CHE
Centre For Health Economics

UNIVERSITY of York

Setting Research Priorities in Global Health: Appraising the Value of Evidence Generation Activities to Support Decision-Making in Health Care

Beth Woods, Claire Rothery, Paul Revill, Timothy Hallett, Andrew Phillips, Karl Claxton

CHE Research Paper 155

Setting research priorities in Global Health:
Appraising the value of evidence generation activities to support decision-making in health care

[1]Beth Woods
[1]Claire Rothery
[1]Paul Revill
[2]Timothy Hallett
[3]Andrew Phillips
[1,4]Karl Claxton

[1]Centre for Health Economics, University of York, UK
[2]The HIV Modelling Consortium, Department of Infectious Disease Epidemiology, Imperial College London, London, UK
[3]Department of Infection & Population Health, University College London, London, UK
[4]Department of Economics and Related Studies, University of York, York, UK

June 2018

Background to series

CHE Discussion Papers (DPs) began publication in 1983 as a means of making current research material more widely available to health economists and other potential users. So as to speed up the dissemination process, papers were originally published by CHE and distributed by post to a worldwide readership.

The CHE Research Paper series takes over that function and provides access to current research output via web-based publication, although hard copy will continue to be available (but subject to charge).

Acknowledgements

This research was funded by the HIV Modelling Consortium through a grant from the Bill & Melinda Gates Foundation to Imperial College London, London, UK. The views expressed are those of the authors and not necessarily those of the HIV Modelling Consortium. Any errors are the responsibility of the authors.

All authors contributed to the conception and design of the research study. BW conducted the analyses with input from CR and KC. BW and CR developed the first draft of the report. All authors contributed to the interpretation of results, critically revised the report, and approved the final version.

No ethical approval was needed.

Further copies

Only the latest electronic copy of our reports should be cited. Copies of this paper are freely available to download from the CHE website www.york.ac.uk/che/publications/. Access to downloaded material is provided on the understanding that it is intended for personal use. Copies of downloaded papers may be distributed to third parties subject to the proviso that the CHE publication source is properly acknowledged and that such distribution is not subject to any payment.

Printed copies are available on request at a charge of £5.00 per copy. Please contact the CHE Publications Office, email che-pub@york.ac.uk, telephone 01904 321405 for further details.

Centre for Health Economics
Alcuin College
University of York
York,
YO10 5DD, UK
www.york.ac.uk/che

Abstract

The allocation of scarce resources among competing health care priorities is a key objective in all jurisdictions, whether in low- and middle-income countries (LMICs) or high-income countries. This involves allocating resources to ensure access to health care programmes, which can deliver improvements in health, but also to managing innovation in the development of new technologies, and investing in evidence generation activities to improve health for future generations. The allocation of health care resources among competing priorities requires an assessment of the expected health effects and costs of investing resources in the different activities and the opportunity costs of these expenditures, as well as an assessment of the uncertainty in health effects and costs. Uncertainty can lead to unintended adverse health consequences, e.g., when expected benefits of an activity are not realised when implemented in practice, or resources committed by an activity are transferred away from other health improving activities.

The consequences of uncertainty can be reduced by investing in evidence generation activities that improve the information available to support future resource allocation decisions. An analytic framework is developed to assess the value of evidence generation activities to support international research funders, who have the responsibility for allocating funds among competing research priorities in Global Health. Within the framework, the costs and health benefits of evidence generation activities are assessed using the same principles as those employed when evaluating the cost-effectiveness of investments in service provision. Metrics of value, founded on an understanding of the health opportunity costs imposed by research expenditure, are used to quantify the scale of the potential global net health impact across all beneficiary populations (in net disability-adjusted life years averted), or the equivalent health care system resources required to deliver this net health impact, and research costs and their potential health opportunity costs.

The framework can be applied to answer key questions such as: whether investment in research activities is worthwhile; which research activities should be prioritised; what type of research activity is necessary and what is the most appropriate design of the research; what are the opportunity costs associated with evidence generation; what is the optimal timing of research; and whether evidence generation activities should be prioritised over investments in service provision or new technology development. An illustrative example is used to demonstrate the application of the framework for informing research priorities in Global Health.

Contents

1. Introduction 1
2. The ecosystem of evidence generation 3
3. Metrics of value for informing research priorities 6
   3.1 Assessment of the global net health effect 6
   3.2 Comparison of the net health effect of research with the opportunity costs of research funding 8
4. Methods for estimating the health benefits of further evidence generation activities 9
   4.1 Heuristic approach to the estimation of the value of research 9
5. Framework for assessing the value of evidence generation activities 13
   5.1 Theme A: Does research seem potentially worthwhile? 14
   5.2 Theme B: What kind of research would best support decision making? 14
   5.3 Theme C: Are there additional opportunity costs associated with generating evidence? 15
6. Illustrating the conceptual framework for the value of evidence generation activities 17
   6.1 Overview of illustrative example 17
   6.2 Theme A: Does research seem potentially worthwhile? 18
   6.3 Theme B: What kind of research would best support decision-making? 29
   6.4 Theme C: Are there additional opportunity costs associated with evidence generation? 31
7. Discussion 33

1. Introduction

Health care decision-makers in all jurisdictions, whether in low- and middle-income countries (LMICs) or high-income countries, face difficult decisions when allocating resources to achieve agreed social objectives. At the highest level, resource allocation decisions are made across different sectors of the economy, e.g., level of spending on health care, education, security and defence, each of which have competing multiple objectives. An important social objective of health care expenditure is to ensure that overall health is improved for the population served. This involves ensuring access to health care interventions, or programmes, which can deliver improvements in health, whilst also managing innovation in the development of new health technologies to improve health for future generations. The challenge for health care decision-makers is to determine how best to allocate limited health care resources among numerous competing priorities.

The allocation of health care resources among competing priorities requires an assessment of the expected costs and health effects of investing resources in individual health technologies (e.g., drugs, diagnostics, medical devices), health care programmes (e.g., public health interventions, health-system strengthening activities) and the opportunity costs of these expenditures. An investment in these activities may be considered to represent value to the health system if it offers an improvement in health that is expected to exceed the health that is forgone elsewhere from diverting resources away from other activities in order to accommodate the additional costs of the investment (i.e., health opportunity costs). In other words, the activity is only expected to represent a cost-effective use of resources if it offers positive *net health effect* to the health system. This holds true in both health care systems where there is an explicit assessment of health opportunity costs that fall on health expenditure, e.g., through a cost-effectiveness threshold, and systems where there is an absence of firm budget constraints, but where opportunity costs manifest in terms of other forms of expenditure, e.g., there is an implicit mechanism for rationing as increases to health care expenditure require increased taxation or co-payments.

The value of an activity in terms of its expected net health impact is based on the balance of evidence currently available. However, uncertainty in the health effects that will be delivered by a programme and its costs (and therefore health opportunity costs) is unavoidable. This means that resource allocation decisions are subject to uncertainty. Uncertainty can lead to unintended effects such as adverse health consequences to individuals, as expected benefits of an activity are not realised, and to the population, as the resources committed by the activity are transferred away from other activities.

Internationally, a significant amount of resource and effort is expended on health-related research activities aimed at supporting improved health care decision making. In HIV, for example, an estimated US$17 billion was invested over the period 2000-2016 on biomedical HIV prevention research and development [1]. Much of this research funding is aimed at better understanding current epidemiological trends and service provision, and how this would be impacted by changes to health care investments. These *evidence generation activities* encompass a wide range of types of studies and may include clinical trials, surveillance, cost studies, morbidity surveys and implementation studies. By improving the information available to support investment decisions, these evidence generation activities have the potential to improve population health by reducing the level of uncertainty in the current evidence base. However, as evidence generation activities are associated with considerable cost, and research funders have constraints on their ability to expand their research budgets, this raises an important question:

How should international research funders assess the value of evidence generation activities and prioritise between competing research studies and other calls on their resources?

This question is pertinent for a wide range of research funding organisations that must prioritise research proposals across diverse clinical areas, types of studies, geographies and target populations. Organisations responsible for this type of research resource allocation include the Bill and Melinda Gates Foundation [2], the UK Department for International Development [3], the European & Developing Countries Clinical Trials Partnership [4], the US National Institute for Health [5], and the Medical Research Council [6].

In this report, a framework is presented for assessing the value of evidence generation activities to support international research funders, who have the responsibility for allocating funds among competing research priorities. Here the focus is on research proposals that are ear-marked for research funding, and expected to have an impact on human health within the next 10-20 years (as opposed to basic physiology research, for example, which is expected to have a long pathway to impact).

Within the framework, the costs and benefits of evidence generation activities are assessed using the same principles as those employed when evaluating the cost-effectiveness of investments in service provision. This involves quantifying the population health benefits of research, and weighing these against the health opportunity costs imposed by research expenditure.

2. The ecosystem of evidence generation

Investments in evidence generation activities may be made at the international, national and sub-national level, and often these investment decisions imply a commitment of resources across the different levels. For example, a clinical trial may be funded by an international organisation, but require local resources in terms of staff-time, overheads, equipment and/or consumables. However, the value of evidence generation activities can only be realised among those populations who can benefit from the improved decisions that result from the new information. An essential first step in assessing the value of an evidence generation activity is to understand the nature of the decisions that it could influence (e.g. a choice between different strategies for HIV outreach testing, or a choice between different packages of prevention care) and the populations whose health could be influenced, which may vary in terms of their size, geographical distribution and other characteristics, e.g., membership of high-risk subgroups.

The value of generating evidence is likely to vary substantively in different populations, and the total value of the evidence generation activity can only be established by estimating the *local net health effect* of research and aggregating the local effects across all beneficiary populations (this is termed the *global net health effect* of research). The breadth of influence of the evidence generation activity, and therefore its value, will also depend on the extent to which the evidence generated in a specific population is considered generalizable across different geographies or different sub-populations defined by disease, behavioural or other characteristics. As information generated by publically funded research is a public good, the health benefits from a local evidence generation effort may be realised over a much broader population, which can substantially increase the value of the research. In some circumstances, an evidence generation activity may offer the potential to inform multiple decisions. For example, improved surveillance data may inform investment decisions across a range of prevention and treatment decisions. The principles outlined in this report also apply to these wide-reaching evidence generation activities as value can be aggregated across the different decisions that will be informed by the research.

The value of research also depends on the type of evidence generation activity. The health benefits derived from the evidence generation activity will depend on the nature of the study design and the degree to which it can reduce uncertainty, or reveal the sources of variation in outcomes. Different types of study designs inform different quantities, which may be important for decision-making about investments in health care. For example, observational cohort studies or surveillance programmes are often used to understand disease progression and outcomes in the absence of intervention, randomised controlled trials (RCTs) and implementation studies are often used to gain an understanding of intervention effectiveness, while cost and morbidity data may be collected from routine data collection exercises or surveys. The value of different study designs will depend both on the breadth of decisions that they could influence and on the likelihood that they could modify decision-making in a way that has substantive implications for population health. Another important way in which evidence generation activities can improve health is by identifying observed characteristics that explain variations in costs or health outcomes, such as epidemiological conditions or disease severity, including surveys aimed at better characterising variation in HIV prevalence across geographies or risk groups. This allows the population to be divided into finer subgroups, based on geography, risk-group or other observed characteristics. This can generate net health effect by allowing interventions to be focused only in those populations in which they deliver the greatest value. Table 1 provides examples of the types of evidence generation activities that are used to support decision-making for HIV treatment and prevention programmes in sub-Saharan Africa.

Table 1: Examples of evidence generation activities that support decision-making for HIV treatment and prevention programmes in sub-Saharan Africa

Types of studies	Key outputs used to inform decision making	Example funders
Phase I-IV clinical trials of medical interventions (drugs, diagnostics, vaccines)	• Effectiveness, e.g., in terms of individual health outcomes or acquisition of HIV; • Measures of feasibility and cost of service implementation; • Formative work to inform potential epidemiological studies, implementation studies and trials.	European & Developing Countries Clinical Trials Partnership (EDCTP) [4] UK Medical Research Council (MRC) [6] Wellcome Trust [7] US National Institutes of Health (NIH) [5]
Implementation studies	• Quasi experimental designs e.g., stepped wedge trials; • Impact of alternative models of delivery of care on programme engagement and costs in different populations and geographies.	US NIH Fogarty International Center [8] Population Council [9] President's Emergency Plan for AIDS Relief (PEPFAR) [10] U.S. Agency for International Development (USAID) [11] Bill and Melinda Gates Foundation [2]
Epidemiological studies (surveillance studies, and longitudinal follow-up)	• Prevalence and incidence of HIV and their variation across time, place and subpopulations; • Behavioural surveillance measures (e.g. number and nature of sexual partners, use of condoms); • Programmatic data on number of individuals receiving specific treatments or prophylaxis; • Response to antiretroviral therapy (viral load, CD4 counts, resistance), rates of clinical events, including mortality, rate of loss to follow-up and re-engagement in care and how these vary across geographies and subpopulations.	Nationally funded programme monitoring and surveillance data. USAID (funds Demographic and Health Surveys, DHS, alongside other international and national funders) [11] US Centers for Disease Control and Prevention (CDC) (funds Population-based HIV Impact Assessments (PHIA) surveys) [12] UK MRC [6] UK Department for International Development [3] Wellcome Trust [7] Bill and Melinda Gates Foundation [2]
Cost studies	• Programme costs and how these vary across geographies, by subpopulation, by programme scale and by service delivery modalities. May be integrated into trials or implementation studies; • Costs of long-term disease management.	Bill and Melinda Gates Foundation [2]
Morbidity surveys	• Disability weights (for computation of disability-adjusted life years, DALYs).	Bill and Melinda Gates Foundation [2]

Investment decisions relating to evidence generation activities will determine how the evidence base evolves over time and its benefits in terms of delivering population net health gains. Decisions regarding investment in evidence generation activities include:

Prioritising investments from within a research budget:

- Identifying research priorities across topics and programmes competing for funding;
- Informing the type of research necessary and the design of the research;
- Informing the timing of research, particularly when additional information that could influence decision-making is expected to become available in the short run; and
- Determining how to allocate a high-level research budget across different funding streams (which may be demarked by types of studies, disease areas, geographical areas or settings).

Accounting for the interaction between research and service provision choices:

- Determining whether an evidence generation activity that would delay routine service implementation is worthwhile.

Prioritising between research and other health-generating activities:

- Determining how to allocate health care resources between research, provision of health care, and investment in the development of new technologies.

These decisions influence the direction of significant global health resources, which impact on population morbidity and longevity. Quantifying the health implications of alternative uses of research resources, therefore, represents an important tool to inform transparent and accountable decision-making. The remainder of this report focuses on how the trade-offs implicit in resource allocation decisions about research priorities can be informed by robust quantitative analysis.

3. Metrics of value for informing research priorities

The value of an investment in research can be assessed by comparing the expected health consequences of uncertainty, with and without the additional information, at the appropriate population level, to the health opportunity costs of acquiring the new information. The opportunity costs associated with research expenditure is the value of the activities that are displaced elsewhere to accommodate the costs of research. Insofar as there is funding dedicated to research activities, the opportunity costs incurred by the research funder is the funding (and associated health benefits) that are diverted away from other types of research competing for the same resources. The cost associated with implementing the research at a local level also incurs opportunity costs. These opportunity costs are different from those incurred by the research funder because the costs are incurred on the ground by the local health care system. The opportunity cost at a local level is the diversion of resources away from service provision, or other types of locally relevant research, which falls on the local health care budget.

In this report, both health and health opportunity costs are represented in terms of disability-adjusted life years (DALYs). Analyses informing resource allocation decisions in LMICs have most commonly used DALYs as a measure of the effect of alternative programmatic choices on morbidity and mortality [13] and this measure of health is generally preferred by international donors such as the World Health Organisation [14] and World Bank [15]. Nonetheless, the approaches presented in this report apply equally to other measures of health, such as the quality-adjusted life year (QALY).

The health opportunity costs of expenditures incurred by local health care systems will vary across settings, with estimates of these opportunity costs now available for most low, middle and high income countries [16–18]. The health opportunity costs of dedicated international research funds are expected to vary substantively across funders according to their specific focus, and whether the budgets held for research are potentially fungible between research and other activities. It is, therefore, essential that any metrics representing the value of investments in evidence generation activities separate out these different types of costs to allow the different opportunity costs to be considered in the research prioritisation process.

Three metrics are proposed when appraising the value of an evidence generation activity: (i) the *global net health effect* across all benefiting local populations, which takes into account health opportunity costs incurred locally; (ii) the *equivalent health care system resources* required to deliver the health effects; and (iii) the *research costs* and their potential health opportunity costs.

Funding decisions may also be informed by the distribution of the benefits of research, which depends on the mandate and priorities of the research funder. For example, in some contexts DALYs averted in very low income settings where the burden of disease is very high may be given a higher weight, but in other contexts, research may be funded where the burden of disease is lower, e.g., if the DALYs averted by research vastly exceed those that could be generated by directly funding services within the local health care system. Therefore, it is recommended that an estimate of the global net health effect is presented, alongside an estimate of the net health effect disaggregated by specific country, risk group or other grouping that is considered important by the research funder (e.g., country income category, Global Alliance for Vaccines and Immunisations (GAVI) eligible, vulnerable or priority populations, or recipient household income).

