



Bill and Melinda Gates Foundation  
Methods for Economic Evaluation Project (MEEP)

**Final Report**

*A partnership between  
Bill and Melinda Gates Foundation, NICE International,  
the Health Intervention and Technology Assessment Program (Thailand), and  
the University of York, Centre for Health Economics*

January 2014

## Process for developing this report

This report was developed through a collaborative partnership commissioned by the Bill and Melinda Gates Foundation and led by NICE International.

Tommy Wilkinson (NICE International) was MEEP project lead and coordinating author of this report. Kalipso Chalkidou (NICE International) oversaw the MEEP project and contributed to all sections of the report. Damian Walker (BMGF) provided overall MEEP project direction and contributed to all sections of the report. Ruth Lopert (George Washington University) provided a technical and editorial review to all sections of the report.

Yot Teerawattananon, Varit Chantarastapornchit, Benjarin Santatiwongchai, Kittiphong Thiboonboon, and Waranya Rattanavipapong of the Health Interventions and Technology Assessment Program (HITAP), Thailand, authored Section One of the report with input from Tommy Wilkinson and Damian Walker.

Tommy Wilkinson and Kalipso Chalkidou (NICE International), and John Cairns (London School of Hygiene and Tropical Medicine) authored Section Two, with inputs from Tony Culyer (Centre for Health Economics, University of York; Institute of Health Policy, Management and Evaluation, University of Toronto) and Amanda Glassman (Centre for Global Development).

Sections Two, Three (the Gates-RC) and Four (Project Recommendations) were informed by the MEEP project workshop held at the Bill and Melinda Gates Foundation headquarters in Seattle in June 2013 and subsequent consultation with workshop attendees (see over for the list of workshop participants, **Appendix 1** for the workshop report and **Appendix 2** for the workshop agenda).

Section Three (the Gates-RC) was drafted by Karl Claxton, Paul Revill and Mark Sculpher (University of York), Tommy Wilkinson (NICE International), John Cairns (London School of Hygiene and Tropical Medicine) and Andrew Briggs (University of Glasgow). The section also benefitted from substantive comments from the following people (in alphabetical order):

- Kalipso Chalkidou - Director, NICE International
- Tony Culyer - Centre for Health Economics, University of York; Institute of Health Policy, Management and Evaluation, University of Toronto
- Ruth Faden - Wagley Professor and Director, Johns Hopkins Berman Institute of Bioethics
- Marthe Gold - Logan Professor of Community Health and Social Medicine, City College NY
- Barbara Jauregui - Technical Officer, Pan American Health Organization
- Kjell Arne Johansson - Associate Professor, Bioethics, Department of Global and Public Health and Primary Care, University of Bergen
- Carol Levin - Clinical Associate Professor, Disease Control Priorities Network, Department of Global Health, University of Washington
- Ruth Lopert - Adjunct Professor, Dept. of Health Policy, George Washington University
- Francis Ruiz - Senior Advisor, NICE International
- Peter Smith - Professor, Imperial College London
- Yot Teerawattananon - Program Leader & Senior Researcher, Health Intervention and Technology Assessment Program, Thailand
- Anna Vassall - Senior Lecturer, Department of Global Health and Development, London School of Hygiene & Tropical Medicine
- Damian Walker - Senior Program Officer, Integrated Delivery, BMGF
- Simon Walker - Research Fellow, University of York

**MEEP Workshop Attendees**  
24<sup>th</sup> and 25<sup>th</sup> June 2013, Seattle

<b>Amanda Adler</b>	Consultant physician Addenbrook's Hospital, Cambridge & Chair of NICE Technology Appraisal Committee	<b>Ruth Lopert</b>	Independent consultant and Adjunct Professor, Dept. of Health Policy, George Washington University
<b>Deborah Atherly</b>	Senior Program Officer, Program Appropriate Technology in Health	<b>Stefan Nachuk</b>	Associate Director, Rockefeller Foundation
<b>Rob Baltussen</b>	Coordinator of the Nijmegen International Center for Health Systems Research and Education, Radboud University Nijmegen MC	<b>Arindam Nandi</b>	Fellow, Center for Disease Dynamics, Economics & Policy
<b>Anirban Basu</b>	Associate Professor, University of Washington	<b>Angeline Nanni</b>	Director, Market Access, Aeras
<b>Andrew Briggs</b>	William R Lindsay Chair in Health Economics, University of Glasgow	<b>Susan Nazzaro</b>	Program Officer, Market Dynamics, BMGF, Seattle
<b>John Cairns</b>	Professor of Health Economics, Department of Health Services Research and Policy, London School of Hygiene & Tropical Medicine	<b>Jodi Nelson</b>	Director, Strategy, Measurement and Evaluation, BMGF
<b>Telma Caldeira</b>	Specialist in Regulation and Health Surveillance Brazilian Health Surveillance Agency (ANVISA)	<b>Peter Neumann</b>	Director of the Center for the Evaluation of Value and Risk in Health, Tufts Medical Center
<b>Kalipso Chalkidou</b>	Director, NICE International	<b>Thomas O'Connell</b>	Senior Health Specialist, Economics and Finance, UNICEF, New York
<b>Angela Chang</b>	Summer intern, BMGF & PhD student at Harvard School of Public Health	<b>Mead Over</b>	Senior Fellow, Center for Global Development
<b>Varit Chantarastapornchit</b>	Researcher, Health Intervention and Technology Assessment Program (HITAP)	<b>Paul Revill</b>	Research Fellow, University of York
<b>Karl Claxton (via TVC)</b>	Professor of Economics, University of York	<b>Nalinee Sangrujee</b>	Chief of Health Economics and Finance Team Centers for Disease Control and Prevention
<b>Ruth Faden</b>	Wagley Professor and Director, Johns Hopkins Berman Institute of Bioethics	<b>Benjarin Santatiwongchai</b>	Researcher, Health Intervention and Technology Assessment Program (HITAP)
<b>Emmanuela Gakidou</b>	Associate Professor, University of Washington	<b>Mark Sculpher (via VTC)</b>	Professor of Health Economics & Director of the Programme on Economic Evaluation and Health Technology Assessment, University of York
<b>Geoff Garnett</b>	Deputy Director, HIV, BMGF, Seattle	<b>Priya Sharma</b>	Policy and Innovative Financing Advisor, USAID
<b>Marthe Gold</b>	Logan Professor of Community Health and Social Medicine, City College New York	<b>Donald Shepard</b>	Professor, Brandeis University
<b>Ulla Griffiths</b>	Lecturer, London School of Hygiene & Tropical Medicine	<b>Ritu Shroff</b>	Senior Officer, Strategy, Measurement and Evaluation, BMGF
<b>Timothy Hallett</b>	Reader, Imperial College	<b>Peter Smith</b>	Professor, Imperial College of London
<b>Michael Hanlon</b>	Assistant Professor, University of Washington	<b>Anderson Stanciole</b>	Program Officer, Strategy Planning & Management, BMGF
<b>Kate Harris</b>	Program Officer, HIV Health Economics, BMGF	<b>Maaya Sundaram</b>	Consultant to BMGF
<b>Robert Hart</b>	Managing Director, Epidemiological Modeling Division, Intellectual Ventures Lab	<b>Tessa Tan-Torres Edejer</b>	Coordinator, Health Systems Financing, World Health Organization
<b>Karen Hofman</b>	Professor, University of Witwatersrand School of Public Health	<b>Fabrizio Tediosi</b>	Swiss Tropical and Public Health Institute
<b>Barbara Jauregui</b>	Technical Officer, Pan American Health Organization	<b>Yot Teerawattananon</b>	Program Leader & Senior Researcher, HITAP
<b>Kjell Arne Johansson</b>	Associate Professor, University of Bergen	<b>Kittiphong Thiboonboon</b>	Researcher, Health Intervention and Technology Assessment Program (HITAP)
<b>Iain Jones</b>	Economic Advisor, Department for International Development, UK	<b>Anna Vassall</b>	Senior Lecturer, Department of Global Health & Development, London School of Hygiene & Tropical Medicine
<b>Alexandre Lemgruber</b>	Regional Advisor in Health Technologies, Pan American Health Organization	<b>Damian Walker</b>	Senior Program Officer, Integrated Delivery, BMGF
<b>Carol Levin</b>	Clinical Associate Professor, Disease Control Priorities Network, Department of Global Health University of Washington	<b>Tommy Wilkinson</b>	Advisor, Health Economics, NICE International

## Table of Contents

Process for developing this report.....	2
Tables and Figures .....	6
Abbreviations.....	7
Glossary.....	8
Executive Summary:.....	9
Introduction: Good decisions, better decisions.....	10
Section One: A review of the evidence – many contexts, multiple methods.....	12
1. Introduction .....	12
2. Objectives .....	12
3. Methodology .....	12
3.1 Scope of review .....	12
3.2 Analytical frameworks.....	12
3.3 Search strategies .....	13
4. Results .....	14
4.1 Overview .....	14
4.2 Variations in methodology of cost-per-DALY studies funded by BMGF .....	20
4.3 Quality of reporting.....	22
4.4 Quality of evidence used.....	24
5. Discussion.....	25
5.1 Key findings .....	25
5.2 Implications for research .....	25
5.3 Implications for BMGF .....	26
5.4 Limitations.....	26
Section Two: The case for a Gates Reference Case .....	27
1. Introduction .....	27
2. Setting the scene .....	27
2.1 What is useful?.....	27
2.2 Usefulness to whom?.....	28
2.3 What is a reference case? .....	30
3. The case for a reference case .....	32
3.1 The advantages of a reference case .....	32
3.2 The disadvantages of a reference case .....	32
4. The case for a Gates Reference Case.....	33
4.1 Opportunities and risks .....	33
4.2 Generalisability-applicability.....	35
Section Three: Meeting the need - the Gates Reference Case.....	37
1. Preamble to the reference case .....	37
1.1 What is the Gates Reference Case? .....	37
1.2 Principles and methods.....	39
1.3 What are good decisions? .....	40
1.4 Who are the decision-makers? .....	40

1.5 Comparability – benefits and limits .....	41
1.6 Gates Reference Case Compliance (Gates-RCC) .....	41
1.7 Gates Reference Case Validation (Gates-RCV) .....	42
2. Principles of Economic Evaluation .....	43
2.1 Transparency .....	44
2.2 Comparators .....	44
2.3. Use of Evidence .....	45
2.4. Measure of outcome .....	46
2.5. Measurement of costs .....	47
2.6. Time horizon for costs and effects .....	49
2.7. Costs and Effects outside health .....	50
2.8. Heterogeneity .....	53
2.9 Uncertainty .....	54
2.10. Impact on other constraints and budget impact .....	56
2.11. Equity implications .....	56
3. Gates-Reference Case Summary Table .....	58
Section Four: Recommendations .....	63
References .....	66

## Tables and Figures

### Tables

Table 1 - Variables for review of cost-per-DALY studies.....	13
Table 2 - Number of economic evaluations identified.....	14
Table 3 - Number of identified economic evaluations by type.....	17
Table 4 - Summary of areas of recommendation from WHO, NICE, US Panel, HITAP .....	31
Table 5 - Perceived advantages and disadvantages of Reference Cases.....	34
Table 6 - Examples of constituencies, investment types and decision-makers.....	38
Table 7 - Principles of Economic Evaluation .....	43

### Figures

Figure 1 - Search period of each systematic review of economic evaluation .....	15
Figure 2 - Details of health economic evaluation funders .....	16
Figure 3 - Number of BMGF and non BMGF-funded economic evaluations by disease .....	18
Figure 4 - Cost-effectiveness league chart .....	19
Figure 5 - Number of study perspectives adopted .....	20
Figure 6 -Time horizon and discount rate used by disease/intervention .....	20
Figure 7 - Number of studies adhering to methodological specifications for DALY calculation.....	21
Figure 8 - Number of studies using different decision analytic model types .....	21
Figure 9 - Number of studies conducting each type of sensitivity analysis .....	22
Figure 10 - Percentage adhering to recommended reporting practices .....	22
Figure 11 - Numbers adopting different approaches for currency conversion .....	23
Figure 12 - Rank of evidence used in in BMGF-funded cost effectiveness analyses .....	24
Figure 13 - Optimisation of usefulness measures.....	28
Figure 15 - The Funder Ecosystem.....	30
Figure 16 - The Gates Reference Case .....	38
Figure 17 - Social and scientific values of BMGF and constituencies .....	41

## Abbreviations

<b>BMGF</b>	Bill and Melinda Gates Foundation
<b>BIA</b>	Budget Impact Analysis
<b>CEA</b>	Cost Effectiveness Analysis
<b>CUA</b>	Cost Utility Analysis
<b>CBA</b>	Cost Benefit Analysis
<b>CPI</b>	Consumer Price Index
<b>DALY</b>	Disability-Adjusted Life Year
<b>Gates-RC</b>	The Bill and Melinda Gates Foundation Reference Case
<b>Gates-RCC</b>	Gates Reference Case Compliant
<b>Gates-RCV</b>	Gates Reference Case Validated
<b>ICER</b>	Incremental Cost Effectiveness Ratio
<b>MEEP</b>	Methods for Economic Evaluation Project
<b>NICE</b>	National Institute for Health and Care Excellence
<b>PPP</b>	Purchasing Power Parity
<b>QALY</b>	Quality-Adjusted Life Year

## Glossary

<b>Constituency</b>	The population whose health improvement is represented in the pre-stated normative assumption: “to improve population health subject to a budget constraint”. Constituency includes all people in a particular region, country, or insurance plan.
<b>Context</b>	The circumstances of a decision, including the population (constituency), investment type and political and other constraints
<b>Element</b>	Any part or aspect of the reference case, including both principles and specifications
<b>Funder (of an economic evaluation)</b>	The party that directly financially supports the production of the Economic Evaluation.
<b>Health Technology</b>	Anything that promotes health or prevents ill-health. Includes health programmes, public health policy initiatives, and privately consumed medical interventions, devices and diagnostics
<b>Principle</b>	A type of reference case element. A general statement that aligns with an overall objective to improve population health related quality and length of life that guides the planning, conduct and reporting of an economic evaluation
<b>Reference Case</b>	A standard set of principles and specifications that an analyst should follow in performing cost-effectiveness analysis - adapted from Culyer (2013), Gold (1996).
<b>Scientific Value Judgements (in health)</b>	Views on the significance and relevance of available scientific, technical, and clinical data
<b>Social Value Judgements (in health)</b>	Individual or communal beliefs, needs and aspirations driving different perceptions of appropriate expenditure priorities
<b>Specification (of a principle)</b>	A type of reference case element. The operationalisation of a principle by specifying a numerical value or defined position.
<b>User (of economic evaluations)</b>	The party that makes use of an economic evaluation to inform decisions. Note that in many instances, there will be multiple direct and indirect potential users of an economic evaluation.



## Executive Summary:

- *Priority setting decisions in health are unavoidable*; priority setting has inevitable consequences in terms of health benefits foregone.
- *Priority setting decisions can be optimised* by ensuring decision-makers have clarity about both the costs and benefits of different options. The use of economic evaluation methods can provide a systematic approach to relating costs and benefits and determining *value*.
- The Bill and Melinda Gates Foundation (BMGF) is one of the largest investors in economic evaluation research on health interventions in low and middle-income countries (LMIC), funding 28% of all published cost effectiveness analyses<sup>1</sup> that use the DALY outcome measure in a LMIC setting. However, to date there has been substantial variation in both the approaches and methodologies used, as well as in the quality of BMGF-funded analyses. Ultimately this limits their usefulness to local policy makers, other stakeholders and BMGF itself.
- While *the development and application of a Gates Reference Case (Gates-RC) by BMGF presents both benefits and risks*, the capacity to improve the utility of economic evaluations in informing decision-making supports its development and adoption.
- This project focuses on the methodological aspects of a reference case, but briefly explores how the use of a reference case would reveal the *social and scientific values* held by the BMGF.
- A *key benefit of a Gates-RC would be greater consistency* in the economic evaluation methodology utilised. This would not only be of benefit to BMGF but also to other institutional funders making decisions that will potentially impact a range of countries and constituencies, and would also assist domestic decision-makers by *improving the overall quality and consistency of economic evaluations*.
- The proposed Gates-RC is founded on *eleven key principles* to guide the planning, conduct and reporting of economic evaluations.
- Each of the eleven key principles is supported by certain methodological specifications and reporting standards that together form a comprehensive template for the conduct and presentation of economic evaluations. When this template is adhered to the resulting economic evaluation may be considered *compliant with the Gates-RC*.
- The terms *Gates-RC Compliant (Gates-RCC)* and *Gates-RC Validated (Gates-RCV)* are proposed to denote a) economic evaluations that *comply with the Gates-RC (Gates-RCC)* and b) economic evaluations that have been *externally validated to comply with the Gates-RC (Gates-RCV)*.
- In order to ensure optimal adoption and application of the Gates-RC, it is recommended that BMGF pursue a number of additional initiatives, including the provision of *training in the use of the Gates-RC* and the establishment of a *repository of economic evaluations*.

---

<sup>1</sup> In the vaccine, malaria, TB and HIV/AIDs programme areas

## Introduction: Good decisions, better decisions

*Primum non nocere* – ‘first do no harm’. It is an unfortunate reality that in priority setting there will invariably be opportunity costs associated with even the most carefully considered decisions. Where resources are finite, investment in an intervention that benefits a particular group will inevitably result in health foregone by another. ‘Harm’, in the form of unrealised health gains or untreated disease, is thus unavoidable. Good decisions in health are those that aim to maximise benefits and minimise harms.

Determining the relative importance of benefits and harms is not the only consideration. Decision-making in health is also inherently value-laden, with individual or collective beliefs, needs and aspirations driving different perceptions of appropriate expenditure priorities. Disregard for these values will give rise to priority setting decisions that fail to reflect societal preferences. Decision-makers, whether local or national policy makers, clinicians, institutions, NGOs or global funding bodies therefore need sound evidence of the likely costs and benefits of their decisions, articulated through a prism of societal values. How then is this evidence to be obtained?

Health economic evaluation represents a set of methods designed to identify resources used against the outcomes of alternative health policy options (1) and thereby improve resource allocation decisions by addressing efficiency in healthcare. Over the past three decades, these methods have gained increasing attention from decision-makers in both resource-rich and resource-poor countries, as well as among global donors (2) (3) (4). Importantly, economic evaluation has been identified as crucial for countries to be able to realise the benefits of universal health coverage, by enabling efficient and equitable allocation of health care and other resources (5).

Established in 2000, the mission of the Bill and Melinda Gates Foundation (BMGF) is to improve health and reduce extreme poverty through a wide-ranging series of Global Programmes (6). BMGF supports its grantees and partners in advancing knowledge, developing innovative approaches and providing services in areas beset by particular health problems, including HIV/AIDS, maternal and child health issues, and vaccine-preventable diseases. In recent years, BMGF has begun funding various health economic evaluation studies to improve decision-making by local and national governments as well as global health funders such as the GAVI Alliance (formerly the Global Alliance for Vaccine and Immunization), the Global Fund and BMGF itself. However to date there has been no consistency in methods or presentation of these analyses, and as a result, there is inconsistency in both the quality and utility of the outputs.

In 1996 the Panel on Cost-effectiveness in Health and Medicine convened by the United States Public Health Service (the US Panel) recommended the use of a ‘reference case’ when conducting cost effectiveness analyses (CEAs) (7). The US Panel had identified an outstanding need for comparability and quality improvement in the conduct and reporting of CEAs, and the use of a reference case was proposed to address this. Since then, many other entities including the World Health Organisation and the National Institute for Health and Care Excellence (NICE) in the UK have introduced the concept of a ‘reference case’ to improve the use of economic evaluations in informing decisions. This report presents the case for the adoption of a reference case by the BMGF.

Section One presents a review of economic evaluations in the program areas of HIV/AIDS, malaria, TB, and vaccines in low and middle income countries (LMIC). Distinction is made between BMGF-funded and other economic evaluations, with a focus on studies that use the Disability-Adjusted Life

Year (DALY) outcome measure. The substantial methodological and reporting variation observed provides empirical support for the adoption of a reference case by BMGF.

Section Two identifies the current landscape for economic evaluations in LMICs, and explores the implications of the adoption of a reference case. This section includes discussion of issues that are of particular relevance to global funders such as BMGF.

Section Three presents the proposed reference case in detail. It provides explanation of the principle-led approach taken, explains why each principle is important, provides methodological specifications, and standards for reporting.

Section Four outlines the recommendations of this report that should accompany the reference case if it were to be adopted by BMGF. The recommendations address the risks and benefits of adoption as outlined in Section Two, and suggest ways in which the use of the reference case could be optimised.

## Section One: A review of the evidence – many contexts, multiple methods

### 1. Introduction

Economic evaluation can support sound policy decisions for health technologies only when performed appropriately and reported with clarity and accuracy. Poor methodological quality is often cited as undermining the utility of economic evaluations (8) (9) (10), and this is particularly relevant in low and middle income countries where methodological guidelines for conducting health economic evaluation may not be established or there is a paucity of reliable data sources (10) (11). In addition, researchers may be insufficiently experienced in undertaking economic evaluations, or lack of methodological guidance may mean economic evaluation is inconsistently performed or reported.

This section presents the findings of a review of published economic evaluations, with the aim of identifying opportunities for improving the quality and relevance of economic evaluations at both global and national levels. The findings from this review have been used to inform the development of a reference case for undertaking health economic evaluations funded by BMGF (see Section Three).

### 2. Objectives

The review aimed to assess the quantity and quality of economic evaluations funded by BMGF, by first identifying health economic evaluation studies funded exclusively or in part by BMGF from among all studies undertaken in low and middle-income countries (LMICs), and in particular disease/intervention areas. The review aimed to examine the quality of evidence used and the standard of reporting, and focused on cost utility analyses using the cost per Disability-Adjusted Life Year (DALY) averted outcome measure.

### 3. Methodology

#### 3.1 Scope of review

The review included published economic evaluations undertaken in LMICs from 2000 onwards in programme areas where BMGF provided substantial funding (6). These include HIV/AIDS, malaria, tuberculosis (TB), and vaccines other than for malaria and TB. Country income levels (i.e. low-income, lower-middle income and upper-middle income countries) were determined using World Bank classifications (12).

The initial review identified all types of published economic evaluations (cost-minimization; cost-effectiveness analysis; cost-utility analysis; and cost-benefit analysis (including willingness to pay studies) (1)). The in-depth analysis of variations in methods, quality of evidence used and standard of reporting was limited to those studies reporting costs per disability adjusted life year (DALY). The health interventions assessed were classified into four categories: health promotion/disease prevention, health screening/disease detection, treatments, and social interventions aimed at mitigating health and social impacts arising from diseases/health problems.

