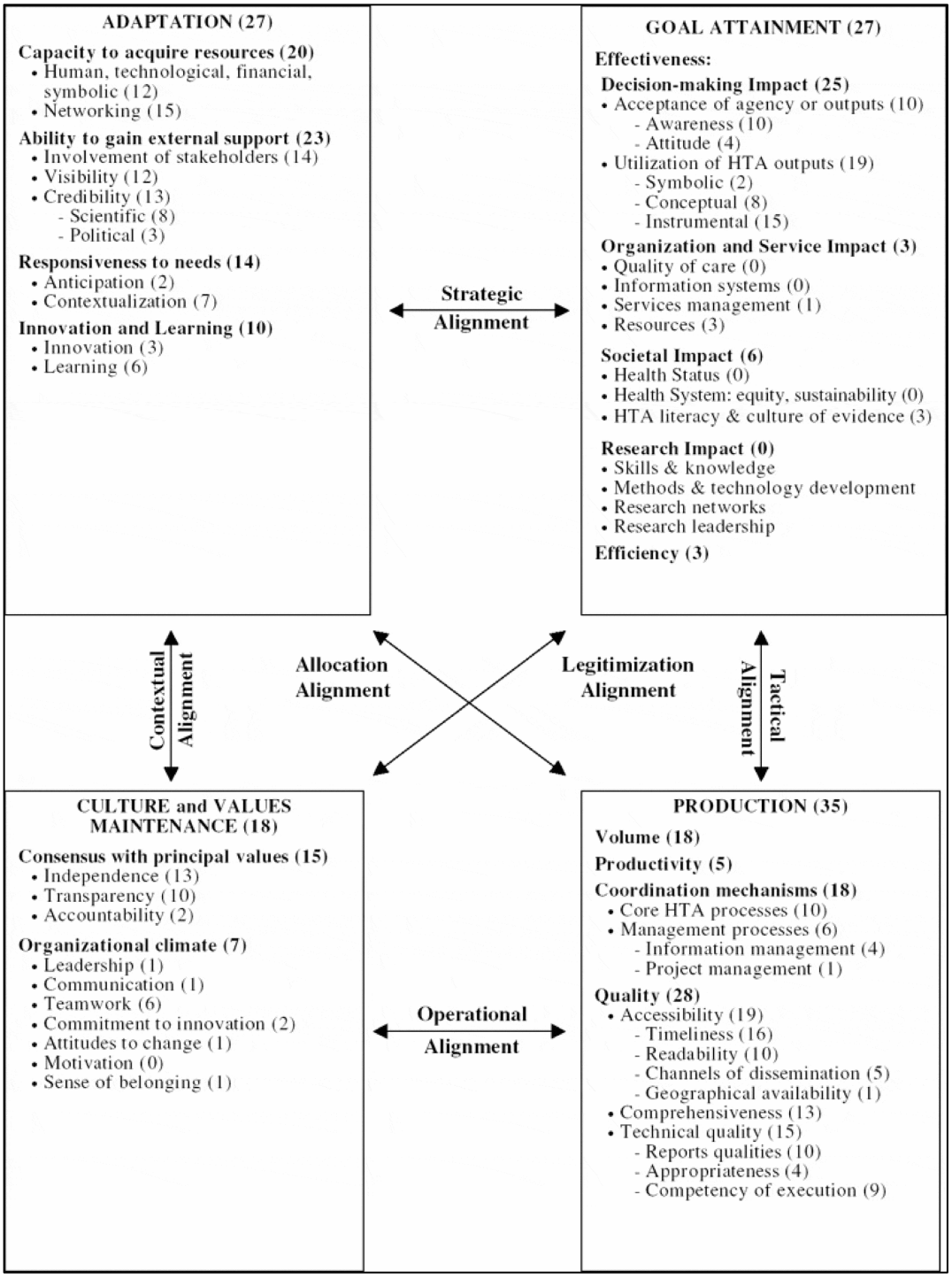
Lafortune L et al.

Lafortune et al(52) were concerned with an HTA agency's performance. Based on Parson's social action theory(67), Lafortune et al propose a conceptual model that includes four functions an organisation needs to balance to perform well: goal attainment, production, adaption to environment and, culture and values maintenance [Figure 2.7]. They identified specific dimensions under each of these functions in relation to performance, and compared this against existing evaluations of HTA agencies thus enabling any gaps to be highlighted. The goal attainment function includes the dimension of 'impact' on decision-making, organisation and service, society, research and efficiency. Lafortune et al's framework captures health outcomes under 'societal impact'.

RAND Europe

RAND Europe evaluated the impact of UK's NIHR HTA Programme 2003-2013(68), with a separate report assessing the potential economic returns(19). The NIHR HTA Programme funds research about the clinical and cost effectiveness, and broader impact of healthcare technologies for those who plan, provide or receive care in the UK's National Health Service (NHS)(69). The RAND Europe evaluation intended to capture the savings to the NHS by estimating the health benefits to patients and converting these into monetary terms. They examined the impact of the HTA programme to understand its potential economic benefits, taking 10 HTA-funded projects that had shown a new treatment could offer benefits if introduced in the NHS, either by saving costs or by improving health through better treatment. They included some additional short illustrative case studies providing more context.

Figure 2-7 Lafortune et al framework



Source: Reproduced from Lafortune L et al. Assessing the performance of health technology assessment organizations: a framework. Int J Technol Assess Health Care. 2008;24(1):76-86(52) with permission from Cambridge University Press and Copyright Clearance Center.

INAHTA impact framework

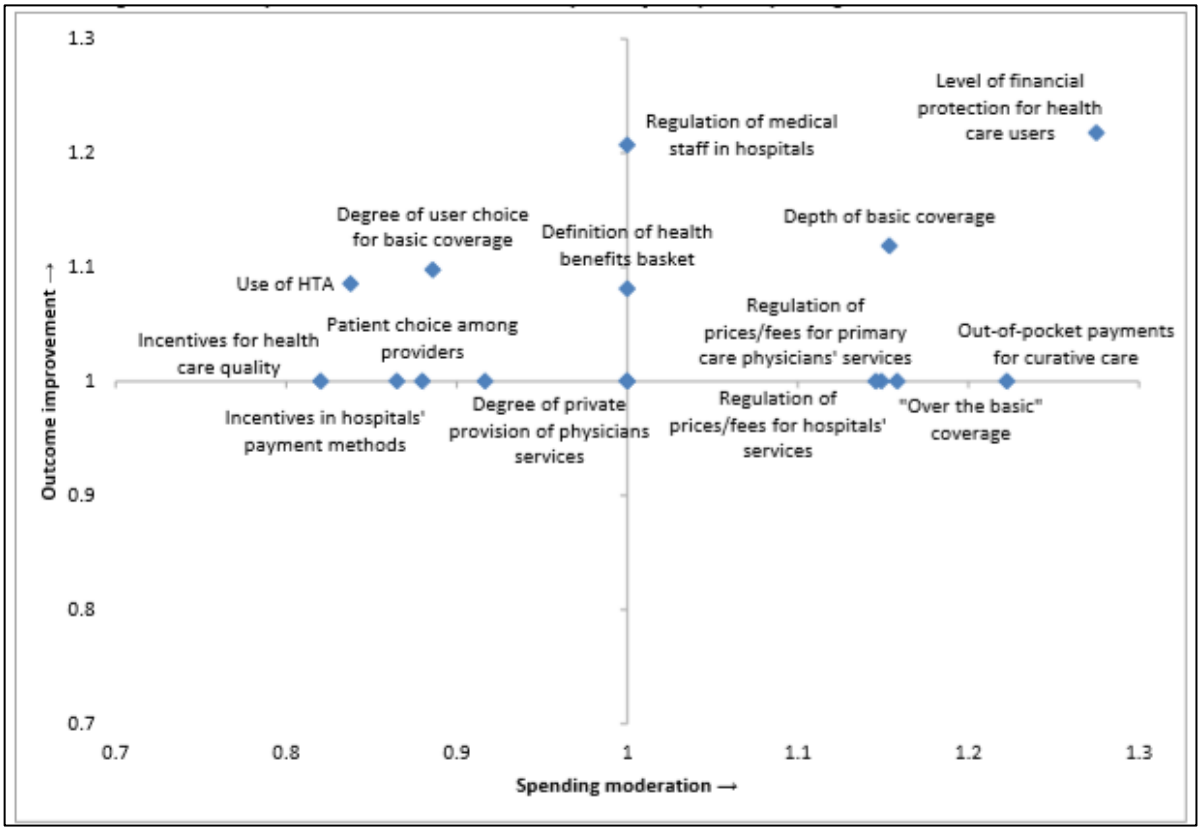
INAHTA has developed a framework for reporting information on htas for which there are some indication of impact on decisions by government at the regional, national or international level. Positive, interim and negative indications of

impact are all to be included. This is a form that is completed by the HTA agency requiring boxes to be ticked on an hta's indications of impact, the agency's and external opinion on the level of impact (from no apparent influence to major influence) along with free text to provide further information and evidence. Indications of impact range from the hta being considered by a decision maker to linked to clinical practice change or changes in health status. The form is available on the INAHTA website(54).

OECD

The OECD aimed to capture the impact of HTA at a country or systems level [Figure 2.8] by measuring its impact on life expectancy and public spending(48). Using a quantitative approach, the OECD investigated how several different policies and institutions helped achieve value for money in healthcare. For HTA, a score was computed based on countries' responses to questions on their structure and capacity for HTA, whether cost-effectiveness and affordability were taken into account, and how the information was used. Regression-based analysis was used to assess the effects of policies and institutions on health spending and life expectancy. They found that HTA's use in providing evidence related to new technologies was likely to both magnify life expectancy gains but also that of spending on health care. Results for HTA showed a statistically significant improvement in health outcomes but also a statistically significant increase in public spending. Given the necessary existence of institutional structures and capacity for technology assessment and its purpose in encouraging the additional use of effective services, HTA's scores were associated with both higher public expenditures and higher care quality as suggested by the positive and significant coefficients(48). This was contrasted with policies which increased life expectancy but also moderated health spending, such as those aimed at increasing the scope of goods and services covered by primary health care. The cost drivers of HTA were mainly related to 'the additional use of effective services, as well as cost increases related to the up-front investment to create, expand and operate health technology assessment agencies'(48). As an example, empirical estimates showed that NICE's recommendations for the adoption of new technologies have cost the NHS an additional GBP 1.65 billion per year(70).

Figure 2-8 OECD Effect of policies on life expectancy & public spending



Source: Reproduced from Lorenzoni, L. et al. (2018), "Which policies increase value for money in health care?", OECD Health Working Papers, No. 104, OECD Publishing, Paris. http://dx.doi.org/10.1787/a46c5b1f-en OECD Fig 2. (48) with permission from OECD as source and copyright owner.

2.3.3 Individual country level HTA impact evaluations

Articles retrieved on countries which have undertaken an impact evaluation of HTA at country level included Austria(71), Catalonia(72), Canada(49), Denmark(73), Iran(74), Malaysia(75), Poland(76), Taiwan(77), the Netherlands(57), Thailand (abstract only retrieved). Each country HTA evaluation is summarised below [Table 2.2].

Table 2-2 Country level HTA impact evaluations

Country and Years of analysis	What was evaluated	Methods	Limitations
Austria 2001 - 2010	A conceptual model developed for assessing the impact of HTA in the Austrian healthcare system(71) used multidimensional aspects of impact, namely, awareness, acceptance, policy process, policy decisions, clinical practice, outcomes and enlightenment.	It was based on indicators developed by Gerhardus et al(78) and Weiss' theory of research utilisation(79).	This model did not address health outcomes due to methodological limitations. The authors make the recommendation that further research should address the methodology on how to improve impact measuring, in particular the relationship between HTA and the overall improvement of health / health care systems.
Catalonia 2000 - 2003	The current and past situation of HTA in Catalonia(72, 80).	An historical review of facts and landmarks; semi-structured interviews.	'We miss the opportunity to measure the impact of our recommendations…By missing this impact analysis, we all miss the opportunity to convince our policy

			makers of the need for HTA and its goodness'.
Canada 63 hta's from 2001 onwards	Impact of various CETS reports. McGill Technology Assessment Unit (TAU) – local hta – covered 63 assessments.	Case study approach to assess the impact on policy. Attribution of HTA to costs to the healthcare system dealt with by sensitivity analysis on upper and lower limits. Impact was measured by the number of recommendations accepted and the dollars spent on new technologies versus those saved through the adoption of cost saving recommendations.	Methods dependent on judgement of the analyst, policy impact not documented, time interval between HTA and its effect, the challenge of what would have happened without the HTA noting that no amount of input from key actors can totally remove the uncertainty regarding this estimate. It is estimated the TAU saved the hospital an average of CAD$1.14 million annually(81). Did not take actual implementation into account.

Denmark Overview from 1997	Individual hta's(73)	Qualitative; individuals' experiences of HTA projects. Descriptive	Experience still too limited to evaluate benefits. 'Above all, the implementation of HTA results will be one of the greatest challenges in the coming years'. (82).
Iran 2007 - 2017	The impact of 23 individual htas, identifying the determinant factors leading to the implementation of report results(74).	A case study approach employing questionnaires and semi-structured interviews with stakeholders informed by the Payback framework.	No mention of health outcomes, only impact on capacity, decision making and knowledge.
The Netherlands(57) 2000 - 2003	The impact of individual HTA projects over 3 years	Payback framework	Deemed too early to determine whether the HTA programme led to actual changes in healthcare policy and practice.
Malaysia (75)	10 years of HTA	IHAHTA framework	Level of influence of reports and decisions are qualitative only.

1997 - 2016			
Poland (76) 2005 - 2015	10 years of HTA	Descriptive analysis of the drug reimbursement criteria before and after HTA.	Descriptive only.
Taiwan(77) 2007 - 2017	10-year implementation of HTA	Descriptive analysis of HTA in Taiwan.	Descriptive only.

2.3.4 Facilitators and barriers to the uptake and impact of HTA

Many articles were found identifying facilitators and barriers[2] in implementing the findings of an 'hta' and how this actually changes practice(83). Drummond summarised the main elements to successful implementation to be: 'a) defining a clear policy question; b) defining a clear research question; c) making recommendations commensurate with the evidence; d) identifying the implementation mechanism; e) paying attention to incentives and disincentives; and f) clarifying the roles and responsibilities of the various parties'(84). Recommendations have been made as to how these factors can then be used to improve impact, notably that key stakeholders (patients, providers and industry) are adequately involved; decision-makers give a prior commitment to use assessment reports (and assessments meet their needs); the necessary resources are available for implementing decisions; there is transparency in the assessment and decision-making processes; and collaboration, knowledge and skills are transferred across jurisdictions(42). Evaluations have been undertaken to establish specifically what HTA agencies need to do to have greater impact in the future(41). Findings were similar to the facilitator and barriers identified in that there is a need to better integrate local practitioners into the HTA process; for greater influence at the policy-making level, assessment and appraisal to be incorporated within a common structure to provide necessary incentives for policy makers to consider the science; and finally, that the HTA agencies would need to be handed greater regulatory powers(85).

Knowledge translation (KT) has emerged as a paradigm to address closing the "know-do" gap [see Figure 1.3]. Facilitators and barriers to knowledge translation in the context of hta as cited by Fournier et al(86) include 'knowledge management (timeliness and relevance, personal contact between researchers and policy makers, inclusion of opinion leaders or knowledge brokers in research planning), organisational and structural characteristics (facilities, resources, financial incentives), as well as personal and professional characteristics (skills, attitudes, experiences, tacit knowledge, clinical judgment)'(25, 86, 87).

[2] 'Facilitators and barriers' came from realist search strategy too – see Annex B

Finally, Straus et al(88) identified over 250 implementation barriers. At the socio-political level, factors relating to the organisation and financing of the health system were found to affect the utilisation of HTA recommendations; at the healthcare organisation level, existing collaborations between the hospital and the HTA agency favoured the integration of recommendations into practice. Formalism in the organisation also influenced the utilisation of HTA recommendations; at the professional level, the high degree of autonomy of specialists, the importance of peers and collegial control, and the definition of professional roles and responsibilities influenced physicians' willingness to integrate HTA recommendations into their practice(80). Unlike other descriptive papers, Cheung et al(89) evaluated the relative importance of each barrier and facilitator. Grouping by themes showed that ‘organisations and resources' and 'policy characteristics' were the most important, with the recommendation made of motivation too to encourage positive attitudes toward its use. The less important groups were 'policy-maker characteristics', 'research and researcher characteristics' and 'contact and collaboration'. A summary of some of the more frequently cited facilitators and barriers in the literature are listed below [Table 2.3].

Table 2-3 Facilitators and barriers to the implementation of hta findings

Barriers	**Facilitators**
Costs of implementation, resourcing and inflexibility of budgets and resources (22, 90-93)	Learning through collaboration and exchange of experience; skills, ability (94); formal links between producers and users (42); Stakeholders are involved and support decisions (61, 95-98) with front line engagement, champions (22, 80, 91)
Uncertainty, weak governance, political constraints (90), leadership (91)	Timeliness (61, 96, 99, 100)
Poor quality of communication and dissemination (61, 91); too few mechanisms to inform; user's (mis)understanding(22, 99)	Accuracy and validity(61, 94, 99); Evidence comes from a trusted source(93)
	Relevance (22, 94-96, 99, 100)
Other Factors Design of health system (90, 93) and broader organisational context (80, 94, 96, 101), including health information systems (102), payment mechanisms (90) and alignment of incentives and support from the top (61, 91) – as well as the interaction at the human level. Knowledge translation (96, 102).	

2.4 Discussion

There has long been a call for the evaluation of HTA organisations(103) yet, relatively little has been published. As Garrido et al(78) state 'the ultimate value of HTA in a health system depends on its contribution to improved health status or increased efficiency rather than to increased knowledge. In this respect, HTA does not differ much from other health technologies and must be subject to the same rigorous standards of evaluation'. Yet, much more was found on reviewing methods for measuring the impact of health research(104). We discuss the findings from the literature review as presented above in terms of HTA impact frameworks, country level evaluations and, facilitators and barriers to implementation leading to impact.

2.4.1 HTA impact frameworks

Each of the frameworks identified stated that they were developed to improve the accountability of public funds and to be able to evaluate the contribution, performance or impact of HTA bodies. Jacob et al state that '(H)owever excellent an HTA may be, if it fails to influence the working of the health care system, it is without impact and must be considered without value (p. 69)'(49). On the other hand, Lavis(105) suggests that moving beyond impact on decision-making to health, economic and social outcomes is best left to a focused evaluation of that specific intervention or policy: 'Moving beyond decision-making outcomes to health, economic and social outcomes, however, is almost certainly asking too much. Research organisations simply want to know whether the research knowledge that they produce is having an impact on decision-making. Tracing the complex pathways though which informed decisions translate into improved implementation or performance and ultimately into better health is best left to stand-alone research initiatives. The same can be said of economic and social outcomes'.

There were a number of limitations in the frameworks found, reflecting the complex nature of the process of HTA as an intervention itself. There was little quantitative data and, from the impact evaluation studies retrieved and the findings of others, most evaluations to-date have made use of only qualitative

methods(71). Wanke et al found the majority of responses they received from INAHTA members were qualitative in nature(53).

The issues of a counterfactual and attribution were also lacking. Wanke et al stated that attribution to HTA can only be partial regarding the acceptance, utilisation and impact of products; and that 'Because the factors influencing population health status and the heath system are numerous, outcomes at these levels cannot, as a general rule, be directly attributed to HTA agencies or programs'(53). Similar to our proposed framework [Chapter 4], RAND Europe used economic analysis to compare potential benefits with the costs, using illustrative case studies. In carrying out their evaluation, RAND Europe were required to make some assumptions, in particular, that the findings were fully implemented across the NHS and that the impact could be wholly attributed to the HTA programme(19). RAND themselves state that many (mostly conservative) assumptions but one big (not conservative) assumption was made with limited formative value[3]. Nevertheless, this gave a 'headline' number in terms of ROI. Assuming 100% implementation and full attribution of the outcomes to the HTA programme, they concluded that 12% of the calculated potential net benefit would cover the total cost of the HTA Programme from 1993 to 2012. Linking this from research for NICE to its recommendations, NICE have carried out audits of the implementation of its guidance(106) but the use of simple before-and-after measures, as applied there and by others too(71), does not provide a credible measure of uptake.

Measuring impact on health outcomes was lacking. Described as 'ultimate HTA outcomes' by Wanke et al(53), this looks beyond the impact of individual HTA outputs and explores the impact of HTA agencies' mandate on the health status of the population or the health system in general(52). The recommendations of a NIHR systematic review(107) support the continued use of the Payback Framework as proposed by Buxton and Hanney to measure the impact of health research(50, 55, 108). The payback approach has been identified too as a key framework for measuring HTA impact(71). The authors of the payback approach found that impact on health was the more challenging to quantify. Furthermore, no existing evaluations of HTA organisations' impact on 'health

[3] Presenter S Guthrie, RAND. ISPOR 20th Annual European Congress, Second Plenary Session

status' or 'health systems' were found by Lafortune et al(52), and there was a dearth of literature retrieved under the category of impact on health status found by Wanke et al(53) [Figures 2.6 and 2.7].

Following on from their impact framework, INAHTA has published a recent report (109) on the practices of HTA impact assessment in INAHTA member agencies with the aim of identifying strategies by which to support HTA agencies to improve their practice. They define HTA impact assessment to be an evaluation of the uptake and the effects of an hta report. They also distinguish between the impact of an HTA report and the impact of the agency. All agencies were invited to participate in semi-structured interviews, with 26 out of 47 taking part. Most reported that they had informal approaches to impact evaluation. Indicators of impact identified by participants could be grouped into five categories related to the report, the agency, the decision maker/policy, health system and outside the health system. Methods of impact assessment included analysis of administrative data, interviews, surveys, audits and documents analysis. It was found that linking impacts that exist 'downstream' from the HTA, such as change in clinical practice can be challenging. For this reason, most agencies chose to measure impact on factors associated with the report and decision-maker – with some stating that their role is not to assess the impacts further downstream in the health system but rather to produce high quality and useful HTAs. The second part of the study explored factors that inhibit or enable HTA impact activities. The state that few methods and strategies have been developed for assessing HTA impact that could assist HTA agencies in understanding and measuring the uptake and influence of their work. Challenges in establishing a causal relationship between an hta and its impacts was noted, as was a reluctance by healthcare providers to open up about their uptake (or not) of the recommendations making data collection difficult. In this same vein, Loblova(110) et al present the intrinsic value of HTA to be its focus on evidence and transparency - despite the absence of empirical evidence on its effects. Indeed, others recommend that UHC will only be achieved through a transparent and participatory process and that states need to institutionalise priority setting with such bodies being accountable to their populations, government and the judiciary(111).

The OECD state that '(A)lthough there is scarce empirical evidence on the impact of HTA processes on efficiency and cost-containment in the health sector as a whole, HTA may be expected to generate overall efficiency gains primarily through its use for an evidence-based definition of the package of interventions publicly funded (and their specific reimbursement levels) within the health system. On the other hand, the creation and operation of an HTA body adds another layer of administrative costs to the health sector and may increase expenditure related to the additional use of effective interventions'(48). Budget constraints and a lack of funding was identified as a key reason for not investing in HTA according to a survey undertaken by the WHO in 2015(16). Loblova et al(110) also highlight the upfront costs of establishing HTA systems which will potentially discourage LMICs from investing in these processes and agencies.

In terms of our proposed framework [Chapter 4], Davies' HTA model(51) most closely resembles this in that it similarly defines impact in terms of improved efficiency through the increased use of efficient health care interventions and decreased use of inefficient healthcare interventions which, in turn, is achieved through this facilitating change in practice. A counterfactual is implicit only. Furthermore, Wanke et al(53) speak to our realist inquiry [Chapters 5 and 6] when they state that 'because socio-political factors play such an important role with respect to HTA, evaluators of HTA agencies, particularly of impact, will need to consider contextual factors influencing the operation and success of the HTA agency/program'.

2.4.2 Individual country level HTA impact evaluations

The approach used by most countries was predominantly the payback method, employing qualitative and descriptive analysis using documents, interviews and questionnaires(105). To-date, most evaluations of HTA have focused on outputs (number of HTA reports, for example) with methodological challenges being cited by some for not evaluating health outcomes(71). A conceptual model developed for assessing the impact of HTA in the Austrian healthcare system(71) used multi-dimensional aspects of impact, namely, awareness, acceptance, policy process, policy decisions, clinical practice, outcomes and enlightenment based on indicators developed by Gerhardus et al(78) and Weiss' theory of research utilisation(79). This model did not address health outcomes due to

methodological limitations, and the authors make the recommendation that further research should address the methodology on how to improve impact measuring, in particular, the relationship between HTA and the overall improvement of health (care systems).

Denmark, yet to carry out a full impact evaluation, stated that the implementation of HTA results will be one of the greatest challenges of the years ahead(73). When agencies are funded publicly, some evaluation of impact is obviously desirable even if, in practice, such evaluations are often extremely difficult to carry out. However, 'failure to make any attempt to assess the impact of HTAs is clearly inconsistent with the concept that decisions should be evidence-based'(49). HTA agencies do recognise the need to think about implementation and impact as illustrated by the INAHTA report(109) and, in particular, CADTH, Canada in its strategic plan 2018-2021 reported: 'There is significant variability in the uptake of health technology assessment recommendations at the policy level compared with clinical practice, and gaps in the resources and capacity required to make better use of evidence…we will engage in a forensic analysis on why sound evidence is ignored or contradicted by policy and in practice'. Indeed, 'the limited number of studies following change in clinical practice and health outcomes indicates that these areas need much more attention in the future. Quality registers and clinical databases are growing rapidly around the world and they could be very useful tools for analysing the influence of HTAs'(112).

2.4.3 Facilitators and barriers to the uptake and impact of HTA

Only when HTA findings and decisions result in implementation and practice change, can better health be achieved. Yet, we know that implementation of HTA findings and decisions are variable(19-23). We come back to facilitators and barriers in the realist synthesis which employs the use of theory to offer greater explanatory power with regard to uptake. The importance of studying theories underlying different approaches to implementing guidelines and changing practice is highlighted by Grol(113) who groups theories in terms of educational, epidemiological, behavioural, social interaction, organisational and coercive approaches(113).

Supply-side constraints relate to the health system, costs of implementation, system interactions, uncertainty, weak governance and politics - as well as demand-side responses by patients. The implementation of recommendations depends on the type of technology, how funding is tied, professional engagement, quality of HTA reports and dissemination and implementation strategies, or lack thereof (41, 42, 84, 90, 114-116). Facilitating factors to the implementation of HTA findings set out by Haan and Rutten(117) *cited in*(118), are divided into regulation by directive and regulation by incentive. Government policy, administrative arrangements and organisational structure can be described as contributing to the context in which decision-making is carried out(39). For the translation of decisions into action, the decision-maker will need to have access to effective policy or administrative instruments, with the will and power to use these. Such supporting structures may include provider payment, regulation, education, patient empowerment and effective leadership. This interaction between context and mechanism is explored further in our realist synthesis [Chapters 5 and 6].

2.5 Conclusion

This review has established a lack of existing evidence and methodology on how best to measure, value and quantify the impact of HTA. The frameworks retrieved illustrate both the dearth of literature on health impact and limitations with existing methodology to measure this. We found some examples of evaluations of HTA agencies in a range of countries and, whilst useful and encouraging in their own right, are mostly only descriptive in nature. None consider what would have happened anyway ie without the HTA process. A comparison group (a counterfactual) enables us to estimate changes in outcome that can we can attribute to an intervention, here the HTA process. A comparator is crucial to impact evaluations as without this, it can lead to erroneous measures and conclusions of impact with attribution which could be wrongly assigned or interpreted. Similarly, none consider the opportunity costs of investing in HTA or attempt to show whether resources are used efficiently. Financial benefit is generally presented as cost savings. This is a narrow interpretation of the role of HTA which aims to ensure an efficient use of resources as distinct from being simply a cost-cutting exercise.

We found many articles listing facilitators and barriers to the uptake of HTA decisions. Only when HTA decisions or recommendations are implemented, does this translate into impact on health. Rather than duplicate this work, we re-analyse this literature using a realist perspective to produce greater explanation as to how and why hta decisions and recommendations are implemented [Chapters 5 and 6].

2.6 Next steps

Given this lack of evidence and methodology in the literature, we have developed a mixed-methods framework to quantify the value of HTA and consider the opportunity costs of establishing HTA processes but which are often overlooked. The focus of our framework is to go beyond 'intermediate' outcomes as to how informed decisions translate into improved implementation, ultimately leading to impact in terms of health gains. Our impact framework and, how it adds to the literature discussed above, is presented in Part 2.

First, we discuss the nature of causality. Impact evaluations address what the difference is between what happened with an intervention, and what would have happened without it. The lack of establishing this in evaluating the impact of HTA is a methodological challenge that does not yet appear to have been fully addressed. We discuss philosophical and methodological issues regarding causality, the different ways of thinking about this, the nature of its measurement and the challenge it presents in the design of impact evaluations, especially of complex interventions such as HTA.

3 Methodological and philosophical issues regarding causation

3.1 Introduction

'The narrow goal of evaluative research is to identify the causal impact of an intervention on outcomes of interest. The broader goal is to understand the mechanisms underlying this impact'(119). An impact evaluation provides information about the impacts produced by an intervention - positive and negative, intended and unintended, direct and indirect. This means that an impact evaluation must establish what has been the cause of observed changes referred to as causal attribution. Without this, evaluations risk making false or assumed attributions that may produce incorrect findings(120). In complex interventions and contexts, the attribution challenge can also limit the outcomes chosen for evaluation, with the evaluation stopping at knowledge or practice outcomes, thus missing opportunities to measure patient or health outcomes – as was apparent in the literature review on HTA impact [Chapter 2]. In this chapter, we describe how we address the challenge of causality and the counterfactual which are fundamental to understanding and measuring impact.

Drummond et al(121) discuss the conceptual and methodological challenges associated with benchmarking HTA and highlight that the key question is whether HTA has improved healthcare provision. However, they recognise that assessing the improvement in healthcare provision is challenging because of the difficulty in specifying the counterfactual(121). Whilst one study found HTA to be cost effective in terms of controlling a rise in health care spending and achieving better health outcomes in life expectancy and mortality in HTA versus non-HTA countries, correlation versus causality could only be considered(122). A study on cancer drugs looked at whether the addition of a complementary HTA and cost-effectiveness evaluation increases the value realised by the drug in practice compared to how it would be used without these assessments(24). This is one of the few studies found addressing the counterfactual in HTA.

We consider first the ontology (nature of reality) and epistemology (nature of knowledge) of different paradigms to 'causation' and their implications for 'the counterfactual'. 'What we consider to be 'real', what we can know about it,

how we might go about valuing it, and how we think programmes or policies cause change – are fundamental questions for evaluators.' (123) Different philosophies make different assumptions about how things are caused. Gates and Dyson(124) state that evaluators ought to consider the relevance of different ways of addressing causality and lay out five ways of thinking about it: '1) a successionist framework that underlies regularity and counterfactual logics; 2) narrative stakeholder accounts; 3) generative accounts of processes and mechanisms; 4) causal packages and contributory accounts; and 5) nonlinear, multidirectional and dynamical accounts of relations as found in complex systems'. We will look at these in more detail, in particular, successionist and generative causation.

Secondly, given we cannot assign 'treatment' – in this case, HTA (or, indeed, an hta) - randomly and compare with 'no treatment/no HTA' in order to provide internally valid estimates of effectiveness, a counterfactual needs to be created. We discuss potential ways to construct a counterfactual drawing on quasi-experimental methods.

Finally, we discuss the use of theory driven evaluation to understand the mechanisms underlying or driving this impact. We commit to the idea that the reality of health systems is complex (different perspectives exist in systems, outcomes are produced by many factors interacting together in an ever-changing way, and the systems themselves are outside the control of any one actor in it). This is our underpinning for a theory-driven approach, and it this translational gap between impact and what happens in practice that has led to calls for the greater use of explicit theory in research. Here, we employ realist inquiry - a complementary approach to the iDSI Theory of Change [see Chapter 1] - to build theory in order to better understand what it is about context that promotes or inhibits uptake and adherence to HTA decisions (as it is only when those decisions are implemented, do we have any impact on health outcomes). In so doing, we aim to offer a generalisable framework that could be applied across different settings.

3.2 Ontology and epistemology

Ontology deals with the nature of reality and is associated with the key question of whether social entities should be perceived as objective or subjective. Accordingly, objectivism (or positivism) and subjectivism can be specified as two important aspects of ontology. Epistemology is the theory and logical analysis of knowledge. Table 3.1 lists perspectives on knowledge which reflect different views of reality(125).

Table 3-1 Perspectives on knowledge

Wholly objective, established by deduction	Rationalism
External to individual, objective and independent of social actors. It is discovered or verified by empirical means or through experience.	Empiricism/ positivism
Objective, exists independently of human thoughts and beliefs or knowledge or their existence (realist) but is interpreted through social conditioning (critical realist).	Realism
Subjective, reflecting personal experiences and values.	Interpretivism
Constructed and reconstructed by the individual	Constructivism

Adapted from 'Access learning in health profession MED5392' (Week 1), Univ. of Glasgow.

The above table implies five schools of thought regarding causation. This includes the positivist tradition which rejects unobservables; Popper's Falsification where theories have falsifiability or refutability if there is the inherent possibility that they can be proven false; reductionists who view multiple causes at different levels down to the micro mediation level; and critical realists (CRs) who assume that causal relationships exist outside of the human mind and that these valid causal relationships cannot be perceived with total accuracy by our imperfect sensory and intellective capacities. CRs also assume causes that we can do something about are the most useful even when the ultimate micro level is not known.

Similarly, as stated above, Gates and Dyson lay out ways of thinking about causality(124) —successionist, narrative, generative, causal package, and complex system. Whilst Table 3-2 presents these as distinct approaches, the authors note elements and assumptions of each are often mixed in methodological approaches.

Table 3-2 Ways of Thinking about Causality

Causal view	Logic of causal argument	Question(s)
• Successionist/regularity: frequency of observation of simultaneous occurrence of independent, single cause and effect	• Simultaneously observe two separate events and (1) show cause temporally prior to effect, (2) a statistical relationship (covariation) between cause and effect, (3) cause is both necessary (cause always present when effect is) and sufficient (effect always present when cause is), (4) rule out other plausible causes, and (5) demonstrate association in high number of cases	• What effects are statistically significantly associated with this intervention?
• Successionist/counterfactual: compare two almost identical cases only differing in cause (the intervention)	• Show that effect follows from the intervention through comparison to a highly similar control group	• Does the intervention work to produce intended effects? • Can we attribute effects to the intervention?
• Narrative: Stakeholders' views on how an intervention has influenced/affected/made a difference in their lives	• Ask participants directly how an intervention influenced their lives, collect evidence verifying observance of these outcomes, and rule out alternative explanations	• According to stakeholders, what influence, effects, and/or difference did the intervention make for their lives?
• Generative: Theory-based explanation of how causal process happens by showing how mechanisms work to generate outcome patterns given contextual factors	• Claim causation by identifying mechanisms that connect two events, empirically verifying theorized causal relations and rejecting alternative explanations	• What works, how, for whom, and under what circumstances? • How and why does the intervention work?
• Causal package: copresence of multiple causes which may or may not be necessary and/ or sufficient for an effect	• Identify a package of multiple causes that work together to produce an effect • Describe the cause as necessary but not sufficient within a causal package that is sufficient • Distinguish ground preparing, triggering, and sustaining contributory causes	• Is it likely that intervention has made a difference? • How does the intervention work in combination with other interventions or factors to make a difference?
• Complex systems: Multiple, interdependent causal factors and nonlinear feedback processes affect the structure and behavior of a system or situation over time	• Build a conceptual model, called a causal loop diagram, of the causal relationships at work in a situation, intervention, or system • Verify this model with empirical evidence for each variable, mathematical formulas, and computer simulation	• How do multiple causal factors and feedback processes affect change in this intervention or situation? • What's working now and how?

Source: Reproduced from Gates E, Dyson L. Implications of the Changing Conversation About Causality for Evaluators. American Journal of Evaluation. 2016;38(1):29-46(124) with permission from SAGE Publications.

We discuss here empiricism and realism as two of these different paradigms we adopt in our impact framework.

Empiricists are aligned to the natural sciences and are generally *positivists* who believe knowledge is external to the individual and can only be discovered by, and verified by, empirical means. The premise of 'causation via succession' is held which is based on the assumption that causation can be determined from high volume correlation of observables(126). Successionist causation is about the description of outcome patterns ie *whether* something has an effect, not *why*(127). Successionist causation is about variables and their association; it is variables that do the explanatory work.

Realists also believe in the existence of an independent, external world but acknowledge that it may not be possible to verify every aspect of that world (125). Their critique of positivism is that this cannot portray the mechanisms and contexts that generate social behaviour. The explanatory analytical process in realist evaluation is based on generative causation which is the assumption that underpinning hidden (in the sense of being below the threshold of observability and empirical measurement) mechanisms are responsible for the manifestation of observable outcomes. Realists see the quest to achieve control in randomised trials is 'to squeeze out of the picture precisely those mechanism and contexts that are required in understanding whether a programme works'(128). As generativism is 'designed to utilise mechanisms and contexts to explain outcome patterns, it provides the most complete approach to causal explanation. Because explanation trades in peoples' choices and societal constraint, it calls on the full repertoire of social research to provide supportive empirical data'(128).

Mahoney et al(129) summarise key aspects of 'causation' for these two paradigms, identifying three fundamental differences that are related to paradigmatic differences in the treatment of causation between post-positivist and realist logic: (1) the construct of mechanism, (2) the relation between mediators and moderators on one hand and, mechanisms and contexts on the other hand, and (3) the variable-oriented approach to analysis of causation versus the configurational approach. Table 3-3 outlines the distinction made between correlation analysis and causal analysis. While the former 'involves

identifying antecedents regularly conjoined with outcomes, causal analysis consists of identifying the mechanism that underlies and generates empirical regularities and outcomes'(130). Indeed, generative analysis aims to explain why correlations exist.

The 'issue of causality in evaluation is complicated by the lack of agreement in philosophy of science about the nature of causality and broader disagreements in the social sciences about how causal claims ought to be warranted'(124). It is recommended that evaluators be 'literate in multiple ways of thinking about causality, be familiar with a range of causal designs and methods [Table 3-4] and, layer theories to explain causality at multiple levels'(124). We look at firstly, quasi-experimental methods and secondly, theory-based approaches in more detail as used in our impact framework.

Table 3-3 Successionist and generative definitions of mechanisms

	Definition #1—"Variables" (successionist mode)	Definition #2—"Theory of change" (successionist mode)	Definition #3—Scientific realism (generative mode)
Definition	An intervening (set of) variable(s) that explain(s) why a correlation exists between an independent and dependent variable	Frequently occurring causal patterns that are triggered under generally unknown conditions and with indeterminate consequences. A mechanism explains by opening up the black box and showing the cogs and wheels of the internal machinery. It provides a continuous and contiguous chain of causal or intentional links between the *explanans* and the *explanandum*	An unobserved entity that, when activated, generates an outcome of interest
Analytical approach	Correlational analysis techniques, such as mediation analysis, are used to identify "mechanisms" that are considered to be mediators of the observed effect	While slightly more broadly defined, this definition is compatible with probabilistic approaches to analysis	Causal analysis consists of identifying the configuration that links the outcome to mechanism(s) triggered by the context, often combining quantitative and qualitative data
Role given to theory	Theories in the form of universal laws can be deduced from empirical research (covering law principle)	Theories in the form of empirical knowledge derived from constant conjunction observations	Research contributes to developing theories of the middle range
Implications	Risk of reduction of mechanisms to measurable indicators, through which dynamic processes of change are reduced to correlations between variables that stand for more complex processes	In this view, and similar to definition 1, causation is reduced to the concatenation of elements in a causal chain. Causation is demonstrated to the degree that empirical regularities can identified	Empirical research allows investigation of a possible mechanism, thus identifying a plausible mechanism and may eventually lead to the identification of the actual mechanism. Research thus contributes to increasing the plausibility of the explanatory hypothesis

Source: Adapted from Mahoney J. Beyond Correlational Analysis: Recent Innovations in Theory and Method. Sociological Forum. 2001;16(3):575-93(130) cited in Van Belle S, Wong G, Westhorp G, Pearson M, Emmel N, Manzano A, et al. Can "realist" randomised controlled trials be genuinely realist? Trials. 2016;17(1):313(129). Use of this image is supported by the Creative Commons CC BY https://creativecommons.org/licenses/.

Table 3-4 Range of Causal Designs and Methodologies

Design approaches	Examples of methodologies	Basis for making causal claims	When/why to use it?
Experimental	Randomized control trial Natural experiments	Comparison to a counterfactual	To generate precise information about whether a particular intervention worked in a particular setting When there is a discrete intervention When a control group and large samples are available and feasible
Quasi-experimental	Propensity score matching Judgmental matching Regression discontinuity Interrupted time series	Comparison to a counterfactual	When precise information about the intervention is needed as with experimental approaches, but there is no random control group When we want to know the effects of particular variables in a large sample When a large sample is available
Theory-based approaches	Realist evaluation Process tracing Contribution analysis Impact pathways analysis	Analysis of causal processes or mechanisms in context	When there is a strong theory of change When it's important to understand how context affects an intervention When it's important to understand how and for whom an intervention works
Participatory approaches	Success case method Most significant change Outcome mapping	Validation by participants that their actions and experienced effects are "caused" by the intervention	To capture multiple, experiential understandings of change and possibly identify unintended consequences For internal needs of an organization (e.g., program improvement) Feasible, timely, and affordable When the sample size is small to medium
Case-based approaches	Within case: analytic induction, network analysis, and process tracing Across case: qualitative comparison case analysis	Analysis of causal processes within a case Presence of causal factors across multiple cases	To identify causal factors within or across multiple cases when known effect(s) have been identified
Systems-based approaches	Causal loop diagramming System dynamics	Build a conceptual model, called a causal loop diagram, of the causal relationships at work in a situation, intervention, or system Verify this model with empirical evidence for each variable, mathematical formulas, and computer simulation	Examine multiple, interdependent causal factors and nonlinear feedback processes that affect the structure and behavior of a situation or system over time To understand a system's dynamical behavior over time To identify unintended, nonlinear and emergent effects

Source: Reproduced from Gates E, Dyson L. Implications of the Changing Conversation About Causality for Evaluators. American Journal of Evaluation. 2016;38(1):29-46(124) with permission from SAGE Publications.

3.3 Positivism

Successionist or 'positivism' has been the dominant framework for thinking about causality and is underpinned by the logics of regularity and the counterfactual(124). We describe positivism in relation to the 'counterfactual' and the use of quasi-experimental methods to construct this.

3.3.1 Causation by counterfactual

The core principle of experimental/quasi-experimental methods is that there is a 'counterfactual' ie observed outcomes from a 'control' group that did not receive the intervention which can be compared to outcomes from the intervention group.(131) Regularity purports there is a statistical and temporal relationship between X and Y; and the counterfactual requires making a comparison between two highly similar situations to estimate what would have happened in the absence of the intervention.

However, a research design such as a randomised controlled trial (RCT) is not suited to the evaluation of many complex or policy interventions, including HTA, as the assignment of individuals to 'HTA' or 'no HTA' is beyond our control, and it is consequently, a non-randomised evaluation. Alternatively, 'the simplest approach is to choose a control area similar to the area where the intervention is being introduced, and compare outcomes in the two areas. The key difficulty is finding a control area sufficiently similar to allow outcome differences to be attributed to the intervention. Statistical methods of adjustment for differences between the areas are limited by the availability of data, and often cannot fully account for all the relevant differences'(132).

3.3.2 Quasi-experimental methods

Quasi-experimental methods are used in such scenarios above to mimic this counterfactual. There are different approaches as to how to do this dependent upon the data available. Individual level data allow for methods such as Regression Discontinuity Design (RDD) to be used where the intervention is allocated on a cut-off point in a continuous variable; or the use of Instrumental Variables (IV) where a variable is found which predicts treatment allocation but is not correlated with outcomes. Similarly, matching treatment and controls on

selected individual characteristics or scoring them on a combination of variables using propensity scores is applicable here. A simple before-and-after comparison is one of the weakest forms of design in terms of providing a credible measure of effect. Interrupted time series (ITS) which applies regression methods to capture time trends, provides a more credible measure. The most basic ITS design requires one experimental group and multiple observations before- and after a treatment. The outcome before the intervention is used as a 'counterfactual' ie there is an assumption that without the intervention the outcome would have remained the same or followed a consistent trend(133, 134).

To overcome lower internal validity associated with such quasi-experimental methods, the use of multiple designs are usually recommended to strengthen the credibility of results(134). As ITS makes it difficult to account for observed and unobserved confounding factors, it is common to add time series data from a non-equivalent comparison group over the same period thus creating a comparative ITS (CITS) design(134). The simplest CITS analysis entails a difference-in-difference estimate where the difference between the pre- and post-intervention means in the comparison group is used as the counterfactual against which the mean difference in the treatment group is evaluated. As such, difference-in-difference additionally brings in an untreated control group and compares the temporal changes in the treated group in the intervention area to those of the control/untreated group over the same period in the control area. This method cannot, however, take account of area-specific trends, ie changes other than those attributable to the intervention that occur in one or other of the areas(132). The key assumption for a difference-in-difference design is that outcomes in both groups would follow parallel trajectories over time in the absence of an intervention.

Where this assumption does not hold, a synthetic controls approach can instead be used. The synthetic control method attempts to overcome this problem by comparing the trend in the outcome of interest in the intervention area with the trend in a synthetic composite area. This uses information on pre-intervention levels on the outcomes of interest and predictors of those levels in other 'unaffected' areas, for example, other states or districts where the intervention

is not being implemented (which form the ‘donor pool’) and which are then used to derive the synthetic control(132). The control is the weighted average of all the areas in the ‘donor pool’ that best mimics the pre-intervention trend in the intervention area. The effect of the intervention is then estimated as the difference between the post-intervention trends in the intervention area and the synthetic control area(132). Thus, in more complex CITS analyses, the means and slopes of the pre-treatment values are used to assess not only changes in mean levels but also changes in trend, in the variation around these trends or in the pattern of temporal variability(134). History can also be examined and the construct validity of the effect enhanced by collecting time series data for some dependent variables that should be affected by a treatment and for others that should not(133).

These quasi-experimental methods are summarised in Table 3-5 along with examples of their use in evaluations in healthcare. We draw on ITS to provide a more credible measure of uptake of a technology following an hta [Chapter 4]. Alternatively, we could simply model scenarios, along with a distribution, to represent a counterfactual level of uptake.

Table 3-5 Quasi-experimental approaches to dealing with the counterfactual

Approach	Method of dealing with confounding	Limitations	Example of use of quasi-exp. method in health
Interrupted time series *Uncontrolled pre- and post-approach* – the outcome before the intervention is used as a 'counterfactual' ie there is an assumption that without the intervention the outcome would have remained the same or followed a consistent trend. *Controlled pre- and post-approach* – the evaluation includes looking at the before and after intervention outcomes for another area (the 'control area') where the intervention did not take place.	Both measured and unmeasured confounding are addressed providing the main characteristics of the group studied remain broadly stable. The similarity in characteristics between the control area and the intervention area should minimise bias.	The changes before and after the implementation of an intervention may be due to the underlying trend rather than the effect of the implementation. Over the medium and long term, the characteristics of the study group may change. It is often difficult to identify a unit that is sufficiently similar to the treated unit.	Sheldon Trevor A, Cullum Nicky, Dawson Diane, Lankshear Annette, Lowson Karin, Watt Ian et al. *What's the evidence that NICE guidance has been implemented? Results from a national evaluation using time series analysis, audit of patients' notes, and interviews BMJ 2004; 329 :999* (135). NICE use ITS in one of its ERNIE Database reports #40 [see Annex C]

Difference in difference (DiD) approach – this method is a version of the before and after approach with comparator where control group is included in the statistical analysis. *Synthetic controls methodology* – a counterfactual is synthesised using a formal statistical approach from a pool of potential controls using characteristics thought to be associated with the outcome.	Provided that the outcome trends are parallel in the pre-intervention period in the intervention and control groups, the analysis will eliminate unmeasured confounding that does not vary over time. The weighting of controls in the synthesised control is determined by a requirement to closely match the pre-treatment outcomes in the intervention area. This weighting is applied to outcome data over time. Unmeasured confounding can be minimised even if it varies over time.	The approach suffers from the same limitations identified above with regard to similarity of characteristics. The parallel trends assumption is often not met. It is important that the characteristics of the units selected as potential controls should be similar to the treated unit.	Consider constructing synthetic controls especially in countries where HTA recommendations are at a sub national / state level, for example, China. Lepine A et al. Free primary care in Zambia: an impact evaluation using a pooled synthetic control method. 2014. Abadie, A et al. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. 2010 (136, 137).

