The Value of Health Technology Assessment: a mixed methods framework

Measuring impact

1 *Health Economics and Health Technology Assessment, Institute of Health and Wellbeing, 1 Lilybank Gardens, University of Glasgow, Glasgow G12 8RZ, UK*

2 *International Decision Support Initiative, Imperial College London, South Wharf Road London, W2 1NY*

Technical report

Eleanor Grieve[1*], Hannah Hesselgreaves[1], Olivia Wu[1], Kalipso Chalkidou[2], Francis Ruiz[2], Peter Smith[2], Ryan Li[2], Laura Morris[2], Andrew Briggs[1]

* *corresponding author Eleanor.grieve@glasgow.ac.uk*

Funded by Bill & Melinda Gates Foundation and the UK Department for International Development

iDSI
Better decisions. Better health.

Table of Contents

Introduction

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 a scientific and fair way (1). Access to good quality, affordable package of services and technologies according to need is a priority. Health technology assessment is one of the tools for priority setting which *"refers to the systematic evaluation of properties, effects, and/or impacts of health technology. It is a multidisciplinary process to evaluate the social, economic, organizational and ethical issues of a health intervention or health technology. The main purpose of conducting an assessment is to inform policy decision making… [and is applied to] medicines, medical devices, vaccines, procedures and systems developed to solve a health problem and improve quality of life (2)."* Whilst developed countries may have led the way, low- and middle-income countries (LMICs) are increasingly beginning to develop HTA processes to assist in their healthcare decision-making. Indeed, the institutionalisation of health technology assessment is moving forward in many countries with a growing global investment in these processes as it has emerged as an important tool for supporting the implementation of Universal Health Coverage (UHC) (3). This has been bolstered by commitments at a global level by the issuance of the HTA resolutions amongst Inter-American countries in 2012 (4), Southeast Asian countries in 2013(5), and in the World Health Assembly in 2014 (6) in recognising the role of HTA and robust priority-setting processes in making fair resource allocation decisions and sustaining health systems' performance (6). Since LMICs face particularly limited resources and, as cost-effectiveness is an integral part of healthcare decision-making, the value of HTA to help make better resource allocation decisions is being recognised.

The benefits concerning the link between health technology assessment and outcomes in terms of health improvements have, however, rarely been quantified. We distinguish here between 'HTA' as a process at the systems level to inform priority-setting and decision-making, and 'hta' as an individual health technology assessment including evidence-based interventions, practices or policies at a practical level. The focus of our research is on the impact of HTA as a tool for priority-setting with its explicit consideration of costs and benefits. Whilst much research has been undertaken on establishing what factors influence improved decision-making including good governance structures, expertise, political and institutional factors, resources and participation (7-10), how such influences on decision-making interact with local context and health systems, leading to impact on health outcomes, is less understood. Where evaluations have been undertaken, they mainly focus on processes or outcomes at the decision-making level, with impact on health outcomes rarely measured. Straus et al (11), as cited in (12), 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/EBM infrastructure and evidence base, the next 10 should focus on the outcomes."* Also, *"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."*(13)

Even in countries where HTA programmes are well established, evidence which identifies their outcomes and impact in terms of health gains is limited. For countries with greater capacity constraints, how decision-making interacts with 'context' leading to health outcomes is even less explored and arguably of critical importance(14): *"... 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."* (15)

It is evident that significant amounts of scare resources are invested in HTA. It has been estimated that NICE depends on about 2000 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 (16). In a recent effort to move beyond impact on decision-making, RAND Europe examined the impact of the National Institute for Health Research (NIHR) HTA Programme to understand its potential economic benefits (17). The NIHR HTA Programme funds research about the clinical and cost effectiveness and broader impact of healthcare treatments and tests for those who plan, provide or receive care in the UK's National Health Service (NHS) (18). This evaluation focused on savings for the NHS and health benefits to patients converted into financial terms. To carry out this evaluation, some assumptions were required, in particular, that the findings were fully implemented across the NHS and the impact could be wholly attributed to the HTA programme (17). Buxton and Hanney's Payback Framework (19-22), previously commissioned to assess the NIHR HTA programme's first ten years (1993–2003) and recognised as the most dominant methodological framework used to assess the impact of healthcare research (23), categorises benefits of health research ranging from traditional knowledge production and research training and targeting, to impacts on policy and product development through to health and economic gains. The impacts primarily focused on the areas of knowledge generation, perceived policy impact and, to some extent, on practice – with the finding that impact on knowledge generation was more easily quantified than that on policy, behaviour or especially, health gain.

The focus of our research is to go beyond decision-making outcomes as to how informed decisions translate into improved implementation, ultimately leading to impact in terms of health gains. As Garrido et al (2008) (24) 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"*. Only when those decisions result in implementation and practice change, can better health be achieved. Yet, we know that implementation of HTA findings is variable. Even where best practice is laid out in guidelines, adoption of those guidelines can be variable. Several studies in the UK and beyond have shown that the implementation of evidence-based guidance is uneven, even across the NHS in the UK despite it being mandatory within 3 months, with estimates for the level of adoption in the range of 25%-67% (17) (25) (26-28).