#### 3.2 Analytical frameworks

The Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement prepared by a task force of the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) (13) and NICE's methodology checklists (14), were reviewed to inform the development of a comprehensive data extraction form. In addition, the review drew on specific analytical frameworks developed by Walker and

Fox-Rushby (15) and Teerawattananon et al (16) for identifying methodological variation in economic evaluations undertaken in LMICs. Table 1 illustrates the modified review frameworks.

**Table 1 - Variables for review of cost-per-DALY studies**

Aspect	Variable
<b>Variation in methods used</b>	Study perspective: Societal; Health system; Third-party payer; Healthcare provider
	Time horizon
	Discount rate (%) for costs and outcomes
	Types of analytical approach: Static Modelling; Dynamic Modelling; did not use modelling
	Types of sensitivity analysis: Univariate analysis; Probabilistic analysis; Threshold analysis; Others
	Methods of currency conversion: Purchasing power parity (PPP); Exchange rate; Other measures
	Methods for DALY calculation: Standard (Japanese) life table; Application of age weighting; Application of discounting
<b>Quality of evidence</b>	Baseline clinical data
	Clinical effect size
	Resource use and costs
	Utility value
<b>Quality of reporting</b>	Description of interventions compared and justification of choice of comparator(s)
	Description of characteristics of the targeted population
	<ul style="list-style-type: none"> <li>Reporting of the date of estimation of resource quantities and unit costs and description of methods for adjusting estimated unit costs to year of reported costs</li> <li>Description of methods for converting costs into a common currency</li> </ul>
	Reporting of values, ranges, references, and, if used, probability distributions for all parameters
	Reporting of results as (point estimate and confidence interval) of ICER
	Discussion of generalizability/transferability of findings
	Description of how the analysis was funded and the role of the funders
	Description of potential for conflicts of interest of study contributors
<b>Others</b>	Methodological guidelines cited in each study

### 3.3 Search strategies

Searches were undertaken using MEDLINE and the Centre for Reviews and Dissemination databases. The search strategy specified systematic reviews of economic evaluation studies published from 2000 to May 2013. The search terms used are listed in **Appendix 3**. Search results were filtered using ‘systematic review’. Individual economic evaluations were then identified from citations in those systematic reviews, and searched manually to identify individual studies undertaken in LMICs from 2000 onwards.

Lastly, Cooper et al (17) proposed a hierarchy of evidence to assess the appropriateness of data sources used in economic evaluation studies (see **Appendix 4**). The Cooper hierarchy was applied in this analysis.

## 4. Results

### 4.1 Overview

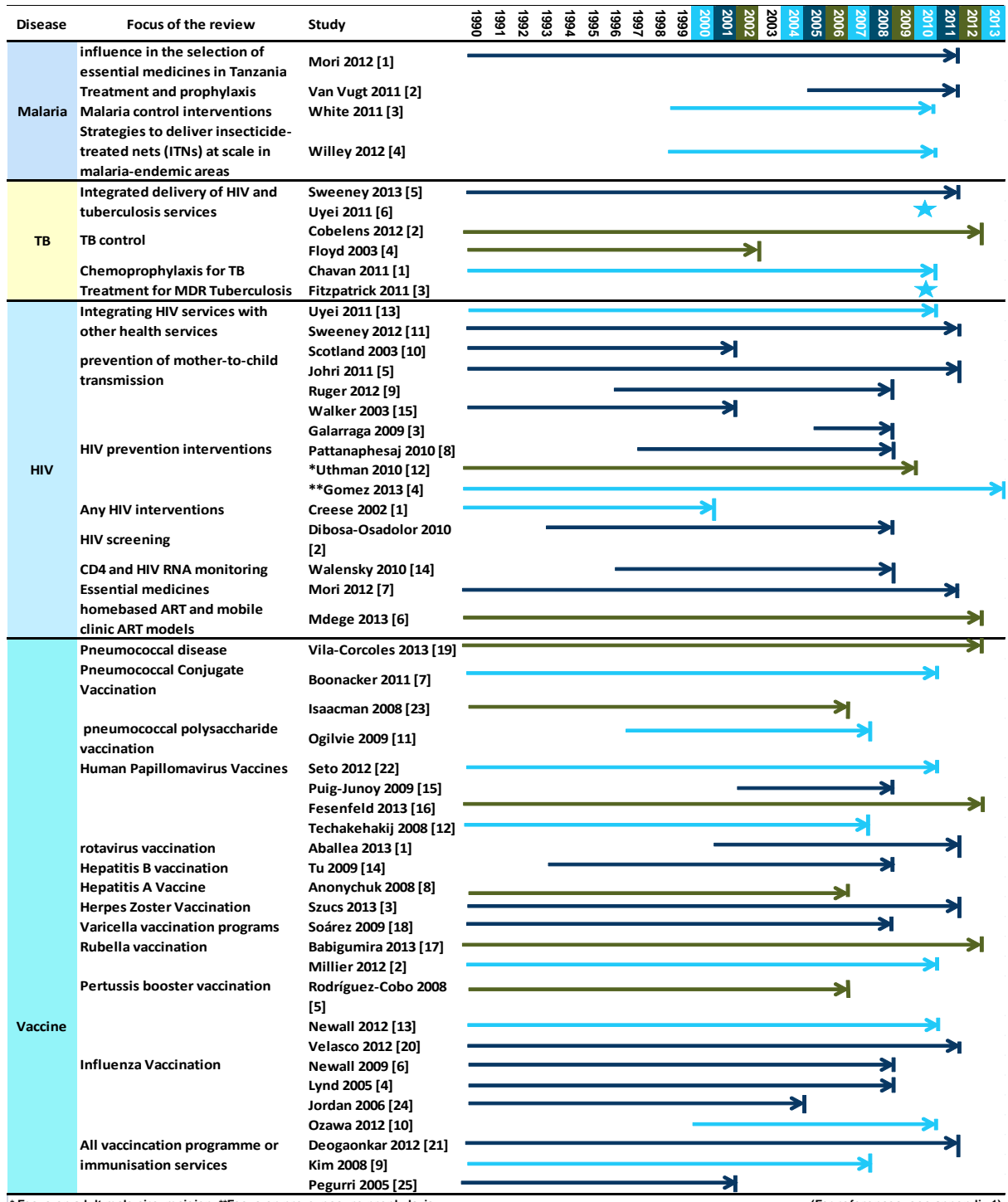
The search identified 504 abstracts, of which 56 were systematic reviews of economic evaluations in low and middle income countries in one of the four programme areas (Table 2 - Number of economic evaluations identified (including cost-per-DALY studies, SR: Systematic Review). From the 56 identified systematic reviews, 224 economic evaluations were found which met the inclusion criteria i.e. undertaken in LMICs, published in or after 2000, and were either cost-minimization analysis, cost-effectiveness analysis; cost-utility analysis, or cost-benefit-analysis. The majority of the economic evaluations were on vaccines (110), followed by HIV/AIDS (58), malaria (41), and TB (15). The lists of included systematic reviews and economic evaluations can be found in **Appendix 5** and **Appendix 6** respectively.

**Table 2 - Number of economic evaluations identified (including cost-per-DALY studies, SR: Systematic Review)**

Programme area	SR abstracts identified	SR matching inclusion criteria	EE in included SRs	EE matching inclusion criteria	Included EE funded by BMGF
Malaria	27	4	166	41	15
<i>using DALY-averted outcome measure</i>				17	9
TB	61	6	419	15	1
<i>using DALY-averted outcome measure</i>				2	1
HIV/AIDS	199	15	350	58	5
<i>using DALY-averted outcome measure</i>				11	1
Vaccines	217	31	899	110	26
<i>using DALY-averted outcome measure</i>				41	9
<b>Total</b>	<b>504</b>	<b>56</b>	<b>1,834</b>	<b>224</b>	<b>47</b>
<i>using DALY-averted outcome measure</i>				<b>71</b>	<b>20</b>

EE – Economic Evaluation; SR – Systematic Review; BMGF - Bill and Melinda Gates Foundation

Figure 1 illustrates the periods covered by the searches utilised by individual systematic reviews. Three systematic reviews did not report the search period used. Most studies did not include economic evaluations published in 2012 and 2013. Topics of the reviews varied largely from cost-effectiveness of strategies for delivery of insecticide-treated nets (18) to cost-effectiveness of interventions related to HIV/AIDS (19). None of the reviews except one (based in Tanzania) focused on a particular setting but included all published literature undertaken in low-, middle-, and high-income countries.



\* Focus on adult male circumcision, \*\*Focus on pre-exposure prophylaxis (For references, see appendix 1)

Figure 1 - Search period of each systematic review of economic evaluation in the areas of malaria, TB, HIV/AIDS, and vaccines★

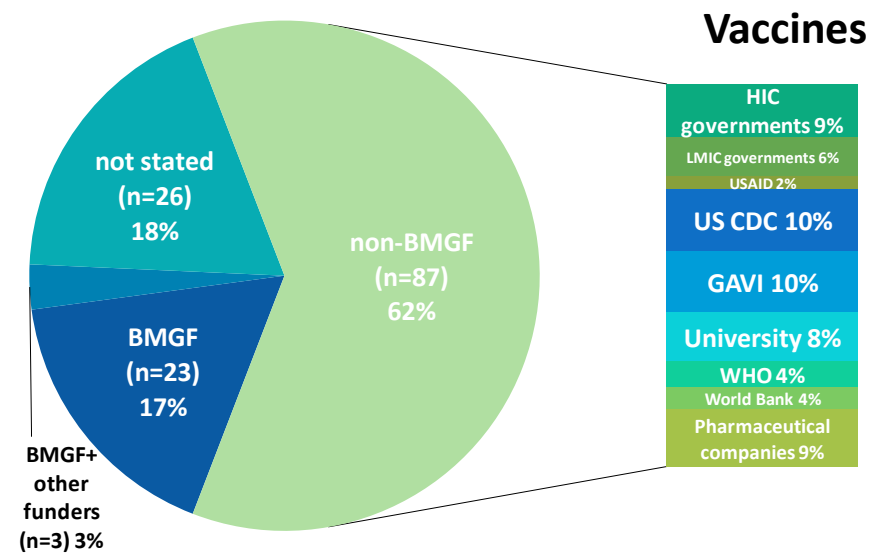
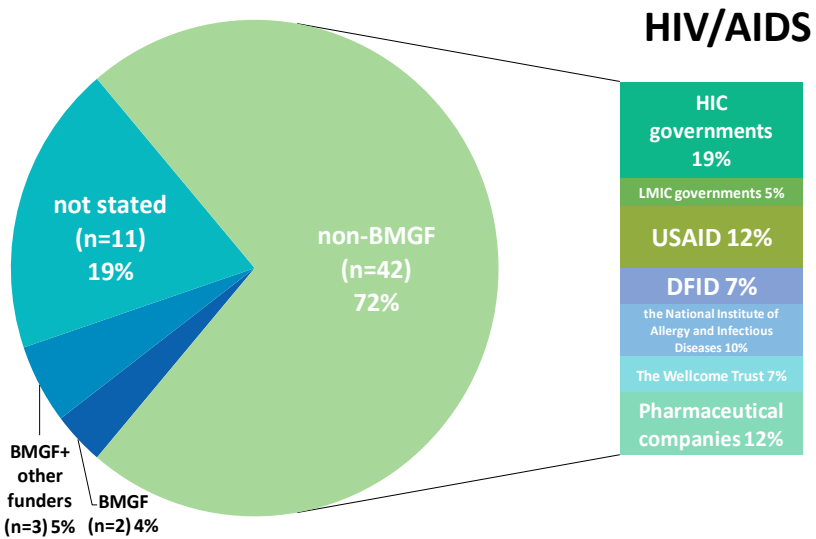
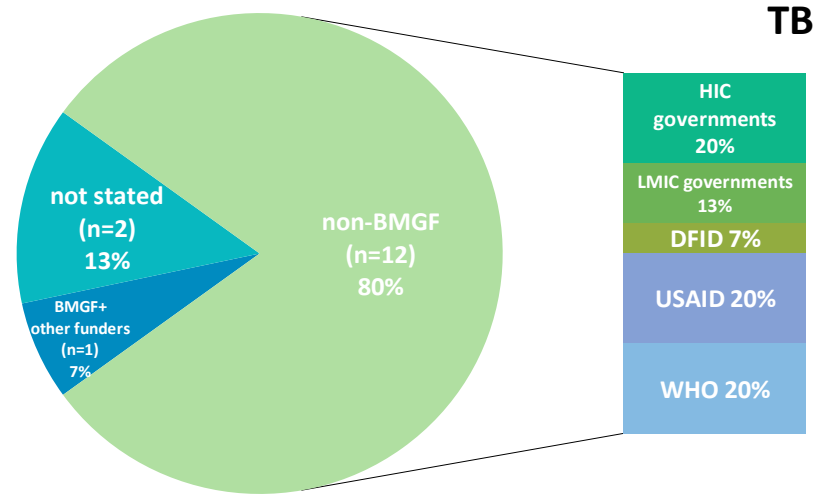
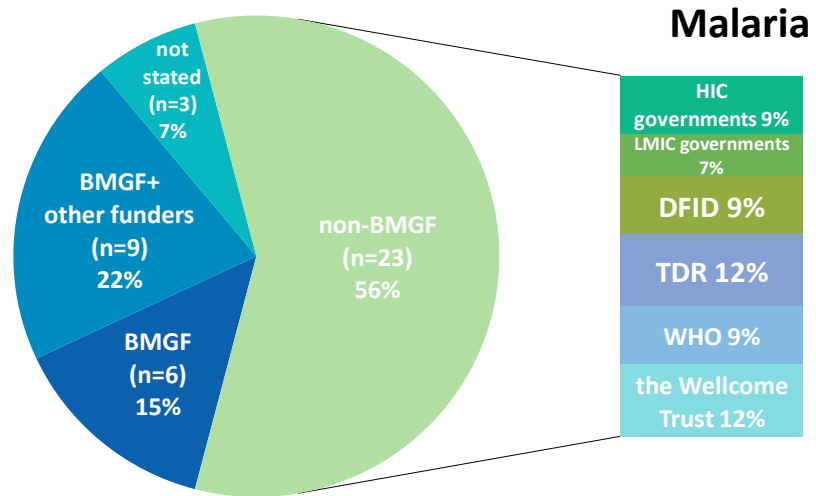


Figure 2 – Details of health economic evaluation funders

\*Governments include those government bodies or agencies other than DFID, USAID, US CDC, TDR, NIAID. HIC: High-income country



Figure 2 – Details of health economic evaluation funders – above shows the distribution of funding sources for the 224 included economic evaluations. While the majority of studies were funded by non-BMGF organisations, BMGF funded 47 economic evaluations and was cited as the funding body more frequently than any other individual organization. Other institutions that funded a notable number of economic evaluations included USAID (n=12), DFID (n=9), the World Health Organisation (n=11), and the GAVI Alliance (n=11). The source of funding was unclear in 42 economic evaluations (19% of all included).

Table 3 illustrates the substantial variation in the settings where BMGF funded economic evaluations in the four programme areas. As shown, most of the malaria studies funded by BMGF were in low-income settings, whereas all HIV/AIDS studies were in upper-middle-income settings. The category 'mixed settings' in Table 3 indicates settings that were a combination of either low-, lower-middle-, or upper-middle-income settings.

**Table 3 - Number of identified economic evaluations by type of funder, country income level of setting where the economic evaluation was undertaken, and area of interest**

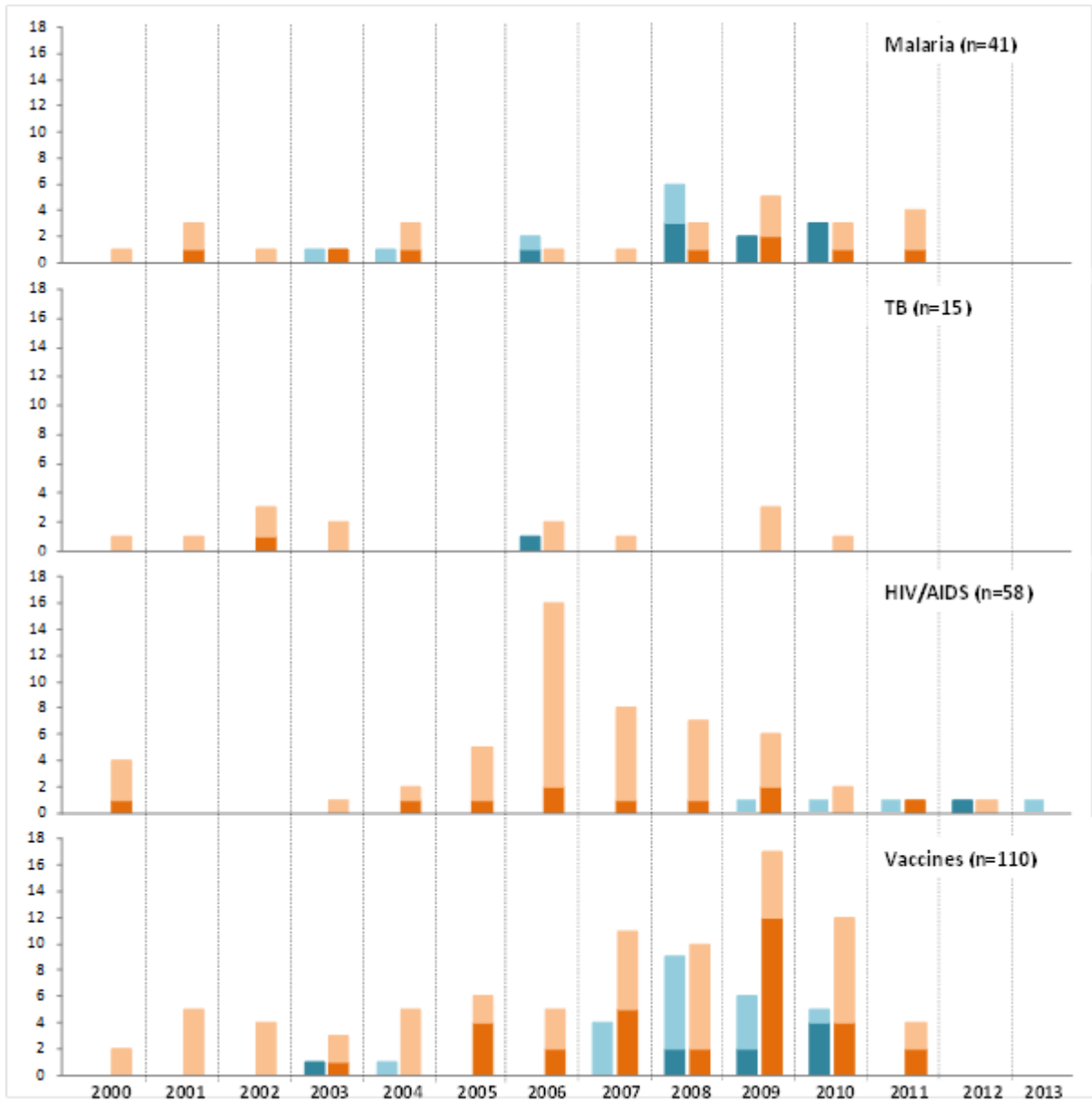
Country income level	BMGF				non-BMGF			
	Malaria	TB	HIV	Vaccine	Malaria	TB	HIV	Vaccine
Low-income	10			7	10	4	18	14
Lower-middle income	2	1		9	9	3	13	15
Upper-middle income			5	4	2	7	17	37
Mixed settings	3			6	5		5	18

Figure 3 displays the number of economic evaluations published by year in the four programme areas, classified by use of the DALY outcome measure and whether studies were BMGF-funded or not. There was a limited number of economic evaluations published prior to 2004 in particular in the HIV/AIDS programme area, and a substantial increase in the use of the DALY from 2007 in both BMGF-funded and non BMGF-funded economic evaluations, notably in the vaccine programme area.

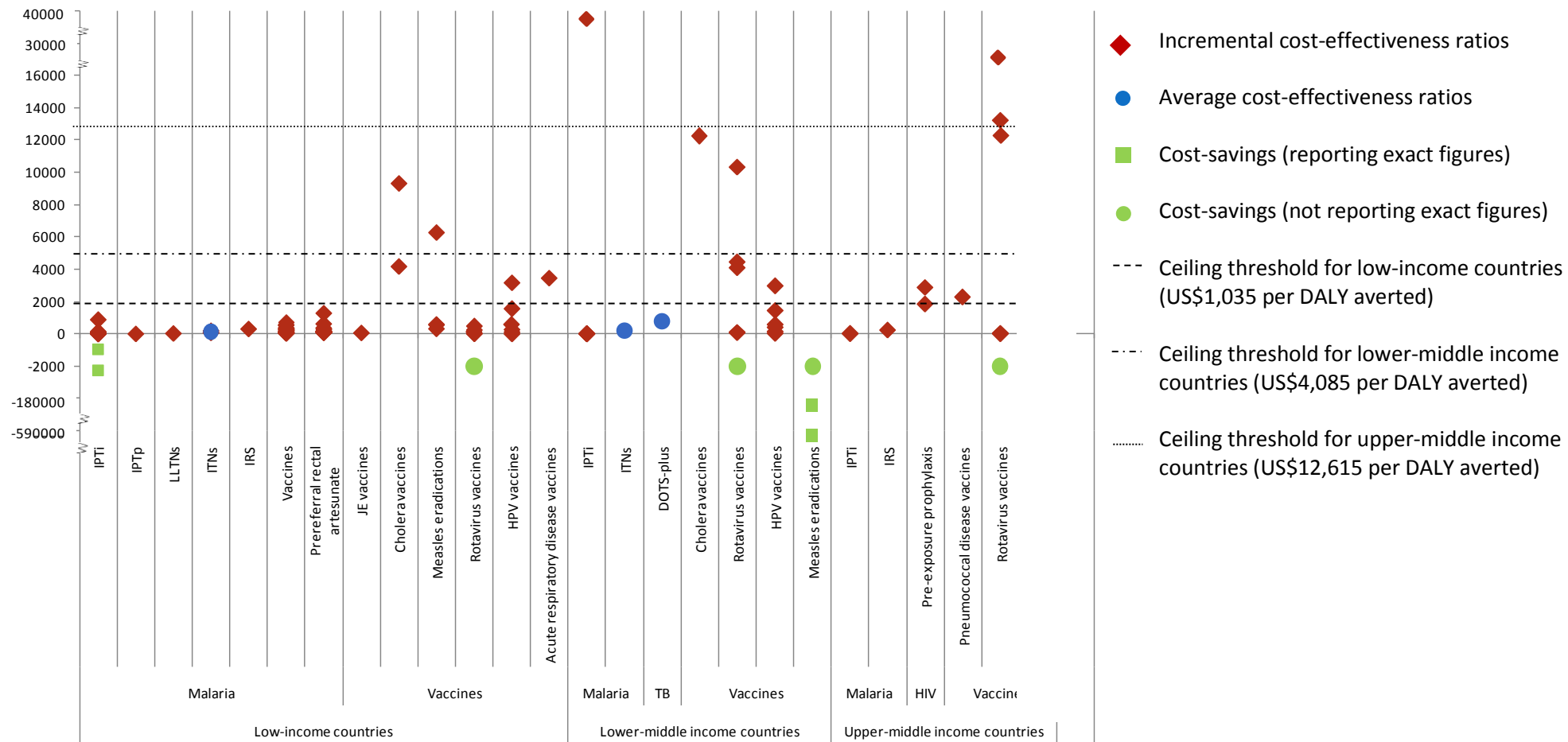
Figure 4 shows the ICERs derived from each of BMGF-funded cost-per-DALY studies. Almost all of the interventions in areas of malaria, TB and HIV/AIDS reflected reasonable value for money when assessed against a threshold equal to Gross National Income per capita as calculated by the World Bank (US\$1,035 for low-income countries) (12). Notably however, each of the results exceeds \$300/DALY<sup>2</sup> – a figure nominated by Bill Gates in a recent interview as a threshold that may be appropriate in poor country health systems (20). In the case of vaccines, different settings yield substantial differences in ICERs largely due to variations in disease burden (21) and methodological approaches, rather than to actual differences in effectiveness and/ or costs between the interventions being evaluated.