Source: Adapted from Bouttell J, Craig P. University of Glasgow(132, 138).

3.4 Theory-based evaluation

Theory-based impact evaluations draw conclusions about an intervention's impact through rigorous testing of whether the causal chains thought to bring about change are supported by sufficiently strong evidence and that alternative explanations can be ruled out(131) Theory-based evaluation is explicitly concerned with both the extent of the change and why change occurs; 'it tries to get inside the black-box of what happens between inputs and outcomes (and outcomes and impacts) and how that is affected by wider contexts'(131). The theory here can be thought of as 'a set of assumptions about how an intervention achieves its goals and under what conditions'(124). It can derive from 'formal, research-based theory or an unstated, tacit understanding about how things work'(139). Theory-based evaluation makes causal claims by considering 'causal mechanisms' and the contextual circumstances in which these mechanisms operate(124). As macro-social structures or systems are not amenable to being examined by experimental methods, researchers have adapted a 'theory of change' perspective developed by the Aspen Institute in the USA(140) and Pawson and Tilley's ideas of 'realistic evaluation'(141). This new conceptualisation of programme theory introduced by Pawson and Tilley(139) has become increasingly used in the evaluation of complex interventions in health. These are described below as two examples of theory-based approaches, and upon which our impact framework draws.

3.4.1 Theory of Change

Theories of change (ToC) is about the identification and confirmation of causal processes – an explanatory pathway of change(139). Developed by the Aspen Institute(140), it is process orientated as it follows the pathway of a programme from its initiation through causal implementation links (to explain how and why the desired change is expected to come about) until intended outcomes are reached. The ToC is developed through collaborative stakeholder engagement. It requires measurement along the pathway of all that must be achieved before the long-term outcome.

3.4.2 Realism

Realism makes a particular set of assumptions about programmes and the nature of reality, causality and evidence. Firstly, realists hold that interventions are 'theories in action' and work through introducing new ideas or resources into existing social systems – and in so doing, change the conditions that shape decision making and behaviour. Secondly, regarding causality, outcomes are caused by mechanisms – they are the choices, actions, reasoning that people make as a result of the resources a programme provides (for example, expertise/information). Mechanisms are triggered when programme resources interact with context (institutional, organisation, individual). Thirdly, realism is methodologically pluralist, and any relevant evidence can used to test theories(128).

Influential thinkers and critical realists who have influenced the realist approach include R. Bhaskar (1975, 1978), M. Archer (1995) and A. Sayer (1984, 1992)(133). Proponents of critical realist evaluation have argued that the central question is not so much whether certain interventions work in a generalisable way but what will work with these stakeholders/actors in this setting at this time. This shifts the focus of evaluation of interventions from a programme-based view of what works to causal pathways(142)…and to recommendations and policy-making that is founded in locality and practice supporting the argument that policy-making should be devolved from national governments so it is informed by practice locally. Critical realism makes a distinction between the real (underlying nature and causal powers of objects/agents), the actual (what happens if/when those powers are activated), and the empirical (what is experienced/observed). This distinction is central to an ontological conviction that there exists a reality distinct from, and greater than, the empirical and that this reality is comprised of structures and mechanisms independent of our perceptions. Mechanisms can coincide under real world conditions to produce emergent properties in time and space. This notion of contingency contrasts with positivist notions of universal logical necessity by highlighting that propositions may hold true only under certain circumstances(142). Context is key. Whilst empiricism/positivist approaches purport that we should just stick to the observable reality, the critical realist approach is that we have to accept we are limited in our perceptive faculties; that there is more to the story, and

our perceptions are always limited. Influenced by critical realists, Pawson(128) compares causation in different groupings as follows.

Successionists – identify causal agents as variables and seek to observe associations between variables by means of experimental / non experimental data. Explanation or causation is about distinguishing between associations that are real as opposed to spurious. As discussed above, this relates to employing experimental (or quasi-experimental methods in the absence of randomisation) in order to obtain a credible empirical measure of effect.

Configurationists – similar to above but using non experimental methods, it promotes approaches that use a variety of microdata sources, statistical methods and behavioural models to compare the outcomes of participants in social programs with those of non participants(127). Possibly relates to econometric methods.

Generativists – similarly, they look for measurable patterns and uniformities but it is assumed these are brought about by the action of underlying mechanisms which are not variables or attributes, and thus not always directly measurable. Rather, they are processes describing human actions. Emergence of social uniformities is highly conditional and causal explanation is about producing theories of the mechanisms that explain both the presence and absence of the uniformity. The key explanatory tool is the generative mechanism which is able to elucidate some of the reasoning and resources and restraints that lead to action. We discuss this in more detail below.

3.4.3 Generative Causation

What distinguishes realism is its particular understanding of how causation works (143). Realism relies on generative causation, a theory-based explanation of how causal processes happen. Generative explanation is pursued by creating and testing theories(128). It assumes multiple possible causal pathways linking an intervention to any outcome which will hold for certain people in certain conditions(124). The addition of a realist theory informed approach is that it seeks to explain why interventions work rather than simply to establish attribution(144). Causal powers are ‘understood as potentials or processes

inherent in the system studied and processes of reasoning'(128). We assume that there is something about A (inherent powers, mechanisms) that is the actual causal force, and that those inherent powers or mechanisms of A were activated in conducive contexts. Realism sits somewhere between positivism and constructivism, 'closer to post-positivism in ontology; closer to some forms of constructivism in epistemology.' (143)

Thus, realism proposes alternatives to 'causation via counterfactual'. Whilst the concept of temporal precedence is associated with successionist causation and explanation is seen to be located in the variables and 'attributes'(145) in realism, causality concerns not a direct relationship between two observable and discrete events but a relationship between 'the causal powers or liabilities of objects or relations, or more generally, their ways-of-acting or mechanisms and the outcomes of those mechanisms'(146). Furthermore, typically in empirical research, variables are aspects of reality that are abstracted quantitatively to fit into numerical analysis. Realist inquiry using generative causation is different as this narrowly defined quantified data may not have enough explanatory power to provide insight into the causal powers of an intervention.(145)

3.5 Discussion

Our proposed impact framework employs both empiricism and realism to address both complexity, and the measurement and attribution of impact. It could be argued that the aim of theory driven approaches is not to provide definitive evidence that the entirety of any measured change can be attributed to an intervention(131, 144). Causal inference is, however, also an important part of theory-driven evaluation in order to be able to draw conclusions(147). Paying attention to congruency, counterfactual comparisons, and critical review can significantly improve the quality of causal inference(147).

There is a lot of debate and controversy about combining realist approaches with RCT design or quasi-experimental methods(148, 149). How congruent or not a counterfactual is to realist thought is a matter of recent debate due to the potential conflation of 'successionist' and 'generative' causation as RCTs and quasi-experiments attempt to isolate cause to that of the intervention alone by controlling out context, namely the very things that realists view to be key in

explaining an intervention's success or value(150). The strength of a realist approach is its ability to examine interventions and their outcomes in their contexts, irrespective of what form the data are in. The potentially 'contentious' issue is that we propose mixing (quasi-) experimental ie positivist methodologies to derive data (on average effects) within a realist evaluation which works by identifying and explaining differences in effects(151) and which is critical of such 'successionist' approaches to causation as they supposedly ignore the complexity of social causation. For example, the concept of mechanism is used across both paradigms but realists use the term in tandem with 'generative causation' and 'ontological depth' as opposed to from an empirical successionist perspective (ie experimental design, dose-response, empirically measurable etc)(152). However, we are of the opinion that it matters less how the data were generated, we are sense making about why and how outcomes are produced so that our epistemology is still in keeping with realism.

Quasi-experimental methods more likely to be suited to the evaluation of HTA is the application of those which would enable us to find variation in uptake or implementation of hta decisions over time, and where aggregated (as opposed to individual level) data suffice. ITS is particularly useful when an intervention is implemented at population level and when there is a clear time point of introduction. Such methods would allow us to explore any temporal change of the uptake of a technology following an hta, using either uncontrolled or controlled pre- and post- measures. This quasi-experimental approach would ask *does* HTA have an effect/impact which, in turn, we quantify as NHB in our proposed framework (Chapter 4).

However, there are limitations to quasi-experimental methods. A major threat to internal validity with most single time series designs is the possibility that forces other than the intervention came to influence the dependent variable(133). Another threat is how data are classified, requiring a constancy of definition that the data may not always reflect(133). Seasonal variation also needs controlled for if known. Delayed effects are also difficult to interpret especially if there is no theoretical specification of how long a delay should elapse before an effect is expected(133). In this regard, no treatment control

groups are generally indispensable(133). Questions can be also raised about external validity, for example, would the same results be obtained in another country? Contextual causal factors are likely to be at play other than just the intervention – and this is where we propose theory-driven approaches will help with providing greater explanatory power.

To this end, whilst both realism and ToC attempt to understand the role of context (albeit in slightly different ways), rather than attempting to exclude contextual influences through controls(153), we have chosen to use realist inquiry as complementary to the ToC approach taken by iDSI. Combining TOC with RE can be useful(153). Both help researchers to think through sets of issues but ToC does not usually address mechanisms, and its strength is the level of detail which help to identify how it may apply or differ in a new context. In realism, the power of the approach is not in the detailed understanding of entire programmes but rather the opportunity to 'cumulate' specific Context-Mechanism-Outcome Configurations (CMOC) [Chapter 5] across different programmes since similar processes of generative causation may apply across different policy areas(153, 154). Moreover, ToC approaches can provide an analysis of programme theory within which realist approaches can be employed to examine the role of particular causal mechanisms and the contexts within which they operate to generate outcomes(153, 155).

Finally, it is worth stating that much of this resonates with some economists (see, for example, Tony Lawson) who work the ideas of critical realism into economics(156) and whose social ontology is concerned with understanding society (what it actually is) before proposing and measuring the effect of interventions. They reject the neoclassical conception of society as consisting of atomistic independent agents interacting in the rational (mechanical) manner envisaged by Walrasian general equilibrium models(157). Society has to be understood as more than just the sum of its parts. They are essentially interdependent, with the concept of 'emergence' recognised where components of a system interacting together may produce effects qualitatively different to what they could alone(157). Lawson critiques mainstream economics for relying predominantly on (mathematical) predictions in economics despite repeated failures and that the underlying theory behind these models of 'rational

behaviour' is flawed and not reflective of the real world. Indeed, John Maynard Keynes observed that economics was to understand how results were 'generated through the behaviour of real human agents, facing an uncertain future and operating under existing institutional conditions, agents whose actions are motivated by hopes, fears, expectations..'(157). Their social ontology is such that macro-micro interaction be brought into economics to more appropriately take account of the nature of social reality. We should give up any attempt to explain macroeconomic phenomena simply in terms of 'micro foundations' - the reality is that they cannot explain everything (for instance, the paradox of thrift / fallacy of composition - what is true of the individual's action is not true of that of the whole group)(157). Also causation may run from the general (macro) to the micro - crowd psychology determining individual behaviour(157) ie there is a causal power that emerges from social structures which cannot be reduced to explanations at the individual level.

3.6 Conclusion

In the literature retrieved on HTA impact evaluations (Chapter 2), the counterfactual is often acknowledged but disregarded as being too challenging to establish given the difficulty of observing what would have happened anyway without the HTA process. As randomisation is infeasible by the nature of our intervention, we conceive our drawing upon quasi-experiment methodology, as well as involving an attempt to understand the generative mechanisms entailed in the intervention, its natural and social context, and those possessed by the actors involved, will provide greater ontological depth. We take a complexity-informed approach, acknowledging that we cannot control the agency of those who implement. So that, in addition to measuring impact, we believe a realist approach could be fruitful. In our framework, we employ these different methodological approaches which offer alternative interpretations to causation. Exploring how these successionist and generative paradigms may work in tandem is a novel aspect to this research(158).

3.7 Next steps

Our methodological HTA impact framework, a key output of this research, is presented in Part 2.

Part 2: Conceptualising, measuring and valuing the impact of HTA.

The objective of Part 2 is to present the mixed methods HTA impact framework including the quantitative ROI and the realist review (or synthesis) and to illustrate the contribution it makes to the current literature on conceptualising, measuring and valuing the impact of HTA. Our framework aims to quantify impact in terms of a Net-Health Benefit Return on Investment (NHB-ROI) but this, of itself, offers little explanatory power. Hence, the realist synthesis – a theory driven approach - to unearth how favourable conditions are created for the uptake of hta decisions as efforts to rationalise the use of resources are only valuable if recommendations are implemented in practice ie everything we are focusing on here is a step on from HTA dissemination. We choose a realist synthesis as a complementary and congruent approach to the iDSI ToC in explaining *how* HTA is impactful. Implementation is context specific, and realism as an approach to developing programme theory is particularly relevant because it focuses specifically on the influence of context as causal mechanisms are activated only in favourable conditions. This second part comprises chapters 4 - 6 as described below.

Chapter 4: presents our NHB-ROI Framework for quantifying the impact HTA and how it addresses limitations discussed in Chapters 2 and 3. We believe this is important as a lack of, in particular, longer-term impact assessment, may undermine the importance and value of HTA. We describe our mixed methods approach to quantify the value of investing in HTA as well as to explain how outcomes and impact is achieved. This framework addresses one of the most overlooked yet, one of the most critical aspects of evaluation, namely the opportunity costs associated with the capital investment and ongoing running costs required to sustain an HTA infrastructure at the systems level(37). If we think priority setting should be better informed by evidence, and use tools such as HTA, this necessarily implies institutional change to establish a sustainable system. Consequently, we have to consider the opportunity costs which are needed for its establishment and ongoing running(38).

Chapter 5 is the realist protocol describing our proposed process of theory-building and refinement. In adopting a realist perspective, we treat the process of HTA as a complex intervention in itself but also as one that is introduced into existing complex systems and contexts.

Chapter 6 is the realist review. A realist review is a theory-led approach to knowledge synthesis that provides an explanatory analysis aimed at discerning what works, for whom, in what circumstances, how and why. Analysis of all data retrieved was undertaken using realist evaluation principles of extracting CMOC of variables at play, and iterative, participative and collaborative approaches to interpretation.

4 The Impact of HTA: a Net Health Benefit-Return on Investment Framework

4.1 Introduction

In this chapter, we present the quantitative part of our HTA impact framework which uses quantitative data to capture an empirical and credible measure of uptake (or stopping) of a technology following an hta recommendation. This is interlinked to an ROI framework which uses this data to estimate the return in investment in HTA. Specifically, this framework will measure the net health benefit (NHB)-ROI in HTA, capturing its health and efficiency impact, and establishing whether the benefits gained are estimated to be greater than the benefits expected to be forgone as a result of other services in the system having to be displaced (opportunity costs).

As explained in Chapter 1, we distinguish between HTA as a process at the systems level to inform decision-making, and hta as an assessment of an individual technology (or technologies). In order to get to the value of investing in health technology assessment at the systems levels (hereafter, referred to as 'HTA'), we need to look at what the process is delivering. In other words, we need to quantify and aggregate the value of individual health technology assessments (hereafter, referred to as 'hta'). Central to understanding the two levels is the 'value of implementation' (VOImp)(159). VOIimp can be defined as the value realised when health technologies are implemented appropriately(160). Dixon(160) recommends payers must invest in implementation activities for each set of guidance they issue, and manufacturers need to set their prices to give payers a margin to invest in this. As otherwise, the current implementation of guidance results in a loss of value to society(160). VOIimp relates directly to individual htas but in aggregating these analyses, we can get to the value of HTA at a systems level. We use case study design with purposive sampling to populate the quantitative framework as illustrative examples (Part 3) as well as a realist synthesis (Chapters 5 and 6) to help theorise the generative forces or mechanisms that lead to health outcomes. The main components or 'building blocks' of the framework are presented in Table 4.1. The quantitative components are discussed in detail below.

Table 4-1 The main building blocks of the proposed HTA impact framework

Building blocks of HTA impact framework	Methods	Data requirements	Potential data sources	Data challenges
Return on investment (ROI) to measure the impact of HTA in terms of Net Health Benefits (NHBs)	ROI, modelling.	Capital and running costs of investing in HTA at a systems level. Realised NHBs of individual htas.	Ministry of Health or local government audit sources. Decision analytic models to combine /supplement routine admin data with economic modelling techniques (into cost-effectiveness, NHB).	Dependent on aggregating NHB for all hta decisions made. Willingness-to-pay threshold assumed if an explicit value does not exist.

Realist synthesis **Quantitative**	Value of implementation (VOI). Quasi-experimental methods to provide measure of uptake (realist outcome).	Longitudinal data on utilisation / drug volume pre- and post-an hta decision to monitor implementation or uptake; total eligible population; disease prevalence.	Administrative health systems' monitoring or audit data.	Need existence of routine monitoring systems to be able to show temporal tends in uptake (stopping) of a technology following an hta recommendation; prevalence of disease often unknown or uncertain, especially in LMIC.
Realist synthesis/evaluation **Qualitative**	Realist synthesis; realist interviews using case study design.	Qualitative data to identify potential mechanisms to produce outcomes and impact.	Relevant stakeholders in HTA.	Availability and willingness of stakeholders to participate; ability to provide insights to help refute or refine candidate theories.

4.2 Assumptions

Key simplifying assumptions made in the framework are stated upfront as follows:

- Whilst we recognise HTA is a process that may include several activities (for example, horizon scanning, pricing), *we consider its overall purpose is to inform decision-making and priority setting.*

- *The hta decision is always the 'better' one on the basis of cost-effectiveness.* There is, however, always the inherent uncertainty of any decision wrongly adopting or rejecting a technology, consequently leading to net benefits forgone [Figure 4.1].

- *Uptake of a technology is linear so that any change in uptake following an hta recommendation is modelled as a step function.* With time lags for implementation, the diffusion of technologies is more realistically known to follow an S-shaped curve [Section 9.2.3].

- *We allocate NHBs proportionately in line with implementation which assumes NHBs are equitably distributed across the eligible population.* This, however, does not consider how the NHBs are distributed across the eligible population. For example, it may be that difficult-to-reach populations who may have the most to gain, are more likely to often be the last to take up the intervention [Section 4.5].

- *We assume uptake and stopping of a technology to be equivalent in our conceptual framework in so far as there might be some natural diffusion of a technology pre- an hta.* However, stopping of a technology here is not equivalent to disinvestment. By 'stopping' we refer to a negative hta recommendation rejecting to invest in a technology in the first place (ie the 'stopping' of a technology that is not actually yet in the system). We differentiate between this and 'disinvesting' in a technology that is in current usage and/or in clinical practice. Our framework deals with the former, not the latter [Section 9.3].

- *We assume the value of HTA at a systems level is quantitatively the aggregate of that of individual htas.* Whilst we are applying our analysis to the individual hta, the OECD have published on estimating the value of HTA directly at country level (Chapter 2).

4.3 Net Health Benefits

We start by introducing NHB, our proposed measure of impact. A healthcare technology is considered beneficial if it provides more overall health than it displaces as a result of its additional cost diverting resources away from other interventions or services. To ensure that the funding of a new intervention is consistent with the objective of maximising health gains subject to a budget constraint, new health care technologies must provide an incremental cost per Quality Adjusted Life Year (QALY) or Disability Adjusted Life Year (DALY) compared to current care less than the cost-effectiveness threshold(161). The calculation of this incremental cost effectiveness ratio (ICER) expressed as QALYs is shown in Equation 4.1, where (λ) represents the cost-effectiveness threshold.

Equation 4-1 Cost per QALY decision rule

$$Cost\ per\ QALY\ standard\ decision\ rule:$$

$$\lambda > \frac{\Delta Cost}{\Delta QALYs}$$

In LMIC, WHO GDP per capita-based 'thresholds' have been widely used to assess cost-effectiveness but empirical evidence suggests these are likely to be significantly higher than actual health opportunity costs(162, 163).

NHB are simply a rearrangement of the more usual cost per QALY or ICER which we compare with our willingness to pay for a QALY. We rearrange the above equation to define net health (or monetary) benefits (NHB/NMBs). Expressing costs in terms of their health equivalence by dividing through by the threshold (λ), allows costs (C) and effect to be combined into a single metric. Equation 4.2 represents the net gain to the healthcare system from introducing a technology. It captures the health gains directly from the technology and

compares it to the health loss from any technologies displaced elsewhere within system to fund the new technology(161). The concept of opportunity costs as expressed through the threshold is central to NHB.

Equation 4-2 Net health benefits

$$Net\ Health\ Benefits = \Delta QALYs - \frac{\Delta Cost}{\lambda}$$

The current value to the health system of the healthcare technology is the value from all the patients who currently receive it and is determined by the NHB (or NMB) of treating each patient ie the health gains minus the opportunity costs. The net population health benefit of introducing a technology is, along with setting a threshold and a function of its incremental costs and effects in comparison with alternative guidance or standard care, the duration of its usage or validity and the size of the patient population served(164). Equation 4.3 represents the value to the healthcare system of patients currently treated with the technology, where $\boldsymbol{n}$ is the total patient population eligible for treatment and **p** is the current utilisation rate of the technology(161).

Equation 4-3 Value to the healthcare system

$$n * p * NHB = current\ value\ to\ the\ healthcare\ system$$

As such, the advantage of using NHB is that they can measure the scale or magnitude of the health benefits at a population level as offered by the intervention(165).

4.4 Return on Investment

We then define ROI. ROI can be expressed as shown in Equation 4.4. Gains are typically measured in monetary terms (financial ROI), or can also be expressed in terms of social values which are then monetarised (a social ROI - SROI); costs remain the same in both cases(166).

Equation 4-4 Return on Investment

$$ROI = \frac{Gain\ from\ investment - Cost\ investment}{Cost\ investment}$$

If we were to apply the concept of a financial ROI to HTA with its focus on maximising financial returns, this would necessarily mean a preference for investing only in cost-saving treatments. Treatment costs stemming from the technology would need to be offset by any net disease cost-savings over the long term. This is not always the case with cost-effective treatments, with new treatments usually leading to incremental costs and effects at the margin. Thus, were we to apply a traditional method or focus of HTA, we may never arrive at a positive financial return, even over a lifetime.

Instead, if we were to apply a SROI, the values of HTA could be maintained and it would be more broadly applicable to HTA. Rather than monetarising these benefits as is the norm in a SROI(167), we propose the use of NHB, expressing costs – as described above - in terms of their health equivalence by dividing through by the willingness-to-pay threshold for a QALY, thereby combining costs and effect into a single metric. By employing a net benefit approach, our framework allows for the explicit consideration of costs and benefits. A technology is considered beneficial if it provides more overall health than it displaces as a result of its additional cost. Currently, NICE takes this value to be between £20,000 - £30,000 per QALY, though research would indicate it is considerably lower(163). This concept of opportunity costs as expressed through the threshold is central to NHB but is also integral to any ROI. Slotting the NHB metric into the original ROI equation, we can express a NHB-ROI in Equation 4.5.

Equation 4-5 NHB-ROI

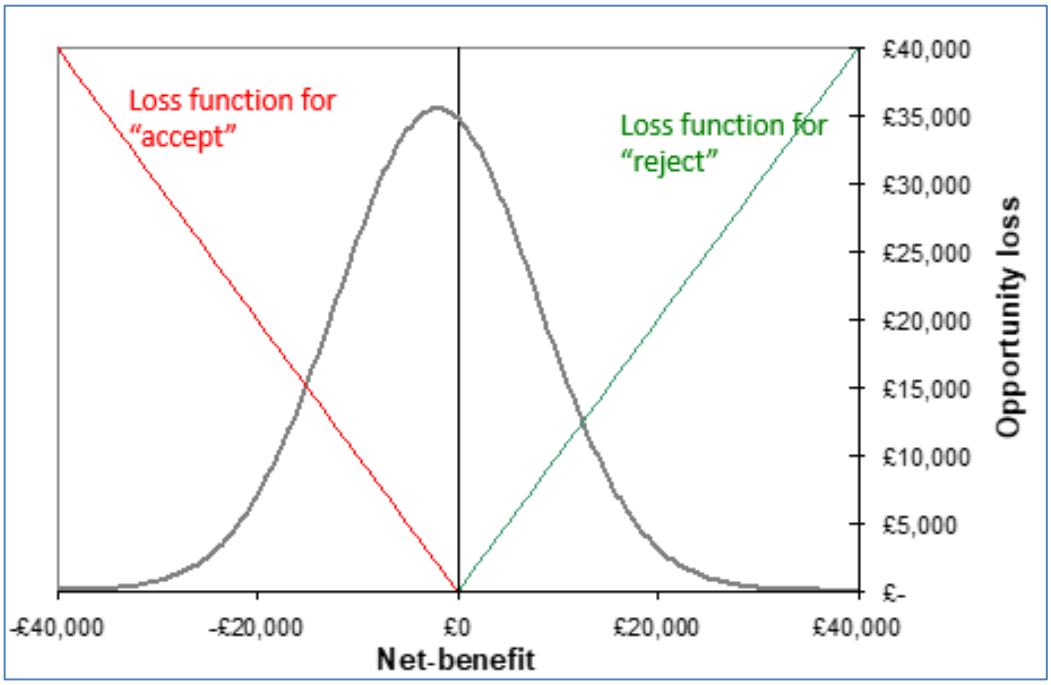
$$NHB\ ROI = \frac{\Delta QALYs + \frac{\left(\Delta C_{Saving} - \Delta C_{technology}\right)}{\lambda} - Cost\ Investment/\lambda}{Cost\ Investment/\lambda}$$

4.5 Realist Synthesis - quantitative

4.5.1 Value of implementation – individual hta

We define the impact of HTA to be achieved through increasing the uptake of net beneficial technologies and decreasing the uptake of non-net beneficial technologies. There is, of course, the inherent uncertainty of any decision not being the 'better' one. We assume cost-effectiveness is the basis but there is always the outside chance the technology recommended was not cost-effective, with that decision not being a true positive (or negative) but rather one which wrongly adopts or rejects a technology, consequently leading to net benefits forgone. Such opportunity cost losses due to the adoption of a 'wrong' decision can be estimated by integrating the distribution of the NHBs associated with an hta with the loss function [Figure 4.1](159, 168).

Figure 4-1 Net Benefit and Loss Function



Source: Reproduced with permission from A Briggs. Briggs A, Sculpher M, Claxton K. Decision Modelling for Health Economic Evaluation: Oxford University Press; 2006

Nevertheless, by employing a VOImp analysis, the objective is to convey the concepts of potential population NHB and realised population NHB. This is depicted in Figures 4.2 and 4.3, and shows that HTA is about both the uptake of cost-effective/beneficial inventions and a decrease in use of those which are not. Implementation is shown across the top and equates to NHB running along

the bottom. The potential population NHB associated with full implementation of a technology is quantified from the use of decision-analytic modelling into long term health and cost outcomes which would normally be undertaken as part of the assessment process of an hta, using the best available evidence at that time. As stated above, the net population health benefit of introducing a technology is, along with setting a cost-effectiveness threshold, a function of its incremental costs and effects in comparison with alternative guidance or standard care, the duration of usage or validity, and the size of the patient population served(164). Realising a net population health benefit then requires using available evidence or assumptions on the degree of uptake (or stopping) of the technology (at any point in time) in order to calculate the value of current implementation. A shortfall between potential and current implementation provides evidence that current care is sub-optimal. Inefficiencies exist in healthcare from the under- or over-use of beneficial interventions as less than optimal adherence implies net benefit forgone.

To what extent we can attribute to the hta the uptake (or stopping) of a technology requires a counterfactual as to what the level of uptake (or stopping) might have been without this. There might have been some natural diffusion or decline in usage anyway which would reduce the overall value and impact of the hta. To estimate the extent to which implementation would have happened without the hta, we propose drawing on methods increasingly applied to natural or quasi-experiments where randomisation is similarly unfeasible(169, 170). We would also propose that randomisation is not just unfeasible; it may not always be the most appropriate approach given the failure of counterfactuals to account for complexity – or for it to be understood (see Chapter 3).

Figure 4-2 The value of implementation (Net beneficial technology)

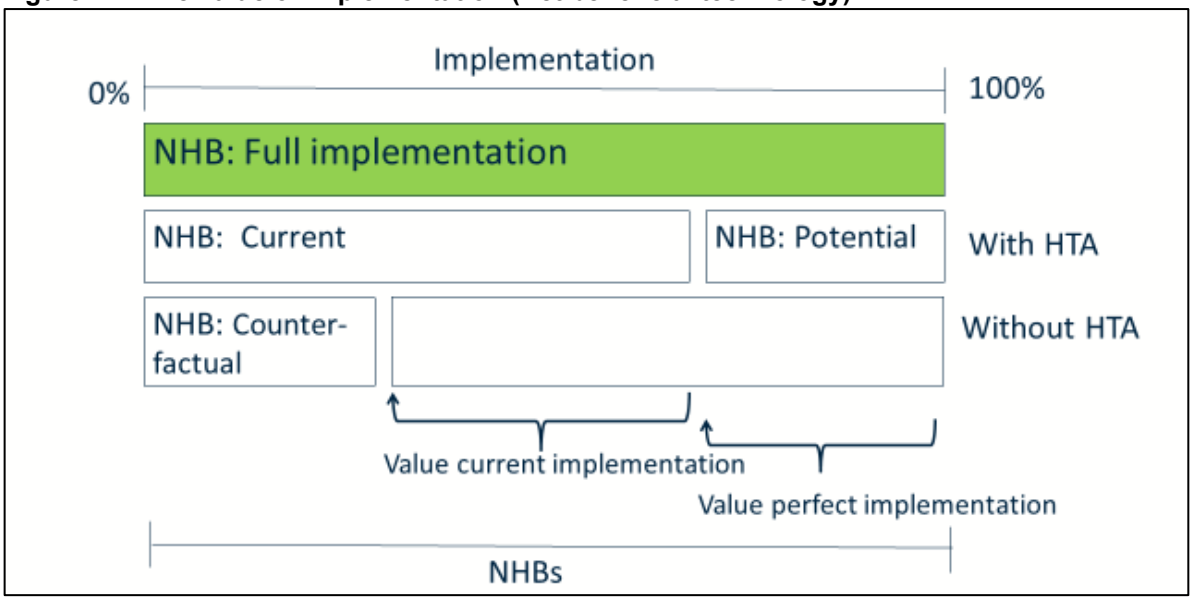
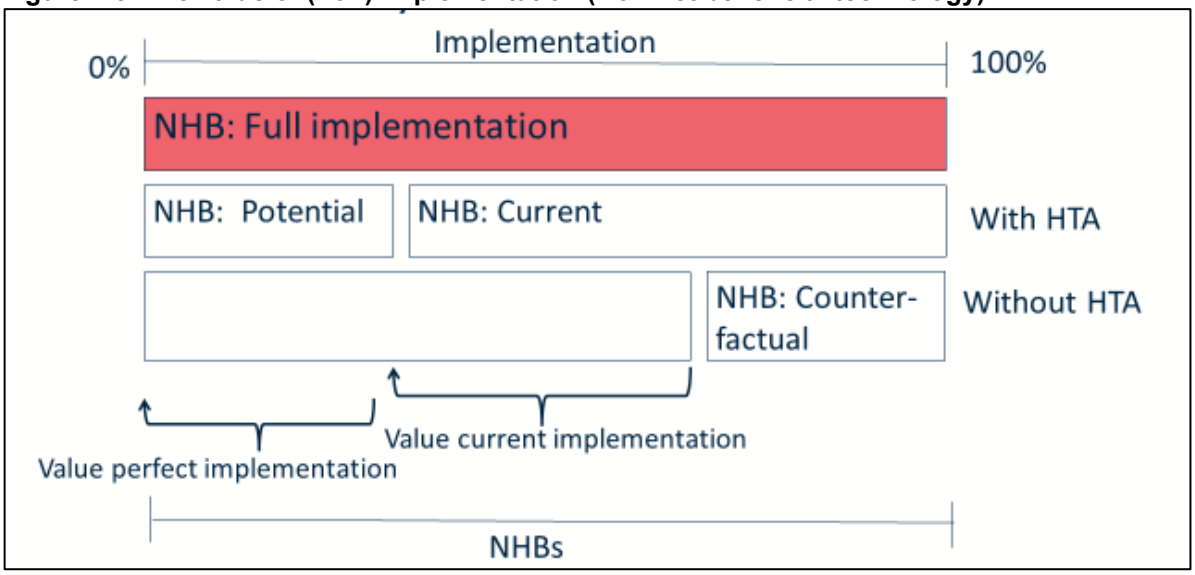


Figure 4-3 The value of (non)-implementation (Non-net beneficial technology)



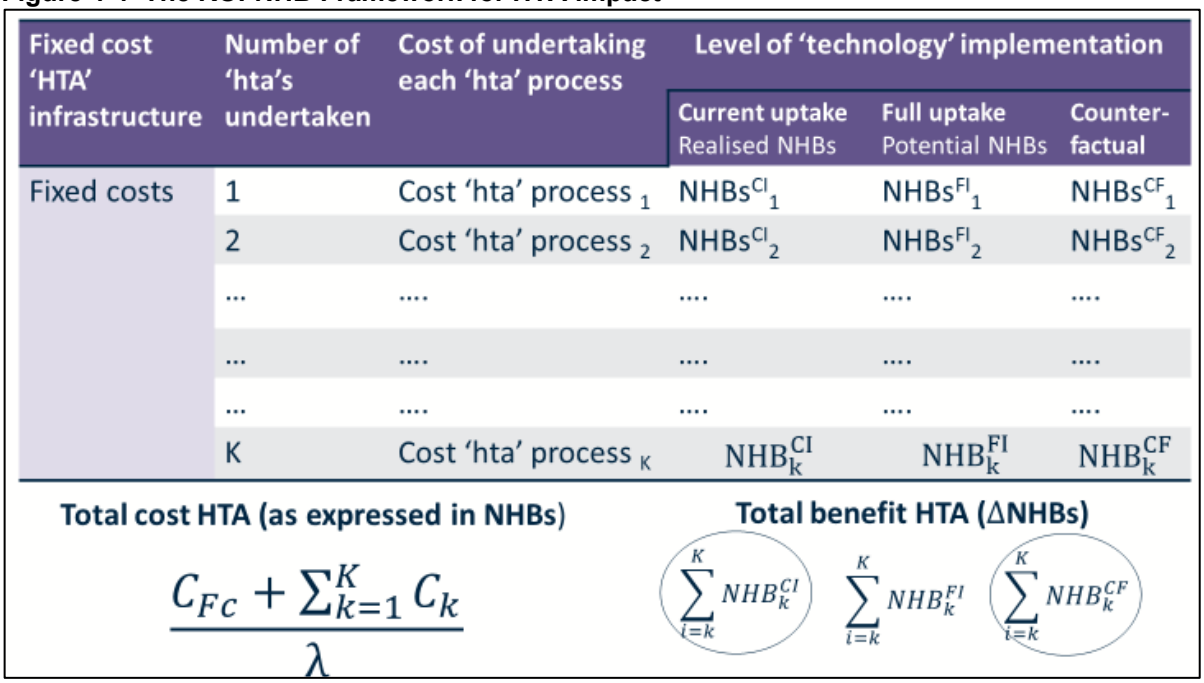
In the literature, VOImp relates at the level of implementing an individual technology. We build on this to employ VoImp at a systems level. We want to measure uptake, for example, the volume purchased of a drug for a specific indication or the adoption of policy or guidance into practice, both following an hta recommendation and, without going through any hta process. By aggregating these htas, we can get to the value of HTA at the systems level. This is described below.

4.5.2 Value of implementation – HTA at the systems level

By aggregating the realised NHB of all hta decisions given the current level of implementation, or as we propose, by taking into account a level of implementation attributable to the hta process itself as measured against a counterfactual, we are able to offset these total benefits against the costs of investing in HTA at a systems level. Such costs would include investing in the HTA infrastructure and the running costs associated with personnel and resources involved in undertaking the assessment and appraisal of each hta.

Figure 4.4 depicts the ROI-NHB impact framework for HTA. Reading from left to right, the first column shows the fixed costs of investing in HTA. In the next, we identify individual htas undertaken as, in order to get to the overall value of investing in HTA at a systems (or country) level, we need to look at what the process is delivering. In other words, we quantify the value stemming from each individual hta. The 3rd column lists the costs of undertaking the assessment and appraisal process for each hta numbered 1-K. Moving onto benefits, the 'level of implementation' relates to the uptake of a technology further to an hta recommendation. Full implementation (potential NHB) is everyone who is eligible receiving it; current implementation (realised NHB) relates to the number of patients actually receiving the treatment; and the counterfactual is what we reckon the situation might have been had the hta not been undertaken. In this way, we can establish what we can attribute to the hta. We show the NHB stemming from each hta, numbered 1-K, associated with full, current and attributable levels of implementation. The summed NHB associated with 'current' and 'counterfactual' levels of implementation are circled as relate to the realised and attributable benefits of the hta (attributable benefits being current NHBs minus the counterfactual). The fixed costs (CFC) of investing in HTA and the running costs (C) are also summed and expressed in terms of their health equivalence by dividing through by the willingness-to-pay threshold (λ). The aggregated costs and benefits of HTA, both expressed in their health equivalence of NHBs, can be directly offset against each other. Net gains or losses can be expressed as a percentage of the initial investment to obtain the ROI.

Figure 4-4 The ROI-NHB Framework for HTA impact



4.6 Discussion

We present here a methodological framework to quantify the impact of HTA expressed in NHBs. The objective is not to rank or score HTA bodies but recognise that evaluation should help to improve (optimise) HTA in a given context. The focus of our framework is on countries where HTA currently exists and on those countries transitioning out of donor support as to what institutional models might be appropriate. If we think priority setting should be better informed by evidence, and use tools such as HTA, this necessarily implies institutional change to establish a sustainable system. Consequently, we have to consider the opportunity costs which are needed for its establishment and ongoing running but which are often overlooked(38). We propose this framework should help with this. Linking to opportunity costs, our proposed use of NHB measures the impact of HTA in terms of health and efficiency gains at a systems level and, importantly, reflects the opportunity costs of investing in individual technologies though the use of a threshold. We recognise most countries do not operate with explicit thresholds and that there is much debate as to how they are to be measured and estimated (either using willingness-to-pay or the marginal productivity of the healthcare system). Where none exist, we would propose doing scenario analyses around plausible values - see for example,

Woods et al(171). Yet, we believe there is learning still to be had for countries where HTA is not currently institutionalised or developed at all, due to limited capacity and lack of awareness of its potential amongst policy makers, by assessing the costs and benefits of different (discrete) processes and mechanisms versus investing in full HTA processes. It should also help build needed political support for improved priority setting processes(172).

We aim to address some of the methodological limitations identified in Chapters 2 and 3; we do not stop at full implementation, we include 'negative' decisions ie not just those which could offer benefits if introduced, and critically, we construct a counterfactual in order to facilitate isolating or disentangling the impact of an hta on outcomes from what would have happened without it. Ideally, we would find a sector or jurisdiction unaffected by an hta to make such a comparison but as groups get increasingly non comparable, threats to internal validity become more plausible. Alternatively, we draw on quasi-experimental methods to construct a 'no hta' control. Quasi-experimental methods, such as ITS or simpler forms more suited to analysing routine data such as segmented regression (see Part 3), can be used to analyse impact at a population level over a defined time period by measuring the adoption of a technology after an hta recommendation against its pre-existing trend of adoption (the counterfactual). We apply this quasi-experimental approach to an individual hta, and propose aggregating the htas to get to the value at a systems level [Figure 4.4]. Note that the OECD have published on the value for money of HTA at a country level on population health outcomes and public spending(48) [see Chapter 2].

Whilst ITS was successfully used previously to assess the implementation of NICE guidance(135), it no longer makes sense to use this on contemporary NICE guidance as these processes have now become institutionalised with hta recommendations made predominantly on newly licensed technologies. This means there are fewer data prior to the hta on prescribing levels or volume against which to measure the influence of the hta recommendation. However, in countries transitioning out of donor support, the intended focus of our impact framework and where HTA processes have not yet been institutionalised, there

should be more data to be able to apply such methods – indeed, as Sheldon et al did when NICE was still only advisory(135).

Finally, as the standard form of cost-effectiveness analysis is indifferent to the distribution of outcomes, there is a need to take account of equity in any value of HTA framework. The QALYs would be modelled as is the norm in any cost-effectiveness analysis as part of an hta, either extrapolating from endpoints in a trial and/or using decision-analytic techniques. However, their distributional impact is of importance in an impact evaluation. Recent methodological developments allow cost-effectiveness considerations to be combined with assessments of equity impacts(173). Extended cost-effectiveness analysis (ECEA) was developed to address both the health and financial consequences of public policies(174). Distributional cost-effectiveness analysis (DCEA), similarly looks at the distribution of costs and benefits but it additionally accounts for the distribution of opportunity costs and weights the trade-off between health and the inequality reduction objective(175). Instead of a traditional CEA, the potential incorporation of ECEA or DCEA could better account for equity(176).

4.7 Limitations

We acknowledge that implementation does not equal impact – but here, impact is given as modelled so we focus on what level of that impact is realised. We purposively do not consider other spillover effects arising from the HTA process itself in order to limit the scope. We do recognise though there are likely to be other externalities to arise from this process including, for example, better bargaining power on price negotiations. However, our understanding is that the impact of HTA on health outcomes is the major gap in the literature. A realist approach [Chapter 5] should help capture other potential outcomes.

A key limitation of using the NHB approach is its reliance on the value of the cost-effectiveness threshold as a representation of the true opportunity costs of resource allocation within a health system. The conceptual basis of thresholds can be either classified as ‘supply side’ (what the health system is ‘able’ to provide given resource constraints) or ‘demand side’ (based upon expressions of

the value of health). The former requires assessment of the opportunity costs of scarce healthcare resources. We recognise most countries do not operate with explicit thresholds and that there is much debate as to how they are to be measured and estimated, either using willingness-to-pay or the marginal productivity of the healthcare system. Where none exist, we would propose doing scenario analyses around plausible values using, for example, published country estimates by Woods et al(171).

Operationalising the framework to assess the ROI of an entire country's HTA programme is unlikely to be feasible. Practically, we can only undertake illustrative case studies. RAND restricted its evaluation to 10 HTA-funded projects framed within, presumably, the more clearly defined boundaries of an audited research programme(19). However, we can scale-up case studies to calculate how many htas may need to be undertaken in order to get a positive ROI.

Finally, the background to this framework in relation to the iDSI ToC has been discussed [Chapter 1]. The iDSI ToC was developed in full consultation with stakeholders and partners in the countries concerned. For our framework, we have taken this forward after its development with stakeholders to undertake illustrative case studies [Part 3] and we propose a future application of the full framework [Part 4].

4.8 Summary and conclusion

The NHB-ROI framework is presented. Ultimately, we hope this ROI-NHB framework will contribute to demonstrating the value of HTA by quantifying the NHBs and opportunity costs of investing in these processes. In turn, we hope this will contribute to generating political will and financial investment in these processes. What our framework adds to existing models and why this is important is summarised in Chapter 10.

4.9 Next steps

We present the realist synthesis in the next two chapters. We envisage this research, by synthesising economic and more qualitative methods, will provide a framework to quantify the value and impact of HTA on health and economic outcomes, as well as evidence informed theory to produce recommendations as how to do HTA by context in order to optimise its impact.

5 The Value of HTA: protocol for a realist synthesis

5.1 Introduction

We present the protocol for a realist synthesis of evidence of HTA value, describing our proposed process of theory-building. Theorisation allows a greater understanding of how interventions work. In adopting a realist perspective, we treat HTA as a complex intervention in itself but also as one that is introduced into existing complex systems and contexts. Indeed, the systems in which HTA will operate already have policies, procedures, communications, cultures, histories and 'ways of doing things'. As in the quantitative framework, we distinguish between hta as an individual assessment of a technology (or technologies) including evidence-based interventions, practices or policies and, HTA at the systems level to inform priority-setting and decision-making. Both are potentially complex, operate at different levels and are likely to interact. The constructs of a realist approach provide a framework to capture and address the complexity and multi-dimensionality of HTA: at the individual technology level, at the systems level and in their interactions. The realist synthesis will follow the steps and procedures outlined in RAMESES publication standards(98, 177).

5.2 Implementation of HTA findings

Efforts to rationalise the use of resources are only valuable if recommendations are implemented in practice(178), as it is only when those decisions result in practice change can better health be achieved. Yet, we know that implementation of HTA findings is variable(19-23). A recent review(179) has highlighted the need to look beyond the development of HTA guidelines to the mechanisms through which guidelines can be implemented. According to this review, HTA has focused on the development of products such as quality standards, guidelines and care pathways but there is less evidence of a systematic approach to thinking through the various institutional mechanisms through which they might be implemented and the associated incentives to do so – thus avoiding a situation where, for example, any change might be short lived

as unless there is investment and commitment to enact upon hta recommendations, any new practices or protocols could ‘snap back’ to the old norm.

A review by Hailey et al (112)on the influence of HTA found that ‘the literature on assessment of HTA influence is still quite limited and there is little on longer term effects on clinical practice and health outcomes’. This is also a likely consequence of a lack of routine administrative monitoring and evaluation data, leading to poor data capture and reporting. A lack of longer-term impact assessment may undermine its importance and value, as hinted by Hailey et al. Assessing the real-world monetary, health and broader societal effects of HTA as currently implemented in countries around the world will require not only the expertise of traditional HTA practitioners but also the unique perspectives of social science(180). By synthesising some of the vast theoretical literature from the social sciences, we aim to promote a better understanding of the implementation of HTA recommendations(181).

5.3 Research aims and objectives

This realist synthesis aims to produce tested and data-driven theory that considers individual, interpersonal, institutional and systems-level components and their interactions on the mechanisms by which HTA impact can be optimised. Our specific research questions are:

- What are the mechanisms that encourage the implementation of hta findings?

- Do the injected resources have the leverage to trigger these mechanisms and counter contextual deficiencies (transform the context)?

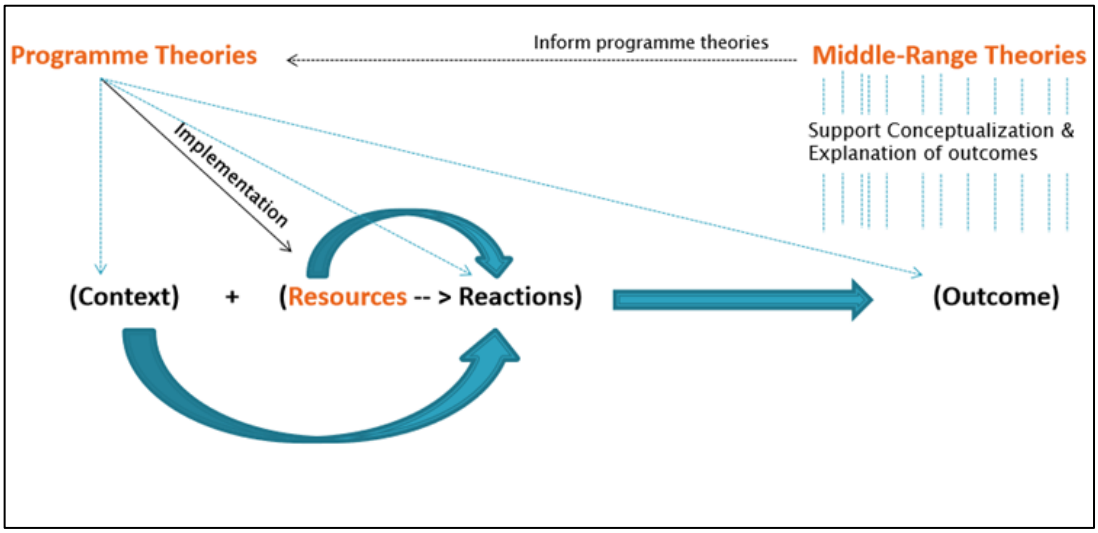
- What are the common predicting mechanisms that transcend contexts?