3.1 Assessment of the global net health effect

DALYs averted by research at a local level are illustrated in Table 2 for an evidence generation activity, which is expected to influence decisions for two distinct subpopulations in two countries. The countries are only expected to differ in the opportunity costs of local health care expenditure.

For example, for country 1, the assessment of health opportunity costs is $200 per DALY averted, reflecting the amount of resource required to deliver one DALY averted in that health care system, while for country 2 the corresponding value is $500 per DALY averted. The stratification of subpopulations within these countries is used to reflect the fact that research is likely to offer different value in different subpopulations and may be used to target services to a particular subpopulation.

The health benefits of research in each subpopulation are presented in column A as the number of DALYs averted by resolving the existing levels of decision uncertainty about the quantity under evaluation. The corresponding local health expenditure that would be required to achieve this health improvement via service provision is given in column B, by multiplying the DALYs averted in column A by the country's estimated local health opportunity costs. The costs of research that fall on the local health care system, e.g., costs associated with implementing the research at a local level, are presented in column C. The corresponding health opportunity costs of the research, expressed in terms of DALYs averted, are presented in column D (research costs divided by the country's estimate of local health opportunity costs). The net health effect in each subpopulation is presented in column E, expressed as net DALYs averted (i.e., the difference between column A and D). The corresponding health expenditure required within the local health care service to achieve the net health effect is presented in column F (i.e., column E multiplied by the country's estimate of local health opportunity costs).

If the local net health effect of research is positive (for example, subpopulation 1 in country 1 and country 2), then further evidence generation is potentially a cost-effective use of health care resources in that setting. In subpopulation 2 of country 2, the expected health benefits of research, equivalent to 30 DALYs averted, are the same as the health opportunity costs associated with the research expenditure at the local level. Therefore, in this setting, research will neither improve nor reduce health outcomes (i.e., the net DALYs averted are zero). In subpopulation 2 of country 1, the health opportunity costs associated with local research expenditure exceed the health benefits of research; therefore, research is expected to reduce health outcomes (i.e., the net DALYs are negative). In other words, the opportunity cost at a local level through the diversion of resources away from service provision or other types of locally relevant research is greater than the health expected to be gained from the research in subpopulation 2 of country 1. The global net health effect is estimated as the total net DALYs averted across the subpopulations (i.e., 1,025 net DALYs averted).

Table 2: The net value of research in different settings

		Health benefits of research*		Costs of research at local level		Net value of research	
Country	**Sub-population**	**DALYs averted** (DALYS) [A]**	**Equivalent health care expenditure (,000 $) [B]**	**Research costs (,000 $) [C]**	**Health opportunity costs (DALYs) [D]**	**Net DALYs averted (DALYS) [E]**	**Equivalent health care expenditure (,000 $) [F]**
1†	1	100	20	5	25	75	15
	2	20	4	10	50	-30	-6
2‡	1	1,000	500	10	20	980	490
	2	30	15	15	30	0	0
Total across subpopulations (global values)						1,025	499

† The assessment of local health opportunity costs is $200 per DALY averted in country 1;

‡ The assessment of local health opportunity costs is $500 per DALY averted in country 2;

* For simplicity, no time horizon for research is considered here; it implicitly assumes that research reports immediately and the costs of research are incurred immediately – this assumption is revisited later in the report;

** DALYs expected to be averted as a result of the research giving greater certainty about the quantities of interest.

3.2 Comparison of the net health effect of research with the opportunity costs of research funding

If an evidence generation activity offers a positive global net health effect, while taking into account local research costs, then the research is potentially worthwhile. However, for the research to be considered net health improving the opportunity costs of the research funds also need to be taken into account. One way to consider this is to directly compare the global net health effect (e.g., global net DALYs averted) per research $ across research topics competing for the same resources. If the research funder's objective is to avert the most DALYs (and a DALY averted is considered the same across jurisdictions and populations), a cost per DALY league table of all research proposals may be used to indicate the relative priority of the research topics. In this way, the global research funder works down the ordered list, funding all proposals until the research resources run out.

The cost per DALY averted of the last funded proposal implies a value for health opportunity costs of global research funds. The relative value of the opportunity costs of research compared to other uses of health care resources, such as service provision, can start to indicate whether the research budget is low, sufficient or whether research funds should be directed towards service provision or other activities instead. For example, if the cost per DALY averted for the research budget is higher than the health opportunity cost thresholds used for service provision in the countries that benefit from the research, this indicates that overall health may be improved by transferring research funding to direct service provision instead of research.

A second way to consider the potential opportunity costs of research funds is to assess whether the global cost per net DALY averted is lower than the cost of averting DALYs in the country or health care system with the lowest cost per DALY threshold, i.e., whether research improves overall health to a greater extent than investments in service provision. For the example in Table 2, if the research costs were $1 million this would generate 5,000 DALYs in country 1 (i.e., $1 million divided by the country's estimate of local health opportunity costs of $200 per DALY averted), which is much greater than the global net DALYs averted by the research (i.e., 1,025 DALYs), suggesting that the research may not be worthwhile compared to other potential uses of the research funds (assuming that there is flexibility to use the funds outside of research). A third approach is to compare the health care expenditure that would be required to avert the same number of DALYs within service provision with the costs of research. Again, if the research costs were $1 million this would appear high compared to the $499,000 required to generate the same amount of health via direct service provision in Table 2, suggesting that research may not be a cost-effective option.

4. Methods for estimating the health benefits of further evidence generation activities

Underpinning the metrics of value for informing research priorities (Section 3) is an understanding of the local net health effect of evidence generation activities. This can be established using methods of Value of Information analysis [19,20], which quantifies the value of further research as the expected improvement in health that could be gained by reducing the consequences of uncertainty in the existing evidence base.

4.1 Heuristic approach to the estimation of the value of research

Once the decisions that will be informed by the research have been identified, it is necessary to understand the quantities that will be estimated from the evidence generation activity. For example, a RCT of an intervention will typically inform the magnitude of effect of the intervention on key clinical outcomes, which is often quantified as a relative treatment effect (e.g., an odds ratio for treatment response, or a hazard ratio for all-cause mortality). An observational study, such as a surveillance programme, might inform the prevalence of HIV in a series of geographical areas and/or risk groups. The expected value of perfect information for a single quantity or group of quantities (known as the Expected Value of Partial Perfect Information, EVPPI), can be used to provide an expected upper bound on the value of eliminating uncertainty in the existing evidence base for the particular quantity (or group of quantities). This is used to start to indicate whether the evidence generation activity is likely to be of value.

In order to estimate the value of resolving uncertainty about a specific quantity, it is necessary to understand two key elements: (i) the current level of uncertainty in the quantity of interest; and (ii) the implications of this uncertainty for the net health effect of the choice between alternative interventions. Note that there is only value associated with reducing uncertainty for quantities that would change the health care decision.

Uncertainty regarding the 'true' value of a quantity of interest can be characterised by assigning a probability distribution to reflect the probability that the quantity may take a range of different values (see Figure 1, panel b). This may be informed by the available data relating to the quantity of interest such as small previous individual studies or a synthesis of evidence from previous studies [19]. For example, if the evidence generation activity is a surveillance programme, then there is likely to be some information from existing smaller surveys regarding HIV prevalence. This data will be available as a proportion, or number, of cases out of a total number of individuals surveyed, and its uncertainty can be represented by a beta distribution. Methods for selecting an appropriate distribution to represent uncertainty and quantify the available information (e.g., mean and confidence interval from existing studies) are reported in Briggs et al (2006) [19]. In some circumstances, there may be no available information on the quantity of interest. In this case, the value of the quantity and an estimate of its uncertainty may rely on expert opinion. Formal methods of expert elicitation may be used to elicit the views of experts and to quantify uncertainty in their beliefs [21,22].

Once the uncertainty around the quantity of interest has been characterised as a distribution, it is possible to identify the probability that the quantity of interest will take a value that will result in a change in decision. Therefore, the second step required to estimate the value of resolving uncertainty about a specific quantity is to identify the value of the quantity beyond which the decision based on current information would switch to a different intervention choice, i.e., the 'trigger point' in the quantity that switches the decision. If the trigger point is considered implausible, then further research will not result in a change in decision and, therefore, research should not be conducted.

A decision between alternative interventions would select the intervention with the highest expected (average) net health effect based on existing information. The trigger point represents the value for the quantity of interest at which this decision would switch and a different intervention would become the net health maximising choice. This is illustrated in Figure 1, panel a. At the mean value of the quantity of interest (representing the 'best guess' of what the quantity value might truly be) the intervention has a negative net health effect and is not considered to be cost-effective, i.e., the health opportunity costs of implementing the intervention are too great to justify the health benefits. However, beyond a certain trigger point the net health effect of the intervention becomes positive. The area to the right of the trigger point (Figure 1, panel b, shaded area) indicates the probability that the results of research would change the decision. The corresponding bars in Figure 1, panel a, represent the potential health gains that could be accrued with the change in decision. The value of resolving uncertainty about the quantity is calculated as the health consequences of the quantity taking a value beyond the trigger point (i.e., the shaded bars in Figure 1, panel a), weighted by the probability of the quantity taking each value beyond the trigger point (i.e., the shaded area in Figure 1, panel b).

If there are multiple interventions there may be multiple trigger points. For example, if the quantity of interest is HIV prevalence, the current assessment of prevalence might support a low-intensity prevention programme, whereas if HIV prevalence fell below a lower trigger point, the decision maximising the net health effect might switch to no prevention, and, if HIV prevalence fell above a higher trigger point, the decision might switch to a high intensity prevention programme.

As well as an estimate of current uncertainties about the quantity of interest, the method outlined above requires an estimate of the (incremental) net health effect of each intervention conditional upon different values of the quantity of interest. This may be informed by an epidemiological or cost-effectiveness model that is able to provide estimates of costs and health outcomes for alternative interventions, or in the absence of a model via expert elicitation [21,22]. When a model is available, the simplest way to estimate net health effect is via one-way sensitivity analysis [19]. This is very similar to the 'one-level' simulation approach to EVPPI proposed in the literature [23]; the only difference is that for the one-level approach, the model is re-run for a series of samples from the quantity distribution rather than formulating the distribution and the conditional net health effect curve separately and then combining them.

The well-documented limitation of these approaches is that they use the mean values for the other quantities in the model when computing net health effect for different values of the quantity of interest [24]. This will produce biased estimates of the net health effect curve if the model is a non-linear function of the other quantities in the model, or correlation between the quantity of interest and the other quantities within the model exists [24].

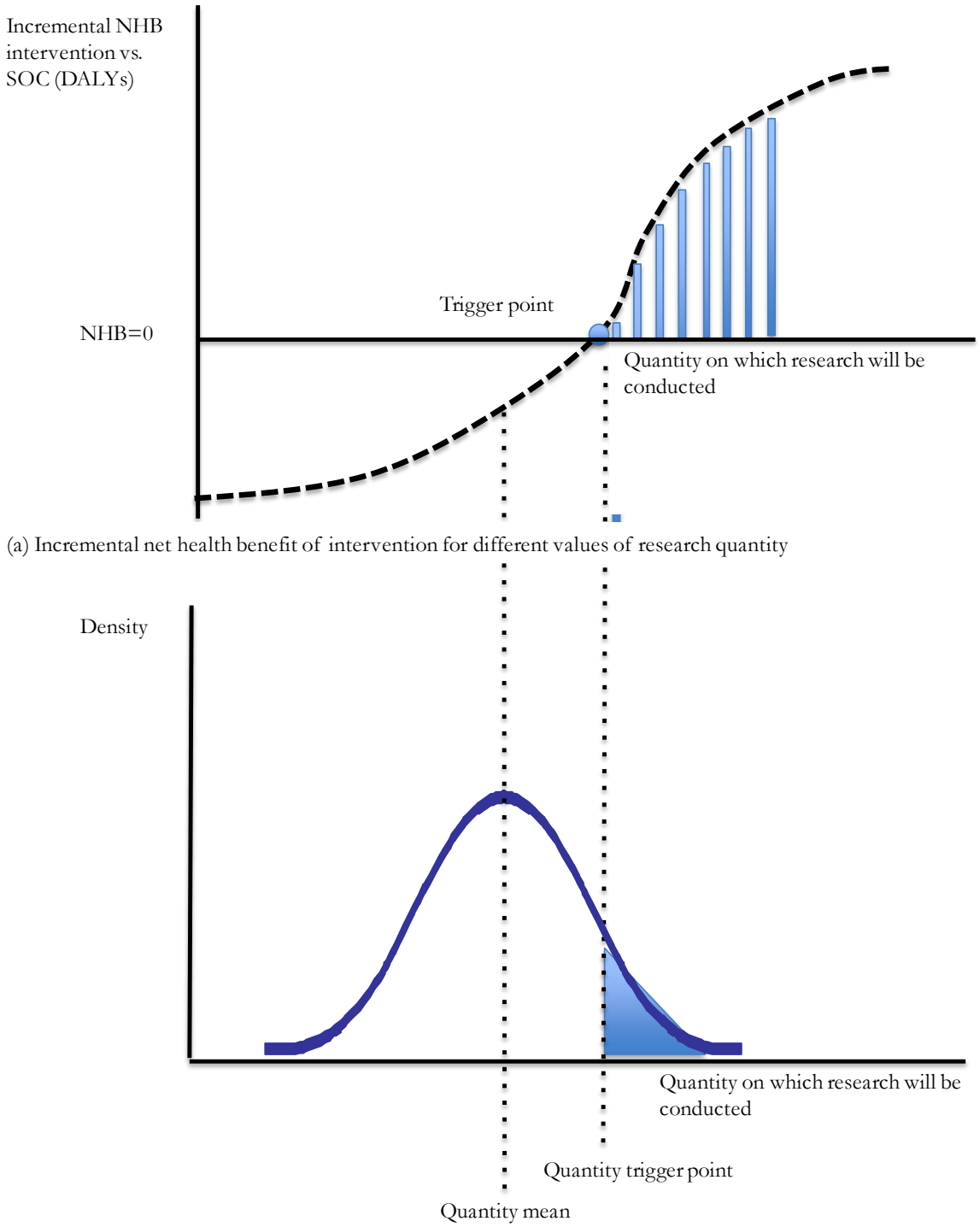


Figure 1: Estimating the value of resolving uncertainty about a quantity of interest

A range of alternative methods have been developed to address this limitation. A number of these methods require evidential uncertainty to be quantified via a probabilistic sensitivity analysis (PSA) [19]. This involves specifying a probability distribution for each of the uncertain quantities and then propagating this uncertainty through the model using Monte Carlo simulation (i.e., drawing a value for each quantity from the specified distributions) to derive the uncertainty around intervention costs and effects, and the corresponding uncertainty around the net health maximising intervention [19]. The variation in the model output over all simulations reflects the impact of uncertainty in all of

the quantities simultaneously. This type of Monte Carlo simulation can be extended to estimate EVPPI by first sampling a value from the quantity of interest, holding that value fixed, and then sampling from the remaining quantities conditional on the value of the quantity held fixed; then repeating this sampling process many times to reflect the impact of uncertainty that is attributable to the specific quantity of interest (often known as two-level simulation) [19]. The two-level simulation approach is computationally very expensive and has been found to be infeasible for computationally expensive models. Methods have been developed to approximate the two-level simulation in order to reduce the associated computational challenges. Two broad approaches have been used: (i) replace the existing model with a meta-model and use the meta-model to generate estimates of uncertainty and value of information [25–28]; and ii) use non-parametric regression methods to approximate EVPPI from the PSA samples [29,30]. However, all of these methods still require a PSA or sufficient simulations to develop a meta-model as a starting point, which may not be available for complex models in HIV. Although it may be feasible for many models, it is likely to be so computationally expensive as to be prohibitive for real-time policy decisions. Therefore, despite the limitations of the heuristic approach, this may represent the only feasible approach in some settings and circumstances.

For some models, calibration is an important part of model parameterisation. Calibration refers to the process of estimating the parameters of the model so that model predictions are consistent with external data (e.g., data on key epidemiological trends), where the external data is often referred to as the calibration target [31]. Calibration is often used when some quantities in the model are difficult to estimate directly, and to ensure that models provide credible representation of all observed data. If Bayesian methods are used for calibration, then these can be used to generate PSA outputs, and subsequently used to generate EVPPI estimates [32]. Application of full Bayesian calibration methods may be infeasible. In these circumstances, application of a more restricted Bayesian calibration approach reflecting uncertainty only in those parameters likely to be most influenced by the calibration target(s) may be feasible. This could be used to provide revised parameter means for the remaining quantities not of direct interest, and may also provide an updated prior for the quantity of interest if this is included within the Bayesian calibration process.