<sup>2</sup> Bill Gates used the term “life year” in the Washington Post interview published on 17<sup>th</sup> May 2013

■ BMGF/DALY   
 ■ BMGF/Not DALY   
 ■ Non-BMGF/DALY   
 ■ Non-BMGF/Not DALY



*Figure 3 - Number of BMGF and non- BMGF-funded economic evaluations by disease published from 2000-2013*



**Figure 4 - Cost-effectiveness league chart showing of ICERs of interventions being evaluated in BMGF-funded cost effectiveness analyses (using DALYs averted outcome measure) (n=20)**

*IPTi: Intermittent preventive treatment for infants; IPTp: Intermittent preventive treatment for pregnant women; LLTNS: Long-lasting treated nets; ITNS: Insecticide treated nets; IRS: Indoor residual spray; JE: Japanese encephalitis; HPV: Human papilloma virus; DOTS: Directly observed treatment, short course; source of consumer price index and purchasing power parity: IMF World economic outlook database*

*The costs in the ICERs are converted to International Dollars using purchasing power parity (PPP) adjustments and inflated to the year 2012 using respective consumer price indices (CPI). The CPIs and PPPs are from IMF World Economic Outlook database.*

## 4.2 Variations in methodology of cost-per-DALY studies funded by BMGF

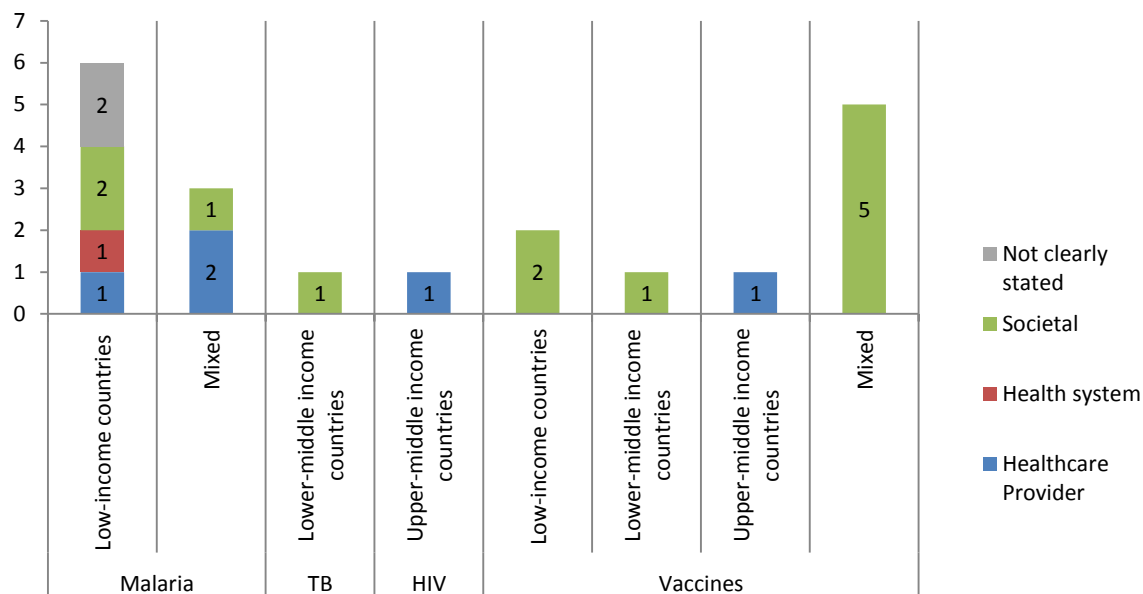


Figure 5 – Number of study perspectives adopted by disease/intervention and study setting in BMGF-funded cost effectiveness analyses (using DALYs averted outcome measure) (n=20)

Most (12 studies) adopted a societal perspective, with 5 studies applying a healthcare provider and one study adopting a health systems perspective. Two studies did not clearly state the perspective used (Figure 5). Seven studies did not clearly state the time horizon, while some studies with time horizons of longer than 1 year did not perform discounting of effects. Where discounting was performed, all studies used a rate of 3% (Figure 6).

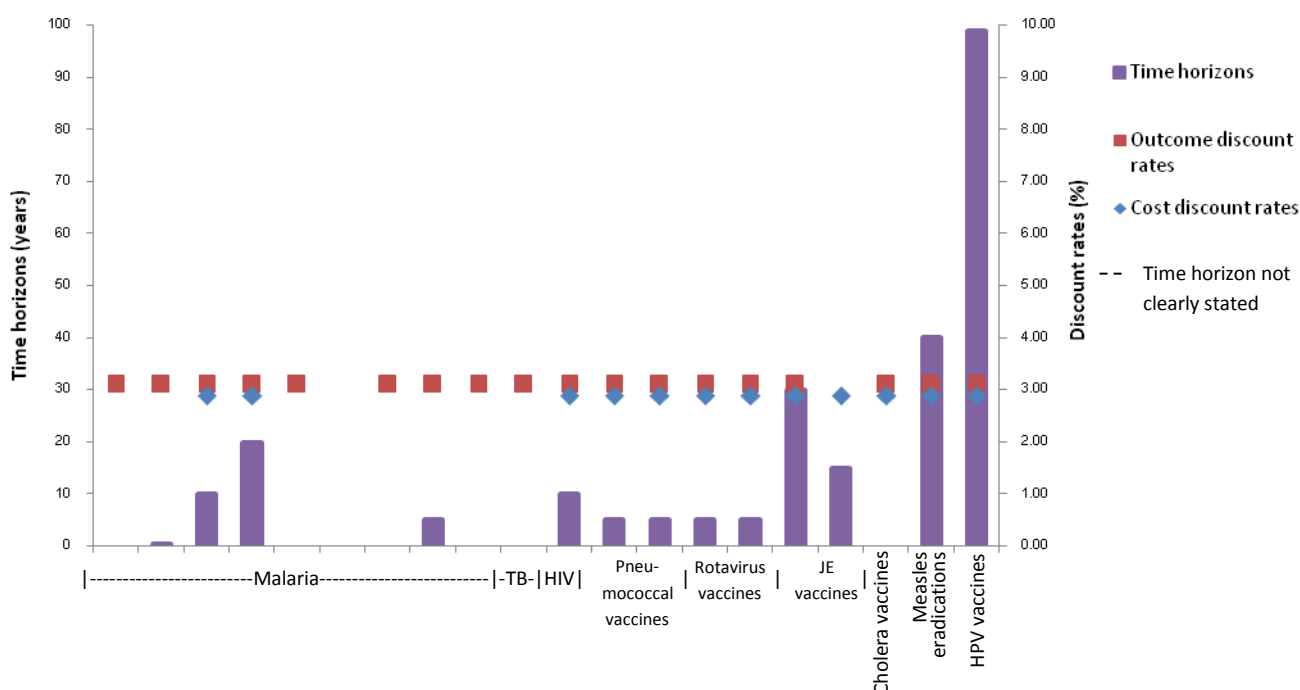
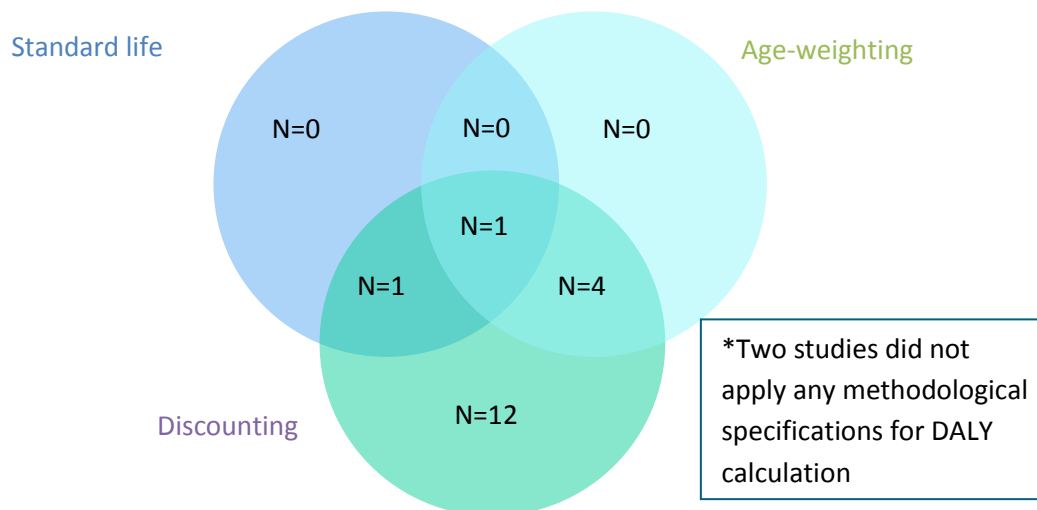


Figure 6 – Time horizon and discount rate used by disease/intervention in BMGF-funded cost effectiveness analyses (using DALYs averted outcome measure) (n=20)

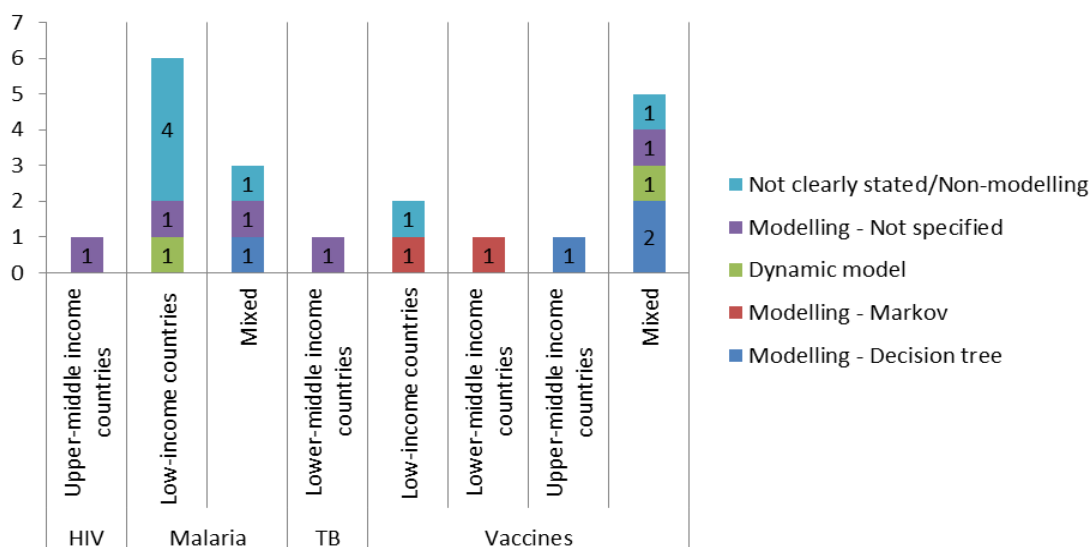
JE = Japanese Encephalitis



**Figure 7 – Number of studies adhering to methodological specifications for DALY calculation made by the Global Burden of Disease Project in BMGF-funded cost effectiveness analyses (n=20)**

Figure 7 shows the method used to calculate DALYs in the relevant studies, with each circle representing a particular method (standard life table, age-weighting, discounting) and overlapping circles representing the use of more than one method.

Most studies did not follow the specific methodological recommendations of the Global Burden of Disease Project (22). Only the study by Mbonye et al (23) adhered to all three major methodological specifications, namely using a standard life table, applying age-weighting, and discounting for future DALYs. One study adopted a standard life table and applied discounting but not age-weighting. Four studies used age-weighting and discounting but not standard life tables. Twelve of the 20 studies discounted future DALYs but did not apply age-weighting or standard life tables. Two studies did not apply any of the recommended methods.



**Figure 8 - Number of studies using different decision analytic model types in BMGF-funded cost effectiveness analyses (using DALYs averted outcome measure) (n=20)**

Among the twenty cost-per-DALY studies identified, 13 studies were model-based economic evaluations and seven studies either did not use modelling in the analysis, or the extent of or type of modelling used was unclear. Four studies used decision tree analysis, and two studies used a Markov model (Figure 8). Of the 20 BMGF-funded cost-per-DALY averted studies, only two studies performed probabilistic sensitivity analysis and 17 performed either univariate or multivariate sensitivity analysis. In some studies, a threshold analysis was also carried out with the sensitivity analyses (Figure 9).

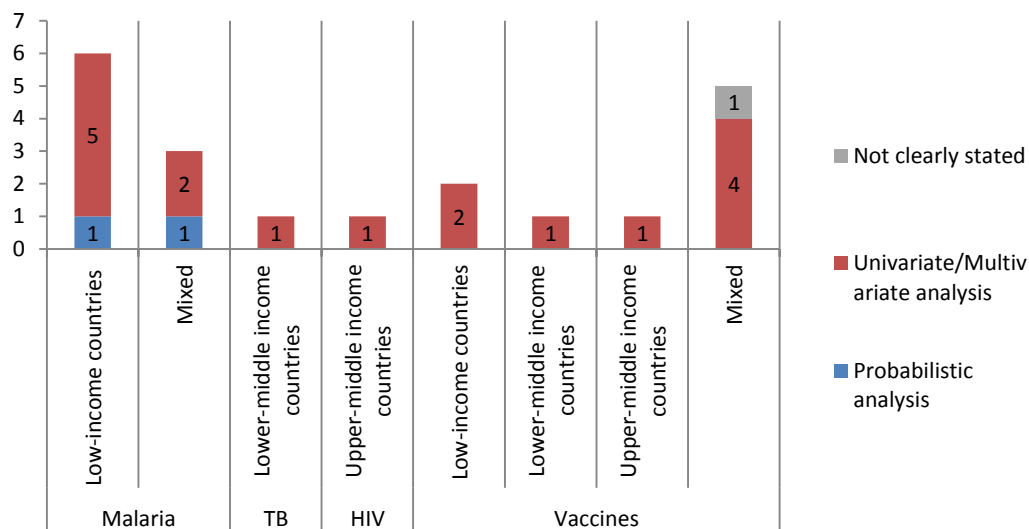


Figure 9 – Number of studies conducting each type of sensitivity analysis by disease/intervention and study setting in BMGF-funded cost effectiveness analyses (using DALYs averted outcome measure) (n=20)

### 4.3 Quality of reporting

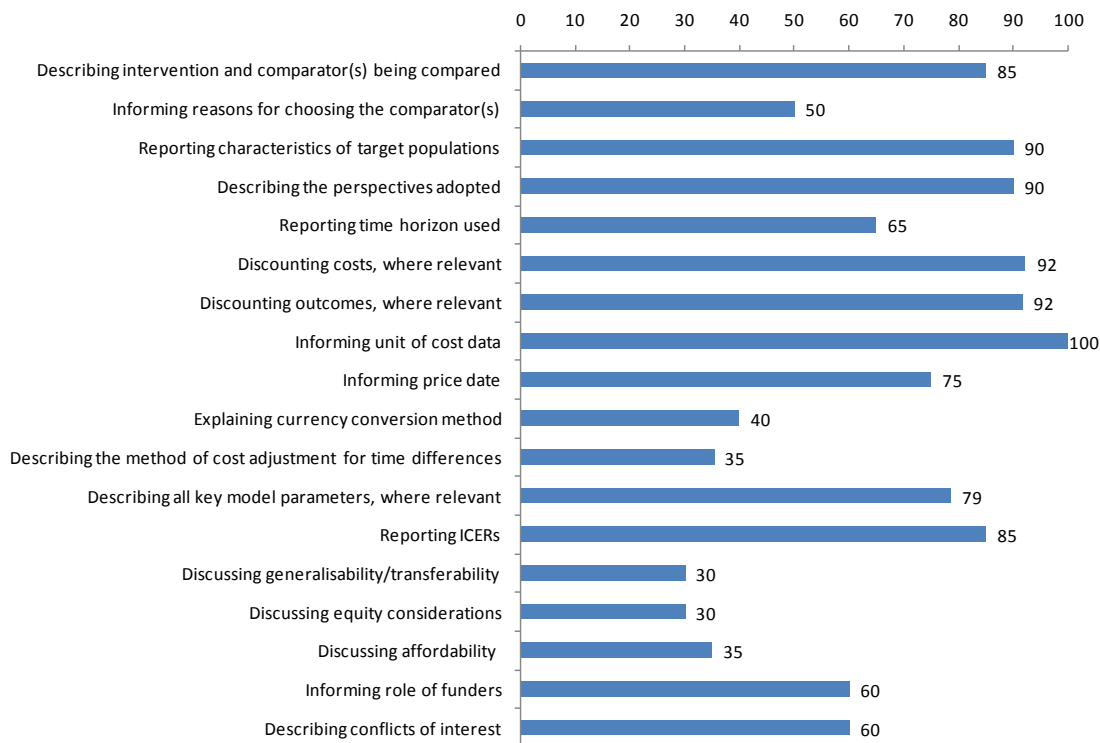
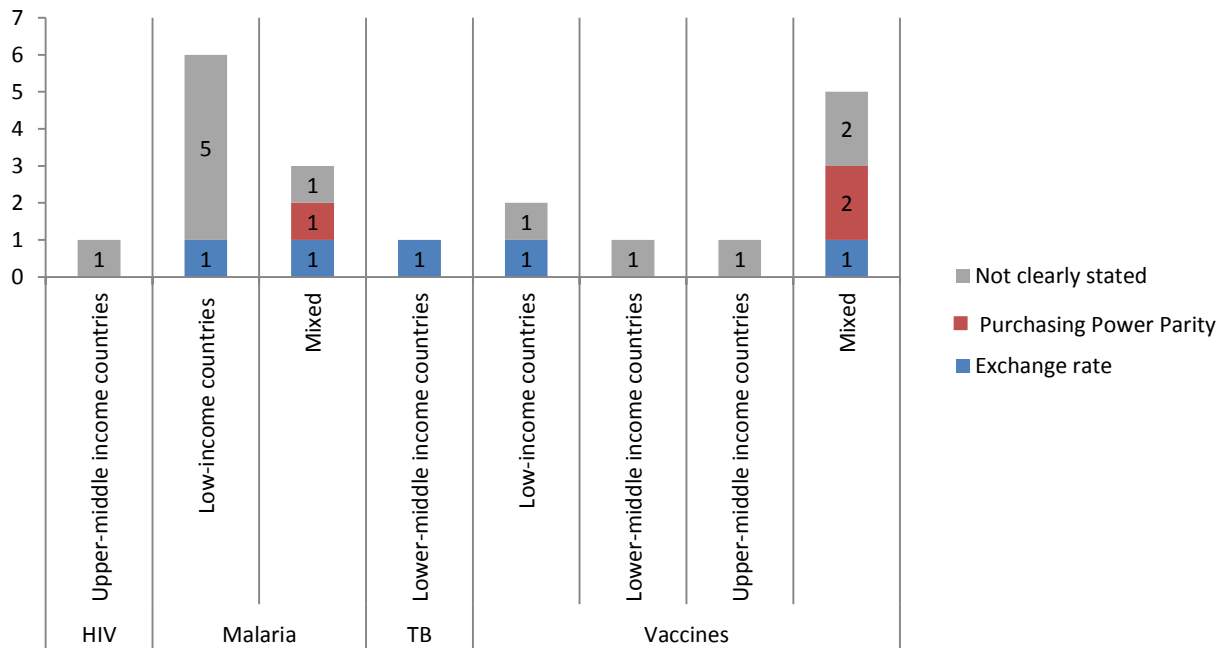


Figure 10 – Percentage of BMGF-funded cost effectiveness analyses (using the DALYs averted outcome measure) adhering to recommended practices for reporting health economic evaluations (n=20)

The proportion of studies adhering to a set of reporting standards is shown in Figure 10. Discussions of generalisability / transferability and equity considerations were the elements most frequently omitted from the studies. Only eleven of the 20 studies discussed the choice of study comparator(s), with the most common rationale cited (in six studies) being current or first-line practice. Reporting of the decision rules applied was variable, with approximately half the studies referring to a ceiling of up to 3 times per capita GDP per DALY gained, as identified by the Commission on Macroeconomics and Health (CMH) (24). Seven of twenty economic evaluations made positive recommendations factoring adoption of the technology.