5.4 Methodology

Realism is a new conceptualisation of programme theory (see Chapter 3), valuable for addressing the complexity of programmes(141). It is a form of theory-driven evaluation based on realist philosophy and which is becoming increasingly used in the evaluation of complex interventions. There are many types of theory driven evaluation but realist evaluation can be differentiated as being underpinned by realist philosophy which makes explicit that there are underlying mechanisms, thus acknowledging different levels or layers of knowledge ie ontological depth(141). This means at the empirical level there are commonly observed outcomes and, at a deeper layer, there are explanatory underpinnings of such outcomes. A realist logic to inquiry attempts to provide an explanatory analysis aimed at discerning what works, for whom, in what circumstances, how and why(141).

Realist theory starts with the basic premise that underlying mechanisms, operating in particular contexts, generate outcomes. Explanation-building using a realist lens is operationalised through investigating CMOC(182). We shall produce CMOC based on the literature and assess to what extent any pre-existing theories fit with these CMOC to produce our preliminary programme theory. Implementation of a programme is dependent on both the context and the intrinsic properties of the intervention itself. A realist perspective treats 'other resources' ie those not introduced by the programme, as context for the programme mechanisms(183). Such contextual elements (values, competencies, assets, deficits, infrastructure) have an impact. It conceives of programme mechanisms as 'reasoning and resources' ie the (new) resources, opportunities or constraints that the programme theory or intervention – here, the hta – introduces, and 'reasoning' (for example, trust-building, motivation to act, realisation of knowledge) as to how people react in response to those resources(184). As such, there is always an interaction between context and mechanism [Figure 5.1](185). Outcomes arise from a combination of the stakeholders' choices (reasoning) and their capacity (resources) to put these into practice(139, 141, 184).

Figure 5-1 Context-Mechanism Interaction

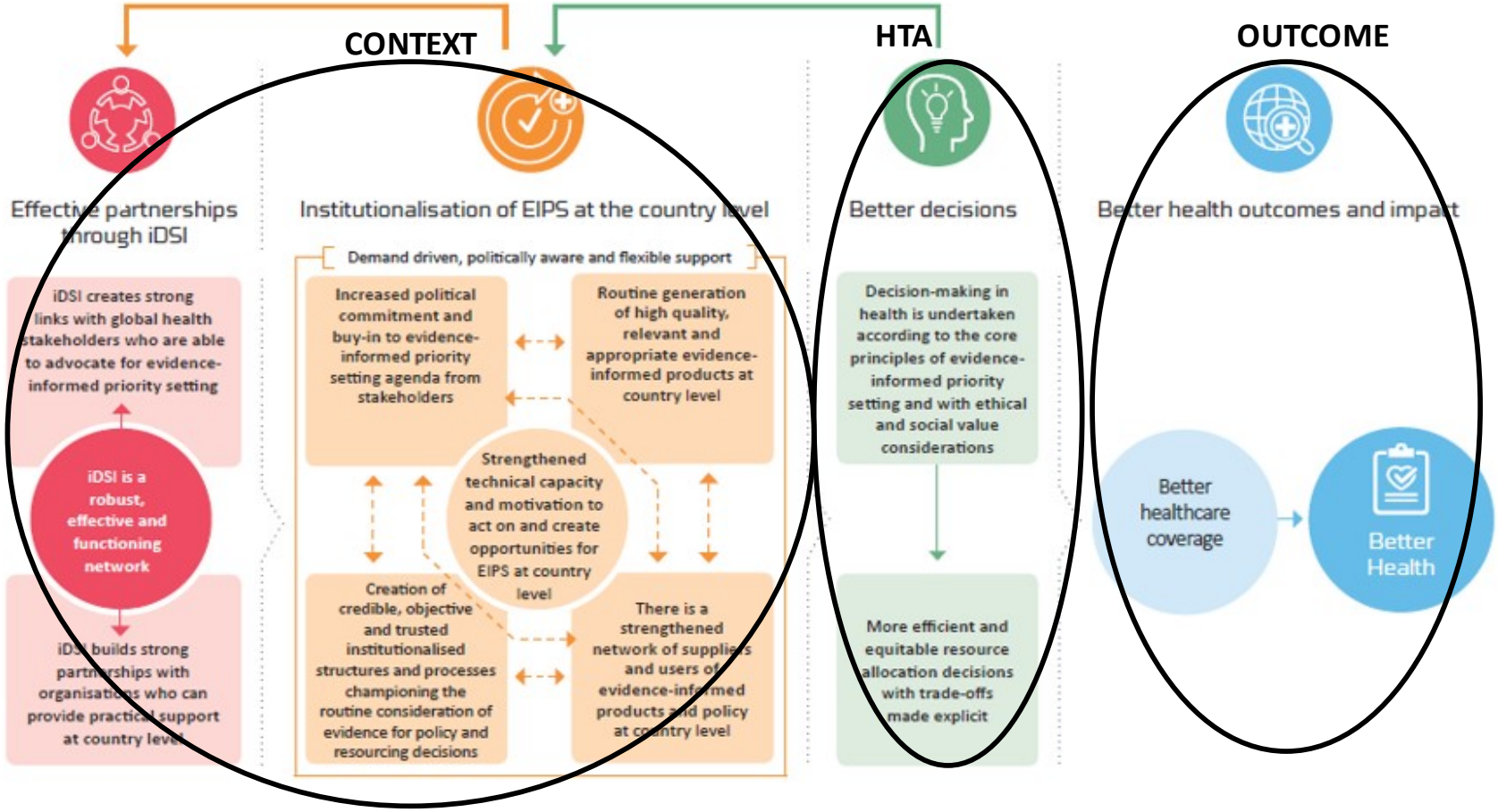


Source: Reproduced with permission from Jagosh J (June 2018). Incorporating Middle-Range Theory in Realist Evaluation and Synthesis. [Webinar]. CARES Webinar Training Series with permission from Justin Jagosh, June 2020.

Whilst iDSI are focusing on institutional aspects around achieving 'better' recommendations/decisions (the 'initiation' context), this realist synthesis will target those institutional characteristics involved in getting hta recommendations into practice (the 'implementation' context) – although clearly there will be some overlap. Barriers and facilitators to implementation can arise at multiple levels - patient, provider, organisational, policy (see Chapter 2). Whilst much research has established what factors influence improved decision-making including good governance structures, expertise, political and institutional factors, resources and participation (41-44), it is less understood how such influences interact with local context and health systems, leading to the improved implementation of evidence-based recommendations and guidance. Indeed, the data for studies of the implementation of reimbursement decisions have been found to be scarce but available evidence indicates that there is a gap between decisions and implementation, resulting in an inefficient use of new technologies and thus a loss of value(24). Our research theorises those elements of the context that are critical to the intervention 'functioning'. At this stage of the hta, there is generally no further injection of resources and instead the reliance is on pre-existing resources for the implementation of recommendations. To what extent can an hta produce a variety of new resources that can leverage against contextual barriers, triggering those mechanisms which optimise uptake of a technology following a

recommendation? Realist inquiry helps to understand the leverage hta/HTA has and this paradigm of realism can help bring new insight to optimise its impact. To illustrate this, we overlay the iDSI ToC with realist terminology: ‘context’, ‘outcome’ and the intervention ‘HTA’ [Figure 5-2], mechanisms being found in an interaction between the context and the HTA resources/responses.

Figure 5-2 iDSI Theory of Change overlaid with CMO configuration



Source: Adapted from report on NICE's engagement in China, 2015(179) and https://www.idsihealth.org/blog/changing-a-theory-of-change-six-important-lessons-from-our-work-with-idsi/ with permission from Itad Ltd.

5.5 Identifying potential theories

A realist search focuses on the theory underpinning the intervention or it can also look for existing evidence how interventions work, building on stakeholders' knowledge of the area or subject to inform the programme theory. By drawing on an iterative analysis of the literature and formal theories, we seek to generate, test and refine explanations for the (non-)implementation of hta recommendations. These formal theories are not specified in realist terms ie with respect to contexts, mechanisms and outcomes. The degree to which hta recommendations are 'taken up' is observable, and therefore lies at the empirical level. Mechanisms, in the sense that realists use the term, are 'underlying' causal processes that cannot usually be directly observed.

An initial scoping of the literature was undertaken to identify existing HTA evaluation or impact frameworks [Chapter 2] as well as theories related to uptake to inform our initial candidate theory. The latter involved a relatively unstructured scan focusing on formal knowledge translation and implementation readiness theories [see Figure 1.3]. We also drew on the growing body of literature on the barriers and facilitators to the implementation of cost-effectiveness recommendations [Chapter 2]. An extensive body of literature identifies various barriers and facilitators to implementing clinical guidelines in general and evidence-based guidance more specifically. Whilst the literature usefully explains implementation in terms of barriers and facilitators, this approach only provides antecedent factors of uptake. By adopting a realist approach, we provide a dynamic framework that offers explanatory power, and thus has the potential to be utilised in evaluation. Finally, as stated above, we are dealing with both hta at the intervention and HTA at a systems' level. Both are potentially complex, operate at different levels and likely to interact. We consider theories at both levels.

5.5.1 Realist approach – individual 'hta'

To make a broader assessment of the role of HTA in an entire health system, we need to understand the value of what that process is actually delivering. Given the value of HTA is dependent on the implementation of a technology itself, this

is about understanding how well an hta works in any given context. How far along the implementation path we get [Figures 4.2 and 4.3] depends on this theory element. It is likely there is relevant theory in a range of different sectors that could be combined or adapted and applied as to how a technology is taken up (or stopped) following an hta recommendation (or rejection) in order to produce an initial programme theory or hypothesis(183). For example, we drew on knowledge translation theories as to how knowledge is utilised as well as the research on the barriers and facilitators to the implementation of cost-effective interventions as useful starting places to theorise the uptake (or stopping), and thus impact, of an 'hta' (41, 42, 84, 90, 114-116). Implementation science, which is 'the study of methods to promote the adoption and integration of evidence-based practices, interventions and policies into routine health care and public health settings'(186), emerged in the wake of evidence-based medicine. It relates, in our case, at the level of implementing an individual technology following an hta recommendation(187). Formal theories from this field could include, for example, those of socio-cognitive behaviour change.

5.5.2 Realist approach – HTA at the systems level

A realist lens is equally valid applied to HTA as a process at the systems' level. Policy implementation is the process of carrying out a government decision(188) and formal theories from this field which apply to more macro or meso institutional levels to inform our theorising could include 'the Diffusion of Innovations', Governance Theory, Institutional Theory and network governance (187). In applying Diffusion of Innovations theory, Yates(189) highlights the importance of both 'macro theory (systemic adoption, that is, organisational and structural change) and micro theory (individual change)'(139), a useful distinction we make here too. Such theories are broadly about the interplay and tensions between knowledge, power and social control, the premise being that organisations do not make decisions but people with biases, motives and histories make the decisions but are required to do so within the confines of power structures (like organisations and governments). This is congruent with the philosophy of realist evaluation.

An overview of the many knowledge translation and implementation theories was facilitated by retrieving existing reviews of frameworks, models and theories(190, 191). Key theories considered included the Consolidated Framework for Implementation Research (CFIR)(192), the Critical Realism and the Arts Research Utilization Model (CRARUM)(142), the Ottawa model of research use (OMRU)(193), the Context and Implementation of Complex Interventions (CICI) framework(194), and another from organisational readiness theory(195). The CICI comprises three dimensions of context, setting and implementation. The OMRU includes six key elements: evidence-based innovation, potential adopters, the practice environment, implementation of interventions, adoption of the innovation, outcomes resulting from implementation of the innovation. The CFIR consolidates multiple different conceptual frameworks relevant to implementation research and proposes that implementation is influenced by intervention characteristics, the outer setting, the inner setting, the characteristics of the individuals involved and the process of implementation. The CRARUM was interesting as an adaptation of the OMRU but borrowing elements from critical realism including the concept of generative mechanisms – and thus relating to our realist inquiry. The CRARUM incorporates critical realism into their knowledge transfer model "to shed light on the structures, powers, generative mechanisms, and tendencies that characterise clinical settings and the agential reflexive capabilities of health care practitioners"(142).

Building on iDSI's ToC which is at a broad macro level, this helped to narrow our realist synthesis to also draw upon theory at an organisational level. In the CRARUM, the unit of analysis is at the individual level, whilst a theory at an organisational/institutional level would keep the focus on HTA at a systems level. Weiner's theory relates behaviour of individuals to knowledge mobilisation at a macro or systems level, reconciling the structural (organisational resources) and psychological (motivations, values) views of organisational readiness which could be seen to relate to the HTA/hta/technology implementation interaction(196). It is noted the application of organisational theory in implementation has to-date been limited(197). Yet, as it is usually government policy which sets the scene for adoption, this public policy/macro perspective is important to uptake.

We consider the Promoting Action on Research Implementation in Health Services Framework (PARiHS)(198-200) model of implementation alongside Weiner's Theory of Organisational Readiness(196), to establish a more thorough framework for expressing our candidate theory. Whilst the CFIR, CICI and OMRU focus is on characteristics of the intervention, PARiHS identifies characteristics of evidence as its focus of implementation. We opted for the latter as whilst we assume the principles of HTA will be the same in LMIC and HIC, a main difference may relate to the availability of evidence. The original PARiHS Framework has been used which proposed that successful implementation is a function of the relationship between evidence, context and facilitation. A later version, the i-PARIHS(201), incorporates the addition of recipients. Although this subsequent inclusion as to how individuals interact with context reflects the realist mechanism of 'reasoning and resources' ie individual's behaviour, intentions and actions, we kept to using the original framework with its focus on evidence as described above.

Together, these theories (PARiHS and Weiner's Theory of Organisational Readiness) provide a model for examining how the uptake of hta recommendations occur successfully. The PARiHS model offers the contextual and intervention perspective, and the theory of organisational readiness considers more fully individuals are agents of change. In so doing, we acknowledge a political/national, organisational and individual spectrum to the model. They are described in more detail below.

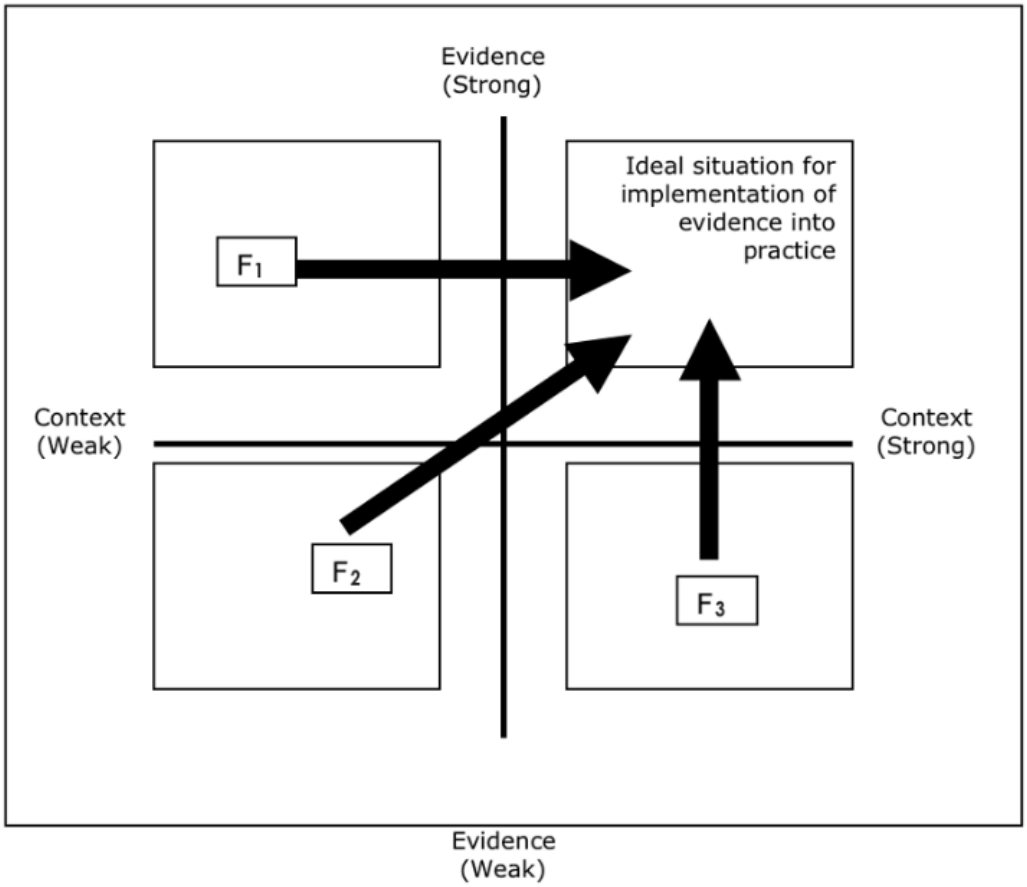
5.5.3 PARiHS model

Knowledge translation (KT) is a term increasingly used in healthcare to represent a process of moving what has been learned through research to the actual applications of such knowledge in a variety of practice settings and circumstances(202). KT encompasses all steps between the creation of new knowledge and its application, and there are a large number of related models and theories(203). However, the final component of the process, the actual use and implementation of knowledge, is not included in all. We reviewed those

models which specifically included implementation and propose the use of the PARiHS model(199, 204-206).

PARiHS is an implementation framework which takes into account the role of context and, therefore, aligns well with a realist approach. Furthermore, the model can operate at an institutional or systems level and contemplates the interplay of evidence, context, and facilitation(207, 208). Whilst we assume the principles of HTA will be the same in LMIC and HIC, a main difference may relate to the availability of evidence. The PARiHS framework was used to help define our initial programme theory by diagnosing the context into which HTA is being introduced. PARiHS can be operationalised through grid-plotting to assess 'readiness'. Each element can be assessed for whether its status is weak or strong with a negative or positive influence on implementation. The PARiHS logic is thus that strong 'Context' plus 'Evidence' results in successful 'Implementation'(202). Our hypothesis is that high- and low-income countries will be positioned in different quadrants of the framework [Figure 5.3]. We shall look for variation in these contexts to test if the theory is transferable. As it stands, the PARiHS model tells us 'why' implementation is achieved but a realist approach will build on this to tell us 'how'. We will interrogate and develop the PARiHS model by adding mechanistic concepts to it.

Figure 5-3 The PARiHS Diagnostic and Evaluative Grid



Source: Kitson A, Rycroft-Malone J, Harvey G, McCormack B, Seer K, Titchen A. Evaluating the successful implementation of evidence into practice using the PARIHS framework: theoretical and practical challenges. Implement Sci. 2008;3(204). Use of this image is supported by the Creative Commons CC BY https://creativecommons.org/licenses/

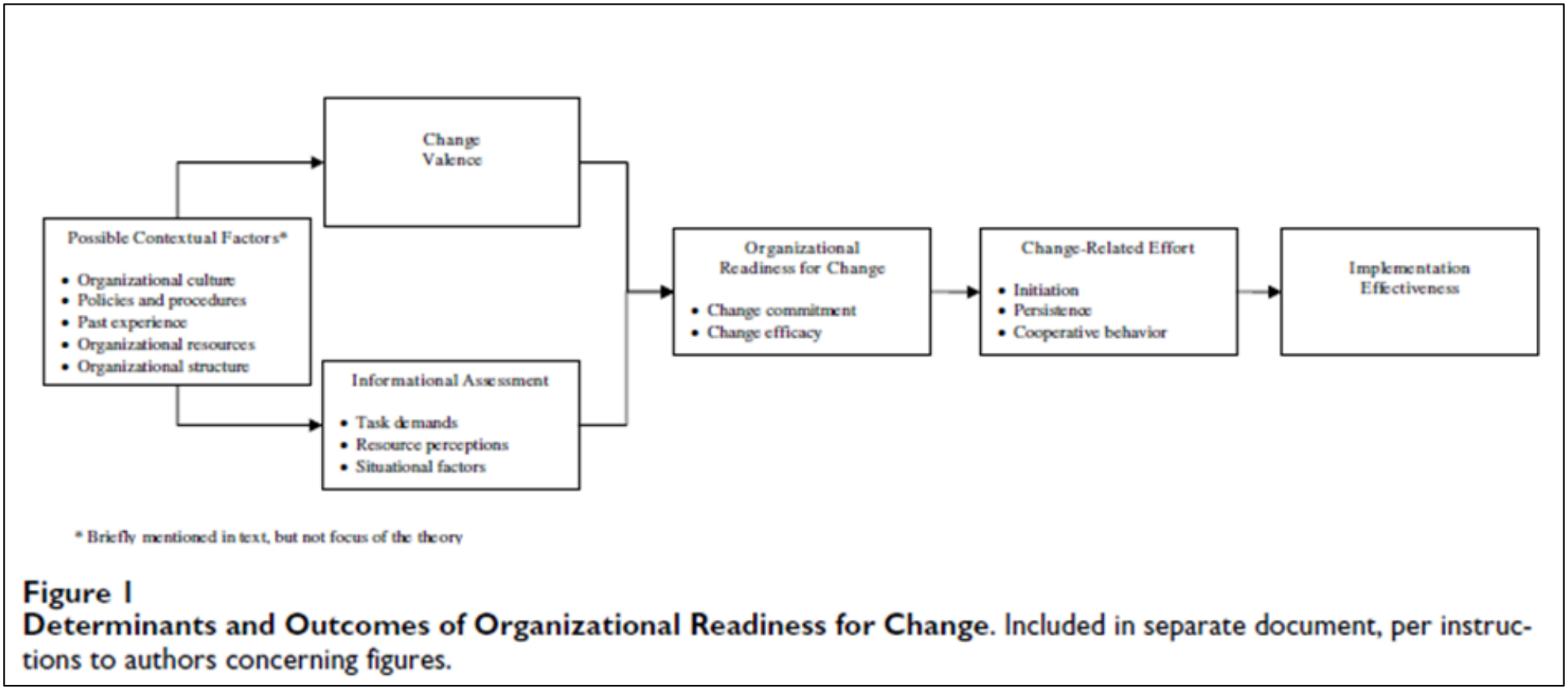
The later addition of the construct of 'recipients' to the framework (201) illustrates how individual level factors, as well as political and structural factors, are also included as influential.

5.5.4 Weiner's Theory of Organisational Readiness

We integrate the individual perspective by drawing upon Weiner's Theory of Organisational Readiness(196) which proposes that 'implementation readiness' depends on collective behaviour change linked to change commitment (wanting to change) and change efficacy (ability to change), thereby making a key distinction between capacity and readiness. Organisational readiness for change varies as a function of how much members value the change and how favourably they appraise implementation capability ie task demands, resource availability and situational factors. When organisational readiness for change is high,

organisational members are more likely to initiate change, exert greater effort, exhibit greater persistence, and display more cooperative behaviour. The result is more effective implementation(196). The theory offers a means of reconciling the structural and psychological views of organisational readiness found in the literature and draws specifically on motivation theory and social cognitive theory(196). Research by Herscovitch and Meyer(209) [cited in(196)] found that organisational members whose commitment to change was determined by 'want to' motives rather than 'need to' or 'ought to' motives, exhibited not only more cooperative behaviour but also championing behaviour ie ‘promoting the value of the change to others, thus creating spread of innovation and ultimately, new cultural practices’. Resources and other structural attributes of organisations are not automatically categorised as features of readiness. Instead, ‘they represent an important class of performance determinants that organisational members consider in formulating change efficacy judgments’(196). Thus, ‘organisations with the same resources, endowments and organisational structures (capacity to implement) can differ in the effectiveness with which they implement the same change depending on how they utilise, combine and sequence organisational resources and routines’(196).

Figure 5-4 Weiner's theory of organisational readiness



Source: Weiner BJ. A theory of organizational readiness for change. Implement Sci. 2009;4 (196) **Use of this image is supported by the Creative Commons CC BY https://creativecommons.org/licenses/**

When the PARiHS model of implementation is considered alongside Weiner's Theory of Organisational Readiness, we establish a more thorough framework for expressing our candidate theory. The PARiHS model offers the contextual and intervention perspective, and the theory of organisational readiness considers that causal mechanisms cannot always be usefully understood in terms of individual reasoning, since they are operating at the level of the organisation (153, 196). Together, they provide a model for examining how the uptake of hta recommendations occur successfully.

5.6 Search strategy

The search strategy proceeded iteratively as follows:

Firstly, a detailed search focused specifically on HTA and 'health outcomes', our impact measure of interest but also HTA and 'uptake/implementation' as a proxy measure given the lack of literature on impact on health outcomes. Search terms, therefore, covered two key domains:

Health Technology Assessment – using MESH terms and expanding this for all relevant terms which could include 'decision-making', 'priority-setting';

Outcomes – and expanding this for all relevant terms which could include 'uptake', 'diffusion', 'adoption', 'implementation', 'health outcomes', 'patient outcomes', 'Net Health Benefits and/or Net Monetary Benefits'.

Secondly, ongoing reference, citation and author tracking to identify the most theory-relevant studies available; and finally, refined searches, operating concurrently with the overall synthesis process, to collect additional materials that may be required to elucidate particular aspects of theory.

Search terms were applied to Ovid MEDLINE and EMBASE databases (see Annex B). Articles were limited to the English language but not limited by timeline or any methodological design. Titles and abstracts were reviewed first, then full papers based on the selection criteria below. Searches of grey literature included the websites of HTA agencies. For example, the UK's National Institute

for Health and Care Excellence (NICE) has databases on shared learning and the implementation of evidence containing 'real-life' examples of how organisations have put NICE guidance into practice (see Annex C).

5.6.1 Study selection criteria

Inclusion criteria were studies that included a measure of implementation of hta decisions or impact on health outcomes. We adopted a definition of evidence of implementation as follows: 'studies that reliably reported consideration by decision-makers of HTA findings and/or recommendations; comparative studies that included relevant measures related to use of a health technology before and after dissemination of an HTA; and studies that reported changes in one or more features that could be credibly linked to information provided by an HTA. Those features were policy related to a health technology, use of a health technology in a healthcare system, relevant health outcomes associated with use of a health technology, and an increased level of research or initiation of research'(112). Studies that reported either a qualitative or quantitative measure of the implementation of HTA findings or its impact on health outcomes were included. We excluded studies that did not provide any measure of implementation or health outcomes or, did so without providing any contextual information thus lacking realist material on contexts or mechanisms. We updated the review undertaken by D Hailey et al, 2016(112) to present day. All their included papers were also reviewed using the theoretical framework.

5.7 Data extraction

For each paper identified as relevant, we extracted contextual data and mechanisms to inform our analysis. Data were extracted from all sections of articles. In appraising the quality of evidence, we critically reflected on all evidence and determined its relevance and robustness for the purposes of answering the review question. The Realist and Meta-narrative Evidence Syntheses: Evolving Standards (RAMESES) group poses two questions in this regard: (1) can the section of data in a given document be used to contribute to theory building/testing (relevance)? and (2) were the methods used to generate this section of data credible and rigorous?

Excel was used to capture the data. Often, the nature or form of data for Contexts (here, of the hta) and Outcomes (here, health outcomes and proxy measures of uptake or implementation) are descriptive and quantitative respectively, and qualitative data allow us to identify potential Mechanisms (currently unknown, and subject to the identification of usable candidate theories). We expected this to be the case in this study, although all data were fitted in the programme theory based on their concept or meaning, not their form. Indeed, the strength of this approach is its ability to examine interventions and their outcomes in their contexts, irrespective of what form the data are in. In addition, assumptions implicit in the expression of authors about HTA may provide useful realist evidence. Thus, the study generated data which could populate the design of an explanatory framework to help optimise the implementation, and thus impact, of hta decisions.

5.8 Data synthesis and theory refinement

Analysis of all the data was linked to the programme theory. Analysis was undertaken using realist evaluation principles of extracting CMOC variables at play, and iterative, participative and collaborative approaches to interpretation. We also considered pre-conditional as well as unintended outcomes. This brought in, for example, issues around quality, uptake by whom – practitioners versus patients, and awareness and acceptance of HTA decisions as stages or pre-conditions before uptake – as well as the accumulation of intellectual, technical and social capital. Using the CMOC to identify (semi-) predictable patterns or pathways in the data, known as demi-regularities, we produced theory to explain the influence of context and mechanisms as to how hta findings are better implemented(141).

5.9 Discussion

We use a realist synthesis as a complementary and congruent approach to our quantitative framework [Chapter 4] to theorise those contextual factors and mechanisms that play a role in the last step of IDSI's ToC. In so doing, we aim to provide greater insight into how 'better decisions' translate into 'health outcomes'. As stated previously, realist synthesis is a relatively new approach

to evaluation in health and an innovative way to review the literature pertaining to HTA outcomes as this is the first time, to our knowledge, that such an approach has been applied to this field. The basis of adopting a realist approach is that we start and end with a theory, thus supposing that the final overarching theory can be translatable to other similar contexts. Our unit of analysis is HTA at the systems level, and we apply a realist lens to HTA at this by looking to theories which apply to more macro or meso institutional levels ie systemic adoption, and organisational and structural change. Such theories offer analysis of the interplay and tensions between knowledge, power and social control, the premise being that organisations do not make decisions but people with biases, motives and histories make the decisions and are required to do so within the confines of power structures (like organisations and governments). This is congruent with the philosophy of realist evaluation which recognises such tensions at different levels.

HTA should seek to harness mechanisms to improve the implementation of findings. In realism, often it is the 'resource' aspect of the mechanism that becomes central to understanding how programmes work. (210) So, for example, in triggering a response/reaction of 'motivation to enact decision', how does the HTA deliver this 'resource'? And, how does that resource manifest differently in different contexts? What are the specific types of resources that are introduced and meaningful that support implementation? The development of robust and evidence-informed products needs to be balanced with work to ensure that there are mechanisms and pathways for implementation and use of products – and without demand for these type of products from many clinicians, and indeed patients, it is unclear how and whether guidance will be implemented once developed(179). While demonstrating the value of guidelines etc may be enough, there seems to be a need to look beyond this to more formal mechanisms of implementation. As such, there is still a recognised need to have better insight into how the parties 'on the front line' - those who have to carry out policy - respond to the policies and guidance that 'come down the line' to them, as well as 'on the ground' service configuration issues. Street level bureaucracy(211) describes the role of discretion in implementation ie the role of agency in delivery. This filtering of policy can inevitably result in differences

between government policy and policy in practice. As these quotes below indicate, there is a need to get to the reasoning of stakeholders ie mechanisms in realist terms:

'Bringing about effective policy change does not simply require good technical design or using evidence to generate policy but must always involve clear attention to the processes by which change is brought about, including concern for the values and interests of the actors with potential to block or subvert policy development and implementation, and for the discourses surrounding policy change processes'(114). This suggests 'the need for policy managers to have a better understanding of the processes of policy development, including insight into the roles of stakeholders, their interests, and interactions with the health system context. However, analysis of health policy is rarely recognized and applied in developing countries' academic institutes and health administrative authorities'(114, 212).

There is much diversity in the role and application of HTA and it could take different forms depending on purpose and context. Such differences reflect not only differences between health systems and their financing but also how well-developed country-specific HTA institutions and processes are, as well as other wider contextual issues such as values and culture. In using a realist perspective, we treat the process of HTA as a complex intervention which is intended to make a change, and would/should take different forms depending on purpose and context. As HTA works differently in different contexts - context being more than locality but embodies, for example, health care systems, resources and stakeholders - and impact is likely to be achieved through different change mechanisms, it is unlikely that the same intervention (ie HTA assessment and appraisal processes) can be replicated from one context to another and achieve the same outcomes. Good understanding about what works, for whom, in what contexts, and how – are, however, portable. We will test these theories with empirical testing using country case studies. We hope this might provide a better understanding as to how HTA needs to (better?) connect and interact with 'context' in translating into health outcomes so that the use of HTA can be optimised as well as understanding the value for money of HTA.

Finally, in HICs, there has been perhaps an element of taking for granted that a fully functional system will pick up hta recommendations and adapt accordingly with little effort required to ensure implementation given our health systems and regulations. For LMICs, we hope this research will offer a forward-looking model that LMICs can point to as a reference for their own implementation. A robust analysis of the context-mechanism association can serve to innovate and transform interventions(213). For example, should elements of the context be determined to be critically important to the successful implementation of hta findings, these may be re-theorised as an intervention component. We aim to produce 'policy-friendly outputs' or guidance in terms of 'what works, where, for whom, and how' – as well as promoting improved monitoring of the adoption of hta recommendations.

6 Realist synthesis

6.1 Introduction

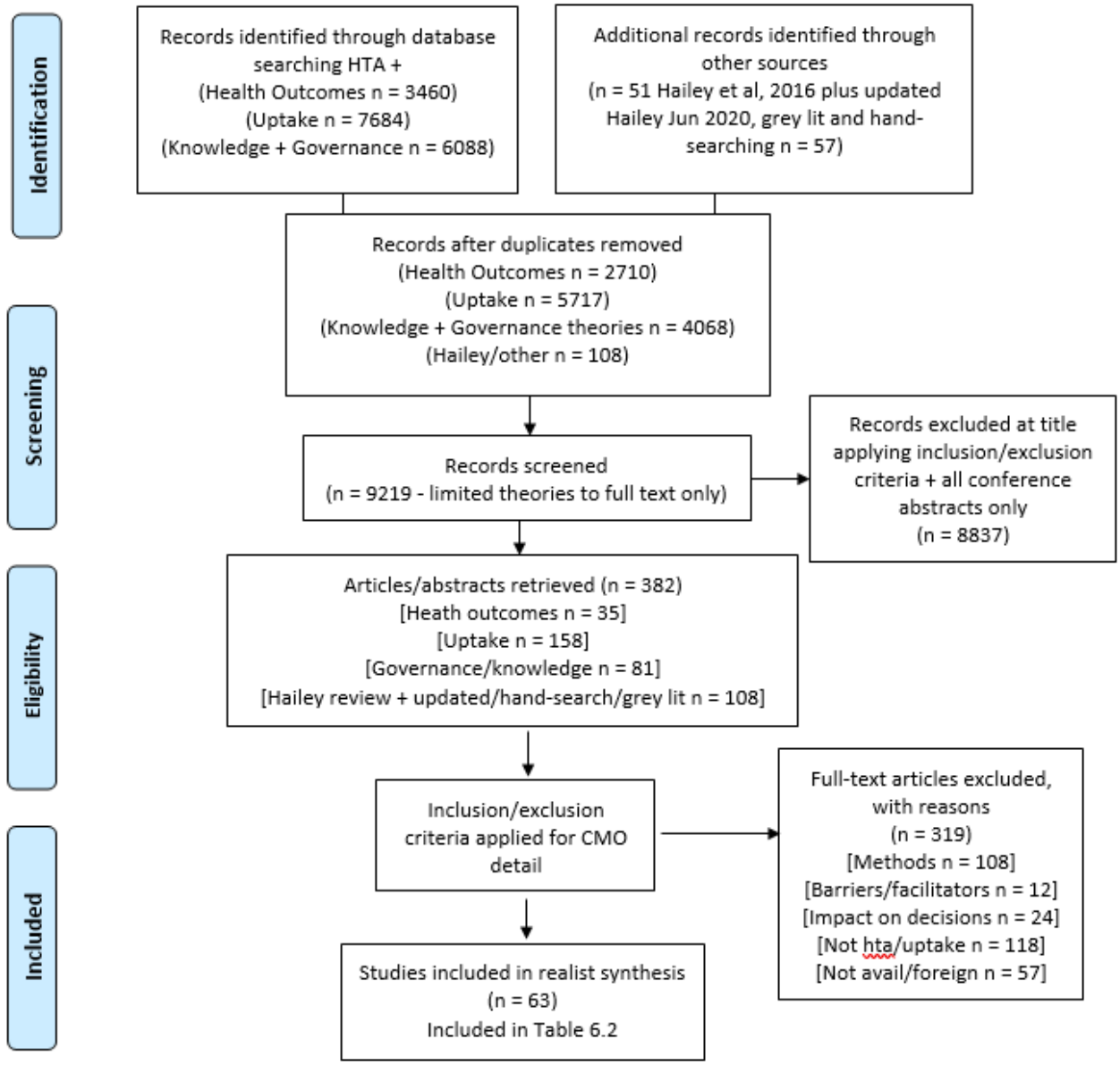
A realist synthesis focuses on the underlying mechanisms and the contexts in which they fire that lead to outcomes. To have impact, hta recommendations need to be realised which requires the opportunities and resources to enact those decisions. HTA is necessarily customised to context and, like all interventions, will introduce both new resources but also catalyse pre-existing resources(182). The question there is 'has the hta suitably harnessed the resources and overcome any deficits to ensure that implementation of its recommendations can occur?' We explore patterns of outcomes and use the evidence to develop and refine our programme theory.

6.2 Articles retrieved

The search was an iterative process across the different themes of HTA and 'health outcomes', 'uptake and implementation', 'governance and knowledge'. The results are shown in the PRISMA flow diagram [Figure 6.1] by search theme as well the updating of a review on the influence of HTA(112). The articles were retrieved from Ovid MEDLINE(R) without Revisions 1996 to November Week 3 2018 and Embase 1996 to 2018 Week 47 then updated to Ovid MEDLINE(R) without Revisions 1996 to June Week 1 2020 and Embase 1996 to 2020 Week 23. Annex B details the search terms. Grey literature searching was also undertaken and Annex C provides details on information retrieved and sources searched.

Many of the articles retrieved were useful in terms of evaluation more generally. Articles, mainly descriptive, on facilitators and barriers were found across all themes. Whilst many of these articles did not necessarily provide suitable material for CMO extraction, these data were usefully mapped to diagnose 'context' as described below.

Figure 6-1 PRISMA flow realist search



6.3 Document characteristics, data extraction and coding

6.3.1 Diagnosis of implementing context

As a first step, we mapped 'barriers and facilitators' to the implementation of hta recommendations onto the PARiHS model to diagnose important aspects of the implementation context [Table 6.1]. These barriers and facilitators were retrieved also from our initial review of the literature to identify existing frameworks or models to measure the impact of HTA [Chapter 2]. Key contextual information was extracted from articles, and mapped to the constructs of 'context', 'evidence' and 'recipient' by their sub-domains as

referred to in the PARIHS model. As failure to implement hta recommendations has been found to emerge from a combination of system (macro), organisation (meso) and individual (micro) factors, this approach to understanding context also determined what sources of data we should include such that the research remained contained.

The mapping to the PARiHS framework identified sub-dimensions under 'context' to be the most frequently cited in terms of barriers and facilitators (these often being the same in reverse) to the uptake of hta recommendations, followed by the constructs of 'evidence' and 'recipient' [Table 6.1]. From a realist perspective, 'evidence' would be considered an element of context whereas 'recipient' refers to the beliefs, attitudes, motivations and values of those implementing, and influenced by, hta recommendations. Evidence, as a context, was informed by the PARiHS model, acting as a (partial) programme theory and analytic framework. The mapping might infer that overcoming deficits in the implementing context pose the greater challenge regarding the uptake of hta decisions rather than the 'buy-in' of those implementing the recommendations. However, frequency does not necessarily equate to what matters most from a realist perspective in terms of generating outcomes. We explore this further through the CMOs.

Table 6-1 Barriers and facilitators mapped to PARIHS

Barriers and facilitators	**PARIHS mapping**
Costs of implementation, resourcing and inflexibility of budgets and resources(22, 90-92, 95).	Context - finances
Uncertainty, weak governance, political constraints(90) including the role of accountability and regulation in affecting behaviour and uptake, leadership(91).	Context – political/leadership

Poor quality of communication and dissemination(61, 91); too few mechanisms to inform; users' (mis)understanding(22, 99).	Context - culture
Learning through collaboration and exchange of experience; skills, ability(94); formal links between producers and users(42); stakeholders are involved and support decisions(61, 95-98) with front line engagement, champions(22, 80, 91).	Evidence – research/experience
Timeliness(61, 96, 99, 100).	Evidence - research
Accuracy and validity(61, 94, 99); evidence comes from a trusted source (95).	Evidence - data
Relevance(22, 94-96, 99, 100).	Evidence - data
Design of health system(90, 95) and broader organisational context(80, 94, 96, 101), including health information systems (102), payment mechanisms (90) and alignment of incentives and support from the top (61, 91).	Context – culture/leadership
Interaction at the human level. Knowledge translation (96, 102).	Recipient construct (PARiHS later version)(200)

6.3.2 Extract of articles and CMO data extraction

In addition to the articles retrieved from our search, we analysed all the studies included in a review on the influence of HTA(112) from a realist perspective and also brought this review up-to-date to June 2020. We examined the conclusions made by Hailey et al(112), extending their review by focusing on mechanisms

explaining *how* HTA is impactful. Hailey et al found few publications assessed the contribution of HTA to changing patient outcomes and, given the limited number of studies following change in clinical practice and health outcomes, suggested that these areas needed much more attention in the future. Taking a realist approach, we aimed to move beyond the useful description presented in their systematic review to interpretation, optimising the evidence about the impact of HTA by examining mechanisms and applying theory to critically appraise the literature. We focused first on those studies employing quantitative methods to measure the implementation of HTA guidance. This related back to our interlinked NHB-ROI framework by using a quantitative measure of uptake. Analysing quantitative data in a realist way often threw up limitations of the data and the reporting given the need for more qualitative detail in the realist synthesis. However, CMOs were able to be identified, especially mechanisms, from several papers [Table 6.2]. These quantitative evaluations included before-and-after study designs, ITS, retrospective longitudinal studies and prospective clinical audits to measure the implementation of interventions pre- and post-hta decisions. All studies were observational in design. The outcome measures were related to the uptake of the technology, compliance with guidance, changes in prescribing rates etc. Data were extracted and entered into Excel under the headings of ‘context’ (of decision-making), ‘mechanisms’ (reactions), ‘intermediate outcomes’ and final outcomes which we refer to here as ‘impact’. This enabled a better unpacking of longitudinal outcomes wherever possible(126). Note that Table 6.2 lists ‘reactions’ under mechanisms only; unless otherwise stated, the hta resource is always in the form of ‘evidence’ or 'information'.

Table 6-2 Extract of articles and data extraction.

Article / study details		Context [in PT, capability is part of context]			Mechanisms	Outcomes	
Article reference [setting]	**Study design**	**hta**	**Initiation context**	**Implementation context**	**Reactions [in PT, willingness is a key mechanism]**	**Intermediate**	**Impact**
Mooney (2011) [2 hospitals] (214)	Retrospective before and after	NICE guidance on emergency head imaging	.	Lack of access to 24 hours CT necessitates the adaptation and evolution of radiology service provision for this to be available in order to comply. Consultant-led publicity programme around the guidance, teaching sessions	Weigh up of risk versus benefit and parent preferences by physicians. Where both CT available 24/7 and consultant publicity undertaken, practitioners' change efficacy increases.	Reluctance to adhere to guidance in children. Formal implementation strategy increased compliance more than site with no formal strategy.	Number of CT scans increased in adults significantly but not in children.
Thavarajah (2012) [1 hospital] (215)	Prospective audit – before and after	NICE guidance on risk assessment for venous thromboembolism	Some issues with the guidance itself / sometimes impractical		Distrust / scepticism / compromised autonomy		Proportion of patients assessed fell short
Bennie (2012) [Scotland] (216)	Longitudinal	Prescribing efficiency of proton-pump inhibitors and statins		Quality & Outcomes Framework targets + demand side measures - education, economics, engineering, enforcement	Cumulative effect on positive influence		Expenditure and number of drugs
Goeree (2006) [Canada provincial health system] (217)	Analysis of admin data	Stents	Relevant, responsive and dynamic process as opposed to a static report - handles uncertainty by turning the process into an iterative data collection,		Enhances credibility and trust as findings are more responsive to needs	Acceptance of recommendations	Reduced time lag between knowledge translation and uptake

			updating and knowledge translation process.				
Article / study details		**Context**			**Mechanisms**	**Outcomes**	
Reference [setting]	**Study design**	**'hta'**	**Initiation context**	**Implementation context**	**Reactions [in PT, willingness is a key mechanism]**	**Intermediate**	**Impact**
Schluessman (2009) [Switzerland national coverage] (218)	Analysis of registry data	Arthroplasty		Establishment of a registry to follow patient outcomes	Creates new attitudes and motivation for HTA through the generation of evidence on outcomes	Transparency / evidence generation	Coverage
Jonsson (2001) [Swedish public hospital] (219)	Analysis of admin data	Screening (various)	Strong cultural acceptance of HTA.	Emphasis on dissemination to utilise results and budgets for this. Believe that dissemination would be enhanced further with links to quality assurance activities too.	Generate greater influence.	Greater utilisation of results.	Decisions and practice patterns following HTA recommendation
Zechmeister (2012) [Austria public health care system] (220)	Analysis of admin data from hospital and health insurance funds	69 hta's		No legally binding mechanisms to use results. Media reporting plays a role.	Trigger for change motivation is as likely to be as much due to the admin/org changes as acceptance of the HTA findings – impact lies at organisational level of health system rather than with the individual.		Cost savings
Bennie (2011) [Scotland] (221)	Analysis of admin data – before and	Medicines rejected by SMC	Delays between medicine launch and initial SMC advice, the publication of conflicting		Complex relationship between HTA advice and clinical	Variable adherence,	Drug volume/prescribing not always showing change in clinical

	after study design		advice from different national bodies. Failure to engage with relevant clinical experts early in the medicine review process.		practice change – inconsistent and conflicting professional advice causing confusion to prescribers		practice following HTA recommendation
Dietrich (2009) [UK] (222)	Analysis of prescribing costs	Medicines rejected by NICE	Lack of education of patients and physicians as to rationale of guidance.	Non compliance not sanctioned. Lack of incentives rather than penalties to adhere.	No incentive or awareness to change practice.	Non acceptance	Prescriptions dispensed (mostly no change)
Rosen (2014) [Sweden] (223)	Documents, before-after surveys and time series register data.	26 hta's	Involvement of experts from early on - effectiveness depends on HTA agency's trustworthiness.	Ability to disseminate information and implementation strategies.	Confidence in HTA body and evidence via colleagues' involvement	Conviction of reliability of the HTA findings	Changes in use and clinical practice

Article / study details		Context			Mechanisms	Outcomes	
Reference [setting]	**Study design**	**'hta'**	**Initiation context**	**Implementation context**	**Reactions [in PT, willingness is a key mechanism]**	**Intermediate**	**Impact**
Platt(224) [UK]	Retrospective multi-site audit	Childhood UTI NICE guidance		Inadequate IT systems – flexible computerised clinical systems needed to support implementation	Complexity of guidance leading to subjective interpretation	Variability in compliance	Under-diagnosis
Sheldon [UK](135)	ITS, audit and interviews	20 NHS hospital trusts	Identification of a lead clinician early in the NICE guidance development and involvement of clinicians in the guideline development process	Implementing deficit emerges from a combination of system, organisational and individual factors.	A commitment to a process for implementing guidance brought about by trust in the evidence.		Trusts with high compliance to NICE guidelines had these common characteristics subject to contextual deficits.
RAND [UK](19)	Economic Net Health Benefits	NIHR research	Lack of practitioners' involvement Capacity/expertise insufficiently accounted for	Implementation costs not considered in UK health system	Existing views and habits of practitioners, personal experience / patient preference not overcome	Poor understanding of research findings	Delays/waiting times in implementation
NICE CVD impact [UK] (225)	Before/after	NICE guidance		Lack of identification of appropriate patients / lack of local referral pathways triggered pro-active networking at local level (via community groups/charities/pharmacy led)	Mechanisms triggered to identify those at risk / local pathways established and/or supported - systems wide approaches	Increase in those who can be diagnosed	Increased uptake
NICE Cancer impact [UK] (226) [this is a series of NICE reports].	Before/after	NICE guidance	Not always timed with marketing authorisation as well in UK as in other countries but changes now to produce earlier draft guidance		Demand generated for such new drugs is 'always' high/exists by patients and clinicians' needs/ "cancer alliances"	Uptake slower than in other countries	Increased uptake

Article / study details		Context			Mechanisms	Outcomes	
Reference [setting]	**Study design**	**'hta'**	**Initiation context**	**Implementation context**	**Reactions [in PT, willingness is a key mechanism]**	**Intermediate**	**Impact**
NICE Maternity impact [UK] (227)	Before/after	NICE guidance		Additional resources such as training, new equipment or pathway reconfiguration required.	Bespoke action plans and support packages / audits (using charities /NHS partnerships / systems partners) to educate/inform/reach out to patients	Patients at risk informed	
Thai HITAP [Thailand](212)	Interviews, review of documentation	Cervical cancer screening		Financial support provided and good infrastructure existed.	Political motivation alienated clinical experts and health professionals in the process of scaling up implementation	Stakeholders' interests and interactions with health system not harmonised.	Top down driven implementation failed
Rosen [Sweden](223)	Before/after, time series		A context where confidence in the HTA agency is high and engaging clinicians is key.		Confidence/trust among health professionals. Confidence in the HTA agencies' evidence creates value among leading healthcare professionals.		Change in clinical practice
Britton [Sweden](228)	7 SBU reports	Various	Collaboration is in evidence.		Ownership is generated.	Generally strong influence	SBU reports on moderately elevated blood pressure, use of neuroleptics, stomach pain, smoking cessation, and preoperative routines have had an impact on clinical practice.