5. Framework for assessing the value of evidence generation activities

This section considers how the estimates of the value of reducing uncertainty about a specific quantity may be used to inform the overall assessment of the value of an evidence generation activity. Whether evidence generation activities deliver value at the local level depends upon the policy options available to decision makers, with and without the evidence generation activity, and how these policies could improve population health [33]. Without further evidence generation activities, decision makers can either implement an intervention ("approve" policy) or retain the current service without implementing an intervention ("reject" policy). When there is uncertainty in costs and health effects and further evidence generation activities are needed to reduce uncertainty, two additional policy options are available. The evidence generation activities could be pursued alongside routine implementation of the intervention ("approval with research", AWR policy). Under this policy choice, the decision maker can withdraw the intervention when the evidence generation activity reports, if continued provision does not appear cost-effective in light of the new evidence. Alternatively, the decision maker can hold back from routine implementation of the intervention whilst the evidence is being generated and then decide whether to implement the intervention once the new evidence is available ("only in research", OIR policy). The potential incremental value of evidence generation activities will depend crucially on the extent to which the policy options that involve evidence generation activities (AWR, OIR) generate additional population health compared to those which do not (approve, reject).

This section of the report sets out a framework for estimating the additional value of policies that involve evidence generation activities, in order to inform evidence-based decision making about investments in research. When assessing the value of evidence generation activities there are three key issues to consider: (i) Does research seem potentially worthwhile?; (ii) What kind of research would best support decision making?; and (iii) Are there additional opportunity costs associated with generating evidence? These considerations are described below and summarised in Box 1.

Box 1: Considerations when assessing the value of evidence generation activities

Theme	Considerations
Theme A: Does research seem potentially worthwhile?	An understanding of how health care decisions could be influenced by the evidence generated.
	Identify quantities that contribute to decision uncertainty and on which further evidence could be obtained, type of evidence generation activities that could inform these quantities, and the value of reducing uncertainties.
	Consider future uncertainties that could impact on the value of evidence generation activities: price changes, entry of other technologies, and availability of ongoing research.
	Consider the degree to which uncertainty is expected to be resolved by the proposed evidence generation activities.
Theme B: What kind of research would best support decision making?	Consider value of study designs that provide information on different quantities: these will include studies looking at different sets of outcomes and in the case of comparative research studies, comparing different sets of interventions.
	Where multiple studies could be commissioned to inform a decision consider whether one study, multiple studies or a sequential approach should be pursued, and consider delaying research until other uncertainties resolve.
Theme C: Are there additional opportunity costs associated with generating evidence?	Consider whether research would require delayed implementation of cost-effective services and the potential for this to reduce the added value of the evidence generation activity.

5.1 Theme A: Does research seem potentially worthwhile?

This theme evaluates whether or not there is the potential for research to generate local and global net health gains, and whether or not these health gains might be considered sufficient to offset the health opportunity costs of the research expenditure. To understand the scale of the potential benefits of evidence generation activities, the number and nature of the decisions that could potentially be informed by further evidence generation activities needs to be understood. The scale of the population that could benefit from improved decision making is an important determinant of the value of research and requires consideration of the populations who are expected to benefit (Section 2) and whether the evidence is likely to be informative for decision making over a long time frame or a relatively short "shelf-life".

Once the potential beneficiaries of improved evidence have been identified, it is necessary to understand which quantities are important drivers of decision uncertainty and the type of evidence generation activities that could reduce this uncertainty. If for all plausible estimates of a quantity, the decision would be the same, then that quantity is not an important driver of decision uncertainty and knowing more about that quantity would not change decisions and cannot, therefore, improve population health. However, if the cost-effective intervention choice does change at different plausible values of the quantity, then knowing more about that quantity has the potential to improve population health and may be a valuable target for evidence generation activities.

Once these assessments have been made it is possible to estimate the expected maximum local value of an evidence generation activity. This requires knowledge about the likelihood that the uncertain quantity takes a range of values, and how those values could alter decision making and, therefore, population health, as shown in Section 4. The total potential value of conducting an evidence generation activity can be estimated by aggregating the local values of the evidence generation activity across those populations who will benefit from the information and comparing this to the opportunity costs imposed by the research expenditure (see Section 3).

In some contexts, future changes that are uncertain and will only resolve over time may substantively modify the value of evidence generation activities. For example, even if the current benefits of research are considerable, if the price of the technology is likely to fall significantly before or shortly after the research reports, or if future innovation makes the current technology obsolete, then the future benefits once the research reports might be very limited. Therefore, the expected impact of future changes over time should be assessed.

Finally, it is necessary to consider the extent to which the proposed evidence generation activity will reduce uncertainty about the quantity of interest. For example, a large and expensive study that provides a more definitive answer regarding the true value of the quantity of interest will represent quite a different value proposition compared to a small-scale relatively cheap study, which will increase our understanding of the quantity but still leave the decision maker with a substantive degree of residual uncertainty. The proposed study (e.g., RCT) may be designed to test the hypothesis that the true value of the quantity is greater than the trigger point that switches the decision.

5.2 Theme B: What kind of research would best support decision making?

Theme A may identify a series of quantities for which research is expected to be potentially worthwhile, and different studies that could support improved understanding of these quantities. This raises two questions: (i) how should individual studies be designed?; and (ii) how should research programmes involving multiple evidence generation activities be designed?

The design of any research study implies a large number of choices, which will have implications for the value of the evidence generation activity. The size of the study, and other factors which could influence its ability to resolve uncertainties about the quantity of interest, will have important implications for both its value and cost. The quantities which can be estimated by the study may also profoundly impact upon the value of the study. Different study designs may allow a comparison of different interventions and may provide data on different biochemical, care-related or economic endpoints. For example, in the context of outreach testing in HIV, two important uncertain quantities are the cost per HIV case identified and the likelihood that individuals newly identified as HIV positive will be linked to appropriate care [34–37]. An appraisal of both a short, cheaper study, looking at the cost per HIV case identified and a longer more expensive study examining this outcome, as well as recording linkage rates, would allow the value of these alternative designs to be compared. This could potentially allow research funds to be used more effectively than if only one study design was considered. However, careful consideration must also be given to the generalisability of the findings from different studies; for example, studies examining the same outcome may vary significantly in terms of the characteristics of the specific setting, the outreach approach, and changes over time. The extent to which outcomes can be generalised across studies will depend, in part, on the outcome examined; for example, a study examining the intrinsic efficacy of a new intervention is less likely to vary across populations compared to studies examining outreach testing in HIV.

Another important aspect of research design is the selection of intervention and comparator study arms in comparative research. For example, if decision makers are currently uncertain about whether a high or low intensity version of an intervention is cost-effective, then this comparison should be the focus of the comparative study, while a study comparing the high intensity intervention to a no intervention option may be of limited value. If there is uncertainty about which of the three interventions is cost-effective (high-, low- or no intervention) then a three-arm study may represent better value than a two-arm study. Again, assessing the value of alternative designs and their costs may allow for more effective use of research funds and enable the identification of high value evidence generation activities for targeted investment.

In some contexts, multiple evidence generation activities may be valuable. For example, a trial focused on short-term outcomes and an observational study linking short-term clinical outcomes to long-term patient morbidity and mortality may appear valuable. In this case a decision maker has the option to commission one or both studies, commission one study and review its results prior to deciding whether to commission the second (a “sequential” research programme), or simply waiting to commission one or both studies until other uncertainties are resolved (a “watchful waiting” approach). Assessing the value of alternative *programmes* of research again offers the potential to identify high value programmes at which to target research funds.

5.3 Theme C: Are there additional opportunity costs associated with generating evidence?

If research is considered valuable, then, in principle, the AWR policy can be pursued when the intervention is cost-effective and the OIR policy pursued when it is not. However, parallel service implementation and research (i.e. AWR) may be impractical for a number of reasons. Firstly, service implementation while research is underway could contaminate the findings of research. Secondly, it can raise ethical issues, and/or make recruitment into the research impracticable. Thirdly, it may not be politically appropriate or practical to withdraw the intervention after the research reports, even if the evidence generated does not support continued implementation (i.e. an AWR policy would actually be an approve decision but with added research costs). Therefore, there may be additional opportunity costs associated with implementing an evidence generation programme if it prohibits wide-scale implementation of the service while the research activity is conducted. If this is the case,

then the health implications associated with not implementing a cost-effective technology while research is underway needs to be accounted for when evaluating the evidence generation activity.

6. Illustrating the conceptual framework for the value of evidence generation activities

A simple illustrative example is used to show how the framework can be applied to provide quantitative assessments of the value of evidence generation activities.

6.1 Overview of illustrative example

The illustrative example considers the decision about whether or not to provide an intervention which aims to prevent HIV infection. Only these two policy options are considered for illustrative purposes; in reality, the decision-maker may face a wider range of alternative options (e.g., in HIV where there may be a need to consider different combinations of interventions and different population coverage levels). Two countries are considering: a low income country (LIC) and a middle income country (MIC). Decision makers in each country face the decision about whether to provide the intervention in a low- and a high-risk subpopulation. A simple decision tree is used to determine the risk of infection, with and without the intervention. This risk depends on the baseline risk of infection and the reduction in infection risk associated with the intervention, which is represented as an odds ratio for the intervention compared to standard of care (SOC). For the illustrative example, the only difference between the LIC and MIC is the assessment of health opportunity costs (i.e., cost-effectiveness threshold) which is $200/DALY for the LIC and $500/DALY for the MIC. The quantities used in the model are presented in Table 3.

Table 3: Quantities used in the illustrative model

Quantity	Value
Probability of infection in low-risk population	0.10
Probability of infection in high-risk population	0.15
Effect of intervention on infection risk (odds ratio for intervention compared to SOC)	0.80
Cost of intervention per individual	$15
Discounted total cost per infected individual	$ 150
Discounted total DALYs per infected individual	1

The implications of resolving uncertainty is explored in three quantities: (i) the probability of infection in low-risk individuals; (ii) the probability of infection in high-risk individuals; and (iii) the effect of the intervention on the risk of infection. Uncertainty about each quantity is characterised as the probability that the quantity may take different values given the current level of evidence, which is formulated as a probability distribution (see Section 4). Uncertainty around the odds ratio (OR) is assigned a log-normal distribution (mean OR on natural scale: 0.80 with 95% CI: 0.61, 1.05) and uncertainty around the annual probability of infection in the low- and high-risk group is assigned a beta distribution (low-risk group: 0.10 with 95% CI: 0.04, 0.16; high-risk group: 0.15 with 95% CI: 0.08, 0.22).

The results of the cost-effectiveness analysis comparing the intervention to SOC are presented in Table 4. The intervention is only cost-effective in the MIC for the high-risk group, as indicated by the incremental cost-effectiveness ratio (ICER) falling below the cost-effectiveness threshold and positive net health effect. In the other groups, the health benefits of the intervention are insufficient to offset the health opportunity costs imposed by the cost of the intervention.

These results represent the *expected* cost-effectiveness of the intervention. However, due to uncertainties in the evidence base, there is a possibility that the intervention is not in fact cost-effective in the MIC for the high-risk group, or that the intervention is cost-effective in the LIC or MIC low-risk groups. These uncertainties in the decision mean that there is potential value to investing in

evidence generation activities to reduce the consequences of uncertainty and improve the likelihood that the health-maximising decision is made.

Table 4: Cost-effectiveness results for the illustrative example

Country	Subpopulation	Incremental cost per individual ($) (intervention vs. SOC)	DALYs averted per individual (intervention vs. SOC)	ICER ($/DALY)	Cost-effectiveness threshold	Net health effect per individual (DALYs)
LIC	Low-risk	$12.24	0.0184	$667	$200	-0.043
	High-risk	$11.06	0.0263	$421	$200	-0.029
MIC	Low-risk	$12.24	0.0184	$667	$500	-0.006
	High-risk	$11.06	0.0263	$421	$500	0.004

6.2 Theme A: Does research seem potentially worthwhile?

6.2.1 How can health care decisions be influenced by further evidence generation activities?

Evidence generation activities offer the potential to influence the decision about whether to implement the new intervention or to retain SOC. The benefits of an improved evidence base accrue each time a more-informed health care decision is made. Therefore, evaluating the potential benefits of an improved evidence base requires information about the prevalence and future incidence of individuals for whom the evidence could improve decision-making. It also requires a judgement about the time horizon over which the decision that will be informed by the evidence is made. For example, if additional evidence is gathered in a clinical area in which the pace of innovation and introduction of new technologies is rapid, evidence may have a relatively short shelf-life (the "technology time horizon"). A judgement is also required for the "model time horizon" over which costs and health benefits are accrued.

For simplicity in the illustrative example, a constant stream of incident cases is assumed, i.e., in each risk group (LIC low-risk, LIC high-risk, MIC low-risk, and MIC high-risk), 500,000 individuals could potentially receive the intervention per year. The intervention is used to treat individuals over a maximum of 10 years. An improved evidence base, therefore, offers the potential to inform 20 million decisions (i.e., whether or not to provide the intervention to an individual to reduce the risk of acquiring HIV) across the four populations over ten years (equivalent to 17.6 million discounted decisions using an annual discount rate of 3%).

6.2.2 Which quantities contribute to decision uncertainty, could feasibly be researched, and what would be the value of improving knowledge in relation to these quantities?

In this example, the three uncertain quantities could feasibly be considered for additional research. The risk of infection could be informed by survey data or an observational cohort study, while the relative effect of treatment could be informed by a RCT, or other study design that can provide information on relative effectiveness.

The value of evidence generation in each population: Figure 2 shows how the principles outlined in Section 4 can be applied to the low-risk populations (LIC and MIC), in order to quantify the health effects of resolving uncertainty relating to the odds ratio describing the reduced risk of infection associated with the intervention compared to SOC. The top panel shows that costs increase and DALYs averted reduce as the odds ratio increases and the intervention becomes less effective compared to SOC. Therefore, the incremental net health effect of the intervention compared with SOC decreases with increasing values of the odds ratio, as shown in the central panel. Although the cost and DALYs averted by the intervention are the same across countries, the net health effect is lower in the LIC because the investment in the intervention implies higher health opportunity costs.

At the mean value for the odds ratio (0.80), the net health effect of the intervention is negative in both the LIC and MIC. This reflects the expected cost-effectiveness results shown in Table 4. For the LIC, the odds ratio would have to take a value of 0.54 or less for the intervention to be considered cost-effective and the likelihood that the odds ratio takes a value this low is very small (probability <0.01). This is shown by the small area to the left of the trigger point in the LIC in the lower panel of Figure 2, which can be interpreted as the probability that resolving uncertainty about this quantity would result in a change in decision. Weighting the potential health consequences of implementing the intervention (as shown in the central panel of Figure 2) by the likelihood that the odds ratio takes a value that would support this decision (as shown in the lower panel of Figure 2) results in an estimate of 57 potential DALYs averted by removing uncertainty about the effectiveness of the intervention. In contrast, for the MIC, if the odds ratio takes a value below the MIC trigger point of 0.75, the intervention would be considered cost-effective. The probability that the odds ratio will take a value that results in the intervention turning out to be cost-effective in the MIC is much higher than in the LIC with a probability of 0.34, as shown by the larger area to the left of the MIC trigger point in the lower panel. Therefore, resolving uncertainty about the effectiveness of the intervention could potentially avert 11,405 DALYs in this group.

Figure 3 shows the corresponding results for the high-risk populations. At the mean value for the odds ratio (0.80), the net health effect of the intervention remains negative in the LIC but is now positive in the MIC due to the higher absolute risk reduction from the intervention. In both countries, the trigger point is closer to the mean than in the low-risk populations (LIC: 0.68; MIC: 0.83) and the probability that the odds ratio could take a value that would support a change in decision is higher (LIC: 0.12; MIC: 0.38). Therefore, resolving uncertainty about the effectiveness of the intervention offers the opportunity to avert more DALYs in the high-risk group (LIC 5,622 DALYs averted; MIC 29,233 DALYs averted) compared to the low-risk group.