**Figure 11 – Numbers adopting different approaches for currency conversion in BMGF-funded cost effectiveness analyses (using DALYs averted outcome measure) (n=20)**

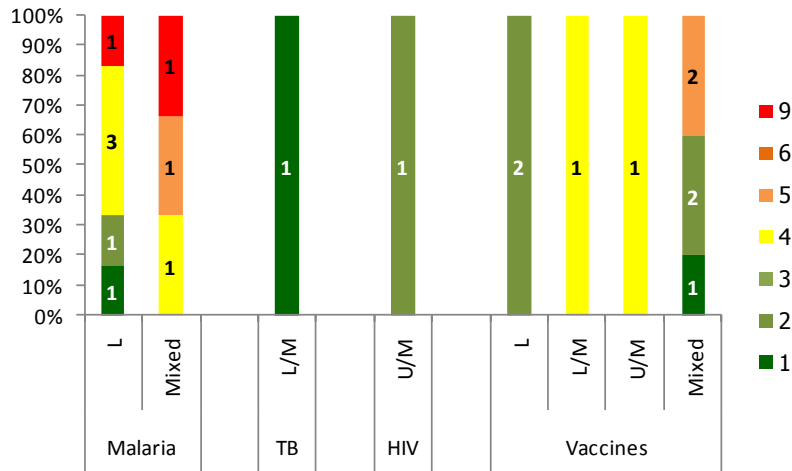
More than half the studies did not describe the method used for currency conversion, despite relying on cost data from sources outside the study settings. Where the method was described, unadjusted exchange rates were more frequently used than purchasing power parity (PPP) adjusted measures (Figure 11).

Figure 12 below shows a ranking of the quality of the sources of (a) baseline clinical data, (b) clinical effect sizes, and (c) costs/resource use data in the selected studies. The method of ranking is adapted from Cooper et al (17) and is shown in **Appendix 4**.

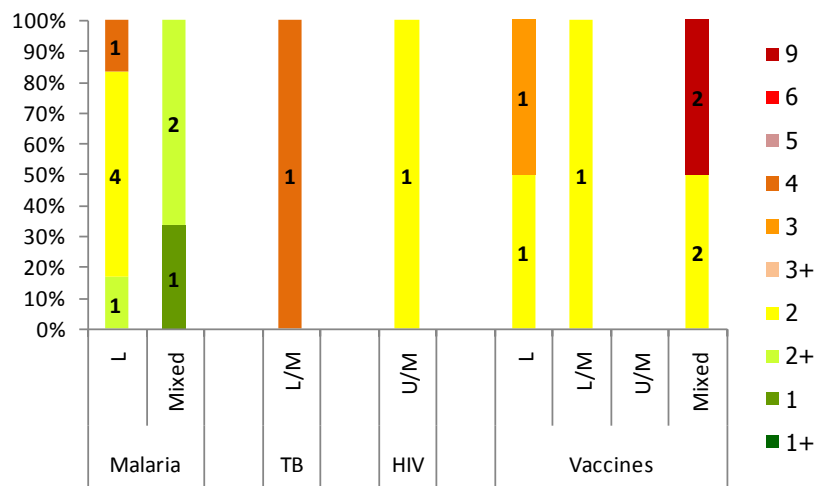
The studies utilised data of a higher quality for cost and resource use parameters with the majority of studies estimating costs by drawing on reliable administrative databases or data sources in the same jurisdiction. Nevertheless, baseline clinical data were often derived from lower quality sources and clinical effect sizes were mainly based on data drawn from single RCTs.

#### 4.4 Quality of evidence used

a) Baseline clinical data (e.g. incidence, prevalence of disease) (n=20)



b) Clinical effect sizes (n = 18)\*



c) Costs/resources used (n = 20)

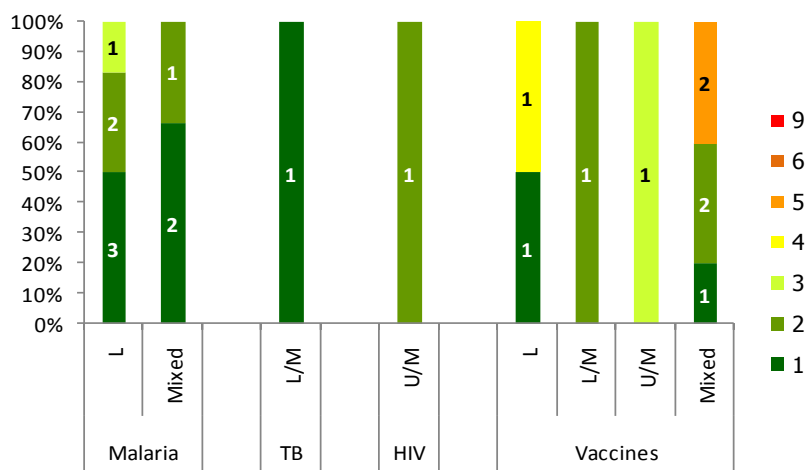


Figure 12 – Ranks of evidence used in in BMGF-funded cost effectiveness analyses (using DALYs averted outcome measure)

\* Although they were properly referenced, two sources of evidence for clinical effect size could not be retrieved



## 5. Discussion

### 5.1 Key findings

The use of economic evaluation in informing resource allocation decision-making in LMICs reflects a growing trend (25) (26), especially in light of increasing investment in this type of analysis by major donors such as BMGF, GAVI Alliance, and the World Health Organization. However, although accepted as a useful tool, economic evaluation can be sub-optimal where methodological approaches and reporting standards are non-uniform. This review summarises key issues identified in cost utility analyses presenting cost-per-DALY measures published in international journals in selected areas.

From the sample of economic evaluations in the four programme areas in LMIC settings and published from 2000 onwards, this review found:

- BMGF funded 21% of all types of economic and 28% of all cost-per-DALY studies. 43% of BMGF funded studies used the DALY averted outcome measure.
- There were significant variations in almost every methodological component.
- The most common analytical perspectives applied were societal and health care providers'. The adoption of different perspectives can easily generate different conclusions even for studies undertaken in similar settings and focusing on the same intervention.
- The majority of studies provided insufficient information about currency conversions and/or methods for adjusting costs to account for temporal disparities. Where information was provided, crude exchange rates were frequently used to convert unit costs drawn from other settings (often resource-rich countries). Purchasing Power Parity (PPP) was used in very few studies.
- Poor adherence to the three key methodological specifications for DALY estimation raises significant concerns as variant approaches to DALY calculation limit comparability between studies.
- Although widely considered as the most comprehensive method of dealing with the various sources of uncertainty in economic evaluations (27), few studies presented probabilistic sensitivity analyses.
- The findings of this review are consistent with previous reviews that found better quality data pertaining to costs and resource utilisation in economic evaluations undertaken in resource-limited settings than for baseline clinical data and clinical effect sizes (16).
- Generalisability/transferability of results and equity implications of evaluated interventions were discussed in less than one-third of all reviewed studies.
- Only 35% of studies discussed the affordability of the interventions being assessed, despite these studies being undertaken in very resource-limited settings.

### 5.2 Implications for research

This review highlights a series of implications for research, especially regarding the quality of reporting and the use of information in the analyses. In particular:

- There is value in adherence to best practice for reporting economic evaluations, with particular emphasis on ensuring clarity in the selection of comparator(s), and in the derivation and adjustment of costs.

- The generalisability and transferability of results, affordability of the intervention, and equity implications of evaluated interventions should be discussed explicitly, in order to make research findings more useful to decision-makers.
- The roles of funders and any actual or perceived potential conflicts of interest should be better addressed in future studies.
- Without standardisation of methods, differences in cost-effectiveness ratios may reflect differences in study methodologies rather than true differences between the interventions being evaluated.
- Overall, there is the pressing need for uniform methodological specifications and reporting standards for conducting health economic evaluations in LMICs to improve the quality and reduce variation in both methods and reporting of results.

### **5.3 Implications for BMGF**

Given that BMGF is a leading individual funding body supporting studies in LMICs, the findings of this review strongly support the introduction of a reference case for health economic evaluations by BMGF.

This review is used not only to justify the importance of developing economic evaluation reference cases, but also to provide valuable inputs for the development of a Gates Reference Case (Gates-RC, Section 3).

A recent focus among key donors such as the Global Fund has been on ensuring the effective use of resources and on generating improved value for money (28). While this review focused on BMGF-funded economic evaluations, it has implications for other global funders who support and use economic evaluations in LMICs.

If a Gates-RC were to be adopted by other entities such as global donors, or governments who lack their own frameworks for economic evaluation, it would enable them to improve the quality of the value-for-money information for various interventions across a range of health problems and settings. Improvement in the conduct and transparency of economic evaluations would enhance both usability and comparability, and ultimately enable decision-makers to apply the results of Gates-RC economic evaluations with confidence when priority setting in health. In order to evaluate impact of a Gates-RC, the review of published economic evaluations in LMICs should be repeated.

### **5.4 Limitations**

This review drew on English language publications only, and excluded conference proceedings and academic theses, as well as 'grey literature' such as government reports. The review included only the malaria, TB, HIV/AIDS, and vaccines programme areas, all of which have received strong support from major global donors, including Global Fund, BMGF, and GAVI Alliance. While it is useful to assess improvements in study quality and reporting over time, the relatively small number of cost-per-DALY studies funded by BMGF limited the comparison of studies undertaken several years ago with those undertaken recently. Lastly, this review may have omitted some studies published in 2012 and 2013.

## Section Two: The case for a Gates Reference Case

### 1. Introduction

This section explores the implications for the Bill and Melinda Gates Foundation of adopting a reference case for the conduct and reporting of economic evaluations. Section One identified substantial variation in published economic evaluations, which lends support for the use of a reference case. However, there is a complex ecosystem shared by the various global funding institutions (such as BMGF) that conduct, commission and report economic evaluations; the users of economic evaluations; and ultimately, the populations and individuals who are the intended beneficiaries and users of the technologies. As a result, it is important to explore the potential for unintended consequences arising from such a major methodological shift. To our knowledge, this is the first time that the adoption of a reference case has been considered by a global health funding body such as BMGF. We therefore seek to explore the following hypothesis:

*“That the usefulness of Bill and Melinda Gates Foundation-funded economic evaluations of health technologies in low and middle income countries would be enhanced by the adoption of a reference case”.*

This section of the report maps the funding ecosystem and addresses the benefits and risks of reference cases in general, and specifically in relation to BMGF. Ways in which BMGF could mitigate any risks and maximize the benefits of adopting a reference case are identified and inform the report recommendations in Section 4.

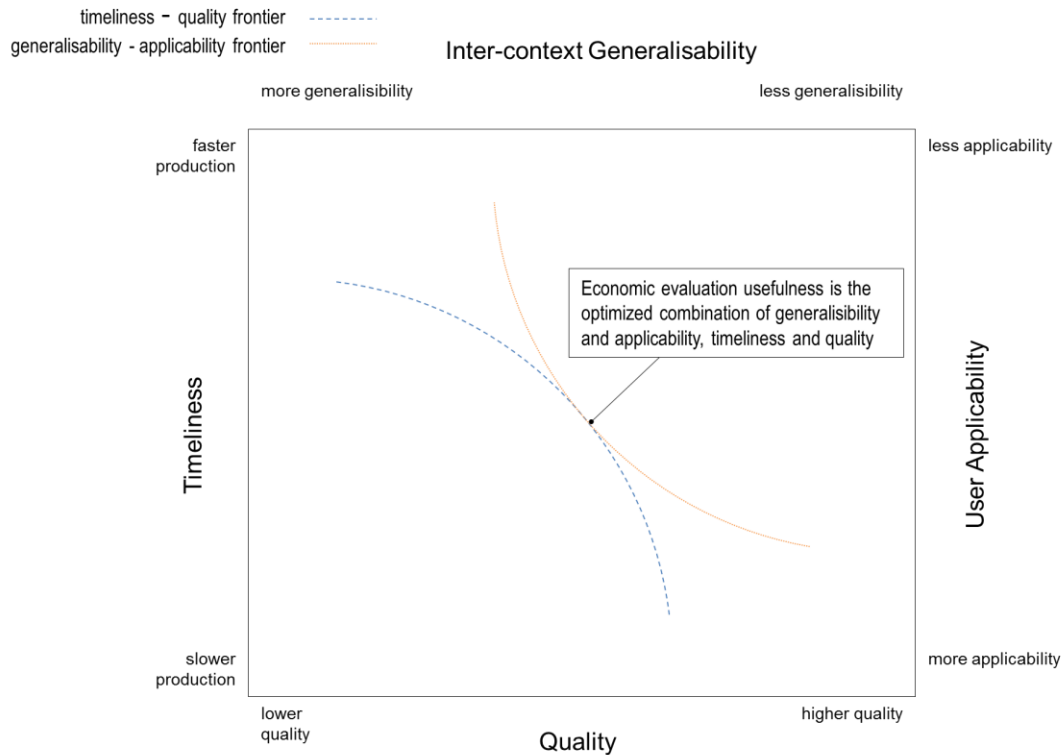
### 2. Setting the scene

#### 2.1 What is useful?

“Usefulness” is the extent to which an economic evaluation is able to inform sound decisions in health and health care resource allocation, where a sound decision is one that meets the normative assumption that the decision-maker is seeking to improve population health within available resources. The principles and specifications that underpin any adopted reference case should also reflect this same normative assumption.

To assess usefulness we apply the following criteria: comparability of results, quality, timeliness, applicability, generalisability, and reproducibility. Reproducibility is a function of how well and comprehensibly an economic evaluation is reported. Comparability is optimal when multiple economic evaluations use similar principles and specifications to construct the analyses; however comparability becomes irrelevant if the methods used are not appropriate to the decision problem. There are often inherent trade-offs between timeliness and quality, and between applicability and generalisability. These trade-offs are illustrated in Figure 13 - Optimisation of usefulness measures where quality is represented on the bottom axis and timeliness represented on the left vertical axis allowing a timeliness-quality frontier to be drawn; and where user applicability is represented on the right vertical axis and inter-context generalisability is represented on the top horizontal axis allowing a generalisability-applicability frontier to be drawn. A key objective in designing a reference case for adoption by BMGF would be to seek an optimal balance of generalisability, applicability, timeliness and quality, while seeking to maximise reproducibility and comparability.

## Usefulness measures optimisation



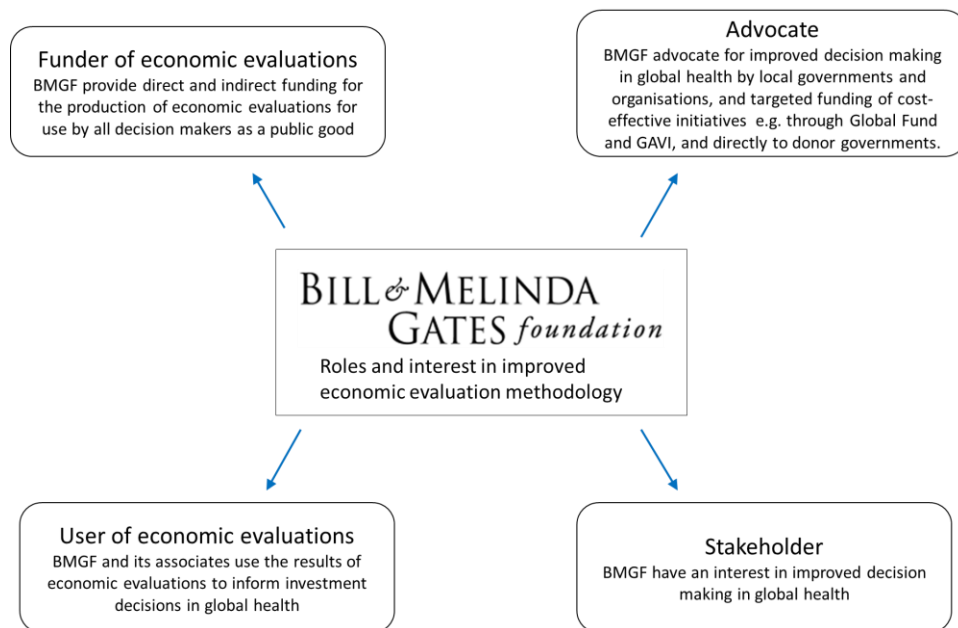
*Figure 13 - Optimisation of usefulness measures*

### 2.2 Usefulness to whom?

The usefulness of an economic evaluation should be judged from the perspective of the decision-maker wishing to use the results to inform an investment decision.<sup>3</sup> Thus in order to assess usefulness, the user must be identified: for the purposes of this report, a user of an economic evaluation can be characterized as (1) a local (within country) decision-maker at a national or sub-national level, and (2) BMGF, a BMGF-partner or a grantee.

BMGF is a well-established funder of economic evaluations in low and middle-income countries, not only utilising these in its own decision-making, but also ensuring that the funded economic evaluations are useful for in-country decision-makers. BMGF is also an advocate for improved decision-making, as well as a stakeholder supporting improved population health at the local level. In each of these roles BMGF's interests are furthered by improving economic evaluation methodology, as shown in Figure 14 below.

<sup>3</sup> Of course, there may be other reasons for undertaking an economic evaluation, such as a post-hoc review to determine if an earlier investment decision was vindicated. In this instance, "usefulness" may be more a function of comprehensive characterization of costs and pre-determined measures of effect, rather than comparability and generalisability.



**Figure 14 - Roles and interests of BMGF in improved economic evaluation methodology**

Figure 15 shows the Funder Ecosystem – a simplified schematic of the funding streams and influences that result in decisions about investment in health technologies in countries and their relationship to population health. The ecosystem shows the potential impact of the adoption of a reference case, as well as the influence BMGF and its partners have on the conduct and reporting of economic evaluations. In theory, the adoption of a reference case would not impact the *process* of decision-making by either BMGF or country-level decision-makers. However, to the extent that the results of economic evaluations are used by BMGF and country-level decision-makers, a reference case would strengthen both their utility and influence.

Crucially, the Funder Ecosystem shows a relationship between the conduct and reporting of economic evaluations, investment decisions by both BMGF and country-level decision-makers, and the length and health-related quality of life (QoL) of local populations. While the displayed relationship between conduct and reporting of economic evaluations and local population health may fail to adequately capture the complexity of the multitude of influences on priority setting decisions in health, it illustrates the central importance of improved economic evaluation methodology.

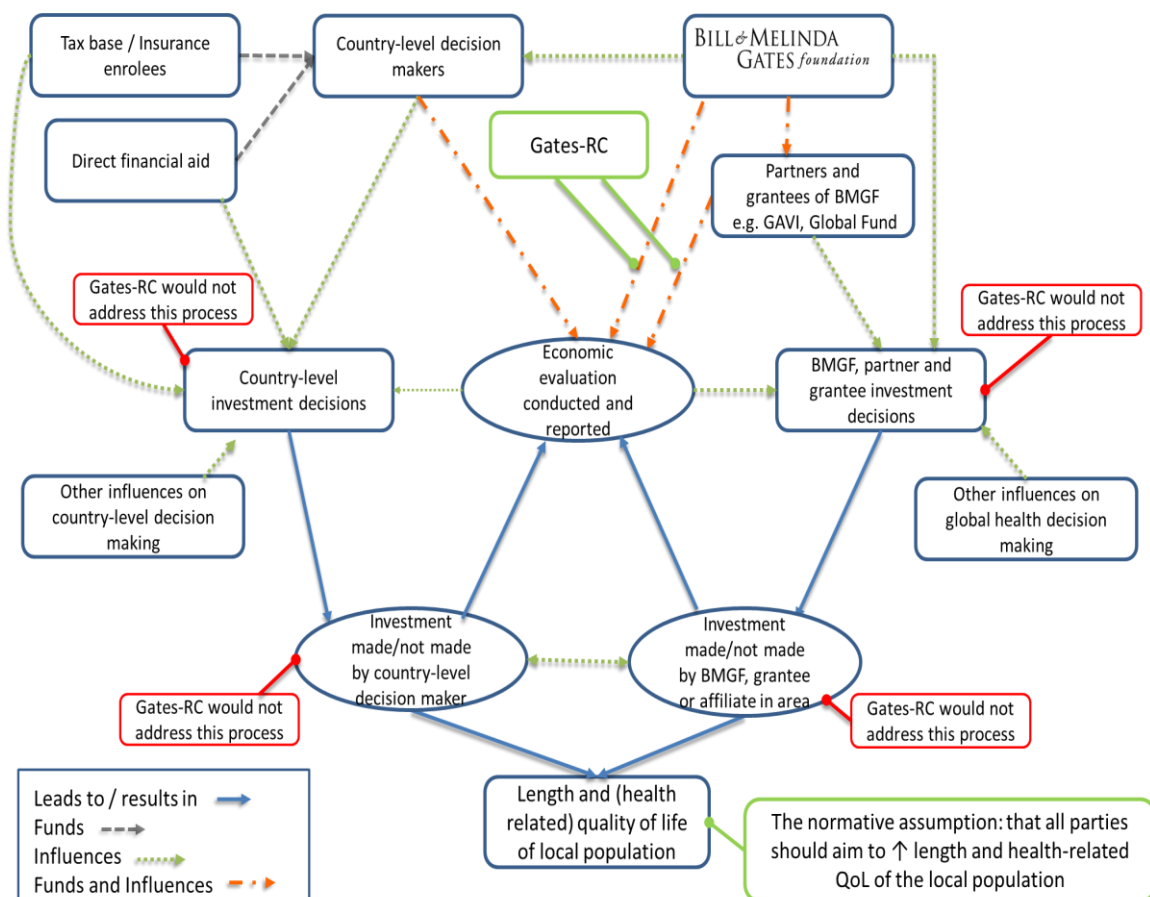


Figure 15 - The Funder Ecosystem

### 2.3 What is a reference case?

The Panel on Cost-Effectiveness in Health and Medicine (the ‘US Panel’) proposed the use of a reference case “to improve comparability of CEAs designed to inform decision-making while allowing analysts the flexibility to design studies that answer issues specific to a particular problem or industry” (7). The underlying principle is that the use of the reference case facilitates the comparison of the results of different studies, while “each study contributes to a pool of information about the broad allocation of resources as well as to the specific questions it was designed to address”.

In 2003, the World Health Organisation (WHO) published a *Guide to Cost-Effectiveness Analysis* (29) in which the authors sought “to provide analysts with a method of assessing whether the current mix of interventions is efficient as well as whether a proposed new technology or intervention is appropriate” and “to maximize the generalizability of results across settings”. Improved comparability of studies was sought less to facilitate consistency in decision-making than to increase generalizability, so that each country need not undertake *de novo* analyses. While this WHO document made no explicit mention it is arguably describing a form of reference case.

Similarly, while the 2008 WHO *Guide for Standardization of Economic Evaluations of Immunization Programmes* (30) did not go as far as defining a reference case, it clearly articulated an endorsement of the reference case approach.