Article / study details		Context			Mechanisms	Outcomes	
Reference [setting]	**Study design**	**'hta'**	**Initiation context**	**Implementation context**	**Reactions [in PT, willingness is a key mechanism]**	**Intermediate**	**Impact**
Axelsson [Sweden](229)	Survey	SBU guidance on tobacco prevention	Accessible, translatable guidance		Awareness and knowledge increased of practitioners	Clinical practice changed	Positive outcomes on smoking cessation but to what extent is this attributable to the guidance?
Solans(230) [Spain]	Interviews	COPD		The barriers and facilitators identified were mostly organisational - in research management, and clinical and healthcare practice	Good collaboration led to increased credibility – with relationships linked to facilitators and barriers	Increased funding (for research, implementation?)	Positive change in healthcare practice
Hanney [UK - Research](231)	Payback - interviews	Various incl. NICE Technology Assessment Reports	Relevance and policy collaboration		Generated policy customers buy in		Programme had perceived impact on policy and to some extent on practice.
Guthrie [UK - Research](68)	Payback – interviews, case studies	Various	Involvement of clinicians in clinical trials		Increased clinical skills	Primary route to impact of programme-funded research on patients is through influence on guidelines.	Implementation of guidance not explored
Schuller, T [Various INAHTA member organisations] (232)	Narrative 'stories'	Various	Early on stakeholder engagement – range of stakeholders including media, politicians, lawyers (+ve) Broadening of scope (+ve) The management and organizational culture do	Quality monitoring (+ve) HTA process and information are totally separated from budgeting and purchasing (-ve)	Difficult to improve adherence to recommendations as process underlying scientific evidence is questioned where lack of (+ve).		A risk that HTA fails to exert any impact on real life decision making.

			not support the use of HTA knowledge as an essential part of decision processes (-ve)		Capability constraints.		
Article / study details		**Context**			**Mechanisms**	**Outcomes**	
Reference [setting]	**Study design**	**'hta'**	**Initiation context**	**Implementation context**	**Reactions [in PT, willingness is a key mechanism]**	**Intermediate**	**Impact**
Frønsdal, K(233) [HTAi Policy Forum]	Descriptive	Various	Whole system consideration - health system structure, tasks (function), people (staff, attitudes, and knowledge) and technology.	Whole system consideration - health system structure, tasks (function), people (staff, attitudes, and knowledge) and technology.	Data from monitoring activities informs practitioners on optimal use by adding to the evidence base. Methods to take into account organisational issues or expand breadth to do this.	Evidence generated on use and contextual barriers overcome using whole systems approach.	Augments uptake but questions whether HTAs take sufficient account of these "organizational issues" as routine.
Lowson, K et al, 2015(234) [NHS Trusts]	Cross-sectional survey	Interventional procedures	Organisational processes, clinician engagement and exist as barriers.	Financing of new procedures	Lack of buy-in and ability (time and resources)		Non implementation
Stemerding D, 2001 (235) [Netherlands]	Descriptive	New serum screening in maternity care	HTA does not incorporate the ethical, social, cultural and political dimensions of health technology.		By stimulating the interplay between different perspectives and preferences among parties involved, social learning could be promoted.	Tension between political and clinical motivations unresolved.	Unsatisfactory outcome to all parties of neither scale up/or scale down.
Gagnon MP, 2006 (80) [Catalonia]	Semi-structured interviews	Insulin pump, cataract surgery, hip/knee replacement	Hospital-based		Considering the dynamics between organisational and human factors generates more meaningful involvement in hta.	Greater acceptance	Greater uptake

Article / study details		Context			Mechanisms	Outcomes	
Reference [setting]	**Study design**	**'hta'**	**Initiation context**	**Implementation context**	**Reactions [in PT, willingness is a key mechanism]**	**Intermediate**	**Impact**
Teerawattananon Y, 2014(236) [Thailand]	Descriptive	National List of Essential Medicines	Strong HTA processes and infrastructure exist		Enforcement of NLEM means obligation to provide/implement		Medicines are made available by law
Corbacho B, 2019 (237) [Spain, England]	Regression on trends 3 months post-guidance	Cancer drugs (NICE, GENESIS)	Number of external factors – not just the guidance in isolation		The more local the decisions, the stronger the adoption – practitioners perceptions key.		+ve recommendations equated to increased uptake but -ve recommendations did not show drop in usage
Cummings G, 2007 (238)	Cross-sectional census	Nursing		Positive culture, strong leadership, evaluation feedback are predictors of greater staffing support and development	This enhances practical support & encourages autonomy.	Greater research utilisation	Better quality care
Dent, T 2002 (239) [UK]	Descriptive	Various	3 years after NICE set up.	Lack of mechanisms for evaluation to feedback and support performance	Need for being informed on what is working and how in clinical process.	Increased capacity by users to feel ownership	Greater uptake of NICE guidance
Fronsdal K, 2010(233)[Multiple countries]	HTAi workshop	Various		Organisational and whole systems approach not incorporated into HTAs – lack of implementation plans	But regulation, funding, monitoring can still augment the evidence base through knowledge generation and ongoing feedback to users	Greater understanding of risks, what works etc	Refine technology utilisation

Poder T, 2018 (240)	Interviews	Various – 20 reports	Complexity of changes on administrative system ignored tying back to HTA process and stakeholder involvement	Complexity of changes on administrative system ignored tying back to HTA process and stakeholder involvement	Validation not generated early on in the process	Discordance between guidance and institutions	Non-implementation of HTA guidance
Sihvo S, 2017(241) [Finland]	Brief survey	Managed uptake of medical methods		HTA process, budgeting and purchasing separate.	Organisational leaders can generate influence and monitoring activities an appreciation by users	HTA importance better acknowledged	Uptake?
Haslam D, 2007(242) – see all plus Chidgey J 2007 (243) [UK]	Descriptive	Various	Uptake is dependent on the HTA process and clinical support		This generates clinical confidence in guidance – but alone not enough without organisational change and support		Change in clinical practice variable
Gallego G, 2011(244) [Australia]	Semi-structured interviews	Medical devices	Fragmented, complex health system		Uptake is dependent on the HTA process but only a facilitator – other factors including context and clinicians' perceptions		Variable uptake
Hastings J, 2006 (245) [Multiple countries]	Survey	PET device		Funding linked to continuous quality improvement	Additional knowledge and monitoring data generated to inform ongoing evidence base		Improved uptake
Rosen R, 1999 (246)	Descriptive	Various	Economic and organisation impact needs assessed = whole system assessment.	Process of hta to implementation is not linear.	Generates openness and interest in guidance	Better preparedness to implement.	A receptive organisational context for change.
Barosi G, 2006 (247) [UK, US]	Review	Clinical guidelines	Guideline-cycle to address current service organisation, beliefs of staff, monitoring		Ability and willingness of users generated where economic and professional resources supplied	Compliance	Improved patient outcomes

Article / study details		Context			Mechanisms	Outcomes	
Reference [setting]	**Study design**	**'hta'**	**Initiation context**	**Implementation context**	**Reactions [in PT, willingness is a key mechanism]**	**Intermediate**	**Impact**
Rosenkotter N, 2011(248)	Review	Genetic tests	Organisational analysis of the technology's interaction with health care delivery and services undertaken plus assessing socio-dynamics of the technology	Organisational analysis of the technology's interaction with health care delivery and services undertaken plus assessing socio-dynamics of the technology	Acceptability and resource feasibility enhanced	Broader and earlier hta shape technology development	Implementation is facilitated
Cacciatore P, 2020 (249) [INAHTA members]	Review	HTA reports		Organisational aspect plays key role in implementation but under-represented in HTA – this, in turn, influences behaviour of users	Generates better understanding of the resources required to be mobilised and co-ordinated following an HTA recommendation and interconnects to user behaviour	User acceptance and capability to implement enhanced	Improved implementation
Owen-Smith A, 2010 (92) [UK]	Qualitative	Guidance in clinical practice	Various settings	Practicalities of funding	Autonomy compromised by rationing	Use is limited	Variable uptake
Hivon M, 2005 (250) [Canada] and Lehoux P, 2009 (251) and Tucker S, 2019 (252)	Qualitative	Case studies		Organisational and material limitations hinder implementation	Shared responsibility / shared commitment between producers and users of HTA leads to greater absorptive capacity	Greater influence over barriers by users	Increased uptake
Pettersson B, 2012 (253) [Sweden]	Segmented regression time series	Lipid modifying therapies		New reimbursement programme. initiatives: education, engineering (organisational), economic and enforcement	External demand created	Improved adherence	Change in prescribing rates
Campillo-Artero C, 2012 (254) and Tucker S, 2019	Descriptive	Various	Behaviours, systems, procedures, resources, IT, evaluation all need	Behaviours, systems, procedures, resources, IT, evaluation all need	Investment in change strategies / multi-faceted strategies	Changed behaviour and	Increased implementation

(252) and Doherty S, 2006 (255) and Mayer J, 2019 (256)			addressed in a culturally appropriate context	addressed in a culturally appropriate context	generate change required across all levels/domains – addressing both capabilities and motivations simultaneously	environmental factors	
Article / study details		**Context**			**Mechanisms**	**Outcomes**	
Reference [setting]	**Study design**	**'hta'**	**Initiation context**	**Implementation context**	**Reactions [in PT, willingness is a key mechanism]**	**Intermediate**	**Impact**
Tian L, 2017 (257)	Before/after survey and interviews	Nursing	Examined facilitators and barriers using PARiHS - both practical resources as well as staff receptiveness to change need addressed	Practical resources	Evidence not enough of itself. Strong leadership generates acceptance, confidence and competence.	Context conducive to change	Variable compliance
Esmail R, 2018 (87)	Theory based	Disinvestment, de-implementation		Usual contextual factors apply as to implementation	Knowledge translation theory to support change as more psychologically challenging to stop rather than take up a technology		Variable change
Saunders H, 2016 (258) [Finland]	Cross-sectional descriptive design	Nursing	Individual and organisational readiness explored	Individual and organisational readiness explored	Self-efficacy needs to be fostered through strong leadership and knowledge	Confidence and self-efficacy lacking within staff	Evidence-based not implemented
Taxman F, 2014 (259)	Trials	Social services/justice	Underlying social structure of whole organisation addressed to affirm value of interventions to users in a whole systems approach	Underlying social structure of whole organisation addressed to affirm value of interventions to users in a whole systems approach	Education/social networking builds expertise and resilience in staff thus improves organisational readiness to implement – less about skills, more about values	Staff and leadership support implementation	Improved receptivity to change given more supportive and committed environment

Article / study details		Context			Mechanisms	Outcomes	
Reference [setting]	**Study design**	**'hta'**	**Initiation context**	**Implementation context**	**Reactions [in PT, willingness is a key mechanism]**	**Intermediate**	**Impact**
Stander J, 2018 (260) [South Africa]	Systematic scoping review	Physiotherapy	Local context and behaviour studied	Local context and behaviour studied	Opinion leaders generate consensus for change – less about knowledge and skills		Barriers in the context overcome
Grenon X, 2016 (261) [Switzerland]	Semi-structured interviews	Hospital-based HTAs	, clinicians' preferences key	Technical difficulties & lack of infrastructure	Compliance was sometimes compromised due to reduced capability and willingness respectively		Variable adherence to guidance
Billsten J, 2018 (262) [Sweden]	Naturalistic design, questionnaires	Psychiatric services		Greater institutional resources and training needs explored	Organisational readiness for change is generated via employees perceived need for change	Organisation readiness to change is a predictor only for implementation	Institutional support and relevance of intervention also required for implementation
Uneke C, 2017 (263) [Nigeria]	Cross-sectional quantitative study	Maternal health		Organisational and individual capacity constraints	IT structures are key to provide support, quality and an enabling environment – lack of organisational enabling environment has a knock-on effect on individuals' motivation	Staff demotivation	Inability to apply evidence-based practices
Sasaki N, 2019 (264) [Japan]	Survey	CPs		IT infrastructure key	Enables accessibility to latest evidence and to disseminate this increasing usage of CPs as less time spent searching for medical info. Information and	CP use promoted	Narrowing of the evidence-practice gap

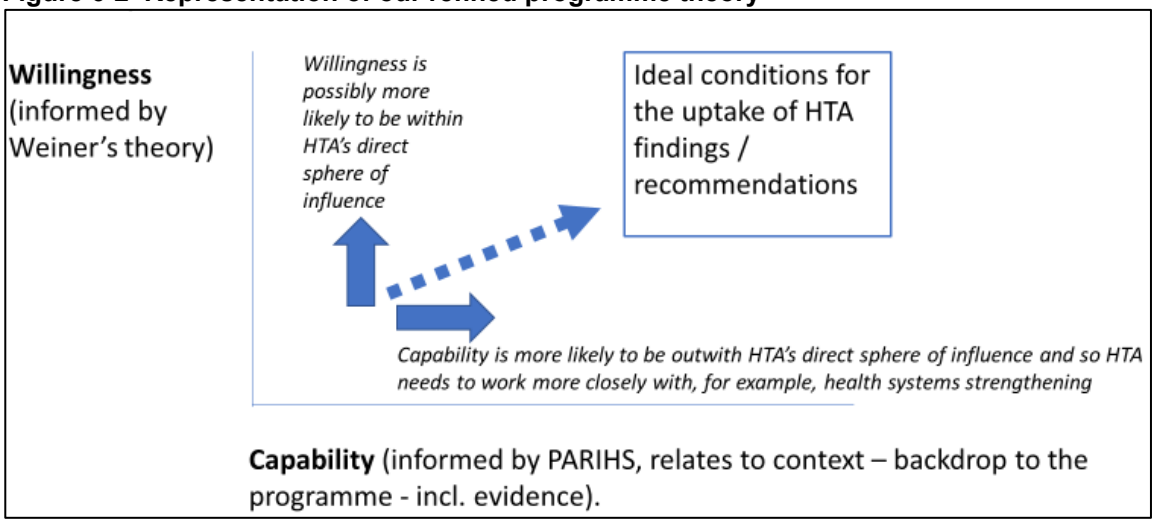
					communication promote autonomy.		
Jolliffe L, 2019 (265) [Australia]	Before and after study	Brain injury rehabilitation CPs	Strong organisational supportive management and leadership culture	Strong organisational supportive management and leadership culture	Feedback and audit empower staff by creating shared goals		Increased adherence
Grove A, 2018 (266) [UK]	Case studies	Orthopaedics	Increased regulatory power .	Targeting users will better enable them to communicate findings via the channels they prefer	Undermines autonomy of users. Greater understanding of how knowledge and evidence accumulated	Micro, meso, macro interlinked as lived experience (micro), organisational constrains (meso) and regulation (macro initiatives) influence uptake.	Variable adherence
Wathen B, 2004 (267) [UK]	Before-and-after + interviews	Drug prescribing		NICE guidance taken in isolation	GP / patient experience influences prescribing through own understanding of risks. Greater transparency and consideration of budget required to sustain any effect.		Unsustainable uptake of prescribing.

6.4 Main findings

6.4.1 Programme Theory

The findings presented below are articulated in the context of the initial programme theory which is used to refine our theory into a working model of how uptake occurs successfully. We employ our programme theory, combining the two formal theories identified, PARIHS and Weiner's theories, to identify and interpret the CMOC. Using similar depiction and language as in the PARIHS framework, we presented this visually to depict our drawn model of theory [Figure 6.2]. Readiness for change is a complex multi-dimensional construct(196) requiring both a willingness and capability to change. These constructs are noted along the X-Y axis, with HTA shown as a direct or indirect facilitator of each. We considered how the CMOs are upheld or challenged by the programme theory. We themed findings by 'capability' and 'willingness', and presented CMOs under each.

Figure 6-2 Representation of our refined programme theory



6.4.2 Willingness

Practitioners' autonomy – CMOs

In 'riskier' or less straightforward situations(214) [C], clinicians are more likely to apply their own rationale or judgement [M] despite complying with guidance at other times and in other populations (for example, in children) resulting in non (consistent) adherence to HTA guidance [O](224, 267). In

situations where the treatment had not resulted in any adverse events or side effects for a period of time [C], physicians may still choose to maintain the original treatment regimen [O](19). The point here being that there is variation on the part of the clinician to use discretion (M). Where there is seen to be no direct evidence of harm [C], although treatment might not be cost effective, there is no change in practice [O] due to the existing habits, views, personal experience and patient preferences [M](19, 214, 221, 222, 224, 237, 244, 267). This is also likely a consequence too of poor understanding of the guidance and/or that the guidance is not suitable in scope to be applied across differences in patient behaviours(215, 221, 224, 232).

Where non-compliance is not sanctioned [C], health professionals are only 'expected' to take guidance into account together with experience and other sources of evidence leading to variable adherence [O](220, 222) – as seen above, subject to the perceived importance/value placed on the guidance as necessary to inform practice [M]. However, incentives rather than penalties [M] were thought to better foster the implementation of guidelines(266) (relates to Weiner's theory on 'want' versus 'ought') together with information at the point of prescribing (for example, providing information about results of cost-effectiveness could be included in electronic prescribing support systems). When hta encompasses an iterative data collection process ie the ongoing generation of evidence and knowledge [C], it enhances acceptance of, and credibility in, the hta [O] through increased motivation of practitioners in their understanding of what is working (217, 218, 233, 239, 245). Education programmes communicating the medical and economic rationale of the guidance can foster the patients' and physicians' acceptance [O](222), especially when they are provided from an independent institution [C] (222) or consultant-led(214) as the credibility of source is taken into account by clinicians when 'deciding' on its implementation [M].

Clinician engagement promotes confidence in HTA - CMOs

Failure to collaborate and engage with clinicians early in the hta comes up time and time again in relation to both negative(221) and positive

recommendations(19, 68, 135, 223, 228, 230, 232, 240). Failure to engage even impacts negatively where capacity in the form of resources and infrastructure are strong (C) (for example, in Thailand with the scaling up of cervical cancer screening(212)). When clinicians have been involved in the production of guidance via an expert advisor(223) or guidance consultee role for NICE [C](135), they and their colleagues are likely to be more engaged (M) with implementing the guidance. More generally, identifying the lead clinician responsible for implementation of guidance prior to guidance publication was reported to aid in clinician engagement(135). However, it was not reported *why* prior notification aids engagement but, given previous reports of time constraints, it is plausible that knowing about the guidance in advance allows for resources to be better planned. In the case of negative decisions, omittance to consult early on with clinicians [C] led to inconsistent professional advice [M] resulting in confusion by prescribers [O](221). Furthermore, the most important means found by one agency (223) for implementing its findings, is the involvement of experts in its project groups. Attending meetings for 2–3 years, reading and assessing relevant articles and reaching agreement on findings and conclusions all serve to create effective advocates for the final report. So, involvement in development of guidance (C) means practitioners have increased ownership and attachment (M) and engagement (M)(234).

Drawing on Weiner's theory of organisational readiness as to how to interpret and articulate all this as a CMOC: Forewarning of the intervention (C) allows planning of resources and leadership (C), which in turn triggers ownership, attachment, and acceptability (M) of the recommendations and increases motivation to implement (O). This is perhaps because an intervention that is implemented in a well-resourced context fosters confidence (including consistency of message, as per above) and encourages staff to get behind it. Moreover, healthcare professionals are more likely to change their behaviour if their colleagues are convinced of the reliability of findings. Rather than this being a practicality about time constraints, we would suggest that where such engagement is carried out [C], legitimacy in the findings is instilled as well as peer trust-building amongst practitioners in the hta [M] leading to

more successful uptake of guidance [O]. It enables practitioners to feel a change is appropriate, as the critical mass reinforces its appropriateness and value. Weiner's theory would suggest that when readiness for change is high ie with high levels of collective change efficacy and commitment, practitioners are more likely to initiate change, exert greater effort, exhibit greater persistence in the face of obstacles or setbacks and display more cooperative behaviour. The key point here is about whole team change, or a climate of agility resulting in more effective implementation(135, 196, 223, 232, 268). Practical means found in the literature used to increase confidence/trust etc include early engagement and external validation (in some contexts).

6.4.3 Capability

Supportive measures must span individual- and system-level- CMOs

Although analysis of routine health administrative data showed clear positive changes in uptake in many instances [O], stakeholders felt that where organisational and administrative changes are triggered and/or justified by hta reports [C], this is more likely to have been the influencing factor rather than health professionals' awareness or acceptance of the htas [M] (220). Contextual change in terms of additional organisational and implementation support accounts for variance in uptake just as much – or more than(220) - professional buy-in (19, 135, 220, 223, 230, 232-234, 240, 242-244, 246-252, 254-256, 261). Several studies identified robust IT systems within the healthcare system as being a key step to improving implementation of guidance (224, 263, 264). Here, HTA is possibly more influential at a (healthcare) systems level rather than at the individual level with IT being an important tangible aspect of an organisation's infrastructure that must be considered; as readiness is gauged, the influence of IT support over-rides the influence of (lack of) individual preference/agency/action.

Although HTA provides resources in terms of guidance, knowledge and sometimes training, additional supportive resources including data management systems(264) and organisational wide processes(252, 254-256)

were often lacking resulting in variable implementation. Similar to the above, this again indicates that knowledge/information-based resources held at the individual level were not enough to produce uptake: context-level resources were required to augment the knowledge and information. A shared responsibility/collaboration between producers and users of hta in addressing the wider causes of, for example, over/under- diagnosis or prescribing, would promote a greater harmonisation of interests among stakeholders across multiple levels of the system [C](230), triggering motivation [M](263) and greater acceptance of the guidance [M](248, 249) and ultimately, greater absorptive capacity(250-252). Supportive measures are seen to be not just about providing capacity in terms of resources; it enables practitioners to feel capable of making the change successful, with greater preparedness and more influence to overcome barriers ie collective change efficacy (M)(230, 246, 249-252). Monitoring systems [C] also reinforced this change efficacy [M] for both practical reasons around the ease of information retrieval but also by continually augmenting the evidence-base(233, 247, 263, 264). Where the process of compliance is turned into an iterative data collection and knowledge translation process(217) [C], there is greater acceptance of recommendations [O] given its perceived relevance and responsiveness by clinicians [M]. Similarly, where monitoring is undertaken [C], this in itself increases coverage [O] as it creates new positive attitudes [M](218).

Furthermore, evidence shows where demand-side measures co-exist in addition to guidance [C], prescribing increases [O] due to a combination of education, economic incentives and enforcement [M](221). Similarly, where there is emphasis on dissemination and budgets for implementation plus links to quality assurance activities (C], implementation of services is likely to face less barriers [O] due to a combination of practical support and enforcement [M] (219). However, the latter example was found in preventative services which may have more influence than other clinical measures. This embellishes the argument about resources being necessary from different levels to leverage collective support to act.

This is again underpinned by Weiner's organisational theory (196) on critical mass at the organisational rather than individual level where the effect is greater when there is shared resolve and a common understanding and valence of change. Weiner's theory is at the supra-individual levels of analysis because many promising approaches to improving healthcare delivery entail collective behaviour change in the form of systems redesign - that is, multiple, simultaneous changes in staffing, work flow, decision making, communication, and reward systems. It depends on collective, coordinated behaviour change. The greater the degree of interdependence in change processes and outcomes, the greater the utility of supra-individual theories of readiness(196). It would appear that where HTA is accompanied by joined up systems-wide approaches (IT, monitoring and administrative structures, funding etc) (C), the change efficacy and change commitment of practitioners is increased (M) leading to their adherence of evidence-based guidance [O]. The above all argue for a complexity-informed approach ie the need to operate across individual, systems/process, and organisational level.

6.4.4 Drawing lines between CMOs and the Programme Theory

Our starting premise is that everything we are focusing on here is a step on from HTA's impact on decision-making. HTA is an input to decision-making but "XYZ" is required for implementation of that decision; HTA supports decisions but not (directly) the implementation of them. This is why this work extends the reach of HTA by an appreciation of the context ('capability' being part of context) and individuals' drivers (willingness, scope in role, change efficacy etc). These can also reinforce or strengthen one another, and thus better transfer decisions into actions. Our focus is, therefore, on specifically looking for mechanisms that link decisions to actions (outcomes). In addition to looking for mechanisms that link decisions to actions, we also consider the characteristics of the recommendations. We can categorise an 'hta' into 3 types(32) [see Chapter 1]:

- Micro: appraisal of individual technologies (or related groups thereof)

- Meso: clinical practice guidelines to manage patient care pathways within a healthcare system

- Macro: efficiency, organisation and strengthening of the healthcare system

The CMOC are tabulated below to demonstrate the links and hypotheses. We also themed these by 'willingness' and 'capability' and indicate where there is an interaction, leading generally to a positive change in implementation. Regarding saturation of the literature, we were finding that mechanisms identified could be broadly placed under our theory around willingness and capability but could be continually modified by new contexts. Finally, we show at which level the hta and any supportive measures were operational [Table 6.3].

Table 6-3 Linking CMOs and Programme Theory

Theme	Context	Mechanism	Outcome	Level	References
Willingness* capability	Risky situations – and also the governance/accountability environment in which they are in	Clinicians apply own rationale – and also may not feel they can make these decisions if context restricts how supportive they feel their infrastructure/ environment is	Non-adherence to HTA	Meso	(214, 232, 267)
Willingness* capability	Evolving alignment of healthcare system practicalities combined with stakeholder-led publicity.	Change efficacy of practitioners increased	Significant increase in compliance	Meso + Macro	(135, 214, 227, 233)
Willingness	Long term treatments	Poor patient acceptability and understanding of	Non acceptance/ potential for medical disputes	Micro	(19, 222, 267)

		rationale for change, preferences			
Willingness	No evidence of harm	Lack of perceived incentive and motivation/perception of 'need'	No change in practice	Micro/meso	(19, 232, 262)
Willingness	Non-compliance not sanctioned	Guidance adds little value to existing knowledge	Variable adherence	Micro /Meso	(220, 222)
Willingness* capability	Organisational buy-in/ convinced of independent, evidence-based need: Education communicating the guidance origins	Clinician acceptance of guidance.	Increased adherence	Meso + Macro	(238, 246)
Willingness	Relevant, responsive and dynamic / monitoring processes	Credibility and trust as findings are more	Reduced time lag between knowledge translation and uptake	Micro / meso	(217, 218, 245)

		responsive and relevant to needs – on an ongoing basis			
Willingness	Low trust settings require external validation (conversely, in existing high trust settings of HTA)	Guidance has perceived legitimacy and credibility	Better communication between patient/practitioner/consultant on patient care	Meso	(214, 222, 240)
Willingness	Co-production of guidance and engagement with end-user (clinician)	Increased ownership / Increased trust	Better implementation/ successful uptake	Meso	(230, 250-252)
Capability * willingness	Forewarning of guidance with capacity to plan organisationally	Increased acceptability as perceived increase in feasibility of enacting upon findings	Change efficacy and commitment of practitioners increases	Meso + Macro	(80, 240, 248, 249)

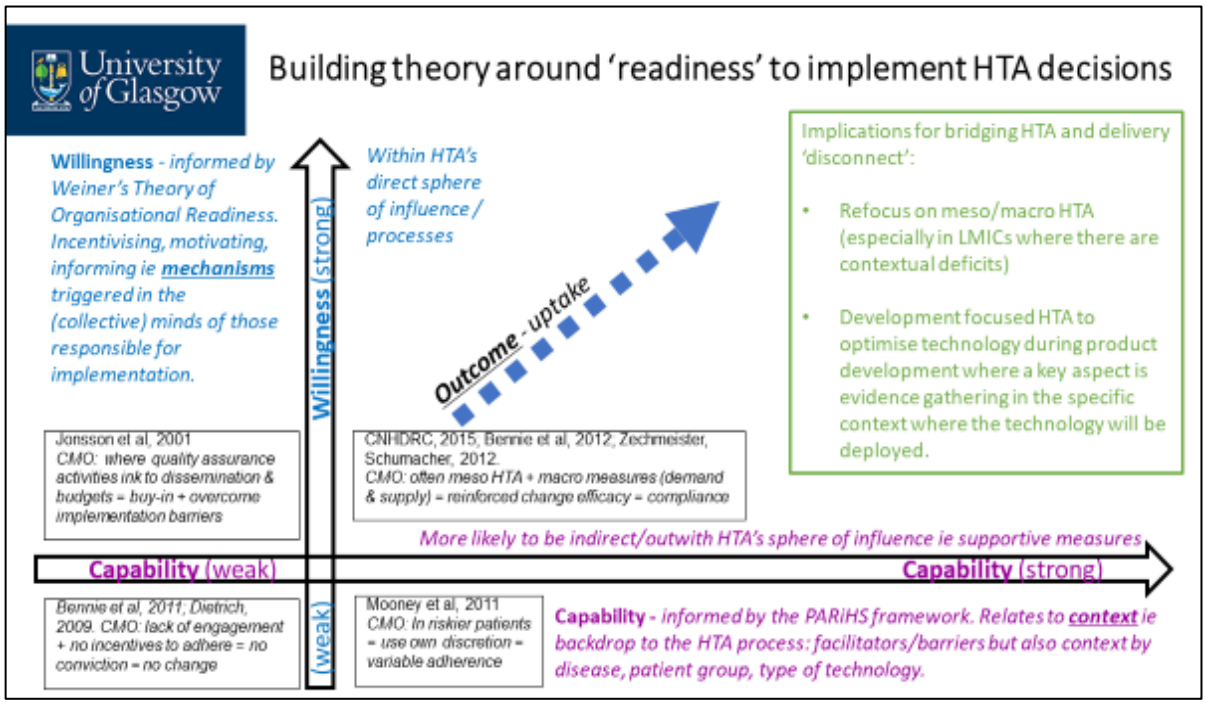
Willingness	Lack of early consultation with clinical experts (conversely involvement of experts from early on)	Inconsistent professional advice (confidence in HTA body and evidence)	Confusion by prescribers / distrust / scepticism (conviction in reliability of the HTA)	Micro	(19, 135, 215, 221, 223, 232, 234, 266)
Capability * willingness	IT / infrastructure reforms / systems-wide support integrated with stakeholder engagement (or a recognised need for)	Alignment with clinical needs and organisational systems	Change efficacy and change commitment of practitioners increases	Meso + Macro	(19, 135, 224, 227, 232-234, 240, 242, 244, 246-249, 252, 254, 256, 261, 263, 264)
Willingness	Non - identification/address of problem (despite good infrastructure)	Non harmonisation of interests, motivation	Non-acceptance of guidance	Meso	(212)

Capability	No legally binding mechanisms to use HTA results. Organisational changes in the system triggered by HTA (converse to the above)	Increased motivation to use HTA resulting from structural changes (rather than acceptance of HTA findings)	Health systems strengthening	Macro	(220)
Capability * Willingness	Demand-side measures co-exist with HTA guidance – organisational readiness needs to be supplemented by uptake on the patient side	Education, economic incentives, enforcement	Increased reach / coverage	Meso	(216, 253)
Capability	Resources - budgets, quality assurance etc	Combination of practical support and enforcement	Fewer barriers to implementation	Macro	(24, 232, 245)
Willingness	Compliance process becomes data collection and KE / monitoring	Increased perceived relevance and responsiveness / new	Greater acceptance of recommendations	Meso/macro	(217, 218, 238, 239, 245, 247)

		attitudes through the generation of evidence			
Capability * Willingness	Strong leadership / cultural acceptance of HTA and support for dissemination of HTA	Increased influence	Increased utilisation of results	Meso & Macro	(219, 232, 238, 241, 257, 258, 260, 265)

Finally, for ease of visual presentation, we adapt Figure 6.2 by plotting here some papers from the realist synthesis as to where they fit regarding the interplay between contextual elements and mechanisms of action. When presenting findings, rather than extensive lists of CMOs, we have presented our programme theory diagrammatically as shown into a 'mid-level of abstraction'.

Figure 6-3 Capability & Willingness - Context & Mechanisms



6.5 Discussion

The NIHR HTA Programme and NICE's guidelines have, according to the authors of an impact study, neglected the final common pathway, namely the HTA and implementation interaction(269). 'The NIHR HTA Programme strives to provide the NHS with the best possible evidence. NICE strives to provide it with the best possible guidance. The interactions of these processes seem to work. What is missing is information on how new evidence or guidance impacts on patient care and what factors influence the implementation of clinical guidelines. These relationships are not nearly so well understood and could be rich areas for future research'(269). We have attempted to explore this through a realist inquiry.

We started by mapping implementing barriers and facilitators to the PARIHS framework. This showed that most facilitators are to do with evidence, with the barriers more likely to be in the context. The CMOs would suggest that willingness (or change commitment) may be predominantly influenced via the hta if the hta process is carried out in accordance to good practice(42, 270). Relating this back to iDSI's ToC with its elements of 'evidence' and 'process', this ToC might be more influential in terms of creating willingness but, of itself, this may not necessarily be sufficient to overcome contextual deficits with regards to the implementation of those decisions.

Weiner refers to a shared resolve to implement a change (change commitment) and a shared belief in collective capacity to do so (change efficacy)(196). He sees this as a function of how practitioners' value the change and determinants of implementation capability (task demands, resource availability and situational factors). As such, it spans both psychological and structural terms. Through the CMOC extracted, we evidence that where there is an interaction between willingness (change commitment) and capability (change efficacy), so that one reinforces the other, we see a more successful change in practice and implementation of hta guidance, decisions, recommendations etc – and that both are necessary but not necessarily sufficient on their own to facilitate the uptake of those decisions. A receptive context in terms of either capability or willingness is not necessarily sufficient by itself. Furthermore, whilst mechanisms (the response/reaction of people to resources) might seem obvious - value, trust, motivation etc all being relatable aspects of the human condition, what is less so is how they are activated by specific resources in specific contexts(213). Again, we attempt to reveal this through the CMOC as to how these mechanisms leading to outcomes can be achieved through different means in different contexts.

Consideration of funding research on the short-term costs of the implementation of new health technologies was a recommendation in a recent NIHR HTA impact evaluation(68). In its HTA process, France takes into consideration other transition costs related to infrastructure modifications for getting the intervention into routine use. The CMOC would indicate it is more complex than

funding alone. NICE started the Health Technologies Adoption Programme in 2014(185)[4] which supports the uptake of new recommended technologies. They facilitate the adoption of prioritised medical technologies across the NHS by: 1) engaging with clinical teams, commissioners, patient groups and industry; 2) gathering real-world experiences from health and social care organisations; 3) identifying adoption barriers and solutions. Their supporting documentation recognises 'readiness'. In assessing readiness, the care pathway is mapped and potential systems effects, the anticipated costs and savings of adoption, change in delivery requirements and monitoring data are highlighted. Whether this mapping of the care pathway adequately addresses complexity is unclear, especially when adopting a whole systems perspective. Such mapping can often result in a set of linear assumptions which will always perpetuate the gap between evidence and delivery/implementation [see Figure 1.3] because the implementation is so context and response dependent. Unless the complexities that our CMOC highlight are accounted for, such linear care pathway mapping will likely keep falling short.

The CMOC support an argument for HTA being seen in a wider systems or organisational context. An organisational level is a useful unit of analysis as collective behaviour change is needed to enact upon hta decisions in order to lead to impact at a population level. Similar findings were found in an evaluation of the Building Capacity to Use Research Evidence (BCURE)[5] where change at one level (for example, organisational) created conducive contexts for change at other levels (individual) and vice versa. This interaction or two-way effect is also highlighted by Currie et al(271) who state that business intelligence structures are necessary but not enough as the mobilisation of knowledge is 'peopled'. Indeed, the CMOC indicate that data infrastructures can facilitate learning where people embed and use this intelligence (rather than just a focus on data per se). In organisational development, socio-technical theory refers to

[4] https://www.nice.org.uk/about/what-we-do/into-practice/adoption-team

[5] https://bcureglobal.wordpress.com/

such interaction between human behaviour and technology in the workplace(272, 273).

Over 20 years ago, Rosen & Gabbay(246) called for technology assessors to branch out from their current focus on clinical outcomes and study outcomes such as staffing implications, impact on related services and set up costs - making their evidence more directly relevant to the decision-makers at whom it is aimed and thus more likely to be used. Whilst guidance exists as to how to understand and address what impact a complex intervention will have in the implementing context(194), this focuses on the technology (rather than the hta) and ends with the hta decision rather than its implementation. A robust analysis of the context-mechanism association can serve to innovate and transform interventions as elements of the context, if determined to be critically important to the successful functioning of the intervention, and may be re-theorised as an intervention component in further iterations of the intervention(213). We suggest HTA frequently needs to come with complementary components to ensure it has value, with examples found of supportive and integrated measures at all levels of the system (organisation, structure, incentives at individual level). Equally, the costs of systems change and service configuration cannot be ignored.

Referring back to the research objectives of the realist synthesis [Chapter 5], we suggest that:

1) the mechanisms that encourage the implementation of hta findings have been themed under willingness (relating to Weiner's collective change commitment) and capability (relating to Weiner's collective change efficacy). The CMOC depict how such responses can be triggered to a greater or lesser extent in different contexts. Understanding these multi-dimensional and multi-level aspects of readiness should help give a more complete picture of readiness to implement(274).

2) The CMOC would indicate that an hta often does not have the leverage to trigger these mechanisms and counter contextual deficiencies without additional

supportive measures. The literature and our CMOC would indicate that whilst acceptance, trust, motivation etc can be influenced by the hta, the systems-wide or organisational domain is under-represented in an hta. The opportunities and resources to enact those decisions are needed but this systems-perspective is often lacking - with a resulting disconnect. This would suggest that the interface between HTA and health delivery systems needs more attention – especially in resource limited contexts, as LMIC face greater complexity when it comes to implementation of evidence(275). We call for greater integration between HTA and health systems strengthening to optimise impact. There is not a clearly defined boundary between HTA and questions relating to service delivery and organisation. Indeed, some would argue that it is a false distinction to make, as technical and organisational or social considerations are so interwoven that they should not be studied in isolation.

3) Common predicting mechanisms that transcend contexts are found in an interaction between collective change efficacy and change commitment or willingness and capability. We evidence that this generally leads to a positive change in implementation of an HTA. Both are necessary, and crucially, one can reinforce the other.

In Chapter 10, we discuss the implications of these findings and make recommendations for stakeholders. To tailor this to LMICs in particular, we suggest this could mean a focus on greater connectedness between HTA and health systems. A practical way to do this, we suggest, is a refocus on meso HTA (for example, guidelines to manage patient care pathways *within a healthcare system*) or macro HTA (efficiency, organisation and strengthening *of the healthcare system*). We also support greater use of early (or development-focused) HTA. In so doing, this research provides a framework for HTA's reach to extend to both readiness and impact – thus stretching both at the beginning and at the end.

6.6 Conclusion

Through the CMO, we explored the mechanisms and pattern of outcomes in the literature on the implementation of hta decisions. Rather than simply summating the literature, by taking a realist approach, we moved beyond a description of facilitators and barriers to an interpretation of uptake in order to produce tested and data-driven theory on the mechanisms by which HTA impact can be optimised. We have analysed and synthesised CMOC, using the programme theory to progress knowledge as a theory or framework.

6.7 Next Steps

In Part 3, we bring together the different paradigms (quantitative and theory driven) into one framework and test the theory with case studies. We look at whether this further evidence supports, refines or refutes the theory developed here.

Part 3: Case studies

This part presents illustrative case studies using the ROI-NHB impact framework [Chapter 4]. In order to get to the value of investing in HTA at the systems level, we need to look at what the process is delivering, in other words, the aggregated value of each individual hta. In both case studies, by offsetting the attributable benefits against the associated costs, we evidence value for money from simply introducing a single 'hta'. We are very grateful to iDSI who helped facilitate access to CNHDRC, and to colleagues in China who kindly agreed to us using their case study. The SHTG case study was facilitated by colleagues at NHS Healthcare Improvement Scotland after I presented on this research at their annual conference. This led to connections being made with the Scottish Health Technologies Group (SHTG) and their recommendation of Freestyle Libre as a potentially impactful case study according to their own impact work.

Chapter 7 presents an hta case study from China. The uptake of new clinical pathways (CP) on stroke management was measured by taking the results of a quasi-experimental analysis carried out by colleagues at the Chinese National Health Development Research Center (CNHDRC), Beijing. We converted this into realised NHB at the population level for the areas of China where the guidance was piloted. We compared those benefits attributable to the hta with the estimated costs of undertaking the process.

Chapter 8 presents an hta case study from Scotland. This case study quantifies the value of investing in HTA using an illustrative case study from the Scottish Health Technologies Group (SHTG). Freestyle Libre (FL) is a technology to enable those with diabetes to continuously measure their glucose levels. Uptake of FL was based on an analysis of routinely collected longitudinal prescribing data. The longer-term impact on health outcomes and resource use was modelled by SHTG as part of the hta. We use these data to estimate realised population NHBs and what we can attribute to the hta recommendation. We compare those benefits to the estimated costs of undertaking the hta.

Chapter 9 reconciles the programme theory with the two case studies. We test our theory using the case studies to provide a continuing test of the same theory with two quite different bodies of data(276). We bring together the qualitative and quantitative methods; quantitative to capture an empirical and credible measure of uptake (or stopping) of a technology following an hta recommendation plus qualitative data to understand what it is about the context that has led to this level of implementation. Qualitative data were collected by itad in the pilot sites in China as part of the project's monitoring(179). This included site visits and the main mode of data collection during these visits was key informant interviews with stakeholders from groups identified as key to the implementation of the pilot. Interviews were conducted according to semi-structured protocols developed in advance of the visits and translated into Chinese. SHTG had undertaken stakeholder analysis but this was unavailable to us, as was the opportunity to undertake our own interviews given the COVID19 pandemic. Instead, we discuss our programme theory in relation to our quantitative results and SHTG's impact findings. These show how the ROI-NHB and realist framework look when they are brought together.

There were several criteria used in selecting the case studies and the rationale for each is described below.

Firstly, there needed to be data on implementation/uptake, for example, prescribing data, both before and after an hta recommendation. This was to allow any change in uptake to be estimated drawing upon quasi-experimental methods. Given that the Scottish Medicines Consortium (SMC) undertakes an hta upon launch of all new medicines, this requirement of pre-hta data meant that, in the case of Scotland, we could only consider non pharmaceutical technologies ie technologies appraised by the SHTG. We note more generally that where HTA has been institutionalised, there are fewer data prior to the hta on prescribing levels or volume against which to measure the influence of the hta recommendation. Thus, whilst ITS was successfully used previously to assess the implementation of NICE guidance (135), it no longer makes sense to use this on contemporary NICE guidance as these processes have now become institutionalised with hta recommendations made predominantly on newly

licensed technologies. However, in countries transitioning out of donor support, the intended focus of our impact framework and where HTA processes have not yet been institutionalised, there should be more data to be able to apply such methods – indeed, as Sheldon et al did when NICE was still only advisory(135).

Secondly, a decision model should ideally have been part of the hta. This meant that we could incorporate the cost-effectiveness results expressed as cost per QALY and/or NHB straight into our framework without having to undertake this modelling ourselves.

Thirdly, accompanying qualitative data relating to the implementation of the technology/hta recommendation were sought. This was to facilitate testing the theory developed as part of the realist inquiry. Ideally, we would have undertaken our own realist interviews with stakeholders but being able to draw upon existing qualitative data meant we had this option within the boundaries of time and other practicalities.

Finally, we wanted to include both an LMIC and HIC country to represent a country where HTA has been institutionalised and one newer to these processes. Although we cannot draw comparison from just two case studies, it seemed appropriate to aim for this contrast.

7 The impact and value of HTA in China: a quantitative case study

7.1 Introduction

This case study quantifies the value of investing in HTA using an illustrative case study on clinical pathways (CP) from China. Uptake of the CP was based on an analysis of routinely collected longitudinal data and the longer-term impact on health outcomes and resource use was modelled. We drew on an hta and evaluation undertaken by colleagues in China to, in turn, estimate realised population NHB and what we could attribute to the hta. By offsetting the attributable benefits against the associated costs, we show high returns from a single hta.

Very few studies to-date have estimated the benefits of HTA in terms of net health gains. However, there is growing interest amongst policy-makers and donors in the value of HTA, especially given its more recent expansion into LMIC. Our literature review [Chapter 2] found the evaluation of HTA to be methodologically lacking with scare empirical evidence of the impact of HTA on health outcomes and costs. In this case study, we apply our NHB-ROI framework to consider the opportunity costs required to undertake an hta and ultimately, to sustain an HTA infrastructure. Since every resource can be put to alternative uses, quantifying the costs and benefits of investing in HTA is important to policy makers. Using this illustrative case study on CP for stroke management which was piloted in four regions in China, we quantify the value of investing in a single hta. We drew on a pre-existing analysis by colleagues at CNHDRC of routinely collected longitudinal data on relevant clinical practices in order to quantify the uptake of the CP. We converted this into NHB at the population level encompassing the four regions of China where the guidance was piloted. We compared the benefits attributable to the hta against the costs associated with undertaking the process. We show these costs to be significantly offset when estimating the financial investment at the provincial levels. We suggest it would require approximately 30 such htas for costs to be offset if investing in HTA at a national level.

7.2 Case study background

Translating evidence into policy and implementation 'on the ground' remains a persistent challenge. Policy decisions derived from HTA include drug reimbursement decisions, development of health benefits packages as well as CP design. Clinical pathways are medical and nursing procedures including details of therapy and consultations designed to reduce variations in clinical practices, optimise patient outcomes, minimise resource utilisation, maximise the clinical effectiveness of process of care, accommodate communication between patient and healthcare professionals and, to improve quality of documentation(277-279). CPs are being increasingly recommended and used in various health settings internationally and are recognised as an effective instrument to facilitate healthcare reform(280, 281).

CNHDRC supported by NICE International recently introduced a CP for stroke covering prevention, treatment and rehabilitation guidance for clinicians at different levels of healthcare delivery(282). In China, non-communicable diseases (NCDs) account for an estimated 87% of annual deaths and 69% of the total disease burden(282). As the leading cause of adult mortality and morbidity in China, the annual stroke death case and mortality rate are approximately 1.6 million, and 157 per 100,000 respectively(283). In addition, China has 2.5 million new stroke cases each year and 7.5 million stroke survivors(284).

Moreover, recent studies have shown that on the ground, over-prescribing has led to dramatic increases in medical costs and an adverse impact upon quality of care. As part of ongoing public hospital reform in China, the development of evidence-based CPs has been prioritised as a potential solution by Chinese stakeholders to contain medical costs whilst improving quality of care in accordance with lessons learned from health sector reforms in other countries. CPs intend to improve medical quality, reduce medical service costs and increase the cost-effectiveness of overall health resource utilisation. For those reasons, we consider the development of CPs as part of the process of HTA (just as NICE has responsibility for clinical guidelines).

For acute stroke patients, including those receiving treatment for Transient Ischaemic Attack (TIA), cerebral haemorrhage and cerebral infarction (CI), the CPs promote the increased use of aspirin and statins, advocate swallowing assessment and early ambulation and, a reduction or avoidance in the use of practice which is ineffective or ambiguous. The CPs were implemented in four pilot counties, a mix of rural and urban areas (Hanbin, Huangdao, Qianjiang and Wenxian) in November 2013. By May 2015, a total of 5,490 patients had been managed following the CPs. The CPs were implemented alongside other health systems-wide interventions, such as payment reform and the development of data management systems and software.