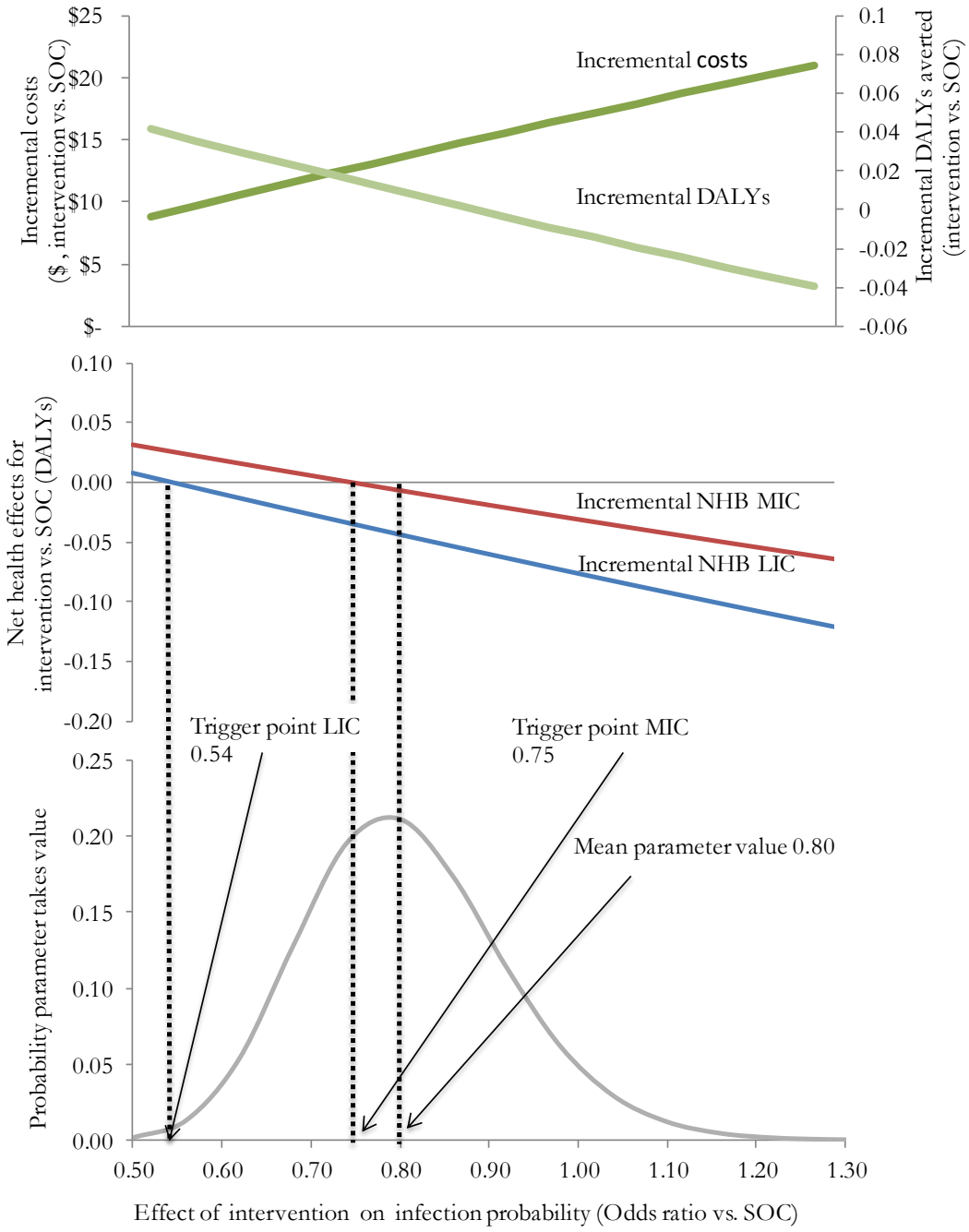


Figure 2: Estimating the value of resolving uncertainty about the effectiveness of the intervention compared to standard of care in low-risk subpopulations

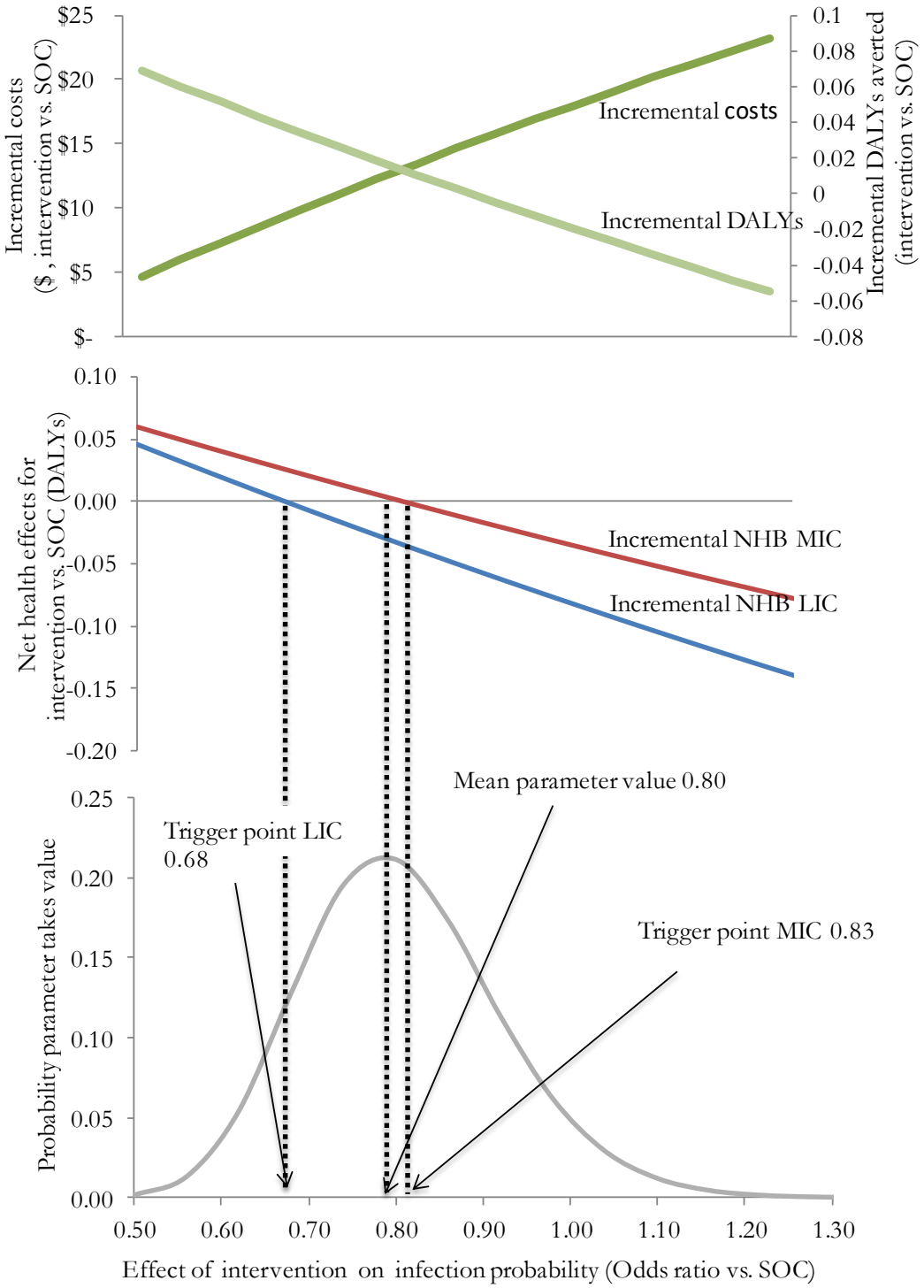


Figure 3: Estimating the value of resolving uncertainty about the effectiveness of the intervention compared to standard of care in high-risk subpopulations

Table 5 shows the results of aggregating the DALYs averted within and across subpopulations (see Section 3). This suggests that a RCT with the potential to generate improved information about the effect of the intervention on infection risk offers the potential to avert 46,317 DALYs globally. The DALY loss imposed by local research costs can be estimated by dividing the research costs by the local cost-effectiveness threshold. Subtracting these DALYs from the DALYs averted via research provides an estimate of the net local DALYs averted (or incurred). Where research offers local value, i.e., in the LIC high-risk, MIC low-risk, and MIC high-risk populations, the local net health effect can be aggregated to estimate the global net health effect associated with research, which is 45,966 DALYs averted in this example. It is informative to understand the local health care resources that would be required to generate these DALYs directly via service provision (see Section 3). In total, approximately $21.3 million would be required to avert the same number of DALYs via direct health care financing. This is estimated by multiplying the DALYs averted by the local cost-effectiveness thresholds.

The above analysis has assumed that the study needs to be run in each subpopulation in order to generate health benefits, i.e. that the evidence is not considered generalisable across subpopulations. However, in practice, evidence on relative treatment effects, such as that obtained via an RCT, is often considered generalizable between jurisdictions and populations with different characteristics [38]. If this is the case, then a decision maker may decide to run the trial in one population and utilise the resulting information to inform decisions made within other populations [39]. This may ultimately reduce research costs, both locally and for the research funder. To maximise the global net health effect of research, the decision maker could look across those contexts in which the study delivers local net health benefits and choose to only run the study in the group in which the DALYs incurred due to local research costs are lowest, i.e., in the MIC low-risk group in this example (note, however, that the decision maker is also likely to factor in their own research costs and how these vary across populations). This would reduce the local health forgone due to research costs by 215 DALYs (i.e., 125 DALYs in the LIC high-risk group plus 90 DALYs in the MIC high-risk group). This would also allow the LIC low-risk group to benefit from the research without incurring any research costs, thus increasing the DALYs averted via improved information by 57 DALYs. The total DALYs averted by research would therefore increase to 46,237 and approximately $21.4 million would be required to avert the same number of DALYs via direct health care financing.

The same approach can be repeated to understand the value of improving the evidence relating to infection risk via additional survey data (Table 6). Data on the risk of HIV acquisition is unlikely to be considered generalizable across populations. Therefore, the research study must be run in each population in order to generate value in the subpopulation, but the research should only be run in subpopulations for which the local net health effect is expected to be positive. The total potential DALYs averted by the survey are 8,062 and the equivalent health care resources required to avert these DALYs via direct service provision are $4.0 million.

These assessments provide an estimate of the aggregate health benefits of potential evidence generation activities taking into account the local opportunity costs of research expenditure. They also show how the population health benefits are distributed across different populations and countries. In this example, the benefits of both the RCT and survey accrue predominantly in the MIC with a comparably low number of DALYs averted by research in the LIC (Table 5 and Table 6).

Table 5: Aggregating the value of an RCT comparing the new intervention to standard of care across populations

Country	Sub-population	Trigger point	Decision error probability*	DALYs avertable via improved information	Equivalent health care expenditure ($)	Local research costs	Local health forgone due to research costs	Net local health effects(DALYs)	Net local value (equivalent health care expenditure, $)
LIC	Low-risk	0.54	0.002	57	$11,348	$20,000	100	-43	-$8,652
	High-risk	0.68	0.124	5,622	$1,124,469	$25,000	125	5,497	$1,099,469
MIC	Low-risk	0.75	0.335	11,405	$5,702,487	$40,000	80	11,325	$5,662,487
	High-risk	0.83	0.384	29,233	$14,616,598	$45,000	90	29,143	$14,571,598
Total				46,317	$21,454,902	$130,000	395	45,922	$21,324,902
Total assuming no generalisability**				46,261	$21,443,554	$110,000	295	45,966	$21,333,554
Total assuming perfect generalisability***				46,317	$21,454,902	$40,000	80	46,237	$21,414,902

* Probability of quantity lying beyond trigger point; ** The trial will not be run in the LIC low-risk group where local net health effect are negative; *** The trial will be run in the MIC low-risk group and the evidence generalised to inform decision-making in all groups.

Table 6: Aggregating the value of a survey of HIV risk across populations

Country	Sub-population	Trigger point	Decision error probability*	DALYs avertable via improved information	Equivalent health care expenditure ($)	Local research costs	Local health forgone due to research costs	Net local health effects(DALYs)	Net local value (equivalent health care expenditure, $)
LIC	Low-risk	0.26	0.000	0	$0	$5,000	25	-25	-$5,000
	High-risk	0.27	0.001	7	$1,343	$5,000	25	-18	-$3,657
MIC	Low-risk	0.14	0.125	2,742	$1,371,071	$10,000	20	2,722	$1,361,071
	High-risk	0.13	0.335	5,360	$2,680,020	$10,000	20	5,340	$2,670,020
Total				8,109	$4,052,434	$30,000	90	8,019	$4,022,434
Total assuming no generalisability**				8,102	$4,051,091	$20,000	40	8,062	$4,031,091

* Probability of quantity lying beyond trigger point; ** The survey will not be run in the LIC low- or high-risk groups where local net health effect are negative.

Accounting for the timing and likelihood of research: The value of evidence will also depend on the time it takes for the research to be conducted and report, and the likelihood that it will report. Research designs that take a long time to complete and report will have a lower value due to the reduced time horizon over which the available evidence can be utilised. Similarly, the less likely a study is to report, the lower its expected value. Both the time taken for research to report and the likelihood that it reports have approximately proportionate effects on the value of evidence (Figure 4)[1]. For example, a study that is considered to have a 80% likelihood of reporting, takes 4 years to report, and could inform decision-making up to year 10 would deliver 48% of the value shown in Table 5 and Table 6. Incorporating the timing and likelihood of research may fundamentally change the value of the research proposal. It may also affect the relative value of alternative uses of research expenditure, tilting decisions towards those proposals that are expected to report quickly and have a high probability of reporting. Understanding the relationship between the time taken for research to report, the likelihood that it will report, and the value of the evidence to future populations can also help inform: (1) investments that might make research findings more quickly available; (2) the trade-off implicit in the choice of alternative research designs; and (3) identification of those areas where, if research is to be undertaken, there must be confidence that it can report quickly [33].

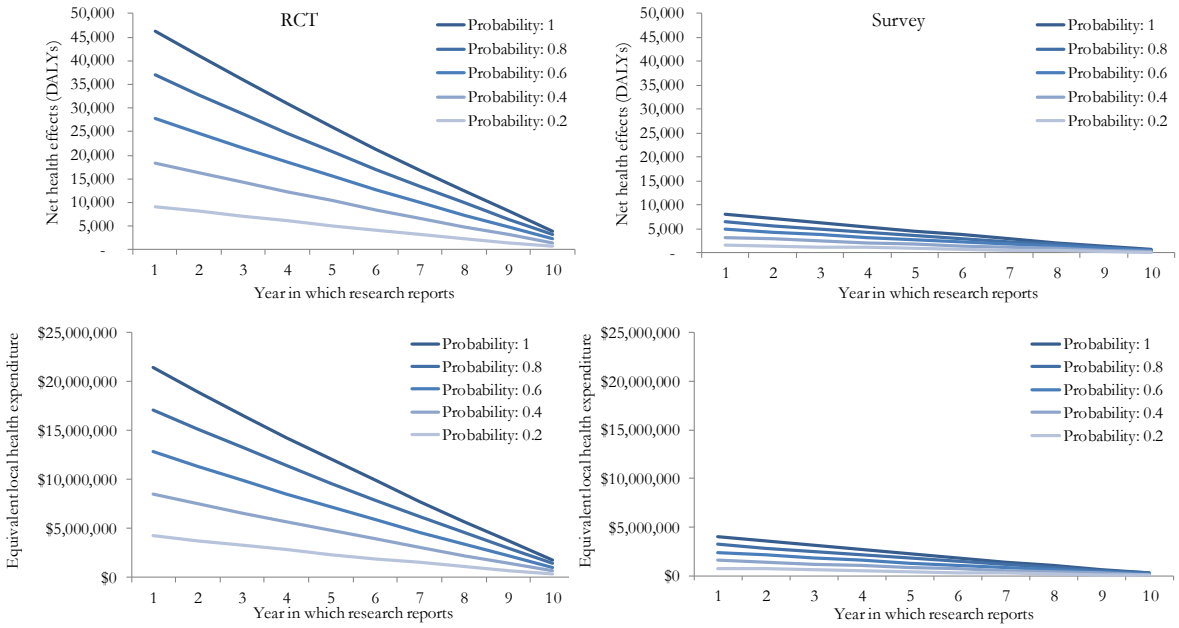


Figure 4: Impact of time taken for research to be conducted and report and the likelihood that it will report on the value of the evidence generated

6.2.3 Are future uncertainties likely to impact on the value of evidence generation activities?

The benefits of further evidence generation activities depend upon the presence of other sources of uncertainty: changes in the prices of the alternative interventions and comparators; the emergence of new technologies that might make existing ones obsolete or change their cost-effectiveness; and other relevant research reporting. The impact of future uncertainties on the value of evidence generation activities should be considered carefully when the future uncertain event is likely to occur and/or will occur before or soon after the research reports. The qualitative effects of future uncertainties are described below. Methods for adjusting the quantitative assessments above for the effects of future uncertainties are presented in Appendix A.

[1] Unless discount rates are very high.

Price changes: Changes in price not only influence expected cost-effectiveness but also uncertainty and the potential benefits of research to future patients. Price changes have had important implications for the investment priorities of HIV programmes and the development of the epidemic. For example, prices for antiretroviral therapies dropped markedly in the 2000s in sub-Saharan Africa as trade rules were relaxed to allow importation of generic forms of these drugs whilst the drugs were within their patent period [40]. A more recent example is the cost of viral load assays which have fallen over time as a result of agreements with manufacturers and in response to the volume of demand [41]. Price *reductions* are discussed here as these are more commonly observed.

If an intervention is expected to be cost-effective then a price reduction will generally reduce the potential benefits of evidence generation activities, since the cost-effectiveness of the intervention will be less uncertain and there may be less to gain from further research. If an intervention is not expected to be cost-effective then a reduction in price will generally initially increase the value of research until the intervention becomes cost-effective, but then eventually reduce as it becomes increasingly likely that at the lower price the intervention would be cost-effective[2]. The value of evidence generation activities will not necessarily fall to zero as price falls to zero if there is a possibility that the intervention may cause harm. In this situation, even at a zero price, there may be value in better understanding the likelihood and magnitude of that harm. In the illustrative example, a price decrease reduces the value of evidence generation in the MIC high-risk population, and potentially increases or decreases the value of evidence generation in the other populations.

Assessing the impact of a price change requires information about when major changes in price are expected, how likely the price change is, and some evidence about the anticipated extent of the price change. Figure 5 illustrates the implications of a 60% price drop at year 4 with a 50% likelihood of occurring. The DALYs averted by both the RCT and the survey increase with the price drop, and the distribution of the benefits of evidence generation changes markedly. Without the price change, benefits are expected to accrue almost exclusively to the populations in the MIC, however, with the price drop the majority of DALYs are averted in the LIC. The price drop means that there is much less uncertainty about whether to implement the intervention in the MIC. Therefore, the benefits of research are focused in the LIC.