The use of the reference case approach by decisions makers was augmented by the adoption of a reference case by National Institute for Health and Care Excellence (NICE) in 2004 (revised most

recently in 2013 (31)). NICE’s analyses and guidances inform resource allocation in the National Health Service for England and Wales, particularly with regard to new technologies and services. A key issue is to facilitate consistency in decision-making in order to maximise the health gains from limited resources. The NICE reference case specifies the methods considered to be most appropriate in estimating clinical and cost effectiveness in technology appraisals. NICE does not preclude the presentation and consideration of non-reference case analyses; however it requires that any departure from the reference case is adequately justified.

**Table 4 - Summary of areas of recommendation from WHO, NICE, US Panel, HITAP**

Areas of recommendation	US Panel	WHO Guide to Cost Effectiveness Analysis	NICE Guide to Technology Appraisal	HITAP (Thailand) Methods Guide
Defining the decision problem	✓		✓	✓
Comparator(s)	✓	✓	✓	✓
Perspective on effects / outcomes	✓	✓	✓	✓
Measuring and valuing health effects	✓	✓	✓	✓
Synthesis of evidence on health effects		✓	✓	✓
Source of data for measurement of health-related quality of life		✓	✓	✓
Source of preference data for valuation of changes in health-related quality of life	✓		✓	✓
Perspective on costs	✓	✓	✓	✓
Evidence on resource use and costs (+/- source of evidence)	✓	✓	✓	✓
Type of economic evaluation	✓	✓	✓	✓
Time horizon	✓	✓	✓	✓
Equity considerations	✓		✓	✓
Discounting	✓	✓	✓	✓
Uncertainty and/or sensitivity analysis	✓	✓	✓	✓
Reporting guidelines / requirements	✓	✓	✓	✓

Table 4 shows a high degree of consensus among the different entities on many of the elements of a reference case<sup>4</sup>. The NICE Guide to Technology Appraisal and the HITAP Methods Guide incorporate all 'areas of recommendation' included in either the US Panel or WHO Guide. The NICE and HITAP reference cases explicitly refer to equity considerations, while the US Panel captures equity considerations indirectly through a requirement to conduct sensitivity analyses in situations where

<sup>4</sup> Non-exclusive “areas of recommendation” classifications were used to encompass the various ways that the different reference cases were presented. Author judgement was required to determine if an element in the above table was either explicitly part of the reference case or a good methodological practice recommendation for the conduct or reporting of economic evaluations.

choice of wage inputs or other characteristics of populations used to calculate costs of an intervention (applying a societal perspective) influence the cost effectiveness ratio. The WHO Guide makes specific reference to equity and ethical considerations resulting from use of CEA to inform decision-making in health, but does not specify an explicit requirement for equity considerations in its reference case. The Gates-Reference Case (presented in Section 3) builds on the lessons drawn from previous initiatives to introduce reference cases and standardised methodologies, and contains many similar aspects to the areas of recommendation in Table 4 **Error! Reference source not found.**

### **3. The case for a reference case**

#### **3.1 The advantages of a reference case**

The overarching advantages of using a reference case are that it can improve not only the quality of individual decisions through robust standards for the planning, conduct and reporting of economic evaluations, but also facilitates greater overall consistency in decision-making over time.

Where the results of multiple studies are to be compared, the use of a reference case facilitates the meaningful and explicit comparison of the analysis and findings. For example, among studies adopting different analytical perspectives, cost data will not be directly comparable even where the interventions and their application are the same. This will be true for comparisons of evaluations of different interventions undertaken within a single country, as well as for those between evaluations of the same intervention in different countries.

#### **3.2 The disadvantages of a reference case**

While few would argue against the benefits of explicitness, comparability and improved quality, the potential disadvantages of using a reference case must also be acknowledged. Economic evaluation of health technologies is a rapidly evolving field; as such there remain several areas where consensus does not yet exist and thus prescriptive guidelines for the conduct and reporting of evaluations may be premature. Indeed it could be argued that progress in the development of methods might be discouraged by the greater use of reference cases.

It is not uncommon for proposals for research funding in the United Kingdom to affirm that economic evaluation will be undertaken in accordance with NICE's methods (ie the NICE reference case (31)). Even where research proposals are not directly intended to inform a NICE appraisal or guidance, use of the NICE reference case is perceived to strengthen credibility and demonstrate rigour, and to ensure that if NICE were to address the same or a related issue subsequently, the research findings would be applicable. The promulgation of a reference case thus carries the risk that certain practices become "set in stone", and in so doing make securing funding for projects or publishing studies that diverge from 'approved' methods more challenging. The Medical Research Council (MRC) in the United Kingdom mitigates this risk by funding (with direct support from NICE) on-going research in methods for economic evaluation.

Use of the reference case has played a central role in the appraisal of health technologies by NICE since 2004. Although empirical evidence of the effectiveness and cost-effectiveness of the reference case approach is lacking (and the counterfactual cannot be observed), there does not appear to have been any brake on the development and application of methods, particularly when considering the nature of the economic evaluations undertaken. While the driver of methodological development



has probably been the contestable nature of the decision-making framework, use the reference case does not appear to have constrained it.

As a reference case is a type of guideline (in the broadest sense), opportunity costs associated with the development of guidelines should be considered. Maynard contended that 'investment in the improvement of guidelines and development of consensus about best practices yields a small marginal product'. This, however, is not an argument against reference cases *per se* but a recognition of the effort required for their development, refinement and implementation. Opportunity costs might be greater if adherence to a reference case were to require costly evaluation. However any such opportunity costs should be balanced against the benefits of improved decision-making and the potential costs of poor allocation of resources in its absence.

Finally, the balance of advantages and disadvantages of the use of a reference case will be strongly influenced by how prescriptive it is. A highly prescriptive reference case will enhance comparability and improve quality but risks imposing methods which are not appropriate to the decision problem, while one that is less prescriptive may be inadequate to facilitate cross study comparisons or reliably inform resource allocation decisions. Importantly however, the issue is not whether to use a reference case but how to determine how prescriptive the reference case should be and how to interpret analyses that diverge from it.

## **4. The case for a Gates Reference Case**

### **4.1 Opportunities and risks**

As identified by Drummond (1), Gold (7), and the World Health Organisation (29), a reference case is essentially a mechanism for enhancing the ability of economic evaluations to inform good decision-making for resource allocation in health. Until recently, the issue of whether to adopt a reference case has been limited to those making multiple health resource allocation decisions that required consistency, transparency, and objectivity. This has been largely the domain of ministries of health and third party payers in various countries or states. To date, mandatory or recommended standards or guidelines for economic evaluation have been published in at least 33 countries, the majority by state institutions in high-income countries (32).

In addition, the publication of the Guide to Cost Effectiveness by the World Health Organisation in 2003 provided a recommended framework for generalised cost effectiveness analysis, and in conjunction with the WHO-CHOICE Programme, has been a valuable tool for the conduct of economic evaluations in low and middle income countries. The question under consideration here is, however, whether BMGF, as a global philanthropic funder with multiple initiatives, stakeholders, and objectives, should also adopt a reference case.

It is important that the term "adoption" as it would apply to a Gates-RC is defined appropriately. Adoption goes beyond promulgating guidance for the conduct of economic evaluations. By adoption we mean that funding and support should be contingent on some level of compliance. Although the nature and extent of adoption may vary depending on the activity to which the reference case is applied, a parallel can be drawn between a Gates-RC and the NICE requirement that for NHS funding decisions to be made, the supporting economic evaluation must follow the NICE reference case, and any departure from it must be justified.

**Table 5 - Perceived advantages and disadvantages of Reference Cases**

	Pros	Cons
Reference Cases in general	<ul style="list-style-type: none"> <li>• Transparency</li> <li>• Quality (minimum standard)</li> <li>• Consistency</li> <li>• Comparability</li> <li>• Generalisability</li> </ul>	<ul style="list-style-type: none"> <li>• Perceived to reduce flexibility and methodological innovation</li> <li>• Bias against difficult to measure parameters and outputs</li> <li>• Unnecessarily onerous (potentially)</li> <li>• Increased time and expense to produce</li> <li>• Misunderstanding of use</li> </ul>
Gates Reference Case	<ul style="list-style-type: none"> <li>• Improves the usefulness of economic evaluations for BMGF the funder, user, advocate, stakeholder.</li> <li>• Increase consistent decision making across investment types</li> <li>• Improves state of the science</li> </ul>	<ul style="list-style-type: none"> <li>• Risk of not achieving optimal generalisability-applicability trade-off</li> <li>• Risk of imposing scientific and social value judgements on local decision makers if RC too prescriptive</li> <li>• Unclear weighting of results of economic evaluation to inform multi-criteria decisions</li> <li>• Limited usefulness of a reference case in data-poor constituencies</li> </ul>
How can BMGF optimise pros and mitigate cons?	<ul style="list-style-type: none"> <li>• Allow contextual differences to determine specifications, but uphold the core principles</li> <li>• Insert recommendations</li> </ul>	<ul style="list-style-type: none"> <li>• Require a reporting template to gather more information</li> <li>• Regular review and update of the reference case elements</li> </ul>

The arguments in support of BMGF adopting a reference case align with the stated values and objectives of BMGF: the practice and advocacy of good decision-making, intelligent use of available data, and a commitment to allocative efficiency. As a major financial supporter donor of economic evaluations in low and middle income countries, BMGF thus has the ability to influence practice as researchers are encouraged to align with the Gates-RC in order to improve both the influence of their analyses and the likelihood of securing funding for them. Over time therefore, the Gates-RC will become a public good, and “Gates-compliant” economic evaluations will become more useful to both BMGF and local decision-makers.

In order to address the potential to stifle methodological innovation, (and given the significant potential influence of any Gates-RC on the state of the science) it is important for BMGF to commit to on-going participation in the development and refinement of the reference case. Notably, NICE recently published the fifth edition of its *Methods for Technology Appraisal*, reflecting the importance not only of keeping up with methodological developments, but also leading their evolution to ensure that the reference case continues to leverage the best possible decisions from available techniques and data sources. In the Gates-RC presented in section three of this report, certain methodological specifications do not have a clear directive. It is likely that the process of developing methodological specifications and continuing to refine existing specifications will promote innovation in LMIC economic evaluation methodology, as has been observed in the UK as a result of NICE.

BMGF, as a funder, is obliged to spend money wisely. The review in section one of this report has shown variation in methodology and quality of BMGF funded economic evaluations. BMGF now has

the ability to introduce a methodology for economic evaluation that allows it to reflect its own social and scientific value judgements.

Arguably, there is an ethical imperative to be transparent about when a social or scientific value judgement is being used to drive funding decisions with local implications. A reference case makes this possible; a Gates-RC would make any social and scientific value judgements explicit. The imperative for requiring a reference case for economic evaluations that inform priority setting decisions for multiple populations is captured in a quote by Melinda Gates: *“Whatever the conditions of people’s lives, wherever they live, however they live, we all share the same dreams”* (33). However, while we may “share the same dreams”, and the improvement in length of life and health-related quality of life of local populations is a justifiable goal, there are likely to be clear differences in the social and scientific judgements made by BMGF and local constituencies. This does not make the adoption of a reference case untenable but it does represent a risk: if a reference case is too methodologically prescriptive it will result in economic evaluations that do not usefully inform local decisions, and are therefore not fit for purpose.

In exploring the adopting of a reference case, BMGF is taking the lead in making transparent the methodological basis of economic evaluations that inform its funding decisions and the value judgements that influence them. It is likely that following the implementation of a Gates-RC, other global funders will acknowledge the importance of the transparency facilitated by the adoption of a reference case. BMGF should be open to establishing links with other global funders for collaboration in this regard.

#### **4.2 Generalisability-applicability**

Of significant importance to BMGF is that economic evaluations are generalisable across different contexts and settings. This contrasts with reference cases adopted by decision-making bodies operating within more narrowly defined contexts, where it is of secondary importance. For example, in developing the NICE reference case it is not relevant whether it enhances the usefulness of economic evaluations for those making funding decisions outside the UK. Further, the NICE reference case makes explicit that elements such as the perspective and health state valuations must be drawn from local contexts. NICE does, however, have an interest in generalisability inasmuch as it improves the applicability of economic evaluations to inform its own decisions where the economic evaluations have been conducted using populations outside England and Wales. NICE has published an “Applicability Checklist” which is a set of standard criteria that may be used to identify whether a particular economic evaluation can be used to inform investment decisions (14).

However, the parallels between NICE and BMGF are limited: NICE is a user of economic evaluations for a defined constituency, whereas BMGF is both a user in multiple and varying constituencies as well as a funder of economic evaluations

BMGF requires economic evaluations to inform its own decisions across multiple contexts, but also needs them to be usable by other decision-makers. This means that the reference case must support as much generalisability as possible. A persuasive case for improved generalisability was made by the ISPOR taskforce (34), which identified clear and consistent reporting as a means for improving transferability across settings. Any reference case adopted by BMGF should build on the ISPOR taskforce recommendations while at the same time striking the right balance in the generalisability-applicability trade-off.

A key risk to BMGF in adopting a reference case is that researchers may choose not to seek funding for economic evaluations that are contingent on the use of a reference case. In this instance, BMGF and partners (such as GAVI Alliance) may find it difficult to secure high-quality researchers. However, this is only likely where researchers feel that the reference case is not improving the state of the science and/or is introducing onerous requirements. As the reference case introduced in the next section has been developed (and will continue to evolve) in conjunction with researchers, the risk of “researcher flight” is low. In addition there are significant benefits for researchers in adopting a reference case, notably the increased likelihood that research will influence policy decisions.

Another potential risk to BMGF is that the adoption of a reference case will incur costs associated with its development, promulgation and maintenance. While it is a matter for BMGF to determine whether the financial imposts of adopting a reference case are worthwhile, the potential value of well-conducted and directly comparable economic evaluations in informing decision-making in local and global health is high. The costs of developing and maintaining a reference case are likely to be relatively modest.

If BMGF were to decide not to mandate the use of a reference case for projects that it funds, it is unlikely any other entity would have an interest in developing a mechanism for improving the usefulness of BMGF studies, or have the capacity to impose a requirement for the use of a reference case to serve as a global public good. In principle, local decision-makers could (and do) devise country-specific reference cases, however enforcing their adoption might be challenging across multiple funders of studies, and even if successful would only contribute to within country comparability. Also, in many LMICs, the funds and expertise to undertake such projects would be very limited.

The case for adoption of a reference case by BMGF is clear. However, the complexity of the funder/user ecosystem in which BMGF operates means that the adoption of a reference case will require a series of initiatives to ensure that any unintended consequences are minimized. There are substantial gains to be made in global health decision-making if BMGF is successful in implementing this initiative. As it is likely that benefits of the adoption and use of a reference case by researchers and policy makers will take time to be realized, long-term funding and support of this initiative will be required for the success of the endeavour.

This section has presented the economic evaluation ecosystem in low and middle income countries, the risks and benefits of reference cases in general, and the specific challenges and opportunities for BMGF in adopting a reference case. These factors have influenced the development of the proposed Gates-RC, which is presented in the next section.

## Section Three: Meeting the need - the Gates Reference Case

### 1. Preamble to the reference case

#### 1.1 What is the Gates Reference Case?

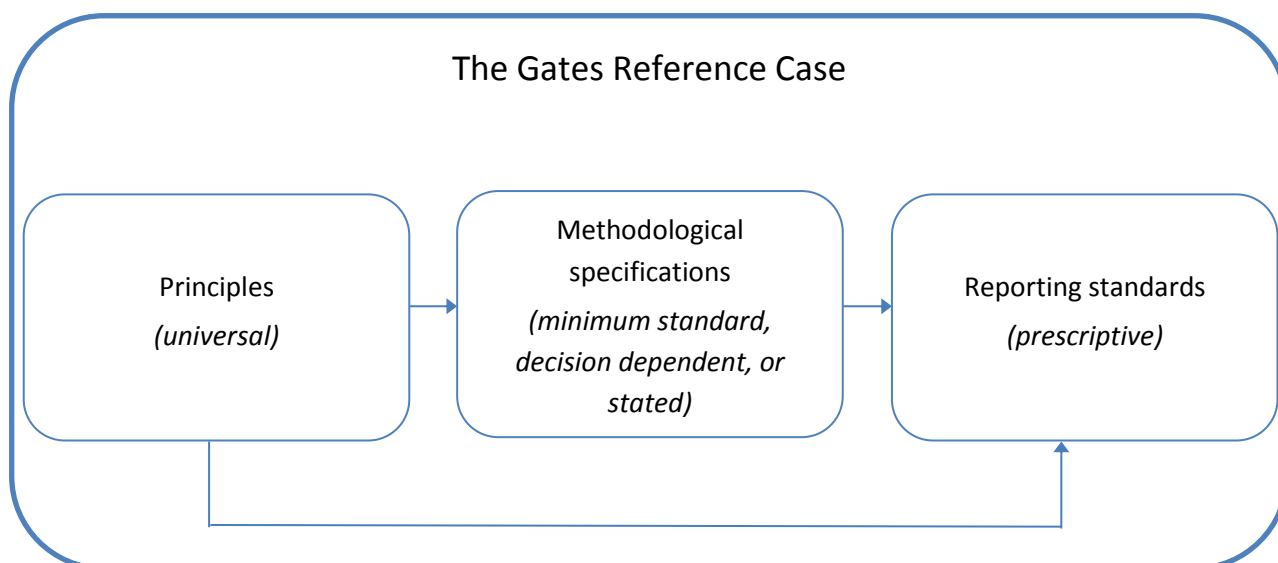
The proposed Gates Reference Case (Gates-RC) is a set of principles, methodological specifications and reporting standards to support health economic evaluations funded by the Bill and Melinda Gates Foundation (BMGF). It is intended that the introduction of the Gates-RC would help to achieve four aims:

- Routine application of certain fundamental principles by researchers and decision-makers to optimise the value of economic evaluations in informing good decisions in health;
- Use of methods that adhere to the same fundamental principles to facilitate comparison of the quality and relevance of economic evaluations when used to inform decisions in different contexts;
- A minimum standard of methodological quality to ensure that economic evaluations are fit for purpose; and
- Minimum reporting standards to ensure clarity and transparency of economic evaluations and to improve the comparability of the content and results of economic evaluations in different contexts.

The process of identifying and assembling common principles to support standards for economic evaluations is not new. The US Panel on Cost-Effectiveness in Health and Medicine first introduced the concept of a 'reference case' – a standard formulation of components of the cost effectiveness ratio which sought to 'improve the comparability of cost effectiveness analyses designed to inform decision-making more broadly, while allowing analysts the flexibility to design studies that answer issues specific to a particular problem or industry' (7). As noted above, several entities have subsequently produced standardised methodologies and reference cases relevant to their objectives. These include the National Institute for Health and Care Excellence (NICE) in England and Wales (31), the Health Interventions and Technology Assessment Program (HITAP) in Thailand (35), and the Guide to Cost-Effectiveness Analysis produced by the World Health Organisation (29).

While the proposed Gates-RC will differ in significant ways, it will nevertheless build on the theoretical foundations of its predecessors. Like other reference cases, the Gates-RC should be considered an aid to thinking, by offering a systematic framework of assessment for making decisions. In this way it should not be applied dogmatically, but considered a tool to support policy level decision-making. It will seek to optimize the use of existing evidence and will prescribe criteria for improving the evidence base; it will minimise the application of specific value judgments and policy parameters, but ensure that relevant values and policy parameters are captured in the decision-making process. The Gates-RC will also expose unresolved methodological and substantive issues for explicit discussion in specific contexts, such as research and administrative feasibility at the local level.

Figure 16 below shows the components of the Gates-RC. The reporting standards are informed by method specifications, which in turn are informed by principles.



**Figure 16 - The Gates Reference Case**

A principle-based approach has been taken for the Gates-RC as it is intended for use across multiple constituencies and investment types, and by multiple decision-makers (Table 6). This enables principles and prescriptive reporting standards to be set, but encourages decision-appropriate methods with minimum standards.

**Table 6 - Examples of constituencies, investment types and decision-makers for which the Gates-RC could potentially be applied**

Decision-makers	Investment types	Constituencies
Governments (ministries of health)	Medical technologies	Supra-national
BMGF (direct investment decisions)	(eg. HIV treatments, vaccines)	Regional
Partners of BMGF (eg. GAVI Alliance)	Public health initiatives	National
Other global or regional funders	Program evaluations	Provincial
		Local

The principles of the reference case describe how to undertake economic evaluations that are fit for purpose, but do not specify particular metrics or parameter values. The methodological specifications list the particular metrics or parameter values, and are a non-exhaustive set of options that enable the economic evaluation to adhere to the principles. While some methodological specifications represent minimum standards to be followed to ensure the economic evaluation is fit for purpose, most are decision dependent and not prescriptive. This allows a choice of methods appropriate to the decision problem in particular instances, such as whether to apply a static or dynamic model in an economic evaluation of an intervention in infectious disease. That said, some methodological specifications denoted as 'stated' are specified explicitly to facilitate comparability across multiple analyses. Examples of stated specifications are the outcome measure in the base case analysis (DALYs averted) or the discount rate for costs and effects (3%). The balance between

minimum standard specifications, decision-dependent specifications, and stated specifications is a careful trade-off that attempts to optimise the potential benefits of adopting a reference case (quality, timeliness, inter-context generalisability, user applicability – see section 2), and minimise the potential risks (methodological rigidity and burden on researchers).

Reporting standards require clear statements on certain aspects of the economic evaluation to be included in the reported analysis or publication and will maximise reproducibility and comparability. The principle *“an economic evaluation should be communicated clearly and transparently to allow the decision-maker to interpret the methods and results”* (see Table 7 below), underpins the reporting standards.

## **1.2 Principles and methods**

The principles stipulate those elements that should be considered when deciding whether and how an intervention should be used in specific populations and contexts. The principles should apply whether the economic evaluation uses highly sophisticated quantitative methods or more limited analyses based on sparse evidence.

No quantitative analysis, however sophisticated or assiduously undertaken, can capture all aspects of social value or reflect all reasonably held views on the significance and relevance of available scientific, technical, and clinical data (scientific value judgements), not least because social value judgements and scientific value judgements are by nature contested. The question is which methods can best inform the decision-maker and which offer the most useful starting point for a deliberation that is accountable to reason, evidence and widely held social values?

Clarity about principles offers many benefits for informing methodological choice. Specifically, being clear about what principles are important provides a means for judging the extent to which particular methods of analysis are decision-appropriate. However, the choice between appropriate methods and the level of sophistication with which they are implemented will depend on the type of evidence and resources available and the feasibility of using particular methods in particular contexts. Further, the selection of methods is often a balance between methods that are ideally suited to the decision problem, and those that are feasible as a result of limited time, resources, evidence and specialist expertise.