7.3 Methods

7.3.1 Markov model

Net health benefits, our proposed measure of impact, were introduced in Chapter 4. To recap, a healthcare technology is considered beneficial if it provides more overall health than it displaces as a result of its additional cost diverting resources away from other interventions or services. To ensure that the funding of a new intervention is consistent with the objective of maximising health gains subject to a budget constraint, new health care technologies must provide an incremental cost per QALY (or DALY averted) compared to current care less than the cost-effectiveness threshold(161) [see Equations 4.1 and 4.2].

The potential population NHB associated with full implementation of the CP was quantified using the results from decision-analytic modelling undertaken as part of the hta by CNHDRC. We converted the lifetime cost per QALY obtained from a Markov model of the CPs [see Box 7.1] into NHBs and scaled this up to the population level using estimates of the eligible population given the prevalence and incidence of stroke(282). One GDP was assumed as the threshold value.

BOX 7.1 Modelling QALYs gained from CPs(285)

7.1.1 A Markov model is an analytical framework frequently used in the economic evaluation of healthcare interventions. Markov models use (mutually exclusive and exhaustive) disease states to represent all possible consequences of an intervention. Individuals transition between disease states as their condition changes over time. Time spent in each disease state for a single model cycle (and transitions between states) is associated with a cost and a health outcome. Costs and health outcomes are aggregated for a cohort of patients over successive cycles to provide lifetime cost-effectiveness against a comparator strategy(286).

7.1.2 A Markov health state transition model was developed to simulate the care trajectory of a cohort of stroke patients. The 3-year period of the evaluation was divided into cycles of 3 months (40 cycles in total), matching the time cycles of the source data. The model was structured into 3 basic modules for each cycle: 1) method of care (CP group versus non-CP group), 2) health states according to residential status (hospital, home, institutions of rehabilitation, death, and stroke-related death), 3) clinical outcome (onset, recurrence, rehabilitation, death). The model was built using Treeage Pro 2017 software. The base case was assumed to be a cohort of over 1000 CI inpatients, all over the age of 60-year-old who were assigned into either the CP group or non-CP group according to the inclusion and exclusion criteria and clinician judgment. The model used a healthcare system perspective and a time horizon of 10 years. The relative risk of stroke-related deaths under the CP was applied as a triggering factor to measure the effectiveness of the intervention. Outcome measures were costs of health states, QALYs and lifetime ICERs.

7.3.1 Value of implementation

As outlined in Chapter 4, we define the impact of HTA to be achieved through increasing the uptake of net beneficial technologies and decreasing the uptake of non-net beneficial technologies [see Figures 4.2 and 4.3]. To recap, the objective is to convey the concepts of population potential, realised and attributable NHBs to the hta (as measured against a counterfactual). We refer

to this approach as the VoImp(159). A shortfall between potential and current implementation provides evidence that implementation is sub-optimal. Inefficiencies would exist in healthcare from inadequate use of CPs, leading to foregone health benefits. The current value to the health system of the CPs is the value from all patients currently managed by the guidance. This is determined by the NHB of treating each patient, in other words, the health gains minus the opportunity costs. The population NHB of introducing a technology is, along with setting a threshold and a function of its incremental costs and effects in comparison with alternative guidance or standard care, the duration of its usage or validity and the size of the patient population served(164). As such, NHB measure the scale or magnitude of the net gains at a population level as offered by the intervention(165). Equation 4.3 defines this value to the healthcare system, where $\boldsymbol{n}$ is the total patient population eligible for treatment and **p** is the current utilisation rate of the technology(161).

Realising the population NHB of the CP required using available evidence on their degree of uptake to calculate the VoImp. We incorporated the findings of an evaluation undertaken by CNHDRC on the uptake of the CPs to estimate realised and attributable NHBs. To what extent we can attribute the uptake of the CPs to the hta and subsequent changes in clinical practice required a counterfactual as to what the level of uptake and change in practice might have been in the absence of the hta. A counterfactual or comparison group enables us to estimate changes in outcome that can we can attribute to an intervention, here the hta. There might have been some natural diffusion or decline in usage anyway of the various clinical practices covered by the CPs which would reduce the overall value and impact of the hta. To estimate the extent to which clinical behaviour change and drug usage as recommended by the CPs would have happened without the hta, CNHDRC drew on a range of quasi-experimental methods, namely interrupted time series, difference-in-difference and the use of matched controls(169, 170). A summary of the CP evaluation is provided in Box 7.2.

BOX 7.2 Estimating the level of implementation of the CP

7.2.1 Evaluation of the uptake of the CPs in 4 pilot sites was undertaken by monitoring change in quality of life, resource use, drug usage and clinical practice. This was measured using a variety of methods including interrupted time series, before-and-after studies with and without a matched control group.

7.2.2 Billing data were deemed to be the most detailed source for medical expenditure and items prescribed providing the best evidence to summarise changes in physicians' behaviour and expenditure patterns. Data were collected for 24 monthly time points before the CPs and 12 monthly time points after their introduction. Data were randomly sourced from 150 patients' records for each condition, at each pilot hospital, from three groups: pre-intervention population, post-intervention population included in CP management, and post-intervention population not included in CP management.

7.2.3 Where this routinely collected data allowed, the effect of the CPs on selected outcomes was assessed using ITS analysis with segmented regression. This can achieve statistical power with a relatively short time series(287, 288). Other analyses were conducted using a before-and-after design in comparison with a paired control group, and for new quality metrics introduced as part of the pathways (ie where there was no routinely collected data prior to their introduction), trends in the period since they were introduced was monitored. There was no pre-existing information available on the extent of any autocorrelation (ie where data items are correlated with lagged versions of themselves in the same time series) and the likely effect size. The Durbin-Watson test was used to test the presence of first order auto-correlation; if auto-correlation was detected in outcomes of interest, a generalised least squares estimator was used to estimate the regression. All analysis was performed using SAS 9.4.

The full evaluation report and results are available in a published report(282).

7.4 Data

7.4.1 Lifetime Quality Adjust Life Years and Costs

The Markov model provided data on the aggregated costs and utilities (based on the EQ-5D-3L) of health states, QALYs and the lifetime ICERs. Only direct medical and intervention costs were included; non-medical and indirect medical costs were excluded from the analysis. A pre-requisite for participation in the pilot study was the implementation of an electronic prescription system within the hospital. The patients' medical records and patients' billing data were collected electronically for data analysis. Thus, the costs of patients' hospitalisation were generated from the hospital information system, and the intervention costs were estimated by the financial department of a local hospital. The cost of care of CI patients in a rehabilitation institution unit and in-home care was estimated from a literature research. The transition probabilities between health states were collected from a literature review. Table 7-1 shows the model inputs and the resulting modelled lifetime QALYs, costs and ICERs.

7.4.2 Level of implementation of the CPs

As the 'attributable' parameter is key to our hta impact calculations, we present and critique the results of the CNHDRC evaluation which estimated the level of implementation of the CPs by looking at changes in clinical practice and drug usage before and after the hta. Analysis of routine clinical data in the pilot sites 12 months after launch suggested that the utilisation and quality of services recommended by the CPs had been strengthened (with fewer significant changes in the use of services not recommended by the CPs). Among the services adopted by the CPs, use of statins and brain imaging (within 24 hours of hospitalisation) showed the largest increase in utilisation – and are likely to be most impactful on health changes. The CNHDRC evaluation of the CP implementation produced the results shown for statins and brain imaging [Table 7.1]. CT and MRI tests within 24 hours was maintained at a relatively high level in both groups. The CP group's utilisation of statin-type drugs substantially increased as compared with usage both prior to the pilot and with the controls.

We show the largest crude before and after change (in bold, Table 7.1). The attributable percentage was calculated as the average of the difference between the pathway group and the controls. The fact that the baseline values are the same for both the intervention and control groups is critiqued as a limitation below. We use this uptake of statin as a proxy for the magnitude of uptake of the CPs and, in turn, to estimate the *realised* and *attributable* NHBs.

7.4.3 Other data

Other quantitative data required to establish the *potential* population NHB of the CP included an estimate of the eligible population given the prevalence and incidence of stroke, and a cost-effectiveness threshold. These were taken from estimates and assumptions provided by CNHDRC. We assume the CP to be valid for 5 years to enable us to calculate the number of eligible patients reached by this one hta. We estimate the investment costs in HTA at a systems or national level to be 0.1% of total healthcare spend. This is based on a relatively consistent percentage spent on HTA across countries(14, 15).

Table 7-1 CP Implementation

	Qianjiang			Hanbin			Huangdao			Wenxian			Before/	Attributable
	Pre%	Post% Pathway	Post% Controls	Pre%	Post% Pathway	Post% Controls	Pre%	Post% Pathway	Post% Controls	Pre%	Post% Pathway	Post% Controls	after %	% Diff-in-Diff
TIA statins	25.30	36.37	28.30	34.38						58.00	70.71	76.61		
CI statins	84.85	92.67	65.22	56.72	64.39	59.68	**24.40**	**95.68**	79.71	64.00	68.24	54.67	**71.28**	**10.65**
CH brain image	89.23	95.45	81.16	73.44	89.55	86.84				70.97	78.52	90.70		
CI brain image	60.61	79.33	69.57	75.37	60.61	82.26	58.93	89.93	68.12	77.42	81.74	70.59		

Adapted from main report(282) (Table 27). Appendix: Wenxian (Tables 9,15,20,21); Qianjiang (Tables 10,12,13); Hanbin (Tables 2,7,8); Huangdao (Tables 13, 16)

Table 7-2 Summary of data

Inputs	Value	Data sources
Prevalence of stroke	2%	Chinese Stroke Prevention and Treatment Report 2015
Population Hanbin	1,010,000	CNHDRC evaluation *to nearest 1000
Population Huangdao	1,470,000	
Population Qianjiang	530,000	
Population Wenxian	420,000	
Incidence	136 – 486 per 100,000 person-years (taken average of regional variation)	https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0078629
Length of time treatment valid	5 years	Assumption
Discount rate	0.035	China Guidelines for Pharmacoeconomic Evaluation
Cost-effectiveness threshold	54000 RMB	2014 China GDP per capita (current US $). [Internet]; [cited Dec 2017]. Available from: https://data.worldbank.org/indicator/NY.GDP.PCAP.CD.
Cost of acute hospitalisation	12595 RMB (range 6922—16516)	Data from pilot county
Cost of rehabilitation	10305 RMB (range 2592—12959)	Thrombolysis Implementation and Monitoring of Acute Ischemic Stroke in China
Cost of intervention	1733 RMB	Data from pilot county
Cost of home care	6773 RMB (range 2028-8639)	China National Stroke Registry

Utilities hospitalisation	0.13	
Utilities rehabilitation	0.56	
Utilities home	0.92	
Relative Risk_hospitalisation_death (CP v.s. non-CP patients)	0.8	Launois, R., Giroud, M., Mégnigbêto, A. C., Le Lay, K., Présenté, G., Mahagne, M. H., Gaudin, A. F. (2004). Estimating the Cost-Effectiveness of Stroke Units in France Compared with Conventional Care. Stroke, 35(3), 770–775. https://doi.org/10.1161/01.STR.0000117574.19517.80.
Mean cost of acute hospitalisation_CP	14328 RMB per patient	Markov model
Mean cost of acute hospitalisation _Non CP	12595 RMB per patient	Markov model
Lifetime Costs non CP pathway	114796	Markov model
QALYs non CP pathway	13.43	Markov model
Lifetime Costs CP pathway	120855	Markov model
QALYs CP pathway	13.88	Markov model
Incremental costs	6059	Markov model
Incremental QALYs	0.45	Markov model
ICER	13318 RMB per QALY	Markov model
Change in uptake/practice	70%	CNHDRC evaluation: average implementation levels post CP [Table 7.1]
Attributable uptake	10%	CNHDRC evaluation: average difference in CP/non CP group [Table 7.1]

7.5 Results

We present the individual NHB stemming from the CPs [Equation 7.1]. The incremental health benefit associated with the CPs compared with usual care was, on average, 0.45 QALYs per person. The incremental costs per person associated with the CPs compared with usual care was 6059 RMB. These costs are converted into their health equivalence by dividing through by the willingness-to-pay threshold and then subtracted from the health gain. The NHB per person is 0.34 QALYs.

Equation 7-1 Individual NHB

$$Net\ Health\ Benefits = \Delta QALYs - \frac{\Delta Cost}{\lambda}$$

= 0.45 – (6059/54000)
= 0.34

This individual NHB of 0.34 QALYs was scaled up to the population level given new and existing cases [Equation 7.2].

Equation 7-2 Prevalence and incidence

$$Prev_y = Prev_{y-1} + I_Y$$

The total eligible population ie those at risk was based on the estimates of the population of the 4 pilot provinces, stroke prevalence and incidence [see Table 7.2], discounted over 5 years. The total population at risk was estimated to be 118325 people [Table 7.3].

Table 7-3 Population at risk in the 4 pilots

Years CP valid	Population	Prevalence	Incidence	Discounted	Cumulative
Year 1	3426400	68528	10656	79184	79184
Year 2	3426400		10656	10296	89480
Year 3	3426400		10656	9948	99427
Year 4	3426400		10656	9611	109039
Year 5	3426400		10656	9286	118325

As we scale up NHB to the provincial level of the 4 pilot sites (rather than to a national level), we consider the costs of HTA as a percentage of the provincial level healthcare spend rather than that at a national level [Table 7.4].

Table 7-4 Provincial annual healthcare spend

Pilot 1	1648.65 (unit 100 million) RMB
Pilot 2	1259.4 (unit 100 million) RMB
Pilot 3	512.03 (unit 100 million) RMB
Pilot 4	730.98 (unit 100 million) RMB

Source: 2013 China healthcare statistics yearbook *more recent versions did not disaggregate by region http://www.nhfpc.gov.cn/htmlfiles/zwgkzt/ptjnj/year2013/index2013.html

The estimated costs of investing in HTA at a regional level based on 0.1% of total healthcare spend are expressed in terms of their health equivalence by dividing through by cost-effectiveness threshold (λ). Using a willingness-to-pay threshold of one GDP, this equates to 7687 NHB.

Table 7-5 Costs of investing in HTA (RMB)

Annual healthcare spend in 4 pilot areas (in 100 million RMB)	Estimated spend on HTA at 0.1% (in 100 million RMB)	Expressed in health equivalence assuming 1 GDP = 54000 RMB
4151 RMB	4.15 RMB	7687 NHB

We present the NHB associated with full, current and the attributable levels of implementation and offset this against the costs of investing in HTA at a systems level. Full implementation was based on all eligible patients receiving treatment. Current implementation was based on before-and-after results from the CP implementation study, and the attributable percentage was taken from the difference-in-difference estimates between the intervention group and the controls. This shows that the NHB attributable to approximately three htas (assuming similar impact) would offset annual investment costs in HTA at a regional level [Table 7-6].

Table 7-6 Results

Population NHB (full implementation)	Realised NHB (Current implementation) 70%	% NHBs attributable to the hta = 10%	Costs of investing in HTA expressed in NHB	No. of similar htas to offset costs
= 118325 x 0.34 = 39970	27979	2798	7687	3

7.5.1 Scenario analyses

We undertook an illustrative two-way scenario analysis by varying the magnitude of uptake and attributable percentages. This scenario analysis can help to examine uncertainty in the implementation and uptake of the CPs [Table 7-7].

Table 7-7 Scenario analysis – implementation

Implementation %	Attributable %	Population NHB	No. hta's to offset regional HTA costs
100%	10%	39970	1.92
90%	20%	39970	1.07
80%	30%	39970	0.80
70%	40%	39970	0.69
50%	50%	39970	0.77
40%	60%	39970	0.80
30%	70%	39970	0.92
20%	80%	39970	1.20
10%	90%	39970	2.14

We also undertook a scenario analysis using national healthcare spend. In 2015, this was 40975(100 million) RMB(289). HTA investment costs were estimated again at 0.1% of total healthcare spend and would include infrastructure and recurring costs of maintaining a sustainable infrastructure. The results of this analysis indicated that around twenty-seven htas (assuming similar impact) would be required to offset these costs at a national level.

Table 7-8 Scenario analysis –offsetting HTA costs at a national level

National healthcare spend (100 million RMB)	Estimated spend on HTA at 0.1% (100 million RMB)	Expressed in health equivalence assuming 1 GDP as threshold		
40974.64	40.97	75879		
Population NHB (Full implementation)	Realised NHB (Current implementation)	% NHBs attributable to hta 10%	HTA as 0.1% of national spend in NHBs	No. of hta's to offset HTA costs at a national level
39970 NHB	27979 NHB	2798 NHB	75879 NHB	Approx. 27

7.6 Discussion

HTA is expanding rapidly in China with recent political commitment for its coordination at a national level(290). However, 'for HTA to be as effective as it should be, researchers need to pay equal attention to the transmission of research findings to potential users and the translation of research into policies, decisions, clinical practice, etc. Without this, HTA is just an academic exercise that is of little value to health care and has little impact on society'(291). For HTA to be more impactful in China, knowledge translation is deemed to be important(292). Furthermore, active producers and users of HTA in China should take every opportunity to showcase the value of HTA(293). This case study illustrates its potential returns by offsetting the benefits of a single hta against typical costs associated with regional and national investment in HTA. The results indicate approximately 30 htas annually, assuming similar impact on average, would offset costs nationally. This could approximate to what

established agencies such as HITAP in Thailand might now average[6], and certainly a lot less than what some HTA agencies carry out[7].

Unlike the RAND impact evaluation of the HTA NIHR programme (19), we estimate realised and attributable NHBs using methods to establish uptake against a counterfactual rather than simply assume full implementation. To do this, we drew upon existing analyses undertaken by colleagues at CHNDRC. There are some limitations to the quasi-experimental methods applied. To measure the change in statin usage, they used a difference-in-difference approach. CNHDRC stated that some analyses were conducted using a before-and-after design in comparison with a paired control group but for new quality metrics introduced as part of the pathways (ie no routinely collected data prior to their introduction), only trends in the period since they were introduced was monitored. As the baseline measurement for both the intervention and the control groups were the same, we took the pooled difference in the post-intervention measures between the two groups. Nevertheless, it illustrates that a control or comparator is crucial to impact evaluations given the potential for a crude before-and-after analysis to lead to erroneous measures and conclusions of impact, with attribution which could be wrongly assigned or interpreted.

Others have estimated the cost-effectiveness of clinical guidelines(294). However, they do not measure implementation or bring in the issue of attribution/the counterfactual as this is about the CP itself rather than the hta. We make the case that this is not just about the CP being cost-effective which, indeed, they appear to be: a French study found similar results of 1395 euros per QALY = 11299 RMB/QALY, and expert opinion has indicated that CPs in stroke units are much more likely to be followed than in other disease areas(295). Our results suggest that the actual hta creates value with benefits of the process offsetting the costs. There is a relatively high level of implementation associated with the CP and realised NHB are estimated to be fairly close to the potential NHB associated with full implementation. However, the observed change in practice from the quasi-experimental analyses found the control group

[6] HITAP list 240 reports on their website over the last 10+ years https://www.hitap.net/en/

[7] The Scottish Medicines Consortium lists ~100 htas in 2019 https://www.scottishmedicines.org.uk/

also experienced a change in practice to higher levels of compliance with the recommended guidance. The fact that the controls also saw a positive change in clinical practice could be due to other factors. This project was implemented under the backdrop of healthcare reform. The impact of the project – and thus the hta – would have likely been influenced by external factors, such as healthcare reform policy, medical insurance policy and hospital management. Indeed, the costs of these supporting measures cannot be ignored. This is discussed in Chapter 9.

We aim to capture here the often over-looked issue of redeploying existing resources from alternative uses (the opportunity costs) in order to ensure a sustainable HTA system be established. If we think priority setting should be better informed by evidence, and use tools such as HTA, this necessarily implies institutional change to establish a sustainable system. Consequently, the opportunity costs that are needed for its establishment and ongoing running have to be considered(38). In order to show the ROI of HTA at a systems level, this would require aggregating all individual htas. By aggregating the realised NHBs of all hta decisions given the current level of implementation, or as we propose, by considering a level of implementation attributable to the hta as measured against a counterfactual, we would be able to offset these aggregated NHB against the costs of investing in HTA at a systems level. A natural progression of this case study would be to apply the full framework. We recognise that the data requirements would not be without their challenges and would likely require purposeful sampling with data collection and not just a reliance on routine administrative systems.

7.7 Limitations

In keeping with the assumptions made by CNHDRC in their original hta, we have used WHO GDP per capita-based thresholds. Although widely used to assess cost-effectiveness, empirical evidence suggests these thresholds are likely to be significantly higher than actual health opportunity costs(162, 171) – and indeed, was not a use for which they were ever intended(296). They depict the value of a statistical life and is not a reflection of opportunity costs. A scenario analysis using more realistic thresholds could be undertaken.

This case study intends to be illustrative only, seeking to use available data and information on the cost, effectiveness and implementation levels of a technology following an hta. We scale this up to make an estimate of how many htas may be required to offset the costs of investing in HTA at a systems level – at both a regional and national level. As stated above, a natural progression of this case study would be to apply the NHB-ROI framework in full or with a wider range of case studies.

The support costs of the payment reforms, IT support etc have not been taken into account in the analysis. Again, a scenario or threshold analysis could be undertaken to incorporate a range of values.

Finally, we recognise we are not capturing spillover benefits of investing in HTA beyond NHB. For example, HTA might facilitate improved monitoring and administrative systems. However, we have used the qualitative data collected from stakeholders in the pilot sites in China which should help with understanding how social, technical, intellectual and political capital may be leveraged by HTA over time as well as the mechanisms underpinning the uptake of the CP [Chapter 9].

7.8 Conclusion

Our study goes beyond showing that the stroke CPs are cost-effective. This case study shows that the hta itself offers value for money to the extent that the benefits attributable to the process offset a significant proportion of its costs.

8 The value of HTA: a quantitative case study on Freestyle Libre Glucose Monitoring System from the Scottish Health Technologies Group

8.1 Introduction

This research intends to quantify the value of HTA using a case study approach from Scotland where HTA has been institutionalised, having been formally established approximately 20 years ago. Using a single illustrative hta on Freestyle Libre (FL) for diabetes management as implemented in Scotland, this case study intends to quantify the value of investing in HTA as a priority setting process. The potential benefits from introducing FL for diabetes management is measured and converted into NHBs at the population level. We use routinely collected longitudinal prescribing data to assess the uptake of FL across all fourteen NHS Boards before-and-after the Scottish Health Technologies Group's (SHTG) guidance; using quasi-experimental methods, we more robustly estimate the benefits attributable to the hta (as compared with a crude before-and-after estimate) by accounting for pre-existing trends in prescribing. This change in uptake is used, in turn, to estimate realised population NHBs – and what we can attribute to the hta itself. We compare those benefits attributable to the hta against its associated costs and show these to be potentially offset in full.

8.2 Case study background

The SHTG is a national HTA agency. SHTG provides evidence, support and advice to NHS Scotland on the use of new and existing health technologies which are not medicines and which are likely to have significant implications for people's care. The role of the SHTG committee is to advise on and support best use of health technology interventions within NHS Scotland. The committee is made up of representatives from NHS boards, clinical and professional networks, academia, National Procurement and Scottish Government. It also includes four volunteer public partners, who represent the views of the Scottish public[8].

[8]http://www.healthcareimprovementscotland.org/our_work/technologies_and_medicines/shtg/about_shtg.aspx

The SHTG undertook an hta of FL in 2018. Diabetes mellitus (DM) is a major cause of morbidity and mortality in Scotland with around 300,000 people diagnosed with the condition. Self-monitoring of blood glucose is essential for people with type 1 (DMT1), and is used by roughly one in 10 people with type 2 (DMT2) to manage glycaemic control and adjust insulin or other medications. Currently, finger-prick self-monitoring of blood glucose (SMBG) using test strips and blood glucose meters is the most frequently used monitoring method(297). DM Freestyle Libre® is a flash glucose monitoring system. The 'game-changing technology' allows people with diabetes to track their blood sugar levels without having to prick their fingers. The flash glucose monitoring system is designed to change how people with diabetes measure their glucose levels and ultimately help them achieve better health outcomes. The system automatically reads glucose levels through a small sensor that is worn on the back of the upper arm for up to fourteen days, eliminating the need for routine finger pricks and user calibration. Two clinical trials and real-world evidence from more than 50,000 users showed that people who used the FL system scanned their glucose levels an average of at least 15 times per day. In real-world use, higher rates of scanning to self-monitor glucose were found to be strongly associated with improved glucose measures, including decreased time in hypoglycaemia and hyperglycaemia(298).

Regulatory approval for FL was obtained in September 2014 in the UK but was available as of then only privately. It was added to the Scottish drugs tariff from November 2017, meaning it would be available for reimbursement across NHS Scotland (as well as the NHS in England and Wales, and Health and Social Care in Northern Ireland) for people with DMT1 and DMT2 who intensively used insulin. The topic was prioritised for inclusion on the SHTG work programme following a referral from the Scottish Diabetes Group and the assessment was intended to provide information to support decision makers within NHS Scotland. In July 2018, an advice statement[9] was issued by the SHTG following an Evidence Note(299). Evidence notes are rapid reviews of the evidence surrounding health technologies that are under consideration by decision makers in NHS Scotland.

---

9 http://www.healthcareimprovementscotland.org/our_work/technologies_and_medicines/topics_assessed/shtg_009-18.aspx

They are intended to provide information quickly to support time-sensitive decisions, usually within 6 months(297). The Evidence Note was based on a European Network for Health Technology Assessment (EUnetHTA) rapid review of published clinical effectiveness and safety literature but also incorporated a range of additional evidence sources including: a Healthcare Improvement Scotland (NHS HIS) economic evaluation, a patient group submission from Diabetes Scotland and NHS Scotland clinical expert input. The advice for NHS Scotland was that flash glucose monitoring with Freestyle Libre® be available for individuals with diabetes who are actively engaged in the management of their diabetes and who intensively manage their condition with multiple daily insulin injections or insulin pump therapy. NHS Scotland is required to consider the SHTG advice but it is not mandatory. Seven of the fourteen NHS health boards in Scotland had already made FL available after launch in November 2017 and had put pathways into practice(300) whilst the rest were awaiting the SHTG recommendation first before prescribing it.

8.3 Methods

8.3.1 Markov model

Our framework utilises the lifetime health economic results from a Markov model [see Chapter 7, Box 7.1] as outlined in SHTG's evidence note. We adapt the information from the published Evidence Note to estimate potential population NHBs [see Chapter 4]. The model is summarised in Box 8.1. For full details, please see SHTG's Evidence Note(299).

BOX 8.1 Modelling QALYs gained from FLs

Economic modelling was undertaken using a Markov Model and considered resource utilisation, utilities and 2 main clinical outcomes, namely the impact of FL on the testing frequency of blood glucose and on the frequency of hypoglycemic events. FL had a statistically significant impact compared with SMBG on both these outcomes using evidence from 2 key randomised controlled trials. A Markov health state transition model was developed to simulate the care trajectory of a cohort of diabetic patients. A simple 2-state structure was applied, separated into sub-models for T1DM and T2DM. A patient can either be

dead or alive in the model (the 2 states), with transition determined by a diabetes-specific mortality rate. One year of living with diabetes is associated with a direct resource use linked to the consumables involved in monitoring blood glucose but also an indirect resource use due to severe hypoglycemic events. A key assumption is that FL monitoring results in a decrease in the number of blood glucose tests used by a patient. The evidence also points to a potential decreased incidence of hypoglycemia with FL compared with SMBG, resulting in an impact on NHS resources for severe hypoglycemic events requiring medical attention. These potential savings are offset by an increased cost associated with FL. Two model structures were used: the first was a restricted model taking into account the relative cost of monitoring and the direct impact of FL on health utility scores; the second ('full' model) also incorporates hypoglycemic events and the associated impact on utility and resource use. The models used a healthcare system perspective and a lifetime horizon. Outcome measures were costs of health states, QALYs and lifetime ICERs(299).

8.3.2 Value of implementation

As outlined previously (see Chapters 4 and 7), the current value to the health system of the healthcare technology is the value from all the patients who currently receive it and is determined by the NHB of treating each patient ie the health gains minus the opportunity costs. To recap, we define the impact of HTA to be achieved through increasing the uptake of net beneficial technologies and decreasing the uptake of non-net beneficial technologies. The objective is to convey the concepts of potential, realised and attributable (ie as measured against a counterfactual) population NHBs. We refer to this approach as VoImp(159).

In order to calculate the potential population NHB associated with FL, we converted the cost-effectiveness results of the lifetime cost per QALY obtained from the Markov model of FL versus SMBG care and scaled this up to the population level given estimates of the prevalence and incidence of diabetes in Scotland. To calculate NHB, we used the lower bound of £20k as the UK threshold. Realising the population NHB then required using available evidence on the degree of uptake of FL as described below.

8.3.3 Interrupted time series – segmented regression

We analysed monthly monitoring data on the prescribing of FL in order to estimate realised and attributable NHB. To estimate the extent to which uptake of FL would have happened without the hta, we used ITS with segmented regression. Uptake of FL was recorded by ISD Scotland who monitored prescribing rates across all NHS boards(301). Prescriptions in the community dataset(302) contained details of items dispensed in the community by General Practices (GP) by NHS Board. Data monitoring started from November 2017 (ie when FL became available on the NHS) and we analysed the data over 24 monthly time points to November 2019 [Table 8.1]. This, therefore, covered 8 months prescribing data pre-hta guidance which was issued in July 2018, and 16 monthly time points post-hta guidance.

Table 8-1 ISD monthly practice level prescribing data by NHS Board

Prescribing by Board	Ayr. Arran[1]	Borders	D&G[2]	Fife	Forth Valley	Grampian	Greater Glasgow & Clyde	Highland	Lanark[3].	Lothian	Orkney	Shet[4].	Tayside	Western Isles	Total	Mile-stones
Nov-17	**4**	**0**	**0**	**6**	**59**	**4**	**14**	**6**	**1**	**8**	**0**	**0**	**0**	**0**	**102**	**NHS**
Dec-17	2	1	2	22	207	8	13	4	9	2	0	0	0	0	**270**	
Jan-18	4	13	9	79	336	18	21	12	18	10	0	0	0	0	**520**	
Feb-18	6	8	10	93	452	6	20	23	12	45	0	0	0	0	**675**	
Mar-18	40	142	5	132	692	6	31	26	20	2755	0	0	0	0	**3849**	
Apr-18	57	182	3	134	878	10	43	22	21	3162	0	0	0	0	**4512**	
May-18	147	81	83	162	1149	18	47	38	70	3005	0	0	0	0	**4800**	
Jun-18	260	242	498	51	1341	4	166	35	229	4109	0	0	0	0	**6935**	
Jul-18	**380**	**140**	**599**	**87**	**1206**	**15**	**1245**	**53**	**308**	**4303**	**0**	**0**	**0**	**0**	**8336**	**HTA**
Aug-18	434	98	677	19	1358	8	1610	82	459	4934	0	0	0	54	**9733**	
Sep-18	504	394	734	138	1339	10	1967	92	516	4433	0	0	0	56	**10183**	
Oct-18	736	383	826	25	1585	277	2351	119	771	4916	0	4	0	70	**12063**	
Nov-18	760	318	854	329	1558	345	2312	170	779	4949	0	45	0	164	**12583**	
Dec-18	915	523	1016	442	1899	718	2460	234	1261	5349	30	58	0	143	**15048**	
Jan-19	947	397	945	528	1758	631	2430	237	1362	5132	58	56	0	163	**14644**	
Feb-19	926	525	934	1213	1792	904	2356	333	408	4971	58	47	989	149	**15605**	
Mar-19	1104	613	1102	1368	2067	973	2668	271	92	5598	49	76	1223	237	**17441**	
Apr-19	1188	524	1156	1491	2106	1219	2413	577	106	5570	74	61	1132	237	**17854**	
May-19	1280	658	1249	1689	2253	1429	2914	856	0	6465	78	89	1432	256	**20648**	
Jun-19	1370	646	1247	1671	2240	1423	3742	840	297	5788	68	71	1453	126	**20982**	
Jul-19	1473	700	1359	1909	2434	1794	3330	949	0	6645	92	78	1705	288	**22756**	
Aug-19	1482	724	1381	1767	2496	1621	3982	965	70	6406	101	90	1775	257	**23117**	
Sep-19	1503	753	1448	1930	2526	1871	3749	1151	419	6586	103	87	1920	294	**24340**	
Oct-19	1720	732	1548	1882	2751	2040	4667	1210	0	6826	93	118	2051	309	**25947**	

1Ayrshire and Arran, 2Dumfries and Galloway 3Lanarkshire, 4 Shetland. Units are the number of prescriptions dispensed for which the dispenser has been reimbursed. [Glossary-of-Terms.pdf (isdscotland.org)]

This routinely collected data allowed the effect of the hta guidance on FL prescribing to be assessed using ITS analysis with segmented regression. This method can achieve statistical power with a relatively short time series (287, 288). ITS studies use routine data collected at equally spaced intervals of time before and after an intervention, and do not necessarily require a control site. Given the limited number of timepoints over which data were collected, segmented regression was a more practical option for analysing and utilising this routinely collected data. The analysis followed the steps as outlined in Lagarde et al, 2012(287).

Equation 8.1 shows the specification of the linear regression.

Equation 8-1 Linear regression

$$Yt = \beta 0 + \beta 1 time + \beta 2 intervention + \beta 3 postslope + \varepsilon t$$

Where Y_t is the outcome variable (ie the number of FL units prescribed) at time t; time is a continuous variable indicating time from the start of the study up to the end of the period of observation. In this model, $\beta 0$ captures the baseline level of the outcome at time 0 (beginning of the period); $\beta 1$ estimates the structural trend or growth rate in utilisation, independently from the intervention; $\beta 2$ estimates the immediate impact of the intervention or the change in level in the outcome of interest after the intervention; and $\beta 3$ reflects the change in trend, or growth rate in outcome, after the intervention.

The intervention (ie the issuance of hta guidance) is coded 0 for pre-intervention time points and 1 for post-intervention time points and, post-slope is coded 0 up to the last point before the intervention phase and coded sequentially from 1 thereafter. The following independent variables ('time', 'intervention', 'post-slope', 'pre-slope') were created [Table 8.2].

Table 8-2 ITS Dataset and variables

Date	Outcome [Table 8.1]	Time	Intervention	Pre-slope	Post-slope
Nov-17	102	1	0	1	0
Dec-17	270	2	0	2	0
Jan-18	520	3	0	3	0
Feb-18	675	4	0	4	0
Mar-18	3849	5	0	5	0
Apr-18	4512	6	0	6	0
May-18	4800	7	0	7	0
Jun-18	6935	8	0	8	0
Jul-18	8336	9	1	8	1
Aug-18	9733	10	1	8	2
Sep-18	10183	11	1	8	3
Oct-18	12063	12	1	8	4
Nov-18	12583	13	1	8	5
Dec-18	15048	14	1	8	6
Jan-19	14644	15	1	8	7
Feb-19	15605	16	1	8	8
Mar-19	17441	17	1	8	9
Apr-19	17854	18	1	8	10
May-19	20648	19	1	8	11
Jun-19	20982	20	1	8	12
Jul-19	22756	21	1	8	13
Aug-19	23117	22	1	8	14
Sep-19	24340	23	1	8	15
Oct-19	25947	24	1	8	16

The analysis was undertaken with ‘outcome’ as the dependent variable in the regression. The Durbin-Watson test was used to test the presence of first order auto-correlation (where values are statistically correlated with past values of the data) using a Prais-Winsten estimator that corrects for data auto-correlation; and a generalised least squares estimator was used to estimate the regression. All analysis was performed using STATA version 14. The STATA code is given below.

Figure 8-1 Stata code for ITS segmented regression

```
Stimport excel ("ITS analysis") firstrow
reg Outcome Time Intervention Postslope
tsset Time
prais Outcome Time Intervention Postslope
predict fitted_values
line fitted_values Time
```

8.3.4 Return on investment

The ROI equation is outlined in Chapter 4 [Equation 4.1]. Slotting the NHBs metric into this original ROI equation, we can express a NHB-ROI [Equation 4.2].

8.4 Data

8.4.1 Lifetime Quality Adjusted Life Years and Costs

The Markov model provided data on the aggregated costs and utilities (based on the EQ-5D) of health states, QALYs and the lifetime ICER. Only direct medical and intervention costs were included, non-medical and indirect medical costs were excluded from the analysis [Table 8.4].

8.4.2 Level of implementation of FL

Analysis of the routine clinical data across all NHS boards 24 months after launch (and 16 months after the hta guidance) showed that the utilisation of FL had been strengthened after the publication of the SHTG HTA Evidence Note. We use these results to inform a magnitude of uptake of FL and, in turn, to estimate the realised and attributable NHBs. The results of the model are presented in Table 8.3.

Table 8-3 Results of ITS segmented regression

Outcome	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Time	1027.365	90.95131	11.3	0	837.6443	1217.087
Intervention	820.4716	490.0288	1.67	0.11	-201.7106	1842.654
Postslope	132.8752	95.19409	1.4	0.178	-65.69615	331.4466
_cons	-1944.49	456.566	-4.26	0	-2896.868	-992.108
rho	-0.18329					

Results indicate a (non significant) increase of 820 FL units per month after the issuance of the hta guidance. We used these results to calculate the relative and absolute effects of the guidance by estimating the expected prescribing uptake of FL 16 months after the hta (24 months after launch). We entered the values of the explanatory variables at month 24 into the estimated regression equation [Equation 8.2].

Equation 8-2 HTA effect on prescribing levels

$$Y24 = \beta 0 + \beta 1\, x\, 24 + \beta 2 + \beta 3\, x\, 16$$

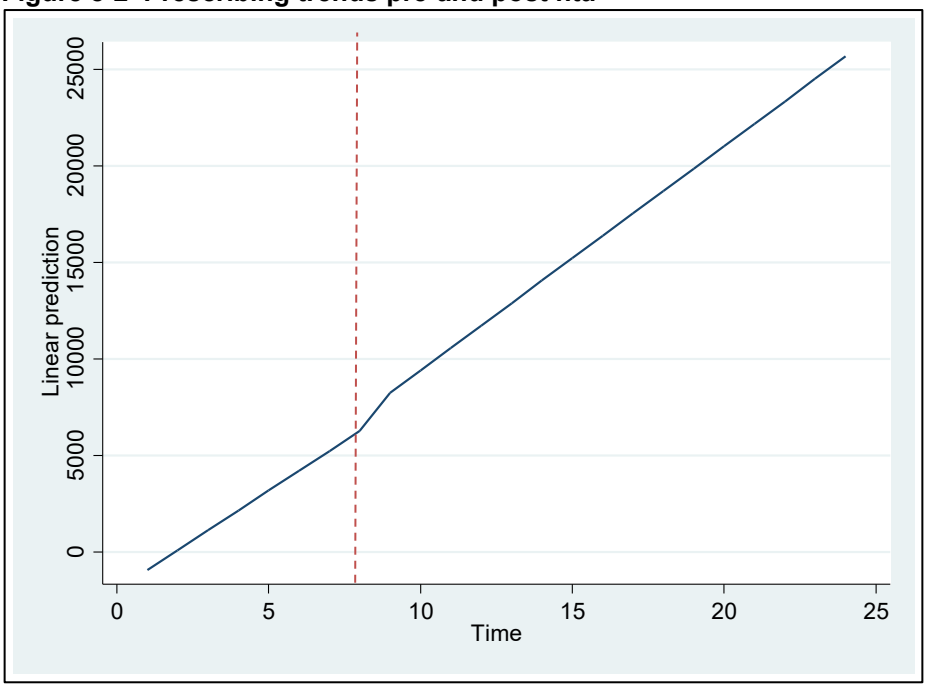
In the absence of a control group, we also used the results of the regression to fabricate a counterfactual and estimate what would have been the level of outcome of prescribing without the HTA [Equation 8.3].

Equation 8-3 Prescribing levels without HTA

$$Y24 = \beta 0 + \beta 1\, x\, 24$$

Using the estimated coefficients, we found that 16 months after the hta guidance there was a predicted monthly prescribing uptake of 25,652 units of FL (actual 25,947 units); while, without the hta, the model predicted that uptake would only have reached 22,704 units. Therefore, the absolute effect of the intervention was an increase of 2,948 units and the relative effect was a 13% increase in prescribing. We show the fitted results with the outcome of interest plotted against time with the hta guidance issued at month 8 indicated by the dashed line [Figure 8.2].

Figure 8-2 Prescribing trends pre-and post hta



8.4.3 Other data

Other quantitative data required to estimate the potential population NHB of FL included an estimate of the eligible population given the prevalence and incidence of diabetes in Scotland, and a cost-effectiveness threshold. These values were taken from the estimates provided in the HTA Evidence Note(299). We assumed FL to be valid for 5 years (for example, before another review or new diagnostic) to enable us to calculate the number of patients potentially able to be reached by this hta. We estimated the investment costs in HTA to be 0.1% of total healthcare spend as a relatively consistent percentage spent on HTA across countries(14, 15)[10].

[10] SHTG were calculating their actual costs for this exercise but was stopped due to COVID19.

Table 8-4 Summary table of data

Inputs	Value	Data sources
Prevalence of diabetes in Scotland	5.4% (n = 291,981) of which 10.6% T1DM 88.3% T2DM <2% other forms of DM	Scottish Diabetes Survey 2016
Eligible population in Scotland meeting FL criteria in Year 1	Over 60,000 patients of which 28,531 have T1DM 32,342 have T2DM	2017 figures/SCI-Diabetes
Incidence (crude rate)	18 cases per 100,000 pop per year (n = 943) T1DM 316 cases per 100,000 pop per year (n = 16,973) T2DM	Evidence Note 81
Length of time diagnostic valid	5 years	Assumption
Discount rate	0.035	UK Treasury recommended rates and as used in SHTG model
Cost-effectiveness threshold	£20,000	HIS/NICE usual lower threshold
Key model inputs- for full details see Evidence Note 81		
SMBG mean no. of blood tests per day	5.5 (T1DM), 3.8 (T2DM)	Markov model input - from RCTs
FL intervention effect on blood tests	Reduction of 90.9% (T1DM), 89.5% (T2DM)	Markov model input - from RCTs
FL intervention effect on severe hypoglycaemic events	Reduction of 55.0% (T1DM), 52.6% (T2DM)	Markov model input - from RCTs
FL cost	£35 each with a lifetime of 14 days use	Markov model input – Scottish Drug Tariff/manufacturer
Cost of severe hypoglycaemic event (SHE)	£1,034 (£855 - £1,253)	Rates of SHE obtained from RCTs/systematic review by SHTG
Full model T1DM base case		
Lifetime costs FL	£18,074	Markov model
Lifetime costs SMBG	£12,860	Markov model
Lifetime QALYs FL	9.73	Markov model
Lifetime QALYs SMBG	7.61	Markov model
Incremental costs	£5,214	Markov model
Incremental QALYs	2.12	Markov model
ICER	£2,459 per QALY	Markov model

Full model T2DM base case		
Lifetime costs FL	£10,450	Markov model
Lifetime costs SMBG	£5,535	Markov model
Lifetime QALYs FL	6.14	Markov model
Lifetime QALYs SMBG	5.04	Markov model
Incremental costs	£4,916	Markov model
Incremental QALYs	1.09	Markov model
ICER	£4,498 per QALY	Markov model
Restricted model T1DM base case		
Lifetime costs FL	£17,010	Markov model
Lifetime costs SMBG	£10,496	Markov model
Lifetime QALYs FL	13.20	Markov model
Lifetime QALYs SMBG	12.67	Markov model
Incremental costs	£6,514	Markov model
Incremental QALYs	0.53	Markov model
ICER	£12,340 per QALY	Markov model
Restricted model T2DM base case		
Lifetime costs FL	£9,837	Markov model
Lifetime costs SMBG	£4,241	Markov model
Lifetime QALYs FL	7.51	Markov model
Lifetime QALYs SMBG	7.20	Markov model
Incremental costs	£5,596	Markov model
Incremental QALYs	0.31	Markov model
ICER	£18,125 per QALY	Markov model
Change in prescribing	From 116 items (November 2017) to 25,947 items (October 2019)	ISD data
Uptake %	25,947 / total eligible Yr 1 n = 60,873	43% assuming units sold equates to the number of patients
Attributable uptake	13%	ITS using segmented regression

8.5 Results

We present the individual NHBs stemming from FL. The incremental health benefit associated with FL compared with usual care was 0.53 and 0.31 QALYs for T1DM and T2DM respectively (restricted model); and 2.12 and 1.09 QALYs for T1DM and T2DM respectively (full model). The incremental costs associated with the FL compared with usual care was £6514 and £5596 for T1DM and T2DM respectively (restricted model); and £5214 and £4916 for T1DM and T2DM respectively (full model). These costs are converted into their health equivalence by dividing through by the willingness-to-pay threshold and then subtracted from the health gain. The NHB per person is 0.2043 and 0.0302 QALYs for T1DM and T2DM respectively (restricted model); and 1.8593 and 0.8442 QALYs for T1DM and T2DM respectively (full model) [Table 8.5].

Table 8-5 NHBs individual level (base case)

Restricted model	**Incremental QALYs**	**Incremental Costs**	**ICER**	**NHBs using £20k threshold**	**NHB**
T1DM	0.53	£6514	£12,340 *	= 0.53 – (6514/20000)	0.2043
T2DM	0.31	£5596	£18,125 *	= 0.31 – (5596/20000)	0.0302
Full model	**Incremental QALYs**	**Incremental Costs**	**ICER**		**NHBs**
T1DM	2.12	£5214	£2,459	= 2.12 – (5214/20000)	1.8593
T2DM	1.09	£4916	£4,498	= 1.09 – (4916/20000)	0.8442

*(rounding)

This is scaled up to the population level in Scotland taking existing and new cases of T1DM and T2DM [Equation 8.4] and discounted over 5 years [Table 8.6].

Equation 8-4 Prevalence and incidence

$$Prev_{y} = Prev_{y-1} + I_{Y}$$

The total population at risk was estimated to be 32593 T1DM and 42291 T2DM people [Table 8.6].

Table 8-6 Cumulative population at risk

CPs valid	Prevalence	Incidence	Discounted	Cumulative
Year 1				
T1DM	28,531	869	29,400	29,400
T2DM	32,342	2129	34,471	34,471
Year 2				
T1DM			840	30,240
T2DM			2057	36,528
Year 3				
T1DM			811	31,051
T2DM			1988	38,516
Year 4				
T1DM			784	31,835
T2DM			1920	40,436
Year 5				
T1DM			758	32,593
T2DM			1855	42,291

The estimated costs of investing in HTA based on 0.1% of total healthcare spend, are expressed in terms of their health equivalence by dividing through by cost-effectiveness threshold (λ). Using a willingness-to-pay threshold of £20,000, this equates to 650 NHB [Table 8.7].

Table 8-7 Costs HTA

National healthcare spend Scotland 2018 (GBP)	Estimated spend on HTA at 0.1% (GBP)	Expressed in health equivalence, £20k threshold
13 billion[11]	13,000,000	650 NHBs

We present the NHB associated with full (potential), current (realised) and the attributable levels of implementation of FL and offset this against the costs of investing in HTA at a systems level. Full implementation was based on all eligible patients receiving treatment. The NHB were scaled up separately by T1DM and T2DM, and then totalled. Current implementation was based on the number of currently prescribed units as a percentage of eligible population, and the attributable percentage was taken from the ITS analysis. This shows that, using the restricted (more conservative) model, the gains attributable to 2 hta

[11] Source: https://fullfact.org/health/nhs-scotland-spending/

(assuming similar impact) would offset the annual investment costs. Using the results of the full model, the costs are offset entirely [Table 8.8].