The global expansion of HTA, its variable implementation resulting in sub-optimal impact, the lack of quantified evidence on health outcomes, along with an increasing investment in these processes at the systems level in LMICs, in particular, has generated greater interest from policy makers and donors about the value and return on investment (ROI) of HTA. A lack of longer-term impact assessment may undermine its importance and value. To address this, we have developed a mixed-methods framework to quantify the value of HTA. Here, we present this framework employing quantitative methods with more qualitative explanatory approaches.

Aim

This research aims to provide a methodological framework and evidence base to i) quantify the returns on investment in HTA and ii) produce explanatory programme theory that considers individual, interpersonal, institutional and systems-level components and their interactions on the mechanisms by which HTA impact can be optimised.

Methods

We use a mixed-methodology aimed at building up a rich picture of process, uptake and impact. We present the work as 1) a mixed methods Realist Evaluation which uses quantitative data to capture an empirical and credible measure of uptake (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; and 2) an interlinked return on investment (ROI) framework which uses the quantitative data to estimate a return on investment in HTA. This ROI framework will measure the net health benefit (NHB) returns on investment in HTA. By employing theory-driven evaluation strategies, we thereby synthesis economic methods with more explanatory approaches.

Throughout the paper, as stated above, we distinguish between 'HTA' as a process at the systems level to inform priority-setting and decision-making, ie as a tool for priority-setting with its explicit consideration of costs and benefits, and 'hta' as an individual health technology assessment including evidence-based interventions, practices or policies. 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) (29). Value of implementation relates directly to individual 'hta's but in aggregating these analyses, we can get to the value of 'HTA' at a systems level. We will use case study design with purposive sampling to collect data to populate the quantitative framework as well as realist synthesis and evaluation 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 1 together with data requirements, their potential sources and the challenges that lie therein. Each component is then discussed in detail below.

Table 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
1. Return on investment (ROI) framework	Decision- analytic modelling; costing of resource use.	Capital and running costs of investing in 'HTA' at a systems level.	Ministry of Health or local government audit sources, national accounts.	Dependent on aggregating NHBs for all 'hta' decisions made. Willingness -to-pay threshold value assumed.
1. Realist evaluation (RE) - Quantitative	Value of implementation (VOI); Interrupted Time Series Analysis (ITS) or similar to provide credible measure of uptake.	Longitudinal data on utilisation with time-based covariates of pre/post uptake of 'hta' (actual VOImp); total eligible population to receive 'hta', disease prevalence (potential VOImp).	Routine administrative health systems monitoring or audit data, for example, prescribing data.	Need existence of routine health services monitoring to be in place to be able to show temporal tends in 'hta' uptake (stopping) before and after HTA recommendation; prevalence of disease often unknown or uncertain, especially in LMICs. Decision analytic model of the 'hta' into health outcomes (DALYs/QALYs) is undertaken as part of the assessment process.
1. Realist evaluation (RE) - Qualitative	Qualitative realist interviews using case study design	Qualitative data; realist synthesis of the literature.	Relevant stakeholders in HTA.	Availability and willingness of stakeholders to participate; ability to provide insights to help refute or refine theories.

1 Return on Investment framework

We start by defining return on investment (ROI) and express this in basic terms as:

ROI= (Gain from investment-Cost investment)/Cost investment

Gains are expressed as a percentage of the initial investment and so a high ROI means the investment gains compare favourably to the cost. These gains are typically measured in monetary terms as a financial ROI, or can also be expressed in terms of social values as a social ROI (SROI), the only difference being how these gains are measured as costs remain the same in both cases (30). 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 'hta' 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 social outcomes and values of 'HTA' could be maintained. It would be more broadly applicable to 'HTA' as it avoids the problem raised above. Rather than monetarising these benefits as is the norm in a SROI (31), we propose the use of net health benefits (NHBs), expressing costs in terms of their health equivalence by dividing through by the willingness-to-pay threshold (λ) for a quality-adjusted life year (QALY), thereby combining costs (C) and effect into a single metric. By employing a total net benefit approach, our framework allows for the explicit consideration of the costs and benefits associated with the current level of uptake of an 'hta'. An 'hta' 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 (32). Any new 'hta' must, therefore, provide an incremental cost per QALY in the UK of less than £20,000-£30,000 to maximise health subject to a budget constraint. This concept of opportunity costs as expressed through the threshold is central to NHBs. This is also integral to any ROI. Slotting the NHBs metric into the original ROI equation, we can express a NHB return on investment (NHB ROI) as follows:

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

2. Realist Framework

2.1 Realist Evaluation - quantitative data

2.1.1 Value of implementation – individual 'hta'

We define the impact of 'HTA' to be achieved through increasing the uptake of net beneficial 'hta's and decreasing the uptake of non-net beneficial 'hta's. As stated above, an 'hta' is considered beneficial if it provides more overall health than it displaces as a result of its additional cost. By employing a VOImp analysis, the objective is to convey the concepts of potential¬ population NHBs and realised population NHBs. This is depicted in Figures 1 – 2. Implementation is shown across the top of the figure which equates to NHBs shown running along the bottom. The potential population NHB associated with full implementation of an 'hta' 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. The net population health benefit of introducing an 'hta' 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 (33). Realising a net population health benefit then requires using available evidence or assumptions on the degree of uptake (stopping) of the 'hta' (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. Indeed, it is recognised that healthcare technologies that are deemed cost-effective (or beneficial) are rarely immediately implemented perfectly into clinical practice (29, 34) because of lack of knowledge, behaviour change resistance or due to wider structural factors (33). Inefficiencies exist in healthcare from the under- or over-use of beneficial interventions as less than optimal adherence implies net benefit forgone. VOImp is consequently about assessing the value in increasing (decreasing) utilisation of the intervention to its optimal level - the value of perfect implementation - as weighed against the costs and benefits of any implementation strategies (33).