The Gates-RC should not and does not specify what the outcome of the trade-off between the ideal and the feasible ought to be in different contexts, but instead defines principles that can guide the researcher and decision-maker when determining whether an economic evaluation will adequately inform the decision problem. Similarly, the Gates-RC does not prescribe a ‘one-size-fits-all’ methodological approach irrespective of the context in which the economic evaluation is being carried out. However it does require explicit justification of the trade-offs made in the planning and conduct of an economic evaluation and how these might impact the estimates of cost-effectiveness.

Applying a principle-based reference case supports:

1. Users (LMICs payers as well as global and national donor agencies) in making judgements as to the extent to which the methods used meet their needs and what other considerations, evidence, or judgements might be required;

2. Researchers in choosing between alternative methods and justifying these choices when designing and conducting research;
3. Funders in determining whether proposed research will meet their needs in a given context, including a judgement (for funders, peer reviewers and users) as to whether the selection of methods represents a reasonable trade-off between the needs of decision-making and the constraints faced. This can also inform what other evidence or analysis might be useful in making proposed research better able to address these needs.

### **1.3 What are good decisions?**

Good resource allocation decisions are those that best meet agreed social objectives using currently available resources. In health, a good decision should involve comparing the additional health benefits of an intervention with the health likely to be lost (or other health benefits foregone) elsewhere as a consequence of any additional costs. It is because of this, the health opportunity cost, that decision-makers should consider the additional costs of an intervention as well as the health consequences of those additional costs. A cost-effectiveness threshold that reflects the health opportunity cost makes this consideration explicit, and although there are a variety of methods used to determine the appropriate threshold in particular contexts (36), a common and robust methodology for determining the benefits and costs of an intervention improves the capacity for good decision-making.

Although health benefits are not the only socially valued objectives, they are the primary objective of health interventions. The Gates-RC allows for consideration of costs and benefits that fall outside health system to be included (but reported separately, see section 2.7), as a good decision will take account of any significant trade-offs between health and other societal costs and benefits.

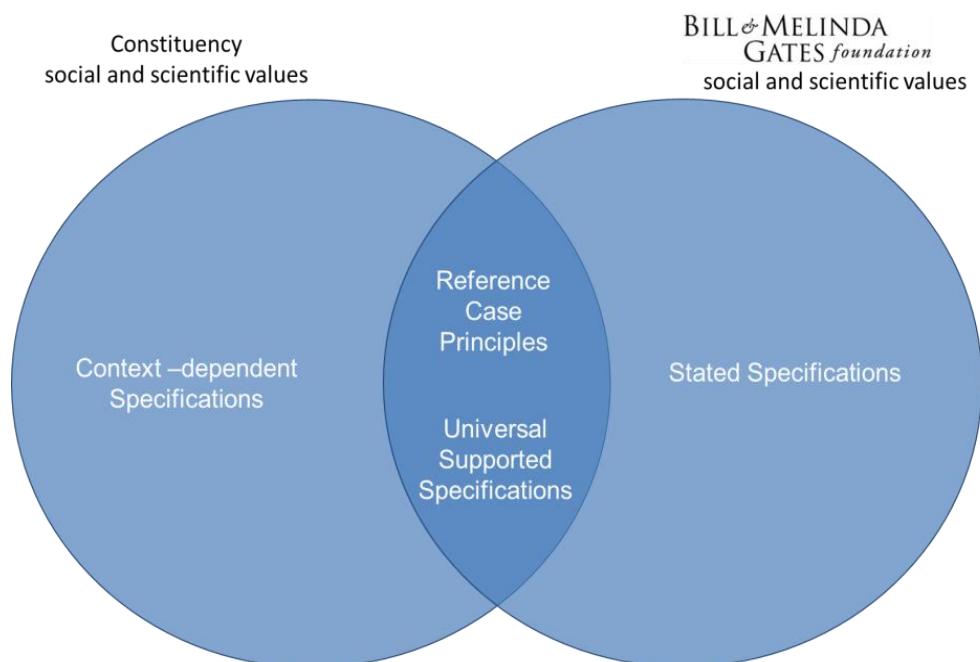
A 'good' decision is also one that is accountable to reason, to scientific value judgements about the evidence available and to the type of social value judgement required when making any decision with societal impact. Thus clear, consistent, and transparent reporting of analyses is necessary so that decisions can be properly scrutinised by relevant stakeholders and constituencies.

### **1.4 Who are the decision-makers?**

A range of potential decision-makers may make use of the Gates-RC (as shown in Table 6 above). Unlike country-specific reference cases, which may support a single decision-maker considering similar types of investment and serving a well-defined constituency (the local population directly affected by the decision), it is intended that the Gates-RC will be used by many decision-makers, across various investment types, and in multiple constituencies.

The value of an intervention will be realised in a specific context for a particular population. Ideally, the constituency of the decision-maker will include the population for whom the intervention will be used. For this reason the Gates-RC ought to reflect the social and scientific values of decision-makers who are directly responsible for and accountable to the population to be affected by a particular decision. See Figure 17 below, which shows how the social and scientific values of BMGF and the values of constituencies have informed elements of the Gates-RC.





**Figure 17 - Social and scientific values of BMGF and constituencies**

When decisions about the availability of an intervention are to be made by other types of decision-makers (eg. a global funder making a decision to fund an HIV treatment program in a particular region) the decisions may impact multiple constituencies. These populations may express different scientific judgements about how evidence should be applied and different social judgements about which values should be taken into account. A decision that impacts multiple constituencies still requires context specific information, because an intervention applied at a national or supranational level is directly impacted by its value at the local level.

That is not to say that separate economic evaluations must be undertaken in all constituencies in order to make a decision that effects multiple populations, but it does mean that there are common elements to the conduct of economic evaluations which should be informed by the needs of local constituencies, irrespective of the ultimate decision-maker.

### **1.5 Comparability – benefits and limits**

Comparability of the results of economic evaluations can be achieved through the rigid application of the same methodology; however, if the methodology is not appropriate to the decision problem then the result may not be informative. Comparability of economic evaluations is valuable only to the extent to which it supports and informs valid and consistent decisions. It is therefore best achieved through the application of common principles and reporting standards rather than by specifying that the same methods be used in all circumstances.

### **1.6 Gates Reference Case Compliance (Gates-RCC)**

Economic evaluations can assert compliance with the Gates-RC if the choice of methodology and design follow the principles of the Gates-RC, apply minimum methodological specifications, and are documented according to specified reporting standards.

Assessing Gates Reference Case Compliance (Gates-RCC) will ensure that BMGF-funded economic evaluations achieve an acceptable standard of usefulness and comparability. While the Gates-RCC

requires a principle-driven approach to the design, conduct and reporting of economic evaluations, it is not intended to be restrictive. Various methodological options may be employed and still reflect adherence to the Gates-RC principles and therefore achieve Gates-RCC. Ultimately, it is envisioned that Gates-RCC will be an indicator of economic evaluations that are methodologically sound and able to inform good decisions.

Researchers will need to consider whether adherence to the various Gates-RC methodological specifications is appropriate to the particular decision problem at hand (eg. whether discounting health effects and costs at a specific rate annually gives an accurate estimate of the incremental costs and effects of an intervention). Although methodological variation is encouraged where appropriate, to claim Gates-RCC it will be necessary to either:

1. present analytical scenario(s) where the stated specifications have been used (the Gates base case); or
2. use alternative specifications, but include an explanation of why the RC specifications were not appropriate, and the projected impact of divergence on the results of the economic evaluation.

Use of the Gates-RC to inform the conduct of economic evaluations not funded by BMGF is welcomed if appropriate to the decision-making process. Economic evaluations should claim Gates-RCC only where researchers can clearly demonstrate that the economic evaluation adheres to the Gates-RC principles and is documented in accordance with Gates-RC reporting standards.

### **1.7 Gates Reference Case Validation (Gates-RCV)**

The limitation of Gates-RCC is that it is self-assessed and reported and could give rise to claims of compliance that might not be borne out if assessed by an independent party. A valuable extension of the self-reported Gates-RCC would therefore be independent validation. This could involve a panel of experts that would assess an economic evaluation for its adherence to the principles of the Gates-RC. A similar framework to assessing Gates-RCC would be used; however Gates-RCV would require substantially more funding and a more comprehensive and consistent assessment process.

## 2. Principles of Economic Evaluation

*Table 7 Principles of Economic Evaluation*

Statement of principles	
1	An economic evaluation should be <b>communicated clearly and transparently</b> to enable the decision-maker(s) to <b>interpret</b> the methods and results.
2	The <b>comparator(s)</b> against which costs and effects are measured should accurately <b>reflect the decision problem</b> .
3	An economic evaluation should consider <b>all available evidence relevant to the decision problem</b> .
4	The <b>measure of health outcome</b> should be <b>appropriate to the decision problem</b> , should capture <b>positive and negative effects on length of life and quality of life</b> , and should be <b>generalisable</b> across disease states.
5	All <b>differences</b> between the intervention and the comparator in <b>expected resource use and costs</b> of delivery to the target population(s) should be incorporated into the evaluation.
6	The <b>time horizon</b> used in an economic evaluation should be of <b>sufficient length</b> to capture all costs and effects <b>relevant to the decision problem</b> ; an appropriate <b>discount rate</b> should be used to <b>discount cost and effects to present value</b> .
7	<b>Non-health effects</b> and <b>costs associated with gaining or providing access to health interventions that don't accrue to the health budget</b> should be identified where <b>relevant to the decision problem</b> . All costs and effects should be <b>disaggregated</b> , either by sector of the economy or to whom they accrue.
8	The <b>cost and effects of the intervention on sub-populations</b> within the decision problem should be <b>explored</b> and the <b>implications</b> appropriately <b>characterised</b> .
9	The <b>uncertainty</b> associated with an economic evaluation should be appropriately <b>characterised</b> .
10	The <b>impact</b> of implementing the intervention on the <b>health budget and on other constraints</b> should be <b>identified clearly and separately</b> .
11	An economic evaluation should explore the <b>equity implications</b> of implementing the intervention.

## 2.1 Transparency

### 2.1.1 *The principle*

*An economic evaluation should be **communicated clearly and transparently** to enable the decision-maker(s) to **interpret** the methods and results.*

### 2.1.2 *Why transparency is important for decision-making*

The aim of an economic evaluation is to inform resource allocation decisions. The most methodologically robust economic evaluation, constructed from the soundest evidence available will not be informative if the conduct and results of the economic evaluation are not reported clearly and transparently. Clear and transparent economic evaluations can also improve the transparency of the decision-making process, and consequently improve the accountability of decision-makers to stakeholders in the decision.

Clear and transparent reporting also improves transferability of economic evaluations, as research undertaken in particular contexts may be used to support decision-making in others. Even where the overall results of the economic evaluation may not be generalisable, aspects of the analysis may still inform analysis in other contexts.

A fundamental element of good scientific practice is that results are reproducible. Clear and transparent reporting enhances the capacity of other researchers to reproduce the results.

### 2.1.3 *Method specification*

Part 3 details the Minimum Reporting Standards, which outline those aspects of an economic evaluation that must be reported to ensure minimal compliance with the transparency principle. Minimum Reporting Standards do not impose any additional methodological burden on researchers as they draw on information and data that must be compiled in the course of the economic evaluation.

## 2.2 Comparators

### 2.2.1 *The principle*

*The **comparator(s)** against which costs and effects are measured should accurately **reflect the decision problem**.*

### 2.2.2 *Why the comparator is important for decision-making*

Identifying the comparator against which costs and effects will be measured is an extremely important element of an economic evaluation as it is critical to ensuring that the economic evaluation accurately informs the decision problem facing the decision-maker.

The choice of comparator determines the comparative costs and benefits associated with the intervention being considered and will therefore drive the incremental cost effectiveness ratio (ICER).

If the comparator does not reflect the decision problem, the economic evaluation will not be applicable to the decision-maker, as it will not represent the information needed to inform the particular decision. Use of a comparator that does not reflect the decision problem could lead to an inaccurate determination of cost-effectiveness, and facilitate bad decisions.

### **2.2.3 Method specification**

The method for determining relevant comparators may include:

- The interventions that are currently available to the population (therapies in routine use) as defined in the decision problem;
- “Do nothing” – i.e. comparing the new intervention to best supportive care (no intervention);
- Current “best practice”;
- The treatment or practice most likely to be replaced if the new intervention is adopted.

Regardless of the choice of comparator, it is imperative that the incremental costs and effects informing the analysis reflect the decision problem. Comparative analysis of therapies currently in routine use should form the base case, with additional analysis exploring “do nothing” as a comparator as a minimum requirement.

The most appropriate comparator is not always immediately obvious. Comparators may not always be alternative interventions, but can include different ways of administering the same intervention (such as different regimens or treatment sequences). The place of an intervention in a care pathway will also influence the choice of relevant comparators.

#### **2.3.1 The principle**

*An economic evaluation should consider **all available evidence relevant to the decision problem.***

#### **2.3.2 Why evidence is important for decision-making**

Failure to draw on all relevant and available evidence when undertaking an economic evaluation will potentially introduce bias of unknown direction, limiting the capacity of the economic evaluation to inform a “good” decision.

Evidence can be broadly defined as factual material necessary to make an informed decision, and encompasses any information that will be used to qualitatively or quantitatively inform the design, results and conclusions of an economic evaluation. Evidence will be used to inform all elements of an economic evaluation, including the unbiased estimate of the mean clinical effectiveness and the costs and resource use of the interventions being compared.

Judgement may however be necessary as to what constitutes 'all relevant and available evidence'. The researcher must apply this judgement in a systematic and transparent way when designing the economic evaluation in order to minimise bias. The decision-maker should also assess whether an economic evaluation contains all relevant and available evidence when deciding if it is applicable to the decision problem.

#### **2.3.3 Method specification**

The approach used to consider the relevance and applicability of available evidence to the decision problem should be determined before evidence gathering begins.

This methodological specification does require all economic evaluations to ensure that all parameters are informed by all available and relevant evidence, where the judgements made to determine

availability and relevance should be made in a transparent and systematic way. It is not intended that this methodological specification would prevent researchers conducting single-study or within-trial based economic evaluations, as many aspects of the decision problem may be limited to a specific trial of an intervention in a particular context and time. Researchers conducting single-study or trial based analyses are still required to outline how the single study or trial considers all available evidence relevant to the stated decision problem.

While the budget and time available for the study are relevant in determining the feasibility of the economic evaluation and scope of the decision problem, these should not influence any determination of the scope of the relevant evidence. That said, while it is important that a systematic review of the literature is undertaken to obtain estimates of the clinical effects of the intervention and its comparator(s), for some other parameters the collection and synthesis of all information may be prohibitively expensive or time-consuming. In these instances a transparent judgement should be made about the likely implications of not including missing information in the economic evaluation. Where feasible, researchers should explore the implications of alternative judgments about the quality and relevance of evidence (eg. disease natural history or progression and treatment effects). This could include presenting different scenarios that represent different judgements about which evidence ought to be included. The justification for each should be clearly expressed so its plausibility can be properly considered. This should draw on the application of accepted principles of clinical epidemiology to the available studies, indications of inconsistency within the analysis, knowledge of the natural history of disease, and reference to other external evidence where appropriate.

Researchers should clearly state when the evidence available to inform aspects of the economic evaluation is weak or unavailable. This allows the decision-maker to make a judgement on the acceptability of the evidence in informing the decision (see section 2.9 on uncertainty).

## **2.4. Measure of outcome**

### **2.4.1 The principle**

*The **measure of health outcome** should be **appropriate to the decision problem**, should capture **positive and negative effects on length of life and quality of life**, and should be **generalisable across disease states***

### **2.4.2 Why the measure of health outcome is important for decision-making**

It is important to use a measure of health outcome that is broad enough to capture all socially valued aspects of health and is applicable across investment types.

In scenarios where the scope of the decision problem is limited to interventions and comparators that impact either length of life *or* quality of life, it is still appropriate to use a measure that captures length *and* quality of life, as this allows proper consideration of the opportunity costs of investing in the intervention.

Using a non-disease specific health outcome measure (i.e. one that is generalisable across disease states) allows consideration of opportunity costs for the entire health sector, and facilitates comparisons across investment types. A disease-specific measure limits the ability of the decision-maker to make reasoned trade-offs between competing investments in different disease states, and can undermine comparability and consistency in decision-making.

### 2.4.3 Method specification

In the Gates-RC, Disability-Adjusted Life Years (DALYs) averted should be used as the measure of health outcome. The DALY provides a measure of both quality and length of life and is generalisable across different disease and therapeutic areas. This metric is frequently used in resource allocation decisions in health in low and middle income countries, and the review reported in section one of this report showed that the DALY averted was the most frequently used length and quality of life metric in economic evaluations funded by the BMGF and published since the year 2000 in LMIC.<sup>5</sup> By contrast, the Quality-Adjusted Life Year (QALY) is a frequently employed outcome metric in economic evaluations conducted in high income countries.

The DALY is the outcome measure required in the Gates-RC<sup>6</sup> to provide continuity with current practice and familiarity to decision-makers in LMICs, and to complement large-scale LMICs analyses funded by the BMGF<sup>7</sup>. Unlike the QALY, the DALY does not require context-relevant health state valuation estimates. If the Gates-RC were to specify the QALY as the preferred outcome measure for all base-case analyses, there is concern that it could undermine the feasibility of economic evaluations in many LMIC settings.

In circumstances where DALYs cannot be estimated, the QALY or other multi-dimensional metric that captures length and quality of life should be used, however clear justification of why DALYs cannot be utilised should be provided. Where other metrics (including the QALY) are available in addition to the DALY, these may be presented as additional (non-Gates-RC) analyses.

While there may be aspects of value that are relevant to the decision problem that are not captured by the DALY measure, other uni-dimensional metrics that capture either length or quality of life, or are disease-specific may be used. However, these should be presented in addition to the DALY, and the limitations of these outcome measures should be considered and discussed.

## 2.5. Measurement of costs

### 2.5.1 The principle

*All differences between the intervention and the comparator in expected resource use and costs of delivery to the target population(s) should be incorporated into the evaluation.*

*(NB this section is concerned with the costs of delivering interventions, regardless of whether they are incurred by public sector bodies (eg. ministries of health), donors, or other organizations involved in the delivery of health interventions. The section should be read in conjunction with Section 2.7 which deals with 'where' costs fall and costs outside the direct costs of health care delivery).*

---

<sup>5</sup> In the vaccine, malaria, TB and HIV/AIDs programme areas

<sup>6</sup> There is uncertainty over which measure of health outcome will be of most use to decision makers (37) and will most fully reflect the principles of the Gates-RC. The DALY has primarily been selected for base-case analysis to allow consistency and generalisability. This stated methodological specification may be amended in future iterations of the Gates-RC subject to the evidence base, expressed needs of decision makers, and the objectives of the Gates-RC.

<sup>7</sup> Such as the work of the Disease Control Priorities Network (DCPN) [www.dcp-3.org](http://www.dcp-3.org)

### **2.5.2 Why costs and resources utilisation are important for decision-making**

Decision-makers need to know the resource use and costs associated with different alternatives because more costly alternatives result in foregone benefits of other interventions (and health opportunity costs), and less costly alternatives can free financial resources for investment in other interventions. Costs and resource use do not need to be included in cases where they do not differ between evaluated alternatives, as this will not impact on the difference in cost between alternatives. Examples of this type of cost could be some central level management costs. However, caution should be taken to ensure there is no significant cost difference before resource use or costs are excluded.

Overall costs of interventions (excluding costs that do not vary across alternatives) should be reported as a key component of cost-effectiveness. Where data are adequate, costs of resource inputs to deliver interventions should also be reported. In addition to reporting costs, quantities of resources should be reported separately from their unit costs/prices to help decision-makers assess whether quantities used are appropriate and valid within their jurisdictions, and whether unit costs/prices used in the evaluation are still relevant at the time a decision is made.

In some cases resource items may have been donated, or their costs may fall on the budgets of more than one organization involved in the delivery of health interventions (including international organizations). All resource items involved in the direct delivery of health interventions should be costed because there will always be opportunity costs, even if these fall in other jurisdictions (eg. if a country attracts international funding for the delivery of an intervention). Decision-makers may also be concerned about the source of funds. See Section 2.7 on reporting costs falling on more than one organization's budget.

The average unit costs of an item of resource use may depend upon the scale at which a new intervention is delivered and the scope of other interventions delivered concomitantly. For example, the cost of each visit to a clinic nurse may differ depending on overall patient throughput in that clinic (scale), as well as on other interventions delivered at the clinic (scope). Average costs that fall (rise) with increasing scale and scope of delivery are called economies (diseconomies) of scale and scope.

Economies of scale and scope may be important and should be incorporated when feasible, particularly when alternatives are likely to differ in their scale and scope of implementation (see Section 2.2 on comparators and Section 2.10 on budget impact). However, in many cases, data from within a jurisdiction will be inadequate to reasonably establish this. Other social objectives may also be important when alternatives involve delivery at different scales and in different locations (eg. if an evaluation involves one comparator being delivered in a community or primary health care setting while another is delivered in a hospital setting). Caution should therefore be applied when applying cost functions if these cannot be supported with reliable evidence, or when other non-health effects may also have social value (see Sections 2.7 on non-health effects and 2.11 on equity).

Primacy should be placed on the transparency, reasonableness and reproducibility of cost estimates, so that different decision-makers can assess whether the results are generalisable to their jurisdictions.

### **2.5.3 Method specification**

Costs should be estimated so that they reflect the resource use and unit costs/prices that are anticipated when interventions are rolled out in real health care settings. Clinical trial protocol-driven



costs not anticipated with rollout should be excluded. Conversely any costs not incurred in a study setting but anticipated in real health care settings should be included.

Overall costs of interventions should be reported as well as costs of resource inputs. In addition, wherever possible it is useful to report quantities of resources separately from their unit costs/prices. In some cases top-down facility level cost estimates provide a useful source of data, particularly if the available resource use and unit cost/price data are insufficiently granular.

Capital and fixed costs can be annualized over the period of implementation, but decision-makers should also consider when costs are likely to be incurred (see Part 2.10 on budget impact).

It is recommended that where possible researchers corroborate their cost estimates with actual costs incurred when implementing the intervention(s) under evaluation, or other similar interventions, in real health care settings, for example using data from feasibility studies or pragmatic trials. This 'reality check' will assist the users of economic evaluations to relate the findings to current practice and costs. Where notable differences between predicted (modelled) and actual costs exist, reasons for these differences should be explored.