Table 8-8 Results

	Population potential NHB (full implement)	Realised NHB (Current implementation 43%)	Attributable NHB = 13%	No. of hta's (assuming similar impact) to offset costs @ 650 NHBs
Restricted model QALYs	7,936 NHB = (6,659 T1DM + 1,277 T2DM)	3,383	439	2
0.2043				
0.0302				
Full model QALYs	96,302 NHB = (60,600 T1 + 35,702 T2DM)	41,049	5,330	Offset ~ 7-fold
1.8593				
0.8442				

Expressing this in terms of an NHB-ROI calculation is shown below [Equation 8.5].

Equation 8-5 NHB-ROI

$$NHB\ ROI = \frac{NHB - Cost\ Investment/\lambda}{Cost\ Investment/\lambda}$$

$$Restricted\ model = \frac{\{439\ NHB - 650\ NHB\}}{650\ NHB} = -32\%$$

$$Full\ model = \frac{\{5330\ NHB - 650\ NHB\}}{650\ NHB} = 720\%$$

This shows the same results as above but expressed as a ROI. The result would need to be positive to show a return on investment. Using the restricted model, it suggests that the NHB associated with one further hta (assuming similar impact) would achieve this. For the full model, it shows the hta yielded a 720% NHB-ROI.

8.6 Discussion

In order to get to the overall value of investing in HTA at a systems (ie here, at a national level), we need to look at what the process is delivering. We used an illustrative example by quantifying the value stemming from an hta of FL. First, we converted the ICER as modelled by SHTG to NHBs and scaled to a population level taking new and existing cases discounted over 5 years (an assumption on the life of the technology). This provided an estimate of the potential population NHBs ie if everyone who was eligible were to receive it. Implementation was monitored by ISD and this routinely collected prescribing data indicated a current uptake of 43% (25,947 units sold with over 60,000 eligible patients in Scotland). This provided an estimate of realised population NHB. Using segmented regression as a simplified form of ITS, we estimated uptake after the issuance of SHTG's guidance against the pre-existing prescribing trend to establish what we could attribute to the hta. This indicated that, although it was predicted that uptake would have increased anyway without the hta, uptake was an additional 13% higher after the advice statement was issued as measured 16 months later. Finally, we offset the attributable benefits against the costs of investing in HTA (expressed in terms of their health equivalence).

The results of the ITS segmented regression undertaken do not detect a statistically significant effect at the conventional 5% level ie it could be considered that the result is inconclusive as to whether the increase in uptake of the technology was due to the hta or not. Yet, for the sake of illustrating the conceptual framework, we incorporate this point estimate of the mean as real and attributable to the hta. The results of the ROI calculation should, therefore, be treated with caution. This effect parameter could be handled probabilistically using the standard error of the estimate from the regression so that the final ROI is expressed with a measure of the degree of

certainty in the result. Additionally, a threshold analysis could be undertaken to establish the minimum level of uptake required in this case study to generate a positive ROI. The latter, however, addresses a slightly different point as results already show a positive ROI on the observed (non significant) effect. Note that when no evidence of effect is found, examining the joint distribution of costs and effects is still valuable for cost-effectiveness analysis(303).

We make the case that this is not just about FL being cost-effective which, indeed, it appears to be with many studies finding similar results(304) but that the hta itself provides value for money – noting the uncertainty discussed above. This case study illustrates the high potential returns and value for money of investing in HTA by offsetting the benefits of a single hta against typical costs associated with HTA investment costs at a systems level(14, 15). Whilst others have estimated HTA impact, they have stopped at the equivalent of what we refer to as potential NHB(19). We estimate impact in terms of realised and attributable NHBs using methods to establish uptake against a counterfactual of what might have happened without the hta. Ramsay et al have critiqued the use of ITS in HTA and have found it has often fallen short; they recommend time series regression techniques when the series is short (305) which we employed. Although subject to uncertainty in the ITS point estimate, our results, based on this being real, suggest that the hta created value, and that these benefits substantially offset the costs.

Finally, there is extensive data linkage in Scotland for those with diabetes to longer term health outcomes. Data are collected from electronic patient records as part of the Scottish Care Information – Diabetes Collaboration[12]. Whilst outcomes have been modelled as part of the hta economic evaluation, there is the potential for longer term follow-up using patient records and actual health outcomes.

8.7 Limitations

Whilst T1DM and T2DM NHBs were able to be split out when calculating population NHB (as we have incidence and prevalence by diabetes type), we do

[12] https://www.sci-diabetes.scot.nhs.uk/

not have prescribing split out by diabetic type. This would provide a more accurate calculation of realised NHB but for the purposes of this example, the case study illustrates the principle and methods which can be applied.

In converting QALYs into their monetary equivalence, we use NICE's lower value of £20k per QALY. However, it should be noted that research would indicate it is considerably lower still(163). Deterministic sensitivity analysis could again be undertaken using more accurate values using specifically, the supply-side estimates by Claxton et al of around £13k per QALY(163). We have, however, refrained from using the upper NICE threshold of £30k or indeed, higher values used for end-of-life treatments by NICE.

In both case studies, the thresholds used have a rationale but both are likely to be over-estimates of the true opportunity costs. Getting the threshold wrong has consequences. The consequences of over-estimating the threshold are arguably worse than under-estimating it[13]. Too high, and this would result in a net reduction in population health as the benefits do not exceed those displaced. Too low, and cost-effectiveness results may recommend interventions are not funded when they could generate population health benefits. In this case, there would be lost opportunities for health improvement. The implications of using too high a threshold in these case studies means the impact of HTA is likely over-estimated. Scenario analysis using lower threshold values to more realistically reflect the true opportunity costs incurred by the health system would provide more conservative estimates of the impact of HTA; the value of NHBs fall and the cost of the hta expressed in health equivalence increases. The ROI is consequently reduced.

Again, similar to Chapter 7, we recognise that doing this for all htas in any given context would be impractical and so we suggest the use of illustrative case studies. A natural progression of this case study would be to apply the full framework. We recognise that the data requirements would not be without their challenges and would likely require purposeful sampling with data

[13] Acknowledge presentations by Prof Mark Sculpher, NUI Galway 2013 on estimating thresholds and Paul Revill, CHE presentation.

collection and not just a reliance on routine administrative systems – though this case study does utilise routinely collected data.

Finally, this framework allows time-series analysis by individual boards so we could have looked at the impact of the hta on those Boards which had already put a care pathway for FL into practice before the SHTG Evidence Note was published. However, this was not entirely clear as the data showed uptake against more than the number of expected Boards waiting for guidance before prescribing FL.

8.8 Conclusion

Our study goes beyond the recommendation that FL is cost-effective by showing that the hta process itself offered value for money, showing potentially high returns with the introduction of a single hta – albeit against a backdrop of uncertainty in the ITS mean value on change in prescribing levels following the hta. To the extent the benefits offset the costs, we quantify the value of investing in such processes. In so doing, we aim to capture the issue of redeploying existing resources from alternative uses (the opportunity costs) to ensure a sustainable HTA system be established(37).

9 Integrating theory-based and quasi-experimental methods for impact evaluation

9.1 Integrating mixed methods

The UK Government's Magenta 2020 book(131) which sets out guidance on what to consider when designing an impact evaluation, groups methods into three types: theory-based impact evaluation methods, experimental and quasi-experimental impact evaluation methods, and value-for-money methods. It states that good quality impact evaluation evidence will be both theoretically driven and provide confidence that the measured outcomes can be attributed to the policy and provide an estimate of the size of that impact. These different methods may be better suited to different types of data analysis and study designs. In this case, measuring the impact of HTA, we suggest that a complexity-informed approach is appropriate. Whilst a dominant empiricist philosophy has been the conventional approach to evaluation, proving causation by quantitative data and statistical relationships, using this alone would not give an understanding of how impact is brought about. Quantitative methods alone would not produce insights about how any measured change comes about, or whether the same outcome would occur if the intervention is tried in another context or at a different scale. Combining experimental or quasi-experimental approaches with theory-based approaches ...can provide this often essential insight.(131) The wider literature also recognises the use of different paradigms, calling for economists who share a broadly 'positivist' perspective to work together with researchers from organisational studies and policy analysis with an interpretivist perspective to study the same change in delivery of a health service. Researchers who have traditionally taken a 'black box' approach focussing on inputs and outputs need to work with researchers who study context and processes, and vice-versa(306).

9.1.1 Applying mixed methods to our research

We incorporate the use of each of these types of methods outlined in the Magenta Book(131) in this thesis to come up with a robust way of quantifying and explaining the impact and value for money of HTA. This 'joining up of methods' allows researchers to think from different perspectives – and, rather than being

a case of either/or, we consider the added value of their combination to be in providing a complexity-informed approach. In our impact framework, the quantitative and qualitative research is positioned under an overarching realist theory driven approach [Table 4.1]. Our framework acknowledges that, at the heart of impact evaluation, is a requirement to explain 'how' and 'why' causes and effects are linked – as well as addressing 'how much' and 'to what extent'. Explanatory methods are used alongside quantitative methods to generate, test and refine explanations for a gap between potential and realised gains in population health further to an hta. In adopting a mixed methods approach, we aim to both demonstrate the value of HTA (quantified in terms of NHB), and to progress knowledge as a theory or framework.

Whilst realist inquiry - and theory-based methods more generally - can be used to explore the causal chains thought to bring about change by an intervention, they do not provide precise estimates of effect sizes(131). Furthermore, a criticism of realism is that it does not always (or often) provide a quantitative effect size(307). Pawson(308) suggests 'that quantitative may mean no more than the ability to count numbers.' Rather than adopting a crude before-and-after count, experimental and quasi-experimental approaches are designed to achieve a robust estimate of the average impact of an intervention. This requires a comparison group. Consideration and quantification of what would have happened without the hta is central to our impact framework. We draw on quasi-experimental methods to estimate how much 'uptake' can be attributed to the hta using (often underutilised) administrative data.

Our framework also utilises a ROI approach which is rooted in a positivist epistemology [Chapter 3]. Some suggest that if methodological issues can be addressed in ROI then it may be possible to arrive at a rigorous and reliable ratio that is representative of social impact(309-311). The use of standard linear (positivist) economic methods as used in (S)ROI has been critiqued on both methodological and philosophical grounds, in so far as it does not account for any relational aspects and the result is reduced to a ratio of financial costs and benefits(309). Here, qualitative methods - or a more complexity-informed approach - may be 'important to understand how an intervention's mechanisms

lead to the desired outcomes as the ratio may take focus away from understanding the processes of achieving outcomes'(309, 310).

HTA is context specific and, realism as an approach to developing programme theory is particularly relevant because it focuses specifically on the influence of context. We use quantitative methods to measure outcomes and qualitative, descriptive methods to develop theory in order to further knowledge. Realism is a broad methodological church(158) and realist evaluation is profoundly multi-method. Hence, the use of the terms realist 'inquiry, principles or approach' - not methodology - as realism favours methodological pluralism. Indeed, Pawson(308) recommends 'qualitative data to investigate mechanisms – you need to get one way or another to the reasoning of stakeholders; quantitative data to get to outcomes…; and comparative data to inspect contexts'. We use the uptake of a technology following an hta as our outcome of interest in the realist inquiry (outcome 'O' in CMO). This enables us to work backwards from outcomes to explore mechanisms and relevant contextual factors, as causal mechanisms are activated only in favourable conditions. By proposing a quasi-experimental method (specifically, ITS) to measure outcomes, we go beyond a mixed methods approach to synthesising a positivist analysis within a realist inquiry.

9.2 Testing the theory

Here, we use the ROI-NHB studies as 'test cases' and show how the ROI-NHB and realist framework looks when brought together. Whilst a realist review is a theory-led approach to knowledge synthesis using secondary sources, realist evaluation is a form of primary research. Although the realist component of this thesis is primarily about developing theory, building this theory from data and evidence in the literature, we also test our programme theory using two quantitative case studies [Chapters 7 and 8] together with accompanying qualitative data which are discussed below for both studies. This provides a continuing test of the same theory with two quite different bodies of data. As such, the case studies are acting as the 'test cases' of the theory resulting from the realist synthesis ie they are an exemplar of how the CMOC work in practice. The case studies show that these approaches can be complementary and how this can work in practice. Through the CMOC, we test the argument

that readiness is a complex multi-dimensional construct(196): it requires both a willingness and capability for successful implementation [Chapter 6].

We note it has been observed that 'HTA reports typically do not define their impact objectives, that is – the effects they would like to achieve (for example, to influence coverage decisions, support guideline formulation or change routine practice)…the stated objectives or research questions are scientific, related to the technology being assessed rather than describing the expected role of the HTA itself'(78). We assume that when a positive hta recommendation is made, 100% implementation is what is aimed for. When considering the case studies – both of which have found the technology in question to be cost-effective (net beneficial) - we suggest that as they show a discernible effect of uptake in the right direction over and above what would have happened without the hta, and that their attributable benefits outweigh the costs, these are successful cases of implementation following an hta (as well as providing value for money of the hta).

9.2.1 CNHDRC case study (China)

The China case study was accompanied by interviews carried out by itad. This qualitative data were collected in the pilot sites as part of the project's monitoring(179). This included site visits and the main mode of data collection during these visits was informant interviews with stakeholders from groups identified as key to the implementation of the pilot. Interviews were conducted according to semi-structured protocols developed in advance of the visits and translated into Chinese. The full report is publicly available(179).

In this case study, the hta was the development of contextually-appropriate CP for patients hospitalised for stroke. Hta resources were guidance, knowledge and training. Over-and under-treatment are common phenomena in China's healthcare system. The causes of such practice not only relate to the absence of appropriately developed clinical guidelines but also because of institutional issues, such as inadequacies in quality monitoring systems, medical insurance reimbursement policies and general compensation mechanisms. The hta was referred to as 'integrated' CP as additional supportive resources, including data management systems and software, the negotiation and development of

remuneration systems for medical personnel to increase their support for clinical pathways and payment reform, were also employed. This addressed institutional inadequacies in terms of quality monitoring systems, medical insurance reimbursement policies and general compensation mechanisms. In particular, robust internal IT systems were identified within the healthcare system as being a key step to improving implementation of the guidance and sites were chosen where such systems were functioning. Such information systems can provide more reliable and timely information for clinicians, support clinical decision-making and medical service behaviour standardisation and improve quality of care.

These meso/macro supportive measures were critical for the uptake of the CP by providing the capacity and practical resources to implement but also, in turn, for installing motivation and willingness to act collaboratively. The hta involved extensive stakeholder engagement with a variety of actors at national and local levels. The reaction of clinicians to buy into the CPs was triggered through the specific delivery of supportive resources (capability) and these resources also activated motivation to act and buy-in too (willingness); it enabled practitioners to feel capable of making the change successful (collective change efficacy) [M]. This indicates that knowledge/information-based resources, held at the individual level, were not enough to produce uptake: context-level resources were required to augment the knowledge and information. Moreover, the scope of training carried out locally was broader than clinical pathways, and included the use of management and information systems, costing, billing etc. By addressing the wider causes (of over/under prescribing), greater harmonisation of interests [C] among stakeholders across multiple levels of the system, triggered motivation [M] and greater acceptance of the guidance [M]. While clinical pathways are an important part, they are not, and cannot be, a standalone reform. There was broad consensus around the need for an integrated solution. This established mechanisms for collaboration and referral across different tiers of the health system – between health insurance, provision of health services and clinical behaviour. All this led to positive changes in clinical practice, prescribing and costs.

There was, in certain situations, nevertheless a display of autonomy and own discretion with variable adherence despite generally good compliance overall: where patients had been receiving the same treatment for a long time [C], the practitioner felt if one was to suddenly change the treatment regimen, patients may not accept it [M], easily leading to medical disputes [O]. External validation and legitimacy were noted in this case study - a low-trust setting [C] – in that external validation was brought about by the hta (as it developed with international support) which led to legitimacy [M] and thus better communication between patients and practitioners on their healthcare [O]. Ultimately, greater buy in, changes in attitudes, increased capacity and collaboration resulted in improved patient communications.

From Weiner's theory of organisational readiness, a shared commitment and shared change efficacy are both necessary. This case study showed that they can reinforce each other. The successful implementation of the integrated CP required working in alignment with health systems strengthening initiatives in order to overcome contextual deficits. The pilot managed to effectively bundle a number of components together that linked clinical pathways to payment reform, management changes and technical support. The case study shows too that a refocus on meso hta [Chapters 1 and 6] (here, guidelines to manage patient care pathways *within* a health system) along with macro interventions (here, payment and IT reform to strengthen a health system) leads to a better integration between the hta and delivery systems - and thus implementation and, ultimately, impact. We propose that it is this interaction between 'willingness and capability' or between change 'commitment and efficacy' that creates an outcome ie a change in clinical practice or prescribing. Thus, we argue that the hta and system-level need to interact to create uptake. A meso level hta necessitating the adaptation of service provision to manage patients' care pathways within a health system – and integrating macro approaches (efficiency, organisation and strengthening of the healthcare system) reinforced the change efficacy of those delivering the services. It enabled practitioners to feel capable of making the change successful ie collective change efficacy was increased which led to their willingness to implement the guidance. It showed a concern for the values and interests of the stakeholders with potential to block

or subvert implementation and their interactions with the health system, with a harmonisation of interests across multiple health system levels.

9.2.2 SHTG case study (Scotland)

HTA in Scotland has been institutionalised for many years. Having found many HTA agencies measure processes rather than outcomes, SHTG have made recent efforts to measure the impact of its work on decision-making and longer-term outcomes. Previous studies have explored the impact of HTA in Scotland, for example, Bennie et al(312) demonstrated the complex relationship between HTA advice and change in clinical practice. Factors identified that may explain the patterns of medicine use included delays between medicine launch and initial advice, the publication of conflicting advice from different national bodies and failure to engage with relevant clinical experts early in the medicine review process(312). SHTG's current impact monitoring is based on the INAHTA impact tool(54) [see Chapter 1] and is largely qualitative. The work relating to FL included stakeholder analysis. The results of that research by SHTG found that the FL hta had a 'major influence on decisions'. We, in turn, quantified this impact in terms of the costs of the hta and its realised NHB.

A context analysis of this case study would support our theory of a positive interaction between willingness (demand-side, at least) and capability leading to successful implementation – indeed, of the hta having 'a major influence' according to SHTG's qualitative work. The topic of this hta was advocated for by patient groups and was prioritised for inclusion on the SHTG work programme following a referral from the Scottish Diabetes Group. They collated information from patient surveys and focus groups on patient experiences. Although use of this device requires regular contact with healthcare professionals to ensure it remains the most appropriate for the patient, it is essentially about patients' home-based management of their diabetes. As such, prescribing and monitoring by the NHS is the main aspect to its implementation. The former could be said to relate to 'willingness', the latter to do with 'capability' within the system. Both appear to have contributed favourably to the implementation of the technology.

Interestingly, whilst some Boards awaited the SHTG's guidance, others had already established a care pathway for FL in practice before the Evidence Note was published. Moreover, some Boards have been slower than others to implement since the guidance was issued. In November 2018, a year on from launch and 3 months after the HTA recommendation, some boards were still not prescribing FL or lacked clear guidance(313). The actual level of prescribing is captured in our quantitative case study by the analysis of the data from each Board over a period of time before and after the hta but a better understanding of the reasons for differences in the timing of implementation could be explored further.

9.2.3 Drawing lines from hta to HTA

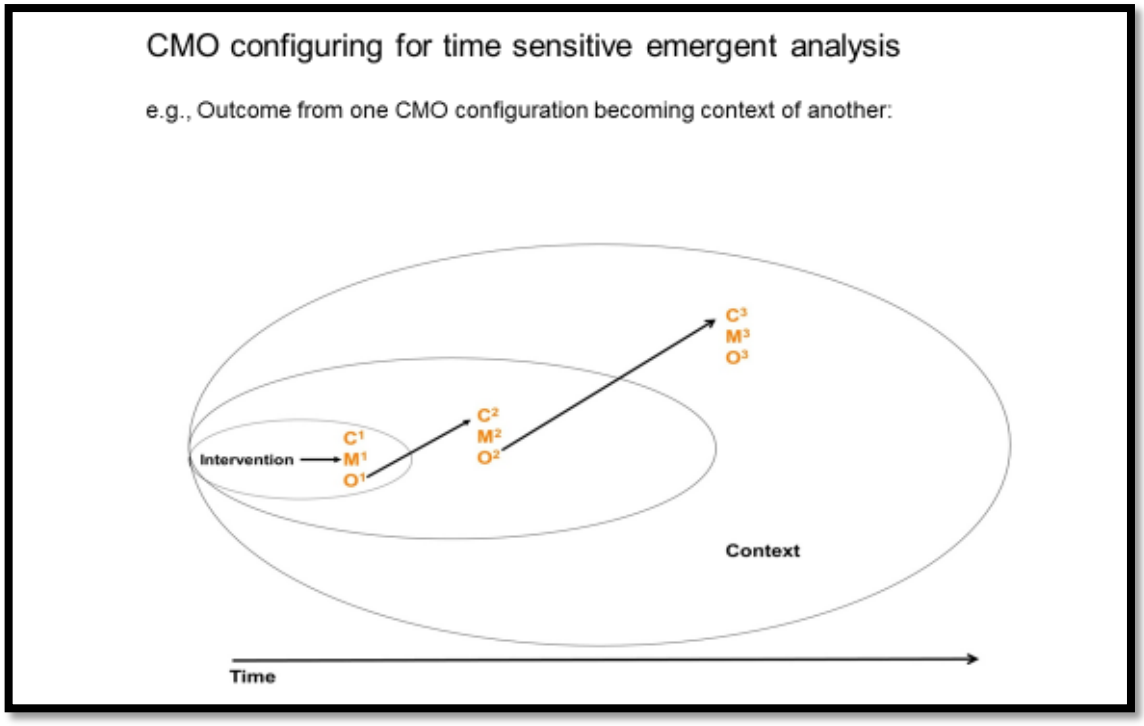
Our understanding is that the impact of HTA on health outcomes is the major gap in the literature and thus, the focus of this research. We recognise though that there are likely to be other externalities to arise from HTA including, for example, better bargaining power on price negotiations. Here, we consider other spillover effects arising from HTA and the building of capital over time.

Through the inclusion of intermediate outcomes [Table 6.2], we show how a generative causal pathway is created of a series of mechanisms activating and interacting in contexts over time(126). The degree to which hta recommendations are 'taken up' is observable. However, if we were to conclude at the empirical, we risk missing a lot of information about what is working (or not) and 'why' regarding the effective implementation, and consequent, impact of HTA. Generative causation would additionally ask *how does* HTA have an effect/impact? For example, there may not yet be a discernible empirical effect on 'uptake' - perhaps there is a time lag for the implementation and, indeed the diffusion of technologies is known to follow an S-shaped curve(314) so that uptake is better modelled as a diffusion curve than a step function. Rather than being seen to be failing, is the hta still having a residual effect in terms of other outcomes such as awareness and/or acceptance in preparing the ground, and what is it about the pre-existing resources in that context and the resources provided by the hta that is driving the implementation of recommendations and decisions? It is about making the ground fertile and receptive to the hta, creating an environment for adoption of HTA to have the broadest scope. An

emergent analysis of context and mechanistic data may allow us to identify that the hta is still having a residual effect in terms of preparatory work and transforming the context over time. Rather than concluding the hta has failed or has produced negligible returns in our interlinked NHB-ROI framework, we can establish how and to what extent this is building social, intellectual and political capital for HTA institutionalisation, capacity and acceptance over time.

In this regard, economies of scale and scope are critical. Scale, as sheer volume of htas will bring cost down per hta but also the potential of reaping economies of scope. This speaks to critical mass by leveraging the technical, intellectual, social and political capital built up over time with each hta. Social capital is, in realist terms, generatively important in that it is adding institutional value which is necessary for other outcomes to emerge(315)[Figure 9.1]. For example, in China, the hta was found to have provided a basis for others to learn from and a policy driven scale-up indicated political support had been developed through the initial hta as well as building capability through linking it to other needed reforms. According to one CNHDRC interviewee, 'overall, the impact of the project has been greater than the impact of the pilots: …(and) has had a large impact in changing ideas at central/policy levels'(179). Promoting such a converging approach between researchers and political stakeholders for future HTA institutionalisation is a key step(316). Thus, the impact is probably greater, and of greater importance than that of the specific pilot experience(179).

Figure 9-1 Modelling realist causation over time



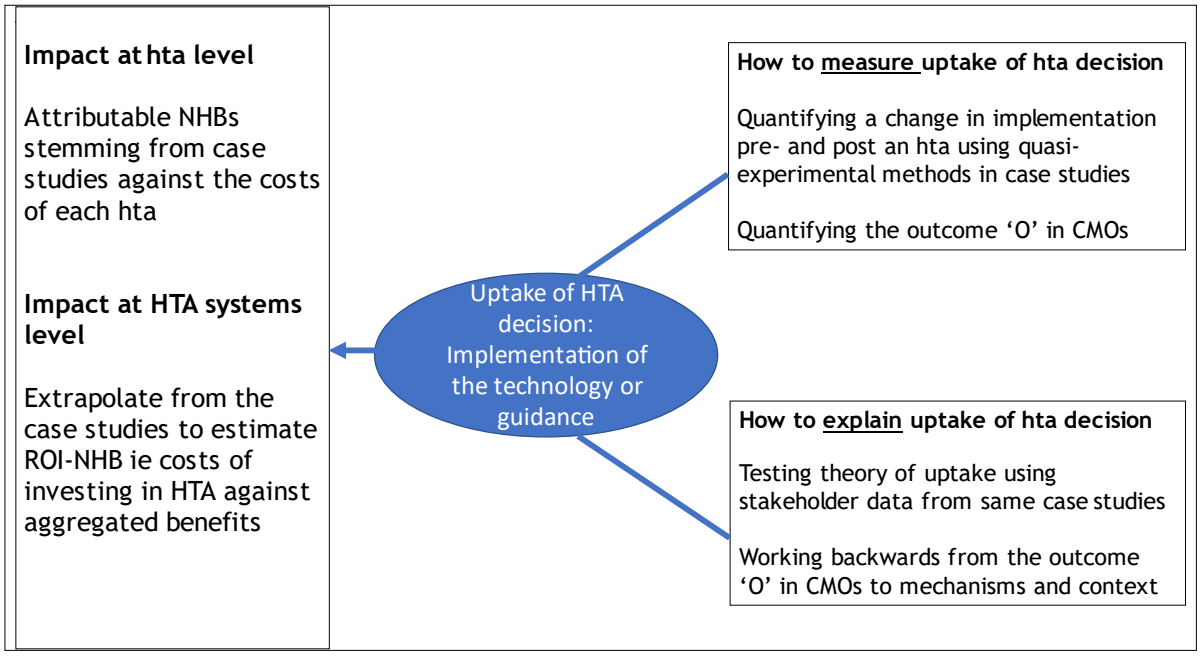
Source: adapted from Pawson R, Tilley N. Realistic evaluation. London: Sage publications; 1997 (141). Permission sought from JJagosh email June 2020.

9.3 Discussion

This chapter integrates the NHB-ROI case study findings and their corresponding stakeholder interviews. We do this by testing the programme theory developed from the realist review with the key informant data from the China and Scotland case studies. This key informant data were purposefully excluded from the realist component in building the programme theory (Chapters 5 and 6) so as to then be able to go on to test the theory with these case studies (Chapters 7-9). Ideally, realist interviewing would have allowed us to explore aspects of the theory in greater depth from perspectives of different stakeholders. However, the existing stakeholder data allowed a continuing test of the same theory with two quite different bodies of data - across the theory building and theory testing components. As such, the case studies are acting as the 'test cases' of the theory resulting from the realist synthesis ie they are an exemplar of how the CMOC work in practice. Furthermore, we use the uptake of each case study's technology following an hta as our outcome of interest (outcome 'O' in CMO) in both the quantitative and qualitative data. This enabled us to work backwards from the quantified outcomes to explore mechanisms and relevant contextual

factors, around uptake. A diagram illustrates how the different elements of the case studies interlink with uptake at its centre.

Figure 9-2: Integration of methods



Garrido et al (78), state 'further research is necessary to consider how to develop and test theoretical frameworks. Successful HTA impact assessment requires the involvement of different disciplines…'. Realism is about testing and refining theory, and as such, the project or intervention is not the object of the evaluation in its own right(317). The basis of adopting a realist approach here is that we start and end with a theory, supposing that the final overarching theory can be translatable to other similar contexts. Generalisability comes from a positivist paradigm that there is a universal truth which can be scaled up or replicated (for example, from a trial) and, as such, can handle contextual differences. Realism, on the other hand, is not looking for the general but transferability and portability. What are the aspects of an intervention that can be transferred to other contexts and expect same results?

Everything we are focusing on in this research is a step on from hta dissemination. If we think of hta as an input to decision making but we need 'complementary' aspects for implementation of that decision to ensure hta has value (ie hta supports decisions but not directly the implementation of them) then by focusing on the context as well as characteristics of the

recommendations, we can help transfer decisions into actions. We tested the CMOC, and the case studies showed that whilst 'willingness' or 'change commitment' can be influenced via the hta itself, this is not necessarily sufficient to overcome contextual deficits. We suggest an hta frequently needs to come with complementary components to ensure it has impact, with examples found of supportive and integrated measures at all levels of the system. Our programme theory, and the testing thereof, would suggest that where an hta is accompanied by joined up systems-wide approaches (IT, payment structures etc) (C), the change efficacy of practitioners is increased (M) leading to their adherence of recommendations [O]. Just as the integration of HTA into the regulatory, reimbursement and decision-making processes is necessary for its findings to translate into policies or decisions(291), implementation of those decisions required a greater interconnectedness between HTA and the delivery system. Our CMOC also evidenced that where there is an interaction between 'capability and willingness', this generally led to a positive change in implementation. A couple of points following on from this are noted below.

Firstly, it should be noted that positivists would want to test the counterfactual situation ie of a specific hta accompanied by systems-wide interventions compared with that same hta not accompanied by systems-wide interventions in comparable contexts. As such, they would see this as a case of the missing counterfactual. As previously discussed [Chapter 3], realism supports generative causation not causation via a counterfactual.

Secondly, whether the same programme theory can be applied to a negative hta decision (ie not to invest in or implement a technology that is not actually yet in the system) and/or disinvestment (ie de-implementation) of a technology is untested. Although there were some articles retrieved on negative decisions or disinvestment(87, 222, 237, 312), they were only a few [Table 6.2]. Indeed, Esmail et al note that knowledge translation theory is needed to support this change as it is more psychologically challenging to stop rather than take up a technology(87).

Finally, in order to assess whether HTA is a good use of resources, we compared its costs and benefits. Whether an intervention represents value for money can

be estimated with relative straightforwardness; so long as one can 'specify the inputs and outcomes with sufficient clarity to ensure that changes in resource use and benefits can be measured and valued, then it is not necessary to understand how the intervention works'(33). There is a growing body of literature on how to address complexity better in economic evaluations and implications for health economists(33, 318, 319), in particular, by applying a realist approach to resource use and cost-effectiveness(319, 320). It is widely recognised that the use, and therefore cost-effectiveness, of new technologies may be adversely influenced by constraints on the demand for and supply of health services in which the technology is placed – and that 'given the extent of constraints, complexity needs to be 'ruled out' rather than 'ruled in', especially in economic evaluations in LMICs'(321). Some have called for health economists to improve the relevance of their findings by distinguishing between context-dependent and context-independent factors in their analysis to close the gap between economic evaluations and clinical practice(322). Vassall et al lay out a conceptual framework for identifying the technology and what constraints there may be in the context of the care pathway(321) – though, to re-iterate, unless complexities are fully accounted for, such linear mapping may fall short. Whilst our research focuses on the evaluation of HTA and the adoption of hta recommendations, rather than the evaluation of technologies, nevertheless as evaluation should be used not just for accountability purposes but also for learning, we hope this research might contribute to the debate as to how and why health economists should broaden the scope of their work by adopting complexity-informed approaches.

9.4 Limitations

Ideally, testing the theory would have provided a continuing test of it by using primary data as this would have allowed us to explore aspects of the theory in greater depth from perspectives of different stakeholders. Although we do use two different bodies of data to develop the theory from that used to test it, we continue to draw on secondary sources. As for the quantitative China case study, we drew on information in the public domain, specifically previous interviews with stakeholders regarding the hta. Whilst this provided a rich source of material to test our programme theory, undertaking our own realist interviewing in China may have enhanced this work. However, it was agreed

that the team had already contributed substantially to monitoring and evaluation learning. Access to the SHTG stakeholder data as well as further realist interviewing with NHS staff to refine the theory further are on hold[14].

We aimed to distinguish between HIC and LMIC. However, there is a dearth of LMIC research [Table 6.2]. Should we have wished to pursue realist primary data collection in LMIC, Gilmore highlights the methodological challenges of undertaking realist research in these contexts(323).

9.5 Conclusions

None of the methods we use are novel per se – although realism is a relatively new approach to evaluation in healthcare and an original way to review the literature pertaining to HTA. Rather, the innovation comes from their combined use, integrating a positivist quasi-experimental approach and an alternative to 'causation via counter-factual' using realist complexity theorising to unpack not only what works but how and why.

9.6 Next steps

We discuss implications of the findings from this research and make recommendations for stakeholders.

[14] Due to the COVID situation and personal circumstances in the last year of my thesis.

Part 4: Synthesis of the research

This final part comprising of Chapter 10 synthesises the findings of the research. We provide practical recommendations for decision-makers and make suggestions for potential areas of future research.

10 Recommendations and Further Research

10.1 What our NHB-ROI framework adds

Many LMIC health systems lack the tools and institutional mechanisms to prioritise the interventions and products that generate the most health for the money, with less than $4 out of every $100 USD in public funds being spent on a health maximising technology(324). The importance of investing in HTA as a means of supporting a better allocation of finite resources and UHC has been recognised by the international community [Chapter 1]. Yet, the benefits concerning the link between HTA as a process at the systems level and impact in terms of net health improvements have rarely been quantified. Furthermore, while there may be long term financial and health benefits to HTA, the need for upfront funding is a major constraint(325). In a similar vein, a recent working paper discusses the importance of the 'impact of impact evaluations' in building a stronger evidence base on effective social interventions in LMIC(326).

We envisage the use of this research will encourage accountability of spending decisions and help to optimise the impact of HTA in an era of investment and expansion in particular, for LMIC, through better understanding of HTA's role in delivering health outcomes and value for money at the system level. This research aims to offer a forward-looking model that LMIC can point to as a reference for their own implementation. Ultimately, we hope this ROI-NHB framework will contribute to demonstrating the value of HTA by quantifying the NHB and opportunity costs of investing in these processes. Indeed, this framework has been cited in the literature(180) as a start to devising empirical strategies to study the broader effects of HTA beyond pricing and reimbursement. In turn, we hope this will contribute to generating political will and financial investment in these processes. We summarise what we consider our impact framework adds to the literature and why we deem this important [Table 10.1].

Table 10-1 What our framework adds

Criteria	Existing models	What our HTA framework adds	Why this is important
Unit of analysis	Evaluation of a specific intervention or hta. Distinction made between impact evaluation of an hta and HTA agencies.	Aggregates the value of each hta to get to the value of HTA at a systems level.	Investment is at a systems level and a lack of longer-term impact assessment at this level may undermine HTA's importance and value.
Outcome measure	Impact on decision-making, outputs or costs/cost savings. Health and costs considered separately. Limited empirical evidence on the impact on health/net health gains.	Impact is expressed in terms of modelled NHBs (or alternatively, NMBs).	The metric of NHBs/NMBs enables us to maintain HTA's broader value of healthcare efficiency rather than just cost-containment. As it is understood that final outcomes are already (usually) modelled as potential NHBs as part of an hta, this allows the focus to be on implementation or uptake of hta decisions – and without which there can be no impact on health.

Criteria	Existing models	What our HTA framework adds	Why this is important
Opportunity costs	Rarely considered	Opportunity costs are made explicit as costs are expressed in terms of their health equivalence using a WTP threshold, combining costs and health into one metric.	Opportunity cost is a key concept when considering efficiency and value for money. It helps policy makers quantify what else they could have done with the money.
Attribution	Rarely considered	Counterfactual modelled via quasi-experimental methods. This enables us to estimate an attributable level of uptake of a technology following an hta to quantify attributable population benefit to the process.	A comparator (even if not observable) is fundamental to measuring impact against what may have happened with and without an intervention, here an hta.
Theory	Rarely used, description only of barriers and facilitators to undertaking HTA and uptake of its findings.	We move beyond describing facilitators and barriers in order to produce tested and data-driven theory on the mechanisms by which HTA impact can be optimised.	Health is impacted only if decisions are implemented. Using theory driven approaches, we generate findings with are portable across contexts and deals with complexity to understand and improve implementation, leading to impact.

10.2 Challenges using the framework

Challenges with implementing our NHB-ROI impact framework at any level are predominantly to do with a lack of monitoring data and information. In order to test the framework convincingly requires greater access to monitoring data. This includes ensuring the quality and frequency of data collection, and ideally data which provides insight into vulnerable populations in order to understand the distribution of costs and benefits. Specifically, this requires administrative longitudinal data on the utilisation or drug volume, or change in clinical practice, pre- and post-an hta to show temporal trends in the uptake (stopping) of a technology following an hta decision or recommendation. Also required are the capital and running costs of investing in HTA at a systems level. It may be that this is currently beyond the data capacity of most countries as part of their routine data or monitoring systems. A 2018 OECD survey of routinely collected data on prescribed and dispensed medicines revealed that countries are at very different stages of data infrastructure and development. Some countries initiated data collections and established prescription databases in the 1990s, while others are only about to start(327).

Operationalising the framework to assess the ROI of an entire country's HTA programme is unlikely to be feasible any context. Practically, we can only undertake illustrative case studies. RAND restricted its evaluation to 10 HTA-funded projects framed within the more clearly defined boundaries of an audited research programme(19). However, we can scale-up the results from case studies to calculate how many htas may be needed in order to get to a positive ROI. Rather than utilising data from existing systems, a full application of the framework would more likely require a specific evaluation or audit to be commissioned to purposely collect the necessary data. That said, in the case studies presented, we have drawn upon existing data and analyses in the public domain to undertake our own further data interrogation and to populate the framework. Such data where it exists are often underutilised(327). Furthermore, the act of monitoring in itself features in the CMOC as a way of creating new attitudes and motivations which can augment the evidence-base and uptake of a technology.

Finally, we make explicit that we present these case studies as illustrative ie as proof-of-method studies rather than for their empirical significance - see key simplifying assumptions made underlying the framework which provides reasons for which this should be clear (Chapter 4). We thus acknowledge uncertainty in the empirical findings of both case studies. We outline below sources of uncertainty in the analyses and propose suggestions of how uncertainty in key parameters could be explored more fully in future research.

In the framework, there exists uncertainty at multiple levels:

- the modelling of the cost-effectiveness of the technology as part of the hta. Sources of uncertainty here could include structural uncertainty (for example, the extent to which a model depicts the condition, intervention, care pathways etc); methodological uncertainty (for example, valuation of quality of life and resource use); data sources on costs and effects etc.

- the reliance of NHBs on threshold values as a representation of the true opportunity costs of resource allocation within a health system is uncertain and a source of potential bias (Chapter 4).

- the estimation of realised NHBs using monitoring data on implementation, prescribing etc. The quality of this data including accuracy, timeliness, completeness, consistency etc is a source of uncertainty.

- the modelling of change in uptake (stopping) using quasi-experimental methods. As well as the data quality issues outlined above, the point estimate result is subject to statistical uncertainty (Chapter 8).

- calculating the costs of HTA/hta. We purposefully avoided a full costing exercise in our case studies for sensitivity reasons, opting to use more general estimates from the literature. Sources of costs and assumptions made while calculating them are further sources of uncertainty if doing a specific costing exercise.

10.3 Implications of this research for decision makers

Our focus has been on developing a framework for evaluating the impact of HTA on net health gains but it is only when those decisions result in implementation and practice change, that the gains can be achieved. Whilst ICERs are scaleless, NHB allows us to scale up the impact by the population. Realising these NHBs is only possible via the implementation of, and access to, beneficial technologies. There has often been a perceived disconnect between the delivery system and HTA, with the uneven implementation of cost-effective interventions. As the evidence shows, even in relatively well-resourced health systems and high income contexts, cost-effective interventions do not always get implemented with recurrent themes around organisational processes, clinician engagement and financing being key barriers(234). Optimisation of health technology utilisation is health system- and health technology-specific, and HTA may be made more relevant by greater explicit consideration of organisational issues (233). Providing evidence about the expected impact of a technology on health system structure, processes and resources might be valuable to inform the construct and recommendations of an hta or develop an implementation plan(233). Yet, this aspect of hta is found to be lacking [Chapter 6]. Cacciatore et al(249) found the assessment of organisational aspects to understand what resources are needed when a new technology is implemented or rejected and the impact on the health system to be lacking in HTA reports – even though this can influence the behaviour of health professionals and may help to overcome barriers to implementation.

Whilst implementation is not strictly considered an integral part of the HTA process, when it is discussed, it is often inconsistent and not formally set out(328). To what extent costs of implementation strategies should be included sequentially or simultaneously(164, 329) as well as analysing how conventional ICERs can be adjusted using methods such as mathematical programming are areas of recent research to take account of specific constraints beyond the budget constraint within health systems(330-332). Hauck, K et al(332) extend the principles of conventional cost-effectiveness analysis to identify an optimal balance between investing in health system strengthening and expenditure on specific interventions. Whilst the standard focus of CEA is on the addition of an incremental intervention, they consider how a range of interventions may

depend on a common service delivery 'platform' as the costs and benefits of interventions are not always independent of each other given this common dependency(332). Revill et al suggest that the field of implementation science with its multi-disciplinary nature, could help inform economic models on the value and opportunity costs of other resources for implementation(331).

However, few practical applications of these approaches have so far been published(333). Indeed, it is acknowledged that such constraints itself signal that the health system needs strengthening in terms of health care inputs(334). Interestingly, some HICs now charge a fee for appraising pharma products including NICE. Rather than adding to this important technical body of work outlined above to address implementation constraints in the healthcare system where the hta is predominantly that of available / existing technologies, we take a different approach. Drawing on our programme theory, we consider the broad practical implications for decision-makers and researchers aiming to achieve greater interconnectedness between HTA and health systems, in LMIC in particular. Firstly, rather than maintaining an emphasis on user-focused hta ie where a synthesis of clinical evidence and economic evaluation of available technologies are the mainstay, we consider the application of development-focused hta where system constraints are addressed whilst the technology is still under development(335). Secondly, although health products have been its more significant focus to-date, HTA could develop to increase its focus on 'technologies applied to health care' ie the regulatory and policy measures for managing and organising health care systems and on policies in non-health care sectors(78). We consider a refocus towards meso and/or macro HTA given the interdependency of HTA and health systems for HTA impact. Both are discussed below.

10.3.1 Early or development focused HTA

For many technologies applied to the system, the question of whether the intervention works, and even the design of the intervention, is essentially context-specific. Innovate early stage HTA has been proposed as a way of identifying critical aspects early on, including organisational-related factors, thus improving and facilitating greater clinical acceptance and implementation of cost-effective technologies(336). Organisational issues are affected by (as

well as affect) the implementation of a technology. Patient-related issues are also important for uptake. A review found these aspects were included rarely and inconsistently in htas as compared with technological, clinical and economic aspects(337). Expert and professional opinion, political judgment, the interpretation of values and traditions, and views from stakeholders and contingencies are all relevant inputs in the decision-making and the formulation of policy. Lomas et al cited in (338) refer to these forms of information as 'colloquial evidence' as opposed to information emanating from the application of scientific methods. Since evidence from research on contextual factors is frequently limited, or sometimes entirely lacking, decision-makers most often rely on such colloquial evidence. 'If HTA does not improve its responsiveness to context-dependent issues, colloquial evidence will remain the main source of information for decision-makers and HTA will fail to establish a robust bridge over this particular knowledge-action gap'(339).

To address this, we propose a greater use of early or development focused HTA. The form of HTA used to inform developers of health technologies has been termed 'early HTA' in the academic literature (340). Bouttell et al prefer to use the label 'development-focused HTA' as it is the audience, rather than the timing of the HTA, which drives many of the differences between this and the more familiar use-focused HTA(335). By drawing on a broad range of multidisciplinary methods, it aims to inform the developers of the technology about a wider range of questions including how the technology should be designed, used and/or priced. As such, it relates to the optimisation of an individual technology during development (rather than an evaluation between available technologies as in more traditional user-focused HTA) and can be considered as contributing to three iterative and interlinked assessments of clinical value, economic value and the business case(335).

A key aspect of development-focused HTA is evidence gathering in the specific context where the technology will be used. By, for example, identifying points along the implementing pathway where things might breakdown when a technology is still at concept stage is a way of being able to deal with contextual-dependent evidence more scientifically, and less 'colloquially'(338) ie by considering contextual issues from the outset/ early on in the process. In

realism, context is key and in LMIC, this is even more crucial. We suggest that development-focused HTA can offer scientific, transparent and systematic methods (such as epidemiological analysis and qualitative methods) to gain a thorough understanding of the human factors, infrastructure and health care organisation of the context where the technology is to be deployed(335). Furthermore, this should also facilitate greater integration and communication between HTA and delivery systems stakeholders as evidence generating methods gather and synthesise many perspectives.

Development-focused HTA is an expanding field, and we believe that the tools of development-focused HTA could be usefully employed to prioritise the development of technologies which impact most upon achieving UHC. To date, we are aware that the Bill and Melinda Gates Foundation have expressed interest in the exploration of these methods. Indeed, in addressing LMIC's need for more of the 'right' new products and fewer 'wrong' products, early HTA has been promoted as a tool to support this(341). This has been exemplified by the Covid-19 pandemic. As LMIC prepare for the COVID-19 outbreak, we propose the use of development HTA where developers need to rapidly evaluate and optimise or adapt existing technologies(335). Furthermore, because LMIC generally lack an effective mechanism to 'pool' their disease burden and resources, the attractive commercial market for pharma is often lacking. To overcome this, Chalkidou et al(342) propose an innovative market-driven, value-based approach whereby richer countries offer advance purchase commitments. In order to ensure local needs and value are reflected in any new technological investments, innovations and prices, Chalkidou et al similarly recommend the use of early hta as an emerging practice and capability. They suggest early or development-focused HTA can 'help shape the development of COVID-19 vaccines by signalling preferable vaccine characteristics (for example, safety and efficacy profiles, mode of administration, number of vaccine doses, vaccine storage requirements, target populations, vaccine purchasing, and delivery costs) that can maximise the public health impact'(343).

10.3.2 Meso and macro HTA

Whilst development HTA is about expanding the breadth of HTA to consider health system constraints whilst a technology is still in development, here, we

consider expanding the remit of HTA by broadening its scope to inform not only technology reimbursement decisions but also health system organisation and service delivery. As Garrido et al state(339), 'although health products and health care services have been its preponderant focus to date, HTA should develop to increase its focus on the 'technologies applied to health care' (ie the regulatory and policy measures for managing and organising health care systems) and on policies in non-health care sectors'. Indeed, the recommendation by the evaluators of the China CP was that the hta should be integrated into the development of other systems in order to change health system functioning, and cannot be a standalone reform(179). We suggest that it is potentially more impactful to focus on hta at the meso and macro level in resource poorer contexts as this may be one way to better integrate HTA within healthcare delivery systems. So, a refocus is not just about how we do HTA (methods etc) but what we evaluate.

It is noteworthy that iDSI state(17) 'HTA is not suitable for addressing all health systems problems. HTA will not help with general health system problems involving financing issues (for example, domestic resource mobilisation) or planning the health workforce'. We would say this is an opportunity overlooked. Towse et al(32, 344) state that macro HTA can support the efficiency, organisation and strengthening of the healthcare system, and define macro technologies as comprising elements of the architecture or framework such as how the system is organised, including number and types of hospitals and physicians. Given the poor infrastructure in many LMIC, 'macro HTA aimed at developing performance in the healthcare system may be of greater importance in this context than in HICs where HTA has had a more traditional micro HTA role of appraisal of single/related technologies'(32). The conventional focus of HTA on technologies that are marginal or incremental to the system is still relevant but a refocus to the process or systems-wide interventions (for example, by supporting regulatory measures for managing and organising health care systems), may result in an improvement in methods used to deliver an existing technology, or help develop innovative ways of overcoming barriers and challenges in adoption behaviour or infrastructure(335). Indeed, the most successful examples found in our realist review are of several meso level hta necessitating the adaptation of service provision to manage patients' care

pathways within a health system – and integrating macro approaches including organisational and structural support. This combination of integrated support provided more than just capacity but reinforced the change commitment and efficacy of practitioners.