To what extent we can attribute the 'HTA' process to the uptake (stopping) of an 'hta' requires a counterfactual as to what the level of uptake (stopping) might have been without this. There might have been some natural diffusion or decline in usage of an 'hta' anyway which would reduce the overall value and impact of the 'HTA' process itself. To estimate the extent to which implementation would have happened without the 'HTA', we draw on methods increasingly applied to natural or quasi-experiments where randomisation is similarly unfeasible. In an interrupted time series (ITS) design, data are collected at multiple instances over time before and after an intervention to detect whether the intervention has an effect significantly greater than the underlying secular trend (35). By applying ITS, we can evaluate the effect of a recommendation arising from the 'HTA' process (the intervention) upon the uptake (stopping) of an 'hta', accounting for pre-intervention trends. A minimum of three time-based covariates are required whose regression coefficients estimate the pre-intervention slope, the change in level at the intervention point, and the change in slope from pre-intervention to post-intervention. The key assumption is that without the intervention ie the 'HTA process', the pre-intervention trend of the uptake (stopping) of an 'hta' would continue unchanged into the post-intervention period and that there are no external factors systematically affecting the trends (36).

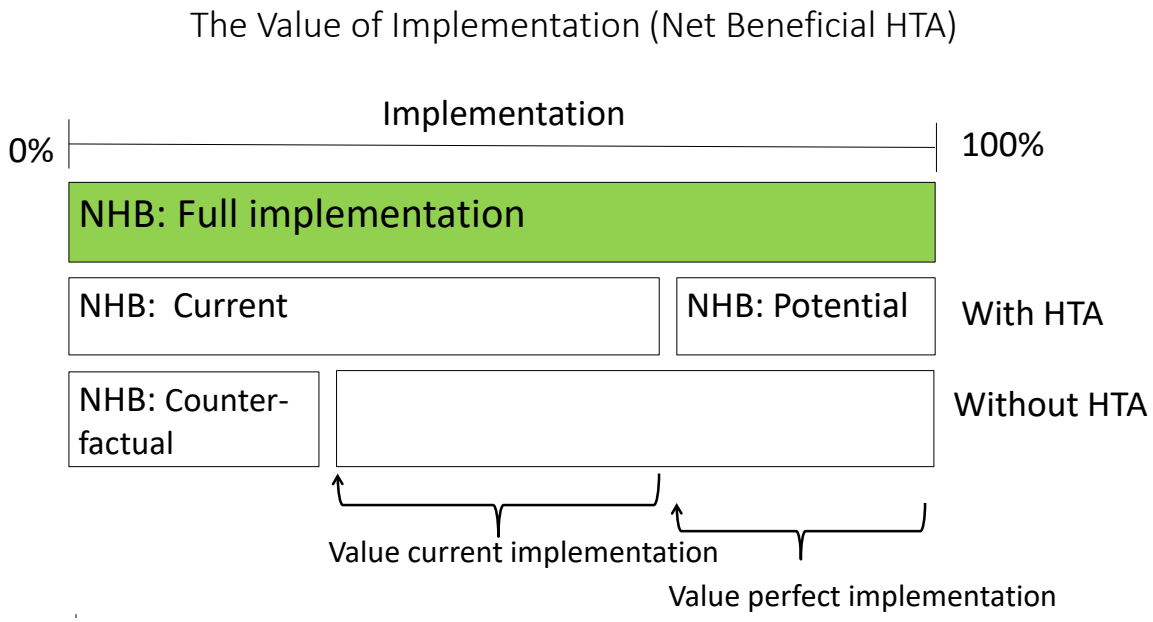

Figure 1: The value of Implementaiton (Net Beneficial 'hta')