All resource items involved in the direct delivery of health interventions that are expected to differ between alternatives should be costed. This includes donated inputs (see Part 2.7 on reporting costs falling on more than one organization's budget). Any resource items that do not differ across alternatives may be excluded.

Economies of scale and scope that are expected with the delivery of interventions should be estimated and incorporated where feasible (NB Part 2.2. recommends interventions are evaluation at different scales of implementation). However, these must be based on reliable data from the jurisdiction concerned. Cost functions should not be imposed if unsupported by reliable evidence.

The mechanism of delivering an intervention is not set exogenously – different delivery mechanisms are usually feasible and the choice of delivery mechanism should be consistent with the overall objectives of the health systems. Researchers should consider heterogeneity of beneficiaries (see Part 2.8), impacts on other budgets, including on individuals (see Parts 2.7, 2.11), and equity considerations (Part 2.10) when using cost functions to evaluate alternative delivery mechanisms.

Costs should be reported in US dollars and in local currency, and any costs that are estimated in other currencies should be converted to US dollars and local currency. The date and source of the exchange rate used should be reported, as well as whether the exchange rate is unadjusted (real) or adjusted for purchasing power parity (PPP). Further iterations of the Gates-RC will contain specifications regarding the use of PPP and real exchange rates.

## **2.6. Time horizon for costs and effects**

### **2.6.1 The principle**

*The **time horizon** used in an economic evaluation should be of **sufficient length** to capture all costs and effects **relevant to the decision problem**; an appropriate **discount rate** should be used to **discount cost and effects to present value***

### **2.6.2 Why time horizon for costs and effects are important for decision-making**

The time horizon is the period over which the costs and effects of the intervention and comparators are calculated. The time horizon used in an economic evaluation is important because any decision made at a point in time will have implications in terms of net intervention effects and resource use extending into the future. The economic evaluation should use a time horizon that is long enough to capture all costs and effects relevant to a decision problem.

When projecting costs and effects into the future, those costs and effects need to be discounted to reflect their value at the time the decision is being made. This ensures that the time preferences of the population affected by the decision are taken into account.

### **2.6.3 Method specification**

The nature of the interventions and comparators in the decision problem will largely define the appropriate time horizon. The time horizon will often be 'lifetime' – ie. the natural length of life of the population cohort in which the analysis is being undertaken. Confirming whether the length of the time horizon is sufficient can be achieved by monitoring the impact of time horizon changes.

It is never appropriate for the time horizon to be determined by the length of time over which evidence is available, as this would lead to incomplete information being made available to the decision-maker. Where data is not available to inform an appropriate time period, some projection of costs and effects into the future will be required.

There are divergent views of the appropriate discount rate to be used in economic evaluations. In addition, the Gates-RC will be used in economic evaluations to inform decisions across constituencies where there will be legitimate and often substantial differences in time preferences for health and wealth. However, to facilitate comparability economic evaluations adhering to the Gates-RC should use an annual discount rate of 3% for both costs and effects as a stated methodological specification.

Use of alternative discount rates is encouraged where appropriate to the decision problem and constituency. These should be presented clearly with justification for their use. In cases where the costs and effects are of particularly long duration (eg. a time horizon of more than 30 years into the future is used) the impact of lower discount rates should be explored in a sensitivity analysis.

## **2.7. Costs and Effects outside health**

### **2.7.1 The principle**

***Non-health effects and costs associated with gaining or providing access to health interventions that don't accrue to the health budget should be identified where relevant to the decision problem. All costs and effects should be disaggregated, either by sector of the economy or to whom they accrue.***

### **2.7.2 Why costs and effects outside health are important for decision-making**

Most economic evaluations of health interventions are concerned with how available healthcare resources (eg. the relevant health budget) can be allocated to maximize gains in health outcomes. This requires estimating the direct health intervention costs (accruing to the available health budget) and outcomes that result from delivery of the alternative interventions being considered. If funding an

intervention generates more 'health' than could be generated from using that funding elsewhere (ie. health opportunity costs) it is considered to be a 'cost effective' use of resources.

In addition to health outcomes and direct costs accruing to the health budget, other costs and consequences of interventions may also be relevant, depending on the context of the decision. They include wider impacts on families, communities, and other sectors of the economy (eg. on educational outcomes). They may also include other (direct and indirect) costs that are incurred in gaining access to an intervention or that result from associated health outcomes. For instance, these may include direct costs falling on individuals and families in accessing health interventions (eg. travel, out-of-pocket and care costs), indirect time costs (eg. relating to the productivity of individuals and informal carers), as well as costs falling on other sectors of the economy.

Non-health effects and costs that fall outside the health budget may be important because alternative interventions may result in different non health effects that have social value. They should therefore be included in the analysis but reported separately, with a justification for the selection of the non-health effects and an explanation of how they may be valued.

Deciding which non-health effects and costs that fall outside the health budget should be included in primary analyses is troublesome as it is not clear which costs and effects are deemed socially valuable. Where there is no consensus on how to codify societal preferences, conflicts between different elements of social value may result. A particular concern is that health resources, primarily intended to generate 'health', may be used to meet other objectives that society may or may not deem to be as valuable as health itself.

As a result of these difficulties in aggregating different effects, primary analyses should only reflect direct costs to the health budget and direct health outcomes. By presenting non-health effects separately, decision-makers are able to draw their own conclusions as to the relative merits of the different effects.

The issue of whether direct costs faced by individuals and their families should be incorporated into an analysis is also relevant. In those health systems in which a significant proportion of healthcare is funded through out-of-pocket (OOP) payments, there may be good reasons to adopt a perspective broader than that of the health care provider when direct OOP costs substitute for costs that would otherwise fall on the health budget. Researchers should take care that alternatives do not shift costs to individuals, and they may choose to incorporate direct OOP costs into primary analyses in such cases. Of central concern is the opportunity costs faced in each case and how these are likely to be valued by society (this may also include concern for financial protection).

Society often values both health and non-health effects differently depending upon who benefits (see Section 2.11 on equity). Similarly, direct health intervention costs may impose different opportunity costs depending upon who is funding the intervention. In many LMICs, health interventions rely on direct funding from multiple sources (for instance national ministries of health may fund recurrent costs; whereas international donors may fund drugs or certain technologies). In these instances donor funds (including the direct provision of drugs and health care materials) may form a significant proportion of the budget available for health. It would therefore be inappropriate for the analysis to disregard the direct impact of an intervention of donor funds; however it is important that recognition is made of the different sources of the available health budget.

For these reasons, it is recommended that direct costs, health effects, non-health effects and costs that fall outside the health sector are disaggregated, so that it is clear who are the beneficiaries and the funders of interventions. This facilitates exploration of health system constraints, budget impacts and opportunity costs (see Part 2.10), and equity issues (see Part 2.11), and enables decision-makers to make assessments of the relative values of each in their own jurisdictions.

Non-health effects can be valued and presented in different units. Valuing non-health effects monetarily has the benefit that both outcomes and costs can be represented in a common metric, but there are contentious methodological issues relating to how to appropriately monetarise outcomes. Alternatively, outcomes can be reported qualitatively or valued in other units, and costs reported monetarily. Thorough exploration of how to value non-health effects is therefore recommended.

### **2.7.3 Method specification**

The reported base case should reflect direct health care costs and health outcomes, and the analysis should adopt a disaggregated societal perspective, so that the funders and beneficiaries of health interventions can be clearly identified. Inclusion of particular costs and effects within the societal perspective may differ depending on the decision problem and context.

Direct costs incurred by funders where these costs would otherwise accrue to government health budgets, should be included in the base case. However, additional analyses should explore the impact of donor funding, and direct health care costs should be disaggregated between funders if it is known that they contribute differentially to the delivery of interventions.

OOP costs falling on individuals can be included if these displace costs that would otherwise fall on the health budget, however, the impact of excluding OOP costs should be included in sensitivity analyses.

Where there are believed to be important non-health effects and costs falling outside the health budget these should be included in the analysis but reported separately, with a justification for their selection and an exploration of the ways they can be valued. Any non-health effects and costs that fall outside the health budget that potentially conflict with other social objectives should be highlighted and discussed. For example, a particular intervention may be expected to have productivity benefits but its adoption may have an adverse impact on population equity.

Decision-makers should be made aware that interventions with positive incremental direct health costs are also likely to impose non-health opportunity costs associated with health interventions that are foregone (as interventions foregone are also likely to have non-health effects). For example, an intervention for HIV/AIDS may have non-health effects but if adopted may displace interventions for maternal health that have equal or even greater claims to generating positive social value.

Researchers should ensure that non-health effects and costs are not double counted, especially in cost-utility analyses. Double counting can occur where a particular effect (or cost) of an intervention relative to a comparator is attributed to more than one outcome measure – for example, there are debates as to the extent that productivity effects are already captured in quality of life measures.

Direct health costs should be disaggregated by funder. Both health and non-health effects should be disaggregated by characteristics of recipients and beneficiaries (see Part 2.11 on equity); and, in the case of non-health effects, the sector or area in which these are incurred.

## 2.8. Heterogeneity

### 2.8.1 *The principle*

*The **cost and effects of the intervention on sub-populations** within the decision problem should be **explored** and the **implications** appropriately **characterised**.*

### 2.8.2 *Why heterogeneity is important for decision-making*

It is important to make a clear distinction between uncertainty, variability and heterogeneity. Uncertainty (see Section 2.9) refers to the fact that we do not know what the expected effects of an intervention will be in a particular population of individuals. This remains the case even if all individuals within this population have the same observed characteristics. Variability refers to the fact that responses to an intervention will differ within the population or even in a sub population of individuals or patients with the same observed characteristics. Heterogeneity refers to those differences in response that can be associated with differences in observed characteristics, ie. where the sources of natural variability can be identified and understood. As more becomes known about the sources of variability the patient population can be partitioned into sub populations or subgroups, each with a different estimate of the expected effect and costs of the intervention, and of the uncertainty associated with these.

An exploration of heterogeneity thus enables decision-makers to consider whether an intervention should be made available to certain groups of individuals with greatest capacity to benefit (or in whom the costs of provision are worthwhile). It means they have the opportunity to make different decisions for different groups of individuals that lead to improved health outcomes overall given the available resources.

There may, however, be good reasons not to make different decisions about provision based on certain types of observed characteristics. These reasons might include: i) the difficulty and/or cost of maintaining differential access; ii) adverse equity implications; or iii) social values that would not support discrimination based on certain types of characteristics. However, even in these circumstances an exploration of heterogeneity is important for a number of reasons:

- 1) The correct assessment of the cost and effects of providing an intervention across some or all subgroups depends on the effects and costs in each group;
- 2) It enables decision-makers to consider the opportunity costs (the health foregone) of acceding to concerns about differential provision across identifiable groups, adding to consistency and accountability in the application of other social values.
- 3) It can provide a foundation for exploring the health equity implications of how an intervention might be provided, and can identify potential trade-offs between equity objectives and overall health benefits ( eg. for equity reasons it might be considered worthwhile providing an intervention to groups where provision is more costly even though this may reduce health benefits overall).
- 4) It can provide a better understanding of the distributional issues associated with an intervention. This can form the basis for further more targeted research and inform other related decisions.

### **2.8.3 Method specification**

Since any observed characteristics that affect health benefits and costs of an intervention are relevant in principle, the exploration of heterogeneity should include subgroups where there is good evidence that the relative effect of the intervention will differ (eg. pre-specified subgroups within a clinical trial). However, subgroup analysis can be considered when external evidence suggests (and there are good reasons to believe) that relative effects differ between subgroups even if they have not been pre-specified.

The exploration of heterogeneity should not be restricted to differences in relative effects between different groups of individuals or patients. It should also include exploration of characteristics that will influence absolute health effects, even where the relative effects are the same, such as differences in base line risk of an event or incidence and prevalence of a disease.

There may also be characteristics which are unrelated to relative effects or baseline risk but nevertheless influence the direct costs of provision or other health intervention costs and benefits, such as geographical location or differences in other health care provision.

The question of which sets of observed characteristics to explore should be informed by

- 1) the evidence base regarding differences in relative effect, baseline risk or other relevant characteristics, and
- 2) whether any differences are likely to have an important influence on costs and effects.

The analysis should justify how the exploration of heterogeneity has been undertaken with respect to these two considerations. A presumption that certain observed characteristics should not (or will not) be used to offer differential access to an intervention is not a justification for failing to explore the implications of heterogeneity (see Part 2.8.2).

## **2.9 Uncertainty**

### **2.9.1 The principle**

*The **uncertainty** associated with an economic evaluation should be appropriately **characterised**.*

### **2.9.2 Why consideration of uncertainty is important for decision-making**

Decisions regarding resource allocation in health are unavoidable. All decisions carry a risk that a more optimal course of action could have been selected, and so in making a decision uncertainty must be acknowledged and measured.

For the chances of reaching a good decision to be optimised, the decision-maker needs to be aware of the magnitude of uncertainty in the results of an economic evaluation. All economic evaluations contain uncertainty, and it is important that all types of uncertainty are appropriately presented to the decision-maker. The types of uncertainty include uncertainty about the source of parameters used in the analysis, the precision of those parameters, and whether models or simulations of how the costs and effects of the intervention and comparators will behave are accurate. The characterisation of this uncertainty enables the decision-maker to make a judgement based not only on a likely estimate of the incremental costs and effects of an intervention, but also on the confidence that those costs and effects reflect reality.

Characterising the uncertainty will also enable the decision-maker to be informed about courses of action that could reduce this uncertainty. This could involve delaying implementation to allow for more evidence to be garnered. In this situation, appropriately characterising uncertainty will allow the decision-maker to make an informed trade-off of the value of new information, the implications of potentially delaying treatment to patients or individuals, and irrecoverable costs associated with implementing funding for an intervention.

### **2.9.3 Method specification**

There are a number of potential selection biases and uncertainties in any economic evaluation, and these should be identified and quantified where possible. There are three types of bias or uncertainty to consider:

Structural uncertainty – for example in relation to the categorisation of different states of health and the representation of different pathways of care. These structural assumptions should be clearly documented and the evidence and rationale to support them provided. The impact of structural uncertainty on estimates of cost effectiveness should be explored by separate analyses of a representative range of plausible scenarios.

Source of values to inform parameter estimates – the implications of different estimates of key parameters (such as estimates of relative effectiveness) must be reflected in sensitivity analyses (for example, through the inclusion of alternative sources of parameter estimates). Inputs must be fully justified, and uncertainty explored by sensitivity analyses using alternative input values.

Parameter precision – uncertainty around the mean health and cost inputs in the model. Distributions should be assigned to characterise the uncertainty associated with the (precision of) mean parameter values. Probabilistic sensitivity analysis (PSA) is preferred, as this enables the uncertainty associated with parameters to be simultaneously reflected in the results of the model. In non-linear decision-models – where there isn't a straight-line relationship between inputs and outputs of a model (such as Markov models) – probabilistic methods provide the best estimates of mean costs and outcomes. Simple decision trees are usually linear. The mean value, distribution around the mean, and the source and rationale for the supporting evidence should be clearly described for each parameter included in the model. Evidence about the extent of correlation between individual parameters should be considered carefully. Assumptions made about the correlations should be clearly presented.

Where lack of evidence restricts reliable estimations of mean values and their distributions, unsupported assumptions or exclusion of parameters in a PSA will limit its usefulness to characterise uncertainty, and may give a false impression of the degree of uncertainty. For this reason, PSA is not explicitly required in all economic evaluations at this time; however any decision not to conduct PSA should be clearly and transparently explained in the analysis. Future iterations of the Gates-RC will provide further specification on the application of PSA.

## **2.10. Impact on other constraints and budget impact**

### **2.10.1 The principle**

*The **impact** of implementing the intervention on the **health budget and on other constraints** should be **identified clearly and separately**.*

### **2.10.2 Why assessing budget impact is important for decision-making**

It is important to determine the net total costs involved in the deployment of a health intervention on a particular scale (see Section 2.8), as these are also a measure of the value of what must be foregone.

The costs of an intervention (even when capital investment is not required) are unlikely to be evenly spread over time, but often have large initial costs offset by later health benefits and at times, cost savings. Decision-makers responsible for annual budgets must assess the timing of the impact as well as the magnitude of the expected incremental costs when deciding if the benefits of an intervention exceed the health opportunity costs. This becomes especially important when later health benefits or cost savings are uncertain, since implementation will require the commitment of resources that may be unrecoverable should subsequent evidence suggests the intervention is not worthwhile (or not cost effective) and should be withdrawn (see Part 2.9).

In addition to expenditure constraints decision-makers may be subject to other infrastructural or resource limitations, such as lack of laboratory capacity or workforce limitations. Decision-makers (at national, regional, or local level) must be able to assess the impact of an intervention in each of these domains to properly determine whether the benefits exceed the health opportunity costs. This may also facilitate some consideration of which constraints have the greatest impact, and the potential value of policies that modified these, such as removing restrictions on the use of donated resources or increasing investment in training health workers.

Since non-health benefits and costs do not impact health budgets or other constraints on health care, they should be assessed separately (see Part 2.7).

### **2.10.3 Method specification**

Budget impact should be presented in a manner that is relevant to the decision problem and the needs of the intended decision-maker. The budget impact should be disaggregated and reflect the costs to all parties as a result of implementation of the intervention (cost outputs). This includes (but is not limited to) impact on government and social insurance budgets, households and direct out of pocket expenses, third-party payers, and external donors. Budget impact should be projected annually for a period appropriate to the decision problem.

## **2.11. Equity implications**

### **2.11.1 The principle**

*An economic evaluation should explore the **equity implications** of implementing the intervention.*

### **2.11.2 Why equity implications are important for decision-making**

The equity implications of implementing an intervention within a given constituency are important, because decisions concerning resource allocation in health frequently reflect considerations other than



efficiency. Important equity considerations may include issues such as whether equal access is given to those in equal need, whether resources are distributed fairly to those with different levels of need, or recognising that interventions such as smoking cessation programmes may simultaneously increase population health and health inequalities. Limiting an economic evaluation to a determination of cost-effectiveness across a population as a whole ignores differences in the capacity to benefit and/or in access to care, and may prevent the decision-maker from appropriately considering the differential impacts of a decision on different subgroups within the population.

It is worth noting that in adhering to certain principles in the Gates-RC some equity implications will be considered implicitly. For example, exploring heterogeneity may involve a consideration of distributional implications of implementing an intervention. However, adherence to the other principles of the Gates-RC will not generally be sufficient to ensure that the equity implications of a decision problem have been adequately explored. For this reason, exploration of equity is a principle that should be addressed in its own right in BMGF-funded economic evaluations.

### ***2.11.3 Method specification***

There are many dimensions to assessing the equity implications of a proposed intervention. Methods employed may be qualitative, such as the seven-step analysis (as used in Miljeteig 2010<sup>8</sup>), or may involve the quantitative assessment of distributive impact and expected trade-offs, using established mechanisms such as the Atkinson index or Gini index. At the most basic level, an exploration of the equity impact may involve a description of particular groups within the constituency that may be disproportionately affected (positively or negatively) by a decision. Adherence to the equity principle is not, however, simply a matter for reporting results. Equity implications should be considered at all stages of an economic evaluation, including the design, analysis and reporting stages. This is important for all types of economic evaluation, including those undertaken along-side clinical trials.

It is not proposed that the Gates-RC be prescriptive about the methods to be used or how equity implications are presented. However, as experience increases, subsequent iterations of the Gates-RC might provide greater guidance for researchers and decision-makers in considering the equity implications of resource allocation decisions in health.

---

<sup>8</sup> Miljeteig I, Johansson K A, Sayeed S A, et al. *J Med Ethics* 2010;36:473-478

### 3. Gates-Reference Case Summary Table

The Summary Table lists the principles, the associated methodological specifications and reporting standards for Gates-RC compliant (Gate-RCC) economic evaluations. The reporting standards adhere to Principle One: that an economic evaluation should be communicated clearly and transparently.

Where a stated specification exists, researchers applying the Gates-RC are required to adopt the position or numerical value stated. However, where appropriate, deviation from the methodological specification is encouraged as long as both a justification for and the implications of the deviation are clearly stated.

The reporting standard below does not replace the responsibility of individual researchers to conduct and report economic evaluations in ways that are appropriate to the context of the decision and the specific decision problem.