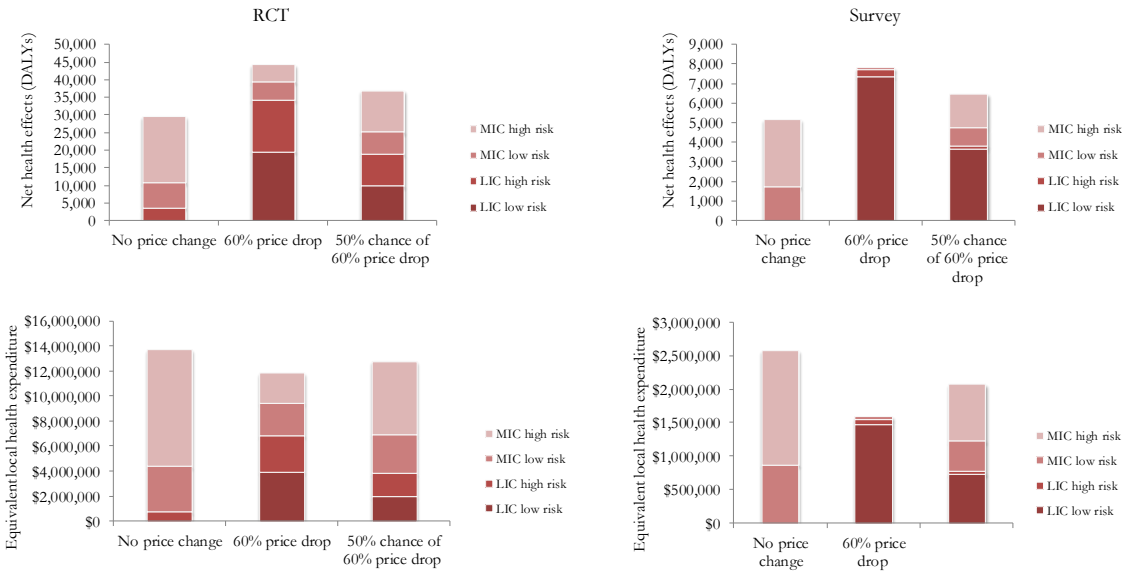


Figure 5: Impact of a future price change on the value of evidence generation activities
Note: the figure assumes that both research studies will report at year 4 and have a 100% likelihood of reporting.

[2] These assessments become more complex when there are three or more interventions, and a geometric assessment may be necessary to determine the direction of effect.

Entry of new interventions: The entry of a new technology will tend to change the relative cost-effectiveness of the alternative interventions and influence uncertainty in the choice between the interventions. The impact of a new intervention on the value of evidence generation will depend on whether the new intervention is expected to be cost-effective, and whether this expected cost-effectiveness changes at different values of the quantity that will be researched via the evidence generation activity (Table 7). If the new intervention is not expected to be cost-effective across plausible values for the quantity under consideration then its availability will not impact upon the value of the evidence generation activity. In other words, there is no result of the research that could result in the new intervention being adopted. If the intervention is not cost-effective on expectation but could be at some values of the quantity of interest, then the value of research will be increased. This is because under some realisations of uncertainty the new intervention offers the potential to generate higher net health effect than current interventions, making the research more worthwhile. If the intervention is cost-effective on expectation, and at all values of the quantity of interest, then the value of the evidence generation activity will be zero. If the intervention is cost-effective on expectation, but not at all values of the quantity of interest, then the effect on the value of evidence generation activity is not clear and understanding the direction of effect may require geometric reasoning or quantitative analysis.

Table 7: Impact of new interventions on the value of evidence generation activities

New intervention cost-effective on expectation	Decision about new intervention switches at some values of the quantity of interest	
	No	**Yes**
No	Value of evidence generation activity **unchanged** by entry of new intervention	Value of evidence generation activity **increased** by entry of new intervention
Yes	Value of evidence generation is **zero**.	Value of evidence generation activity **may increase or decrease** with entry of new intervention.

Other research reporting: Research that is already under way, commissioned, or likely to be undertaken is relevant as there is a chance that it will change the estimates of cost-effectiveness and resolve some of the current uncertainties. The value of evidence generation activities may be reduced if decision uncertainty is likely to be resolved in the near future when other research reports[3]. The impact of ongoing research on the value of the current evidence generation activity is difficult to predict as it depends upon how the value of the evidence generation activity would be modified by different possible results of the ongoing research. Appendix A shows how this can be assessed quantitatively. In practice, in most instances the impact of ongoing research on the value of evidence generation activities is likely to be modest unless the ongoing research has very similar aims and scope to the evidence generation activity under evaluation.

6.2.4 The degree to which uncertainty is expected to be resolved by the proposed evidence generation activities

No evidence generation activity can hope to resolve all uncertainty about a quantity of interest. Uncertainties will remain as any evidence generation activity will have a finite sample size and, therefore, when combined with the available prior information will not provide a single definitive estimate of the quantity of interest. Uncertainties will also always remain due to the possibility of biases and imperfect exchangeability between the study setting and the setting in which the research findings are to be implemented.

[3] Though it is also possible that accounting for ongoing research could increase the value of the evidence generation activity under evaluation.

The quantitative summaries above have assumed that the evidence generation activity will resolve all uncertainty and, therefore, the results represent an expected upper bound on the value of additional research. It is useful to consider the situations in which this may/may not represent a good approximation to the value of a new study (Figure 6). If prior information is relatively weak (i.e. we are highly uncertain about the quantity of interest) and the study planned is both large and expected to provide a reliable estimate of the quantity for the context of interest, then the quantitative estimates shown above are likely to represent reasonable approximations to the value of the study. If prior information is relatively weak but a smaller, less definitive study is likely to be conducted then the estimates presented above will overestimate the value of the evidence generation activity. This doesn't mean that the study would not be worthwhile since a smaller-scale cheap study could deliver high value (i.e. a low $/DALY averted). If prior information is strong (i.e. we are not particularly uncertain about the quantity of interest) then the estimates above may not provide a good approximation to the value of the evidence generation activity. However, in this situation the research is unlikely to be worthwhile and evidence generation expenditures should be focused elsewhere. Methods are available to estimate the value of evidence generation activities with different sample sizes, and recent advances have enabled these calculations to be run using the outputs of a PSA [29,42,43] (see Rabideau et al (2017) [44] for an application of these methods in the context of the Cost-Effectiveness of Preventing AIDS Complications (CEPAC) HIV model).

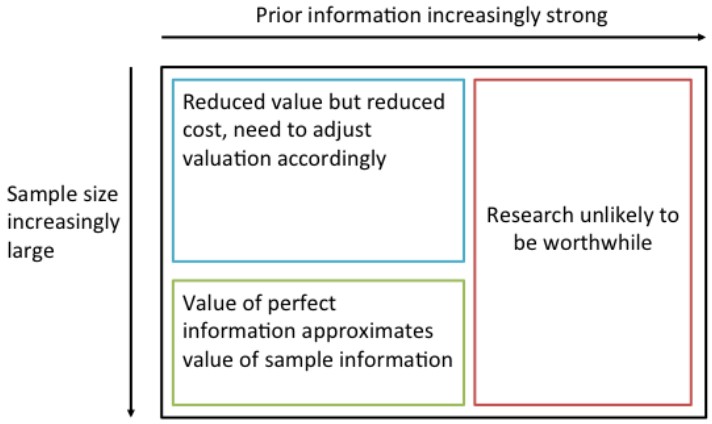


Figure 6: Impact of prior information and sample size on the value of information

It is also possible to estimate the maximum study sample size that could prove to represent a cost-effective use of research resources. This can be calculated by using EVPPI estimates to understand the maximum value of the research in terms of DALYs averted, alongside information on the fixed costs of the study and the costs of enrolling each individual in the study (sometimes referred to as the marginal sampling cost of the study) and their health opportunity costs. The maximum sample size can be calculated by subtracting the health opportunity costs associated with the fixed research costs (e.g. fixed trial costs) from the global net health effect associated with the study, and dividing the resulting amount by the health opportunity cost per individual enrolled in the study. This maximum sample provides an indication of whether a study design in its current form could potentially be of value which requires that the proposed sample size is lower than the maximum sample size.

6.2.5 A base case for the value of evidence generation activities

A "base case" analysis is often presented for a cost-effectiveness study. This represents a set of judgements with respect to the most plausible model structure and input quantity values. A similar concept can be applied to estimating the value of evidence generation activities. This involves

identifying the most plausible values for the populations who can benefit from the improved information, the time horizon over which the evidence will be used to inform decision-making, the timing and likelihood of the research, whether any future events are considered to be important, and whether there is a need to adjust the value of the evidence generation activity to reflect the extent to which it will be able to resolve uncertainty. An example set of base case judgements are shown in Table 8 and the corresponding implications for the value of evidence generation activities shown in Figure 7. These judgements may be subject to sensitivity analysis to assess the impact of alternative plausible judgements on the value of evidence generation.

Table 8: Base case judgements relating to evidence generation activities

Attribute	RCT	Survey
Probability of successful completion of research	80%	60%
Timing of reporting	At year 4	At year 2
Future uncertainties	Price change at year 4: 50% chance of 60% price drop	
Likely reduction in uncertainty	Weak prior information and large study planned, use EVPPI	Weak prior information and large study planned, use EVPPI

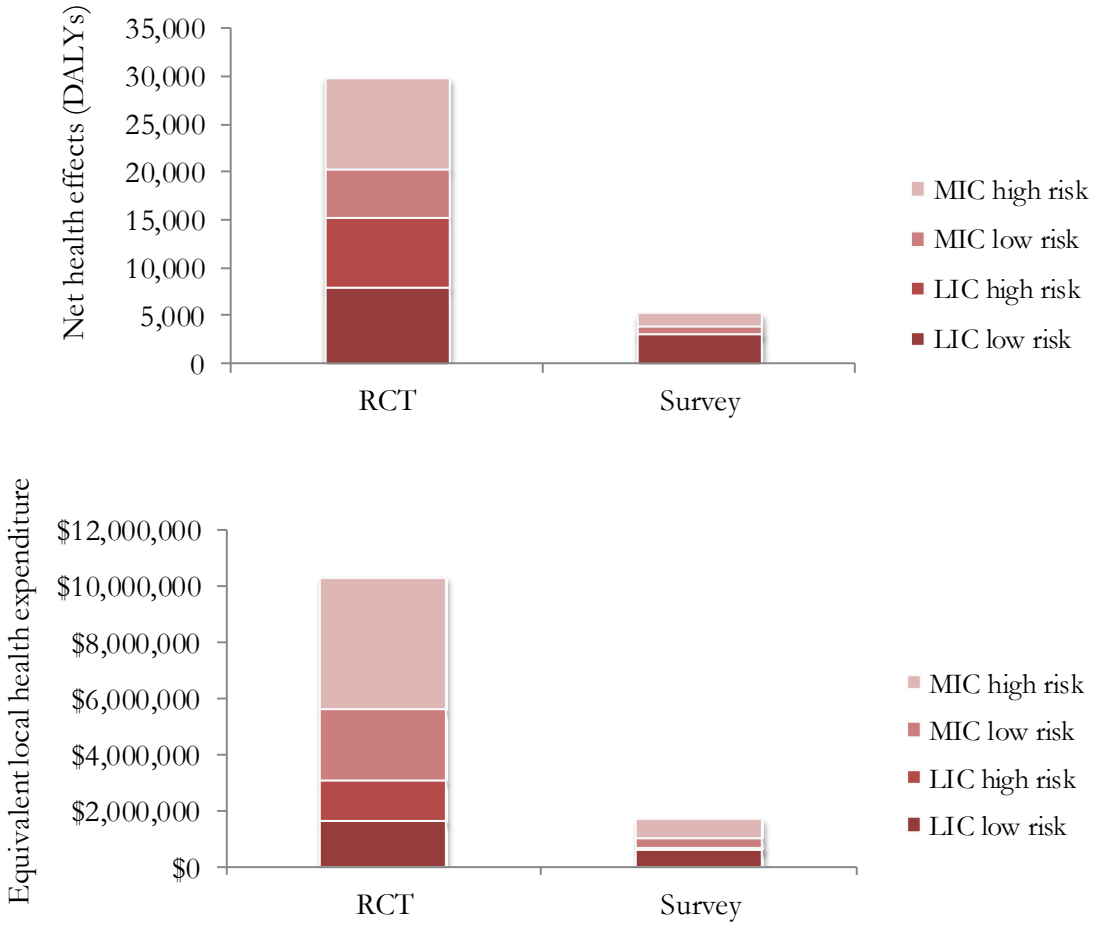


Figure 7: Value of evidence generation activities based on base case judgements

6.3 Theme B: What kind of research would best support decision-making?

6.3.1 Consider the value of alternative study designs

Different study designs may allow a comparison of different interventions, and may provide data on different biochemical, care-related or economic endpoints. All of these differences in study design imply that information will be collected on different quantities, and this data collection may profoundly impact both the value of the evidence generation activity and its costs. When more than one quantity that informs a decision could be informed by an evidence generation activity, the value of the evidence generation activity is not simply the sum of the value of reducing uncertainty around each quantity [45]. Instead, it is the value of the combined improved information about the quantities. In some cases, this joint value may be much greater than the sum of the individual values. This would be the case if the decision only changed when both quantities took extreme values (e.g., the decision only changed if a service was found to be both cheaper to provide and offered higher adherence rates than expected). It is also possible that the joint value will be much lower than the sum of the individual values. This would be the case if the resolution of uncertainty in one of the quantities meant that the other quantity had little bearing on the decision. The new study may also inform other quantities that might not have been considered in the original model. The same considerations apply regardless of whether the study includes an additional intervention arm or provides information on additional endpoints. Methods for estimating the value of different research designs are described in Appendix B.

6.3.2 Consider the value of alternative programmes of research

Where multiple studies could be commissioned to inform a decision, different types of research programmes can be envisaged which may involve commissioning one study, multiple studies simultaneously, a sequential approach, whereby decisions about one or more evidence generation activities are only made once the results of earlier studies become available, or a watchful waiting approach, whereby a decision maker plans to wait until other uncertainties have resolved before deciding which studies to commission.

The value of different research programmes will depend on: (1) the joint value of the information across the research programmes, which will not simply be a sum of the value of the individual studies; (2) the timing of when different studies will report and the implications of this for the time horizon over which these studies are expected to contribute to improved decision-making; and (3) whether delaying studies via sequential or watchful waiting approaches can ensure that studies are only funded when they are most valuable. In general, different research programmes will imply different trade-offs. Figure 8 shows five different research programmes that could be commissioned in the context of the illustrative example. This shows that upfront commissioning of the RCT, survey, or both studies, allows the benefits of these studies to be realised over 8 years for the survey and 6 years for the RCT. Although the sequential studies offer the potential to avoid commissioning a second study where it would not be worthwhile, there are opportunity costs associated with delaying the initiation and, therefore, reporting of the second study. If the RCT is commissioned first, the survey will only inform decision-making over 4 years. Similar trade-offs need to be considered when adopting a watchful waiting approach.

Sequential research programmes are anticipated to be most valuable when the first study will report quickly and when it is likely that the results of the first study could impact upon the decision about whether to conduct the second study. Similarly, a period of watchful waiting prior to commissioning research is likely to deliver most value if the uncertain event is expected to occur early and modify the decision about whether to commission research. Methods for formally evaluating different research programmes are presented in Appendix C.