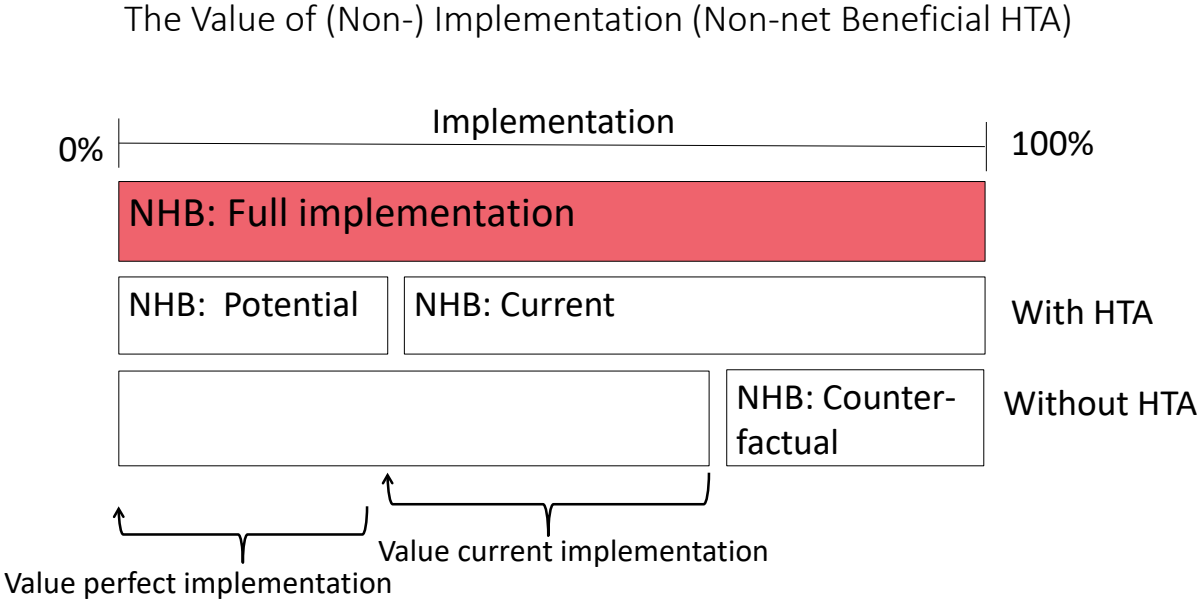

Figure 2: The value of (Non-)implementaiton (Non-net Beneficial 'hta')

2.1.2 Value of implementation – HTA at the systems level

By aggregating the realised NHBs 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 as measured against a counterfactual, we are able to offset these total net benefits against the total 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 3 depicts the ROI-NHB impact framework for HTA. It shows the NHBs stemming from each individual 'hta', numbered 1-K, associated with current and full levels of implementation. The costs of undertaking the assessment and appraisal process for each 'hta' numbered 1-K are also shown. The summed NHBs associated with 'current' and 'counterfactual' levels of implementation are circled,

those being the 'realised' and 'realised and attributable' benefits of the 'HTA' process respectively. The expected value of perfect implementation is the difference between the total net benefit of an 'hta' with perfect use (full implementation) and that with its current use. The fixed costs (CFC) of investing in 'HTA' and the running costs (C) are then summed. These total costs associated with the 'HTA' process can similarly be expressed in terms of their health equivalence by dividing through by the willingness-to-pay threshold (λ).

Finally, the aggregated total costs and benefits of 'HTA', both expressed in NHBs, can be directly offset against each other. Net gains or losses can be expressed as a percentage of the initial investment to obtain a return on investment.

Fixed cost 'HTA' infrastructure	Number of 'hta's undertaken	Cost of undertaking each 'hta' process	Level of 'hta' implementation		
			Current	Full	Counter-factual
Fixed costs	1	Cost 'hta' process $_1$	NHB_1^{CI}	NHB_1^{FI}	NHB_1^{CF}
	2	Cost 'hta' process $_2$	NHB_2^{CI}	NHB_2^{FI}	NHB_2^{CF}
	…	….	….	….	….
	…	….	….	….	….
	…	….	….	….	….
	K	Cost 'hta' process $_K$	NHB_K^{CI}	NHB_K^{FI}	NHB_K^{CF}

Total cost HTA (as expressed in NHBs)

$$\frac{C_{Fc} + \sum_{k=1}^{K} C_k}{\lambda}$$

Total benefit HTA (ΔNHBs)

$$\sum_{i=k}^{K} NHB_k^{CI} \quad \sum_{i=k}^{K} NHB_k^{FI} \quad \sum_{i=k}^{K} NHB_k^{CF}$$

Figure 3: The ROI-NHB Framework for HTA impact

2.2. Realist Evaluation – qualitative data

At the heart of impact evaluation is a requirement to link cause and effect in order to explain 'how' and 'why' – as well as addressing 'how much' and 'to what extent'. Theory driven methods will be used alongside the quantitative framework described above to generate, test and refine explanations for a gap between potential and realised gains in population health. As macro-social structures or systems are not amenable to being examined by experimental methods, researchers have adapted Pawson and Tilley's (1997) (37) ideas of 'realistic evaluation'; and a 'theory of change' perspective developed by the Aspen Institute in the USA (Judge et al., 1999). These approaches acknowledge social programmes as complex, open systems where the potential for traditional evaluation methods to explain how or why the programme works is viewed as weak, particularly given that the variety of variables cannot be controlled. "A program theory is an explicit theory or model of how an intervention, such as a project, a program, a strategy, an initiative, or a policy, contributes to a chain of intermediate results and finally to the intended or observed outcomes" (38). It can derive from formal, research-based theory or an unstated, tacit understanding about how things work. Realism was a new conceptualisation of program theory, valuable for addressing the complexity of programmes(37, 39). It is a form of theory-driven evaluation based on realist philosophy and which is becoming increasingly used in the evaluation of complex interventions(37). In adopting a realist perspective, we treat the process of 'HTA' as a complex intervention. There is much diversity in the role and application of HTA. Such