Finally, Garrido et al(339) recommend that 'countries embarking on HTA should not consider establishing separate agencies for HTA, quality development, performance measurement, and health services development but should rather combine these functions and goals into a common knowledge strategy for evidence-informed decision-making on health care and the health system'. Thomas and Chalkidou(345) discuss the benefits and challenges of applying economic evaluation at the macro level. Nevertheless, HTA's rigorous and transparent way of synthesising evidence should be followed when addressing health system issues despite any associated methodological challenges(339).

10.4 Future research

10.4.1 More empirical case studies

There is potential for further illustrative empirical case studies through two direct contacts:

1) INAHTA invitation (June 2018) to follow up for future collaboration on impact. This would more likely veer toward HIC case studies but the framework is applicable to all countries looking to measure impact and value for money. INAHTA have recently published a report (February 2020) on their members' practices relating to impact assessment(109). This offers a potential platform on which to build and take this framework forward.

2) the Agency for Care Effectiveness (ACE), Healthcare Performance Group, Ministry of Health, Singapore are exploring the feasibility of using aggregated and patient-level drug prescription/dispense data to monitor the impact of subsidy decisions. Our contact is the Lead Specialist (Economic Modelling & Outcome Evaluation) and the CEO. They are happy to share more as they progress.

In any future case studies, in order to assess the robustness of results, different methods could be employed. For the modelling of a technology's cost-effectiveness as part of the hta, assumptions on model parameters are typically handled probabilistically ie each input parameter is assigned a specific sampling distribution. The framework parameters could also be handled in this way with distributions so that the NHB-ROI is expressed as a measure of the degree of certainty in the final results. In order to handle the impact assessment probabilistically would require the use of probabilistic assessments of all of the underlying htas. Treating the impact assessment parameters probabilistically with fixed underlying cost-effectiveness from the htas would not fully represent uncertainty. Aggregate statistics for probabilistic NHB of each underlying hta could be used to simplify this in the (likely) event of not having a fully probabilistic model for each hta.

However, a full probabilistic assessment is unlikely to be feasible – hence a focus on the use of standard deterministic sensitivity analysis to represent uncertainty in the overall system level HTA could be more straightforwardly used. One-way or multi-way deterministic sensitivity analysis could be utilised to explore the impact of parameters on the NHB-ROI result either by varying the parameters one at a time or analysing how combinations of variations affect the results respectively. This could include varying parameters on the value of the threshold, uptake of the intervention, costs of the hta etc. The use of tornedo diagrams showing the relative importance and impact of different variables could be used to then depict the results visually. Scenario analysis, based on subsets of these analyses, can be presented. A threshold analysis to establish the critical value/s of a parameter or parameters for a positive NHB-ROI could also be identified.

10.4.2 Full application of the ROI framework

A natural progression of the case studies would be to apply the full framework to a given jurisdiction. For example, HITAP in Thailand, has considered lending itself as an application of the wider framework whereby a representative sample of htas could be evaluated. Operationalising the framework to assess the ROI of an entire country's HTA programme is unlikely to be feasible in any context.

10.4.3 Data linkage

Data linkage to health outcomes could be explored to validate this work. A 2018 OECD survey on routine data collection states the richest data sources extend beyond prescribing and dispensing to information on utilisation of health services and outcomes of treatment(327). Although this survey identified at least 25 OECD and EU Member States that collect routine data on prescribed and dispensed medicines, linkage with other health or health care datasets is less common. The OECD recommends that methods to generate evidence from routinely collected data….need to be further developed and gain greater legitimacy and recognition from HTA agencies(327). Countries with good data linkage and the established use of HTA might be Scotland (linkage to Scottish Morbidity Records), or possibly Taiwan. This is only likely to be available in HICs. Nevertheless, it still comes with wider learning opportunities to other countries.

10.4.4 Realist evaluation

Realist interviewing could help to further test and refine theory by collecting primary qualitative data. A continuing test of the same theory using primary data would allow us to explore aspects of the theory in more depth with different stakeholders. This could tie into further empirical case studies.

10.4.5 Expansion to Health Systems Strengthening

Health systems strengthening (HSS) is very central in global health discussions. In theory, this NHB-ROI framework could be extended to apply not just to assessing HTA impact but also a way of capturing the impact and value for money of HSS initiatives with the same focus on efficiency and health outcomes. Our impact framework measures impact as potential, realised and attributable NHBs. These could be written up as a typology of HSS estimates. A potential starting place might be collaborating with iHEA special interest group, and could complement existing methods to measure health systems efficiency(345).

10.5 Final reflections

In this chapter, we outline the main strengths (Table 10-1 "What our framework adds") and challenges (Section 10-2). We also draw out implications of this research for decision-makers and researchers (Section 10-3) and recommendations for future research (Section 10-4). Rather than risk repeat what is stated, here I take a more reflective review of the main strengths and weaknesses of the whole body of research.

Our framework uniquely brings together the costs of HTA/hta as well as quantifying and understanding impact of HTA/hta as a process or complex intervention of itself – and thus, goes beyond the impact of the costs and benefits of the technology in question. Whilst simplifying assumptions were required, we consider these to go beyond those made by others in that we have attempted to address the counterfactual, attribution and actual implementation which we express as realised and attributable NHBs. Furthermore, expressing impact in NHB (or NMB) allows us to go beyond thinking solely in terms of cost-savings but to capture health and efficiency gains - with the realist approach capturing other potential outcomes. NHB importantly reflect the opportunity costs of investing in individual technologies as well as in HTA at a systems level. Using modelled health outcomes allows us to focus on improving implementation over which we have direct influence and thereby, to realise health gains.

As a researcher, I have learnt a huge amount about different disciplines and methods across the evaluation spectrum over the course of this work. Expanding on the usual health economist's repertoire has been invaluable. Undertaking a realist synthesis as a complementary and congruent approach to iDSI's ToC and then drawing on quasi-experimental methods and finally, linking this to an ROI framework mirrors the evaluation method groups as set out by the Magenta 2020 book(131). These are namely, theory-based impact evaluation methods, experimental and quasi-experimental impact evaluation methods and value-for-money methods. As such, the novel aspect could be seen to be less about using mixed methods but rather that we combine positivist and quasi-experimental approaches with complexity theorising.

The strength, however, of developing a new conceptual framework has led, perhaps, to a weakness in this research in that illustrations of its application are proof-of-methods case studies only. A natural progression of these applications would be to apply the recommendations made around handling uncertainty in the framework as well as to apply the full framework in collaboration with a specific HTA agency or agencies. As stated before, we recognise that the data requirements would not be without their challenges. However, this would be an opportunity to take this framework forward with stakeholders. The iDSI ToC was developed in full consultation with stakeholders and partners in the countries concerned. A weakness is that our framework might be seen as a somewhat academic exercise but we have potential for collaborations to take this forward with our HTA partners. Furthermore, the recent INAHTA Impact report (February 2020)(109) provides a relevant and timely platform upon which to build.

As I was writing this section, I received two relevant emails. One was on REF and the latest thinking and evidence on effective impact assessment. The other was NICE's newsletter, entitled "Making a difference: the impact of NICE's work". To quote NICE's December 2020 newsletter: "Over the last 20 years, NICE has established itself as a global leader in the development of evidence-based guidance for health and social care, using robust, transparent methodologies and processes. We have one of the biggest guideline programmes in the world, with topics covering a range of clinical conditions, social care and public health. But it is only by putting this advice into practice that it will make a difference to real people, to health outcomes and to equitable access to services". NICE discuss their National Clinic Audit Programme, a new version of their Innovation Scorecard which reports on the use of medicines and medical technologies in the NHS in England which have been positively appraised, NICE Impact Reports now covering 17 topics, and shared learning case studies. The latter two especially contain a wealth of quantitative data on uptake following guidance and qualitative data on implementation. Indeed, these were reviewed as sources of information for this research and offer potential for future case studies.

Furthermore, commentary and analysis on CGD's website (30 November 2020) highlights that with the onset of COVID-19, SSA countries face economic contraction compounding existing pressures on healthcare budgets. Institutionalised priority setting processes are needed to ensure that investments in health continue to strive for improving overall population health, protecting citizens from impoverishment, and enhancing the quality of services. Leveraging expertise and resources across the global movement of priority setting in health to help support all countries in their COVID-19 recovery is paramount[15].

These current excerpts illustrate the importance of measuring and optimising HTA's impact in the current climate. Highlighting the implications drawn from this research on macro and development-focused hta to strengthen health systems and embed implementation issues in hta from the outset respectively can, hopefully, be seen as timely and useful contributions to this field.

[15] Afro-European Partnerships in Health: Accelerating Better Efficiency of Health Spending | Center For Global Development (cgdev.org)

Annexes

Annex A: HTA impact articles included in Chapter 2 review

Annex B: Realist synthesis search terms and strategy

Annex C: Grey literature search results

Annex A: HTA impact articles retrieved

TOWARD INTEGRATION IN THE CONTEXT OF HEALTH TECHNOLOGY ASSESSMENT: THE NEED FOR EVALUATIVE FRAMEWORKS.	International Journal of Technology Assessment in Health Care. 33(5):586-590, 2017 Jan.	van der Wilt GJ, Gerhardus A, Oortwijn W
INSIGHTS FROM THE FRONT LINES: A COLLECTION OF STORIES OF HTA IMPACT FROM INAHTA MEMBER AGENCIES.	International Journal of Technology Assessment in Health Care. 33(4):409-410, 2017 Jan.	Schuller T, Soderholm Werko S
Most important barriers and facilitators of HTA usage in decision-making in Europe.	Expert Review of Pharmacoeconomics & Outcomes Research. 18(3):297-304, 2018 Jun.	Cheung KL, Evers SMAA, De Vries H, Levy P, Pokhrel S, Jones T, Danner M, Wentlandt J, Knufinke L, Mayer S, Hiligsmann M
MOST IMPORTANT BARRIERS AND FACILITATORS REGARDING THE USE OF HEALTH TECHNOLOGY ASSESSMENT.	International Journal of Technology Assessment in Health Care. 33(2):183-191, 2017 Jan.	Cheung KL, Evers SMAA, de Vries H, Hiligsmann M
Economic evaluations of health technologies in Dutch healthcare decision-making: a qualitative study of the current and potential use, barriers, and facilitators.	BMC Health Services Research. 17(1):89, 2017 Jan 26.	Roseboom KJ, van Dongen JM, Tompa E, van Tulder MW, Bosmans JE

INFLUENCE OF HEALTH TECHNOLOGY ASSESSMENT AND ITS MEASUREMENT. [Review]	International Journal of Technology Assessment in Health Care. 32(6):376-384, 2016 Jan.	Hailey D, Werko S, Rosen M, Macpherson K, Myles S, Gallegos Rivero V, Hipolito-Olivares C, Sihvo S, Pwu J, Yang WW, Chen YC, Perez Galan A, Aleman A, Villamil E
Developing an evidence-based methodological framework to systematically compare HTA coverage decisions: A mixed methods study. [Review]	Health Policy. 120(1):35-45, 2016 Jan.	Nicod E, Kanavos P
The impact of the National Institute for Health Research Health Technology Assessment programme, 2003-13: a multimethod evaluation.	Health Technology Assessment (Winchester, England). 19(67):1-291, 2015 Aug.	Guthrie S, Bienkowska-Gibbs T, Manville C, Pollitt A, Kirtley A, Wooding S
Impact of NIHR HTA Programme funded research on NICE clinical guidelines: a retrospective cohort.	Health Research Policy & Systems. 13:37, 2015 Aug 22.	Turner S, Bhurke S, Cook A
Assessing the impact of health technology assessment on the Austrian healthcare system.	International Journal of Technology Assessment in Health Care. 29(1):84-91, 2013 Jan.	Schumacher I, Zechmeister I
Can we reliably benchmark health technology assessment organizations?.	International Journal of Technology Assessment in Health Care. 28(2):159-65, 2012 Apr.	Drummond M, Neumann P, Jonsson B, Luce B, Schwartz JS, Siebert U, Sullivan SD

The impact of health technology assessment reports on decision making in Austria.	International Journal of Technology Assessment in Health Care. 28(1):77-84, 2012 Jan.	Zechmeister I, Schumacher I
Health technology assessment to optimize health technology utilization: using implementation initiatives and monitoring processes. [Review]	International Journal of Technology Assessment in Health Care. 26(3):309-16, 2010 Jul.	Fronsdal KB, Facey K, Klemp M, Norderhaug IN, Morland B, Rottingen JA
Developing Health Technology Assessment to address health care system needs. [Review] [34 refs]	Health Policy. 94(3):196-202, 2010 Mar.	Velasco Garrido M, Gerhardus A, Rottingen JA, Busse R
Assessing the impact of England's National Health Service R&D Health Technology Assessment programme using the "payback" approach.	International Journal of Technology Assessment in Health Care. 25(1):1-5, 2009 Jan.	Raftery J, Hanney S, Green C, Buxton M
Assessing the impact of health technology assessment in The Netherlands.	International Journal of Technology Assessment in Health Care. 24(3):259-69, 2008.	Oortwijn WJ, Hanney SR, Ligtvoet A, Hoorens S, Wooding S, Grant J, Buxton MJ, Bouter LM
Assessing the performance of health technology assessment organizations: a framework. [Review] [87 refs]	International Journal of Technology Assessment in Health Care. 24(1):76-86, 2008.	Lafortune L, Farand L, Mondou I, Sicotte C, Battista R

The role of health technology assessment in the comprehensive evaluation of the impact of immunotherapy on real practice.	European Annals of Allergy & Clinical Immunology. 39 Spec No:4-6, 2007.	Frati F, Ariano R, Cadario G, Ortolani C, Passalacqua G
An assessment of the impact of the NHS Health Technology Assessment Programme. [Review] [103 refs]	Health Technology Assessment (Winchester, England). 11(53):iii-iv, ix-xi, 1-180, 2007 Dec.	Hanney S, Buxton M, Green C, Coulson D, Raftery J
Integration of health technology assessment recommendations into organizational and clinical practice: A case study in Catalonia.	International Journal of Technology Assessment in Health Care. 22(2):169-76, 2006.	Gagnon MP, Sanchez E, Pons JM
Dissemination of health technology assessments: identifying the visions guiding an evolving policy innovation in Canada.	Journal of Health Politics, Policy & Law. 30(4):603-41, 2005 Aug.	Lehoux P, Denis JL, Tailliez S, Hivon M
End-user involvement in health technology assessment (HTA) development: a way to increase impact.	International Journal of Technology Assessment in Health Care. 21(2):263-7, 2005.	McGregor M, Brophy JM
Quality improvement in an outpatient department for subacute low back pain patients: prospective surveillance by outcome and performance measures in a health technology assessment perspective.	Spine. 29(8):925-31, 2004 Apr 15.	Johansen B, Mainz J, Sabroe S, Manniche C, Leboeuf-Yde C

Impact of health technology assessments. Some experiences of SBU.	International Journal of Technology Assessment in Health Care. 18(4):824-31, 2002.	Britton M, Jonsson E
Impact of health technology assessment on preventive screening in Belgium: case studies of mammography in breast cancer, PSA screening in prostate cancer, and ultrasound in normal pregnancy.	International Journal of Technology Assessment in Health Care. 17(3):316-28, 2001.	Vermeulen V, Coppens K, Kesteloot K
The use and impact of rapid health technology assessments.	International Journal of Technology Assessment in Health Care. 16(2):651-6, 2000.	Hailey D, Corabian P, Harstall C, Schneider W
Prioritizing investments in health technology assessment. Can we assess potential value for money?.	International Journal of Technology Assessment in Health Care. 16(1):73-91, 2000.	Davies L, Drummond M, Papanikolaou P
Assessing the impact of health technology assessment.	International Journal of Technology Assessment in Health Care. 13(1):68-80, 1997.	Jacob R, McGregor M
What has health technology assessment ever done for us?.	Journal of Health Services Research and Policy. 23 (2) (pp 134-136), 2018. Date of Publication: 2018.	Loblova O.
Influence of a health technology assessment on the use of paediatric cochlear implantation in Kazakhstan.	Health Policy and Technology. 7 (3) (pp 239-242), 2018. Date of Publication: September 2018.	Kosherbayeva L., Medeulova A., Hailey D., Yermukhanova L., Uraz R., Aitmanbetova A.

Impact assessment of Iran's health technology assessment programme.	Health Research Policy and Systems. 16 (1) (no pagination), 2018. Article Number: 15. Date of Publication: 22 Feb 2018.	Yazdizadeh B., Mohtasham F., Velayati A.
The impact of parallel regulatory-health technology assessment scientific advice on clinical development. Assessing the uptake of regulatory and health technology assessment recommendations.	British Journal of Clinical Pharmacology. 84 (5) (pp 1013-1019), 2018. Date of Publication: May 2018.	Tafuri G., Lucas I., Estevao S., Moseley J., d'Andon A., Bruehl H., Gajraj E., Garcia S.Hedberg N.,Massari M., Molina A.,Obach M., Osipenko L.,Petavy F.,Petschulies M., Pontes C, Russo P, Schiel A.,Van de Casteele M.,Zebedin-Brandl E.-M.,Rasi G,Vamvakas S.
Best-worst scaling to assess the most important barriers and facilitators for the use of health technology assessment in Austria.	Expert Review of Pharmacoeconomics and Outcomes Research. 18 (2) (pp 223-232), 2018. Date of Publication: 04 Mar 2018.	Feig C., Cheung K.L., Hiligsmann M., Evers S.M.A.A.,Simon J., Mayer S.
Models and applications for measuring the impact of health research: Update of a systematic review for the health technology assessment programme.	Health Technology Assessment. 20 (76) (no pagination), 2016. Date of Publication: October 2016.	Raftery J., Hanney S., Greenhalgh T., Glover M., Blatch-Jones A.
Value of health technologies, HTA and decisions.	Global and Regional Health Technology Assessment. 4 (1) (pp e41-e43), 2017. Date of Publication: 2017.	Lucioni C., Jommi C.

EMERGENCE of DRIVERS for the IMPLEMENTATION of HEALTH TECHNOLOGY ASSESSMENT.	International Journal of Technology Assessment in Health Care. 32 (4) (pp 300-306), 2016. Date of Publication: 2016.	Jaramillo H.E.C., Moreno-Mattar O.,Osorio-Cuevas D.
An assessment of the impact of the NHS Health Technology Assessment Programme.	Health Technology Assessment. 11 (53) (pp iii-82), 2007. Date of Publication: December 2007.	Hanney S., Buxton M.,Green C., Coulson D.,Raftery J.
Health technology assessment within a public accountability framework.	Clinical Governance. 9 (1) (pp 51-58), 2004. Date of Publication: 2004.	Kazanjian A., Green C.J.
Does health technology assessment benefit health services and politics? The experiences of an established HTA institution: The Danish centre for evaluation and HTA.	European Journal of Health Economics. 3 (1) (pp 54-58), 2002. Date of Publication: 2002.	Sigmund H., Kristensen F.B.
Implementing the findings of health technology assessments: If the CAT got out of the bag, can the TAIL wag the dog?.	International Journal of Technology Assessment in Health Care. 16 (1) (pp 1-12), 2000. Date of Publication: Winter 2000.	Drummond M., Weatherly H.
From Health Technology Assessment to research on the organisation and delivery of health services: addressing the balance		N. Fulop, P. Allen, A. Clarke, N. Black

An Exploratory Review of Evaluations of Health Technology Assessment Agencies		Margaret Wanke Don Juzwishin Richard Thornley Liza Chan
IMPACT of HEALTH TECHNOLOGY ASSESSMENT REPORTS on HOSPITAL DECISION MAKERS - 10-YEAR INSIGHT from A HOSPITAL UNIT in SHERBROOKE, CANADA: IMPACT of HEALTH TECHNOLOGY ASSESSMENT on HOSPITAL DECISIONS.	International Journal of Technology Assessment in Health Care. 34 (4) (pp 388-392), 2018. Date of Publication: 2018.	Poder T.G., Bellemare C.A.Bedard S.K.,Fisette J.-F. Dagenais P.
Outcomes and Impacts of 10-Year HTA Implementation in Taiwan.	International Journal of Technology Assessment in Health Care. 35 (6) (pp 441-445), 2019. Date of Publication: 2019.	Kao K.-L., Huang L.-Y., Wu Y.-H., Gau C.-S.
Threats to the value of Health Technology Assessment: Qualitative evidence from Canada and Poland.	Health Policy. 123 (2) (pp 191-202), 2019. Date of Publication: February 2019.	Wranik W.D., Zielinska D.A., Gambold L., Sevgur S.
Health Technology Assessment as Part of a Broader Process for Priority Setting and Resource Allocation.	Applied Health Economics and Health Policy. 17 (5) (pp 573-576), 2019. Date of Publication: 01 Oct 2019.	Mitton C., Seixas B.V., Peacock S., Burgess M., Bryan S.
An impact analysis of the implementation of health technology assessment for new treatment of orphan diseases in Japan.	Expert Review of Pharmacoeconomics and Outcomes Research. (no pagination), 2019. Date of Publication: 2019.	Kogushi K., Ogawa T., Ikeda S.

Impact of health technology assessment and managed entry schemes on reimbursement decisions of centrally authorised medicinal products in Belgium.	European Journal of Clinical Pharmacology. (no pagination), 2019. Date of Publication: 2019.	Van Wilder P., Pirson M., Dupont A.
The implementation of HTA in medicine pricing and reimbursement policies in Indonesia: Insights from multiple stakeholders.	PLoS ONE. 14 (11) (no pagination), 2019. Article Number: e0225626. Date of Publication: 2019.	Wasir R., Irawati S., Makady A., Postma M., Goettsch W., Feenstra T., Buskens E.
Does the use of health technology assessment have an impact on the utilisation of health care resources? Evidence from two European countries.	European Journal of Health Economics. 21 (4) (pp 621-634), 2020. Date of Publication: 01 Jun 2020.	Corbacho B., Drummond M., Santos R., Jones E., Borras J.M., Mestre-Ferrandiz J., Espin J., Henry N., Prat A.
Returns on Research Funded Under the NIHR Health Technology Assessment (HTA) Programme: Economic Analysis and Case Studies.	Rand Health Quarterly. 2016;5(4):5.	Guthrie S, Hafner M, Bienkowska-Gibbs T, Wooding S.
How Can Payback from Health Services Research Be Assessed?	Journal of Health Services Research. 1996;1(1):35-43.	Buxton M, Hanney S.
"Which policies increase value for money in health care?,"	OECD Health Working Papers 104, OECD Publishing. 2018	Luca Lorenzoni & Fabrice Murtin & Laura-Sofia Springare & Ane Auraaen & Frederic Daniel

Health Technology Assessment in Malaysia: Past, Present, and Future.	International Journal of Technology Assessment in Health Care. 2019;35(6):446-51.	Roza S, Junainah S, Izzuna MMG, Ku Nurhasni KAR, Yusof MAM, Noormah MD, et al.
+ HTA Impact Framework + HTA impact assessment study February 2020, Parts I - II	INAHTA – see grey lit review [Annex C]	INAHTA members

Annex B: Realist Synthesis Search terms and strategy

The search strategy proceeded iteratively in four phases:

1. Initial search to identify existing frameworks or models to measure the impact of HTA [as per Chapter 2] and a relatively unstructured scan to develop tentative theories on HTA impact. Search terms for this initial search were first run with the terms shown in Box 1 but given the large number of articles returned, it was re-run with the adjacent term "*adj5 health technolog$ or HTA"* in lines 4 and 5 (Box 2). Box 2 results will be a subset of the results retrieved in Box 1.

Run on Ovid MEDLINE(R) without revisions, Ovid MEDLINE In-Process & Other Non-Indexed Citations, EMBASE 1996 to 1st June 2020.

Box 1
1. exp Technology Assessment, Biomedical/
2. ((health technolog$ adj3 assessment$) or HTA).mp.
3. 1 or 2
4. ("evaluat$" or "method$" or "framework$" or "model$" or "concept$" or "empirical research" or "theor$" or "cost benefit analys$").tw.
5. ("impact$" or "value" or "gain$" or "benefit$" or "influence$" or "return$ on investment" or "ROI" or "social adj3 return$" or "economic adj3 return$").tw.
6. 4 or 5
7. 3 and 6

Box 2
1. exp Technology Assessment, Biomedical/
2. ((health technolog$ adj3 assessment$) or HTA).mp.
3. 1 or 2
4. (("evaluat$" or "method$" or "framework$" or "model$" or "concept$" or "empirical research" or "theor$" or "cost benefit analys$") adj5 (health technolog$ or HTA)).tw.
5. (("impact$" or "value" or "gain$" or "benefit$" or "influence$" or "return$ on investment" or "ROI" or "social adj3 return$" or "economic adj3 return$") adj5 (health technolog$ or HTA)).tw.
6. 4 or 5
7. 3 and 6

2. Detailed search to identify literature that could be used to elucidate, test and refine those tentative theories. This focuses specifically on i) HTA & health outcomes; ii) HTA & 'uptake/implementation' as a proxy measure given the identified lack of literature on impact on health outcomes; and iii) governance and knowledge transfer/utilisation theories with respect to HTA.

Run on Ovid MEDLINE(R) without revisions, Ovid MEDLINE In-Process & Other Non-Indexed Citations, EMBASE 1996 to 1st June 2020.

HTA and health outcomes

1. exp Technology Assessment, Biomedical/
2. ((health technolog$ adj3 assessment$) or hta).mp.
3. 1 or 2
4. (("health" or "patient") adj5 ("gain$" or "outcome$" or "benefit$")).tw.
5. ("net health benefit$" or "net monetary benefit$" or "NMB$" or "NHB$").tw.
6. exp Health Impact Assessment/
7. "Outcome Assessment (Health Care)"/
8. 3 and (4 or 5 or 6 or 7)

HTA and uptake

1. exp Technology Assessment, Biomedical/
2. ((health technolog$ adj3 assessment$) or hta).mp.
3. 1 or 2
4. ("implement$" or "uptake" or "diffusion" or "utilis$" or "utiliz$" or "adopt$" or "practice").tw.
5. 3 and 4

HTA, or health outcomes or evidence and theory of governance/knowledge

1. exp Technology Assessment, Biomedical/
2. ((health technolog$ adj3 assessment$) or hta).mp.
3. (("health" or "patient") adj5 (gain$ or outcome$ or benefit$)).mp.
4. ("evidence" adj5 ("implement$" or "uptake" or "diffusion" or "utilis$" or "utiliz$" or "adopt$" or "practice")).mp.
5. 1 or 2 or 3 or 4
6. (governance adj2 (theor$ or model$ or concept$ or framework$)).mp.
7. (knowledge adj2 (theor$ or model$ or concept$ or framework$)).mp.
8. 5 and (6 or 7)

1. exp Technology Assessment, Biomedical/ adj2 (("knowledge" or "governance") and (theor$ or model$ or concept$ or framework$)).mp.
2. (((health technolog$ adj3 assessment$) or HTA) adj2 (("knowledge" or "governance") and (theor$ or model$ or concept$ or framework$))).mp.
3. (("health" or "patient") adj5 (gain$ or outcome$ or benefit$) adj2 (("knowledge" or "governance") and (theor$ or model$ or concept$ or framework$))).mp.
4. ("evidence" adj5 ("implement$" or "uptake" or "diffusion" or "utilis$" or "utiliz$" or "adopt$" or "practice") adj2 (("knowledge" or "governance") and (theor$ or model$ or concept$ or framework$))).mp.

3. Ongoing reference, citation and author tracking to identify the most theory relevant studies available.

4. Finally, very refined searches, again operating concurrently with the synthesis process, to collect additional materials that may be required to elucidate particular aspects of theory. This included expanding the search to other areas, for example, in organisational theory where it was inferred that similar mechanisms around willingness and capability might be activated.

Annex C: Grey literature search results

The following list of websites were searched using the keywords: HTA impact and implementation

Website	Webpage url if specific sections were searched	Downloaded
NICE ERNIE database	http://webarchive.nationalarchives.gov.uk/20080611115838/ http://www.nice.org.uk/usingguidance/evaluationandreviewofniceimplementationevidenceernie/ niceimplementationuptakecommissionedreports/ nice_commissioned_surveys_and_reports.jsp http://webarchive.nationalarchives.gov.uk/20080611122835/ https://www.nice.org.uk/usingguidance/evaluationandreviewofniceimplementationevidenceernie/ niceimplementationuptakecommissionedreports/nice_implementation_uptake__commissioned_reports.jsp	All NICE internal and externally commissioned reports downloaded – see table below
NICE shared learning	https://www.nice.org.uk/localPractice/collection	Case studies
EUnetHTA / Evidence Database on New Technologies	Database part of EUnetHTA Joint Action on HTA 3	EVIDENT Not accessible
CADTH	Strong emphasis in latest strategy on implementation	Latest strategy documentation
TRIP	HTA impact and implementation	No access to advanced search
INAHTA	HTA Impact and Influence https://www.inahta.org/hta-tools-resources/hta-impact-influence/#IAStudy	HTA impact framework & Report 2020

ERNIE Database

Report	Key findings
Pettit, JL et al. Assessing the implementation of NICE guidance: is there a correlation between recommendations and uptake in clinical practice?	Uses ERNIE uptake against NIICE projections and finds in the whole a correlation. This is a before/after study.
NICE Implementation – Breakthrough Breast Cancer	Lack of forward planning, resources, awareness, general need for monitoring systems.
Variation in usage of Cancer drugs approved by NICE. Report of the Review undertaken by the National Cancer Director.	Capacity issues, clinical factors. Solutions – information, prospective audit and feedback.
Implementation of NICE Guidance – An ABPI perspective	Generally compared data on uptake against NICE targets. Guidance and implementation planning should be linked, accountability, awareness, education, audit, monitoring, and incentives.
Audit in the implementation of NICE Guidance for Roche Drugs	Compared audit data on uptake against potential eligible patient population. Challenging to do so.
Implementation of NICE Guidance # 49	Abacus International posted questionnaire to anaesthetists. Disclaimer that 'no statistical analysis has been conducted as this survey is simply designed to give a top line picture of the impact and implementation of guidance #49'. Illustrates not all practitioners agree/concur with guidance. 46% had no access/little access to the required technology required to implement the guidance. Barriers included infrastructure issues like training ie not just funding and access to necessary equipment. Some no change to practice as they were already complying.
Implementation of NICE Guidance # 24	Abacus Int and disclaimer as above. Not all felt guidance had an impact and sometimes things in place anyway but had not been implemented as a result of NICE recommendations. Guidance been read, understood but not brought about a major change. One of the issues was that the guidance did not make specific recommendations and was described as woolly: "there is nothing new in it", guidance too narrow in scope or simply matched current practice and would therefore not expect any dramatic changes.
Abacus International developed a methodology to measure the	Identify data, consult with NICE to establish the expected change in HTA utilisation against what NICE had predicted. Infrastructure barriers, training and resources, guidance sometimes endorsing

impact of 28 pieces of NICE guidance.	current practice. "It is difficult to interpret the impact of NICE guidance because it was issues at time of X launch". Some show pre-guidance uptake and post-guidance uptake but may have been on an up/downward trend anyway. Predominantly before/after studies.
A survey measuring the impact of NICE Guidance # 11 Abacus Int	68% found not easy to implement. Some disagreement with recommendations usually due to a lack of perceived evidence base. Belief that it is a financial not clinical perspective driving these recommendations. Should be some element of flexibility. Lack of funding/pressure. PCT has competing priorities. Majority felt they were already implementing some of the key priorities prior to guideline publication. 62% felt that the Govt pressures had a high/very high impact on the implementation of this guideline. Less than half of all PCTs reported patient group pressure having the same level of impact. Other external pressures – patients, consultants, public health, legal depts., SHA, media, local taskforce and professional best practice.
NICE Guidance – Ritizole	Pre/post uptake shows an upward trend anyway. ITS would have provided estimates of significant change in trend.
NICE Guidance – Cox II	Provide national trends and activity
NICE Guidance - insulin glargine	Uptake greater than anticipated in NICE guidance as now more/better data on prevalence.
NICE Guidance – Attention Deficit hyperactivity Disorder	"It is not anticipated that this guidance will result in a major increase over current trends in the rate of prescribing for ADHD" – data shows steadily increasing as part of a longer term trend.
NICE Guidance – Insomnia: newer hypnotic drugs	The long term downward trend in this continues. Both trends pre-date the publication of NICE guidance. Whilst trend lines drawn on, no ITS undertaken. General lack of contextual detail available in these reports though some attempt to provide contextual info from a sample of patients' records from GP to gain insight into hypertension prescribing by age.
NICE Guidance – epilepsy: newer drugs	"Guidance expected to have a neutral impact on prescribing trends". The newer drugs are increasing at a faster rate than the older drugs.
NICE Guidance – lithium	Latest trend show that the prescribing rate is steady and in line with expectations. Lithium is an established treatment and continues to be widely prescribed. So a change in uptake not always expected.
NICE Guidance – statins	Estimated an additional 1.7million people will present for treatment as a result of guidance. There was though already a consistent upwards trend in statin usage.
NICE Guidance – laparoscopic surgery for colorectal cancer	The publication of NICE guidance appears to correspond with a further increase in the rate of uptake. It is too early to confirm a statistical link between guidance publication and change in

	uptake. The main limiting factor in uptake is a recognised shortage of surgeons skilled in this technique and patient choice.
NICE Guidance # 40 – Drugs in the management of urinary incontinence	This does use interrupted time series analysis. "This upturn was subjected to time series analysis to see if the apparent correspondence between the slight increase in uptake and the guidance publication could be confirmed statistically. The ARIIMA testing technique could not confirm a statistical link between guidance publication and change in uptake. However, this null finding may be because, at this stage, there are an insufficient number of monthly data points for oxybutynin usage since the guidance was published (sub-section 1.2.6 drug therapies).
NICE Guidance – Adefair dipivoxil to treat Hep B	Too early to confirm a statistical link between guidance publication and change in uptake. Modelled upper and lower trajectories to offer a basic way of understanding the potential uptake over time.

List of References

1. Luz A, Santatiwongchai B, Pattanaphesaj J, Teerawattananon Y. Identifying Priority Methodological Issues in Economic Evaluation in Low- and Middle-Income Countries: Finding the Holy Grail http://gear4healthcom/gear. 2017.
2. WHO. World Health Organisation: Health Technology Assessment. http://wwwwhoint/medical_devices/assessment/en/. Accessed Jun 2017.
3. WHO. Research for Universal Health Coverage 2013 [Available from: http://www.who.int/whr/en/.
4. PAHO. Resolution CSP28.R9: Health Technology Assessment and Incorporation into Health Systems. 2012.
5. WHO. Resolution SEA/RC66/R4 : Health Intervention and Technology Assessment in Support of Universal Health Coverage. 2013.
6. WHA. Resolution WHA67.3: Health Interventions and Technology Assessment in Support of Universal Health Coverage. 2014.
7. Piemonte C, Cattaneo O, Morris R, Pincet A, Poensgen K. Transition Finance: Introducing a new concept. Paris; 2019.
8. Bouttell J, Grieve E, Hawkins N. T he Role of Development-Focused Health Technology Assessment in Optimising Science, Technology and Innovation to Achieve Sustainable Development Goal 3. Oxford University Press; 2020 (in press).
9. Hollingworth S, Gyansa-Lutterodt M, Dsane-Selby L, Nonvignon J, Lopert R, Gad M, et al. Implementing health technology assessment in Ghana to support universal health coverage: building relationships that focus on people, policy, and process. International Journal of Technology Assessment in Health Care. 2020;36(1):8-11.
10. Gad M, Lord J, Chalkidou K, Asare B, Lutterodt MG, Ruiz F. Supporting the Development of Evidence-Informed Policy Options: An Economic Evaluation of Hypertension Management in Ghana. Value in Health. 2020;23(2):171-9.
11. iDSI. international Decision Support Initiative [Available from: https://idsihealth.org/.
12. Wirtz VJ, Hogerzeil HV, Gray AL, Bigdeli M, de Joncheere CP, Ewen MA, et al. Essential medicines for universal health coverage. The Lancet. 2017;389(10067):403-76.
13. Airoldi M, Morton A, Smith JA, Bevan G. STAR--people-powered prioritization: a 21st-century solution to allocation headaches. Medical decision making : an international journal of the Society for Medical Decision Making. 2014;34(8):965-75.
14. Glassman A, Chalkidou K. Priority-Setting in Health: Building Institutions for Smarter Public Spending. eSocialSciences; 2012.
15. Glassman A, Chalkidou K, Giedion U, Teerawattananon Y, Tunis S, Bump JB, et al. Priority-Setting Institutions in Health: Recommendations from a Center for Global Development Working Group. Global Heart. 2012;7(1):13-34.
16. World Health Organisation. 2015 Global Survey on Health Technology Assessment by National Authorities. 2015.
17. Jeffery M, Chi Y, Stewart M. iDSI Health Technology Assessment Toolkit. 2018. Contract No.: https://doi.org/10.7490/f1000research.1116126.1.
18. Lloyd R, Gleed G, Wallach S. iDSI Theory of Change Review Report. *F1000Research*; 2018.

19. Guthrie S, Hafner M, Bienkowska-Gibbs T, Wooding S. Returns on Research Funded Under the NIHR Health Technology Assessment (HTA) Programme: Economic Analysis and Case Studies. Rand Health Quarterly. 2016;5(4):5.
20. Timmermans S, Mauck A. The promises and pitfalls of evidence-based medicine. Health affairs (Project Hope). 2005;24(1):18-28.
21. Pettit L, Leonard S, Brooks-Rooney C, Hamerslag L, Kusel J. Assessing the Implementation of NICE Guidance: Is There a Correlation Between Recommendations and Uptake in Clinical Practice. 2013.
22. AuditCommission. Managing the Financial Implications of NICE Guidance. 2005.
23. Grol R. Successes and failures in the implementation of evidence-based guidelines for clinical practice. Med Care. 2001;39.
24. Jönsson B, Ramsey S, Wilking N. Cost effectiveness in practice and its effect on clinical outcomes. Journal of Cancer Policy. 2014;2(1):12-21.
25. Braithwaite J, Churruca K, Long JC, Ellis LA, Herkes J. When complexity science meets implementation science: a theoretical and empirical analysis of systems change. BMC Medicine. 2018;16(1):63.
26. Cooksey D. A Review of UK health research funding. UK; 2006.
27. ISPOR. [Available from: https://www.ispor.org/.
28. O'Rourke B, Oortwijn W, Schuller T. The new definition of health technology assessment: A milestone in international collaboration. International Journal of Technology Assessment in Health Care. 2020:1-4.
29. Grieve E, Hesselgreaves H, Wu O, Briggs A, iDSI. The Value of Health Technology Assessment: a mixed methods framework. https://doi.org/10.7490/f1000research.1115169.1; 2017.
30. Glassman A, Giedion U, Smith PC. What's In, What's Out: Designing Benefits for Universal Health Coverage. Washington DC Center for Global Development; 2017.
31. Handley MA, Gorukanti A, Cattamanchi A. Strategies for implementing implementation science: a methodological overview. Emergency Medicine Journal. 2016;33(9):660.
32. Towse A. HTA's Macro role in Health Care systems https://www.slideshare.net/OHENews/macro-hta-presentation-ispor-beijing-towse-20142014 [
33. Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. BMJ. 2008;336(7656):1281.
34. WK Kellogg Foundation. Logic Model Development Guide. Michigan; 2004.
35. Westhorp G. Email correspondence from member of the RAMESES JISC online research community. RAMESES List at WWWJISCMAILACUK2020.
36. Research Excellence Framework 2020 [Available from: https://impact.ref.ac.uk/casestudies/.
37. Cook DA, Ellaway RH. Evaluating technology-enhanced learning: A comprehensive framework. Medical Teacher. 2015;37(10):961-70.
38. Chang H-J. Institutions and economic development: theory, policy and history. Journal of Institutional Economics. 2010;7(4):473-98.
39. INAHTA. Published Evidence on the Influence of Health Technology Assessment. wwwinahtaorg. 2014.
40. Rudan I, Kapiriri L, Tomlinson M, Balliet M, Cohen B, Chopra M. Evidence-Based Priority Setting for Health Care and Research: Tools to Support Policy in Maternal, Neonatal, and Child Health in Africa. PLoS Medicine. 2010;7(7):e1000308.

41. Drummond M. Twenty years of using economic evaluations for drug reimbursement decisions: what has been achieved? J Health Polit Policy Law. 2013;38(6):1081-102.
42. Sorenson C, Drummond M, Kristensen F, Busse R. How can the impact of health technology assessments be enhanced? : WHO Regional Office for Europe, European Observatory on Health Systems and Policies; 2008.
43. Williams L, Brown H. Factors influencing decisions of value in healthcare: a review of the literature. Health Services Management Centre: University of Birmingham; 2014.
44. Goodman C. HTA 101: IX Monitor Impact of HTA 2014 [3rd:[Available from: www.nlm.nih.gov/nichsr/hta101/ta101011.html.
45. Glassman A, Fan V, Over M. More Health for the Money: Putting Incentives to Work for the Global Fund and Its Partners. 2013 2013.
46. Straus SE, Jones G. What has evidence based medicine done for us?: It has given us a good start, but much remains to be done. BMJ : British Medical Journal. 2004;329(7473):987-8.
47. O'Donnell JC, Pham SV, Pashos CL, Miller DW, Smith MD. Health Technology Assessment: Lessons Learned from Around the World—An Overview. Value in Health. 2009;12:S1-S5.
48. Lorenzoni Lea. "Which policies increase value for money in health care?". OECD Publishing, Paris; 2018.
49. Jacob R, McGregor M. Assessing the impact of health technology assessment. Int J Technol Assess Health Care. 1997;13(1):68-80.
50. Buxton M, Hanney S. How Can Payback from Health Services Research Be Assessed? Journal of Health Services Research. 1996;1(1):35-43.
51. Davies L, Drummond M, Papanikolaou P. Prioritizing investments in health technology assessment. Can we assess potential value for money? Int J Technol Assess Health Care. 2000;16(1):73-91.
52. Lafortune L, Farand L, Mondou I, Sicotte C, Battista R. Assessing the performance of health technology assessment organizations: a framework. Int J Technol Assess Health Care. 2008;24(1):76-86.
53. Wanke M. An Exploratory Review of Evaluations of Health Technology Assessment Agencies. 2006;Alberta Heritage Foundation for Medical Research (AHFMR) HTA Initiative #16.
54. INAHTA. INAHTA Impact Framework http://www.inahta.org/hta-tools-resources/briefs/#Impact [
55. Raftery J, Hanney S, Green C, Buxton M. Assessing the impact of England's National Health Service R&D Health Technology Assessment program using the "payback" approach. Int J Technol Assess Health Care. 2009;25(1):1-5.
56. Cruz Rivera S, Kyte DG, Aiyegbusi OL, Keeley TJ, Calvert MJ. Assessing the impact of healthcare research: A systematic review of methodological frameworks. PLoS Med. 2017;14(8):e1002370.
57. Oortwijn WJ, Hanney SR, Ligtvoet A, Hoorens S, Wooding S, Grant J, et al. Assessing the impact of health technology assessment in The Netherlands. Int J Technol Assess Health Care. 2008;24(3):259-69.
58. Stufflebeam D. Evaluation Models. New Directions for Evaluation. 2001;2001(89):7-98.
59. Hailey D, Alberta Heritage Foundation for Medical R. Elements of effectiveness for health technology assessment programs. 2003.
60. Hailey DM, Cowley DE, Dankiw W. THE IMPACT OF HEALTH TECHNOLOGY ASSESSMENT. Community Health Studies. 1990;14(3):223-34.

61. Drummond M. Evaluation of health technology: Economic issues for health policy and policy issues for economic appraisal. Social Science & Medicine. 1994;38(12):1593-600.
62. Lehoux P, Battista R, Lance J. Monitoring health technology assessment agencies2000; 15.
63. Gerhardus A, Dintsios C-M. The impact of HTA reports on health policy: a systematic review. GMS health technology assessment. 2005;1:Doc02-Doc.
64. Battista RN, Feeny DH, Hodge MJ. Evaluation of the Canadian Coordinating Office for Health Technology Assessment. Int J Technol Assess Health Care. 1995;11(1):102-16.
65. Jacob R, Battista RN. Assessing technology assessment. Early results of the Quebec experience. Int J Technol Assess Health Care. 1993;9(4):564-72.
66. 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.
67. Robertson R, Turner BS. Talcott Parsons and Modern Social Theory — An Appreciation. Theory, Culture & Society. 1989;6(4):539-58.
68. Guthrie S, Bienkowska-Gibbs T, Manville C, Pollitt A, Kirtley A, Wooding S. The impact of the National Institute for Health Research Health Technology Assessment programme, 2003-13: a multimethod evaluation. Health Technol Assess. 2015;19(67):1-291.
69. NIHR. https://www.nihr.ac.uk/funding-and-support/funding-for-research-studies/funding-programmes/health-technology-assessment/ [Available from: https://www.nihr.ac.uk/funding-and-support/funding-for-research-studies/funding-programmes/health-technology-assessment/.
70. Chalkidou K, Tunis S, Lopert R, Rochaix L, Sawicki PT, Nasser M, et al. Comparative effectiveness research and evidence-based health policy: experience from four countries. The Milbank quarterly. 2009;87(2):339-67.
71. Schumacher I, Zechmeister I. ASSESSING THE IMPACT OF HEALTH TECHNOLOGY ASSESSMENT ON THE AUSTRIAN HEALTHCARE SYSTEM. International Journal of Technology Assessment in Health Care. 2012;29(1):84-91.
72. de Sola-Morales O, Granados A. Health technology assessment in Catalonia: an overview of past and future perspectives. Int J Technol Assess Health Care. 2009;25 Suppl 1:88-93.
73. Sigmund H, Kristensen FB. Does health technology assessment benefit health services and politics? The experiences of an established HTA institution: the Danish Centre for Evaluation and HTA. The European journal of health economics : HEPAC : health economics in prevention and care. 2002;3(1):54-8.
74. Yazdizadeh B, Mohtasham F, Velayati A. Impact assessment of Iran's health technology assessment programme. Health Res Policy Syst. 2018;16(1):15.
75. Roza S, Junainah S, Izzuna MMG, Ku Nurhasni KAR, Yusof MAM, Noormah MD, et al. Health Technology Assessment in Malaysia: Past, Present, and Future. International Journal of Technology Assessment in Health Care. 2019;35(6):446-51.
76. Lipska I, McAuslane N, Leufkens H, Hövels A. A DECADE OF HEALTH TECHNOLOGY ASSESSMENT IN POLAND. Int J Technol Assess Health Care. 2017;33(3):350-7.
77. Kao K-L, Huang L-Y, Wu Y-H, Gau C-S. Outcomes and Impacts of 10-Year HTA Implementation in Taiwan. International Journal of Technology Assessment in Health Care. 2019;35(6):441-5.