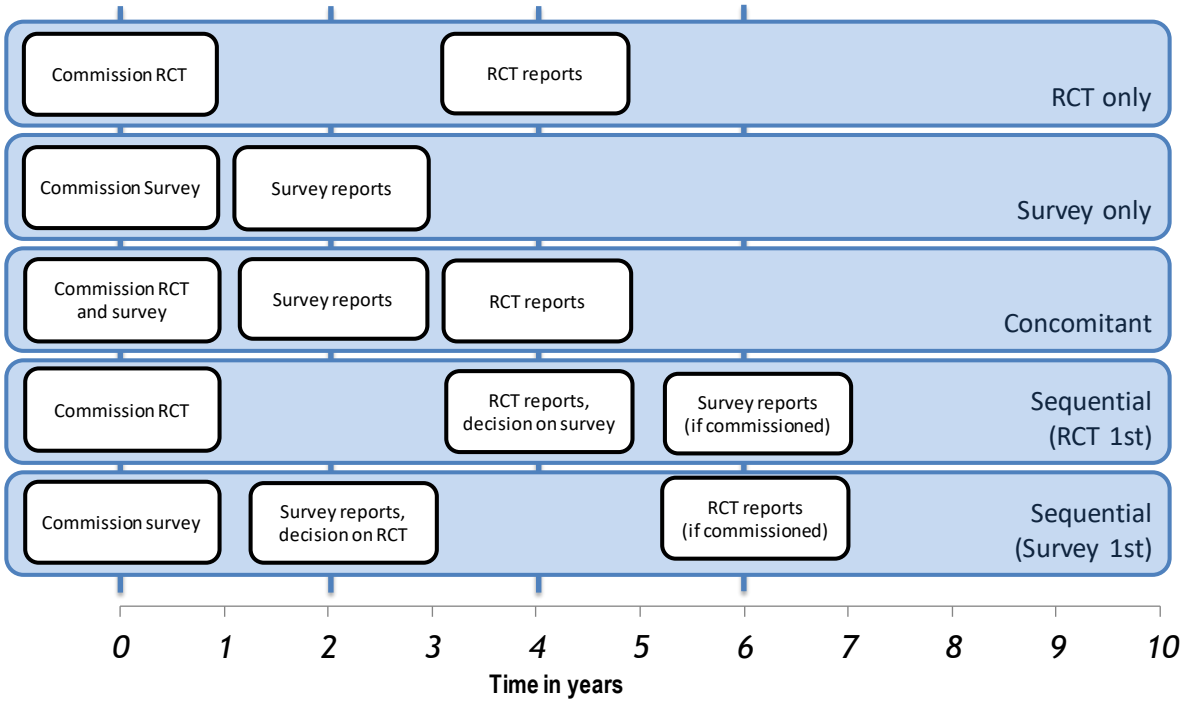


Figure 8: Implications of alternative research designs for the time horizon over which research can inform decision-making

The value of alternative research designs is presented in Table 9 for the illustrative example. The preferred research design will depend upon the cost of the RCT and survey, and the opportunity costs of the research expenditure. However, some general points are worth noting. The sequential design that involves commissioning the survey first is unlikely to be considered a valuable proposition. This design reduced the global net health effect delivered by the research, due to the delay in reporting from year 4, if the RCT is commissioned now, to year 6 if it is commissioned following results of the survey. The design is not accompanied by a reduction in costs, since regardless of the survey results, the RCT is almost always commissioned (probability of commissioning = 99%) as it remains of high value for almost all possible outcomes of the survey. The sequential design, which involves commissioning the RCT first, generates considerably more value than commissioning the RCT alone, but only requires the survey to be run (and it's costs incurred) with a probability of 60%. Further consideration of the research costs and the distribution of DALYs is likely to be necessary in order to inform decision-making with respect to this research design.

Table 9: Value of alternative research designs and the probability that research will be commissioned

	DALYs averted via improved information	Equivalent health care expenditure ($)	Probability RCT commissioned	Probability survey commissioned
RCT only	20,827	$9,644,617	100%	0%
Survey only	3,731	$1,865,417	0%	100%
Concomitant	24,927	$11,242,156	100%	100%
Sequential (RCT 1st)*	23,033	$10,436,662	100%	60%
Sequential (Survey 1st)*	16,555	$7,781,150	99%	100%

* It is necessary to stipulate a "decision rule" regarding commissioning of the second research phase when evaluating the sequential research designs. Here it is assumed that: (1) the survey will only be commissioned if the additional DALYs generated would have cost at least $500,000 to generate locally via direct health care financing; and (2) the RCT will only be commissioned if the additional DALYs generated would have cost at least $2,000,000 to generate locally. **Analysis uses base case settings outlined in Table 8 but assuming no price change.

6.4 Theme C: Are there additional opportunity costs associated with evidence generation?

The analysis presented above assumes that all policy options are available to decision makers (adopt, reject, OIR, AWR). The value of evidence generation activities reflect the added value of research assuming that the intervention will not be implemented whilst research is conducted in the LIC low-risk, LIC high-risk and MIC low-risk groups, i.e., by comparing the net health effect of an OIR policy to those generated by a reject policy. For the MIC high-risk group, the value of evidence generation activities reflect the added value of research, assuming that the intervention will be implemented whilst research is conducted and then potentially continued or withdrawn, depending upon the results of the research, i.e., by comparing the net health effect of an AWR policy to an adopt policy. Parallel service implementation and research may be impractical for the reasons described previously (e.g., difficulty in recruiting to an RCT once an intervention is widely available, ethical concerns, difficulty in withdrawing an intervention when research reports and contamination of survey results). If these considerations rule out an AWR policy, then the health implications of not implementing a cost-effective technology whilst research takes place needs to be accounted for when evaluating the evidence generation activity.

If research prohibits adoption, then the value of the evidence generation activity may be reduced markedly. For example, if commissioning the RCT means that the intervention cannot be adopted in the MIC high-risk subpopulation until the research reports, then this subpopulation cannot accrue the health benefits of adopting the intervention during the 4 years research takes to conduct and report (Figure 9). The greater the expected net health effect of the intervention, and the longer a study will take to report, the larger the reduction in the value of commissioning the RCT. If RCT evidence is considered generalizable, it may be feasible to avoid this loss of health by running the study in one of the other subpopulations whilst implementing the service within the MIC high-risk group. This would avoid the opportunity costs associated with delaying implementation [39].

When assessing the opportunity costs associated with research policies, it is also important to consider the time profile of the net health effect of service investments. For many service investments there is an initial high commitment of resources. This may occur due to high upfront costs associated with setting up a service (e.g., costs of setting up or reconfiguring facilities, or expenditure on equipment). If this is the case, an AWR policy is less advantageous because the initial upfront costs may be irrecoverable if the research shows that the service is not cost-effective (i.e., sunk upfront costs are incurred). Even if set-up costs are not a major consideration, a common pattern with interventions in HIV and other long-term conditions is that the costs of rolling out a programme of prevention or treatment occur early, whereas the health benefits for recipients often take much longer to accrue. In these instances, the benefits of an OIR policy will be increased if the intervention can be given once research reports (i.e., research confirms that the intervention is cost-effective).

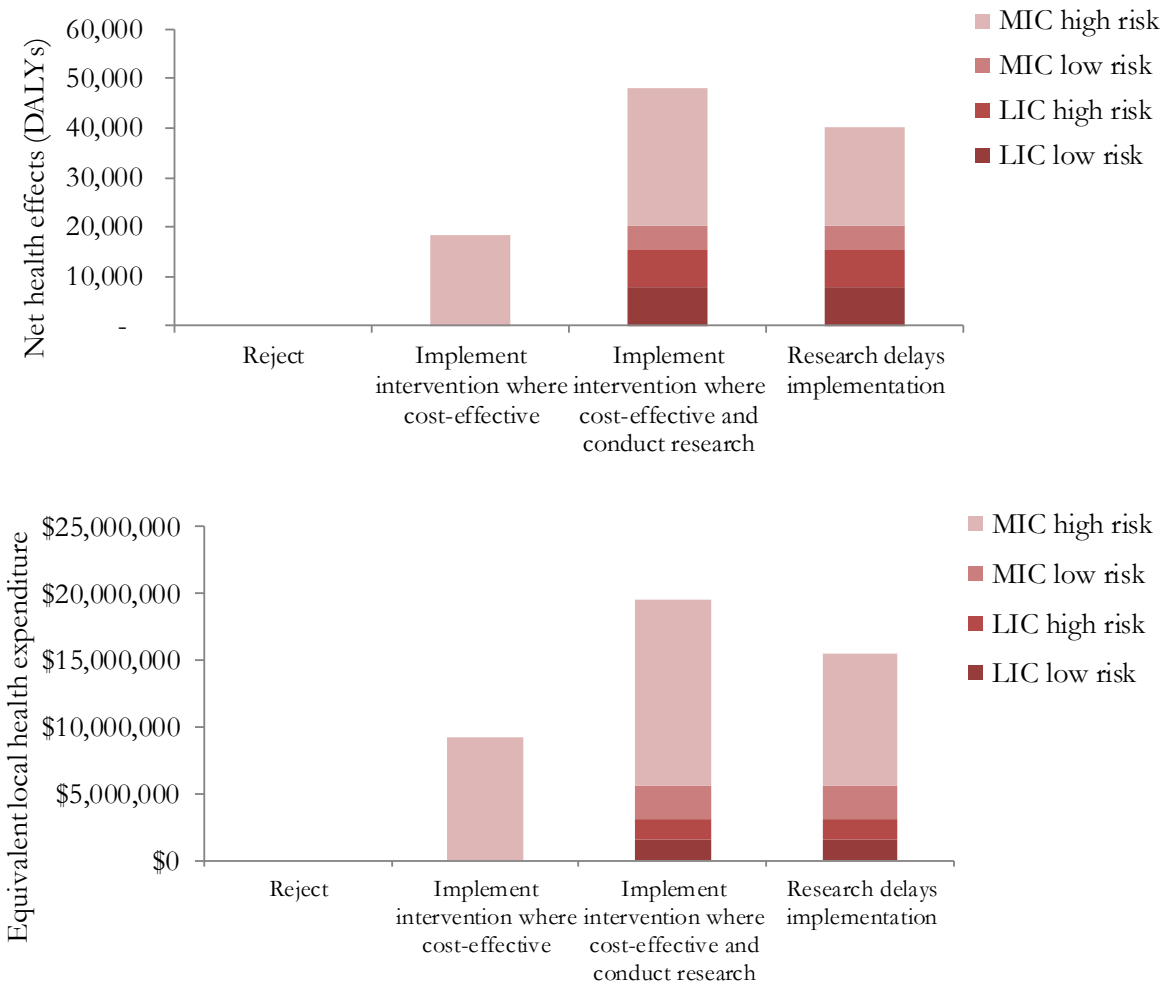


Figure 9: Impact of research delaying implementation on the value of evidence generation

7. Discussion

This report sets out a framework for evaluating investments in evidence generation activities by international research funders. The methods presented have the potential to be used to inform decisions influencing the direction of a huge volume of global health resources. By quantifying the health implications of alternative uses of research resources, the framework represents an important tool for transparent and accountable decision-making. The framework has been illustrated using examples from HIV, although the methods can be applied to any decision relating to investments in evidence generation activities in low-, middle- or high-income countries.

The approach takes into account both the potential health benefits in local populations who will benefit from the evidence, as well as the health opportunity costs of the required research expenditure. The framework shows that the local health effects of investments in evidence generation activities are quantifiable and depend crucially on the scale of the populations who stand to benefit from improved information, the ability of the research study (or programme) to provide information that could change decision-making, and the magnitude of any local research costs. The framework also emphasises the need to take account of the health opportunity costs of international research expenditure, which will depend on the potential alternative uses for those research funds which may include other types of research and depending on the remit of the funder, non-research investments competing for the same resources.

In this report, the assessments required are illustrated using a simple didactic example. One feature of this example is that the costs and outcomes of individuals treated in each year are assumed to be independent of the policy choices in previous or subsequent years. This assumption is likely to be applicable for the majority of non-communicable disease settings and some infectious disease settings; however, for many infectious disease settings this assumption is not considered tenable due to the existence of transmission dynamics which have substantive implications for the impact of policy interventions. Where these disease dynamics are important, the impact of improved evidence for those treated within a given year will depend upon: (i) when the evidence becomes available (as the epidemic changes over time); (ii) whether an OIR or AWR policy was pursued as the initial adoption or rejection of an intervention as this may modify the epidemic; and (iii) the policy choice made beyond the time horizon over which the research is considered valuable. Due to these considerations a more complete consideration of the implications of OIR and AWR policies would account for the full policy trajectory within a dynamic model. This would involve modelling three periods: (i) the period prior to the research reporting in which the decision to adopt or reject the intervention depends on whether an OIR or AWR policy is being evaluated; (ii) the period beyond the research reporting during which the decision is revised in light of the evidence generated; and (iii) the policy choice expected beyond the period over which the research is anticipated to inform decision-making. Further work to demonstrate the implementation of this type of modelling is required.

In assessing the value of evidence generation activities, the focus has been placed on quantifying the benefits of improved information as a way of better understanding the costs and effects of alternative investment choices and, therefore, improving programmatic choices about the implementation of services. A second way in which evidence generation activities can generate value is by ensuring that budget allocations at the regional, programme, and intervention level are closely aligned with the funds required to deliver planned services. The extent to which this will be an important source of value will depend upon the way in which decision makers manage funds and service delivery in response to cost variances. For example, the value of having more robust evidence, particularly in relation to costs, is likely to be higher in a decentralised system with limited transferability of funds between geographical areas and services, than in a more centralised system

where decision makers have the flexibility to manoeuvre funds in response to unfolding events [46]. Further work to develop simple metrics to reflect these considerations when estimating the value of evidence generation activities is warranted.

References

1 The Resource Tracking for HIV Prevention R&D Working Group. Investment priorities to fund innovation in a challenging global health landscape 2000-2016. http://www.ghtcoalition.org/pdf/HIV-Prevention-Research-Development-Funding-Trends-2000-2015.pdf (accessed 17 May 2018).

2 Bill and Melinda Gates Foundation. https://www.gatesfoundation.org/ (accessed 17 May 2018).

3 Department for International Development - GOV.UK. https://www.gov.uk/government/organisations/department-for-international-development (accessed 17 May 2018).

4 European and Developing Countries Clinical Trials Partnership. http://www.edctp.org/ (accessed 17 May 2018).

5 US National Institutes of Health (NIH) | Turning Discovery Into Health. https://www.nih.gov/ (accessed 17 May 2018).

6 Medical Research Council. https://mrc.ukri.org/ (accessed 17 May 2018).

7 Wellcome Trust. https://wellcome.ac.uk/ (accessed 18 May 2018).

8 Fogarty International Center. https://www.fic.nih.gov/Pages/Default.aspx (accessed 18 May 2018).

9 Population Council. http://www.popcouncil.org/ (accessed 18 May 2018).

10 PEPFAR: Accelerating Progress Toward HIV/AIDS Epidemic Control. https://www.pepfar.gov/ (accessed 18 May 2018).

11 U.S. Agency for International Development. https://www.usaid.gov/ (accessed 18 May 2018).

12 Centers for Disease Control and Prevention. https://www.cdc.gov/ (accessed 18 May 2018).

13 Neumann PJ, Anderson JE, Panzer AD, *et al.* Comparing the cost-per-QALYs gained and cost-per-DALYs averted literatures. *Gates Open Res* 2018;**2**:5. doi:10.12688/gatesopenres.12786.1

14 WHO | World Health Organization. http://www.who.int/ (accessed 18 May 2018).

15 World Bank Group - International Development, Poverty, and Sustainability. http://www.worldbank.org/ (accessed 18 May 2018).

16 Ochalek J, Lomas J, Claxton K. Cost Per DALY Averted Thresholds for Low-and Middle-Income Countries: Evidence From Cross Country Data. *Cent Heal Econ Res Pap 122, Univ York* Published Online First: 2015.https://www.york.ac.uk/media/che/documents/papers/researchpapers/CHERP122_cost_DALY_LMIC_threshold.pdf (accessed 17 May 2018).

17 Woods B, Revill P, Sculpher M, *et al.* Country-Level Cost-Effectiveness Thresholds: Initial Estimates and the Need for Further Research. *Value Heal* 2016;**19**:929–35. doi:10.1016/J.JVAL.2016.02.017

18 Claxton K, Ochalek J, Revill P, *et al.* Informing Decisions in Global Health: Cost Per DALY Thresholds and Health Opportunity Costs. 2016. https://www.york.ac.uk/media/che/documents/policybriefing/Cost per DALY thresholds.pdf (accessed 17 May 2018).

19 Briggs AH, Claxton K, Sculpher MJ. *Decision Modelling for Health Economic Evaluation*. Oxford University Press 2006.

20 Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999;**18**:341–64. http://www.ncbi.nlm.nih.gov/pubmed/10537899 (accessed 17 May 2018).