differences reflect not only differences between health systems and their financing but also how well-developed country-specific HTA agencies and processes are, and other wider contextual issues. Thus, in theorising how the basic nature of HTA works, we need to take account that HTA is an open-ended intervention ie it is already customised and needs to be highly adapted to context (40). Theorisation allows a greater understanding of how programmes work and as such a realist evaluation asks not 'what works?' but 'how or why does this work, for whom, in what circumstances?' (37) (41). Realist theory starts with the basic premise that underlying mechanisms operating in particular contexts generate outcomes and makes explicit use of a 'context-mechanism-outcome configuration' (CMOC) to aid theory development. According to realism principles, it is actors – not the programme – who make the outcomes possible through various mechanisms. Programmes are understood to provide resources or to change contexts to which participants respond: it is the interaction between resources and response that creates outcomes. Outcomes arise from a combination of the stakeholders' choices (reasoning) and their capacity (resources) to put these into practice (38, 42).

Realism works at a middle level of abstraction, and it is likely there is relevant theory in a range of different sectors that could be combined or adapted and applied as to how HTA is uptaken leading to health outcomes in order to produce an initial program theory or hypothesis (43). For example, we could potentially draw on the knowledge translation theories outlined below as to how knowledge is utilised as well as a growing body of research on the barriers and facilitators to the implementation of cost-effective interventions as a useful place to start to theorise the uptake, and thus impact, of HTA (7, 8, 34, 44-47). Battista and Hodge (48) in their 'natural' history of HTA development outline the ways in which HTA has evolved in different health care systems around the world. They identify knowledge transmission strategies to relate to HTA dissemination and implementation. Knowledge transfer exchange models have also been used in capacity building for HTA initiatives . We have undertaken an initial scoping of the literature in these areas to inform our initial programme theory.

Also forming part of our initial realist programme theory, is a theory of change produced as part of the monitoring work of the international Decision Support Initiative (iDSI) , a global partnership of leading government institutes, universities and thinktanks established in 2013 to support policy makers in priority-setting for UHC. The iDSI theory of change (ToC) shows the central processes by which change comes about; it 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. Rather than developing an HTA agency in and of itself, 'institutionalising' priority-setting and HTA is about the importance of developing accepted norms and rules, and sustaining effective working relationships between relevant policymakers and research institutions(49) . This recognises that the successful implementation of evidence-based decisions depends, in large, on the decision-making processes itself and, as such, is necessary – though not necessarily, sufficient – to facilitate an impact in terms of health gains. IDSI are concerned with the initiation context for HTA, focusing on 'institutional' aspects around achieving 'better' recommendations or decisions, whilst the work presented here is about diagnosing the implementation context, targeting those 'institutional' characteristics involved in getting those recommendations or decisions into practice – its receptiveness and 'readiness' for change. Both contexts are closely aligned and influence each other.

Realism is complementary approach that focuses on one or more causal mechanisms in a program theory and explores what it is about the program that makes this causal mechanism work (38). We shall produce a number of CMOCs based on iDSI's ToC key assumptions as to how this, in turn, leads to health outcomes, namely; decisions are implemented; health practitioner behaviour follows evidence and policy; beneficiaries choose to access healthcare when it is available; lack of healthcare or poor quality of healthcare is a key driver for poor health outcomes. Empirical testing would be required, necessarily involving a thorough review of current evidence, as well as the collection of primary data. 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 (37). Realist evaluation, on the other hand, is a form of primary research. We propose to do both by collecting primary data following a secondary analysis in order to provide a continuing test of the same theories with two quite different bodies of data (secondary and primary)(50). Analysis of all data collected would be undertaken using realist evaluation principles of extracting context-mechanism-outcome configurations of variables at play, and iterative, participative and collaborative approaches to interpretation. A protocol for our realist evaluation describing the proposed process of theory-building and refinement has been produced separately.

2.2.1 Realist approach – 'hta'

As stated above, we are dealing with both 'hta' as the intervention and 'HTA' at a systems' level. Both are potentially complex, operate at different levels and likely to interact. 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 an intervention or technology itself, this is about understanding how well an 'hta' works in any given context. How far along the implementation path we get (Figures 1-2) depends on this theory element. 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' (51), emerged in the wake of evidence-based medicine. It relates in our case at the level of implementing an individual 'hta' (52). Formal theories from this field could include, for example, those of socio-cognitive behaviour change.

2.2.2 Realist approach – HTA at the systems level

A realist lens is equally valid applied to 'HTA' as a process at the systems' level, albeit perhaps at a more abstract level. Policy implementation is the process of carrying out a government decision (Berman, 1978) 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 (52). 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.

Discussion

We present here a methodological framework to quantify the impact of HTA. None of the methods discussed are novel per se – although realist evaluation is a relatively new approach to evaluation in healthcare and an innovative way to systematically review the literature pertaining to HTA outcomes as this is the first time, as far as we are aware, such an approach has been applied to this field. The innovation comes, rather, from their combined use with the objective to calculate the return in investment in HTA. We emphasise that we are not questioning the intrinsic logic of HTA but rather how, via its effective uptake, it leads to impact in health gains. Indeed, a mid-term review of iDSI last year highlighted the need to look beyond guideline development to implementation mechanisms. Implementation is context specific, and realism as an approach to developing programme theory is particularly relevant as it focuses specifically on the influence of context because causal mechanisms are activated only in favourable conditions. We will use this to explore whether HTA can produce the intended impacts on health without assistance from re-theorising the programme or an injection of additional resources.