	Statement of principle	Methodological Specifications	Reporting Standards
1	An economic evaluation should be <b>communicated clearly and transparently</b> to enable the decision-maker(s) to <b>interpret</b> the methods and results.	<ul style="list-style-type: none"> <li>The decision problem must be fully and accurately described and the economic evaluation characterised</li> <li>Limitations of the economic evaluation in informing policy should be characterised</li> <li>Declarations of interest should be made</li> </ul>	<ul style="list-style-type: none"> <li>The decision problem should be stated, clearly identifying:               <ul style="list-style-type: none"> <li>Population (description and characteristics) in which the intervention would be used</li> <li>Intervention(s) that are being evaluated and its Comparator (see principle 2)</li> <li>Outcome that is being assessed (see principle 4)</li> </ul> </li> <li>The characteristics of the economic evaluation should be stated, clearly identifying:               <ul style="list-style-type: none"> <li>The relevance for health practice and policy decisions</li> <li>The constituency that the economic evaluation would seek to inform</li> <li>The intended user of the economic evaluation</li> </ul> </li> <li>The limitations of the economic evaluation should be transparent, including               <ul style="list-style-type: none"> <li>Limitations in the design, analysis and results</li> <li>Aspects of the economic evaluation that would limit generalisability of results to other constituencies</li> </ul> </li> <li>Declarations of interests should be reported that include:               <ul style="list-style-type: none"> <li>All pecuniary and non-pecuniary interests of the study contributors</li> <li>All sources of funding that supported conduct of the economic evaluation</li> <li>Nonmonetary sources of support for conduct of the economic evaluation</li> </ul> </li> </ul>

2	<p>The <b>comparator(s)</b> against which costs and effects are measured should accurately <b>reflect the decision problem</b>.</p>	<ul style="list-style-type: none"> <li>• At a minimum, the following comparative analysis should be undertaken: <ul style="list-style-type: none"> <li>○ The intervention(s) that is(are) currently offered to the population as defined in the decision problem as the base case comparator</li> <li>○ A “do nothing” analysis representing best supportive (non-interventional care) for the population as additional analysis</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Clear description of comparator(s) that includes: <ul style="list-style-type: none"> <li>○ Basic descriptive information including setting where comparator is administered (especially if setting is different to the intervention)</li> <li>○ Statement of availability of the comparator across the population being considered</li> </ul> </li> <li>• The difference between mean costs and effects of the intervention and chosen comparators should be reported as incremental cost effectiveness ratios</li> </ul>
3	<p>An economic evaluation should consider <b>all available evidence relevant to the decision problem</b>.</p>	<ul style="list-style-type: none"> <li>• A systematic and transparent approach should be taken to obtain evidence and make judgements about evidence exclusion</li> <li>• Estimates of clinical effect of intervention and comparator(s) should be informed by a systematic review of the literature</li> <li>• Single-study or trial-based analyses should outline how the single study or trial is a sufficient source of evidence and should ensure that the stated decision problem is specific to particular context and time of the study or trial</li> <li>• Budget and time allocated to perform an economic evaluation should not determine selection of evidence</li> </ul>	<ul style="list-style-type: none"> <li>• Describe approach used to obtain included evidence</li> <li>• Systematic review protocol and evidence search strategies should be made available</li> <li>• List sources of all parameters used in economic evaluation</li> <li>• Describe areas where evidence is incomplete or lacking</li> </ul>
4	<p>The <b>measure of health outcome</b> should be <b>appropriate to the decision problem</b>, should capture <b>positive and negative effects on length of life and quality of life</b>, and should be <b>generalisable</b> across disease states</p>	<ul style="list-style-type: none"> <li>• Disability-Adjusted Life Years (DALYs) averted should be used [<b>stated methodological specification</b>]</li> <li>• Other generic measures that capture length and quality of life (e.g. the QALY) can be used in separate analysis where information is available</li> </ul>	<ul style="list-style-type: none"> <li>• Clear description of method of weighting used to inform the DALY</li> <li>• Discussion of any important outcomes insufficiently captured by the DALY</li> <li>• If DALY not used, clear justification should be provided with description of impact of using the alternative outcome measure</li> </ul>

5	<p>All <b>differences</b> between the intervention and the comparator in <b>expected resource use and costs of delivery</b> to the target population(s) should be incorporated into the evaluation.</p>	<ul style="list-style-type: none"> <li>• Estimates should reflect the resource use and unit costs/prices that are expected if the intervention were to be rolled out to the population defined in the decision problem</li> <li>• Costs not included in study settings used to inform the analysis but would be incurred if the intervention was rolled out should be included in the base case analysis</li> <li>• All resource implications relevant to the decision problem should be costed, including donated inputs and out of pocket inputs from individuals</li> <li>• Analysis should include estimation of changes in costs estimates due to scalability</li> </ul>	<ul style="list-style-type: none"> <li>• Quantities of resources should be reported separately from their unit costs/prices</li> <li>• Capital and fixed costs should be annuitized over the period of implementation</li> <li>• Description of how the costs have been validated ( e.g. corroboration with other similar interventions in similar settings)</li> <li>• Any major differences between predicted (modelled) and realized costs should be explained</li> <li>• Implications of changes in costs due to scalability of the intervention should be reported</li> <li>• Costs should be reported in local currency and in United States dollars.</li> <li>• Costs should be converted to US\$ and local currency. The date and source of the exchange rate should be reported.</li> </ul>
6	<p>The <b>time horizon</b> used in an economic evaluation should be of <b>sufficient length</b> to capture all costs and effects <b>relevant to the decision problem</b>; an appropriate <b>discount rate</b> should be used to <b>discount cost and effects to present value</b>.</p>	<ul style="list-style-type: none"> <li>• Lifetime time horizon should be used in first instance.</li> <li>• Shorter time horizon can be used where shown that all costs and effects that are relevant to the decision problem have been captured.</li> <li>• A 3% annual discount rate for costs and effects should be used in base case analysis [<b>stated methodological specification</b>]. Additional analysis exploring differing discount rates appropriate to the decision problem should be used</li> <li>• Additional analysis should explore an annual discount rate that reflects the rate at which the government can borrow funds on the international market</li> <li>• Where time horizon used is greater than 30 years, the impact of lower discount rates should be explored in sensitivity analysis</li> </ul>	<ul style="list-style-type: none"> <li>• Clearly state the time horizon over which costs and effects are being evaluated, including additional analysis if different time horizons have been explored.</li> <li>• If lifetime time horizon not used, justification of why and impact of the different time horizon should be reported</li> <li>• Clearly state the discount rate used for both costs and effects, and include additional analysis performed with different discount rates.</li> <li>• If a 3% annual discount rate is not used, justification of why and impact of the different discount rate(s) should be reported</li> </ul>

7	<p><b>Non-health effects and costs associated with gaining or providing access to health interventions that don't accrue to the health budget</b> should be identified where <b>relevant to the decision problem</b>. All costs and effects should be <b>disaggregated</b>, either by sector of the economy or to whom they accrue.</p>	<ul style="list-style-type: none"> <li>• Base case analysis should reflect direct health costs and health outcomes; however the analysis should adopt a disaggregated societal perspective</li> <li>• Non-health effects and costs that fall outside the health budget should be included in additional analysis; the mechanism of inclusion will differ depending on the decision problem and context.</li> <li>• Where external funding or individual OOP payments substitute for costs that would otherwise fall on a health budget, these costs should be included in the base case analysis, however the impact of excluding these payments must be explored in sensitivity analysis</li> </ul>	<ul style="list-style-type: none"> <li>• Clear description of the result of the base case analysis</li> <li>• Alternative analysis exploring impact of individual out of pocket payments and external funding should be explored</li> <li>• Non-health effects and costs that fall outside the health sector should be reported and the mechanisms used to reported impact of these cost and effects should be explained and justified</li> <li>• If non-health effects and costs that fall outside the health sector are not included in the economic evaluation, justification of reasons for exclusions should be reported and estimations of the potential impact of these exclusions made</li> </ul>
8	<p>The <b>cost and effects of the intervention on sub-populations</b> within the decision problem should be <b>explored</b> and the <b>implications</b> appropriately <b>characterised</b>.</p>	<ul style="list-style-type: none"> <li>• Heterogeneity should be explored in subgroups of the population identified in the decision problem, where subgroup formation should be informed by: <ul style="list-style-type: none"> <li>○ Relevant effect of the intervention differs in different populations</li> <li>○ Characteristics of different populations that may influence the absolute health effects</li> <li>○ Characteristics that influence direct costs of provision or other associated costs such as geographical location across the constituency</li> </ul> </li> <li>• Subgroup analysis should always be determined by: <ul style="list-style-type: none"> <li>○ The evidence base regarding differences in relative effect, baseline risk or other characteristics</li> <li>○ Whether the differences are likely to have an important influence on costs and effects</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Clear reporting of: <ul style="list-style-type: none"> <li>○ subgroup characteristics and justification of why particular groups are chosen for subgroup analysis</li> <li>○ evidence base used to determine subgroup formation</li> <li>○ the costs effectiveness of the intervention in the different subgroups</li> <li>○ subgroups that have potentially important differences in costs and effects but have not been included in analysis due to lack of evidence</li> </ul> </li> </ul>

9	The <b>uncertainty</b> associated with an economic evaluation should be appropriately <b>characterised</b> .	<ul style="list-style-type: none"> <li>• The economic evaluation should explore: <ul style="list-style-type: none"> <li>○ Uncertainty in the structure of the analysis</li> <li>○ Uncertainty due to source of parameters</li> <li>○ Uncertainty due to precision of parameters</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• The effects of all types of uncertainty should be clearly reported, noting impact on final results.</li> <li>• Uncertainty due to precision of parameters should be characterised using sensitivity analysis appropriate to the decision problem.</li> <li>• The likelihood of making the wrong decisions given the existing evidence should be addressed</li> </ul>
10	The <b>impact</b> of implementing the intervention on the <b>health budget and on other constraints</b> should be <b>identified clearly and separately</b> .	<ul style="list-style-type: none"> <li>• Budget impact analysis should be performed that provides an estimate of the implications of implementing the intervention on various budgets</li> <li>• Budget impact analysis should reflect the decision problem and the constituency in which the intervention will be implemented.</li> </ul>	<ul style="list-style-type: none"> <li>• A disaggregated and annualised budget impact analysis should be reported that shows budget implications of the intervention on the following parties: <ul style="list-style-type: none"> <li>○ Government and social insurance budgets</li> <li>○ Households and out of pocket expenses</li> <li>○ Third-party payers</li> <li>○ External donors</li> </ul> </li> </ul>
11	An economic evaluation should explore the <b>equity implications</b> of implementing the intervention.	<ul style="list-style-type: none"> <li>• There are various mechanisms available for how the equity implications of an intervention should be assessed. The method chosen should be appropriate to the decision problem and justifiable to the decision-maker</li> <li>• Equity implications should be considered at all stages of the economic evaluation, including design, analysis and reporting</li> </ul>	<ul style="list-style-type: none"> <li>• The method used to incorporate equity implications should be clearly and transparently explained.</li> <li>• A minimum level of reporting should include a description of particular groups within the constituency that may be disproportionately positively or negatively affected by a decision to implement (or not implement) the intervention.</li> </ul>

## Section Four: Recommendations

The section of the report presents recommendations to BMGF arising from the various components of the Methods of Economic Evaluation Project (MEEP).

The first – and overarching – recommendation, that the BMGF should adopt a reference case, is supported by a further eight recommendations regarding initiatives to realise the gains from the reference case described in this report. The potential risks and benefits of adopting a reference case identified in section two of the report feed into recommendations two to six, which propose ways to mitigate risks and optimise benefits. Recommendations seven to nine focus on the integral role of decision-makers, and encompass initiatives to build understanding and capacity among them in order to realise the benefits of a reference case.

### **1. BMGF should adopt a reference case**

A reference case will provide substantial benefits to BMGF, in its roles as both a funder and user of economic evaluations to inform investment decisions, as an advocate for better decision-making in global health, and as a stakeholder in improving length and quality of life for the global population. The principal benefits of adopting a reference case are that it will improve the quality and comparability of results across constituencies and investment types.

It will be important for BMGF to consider the scope of deployment of the RC. This could range from issuing general guidance to researchers, to requiring all funded research to be Gates Reference Case Compliant (Gates-RCC). In addition, the detail and prescriptiveness of the reference case must be determined, in doing so striking an appropriate balance among the many attributes that contribute to usefulness in an economic evaluation: global generalisability vs local applicability, quality vs timeliness, as well as researcher flexibility and comparability.

### **2. BMGF should adopt a template for reporting economic evaluations**

A reporting template is a key component of a reference case. It promotes standardised reporting of key components of an economic evaluation and minimises administrative burden in summarising multiple economic evaluations for researchers and decision-makers. It may be applied to both existing and future evaluations and its use will facilitate on-going data collection (in addition to the Methodological Review) to inform the reference case specifications.

### **3. BMGF should establish a technical panel for submission-stage and/or for validation of economic evaluations, to check compliance with the Gates-RC and validate completed economic evaluations**

A technical panel would provide a means of ensuring appropriately qualified technical experts are checking compliance with the reference case, but would also serve as a forum for adjudicating whether departures from the reference case are reasonable. There may be instances where an economic evaluation cannot conform to all reference case specifications. In this case, technical experts could be involved in the initial stages (before funding is approved) – either to endorse departures from the reference case, (balancing loss of comparability) or to suggest ways in which researchers could comply with the reference case within context-specific constraints.

Consideration would need to be given to the make-up of the technical panel and the timeframes under which it would operate. It is not the intention of this recommendation to introduce overly burdensome requirements on researchers. However, planned appropriately, a technical panel could provide useful guidance to researchers and technical panel review could become an integral step in the funding process.

#### **4. BMGF should introduce a process or panel for monitoring and evaluating the benefits of the reference case, including defining metrics for success**

This report has identified various perceived risks and benefits to BMGF in adopting a reference case; these should be formally and quantitatively considered where feasible. A monitoring and evaluation panel would complement the work of the technical panel, and it may be appropriate for members to be common to both. It is worth noting that it may take a several years to accrue a sufficient number of economic evaluations to allow a meaningful assessment of the anticipated benefits (such as improved comparability); however others, such as usefulness to decision-makers or variation of methods, may be discernible more quickly.

#### **5. BMGF should introduce a process for on-going refinement and review of the reference case, with involvement of key experts**

The intention of this recommendation is to engage the research community in further refinements to the reference case and reporting standards. This stems from the experience of the National Institute for Health and Care Excellence (NICE). NICE published its fifth edition of its Methods for Technology Appraisal in April 2013, acknowledging that as methodologies, statistical techniques, data availability and needs of decision-makers evolve, a reference case must also evolve to support improved decision-making. Implementing this recommendation will be important to address a potential risk of the reference case: that imposing methodological orthodoxy could stifle methodological evolution. The reference case must evolve in order to optimise its relevance and usefulness.

#### **6. BMGF should develop a library of economic evaluations and models used in BMGF-funded analyses**

Timeliness in meeting the needs of decision-makers is an important measure of the usefulness of economic evaluations. In addition, an objective in adopting a reference case is to improve comparability. It is therefore important at a practical level for economic evaluations to be made available to other researchers, to avoid costly duplication and facilitate comparisons and consistency in decision-making.

While the inclusion of decision analytic models in the repository would require consultation with researchers to ensure that it is not unnecessarily burdensome, a repository with full working models has the potential to be an extremely valuable public good for researchers in low and middle income countries.

In order to avoid duplication, BMGF should discuss the potential for synergies with existing initiatives such as the Cost-Effectiveness registry ([www.cearegistry.org](http://www.cearegistry.org)) at the Center for the Evaluation of Value and Risk in Health at the Institute for Clinical Research and Health Policy Studies



at Tufts Medical Center, or the Centre for Reviews and Dissemination ([www.crd.york.ac.uk](http://www.crd.york.ac.uk)) at the University of York.

#### **7. BMGF should support training in health technology assessment in low and middle income countries, and encourage the use of a reference case**

Ultimately the reference case will be of limited usefulness if skilled resources necessary to undertake economic evaluations and interpret the results are unavailable. It is also important that economic evaluations are not used inappropriately, for example by making inappropriate generalisations across constituencies. It is therefore necessary for appropriate and targeted training to be made available.

#### **8. BMGF should work with low and middle income countries to build priority-setting institutions**

This recommendation links with HTA training support in LMICs, but the focus here is on institution building. It will therefore require the engagement of policy makers in LMICs, and the development of a longer-term engagement strategy. Priority-setting institutions could take many forms including within-country institutions, regional entities, or entities with a global focus. A potential partner in the development of priority-setting institutions is the Centre for Global Development ([www.cgdev.org/page/priority-setting-institutions-global-health](http://www.cgdev.org/page/priority-setting-institutions-global-health)).

#### **9. BMGF should develop guidance on the process of using economic evaluations to inform priority setting decisions in LMICs**

The scope of the Methods for Economic Evaluation Project (MEEP) was necessarily limited to methodologies for researchers conducting economic evaluations. However, feedback from consultation stressed the importance of coordinating improved decision-making processes with improved methodology, as the usefulness of an economic evaluation is highly contingent on the decision context within which it is used.

Therefore, to optimise the benefit that is possible from improved methodology in the conduct of economic evaluations through the use of a Gates-RC, it is necessary to seek to address processes for consideration of economic evaluations by decision-makers in LMICs, as well as those in global health that impact LMIC constituencies. The foundational work of MEEP will provide a starting point for developing guidance on the process of consideration of economic evaluations; however this represents a substantial body of work with large potential for global impact. A key partner in process development is likely to be the Centre for Global Development linking to recent BMGF-funded projects (e.g. [www.morehealthforthemoney.org](http://www.morehealthforthemoney.org))

## References

1. Drummond M, Sculpher M, Torrance G, O'Brian B, Stoddart G. *Methods for the Economic Evaluation of Health Care Programs*. (Oxford: Oxford University Press, 2005)
2. Ross J. The use of economic evaluation in health care: Australian decision-makers' perceptions. *Health Policy* 1995; 31( 2):103-10.
3. Yothasamut J, Tantivess S, Teerawattananon Y. Using economic evaluation in policy decision-making in Asian countries: mission impossible or mission probable? *Value in Health* 2009;12(Suppl3):S26-30.
4. Buxton M. Economic Evaluation and decision-making in the UK. *Pharmacoeconomics* 2006;24(11): 1133-42.
5. Chalkidou K, Marten R, Cutler D, Culyer T, Smith R, Teerawattananon Y, et al. Health Technology Assessment in Universal Health Coverage. *Lancet* 2013; 382(9910): 48-49.
6. Bill & Melinda Gates Foundation. *What we do - Bill & Melinda Gates Foundation*. 17 May 2013. At [www.gatesfoundation.org/What-We-do](http://www.gatesfoundation.org/What-We-do).
7. Gold M, Siegel J, Russell L, Weinstein M. *Cost-effectiveness in Health and Medicine*. (New York : Oxford University Press, 1996).
8. Drummond M, Sculpher M. Common Methodological Flaws in Economic Evaluations. *Medical Care* 2005; 43(7):II-5-II-14.
9. Jefferson T, Demicheli V. Quality of economic evaluations on health care: it is time for action to ensure higher methodological quality. *BMJ* 2002; 324(7333): 313–314
10. Al-Aqeel SA. State of health economic evaluation research in Saudi Arabia: a review. *Journal of Clinicoeconomic and Outcomes Research* 2012;4:177-84.
11. Hoque M, Khan J, Hossain S, Gazi R, Rashid H-A, Keohlmoos T, et al. A systematic review of economic evaluations of health and health-related interventions in Bangladesh. *Cost Effectiveness and Resources and Resource Allocation* 2011;9:12.
12. World Bank Group. *How We Classify Countries*, 20 April 2013. At <http://data.worldbank.org/about/country-classifications>.
13. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *BMC Medicine*. 2013;11:80.
14. National Institute for Health and Care Excellence (NICE). *Appendix H: Methodology Checklist - economic evaluations*. 2013.
15. Walker D, Fox-Rushby JA. Economic Evaluation of Communicable Disease Interventions in Developing Countries: a critical review of the published literature. *J Health Econ* 2000;9(8):681-98.
16. Teerawattananon Y, Russell S, Mugford M. A systematic review of economic evaluation literature in Thailand: Are the data good enough to be used by policy makers? *Pharmacoeconomics* 2007;25(6):467-79.

17. Cooper N, Coyle D, Abrams K, Mugford M, Sutton A. Use of evidence in decision-models: an appraisal of health technology assessments in the UK since 1997. *J Health Services Research and Policy* 2005;10(4):245-50.
18. Willey B, Paintain L, Mangham L, Car J, Schellenberg J. Strategies for delivering insecticide-treated nets at scale for malaria control: a systematic review. *Bulletin of the World Health Organisation* 2012;90(9):672-84.
19. Creese A, Floyd K, Alban A, Guinness L. Cost-effectiveness of HIV/AIDS interventions in Africa: a systematic review of the evidence. *Lancet* 2002;359(9318):1635-43.
20. Klein E. Bill Gates: *Death is something we really understand extremely well*. The Washington Post, 17 May 2013. At [www.washingtonpost.com/blogs/wonkblog/wp/2013/05/17/bill-gates-death-is-something-we-really-understand-extremely-well](http://www.washingtonpost.com/blogs/wonkblog/wp/2013/05/17/bill-gates-death-is-something-we-really-understand-extremely-well).
21. Goldie SJ, O'Shea M, Campos NG, Diaz M, Sweet S, Kim SY. Health and economic outcomes of HPV 16,18 vaccination in 72 GAVI-eligible countries. *Vaccine* 2008;26(32):4080-93.
22. Murray CJL, Lopez AD, Harvard School of Public Health, World Health Organisation, World Bank. *The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*. (Cambridge: Harvard University Press, 1996).
23. Mbonye AK, Hansen KS, Bygbjerg IC, Magnussen P. Intermittent preventative treatment of malaria in pregnancy: the incremental cost-effectiveness of a new delivery system in Uganda. *Trans R Soc Trop Med Hyg* 2008;102(7):685-93.
24. World Health Organisation. Cost effectiveness thresholds. At [www.who.int/choice/costs/CER\\_thresholds](http://www.who.int/choice/costs/CER_thresholds), (accessed 5 September 2013).
25. Ozawa S, Mirelman A, Stack ML, Walker DG, Levine OS. Cost-effectiveness and economic benefits of vaccines in low- and middle-income countries: a systematic review. *Vaccine* 2012;31(1):96–108
26. Guinness L, Harker M. *The Use of QALYs in economic evaluation in low- and middle-income countries: a systematic review*. (Sydney : iHEA 9<sup>th</sup> World Congress on Health Economics, 2013)
27. Briggs A, Sculpher M, Buxton M. Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. *Health Econ* 1994;3(2):95-104.
28. Teerawattananon Y, McQueston K, Glassman A, Yothasamut J, Myint CY. Health technology assessments as a mechanism for increased value for money: recommendations to the global fund. *Globalization and Health* 2013;9:35
29. Tessa Tan-Torres Edejer, ed. *WHO guide to cost-effectiveness analysis*. s.l. : World Health Organisation, 2003.
30. Walker D, Hutubessy R, Beutels, P. WHO Guide for standardisation of economic evaluations of immunization programmes. *Vaccine* 2010; 28(11):2356-2359.
31. National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal*. s.l. : NICE, 2013.
32. International Society for Pharmacoeconomics and Outcomes Research. [www.ispor.org/PEguidelines/index.asp](http://www.ispor.org/PEguidelines/index.asp). (Accessed 05 May 2013)

33. Bill and Melinda Gates Foundation. *Who we are - Bill and Melinda Gates Foundation*. [www.gatesfoundation.org/who-we-are](http://www.gatesfoundation.org/who-we-are). (Accessed 05 May 2013)
34. Drummond, M et al. Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force report. *Value in Health* 2009;12(4):409-418.
35. Health Intervention and Technology Assessment Program. *Health Technology Assessment Process Guidelines*. At: [http://www.hitap.net/system/files/health\\_technology\\_assessment\\_process\\_guidelines\\_english\\_finalversion\\_01march2012.pdf](http://www.hitap.net/system/files/health_technology_assessment_process_guidelines_english_finalversion_01march2012.pdf)
36. Shillcutt SD1, Walker DG, Goodman CA, Mills AJ. Cost-effectiveness in Low- and Middle-Income Countries. *Pharmacoeconomics* 2009, 27(11):903-917.
37. Gold, M. R., Stevenson, D., & Fryback, D. G. (2002). HALYS and QALYS and DALYS, Oh My: similarities and differences in summary measures of population Health. *Annual Review of Public Health*, 23(1), 115-134.