21 Bojke L, Grigore B, Jankovic D, *et al.* Informing Reimbursement Decisions Using Cost-Effectiveness Modelling: A Guide to the Process of Generating Elicited Priors to Capture Model Uncertainties. *Pharmacoeconomics* 2017;**35**:867–77. doi:10.1007/s40273-017-0525-1

22 Soares MO, Dumville JC, Ashby RL, *et al.* Methods to Assess Cost-Effectiveness and Value of Further Research When Data Are Sparse. *Med Decis Mak* 2013;**33**:415–36. doi:10.1177/0272989X12451058

23 Oostenbrink JB, Al MJ, Oppe M, *et al.* Expected Value of Perfect Information: An Empirical Example of Reducing Decision Uncertainty by Conducting Additional Research. *Value Heal* 2008;**11**:1070–80. doi:10.1111/j.1524-4733.2008.00389.x

24 Brennan A, Kharroubi S, O'Hagan A, *et al.* Calculating Partial Expected Value of Perfect Information via Monte Carlo Sampling Algorithms. *Med Decis Mak* 2007;**27**:448–70. doi:10.1177/0272989X07302555

25 Oakley JE. Decision-Theoretic Sensitivity Analysis for Complex Computer Models. *Technometrics* 2009;**51**:121–9. doi:10.1198/TECH.2009.0014

26 Oakley JE. Modelling with Deterministic Computer Models. In: *Simplicity, Complexity and Modelling*. Chichester, UK: : John Wiley & Sons, Ltd 2011. 51–67. doi:10.1002/9781119951445.ch4

27 Rojnik K, Naveršnik K. Gaussian Process Metamodeling in Bayesian Value of Information Analysis: A Case of the Complex Health Economic Model for Breast Cancer Screening. *Value Heal* 2008;**11**:240–50. doi:10.1111/j.1524-4733.2007.00244.x

28 Stevenson MD, Oakley J, Chilcott JB. Gaussian Process Modeling in Conjunction with Individual Patient Simulation Modeling: A Case Study Describing the Calculation of Cost-Effectiveness Ratios for the Treatment of Established Osteoporosis. *Med Decis Mak* 2004;**24**:89–100. doi:10.1177/0272989X03261561

29 Strong M, Oakley JE, Brennan A, *et al.* Estimating the Expected Value of Sample Information Using the Probabilistic Sensitivity Analysis Sample. *Med Decis Mak* 2015;**35**:570–83. doi:10.1177/0272989X15575286

30 Heath A, Manolopoulou I, Baio G. A Review of Methods for Analysis of the Expected Value of Information. *Med Decis Mak* 2017;**37**:747–58. doi:10.1177/0272989X17697692

31 Vanni T, Karnon J, Madan J, *et al.* Calibrating Models in Economic Evaluation. *Pharmacoeconomics* 2011;**29**:35–49. doi:10.2165/11584600-000000000-00000

32 Menzies NA, Soeteman DI, Pandya A, *et al.* Bayesian Methods for Calibrating Health Policy Models: A Tutorial. *Pharmacoeconomics* 2017;**35**:613–24. doi:10.1007/s40273-017-0494-4

33 Claxton K, Palmer S, Longworth L, *et al.* Informing a decision framework for when NICE should recommend the use of health technologies only in the context of an appropriately designed programme of evidence development. *Health Technol Assess (Rockv)* 2012;**16**:1–323. doi:10.3310/hta16460

34 Cambiano V, Mavedzenge SN, Phillips A. Modelling the Potential Population Impact and Cost-Effectiveness of Self-Testing for HIV: Evaluation of Data Requirements. *AIDS Behav* 2014;**18**:450–8. doi:10.1007/s10461-014-0824-x

35 Smith JA, Sharma M, Levin C, *et al.* Cost-effectiveness of community-based strategies to strengthen the continuum of HIV care in rural South Africa: a health economic modelling analysis. *Lancet HIV* 2015;**2**:e159–68. doi:10.1016/S2352-3018(15)00016-8

36 Bassett I V., Govindasamy D, Erlwanger AS, *et al.* Mobile HIV Screening in Cape Town, South Africa: Clinical Impact, Cost and Cost-Effectiveness. *PLoS One* 2014;**9**:e85197. doi:10.1371/journal.pone.0085197

37 Sharma M, Farquhar C, Ying R, *et al.* Modeling the Cost-Effectiveness of Home-Based HIV Testing and Education (HOPE) for Pregnant Women and Their Male Partners in Nyanza Province, Kenya. *JAIDS J Acquir Immune Defic Syndr* 2016;**72**:S174–80. doi:10.1097/QAI.0000000000001057

38 Drummond M, Barbieri M, Cook J, *et al.* Transferability of Economic Evaluations Across Jurisdictions: ISPOR Good Research Practices Task Force Report. *Value Heal* 2009;**12**:409–18. doi:10.1111/j.1524-4733.2008.00489.x

39 Eckermann S, Willan AR. Globally optimal trial design for local decision-making. *Health Econ* 2009;**18**:203–16. doi:10.1002/hec.1353
40 Novak K. The WTO's balancing act. *J Clin Invest* 2003;**112**:1269–73. doi:10.1172/JCI20177
41 Barnabas R V, Revill P, Tan N, *et al.* Cost-effectiveness of routine viral load monitoring in low- and middle-income countries: a systematic review. *J Int AIDS Soc* 2017;**20**:e25006. doi:10.1002/jia2.25006
42 Heath A, Manolopoulou I, Baio G. Efficient Monte Carlo Estimation of the Expected Value of Sample Information Using Moment Matching. *Med Decis Mak* 2018;**38**:163–73. doi:10.1177/0272989X17738515
43 Jalal H, Alarid-Escudero F. A Gaussian Approximation Approach for Value of Information Analysis. *Med Decis Mak* 2018;**38**:174–88. doi:10.1177/0272989X17715627
44 Rabideau DJ, Pei PP, Walensky RP, *et al.* Implementing Generalized Additive Models to Estimate the Expected Value of Sample Information in a Microsimulation Model: Results of Three Case Studies. *Med Decis Mak* 2018;**38**:189–99. doi:10.1177/0272989X17732973
45 Griffin S, Welton NJ, Claxton K. Exploring the Research Decision Space: The Expected Value of Information for Sequential Research Designs. *Med Decis Mak* 2010;**30**:155–62. doi:10.1177/0272989X09344746
46 Woods B, Rothery C, Anderson S-J, *et al.* Appraising the value of evidence generation activities: An HIV Modelling Study. *Submiss to BMJ Glob Heal* 2018.

Appendix A: Methods for assessing the impact of uncertain future events on the value of evidence generation activities

The implications of future events for the value of evidence generation activities will depend upon the extent to which the future events are likely to modify the per-period value of the evidence generation activity, when the events are expected to occur and their likelihood of occurring. The sections below show how to reflect different types of uncertain event when calculating the per-period value of the evidence generation activity. If we call the original per-period value of the evidence generation activity VOI (no future change) and the value reflecting the future change VOI (future change) then the overall value of the evidence generation activity is:

$$VOI(no\ future\ change) \cdot (t) + VOI(no\ future\ change) \cdot (T - t) \cdot (1 - p) + VOI(future\ change) \cdot (T - t) \cdot p$$

Where *T* is the time horizon over which the evidence could potentially inform decision-making, *t* is the time point at which the future change is likely to occur, and *p* is the probability that the future change does occur. Discounting should also be applied appropriately but is not shown here for simplicity.

Appendix A.1 Assessing the impact of a price change

The impact of a potential future price change can be quantified by modifying the net health effect curve to reflect the new price. An example of the implications of a 60% price reduction for the value of an RCT in the LIC high-risk population is shown in Figure 10. The price reduction results in the intervention becoming cost-effective on expectation, increases the trigger point value for the odds ratio at which the decision would change from 0.68 to 0.88, increases the error probability from 0.124 to 0.248, and the health benefits of the RCT (from 5,622 DALYs averted to 22,724 DALYs averted). The future price change is also uncertain in that it may or may not occur. Therefore, the value of evidence generation, with and without the price change, should be weighted to reflect a judgement about the likelihood that the price change will and will not occur.

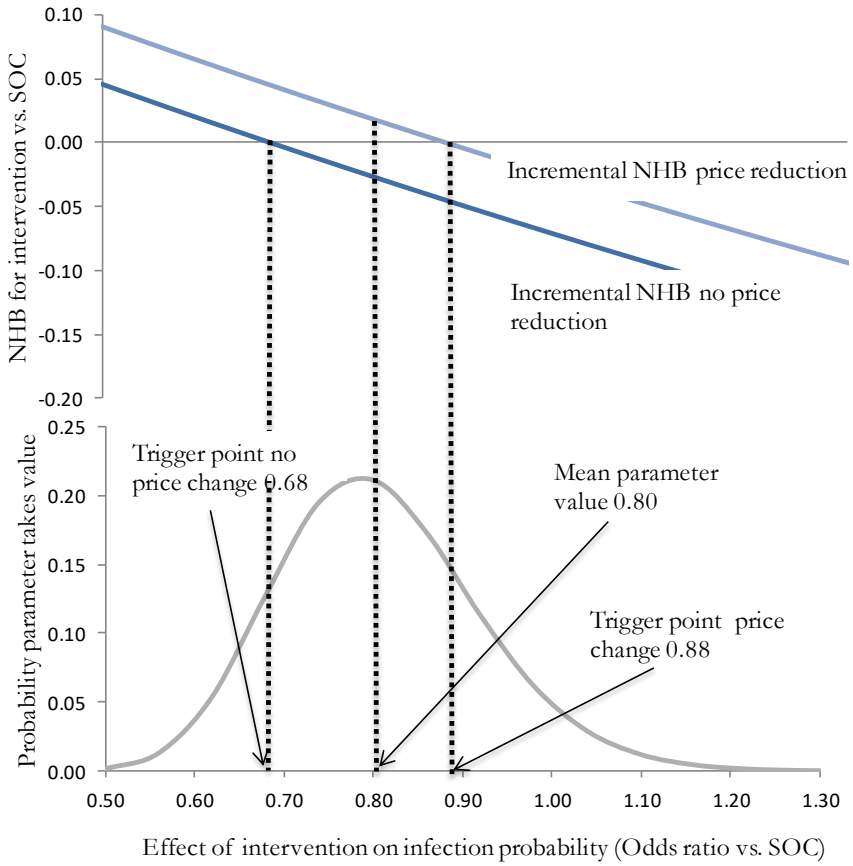


Figure 10: Impact of price change on value of evidence generation: value of RCT in LIC high-risk population

Appendix A.2 Assessing the impact of the entrance of a new technology

The impact of the entrance of a new technology can be quantified by understanding the net health effect curve for the new technology and how it changes with the quantity of interest. An example of the implications of a new technology for the value of an RCT (comparing the original intervention under evaluation to SOC) is shown in Figure 11 for the MIC low-risk population. The new comparator is expected to be cost-effective as indicated by the positive net health effect at the mean value of the quantity. The trigger point is shifted left from 0.75 to 0.69 since the original intervention has to be even more effective to offer a higher net health effect than the new technology, the error probability reduces from 0.34 to 0.14, and the health benefits of the RCT also reduce from 11,405 DALYs averted to 3,583 DALYs averted. Again, if the entry of the new technology is uncertain, the value of evidence generation, with and without the new entrant, should be calculated and weighted to reflect a judgement about the likelihood that the new entrant will become a viable programmatic choice. This shows that innovation can reduce the probability that an innovation will be cost-effective and, therefore, reduce the value of further research on this comparator. This illustration also demonstrates that if the new intervention is currently a relevant comparator it should be included when estimating the value of the evidence generation activity. Omission of relevant comparators will bias estimates of the value of further research.

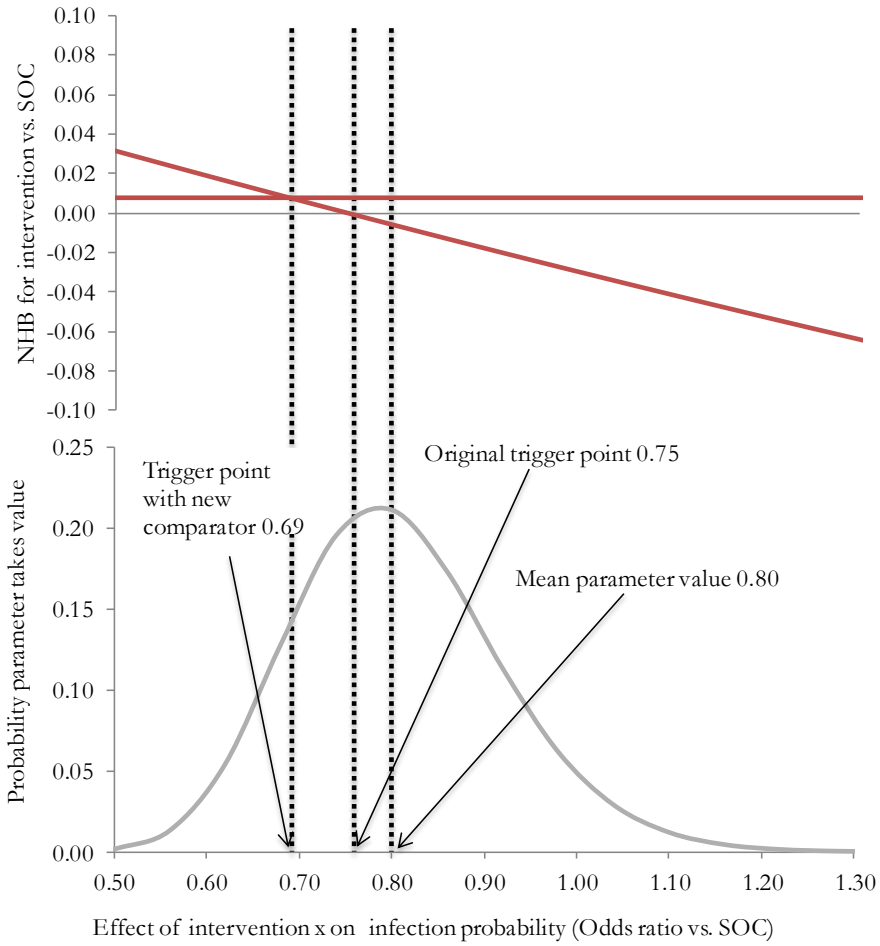


Figure 11: Impact of new technology on value of evidence generation: value of RCT in MIC low-risk population

Appendix A.3 Assessing the impact of ongoing research

It is slightly more challenging to quantify the implications of other ongoing research reporting for the value of an evidence generation activity. This requires computation of the value of the evidence generation activity of interest for each possible outcome of the ongoing research. For example, consider the implications of improved evidence on the DALYs associated with infection (due to improved morbidity or mortality data) for the value of the RCT in the MIC low-risk population. The value of the RCT can be estimated using the methods described in the main text but varying the DALYs associated with infection across the plausible range for this quantity (i.e., the range of DALY values we anticipate that the ongoing research could feasibly generate). This can be used to generate the upper panel of Figure 12. This panel shows that the value of the RCT increases as the DALYs per infection increases up to approximately 1.40 DALYs lost per infection and then decreases beyond this point. As the DALYs lost through infection increase, the decision about implementing the intervention becomes more uncertain, and so the value of understanding intervention effectiveness via the RCT increases. Eventually, when the DALYs lost through infection become very high, it becomes less likely that the RCT will produce a result that would not support implementation of the intervention and the value of the RCT therefore diminishes. As we don't know what the results of the ongoing research will be, the value of the RCT as shown in the upper panel of Figure 12 needs to be weighted to reflect the likelihood of different results from the ongoing research, as shown by the probability distribution in the lower panel of Figure 12. The DALYs averted by the RCT once ongoing research is accounted for are 12,669 compared to 11,405 in the absence of additional information on DALYs due to infection. The difference between these values is driven by the non-linearity in the

curve shown in the upper panel of Figure 12. If this curve is broadly linear, then the ongoing research will only have a small impact on the value of the evidence generation activity, as seen here.

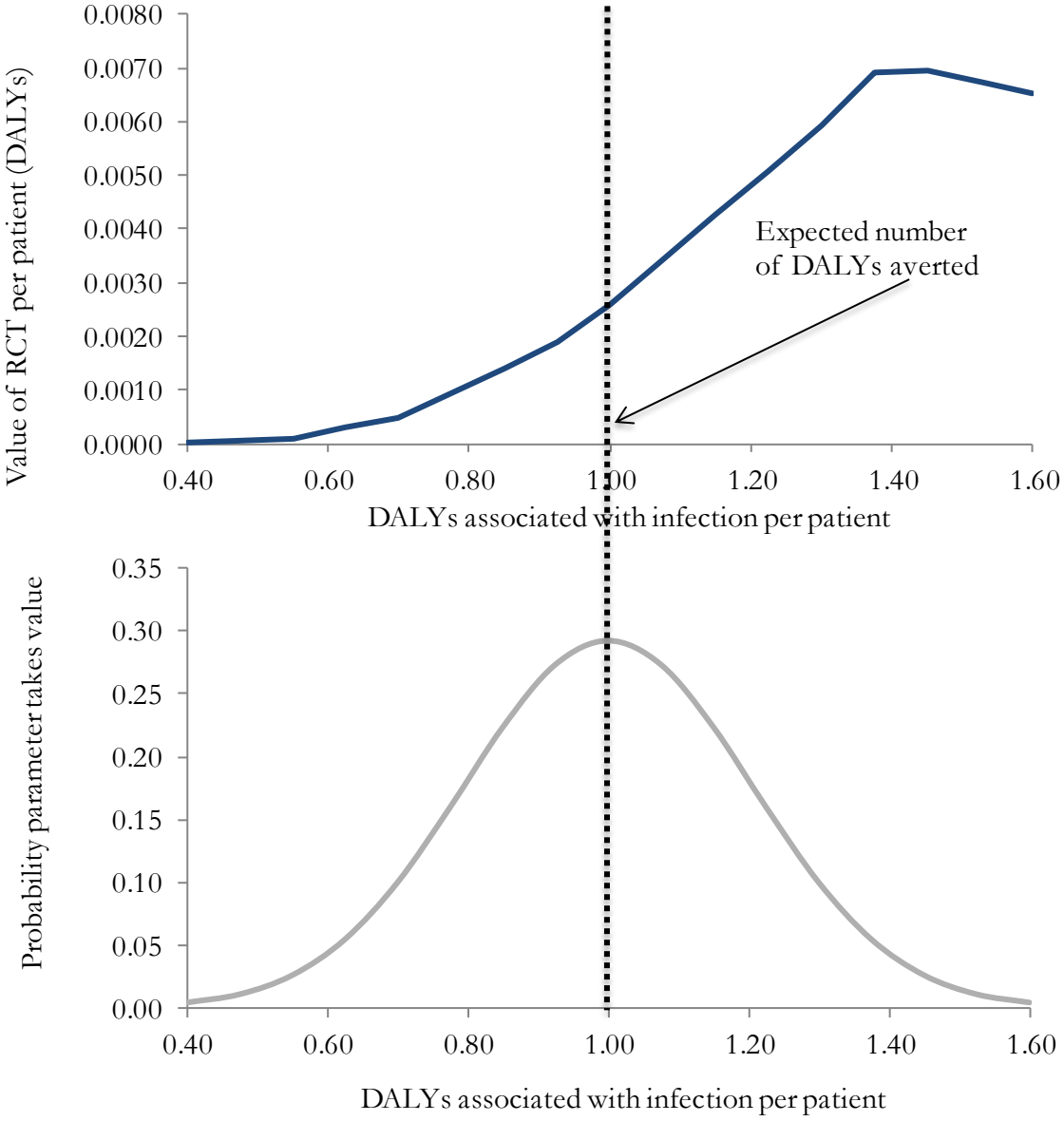


Figure 12: Impact of ongoing research on value of evidence generation: value of RCT in MIC low-risk population

Appendix B: Methods for assessing the value of alternative research designs

Many evidence generation activities could be designed to collect data on one, two or more quantities. The choice of which data to collect has implications for research design, the value of the research, and the costs of research.