There have been several impact evaluations of HTA or HTA research to-date although Lavis (53) 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 (cited Lavis 2002). The same can be said of economic and social outcomes."* We would draw a parallel with this given the requirement in our framework to evaluate and aggregate the impact of 'hta's in order to get to the value of the 'HTA' process. Most evaluations to-date have made use of qualitative methods (54). The findings of a NIHR systematic review (55) support the continued use of the Payback Framework as proposed by Buxton and Hanney to measure the impact of health research (20) (19) (22). The payback approach has been identified too as a key framework for measuring HTA impact as based on a review of HTA impact models (54). It has five categories of impact: knowledge production, research targeting and capacity building, informing policy and product development, health and health sector benefit, and broader economic benefit. Others have added additional levels of impact to their work (56). A conceptual model developed for assessing the impact of HTA in the Austrian healthcare system (54) 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 (24) and Weiss' theory of research utilisation (57). 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). Indeed, Jacob and McGregor, in one of the earliest HTA impact frameworks (58), define the impact of health technology assessments 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 and technology diffusion and utilisation, and found cost-savings into the millions through systematic documentation of its effects. RAND Europe examined the UK's NIHR Health Technology Assessment Programme, 2003-2013 (59), with a separate report assessing the potential economic returns (17). 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. NICE have also carried out audits of the implementation of its guidance as published in its, now archived, Evaluation and Review of NICE Implementation evidence (ERNIE) database but the use of simple before-and-after measures, as

applied here and by others too (54), does not provide a credible measure of uptake.

We aim to address some of the methodological limitations above; we do not stop at full implementation, we include 'negative' decisions ie not just those which could offer benefits if introduced in the NHS, either by saving costs or by improving health through better treatment, and critically, we construct a counterfactual. Ideally, we would find a sector unaffected by the HTA to make such a comparison but as groups get increasingly non comparable threats to internal validity become more plausible. Alternatively, we propose drawing on quasi-experimental methods to construct a 'no treatment' control. In particular, we propose the use of those quasi-experimental methods which would enable us to a) find variation over time and b) where aggregated data is more likely to suffice as HTA is an area-based intervention. Interrupted time series analysis (ITS) (35) is used in quasi-experiments where randomisation is unfeasible and which can provide an estimate of attribution. Additionally, ITS can also include looking at the before and after intervention outcomes for another area (the 'control area') where the intervention did not take place. This includes the difference-in-difference method or the more sophisticated synthetic controls approach (60). Each of these methods would allow us to explore any temporal change of the uptake of an 'hta' following an 'HTA' decision, using either uncontrolled or controlled pre and post measures. Other quasi-experimental methods, such as regression discontinuity design (61) cannot be applied as we do not have individual level variation. Similarly, matching treatment and controls on selected individual characteristics or scoring them on a combination of variables using propensity scores is also not applicable here (62). Interrupted time series has been used previously by Sheldon et al to explore temporal variation in the implementation of HTA guidance where rates of prescribing and the appropriate use of procedures and medical devices relative to evidence based guidance including both stopping and uptake messages was estimated (63). Their findings as to whether HTA guidance had a discernible additional effect over and above the natural diffusion of technologies was mixed.

The natural diffusion of technologies has been shown to follow an S-shaped implementation curve as predicted by Diffusion of Innovation theory (64). In applying this theory, Yates (65) highlights the importance of both "macro theory (systemic adoption, that is, organisational and structural change) and micro theory (individual change)" (38), a distinction we make too, as does a realist approach which focuses its explanatory power on individual's decisions and actions, as well the wider environment. Practically, the construction of a counterfactual to take into account the natural diffusion (stopping) of a health technology is perhaps the most challenging aspect of this research. Whilst ITS was successfully used previously to assess the implementation of NICE guidance (63), 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 'hta's. This means there is fewer data prior to the HTA decision on prescribing levels and usage against which to measure the effect or influence of the HTA recommendation. Indeed, Sheldon et al applied this method prior to NICE recommendations becoming mandatory in 2004. However, in LMICs, the intended focus of our impact framework and where these processes have not yet been institutionalised, there should be more data to be able to apply such methods as Sheldon et al did when NICE was still only advisory. Finally, we build on such existing work by incorporating these estimates of uptake into a larger ROI framework in order to show impact in NHBs.