78. Garrido M, Kirstenen F, Nielsen C, Busse R. Health Technology Assessment and Health Policy-Making in Europe. In: eunethta, editor. Observatory Study Series No 142008.
79. Weiss CH. The many meanings of research utlisation. Public Administration Review. 1979;39(5):426-31.
80. Gagnon MP, Sanchez E, Pons JM. Integration of health technology assessment recommendations into organizational and clinical practice: A case study in Catalonia. Int J Technol Assess Health Care. 2006;22(2):169-76.
81. Martin J, Polisena J, Dendukuri N, Rhainds M, Sampietro-Colom L. LOCAL HEALTH TECHNOLOGY ASSESSMENT IN CANADA: CURRENT STATE AND NEXT STEPS. International Journal of Technology Assessment in Health Care. 2016;32(3):175-80.
82. Sigmund H, Kristensen FB. Does health technology assessment benefit health services and politics? The European Journal of Health Economics. 2002;3(1):54-8.
83. Feig C, Cheung KL, Hiligsmann M, Evers S, Simon J, Mayer S. Best-worst scaling to assess the most important barriers and facilitators for the use of health technology assessment in Austria. Expert review of pharmacoeconomics & outcomes research. 2018;18(2):223-32.
84. Drummond M, Weatherly H. Implementing the findings of health technology assessments. If the CAT got out of the bag, can the TAIL wag the dog? Int J Technol Assess Health Care. 2000;16(1):1-12.
85. Oliver A, Mossialos E, Robinson R. Health technology assessment and its influence on health-care priority setting. Int J Technol Assess Health Care. 2004;20(1):1-10.
86. Fournier MF. Knowledge mobilization in the context of health technology assessment: an exploratory case study. Health Res Policy Syst. 2012;10:10.
87. Esmail R, Hanson H, Holroyd-Leduc J, Niven DJ, Clement F. Knowledge translation and health technology reassessment: identifying synergy. BMC Health Serv Res. 2018;18(1):674.
88. Straus SE, Tetroe J, Graham I. Defining knowledge translation. Canadian Medical Association Journal. 2009;181(3-4):165.
89. Cheung KL, Evers S, de Vries H, Hiligsmann M. Most Important Barriers and Facilitators regarding the Use of Health Technology Assessment. Int J Technol Assess Health Care. 2017;33(2):183-91.
90. Hauck K, Thomas R, Smith PC. Departures from Cost-Effectiveness Recommendations: The Impact of Health System Constraints on Priority Setting. Health Systems & Reform. 2016;2(1):61-70.
91. Holmes BJ, Best A, Davies H, Hunter D, Kelly MP, Marshall M, et al. Mobilising knowledge in complex health systems: a call to action. Evidence & Policy: A Journal of Research, Debate and Practice. 2017;13(3):539-60.
92. Owen-Smith A, Coast J, Donovan J. The usefulness of NICE guidance in practice: Different perspectives of managers, clinicians, and patients. International Journal of Technology Assessment in Health Care. 2010;26(3):317-22.
93. OECD. Health Project 2005. OECD iLibrary | The OECD Health Project (oecd-ilibrary.org); 2005.
94. Foy R, Walker A, Penney G. Barriers to clinical guidelines: the need for concerted action. British Journal of Clinical Governance. 2001;6(3):166-74.
95. OECD. Health Project 2005. 2005.
96. Hunter DJ, Marks L, Brown J, Scalabrini S, Salway S, Vale L, et al. The potential value of priority-setting methods in public health investment decisions:

qualitative findings from three English local authorities. Critical Public Health. 2016;26(5):578-87.
97. Hunter David J, Kieslich K, Littlejohns P, Staniszewska S, Tumilty E, Weale A, et al. Public involvement in health priority setting: future challenges for policy, research and society. Journal of Health Organization and Management. 2016;30(5):796-808.
98. Saul JE, Willis CD, Bitz J, Best A. A time-responsive tool for informing policy making: rapid realist review. Implementation Science. 2013;8(1):103.
99. Merlo G, Page K, Ratcliffe J, Halton K, Graves N. Bridging the gap: exploring the barriers to using economic evidence in healthcare decision making and strategies for improving uptake. Applied health economics and health policy. 2015;13(3):303-9.
100. Hunter DJ, Visram S. Better evidence for smarter policy making. BMJ. 2016;355:i6399.
101. Gallego G, van Gool K, Kelleher D. Resource allocation and health technology assessment in Australia: Views from the local level. International Journal of Technology Assessment in Health Care. 2009;25(2):134-40.
102. Organisation WH. Action plan to strengthen the use of evidence, information and research for policy-making in the WHO European Region. 2016.
103. Battista RN. Expanding the scientific basis of health technology assessment: a research agenda for the next decade. Int J Technol Assess Health Care. 2006;22(3):275-80; discussion 80-2.
104. Hanney S, Greenhalgh T, Blatch-Jones A, Glover M, Raftery J. The impact on healthcare, policy and practice from 36 multi-project research programmes: findings from two reviews. Health Research Policy and Systems. 2017;15(1):26.
105. Lavis JN, Robertson D, Woodside JM, McLeod CB, Abelson J, Knowledge Transfer Study G. How Can Research Organizations More Effectively Transfer Research Knowledge to Decision Makers? The Milbank Quarterly. 2003;81(2):221-48.
106. NICE. Evaluation and Review of NICE Implementation Guidance [Available from: http://webarchive.nationalarchives.gov.uk/20080611105936/http://www.nice.org.uk/usingguidance/evaluationandreviewofniceimplementationevidenceernie/evaluation_and_review_of_nice_implementation_evidence_ernie.jsp.
107. Raftery J, Hanney 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 Technol Assess. 2016;20(76):1-254.
108. Hanney SR, Grant J, Wooding S, Buxton MJ. Proposed methods for reviewing the outcomes of health research. the impact of funding by the UK's 'Arthritis Research Campaign'. 2004;2(1).
109. INAHTA. INAHTA Impact Assessment Study 2020. Canada; 2020.
110. Löblová O. What has health technology assessment ever done for us? Journal of Health Services Research & Policy. 2017;23(2):134-6.
111. Rumbold B, Baker R, Ferraz O, Hawkes S, Krubiner C, Littlejohns P, et al. Universal health coverage, priority setting, and the human right to health. The Lancet. 2017;390(10095):712-4.
112. Hailey D, Werko S, Rosen M, Macpherson K, Myles S, Gallegos Rivero V, et al. INFLUENCE OF HEALTH TECHNOLOGY ASSESSMENT AND ITS MEASUREMENT. Int J Technol Assess Health Care. 2016;32(6):376-84.
113. Grol R. Personal paper: Beliefs and evidence in changing clinical practice. BMJ. 1997;315(7105):418.

114. Gilson L, Raphaely N. The terrain of health policy analysis in low and middle income countries: a review of published literature 1994-2007. Health policy and planning. 2008;23(5):294-307.
115. Kapiriri L, Martin DK. Successful Priority Setting in Low and Middle Income Countries: A Framework for Evaluation. Health Care Analysis. 2010;18(2):129-47.
116. Smith P. The Politics of Priority Setting in Health: A Political Economy Perspective. Working paper 414. 2015.
117. Haan G, Rutten F. Economic appraisal, health service planning, and budgetary management of health technologies. In: Drummond M, editor. Economic Appraisal of Health Technology in the European Community. Oxford: Oxford University Press; 1987.
118. Drummond MF. Economic Appraisal of Health Technology in the European Community: Oxford Medical Publications; 1987.
119. Glied S, Smith P. The Oxford Handbook of Health Economics. Oxford: Oxford University Press; 2011.
120. Impact Evaluation [Available from: https://www.betterevaluation.org/en/themes/impact_evaluation.
121. Drummond M, Neumann P, Jonsson B, Luce B, Schwartz JS, Siebert U, et al. Can we reliably benchmark health technology assessment organizations? Int J Technol Assess Health Care. 2012;28(2):159-65.
122. Armstrong S, Garfield S. Assessing Assessment: Does Health Technology Assessment Do Its Job of Controlling Costs Without Compromising Quality? Value in Health. 2014;17(7):A420.
123. RAMESES. Theory in Realist Evaluation. RAMESES_II_Theory_in_realist_evaluation.pdf (ramesesproject.org); 2017.
124. Gates E, Dyson L. Implications of the Changing Conversation About Causality for Evaluators. American Journal of Evaluation. 2016;38(1):29-46.
125. Ontology [Available from: https://research-methodology.net/research-philosophy/ontology/.
126. Jagosh J. Evidencing Generative Causal Pathways: Future directions in Realist Inquiry Newcastle2018 [Available from: https://www.youtube.com/watch?time_continue=4&v=jjF-NH4mfo0.
127. Heckman JJ, Smith JA. Assessing the Case for Social Experiments. The Journal of Economic Perspectives. 1995;9(2):85-110.
128. Pawson R. Causality for Beginners. In: NCRM, editor. National Centre for Research Methods Festival 2008; NCRM: NCRM; 2008.
129. Van Belle S, Wong G, Westhorp G, Pearson M, Emmel N, Manzano A, et al. Can "realist" randomised controlled trials be genuinely realist? Trials. 2016;17(1):313-.
130. Mahoney J. Beyond Correlational Analysis: Recent Innovations in Theory and Method. Sociological Forum. 2001;16(3):575-93.
131. HM Treasury. Magenta Book. Treasury H, editor. London: UK Gov.; 2020.
132. Craig P. Synthetic controls: a new approach to evaluating interventions. What Works Scotland2015.
133. Cook T, Campbell D. Quasi-Experimentation Design & Analysis Issues for Field Settings. USA: Houghton Mifflin Company; 1979.
134. St.Clair T, Hallberg K, Cook TD. The Validity and Precision of the Comparative Interrupted Time-Series Design. Journal of Educational and Behavioral Statistics. 2016;41(3):269-99.
135. Sheldon TA, Cullum N, Dawson D, Lankshear A, Lowson K, Watt I, et al. What's the evidence that NICE guidance has been implemented? Results from a

national evaluation using time series analysis, audit of patients' notes, and interviews. BMJ. 2004;329(7473):999.
136. Abadie A, Diamond A, Hainmueller J. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. Journal of the American Statistical Association. 2010;105(490):493-505.
137. Lepine A, Largarde M, LeNestour A. Free primary care in Zambia: an impact evaluation using a pooled synthetic control method. York; 2015.
138. Bouttell J, Craig P, Lewsey J, Robinson M, Popham F. Synthetic control methodology as a tool for evaluating population-level health interventions. Journal of Epidemiology and Community Health. 2018;72(8):673.
139. Funnell S, Rogers P. Purposeful Program Theory: Effective Use of Theories of Change and Logic Models: Wiley; 2011.
140. Connell J. New Approaches to Evaluating Community Initiatives. Institute A, editor. New York1995.
141. Pawson R, Tilley N. Realistic Evaluation: Sage; 1997.
142. Kontos PC, Poland BD. Mapping new theoretical and methodological terrain for knowledge translation: contributions from critical realism and the arts. Implement Sci. 2009;4:1.
143. RAMESES. Philosophies and Evaluation Design. RAMESES_II_Philosophies_and_evaluation_design.pdf (ramesesproject.org); 2017.
144. Punton M, Vogel I, Leavy J, Michaelis C, Boydell E. Reality Bites: Making Realist Evaluation Useful in the Real World, CDI Practice Paper 22, Brighton: IDS. 2020.
145. Jagosh J. Email correspondence from member of the RAMESES JISC online research community. RAMESES List at WWWJISCMAILACUK. 11.09.2019 ed. https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=RAMESES2019.
146. Sayer A. Method in Social Science: a realist approach. London: Routledge; 1992.
147. Funnell S, Rogers P. Purposeful program theory: effective use of theories of change and logic models. San Francisco: Jossey-Bass/Wiley; 2011.
148. Jamal F, Fletcher A, Shackleton N, Elbourne D, Viner R, Bonell C. The three stages of building and testing mid-level theories in a realist RCT: a theoretical and methodological case-example. Trials. 2015;16:466.
149. Blackwood B, O'Halloran P, Porter S. On the problems of mixing RCTs with qualitative research: the case of the MRC framework for the evaluation of complex healthcare interventions. Journal of Research in Nursing. 2010;15(6):511-21.
150. Greenhalgh J. Complexity, Power and Evidence in the UK Healthcare Sector 2016 [Available from: https://www.youtube.com/watch?v=tvovm0BzmHI&feature=youtu.be.
151. Westhorp G. Email correspondence from member of the RAMESES JISC online research community https://wwwjiscmailacuk/cgi-bin/webadmin?A0=RAMESES 2018.
152. Lewis CC, Boyd MR, Walsh-Bailey C, Lyon AR, Beidas R, Mittman B, et al. A systematic review of empirical studies examining mechanisms of implementation in health. Implementation Science. 2020;15(1):21.
153. Rolfe S. Combining Theories of Change and Realist Evaluation in practice: Lessons from a research on evaluation study. Evaluation. 2019;25(3):294-316.
154. Pawson R. Evidence-based Policy: a realist perspective: SAGE Publishing; 2006.
155. Blamey A, Mackenzie M. Theories of Change and Realistic Evaluation: Peas in a Pod or Apples and Oranges? Evaluation. 2007;13(4):439-55.

156. Lawson T. Comparing Conceptions of Social Ontology: Emergent Social Entities and/or Institutional Facts? Journal for the Theory of Social Behaviour. 2016;46(4):359-99.
157. Grieve RH. Sheila Dow, Jesper Jerspersen, and Geoff Tily, eds., The General Theory and Keynes for the 21st Century (Cheltenham: Edward Elgar, 2018), pp. ix + 208, $130 (hardcover). ISBN: 9781786439871. Journal of the History of Economic Thought. 2019;41(3):469-74.
158. Porter S. Email correspondence from member of the RAMESES JISC online research community. . https://wwwjiscmailacuk/cgi-bin/webadmin?A0=RAMESES 2017.
159. Fenwick E, Claxton K, Sculpher M. The value of implementation and the value of information: combined and uneven development. Medical decision making : an international journal of the Society for Medical Decision Making. 2008;28(1):21-32.
160. Dixon S. Estimating the value of implementing reimbursement decisions: What can payers and manufacturers learn from economics? Healthcare Management Forum. 2019;32(6):303-6.
161. Walker S, Faria R, Whyte S, Dixon S, Palmer S, Sculpher M. Getting Cost-Effective Technologies into Practice: The Value of Implementation. Report on Framework for Valuing Implementation Initiatives. 2014.
162. Ochalek J, Lomas J, Claxton K. Estimating health opportunity costs in low-income and middle-income countries: a novel approach and evidence from cross-country data. BMJ Global Health. 2018;3(6):e000964.
163. Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S, et al. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. Health Technol Assess. 2015;19(14):1-503, v-vi.
164. Hoomans T, Ament AJ, Evers SM, Severens JL. Implementing guidelines into clinical practice: what is the value? J Eval Clin Pract. 2011;17(4):606-14.
165. Ochalek J, Revill P, Manthalu G, McGuire F, Nkhoma D, Rollinger A, et al. Supporting the development of a health benefits package in Malawi. BMJ Global Health. 2018;3(2):e000607.
166. SROINetwork, Aitken H. Starting out on Social Return on Investment. 2014.
167. Ismail N. Why is it difficult to evaluate the effectiveness and cost-effectiveness of complex public health interventions in the community? A health economics perspective. Perspectives in Public Health. 2017;137(4):206-7.
168. Briggs A, Sculpher M, Claxton K. Decision Modelling for Health Economic Evaluation: Oxford University Press; 2006 2006.
169. Ramsay CR, Matowe L, Grilli R, Grimshaw JM, Thomas RE. Interrupted time series designs in health technology assessment: lessons from two systematic reviews of behavior change strategies. Int J Technol Assess Health Care. 2003;19(4):613-23.
170. Kontopantelis E, Doran T, Springate DA, Buchan I, Reeves D. Regression based quasi-experimental approach when randomisation is not an option: interrupted time series analysis. BMJ : British Medical Journal. 2015;350.
171. Woods B, Revill P, Sculpher M, Claxton K. Country-Level Cost-Effectiveness Thresholds: Initial Estimates and the Need for Further Research. Value in Health. 2016;19(8):929-35.
172. Doherty JE, Wilkinson T, Edoka I, Hofman K. Strengthening expertise for health technology assessment and priority-setting in Africa. Global health action. 2017;10(1):1370194.

173. Asaria M, Ali S, Doran T, Ferguson B, Fleetcroft R, Goddard M, et al. How a universal health system reduces inequalities: lessons from England. Journal of Epidemiology and Community Health. 2016;70(7):637.
174. Verguet S, Kim JJ, Jamison DT. Extended Cost-Effectiveness Analysis for Health Policy Assessment: A Tutorial. PharmacoEconomics. 2016;34(9):913-23.
175. Asaria M, Griffin S, Cookson R. Distributional Cost-Effectiveness Analysis. Medical Decision Making. 2015;36(1):8-19.
176. Cookson R, Mirelman AJ, Griffin S, Asaria M, Dawkins B, Norheim OF, et al. Using Cost-Effectiveness Analysis to Address Health Equity Concerns. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2017;20(2):206-12.
177. Wong G, Greenhalgh T, Westhorp G, Buckingham J, Pawson R. RAMESES publication standards: realist syntheses. BMC Medicine. 2013;11(1):21.
178. Auraaen A, Fujisawa R, Lagasnerie Gd, Paris V. How OECD health systems define the range of good and services to be financed collectively. 2016.
179. itad. Report on NICE International's Engagement in China. https://www.idsihealth.org//wp-content/uploads/2016/05/NICE-ITAD-China-Report.pdf; 2015.
180. Löblová O, Trayanov T, Csanádi M, Ozierański P. The Emerging Social Science Literature on Health Technology Assessment: A Narrative Review. Value in Health. 2020;23(1):3-9.
181. Ridde V. Need for more and better implementation science in global health. BMJ Global Health. 2016;1(2).
182. Jagosh J. The 'Context + Mechanism' Association: Mastering a Key Heuristic in Realist Evaluation for Innovating Complex Programmes and Policy. . In: CECAN, editor. 2017.
183. Westhorp G. Email correspondence from member of the RAMESES JISC online research community. https://wwwjiscmailacuk/cgi-bin/webadmin?A0=RAMESES 2017.
184. Lacouture A, Breton E, Guichard A, Ridde V. The concept of mechanism from a realist approach: a scoping review to facilitate its operationalization in public health program evaluation. Implementation Science. 2015;10(1):153.
185. Westhorp G. Community Matters [Available from: http://www.communitymatters.com.au/gpage1.html.
186. Fogarty. https://www.fic.nih.gov/researchtopics/pages/implementationscience.aspx [
187. Nilsen P, Ståhl C, Roback K, Cairney P. Never the twain shall meet? - a comparison of implementation science and policy implementation research. Implementation Science. 2013;8(1):63.
188. Berman P. The study of macro- and micro-implementation. Public policy. 1978;26(2):157-84.
189. Yates B. Applying Diffusion Theory: Adoption of Media Literacy Programs in Schools. SIMILE Studies In Media & Information Literacy Education 2001;4(2).
190. Nilsen P. Making sense of implementation theories, models and frameworks. Implementation Science. 2015;10(1):53.
191. Birken SA, Powell BJ, Shea CM, Haines ER, Alexis Kirk M, Leeman J, et al. Criteria for selecting implementation science theories and frameworks: results from an international survey. Implementation Science. 2017;12(1):124.
192. Damschroder LJ, Aaron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci. 2009;4.

193. Logan JO, Graham ID. Toward a Comprehensive Interdisciplinary Model of Health Care Research Use. Science Communication. 1998;20(2):227-46.
194. Pfadenhauer LM, Gerhardus A, Mozygemba K, Lysdahl KB, Booth A, Hofmann B, et al. Making sense of complexity in context and implementation: the Context and Implementation of Complex Interventions (CICI) framework. Implement Sci. 2017;12(1):21.
195. Lehman WE, Greener JM, Simpson DD. Assessing organizational readiness for change. J Subst Abuse Treat. 2002;22(4):197-209.
196. Weiner BJ. A theory of organizational readiness for change. Implement Sci. 2009;4.
197. Birken SA, Bunger AC, Powell BJ, Turner K, Clary AS, Klaman SL, et al. Organizational theory for dissemination and implementation research. Implementation Science. 2017;12(1):62.
198. Kitson A, Harvey G, McCormack B. Enabling the implementation of evidence based practice: a conceptual framework. 1998;7:149-58.
199. Rycroft-Malone J, Harvey G, Seers K, Kitson A, McCormack B, Titchen A. An exploration of the factors that influence the implementation of evidence into practice. Journal of Clinical Nursing. 2004;13.
200. Harvey G, Kitson A. PARIHS revisited: from heuristic to integrated framework for the successful implementation of knowledge into practice. Implementation Science. 2016;11(1):33.
201. Harvey G, Kitson A. PARIHS re-visited: introducing i-PARIHS. In: Harvey G, Kitson A, editors. Implementing evidence-based practice in health care: a facilitation guide. Abingdon, Oxon: Routledge; 2015.
202. Sudsawad P. Knowledge translation: Introduction to models, strategies and measures. Austin TX: Southwest Educational Development Laboratory, National Center for the Dissemination of Disability Research. 2007.
203. Ward V, House A, Hamer S. Developing a framework for transferring knowledge into action: a thematic analysis of the literature. Journal of health services research & policy. 2009;14(3):156-64.
204. Kitson A, Rycroft-Malone J, Harvey G, McCormack B, Seer K, Titchen A. Evaluating the successful implementation of evidence into practice using the PARIHS framework: theoretical and practical challenges. Implement Sci. 2008;3.
205. Rycroft-Malone J, Kitson A, Harvey G, McCormack B, Seers K, Titchen A, et al. Ingredients for change: revisiting a conceptual framework. Quality & Safety in Health Care. 2002;11.
206. Kitson A, Harvey G, McCormack B. Enabling the implementation of evidence based practice: a conceptual framework. Qual Health Care. 1998;7.
207. Kitson AL, Rycroft-Malone J, Harvey G, McCormack B, Seers K, Titchen A. Evaluating the successful implementation of evidence into practice using the PARiHS framework: theoretical and practical challenges. Implementation Science. 2008;3(1):1.
208. Stetler CB, Damschroder LJ, Helfrich CD, Hagedorn HJ. A Guide for applying a revised version of the PARIHS framework for implementation. Implementation Science. 2011;6(1):99.
209. Commitment to organizational change: Extension of a three-component model, (2002).
210. Jagosh J. Email correspondence from member of the RAMESES JISC online research community. RAMESES List at WWWJISCMAILACUK. 5.7.2018 ed. https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=RAMESES2018.
211. Cooper MJF, Sornalingam S, O'Donnell C. Street-level bureaucracy: an underused theoretical model for general practice? The British journal of general

practice : the journal of the Royal College of General Practitioners. 2015;65(636):376-7.
212. Yothasamut J, Putchong C, Sirisamutr T, Teerawattananon Y, Tantivess S. Scaling up cervical cancer screening in the midst of human papillomavirus vaccination advocacy in Thailand. BMC Health Services Research. 2010;10(Suppl 1):S5-S.
213. Jagosh J. CARES Training workshop. 2016.
214. Mooney JS, Yates A, Sellar L, Shipway T, Roberts C, Parris R, et al. Emergency head injury imaging: implementing NICE 2007 in a tertiary neurosciences centre and a busy district general hospital. Emergency medicine journal : EMJ. 2011;28(9):778-82.
215. Thavarajah D, Wetherill M. Implementing NICE guidelines on risk assessment for venous thromboembolism: Failure, success and controversy. International Journal of Health Care Quality Assurance. 2012;25(7):618-24.
216. Bennie M, Godman B, Bishop I, Campbell S. Multiple initiatives continue to enhance the prescribing efficiency for the proton pump inhibitors and statins in Scotland. Expert review of pharmacoeconomics & outcomes research. 2012;12(1):125-30.
217. Goeree R, Levin L. Building bridges between academic research and policy formulation: the PRUFE framework - an integral part of Ontario's evidence-based HTPA process. Pharmacoeconomics. 2006;24(11):1143-56.
218. Schluessmann E, Diel P, Aghayev E, Zweig T, Moulin P, Roder C. SWISSspine: a nationwide registry for health technology assessment of lumbar disc prostheses. European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society. 2009;18(6):851-61.
219. Jonsson E, Banta HD, Schersten T. Health technology assessment and screening in Sweden. Int J Technol Assess Health Care. 2001;17(3):380-8.
220. Zechmeister I, Schumacher I. The impact of health technology assessment reports on decision making in Austria. Int J Technol Assess Health Care. 2012;28(1):77-84.
221. Bennie M, Dear J, Hems S, Black C, McIver L, Webb DJ. An investigation into the effect of advice from the Scottish Medicines Consortium on the use of medicines in Scotland's Health Service. British journal of clinical pharmacology. 2011;71(2):283-8.
222. Dietrich ES. Effects of the National Institute for Health and Clinical Excellence's technology appraisals on prescribing and net ingredient costs of drugs in the National Health Service in England. International Journal of Technology Assessment in Health Care. 2009;25(3):262-71.
223. Rosen M, Werko S. Does health technology assessment affect policy-making and clinical practice in Sweden? Int J Technol Assess Health Care. 2014;30(3):265-72.
224. Platt C, Larcombe J, Dudley J, McNulty C, Banerjee J, Gyoffry G, et al. Implementation of NICE guidance on urinary tract infections in children in primary and secondary care. Acta paediatrica (Oslo, Norway : 1992). 2015;104(6):630-7.
225. NICE. NICE Impact CVD prevention. https://www.nice.org.uk/about/what-we-do/into-practice/measuring-the-uptake-of-nice-guidance/impact-of-guidance; 2018.
226. NICE. NICE Impact Cancer. https://www.nice.org.uk/about/what-we-do/into-practice/measuring-the-uptake-of-nice-guidance/impact-of-guidance; 2018.

227. NICE. NICE Impact Maternity. https://www.nice.org.uk/about/what-we-do/into-practice/measuring-the-uptake-of-nice-guidance/impact-of-guidance; 2018.
228. Britton M, Jonsson E. Impact of health technology assessments. Some experiences of SBU. Int J Technol Assess Health Care. 2002;18(4):824-31.
229. Axelsson S, Helgason AR, Lund KE, Adolfsson J. Disseminating evidence from health technology assessment: the case of tobacco prevention. Int J Technol Assess Health Care. 2006;22(4):500-5.
230. Solans-Domenech M, Adam P, Guillamon I, Permanyer-Miralda G, Pons JM, Escarrabill J. Impact of clinical and health services research projects on decision-making: a qualitative study. Health Res Policy Syst. 2013;11:15.
231. Hanney S, Buxton M, Green C, Coulson D, Raftery J. An assessment of the impact of the NHS Health Technology Assessment Programme. Health Technol Assess. 2007;11(53):iii-iv, ix-xi, 1-180.
232. Schuller T, Soderholm Werko S. INSIGHTS FROM THE FRONT LINES: A COLLECTION OF STORIES OF HTA IMPACT FROM INAHTA MEMBER AGENCIES. Int J Technol Assess Health Care. 2017;33(4):409-10.
233. Fronsdal KB, Facey K, Klemp M, Norderhaug IN, Morland B, Rottingen JA. Health technology assessment to optimize health technology utilization: using implementation initiatives and monitoring processes. Int J Technol Assess Health Care. 2010;26(3):309-16.
234. Lowson K, Jenks M, Filby A, Carr L, Campbell B, Powell J. Examining the implementation of NICE guidance: cross-sectional survey of the use of NICE interventional procedures guidance by NHS Trusts. Implementation science : IS. 2015;10:93-.
235. Stemerding D, van Berkel D. Maternal serum screening, political decision-making and social learning. Health Policy. 2001;56(2):111-25.
236. Teerawattananon Y, Tritasavit N, Suchonwanich N, Kingkaew P. The use of economic evaluation for guiding the pharmaceutical reimbursement list in Thailand. Z Evid Fortbild Qual Gesundhwes. 2014;108(7):397-404.
237. Corbacho B, Drummond M, Santos R, Jones E, Borràs JM, Mestre-Ferrandiz J, et al. Does the use of health technology assessment have an impact on the utilisation of health care resources? Evidence from two European countries. The European Journal of Health Economics. 2020;21(4):621-34.
238. Cummings GG, Estabrooks CA, Midodzi WK, Wallin L, Hayduk L. Influence of organizational characteristics and context on research utilization. Nurs Res. 2007;56(4 Suppl):S24-39.
239. Dent THS, Sadler M. From guidance to practice: Why NICE is not enough. BMJ (Clinical research ed). 2002;324(7341):842-5.
240. Poder TG, Bellemare CA, Bédard SK, Fisette J-F, Dagenais P. IMPACT OF HEALTH TECHNOLOGY ASSESSMENT REPORTS ON HOSPITAL DECISION MAKERS – 10-YEAR INSIGHT FROM A HOSPITAL UNIT IN SHERBROOKE, CANADA: IMPACT OF HEALTH TECHNOLOGY ASSESSMENT ON HOSPITAL DECISIONS. International Journal of Technology Assessment in Health Care. 2018;34(4):393-9.
241. Sihvo S, Ikonen T, Mäkelä M. IMPLEMENTING HEALTH TECHNOLOGY ASSESSMENT-BASED RECOMMENDATIONS IN FINLAND: MANAGED UPTAKE OF MEDICAL METHODS. International Journal of Technology Assessment in Health Care. 2017;33(4):430-3.
242. Haslam D, Rao S, Leng G, Steele J, Harris S, McKeon A, et al. Implementing NICE guidance. Roundtable discussion. Talking cure. Health Serv J. 2005;115(5984):suppl 4-7.

243. Chidgey J, Leng G, Lacey T. Implementing NICE guidance. Journal of the Royal Society of Medicine. 2007;100(10):448-52.
244. Gallego G, Casey R, Norman R, Goodall S. Introduction and uptake of new medical technologies in the Australian health care system: A qualitative study. Health Policy. 2011;102(2):152-8.
245. Hastings J, Adams EJ. Joint project of the international network of agencies for health technology assessment—Part 1: Survey results on diffusion, assessment, and clinical use of positron emission tomography. International Journal of Technology Assessment in Health Care. 2006;22(2):143-8.
246. Rosen R, Gabbay J. Linking health technology assessment to practice. Interview by Abi Berger. Bmj. 1999;319(7220):1292.
247. Barosi G. Strategies for dissemination and implementation of guidelines. Neurological Sciences. 2006;27(3):s231-s4.
248. Rosenkötter N, Vondeling H, Blancquaert I, Mekel OCL, Kristensen FB, Brand A. The Contribution of Health Technology Assessment, Health Needs Assessment, and Health Impact Assessment to the Assessment and Translation of Technologies in the Field of Public Health Genomics. Public Health Genomics. 2011;14(1):43-52.
249. Cacciatore P, Specchia ML, Solinas MG, Ricciardi W, Damiani G. The organizational domain in HTA reports: towards a technology-oriented assessment. Eur J Public Health. 2020;30(2):219-23.
250. Hivon M, Lehoux P, Denis J-L, Tailliez S. Use of health technology assessment in decision making: Coresponsibility of users and producers? International Journal of Technology Assessment in Health Care. 2005;21(2):268-75.
251. Lehoux P, Hivon M, Denis J-L, Tailliez S. What medical specialists like and dislike about health technology assessment reports. Journal of Health Services Research & Policy. 2009;14(4):197-203.
252. Tucker S, O'Brien K, Brown HM. Using health technology assessment for better healthcare decisions. Healthcare Management Forum. 2019;32(4):197-201.
253. Pettersson B, Hoffmann M, Wändell P, Levin L-Å. Utilization and costs of lipid modifying therapies following health technology assessment for the new reimbursement scheme in Sweden. Health Policy. 2012;104(1):84-91.
254. Campillo-Artero C. When health technologies do not reach their effectiveness potential: A health service research perspective. Health Policy. 2012;104(1):92-8.
255. Doherty S. Evidence – based implementation of evidence – based guidelines. International Journal of Health Care Quality Assurance. 2006;19(1):32-41.
256. Mayer J, Kipps C, Cock HR. Implementing clinical guidelines. Pract Neurol. 2019;19(6):529-35.
257. Tian L, Yang Y, Sui W, Hu Y, Li H, Wang F, et al. Implementation of evidence into practice for cancer-related fatigue management of hospitalized adult patients using the PARIHS framework. PLoS One. 2017;12(10):e0187257.
258. Saunders H, Stevens KR, Vehviläinen-Julkunen K. Nurses' readiness for evidence-based practice at Finnish university hospitals: a national survey. Journal of Advanced Nursing. 2016;72(8):1863-74.
259. Taxman FS, Henderson C, Young D, Farrell J. The Impact of Training Interventions on Organizational Readiness to Support Innovations in Juvenile Justice Offices. Administration and Policy in Mental Health and Mental Health Services Research. 2014;41(2):177-88.

260. Stander J, Grimmer K, Brink Y. Training programmes to improve evidence uptake and utilisation by physiotherapists: a systematic scoping review. BMC Med Educ. 2018;18(1):14.
261. Grenon X, Pinget C, Wasserfallen JB. HOSPITAL-BASED HEALTH TECHNOLOGY ASSESSMENT (HB-HTA): A 10-YEAR SURVEY AT ONE UNIT. Int J Technol Assess Health Care. 2016;32(3):116-21.
262. Billsten J, Fridell M, Holmberg R, Ivarsson A. Organizational Readiness for Change (ORC) test used in the implementation of assessment instruments and treatment methods in a Swedish National study. Journal of Substance Abuse Treatment. 2018;84:9-16.
263. Uneke CJ, Sombie I, Keita N, Lokossou V, Johnson E, Ongolo-Zogo P, et al. Assessment of policy makers' individual and organizational capacity to acquire, assess, adapt and apply research evidence for maternal and child health policy making in Nigeria: a cross-sectional quantitative survey. Afr Health Sci. 2017;17(3):700-11.
264. Sasaki N, Yamaguchi N, Okumura A, Yoshida M, Sugawara H, Imanaka Y. Does hospital information technology infrastructure promote the implementation of clinical practice guidelines? A multicentre observational study of Japanese hospitals. BMJ Open. 2019;9(6):e024700.
265. Jolliffe L, Morarty J, Hoffmann T, Crotty M, Hunter P, Cameron ID, et al. Using audit and feedback to increase clinician adherence to clinical practice guidelines in brain injury rehabilitation: A before and after study. PLoS One. 2019;14(3):e0213525.
266. Grove A, Clarke A, Currie G. How are evidence and knowledge used in orthopaedic decision-making? Three comparative case studies of different approaches to implementation of clinical guidance in practice. Implement Sci. 2018;13(1):75.
267. Wathen B, Dean T. An evaluation of the impact of NICE guidance on GP prescribing. The British journal of general practice : the journal of the Royal College of General Practitioners. 2004;54(499):103-7.
268. Weiner B. A theory of organizational readiness for change. Implement Sci. 2009;4.
269. Turner S, Bhurke S, Cook A. Impact of NIHR HTA Programme funded research on NICE clinical guidelines: a retrospective cohort. Health Research Policy and Systems. 2015;13(1):37.
270. Drummond MF, Schwartz JS, Jönsson B, Luce BR, Neumann PJ, Siebert U, et al. Key principles for the improved conduct of health technology assessments for resource allocation decisions. Int J Technol Assess Health Care. 2008;24(3):244-58; discussion 362-8.
271. Currie G, Kiefer T, Spyridonidis D. From what we know to what we do: enhancing absorptive capacity in translational health research. BMJ Leader. 2020;4(1):18.
272. Carayon P. Human factors of complex sociotechnical systems. Applied Ergonomics. 2006;37(4):525-35.
273. Baxter G, Sommerville I. Socio-technical systems: From design methods to systems engineering. Interacting with Computers. 2011;23(1):4-17.
274. Holt DT, Helfrich CD, Hall CG, Weiner BJ. Are you ready? How health professionals can comprehensively conceptualize readiness for change. J Gen Intern Med. 2010;25 Suppl 1:50-5.
275. Puchalski Ritchie LM, Khan S, Moore JE, Timmings C, van Lettow M, Vogel JP, et al. Low- and middle-income countries face many common barriers

to implementation of maternal health evidence products. J Clin Epidemiol. 2016;76:229-37.
276. Pawson R. Email correspondence from member of the RAMESES JISC online research community. https://wwwjiscmailacuk/cgi-bin/webadmin?A0=RAMESES 23.2.16. 2016.
277. Rotter T, Kinsman L, James E, Machotta A, Gothe H, Willis J, et al. Clinical pathways: effects on professional practice, patient outcomes, length of stay and hospital costs. The Cochrane database of systematic reviews. 2010(3):Cd006632.
278. Kwan J, Sandercock P. In-hospital care pathways for stroke. The Cochrane database of systematic reviews. 2002(2):Cd002924.
279. Vanhaecht K, De Witte K, Depreitere R, Sermeus W. Clinical pathway audit tools: a systematic review. J Nurs Manag. 2006;14(7):529-37.
280. Crummer MB, Carter V. Critical pathways--the pivotal tool. The Journal of cardiovascular nursing. 1993;7(4):30-7.
281. Weingarten S. Critical pathways: what do you do when they do not seem to work? The American journal of medicine. 2001;110(3):224-5.
282. CNHDRC. Strengthening evidence-based policy making in support of universal healthcare: Introducing evidence-based clinical pathways for stroke and COPD in rural China. https://www.idsihealth.org//wp-content/uploads/2016/05/Introducing-evidence-based-clinical-pathways-for-stroke-and-COPD-rural-China.pdf; 2015.
283. Wu X, Zhu B, Fu L, Wang H, Zhou B, Zou S, et al. Prevalence, incidence, and mortality of stroke in the chinese island populations: a systematic review. PLoS One. 2013;8(11):e78629.
284. Chen Z. The mortality and death cause of national sample areas. In: Chen Z, editor. The Third National Survey on the Cause of Death. Beijing: Peking Union Medical University Press; 2008.
285. Guo W. Economic Evaluation of Clinical Pathways, China. 2017.
286. YHEC. Markov Model York: University of York; 2016 [Available from: http://www.yhec.co.uk/glossary/markov-model/.
287. Lagarde M. How to do (or not to do) … Assessing the impact of a policy change with routine longitudinal data. Health policy and planning. 2012;27(1):76-83.
288. Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. Journal of Clinical Pharmacy and Therapeutics. 2002;27(4):299-309.
289. NBS. China Statistical Year Book http://www.stats.gov.cn/tjsj/ndsj/2016/indexeh.htm2016 [Available from: http://www.stats.gov.cn/tjsj/ndsj/2016/indexeh.htm.
290. MacQuilkan K, Baker P, Downey L, Ruiz F, Chalkidou K, Prinja S, et al. Strengthening health technology assessment systems in the global south: a comparative analysis of the HTA journeys of China, India and South Africa. Global health action. 2018;11(1):1527556.
291. Liu W, Shi L, Pong RW, Dong H, Mao Y, Tang M, et al. Determinants of knowledge translation from health technology assessment to policy-making in China: From the perspective of researchers. PLoS One. 2018;13(1):e0190732.
292. Liu W, Shi L, Pong RW, Dong H, Mao Y, Tang M, et al. Determinants of knowledge translation from health technology assessment to policy-making in China: From the perspective of researchers. PloS one. 2018;13(1):e0190732-e.
293. Chen Y, He Y, Chi X, Wei Y, Shi L. Development of health technology assessment in China: New challenges. BioScience Trends. 2018;12(2):102-8.

294. Gandjour A, Lauterbach KW. A METHOD FOR ASSESSING THE COST-EFFECTIVENESS AND THE BREAK-EVEN POINT OF CLINICAL PRACTICE GUIDELINES. International Journal of Technology Assessment in Health Care. 2001;17(4):503-16.
295. HTAi. HTAI conference. Vancouver2018.
296. Bertram MY, Lauer JA, De Joncheere K, Edejer T, Hutubessy R, Kieny M-P, et al. Cost-effectiveness thresholds: pros and cons. Bulletin of the World Health Organization. 2016;94(12):925-30.
297. SHTG. Freestyle Libre flash glucose monitoring Scotland: Healthcare Improvement Scotland; 2018 [Available from: http://www.healthcareimprovementscotland.org/our_work/technologies_and_medicines/shtg_advice_statements/advice_statement_009-18.aspx.
298. DiabetesTimes. NHS Drug Tariff approval for Freestyle Libre System 2019 [Available from: https://diabetestimes.co.uk/nhs-drug-tariff-approval-freestyle-libre-system/.
299. SHTG. Evidence Note 81 Scotland: HIS; 2018 [Available from: http://www.healthcareimprovementscotland.org/our_work/technologies_and_medicines/shtg_advice_statements/advice_statement_009-18.aspx.
300. HIS NHS. SHTG recommends FL monitoring device 2018 [Available from: http://www.healthcareimprovementscotland.org/news_and_events/news/news_freestyle_libre.aspx.
301. NHSScotland. Prescribing and Medicines [Available from: https://www.isdscotland.org/Health-topics/Prescribing-and-medicines/) .
302. NHSScotland. Prescriptions in the Community 2020 [Available from: https://www.opendata.nhs.scot/dataset/prescriptions-in-the-community.
303. Petrou S, Gray A. Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting. BMJ. 2011;342:d1548.
304. Levrat-Guillen F. Role of HTA in the Adoption of New Technologies in Diabetes Care: A Review of HTA Reports on Freestyle Libre. Pharmacoeconomics. 2019;4(1).
305. Ramsay CR, Matowe L, Grilli R, Grimshaw JM, Thomas RE. INTERRUPTED TIME SERIES DESIGNS IN HEALTH TECHNOLOGY ASSESSMENT: LESSONS FROM TWO SYSTEMATIC REVIEWS OF BEHAVIOR CHANGE STRATEGIES. International Journal of Technology Assessment in Health Care. 2003;19(4):613-23.
306. Fulop N, Allen P, Clarke A, Black N. From health technology assessment to research on the organisation and delivery of health services: addressing the balance. Health Policy. 2003;63(2):155-65.
307. Booth A, Briscoe S, Wright JM. The "realist search": A systematic scoping review of current practice and reporting. Research Synthesis Methods. 2020;11(1):14-35.
308. Pawson R. Email correspondence from member of the RAMESES JISC online research community. https://wwwjiscmailacuk/cgi-bin/webadmin?A0=RAMESES 16.5.16. 2016.
309. Farr M, Cressey P. The social impact of advice during disability welfare reform: from social return on investment to evidencing public value through realism and complexity. Public Management Review. 2019;21(2):238-63.
310. Maier F, Schober C, Simsa R, Millner R. SROI as a Method for Evaluation Research: Understanding Merits and Limitations. VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations. 2015;26(5):1805-30.
311. Banke-Thomas AO, Madaj B, Charles A, van den Broek N. Social Return on Investment (SROI) methodology to account for value for money of public health interventions: a systematic review. BMC Public Health. 2015;15(1):582.

312. Bennie M, Dear J, Hems S, Black C, McIver L, Webb DJ. An investigation into the effect of advice from the Scottish Medicines Consortium on the use of medicines in Scotland's Health Service. Br J Clin Pharmacol. 2011;71(2):283-8.
313. iPAG. Abbott Freestyle Libre Prescribing Around Scotland Scotland: Insulin Pump Awareness Group; 2018 [Available from: https://www.ipag.co.uk/abbott-freestyle-libre-prescribing-around-scotland/.
314. Whyte S, Dixon S, Faria R, Walker S, Palmer S, Sculpher M, et al. Estimating the Cost-Effectiveness of Implementation: Is Sufficient Evidence Available? Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2016;19(2):138-44.
315. Email correspondence from member of the RAMESES JISC online research community 24.9.18 [Internet]. RAMESES. 2018. Available from: https://wwwjiscmailacuk/cgi-bin/webadmin?A0=RAMESES
316. Rajan A, Gutierrez-Ibarluzea I, Moharra M. Addressing issues in health technology assessment promotion: Motives, enablers, and barriers. International Journal of Technology Assessment in Health Care. 2011;27(1):55-63.
317. Deaton AS. Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development. National Bureau of Economic Research Working Paper Series. 2009;No. 14690.
318. Meacock R. Methods for the economic evaluation of changes to the organisation and delivery of health services: principal challenges and recommendations. Health Economics, Policy and Law. 2019;14(1):119-34.
319. Anderson R, Hardwick R. Realism and resources: Towards more explanatory economic evaluation. Evaluation (London, England : 1995). 2016;22(3):323-41.
320. Doing Realist Research. 2018 2020/07/09. 55 City Road, London: SAGE Publications Ltd. Available from: https://methods.sagepub.com/book/doing-realist-research.
321. Vassall A, Mangham-Jefferies L, Gomez GB, Pitt C, Foster N. Incorporating Demand and Supply Constraints into Economic Evaluations in Low-Income and Middle-Income Countries. Health Economics. 2016;25(S1):95-115.
322. Rolden HJA, van der Wilt GJ, Maas A, Grutters JPC. THE GAP BETWEEN ECONOMIC EVALUATIONS AND CLINICAL PRACTICE: A SYSTEMATIC REVIEW OF ECONOMIC EVALUATIONS ON DABIGATRAN FOR ATRIAL FIBRILLATION. Int J Technol Assess Health Care. 2018;34(3):327-36.
323. Gilmore B. Realist evaluations in low- and middle-income countries: reflections and recommendations from the experiences of a foreign researcher. BMJ Global Health. 2019;4(5):e001638.
324. CGD. [cited 2020]. Available from: https://www.cgdev.org/blog/priority-setting-better-health-international-decision-support-initiative?utm_source=200428&utm_medium=cgd_email&utm_campaign=cgd_weekly.
325. Surgey G, Chalkidou K, Reuben W, Suleman F, Miot J, Hofman K. Introducing health technology assessment in Tanzania. International Journal of Technology Assessment in Health Care. 2020;36(2):80-6.
326. The impact of impact evaluation. Helsinki, Finland: UNU-WIDER.
327. OECD. Using routinely collected data to inform pharmaceutical policies. 2019.
328. Heggie R, Wu O, Boyd K. *Why health economists should care about implementation: a rapid review*. 2020.

329. Mason J, Freemantle N, Nazareth I, Eccles M, Haines A, Drummond M. When is it cost-effective to change the behavior of health professionals? Jama. 2001;286(23):2988-92.
330. van Baal P, Morton A, Severens JL. Health care input constraints and cost effectiveness analysis decision rules. Soc Sci Med. 2018;200:59-64.
331. Revill P, Walker S, Cambiano V, Phillips A, Sculpher MJ. Reflecting the real value of health care resources in modelling and cost-effectiveness studies-The example of viral load informed differentiated care. PLoS One. 2018;13(1):e0190283.
332. Hauck K, Morton A, Chalkidou K, Chi YL, Culyer A, Levin C, et al. How can we evaluate the cost-effectiveness of health system strengthening? A typology and illustrations. Social Science & Medicine. 2019;220:141-9.
333. Love-Koh J, Walker SM, Kataika E, Sibandze S, Arnold M, Ochalek JM, et al. Economic Analysis for Health Benefits Package Design. [Discussion paper]. In press 2019.
334. Morton A, Thomas R, Smith PC. Decision rules for allocation of finances to health systems strengthening. Journal of Health Economics. 2016;49:97-108.
335. Bouttell J, Grieve E, Hawkins N. The role of development-focused health technology assessment in optimising science, technology and innovation to achieve SDG3 Science, Technology and Innovation for Meeting Sustainable Development Goals: OUP; in press.
336. van Harten WH, Retèl VP. Innovations that reach the patient: early health technology assessment and improving the chances of coverage and implementation. Ecancermedicalscience. 2016;10:683-.
337. Lee A, Skött LS, Hansen HP. Organizational and patient-related assessments in HTAs: state of the art. Int J Technol Assess Health Care. 2009;25(4):530-6.
338. Kowalewski K, Lavis JN, Wilson M, Carter N. Supporting evidence-informed health policy making: the development and contents of an online repository of policy-relevant documents addressing healthcare renewal in Canada. Healthcare policy = Politiques de sante. 2014;10(2):27-37.
339. Velasco Garrido M, Gerhardus A, Rottingen JA, Busse R. Developing Health Technology Assessment to address health care system needs. Health policy (Amsterdam, Netherlands). 2010;94(3):196-202.
340. Ijzerman MJ, Steuten LM. Early assessment of medical technologies to inform product development and market access: a review of methods and applications. Appl Health Econ Health Policy. 2011;9(5):331-47.
341. Towse A. OHE. 2020. [cited 2020]. Available from: *https://www.ohe.org/news/unpacking-black-box-payer-policy-demand-side-approach-equitable-uptake-cost-effective-health-0*.
342. Chalkidou K, Towse A, Silverman R, Garau M, Ramakrishnan G. Market-driven, value-based, advance commitment (MVAC): accelerating the development of a pathbreaking universal drug regimen to end TB. BMJ Global Health. 2020;5(4):e002061.
343. Painter C, Teerawattananon Y, Chalkidou K, Clapham H. CGD, editor2020. Available from: https://www.cgdev.org/blog/early-health-technology-assessment-covid-19-vaccine.
344. Towse A, Devlin N, Hawe E, Garrison L. The Evolution of HTA in Emerging Markets Health Care Systems: Analysis to a Policy Response. 2011.
345. Cylus J, Papanicolas I, Smith P. Health System Efficiency: How to make measurement matter for policy and management. In: WHO, editor. London2016.