To estimate the value of these alternative designs, the process described in Section 3 can be expanded to consider the value of improved information on multiple quantities. If we consider two quantities, then this requires an estimate of the net health effect of each intervention conditional upon combinations of feasible values of both quantities. The trigger point now becomes a trigger boundary, i.e., a set of combinations of the quantities at which the decision would change from the decision selected at the mean value of each parameter. The value of a decision change can then be weighted by the joint likelihood of both quantities taking values beyond the trigger boundary.

This can be illustrated by considering a three-arm trial which can inform two quantities in the model: the relative effectiveness of two of the comparators, each compared to a single comparator that is considered to represent the "baseline" intervention. For the purposes of this illustration, we extend the calculations above to a situation in which there is a third comparator (new intervention) which is cheaper ($5 per treated individual) but less effective than the original intervention considered (odds ratio for intervention 2 compared to SOC: 0.85 (95% CI: 0.77, 0.94)). This comparator is the "new intervention" illustrated in Figure 11 and is cost-effective on expectation. The value of the RCT in this context will therefore be realised if it provides estimates of effectiveness that suggest that the new intervention is not cost-effective, and that the original intervention or standard of care is cost-effective. This is shown in the upper panel of Figure 13. The outlined cell shows that at the expected value of the two quantities the new intervention has the highest net health effect. At points beyond the trigger boundary (shown by the dashed line) the decision switches either to the original intervention or to SOC depending on the results of the research. The value of improved information is calculated by weighting the gain in net health effect associated with the change in decision by its likelihood. For example, if the trial result found an odds ratio of 0.60 for the intervention and 0.80 for the new intervention the decision would switch to the original intervention. This would result in a gain in net health effect per individual of 0.010 DALYs averted (0.024 minus 0.014) which would be weighted by the probability of this outcome, which as shown in the lower panel of Figure 13 is 0.004 (0.013 multiplied by 0.0293).

			Effectiveness of new intervention															
			0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00	1.05	1.10	1.15	1.20	1.25	1.30	1.35
			Net health effect of new intervention																
			0.052	0.044	0.036	0.029	0.021	0.014	0.007	-0.001	-0.008	-0.015	-0.022	-0.028	-0.035	-0.042	-0.048	-0.055	-0.061
Effectiveness of intervention	0.55	Net health effect of intervention	0.032	New	New	New	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int
	0.60		0.024	New	New	New	New	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int
	0.65		0.016	New	New	New	New	New	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int
	0.70		0.009	New	New	New	New	New	New	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int
	0.75		0.001	New	New	New	New	New	New	New	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int
	0.80		-0.006	New	New	New	New	New	New	New	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC
	0.85		-0.013	New	New	New	New	New	New	New	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC
	0.90		-0.021	New	New	New	New	New	New	New	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC
	0.95		-0.028	New	New	New	New	New	New	New	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC
	1.00		-0.035	New	New	New	New	New	New	New	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC
	1.05		-0.042	New	New	New	New	New	New	New	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC
	1.10		-0.048	New	New	New	New	New	New	New	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC
	1.15		-0.055	New	New	New	New	New	New	New	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC
	1.20		-0.062	New	New	New	New	New	New	New	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC
	1.25		-0.068	New	New	New	New	New	New	New	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC
	1.30		-0.075	New	New	New	New	New	New	New	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC
	1.35		-0.081	New	New	New	New	New	New	New	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC	SOC

			Effectiveness of new intervention															
			0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00	1.05	1.10	1.15	1.20	1.25	1.30	1.35
			Probability quantity takes this value																
			0.000	0.000	0.000	0.000	0.024	0.293	0.504	0.166	0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Effectiveness of intervention	0.55	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.60	0.013	0.000	0.000	0.000	0.000	0.000	0.004	0.007	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.65	0.054	0.000	0.000	0.000	0.000	0.001	0.016	0.027	0.009	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.70	0.128	0.000	0.000	0.000	0.000	0.003	0.037	0.065	0.021	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.75	0.195	0.000	0.000	0.000	0.000	0.005	0.057	0.099	0.032	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.80	0.211	0.000	0.000	0.000	0.000	0.005	0.062	0.107	0.035	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.85	0.173	0.000	0.000	0.000	0.000	0.004	0.051	0.087	0.029	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.90	0.113	0.000	0.000	0.000	0.000	0.003	0.033	0.057	0.019	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.95	0.062	0.000	0.000	0.000	0.000	0.001	0.018	0.031	0.010	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1.00	0.029	0.000	0.000	0.000	0.000	0.001	0.009	0.015	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1.05	0.012	0.000	0.000	0.000	0.000	0.000	0.004	0.006	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1.10	0.005	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1.15	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1.20	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1.25	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1.30	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1.35	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Figure 13: Evaluating a three-arm research design: (a) Choice of intervention according to effectiveness of both interventions; and (b) probability of different trial results

This example raises the wider question of the value of RCT designs with different permutations of study arms. The main text focuses on comparing the intervention to SOC, and shows how this can be extended to the three-arm trial cases. However, it is also appropriate to consider the value of a trial comparing the new intervention to SOC, and a trial comparing the original intervention to the new intervention (i.e., there are four possible comparator sets for the RCT). The methods described in the main text can be applied to evaluate an RCT comparing the new intervention to SOC. A trial comparing the new intervention to the existing intervention will generate improved information on the odds ratio comparing the new intervention to the original intervention. Resolving the uncertainty around this quantity will reduce the uncertainty around the two quantities in the model (the odds ratios comparing the original intervention and new intervention to SOC) but it will not resolve it completely. To estimate the value of resolving this uncertainty, we require an estimate of the net health effect of the original and new intervention for each plausible value of the odds ratio comparing the new intervention to the original intervention. We propose that this is calculated by inputting the expected mean of the odds ratios comparing the intervention to SOC and the new intervention to SOC conditional upon the odds ratio comparing the new intervention to the original intervention. These expected means can be calculated by assuming multivariate normality on the log-odds ratio (LOR) scale. The computations are as follows replacing new intervention with the label Y and the original intervention with the label X:

$$E(LOR_{Xvs.SOC}|LOR_{Xvs.Y} = lor_{Xvs.Y}) = E(LOR_{Xvs.SOC}) + \frac{cov(LOR_{Xvs.SOC}, LOR_{Xvs.Y})}{var(LOR_{Xvs.Y})}(lor_{Xvs.Y} - E(LOR_{Xvs.Y}))$$

$$E(LOR_{Yvs.SOC}|LOR_{Xvs.Y} = lor_{Xvs.Y}) = E(LOR_{Yvs.SOC}) + \frac{cov(LOR_{Yvs.SOC}, LOR_{Xvs.Y})}{var(LOR_{Xvs.Y})}(lor_{Xvs.Y} - E(LOR_{Xvs.Y}))$$

Where the treatment effect estimates have been obtained from a mixed treatment comparison or network meta-analysis, the required covariances can be obtained directly from that analysis. Where the treatment effect estimates have been obtained from individual trials the covariances are:

$$cov(LOR_{Xvs.SOC}, LOR_{Xvs.Y}) = var(LOR_{Xvs.SOC})$$
$$cov(LOR_{Yvs.SOC}, LOR_{Xvs.Y}) = -var(LOR_{Yvs.SOC})$$

These estimates can then be plugged in for each value of the odds ratio comparing the new intervention to the original intervention in order to calculate the net health effect of each intervention.

When there are more than two quantities of interest the problem becomes more difficult to visualise but a similar process can be adopted as shown in Table 10. We assume here that the new intervention is cost-effective on expectation. In order to calculate the value of the evidence generation activity the values in column [G] are weighted by the probability that quantities 1-3 will take the values in columns [A-C] i.e. the probability in column [D].

Table 10: Extending the valuation of evidence generation activities to designs that will inform three or more quantities that are key for decision-making

Quantity 1 [A]	Quantity 2 [B]	Quantity 3 [C]	Joint probability of quantities taking values [D]	Incremental net health effect of new intervention conditional upon values of quantities 1-3 [E]	Incremental net health effect of original intervention conditional upon values of quantities 1-3 [F]	Value of improved information [G]
1	1	1	0.01	0.03	0.00	0.00
1	1	2	0.02	0.03	0.02	0.00
1	1	3	0.01	0.03	0.05	0.02
1	2	1	0.02	0.05	0.02	0.00
1	2	2	0.03	0.05	0.04	0.00
.	.	.	.	.	.	.
.	.	.	.	.	.	.
10	5	3	0.01	0.08	0.09	0.01

Appendix C: Methods for comparing research programmes that include multiple evidence generation activities

Where multiple studies could be commissioned to inform a decision, different types of research programmes can be evaluated. In order to evaluate a programme of multiple research studies, the joint benefits of the different types of evidence that will be generated need to be estimated. For example, in order to evaluate the value of the concomitant programme shown in Figure 8, the value of having survey information alone for years 2-4 would need to be added to the value of having both survey and RCT data from years 4-10 (with appropriate discounting). The value of having both survey and RCT data can be calculated using the methods described in Appendix B. Although this focuses on the value of a single study containing multiple quantities, exactly the same principles apply to multiple studies collecting data on one or more quantities. The options available to decision makers are actually somewhat broader than the summary shown in Figure 8. The concomitant design could comprise a decision maker running the RCT (the results of which are considered to generalise across populations) and a survey in one, two, three or all four of the populations. The concomitant design, therefore, actually comprises 20 alternative designs. As the value of the survey in each subpopulation is independent of whether surveys have been conducted in other subpopulations, there is no need to consider the value of each design separately. Instead, the additional local net health effect of having RCT and survey data (compared to having RCT data alone) in each subpopulation can be assessed, and compared to the opportunity costs of the additional international research funds required to extend the survey to that population. If the local net health effect outweighs the health opportunity costs of the international research funds for a specific subpopulation, then the subpopulation should be included within the survey.

For sequential research designs, such as those outlined in Figure 8, additional analyses are required. We first consider a sequential design whereby the RCT is commissioned first, and then depending on the results of the RCT a decision is made with respect to whether to commission the survey. From years 4-6 the value of improved evidence from the RCT only is accrued. For years 6-10 the additional expected value and costs of the survey must be considered, taking into account the fact that for some results of the RCT the survey will not be considered worthwhile in some or all of the populations. This process is shown in Figure 14. For each plausible result of the RCT (simplified here to three possible results), the local net health effect are estimated. The total value of the survey across populations can then be calculated by aggregating local net health effect where these are positive (as in the main text we assume that the survey would not be conducted where it would impose net health loss). An assessment is then made as to whether the net DALYs averted are sufficient to offset the opportunity costs of the international research funds. When the RCT results are such that the the global net DALYs averted by the survey exceed the health opportunity costs of the international research funds, the survey is commissioned and its value accrued. For RCT results where this is not the case the survey is not commissioned and no value (or costs) accrue. The resulting values can then be weighted to reflect the likelihood of different RCT results and, therefore, the expected additional value of the survey.

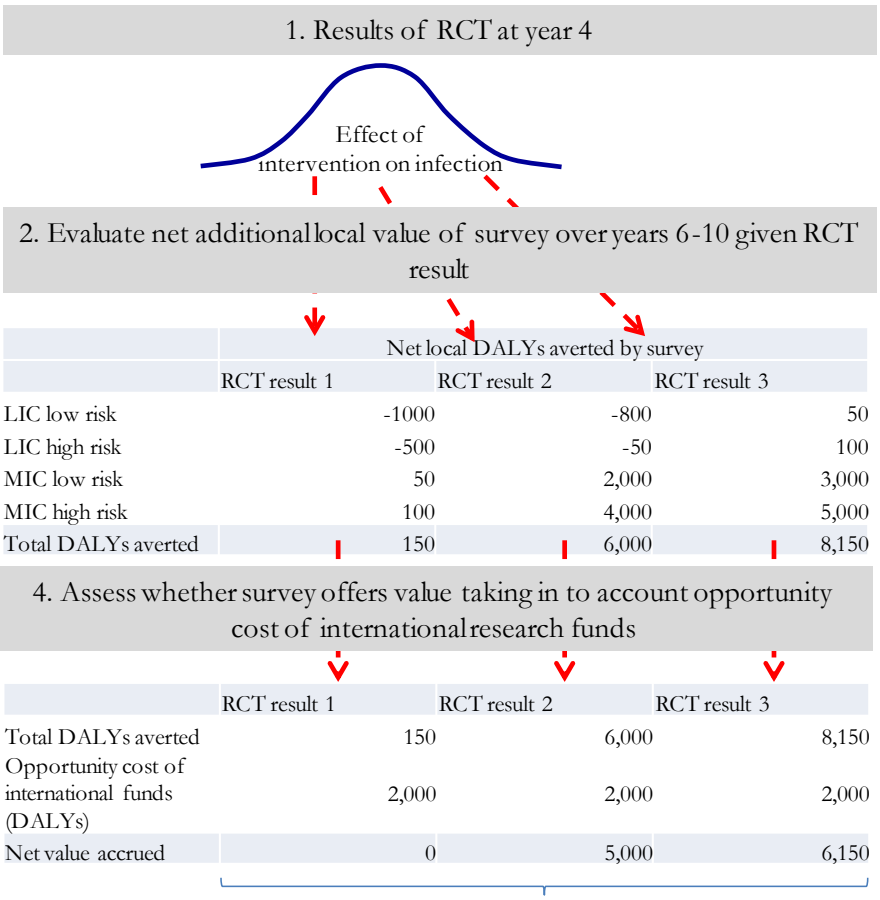


Figure 14: Estimating the value of sequential research designs: RCT followed by survey

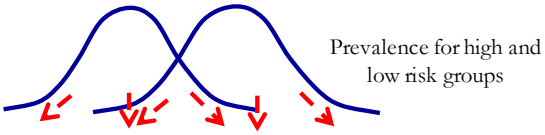
Note: For simplicity, this schematic assumes that the international research funding is the same regardless of the subpopulations included in the survey. More complex cost functions can, however, be accommodated by the approach.

If there is a possibility that the RCT will not report, some assumptions need to be made about what would happen to the commissioning of the survey in this instance. For example, it may be reasonable to assume that the survey decision would be made based on currently available evidence on the relative effectiveness of the intervention. In this case, the value from years 4-10 would reflect the value of the sequential design, weighted to reflect the likelihood that the RCT does report, and the value of conducting the survey (assuming it has a positive net value taking into account the opportunity costs of international research funds) weighted to reflect the likelihood that the RCT does not report.

For a sequential design, whereby the survey is commissioned first and depending on the results of the survey a decision is made with respect to whether to commission the RCT, the process is outlined in Figure 15. Again, there are potentially a large number of sequential strategies as the survey could be run in one, two, three or all four populations, and the value of the subsequent RCT will depend upon the where the surveys were conducted. A pragmatic approach to this issue is to assume that decisions about where to run the surveys could be based on the value of the survey independent of the RCT. This will provide information on the value of this research programme, though may mean that another research programme that provided more value is missed. The main differences compared to Figure 14 are that the results of the survey are independent across populations, so each survey result actually represents a joint set of survey results across populations. A second difference is that due to the generalisability of the RCT evidence, the RCT need only be run in the population in which it imposes the lowest DALY opportunity cost.

The same principles can be used to evaluate a watchful waiting strategy, but instead of an initial research phase there is a costless evidence generation process which provides information on the occurrence and nature of future uncertain events.

1. Results of surveys at year 2



2. Evaluate net additional local value of RCT over years 6-10 given survey results

	DALYs averted by RCT			Local DALY cost of RCT
	Survey results 1	Survey results 2	Survey results 3	
LIC low risk	10	50	1,000	100
LIC high risk	5,000	1,000	1,000	125
MIC low risk	15,000	5,000	10,000	80
MIC high risk	25,000	30,000	20,000	90
Total DALYs averted	**44,390**	**35,970**	**31,920**	

4. Assess whether RCT offers value taking in to account opportunity cost of international research funds

DALYs averted	**Survey results 1**	**Survey results 2**	**Survey results 3**
Total DALYs averted	44,390	35,970	31,920
Opportunity cost of international funds (DALYs)	6,000	6,000	6,000
Net value accrued	38,930	29,970	25,920

Expected value of RCT = 31,607

Figure 15: Estimating the value of sequential research designs: survey followed by RCT