As ITS can only infer causality and produces an average measure of effect using, in this case, aggregated data, we also bring in realism for a more explanatory, theory-building approach. 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 randomised controlled trials or quasi-experiments attempt to isolate cause to that of the intervention alone by controlling out context ie the very things that realists view to be key in explaining an intervention's success or value . A positivist or successionist paradigm, rather than explaining those causes, controls for them to isolate that one cause

ie it is deemed possible to control for and test in that open system. The question is can we ignore those causes? Causal inference is also an important part of program theory evaluation in order to be able to draw conclusions (38). Realist inquiry is operationalised through investigating generative causation via the CMO heuristic. Generative causation is understanding what underpins outcomes. Outcomes are the product of a context-mechanism association, with the nature of the context-mechanism association dependent upon the architecture of the programme in question (40). By conceiving of programme mechanisms as 'reasoning and resources' ie the (new) resources, opportunities or constraints that the programme or intervention – in our case, the 'HTA' process - introduces and 'reasoning' (for example, trust-building, motivation to act, realisation of knowledge) as to how people react in response to those resources which is influenced by context (42), there is always an interaction between context and mechanism. It is that interaction which creates outcomes (66). 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. As randomisation is unfeasible by the nature of our intervention, we conceive our drawing upon natural experiment methodology (ie utilising a measure of effect as generated by a 'successionist' design) as also involving an attempt to understand the generative mechanisms entailed in the intervention, its natural and social context as well as those possessed by the actors involved so that in addition to measuring its impact, we believe a realist approach could be fruitful. Realism is a broad methodological church(67); exploring how these successionist and generative paradigms may work in tandem is another novel aspect to this research.

Indeed, we recognise there is such a big leap between 'HTA' and better health; it is a complex process, dependent on many assumptions about local factors and systems, including linkage between decisions and budgets, delivery, implementation and data accuracy – and, alluding to realism, mechanisms. As Gilson et al state: *"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"(44). "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."*(44, 68).

Finally, as the standard form of cost-effectiveness analysis is indifferent to the distribution of health and economic outcomes, there is a need to take account of equity in any value of 'HTA' framework. Extended cost-effectiveness analysis (ECEA) was developed to address health policy assessment, specifically to evaluate the health and financial consequences of public policies in four domains: (1) the health gains; (2) the financial risk protection benefits; (3) the total costs to the policy makers; and (4) the distributional benefits (69). 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 (70). Instead of a traditional CEA, the potential incorporation of ECEA would account for financial risk, whilst DCEA could account for the distribution of opportunity costs and quantify the trade-off of health and inequality (71).

Limitations/challenges

We evidence outcomes using (a continuum of) 'uptake' as a proxy for modelling impact on health. We acknowledge that implementation does not equal impact. We also limit the framework to those benefits resulting from the 'hta's being assessed and not from the process of undertaking 'HTA' itself, although we do include its associated costs. We recognise there are likely to be other externalities (both intrinsic and instrumental benefits) to arise from 'HTA', for example, better information, better administration and better payment mechanisms. We purposively do not consider other externalities or spill-over effects arising from the HTA process in order to limit the scope. Furthermore, our understanding is that the impact of 'HTA' on health outcomes is the major gap in the literature.

Operationalising the framework to assess the return on investment 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. However, we can scale-up case studies to calculate how many 'hta's may need to be undertaken in order to get a positive return on investment.

Conclusions

We envisage the use of this research to support learning and to help optimise the impact of 'HTA' in an era of investment and expansion, in particular, into LMICs, through better understanding of its translation into health outcomes and estimates of its value for money. In particular, for LMICs, we want them to have a forward-looking model in the way that high-income countries have perhaps taken implementation and outcomes for granted. 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 and recommendations to produce guidance as how to do 'HTA' by context in order to optimise its impact on health.

Funding

This working paper 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 and the UK Department for International Development.

Acknowledgements

Mark Sculpher and Paul Revill for their support with the initial research.

Itad (http://itad.com/) for developing iDSI's Theory of Change

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. Asia. WRCfS-E. 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. 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.

8. 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.

9. Williams L, Brown H. Factors influencing decisions of value in healthcare: a review of the literature. Health Services Management Centre: University of Birmingham; 2014.

10. Goodman C. HTA 101: IX Monitor Impact of HTA 2014 [3rd:[Available from: www.nlm.nih.gov/nichsr/hta101/ta101011.html.

11. 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.

12. 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.

13. INAHTA. Published Evidence on the Influence of Health Technology Assessment. wwwinahtaorg. 2014.

14. Glassman A, Fan V, Over M. More Health for the Money: Putting Incentives to Work for the Global Fund and Its Partners. 2013 2013.

15. 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.

16. 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.

17. 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.

18. 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/.

19. Buxton M, Hanney S. How Can Payback from Health Services Research Be Assessed? Journal of Health Services Research. 1996;1(1):35-43.

20. 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).

21. Buxton MJ, Hanney S. [Developing and applying the Payback Framework to assess the socioeconomic impact of health research]. Medicina clinica. 2008;131 Suppl 5:36-41.

22. 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.

23. 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.

24. Garrido M, Kirstenen F, Nielsen C, Busse R. Health Technology Assessment and Health Policy-Making in Europe. 2008.

25. Timmermans S, Mauck A. The promises and pitfalls of evidence-based medicine. Health affairs (Project Hope). 2005;24(1):18-28.

26. 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.

27. AuditCommission. Managing the Financial Implications of NICE Guidance. 2005.

28. Grol R. Successes and failures in the implementation of evidence-based guidelines for clinical practice. Med Care. 2001;39.

29. 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.

30. SROINetwork, Aitken H. Starting out on Social Return on Investment. 2014.

31. 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.

32. 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.

33. 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.

34. 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.

35. 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.

36. 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.

37. Pawson R, Tilley N. Realistic evaluation. London: Sage publications; 1997.

38. Funnell S, Rogers P. Purposeful Program Theory: Effective Use of Theories of Change and Logic Models: Wiley; 2011.

39. e-mops. Ethics & Engagement across the Wellcome Trust Major Overseas Programmes [Available from: http://e-mops.ning.com/page/realist-evaluation-introductory-resources.

40. Jagosh J. The 'Context + Mechanism' Association: Mastering a Key Heuristic in Realist Evaluation for Innovating Complex Programmes and Policy. . In: CECAN, editor. 2017.

41. Westhorp G. Realist Impact Evaluation: an Introduction. 2014.

42. 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.

43. Westhorp G. Email correspondence from member of the RAMESES JISC online research community. https://wwwjiscmailacuk/cgi-bin/webadmin?A0=RAMESES 2017.

44. 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.

45. 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.

46. 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.

47. Smith P. The Politics of Priority Setting in Health: A Political Economy Perspective. Working paper 414. 2015.

48. Battista RN, Hodge MJ. The "natural history" of health technology assessment. Int J Technol Assess Health Care. 2009;25 Suppl 1:281-4.

49. Li R, Ruiz F, Culyer AJ, Chalkidou K, Hofman KJ.

50. Pawson R. Email correspondence from member of the RAMESES JISC online research community. https://wwwjiscmailacuk/cgi-bin/webadmin?A0=RAMESES 23.2.16. 2016.

51. Fogarty. https://www.fic.nih.gov/researchtopics/pages/implementationscience.aspx [

52. 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.

53. 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.

54. 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.

55. 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.

56. M G, Kristensen F, Nielsen C, Busse R. Health Technology Assessment and Health Policy-Making in Europe. 2008.

57. Weiss CH. The many meanings of research utlisation. Public Administration Review. 1979;39(5):426-31.

58. Jacob R, McGregor M. Assessing the impact of health technology assessment. Int J Technol Assess Health Care. 1997;13(1):68-80.

59. 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.

60. Craig P. Synthetic controls: a new approach to evaluating interventions. What Works Scotland2015.

61. Bor J, Moscoe E, Mutevedzi P, Newell M-L, Bärnighausen T. Regression Discontinuity Designs in Epidemiology: Causal Inference Without Randomized Trials. Epidemiology (Cambridge, Mass). 2014;25(5):729-37.

62. O'Neill S, Kreif N, Grieve R, Sutton M, Sekhon JS. Estimating causal effects: considering three alternatives to difference-in-differences estimation. Health services & outcomes research methodology. 2016;16:1-21.

63. 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.

64. Rogers E. Diffusion of Innovations. New York: Simon and Schuster; 2003.

65. Yates B. Applying Diffusion Theory: Adoption of Media Literacy Programs in Schools. SIMILE Studies In Media & Information Literacy Education 2001;4(2).

66. Westhorp G. Community Matters [Available from: http://www.communitymatters.com.au/gpage1.html.

67. Porter S. Email correspondence from member of the RAMESES JISC online research community. . https://wwwjiscmailacuk/cgi-bin/webadmin?A0=RAMESES 2017.

68. 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.

69. Verguet S, Kim JJ, Jamison DT. Extended Cost-Effectiveness Analysis for Health Policy Assessment: A Tutorial. PharmacoEconomics. 2016;34(9):913-23.

70. Asaria M, Griffin S, Cookson R. Distributional Cost-Effectiveness Analysis. Medical Decision Making. 2015;36(1):8-19.

71. 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.

About iDSI

The International Decision Support Initiative (iDSI) is a global network of health, policy and economic expertise, working to achieve Universal Health Coverage and the health Sustainable Development Goal (SDG 3). We support countries to make better decisions about how much public money to spend on healthcare and how to make that money go further. We believe everyone should have fair access to health, receiving the right treatment and the right medicines at the right time.

A global network

iDSI forges regional and global partnerships that share the knowledge and support needed to achieve real world health gains. We focus on building institutional knowledge within existing health systems so countries can lead their own progress towards UHC.

Evidence-based decision making

Our work is underpinned by robust evidence, analysis and decision-making that policymakers, funders and researchers can use to balance trade-offs between different policy options and model potential results to make the best choice available. As a result, health ministries are equipped to make persuasive demands on public and donor spending that will save lives.

Long-term partnerships

Our work focuses on building the skills and expertise of those involved in national health systems so they can make the best use of finite resources to solve problems for current and future generations. As this expertise grows, we facilitate regional cooperation to increase contextual knowledge and skills sharing that will improve the impact and value of healthcare spending.

Our values

- Everyone should have fair access to health, receiving the right treatment and the right medicines at the right time.
- Health systems need to develop and maintain their own skills so they can make the best use of finite resources to solve problems for current and future generations.
- Sustainable and progressive health systems are only possible when they engage and involve those with a stake in its success, from the public through to funders.

iDSI Secretariat, Global Health and Development Group
Imperial College London
St Mary's Hospital
10th Floor QEQM Wing
South Wharf Road
London, W2 1NY

info@idsihealth.org
www.idsihealth.org
@idsihealth

iDSI
Better decisions Better health