Journal of Theoretical Biology 611 (2025) 112197

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/jtbi

Uncertainty quantification in cost-effectiveness analysis for stochastic-based infectious disease models: Insights from surveillance on lymphatic filariasis

Mary Chriselda Antony Oliver [a,b,*], Matthew Graham [b], Ioanna Manolopoulou [c], Graham F. Medley [d], Lorenzo Pellis [e], Koen B. Pouwels [f,1], Matthew Thorpe [g,1], T. Deirdre Hollingsworth [b,1]

[a] University of Cambridge, Department of Applied Mathematics and Theoretical Physics, Wilberforce Road, Cambridge, CB3 0BN, UK
[b] University of Oxford, Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Oxford, OX3 7LF, UK
[c] University College London, Department of Statistical Science, Gower Street, London, WC1E 6BT, UK
[d] London School of Hygiene and Tropical Medicine, Centre for Mathematical Modelling of Infectious Disease and Department of Global Health and Development, Keppel Street, London, WC1E 7HT, UK
[e] University of Manchester, Department of Mathematics, Oxford Road, Manchester, M13 9PL, UK
[f] University of Oxford, Nuffield Department of Primary Care Health Sciences, Oxford, OX2 6HT, UK
[g] University of Warwick, Department of Statistics, Coventry, CV4 7AL, UK

ARTICLE INFO

Keywords:
Lymphatic filariasis
Mathematical modelling
Stopping threshold
Cost-effective analysis
Optimal transport

ABSTRACT

Cost-effectiveness analyses (CEA) typically involve comparing the effectiveness and costs of one or more interventions compared to the standard of care, in order to determine which intervention should be optimally implemented to maximise population health within the constraints of the healthcare budget. Traditionally, cost-effectiveness evaluations are expressed using incremental cost-effectiveness ratios (ICERs), which are compared with a fixed willingness-to-pay (WTP) threshold. Due to the inherent uncertainty in intervention costs and the overall burden of disease, particularly with regard to diseases in populations that are difficult to study, it becomes important to consider uncertainty quantification while estimating ICERs.

To tackle the challenges of uncertainty quantification in CEA, we propose an alternative paradigm utilizing the Linear Wasserstein framework combined with Linear Discriminant Analysis (LDA) using a demonstrative example of lymphatic filariasis (LF). This approach uses geometric embeddings of the overall costs for treatment and surveillance, disability-adjusted life-years (DALYs) averted for morbidity by quantifying the burden of disease due to the years lived with disability, and probabilities of local elimination over a time-horizon of 20 years to evaluate the cost-effectiveness of lowering the stopping thresholds for post-surveillance determination of LF elimination as a public health problem. Our findings suggest that reducing the stopping threshold from $< 1\,\%$ to $< 0.5\,\%$ microfilaria (mf) prevalence for adults aged 20 years and above, under various treatment coverages and baseline prevalences, is cost-effective. When validated on 20 % of test data, for 65 % treatment coverage, a government expenditure of WTP ranging from \$500 to \$3000 per 1 % increase in local elimination probability justifies the switch to the lower threshold as cost-effective.

Stochastic model simulations often lead to parameter and structural uncertainty in CEA. Uncertainty may impact the decisions taken, and this study underscores the necessity of better uncertainty quantification techniques within CEA for making informed decisions.

* Corresponding author at: DAMTP, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK.
E-mail address: mca52@cam.ac.uk (M.C. Antony Oliver).
[1] Equal contribution.

https://doi.org/10.1016/j.jtbi.2025.112197
Received 30 July 2024; Received in revised form 9 April 2025; Accepted 30 May 2025
Available online 20 June 2025

1. Introduction

1.1. Health economics motivation

Global health systems face enormous challenges due to the rising demand for healthcare services and the finite resources available to them. While other factors, such as equity may play a role, a common aim for governments is to maximise overall population health within the constraints of the available healthcare budget. Planning, managing, and assessing health systems heavily rely on economic factors. The best use of limited resources is guided by health economic analyses, which provide cohesive techniques for evaluating the cost-effectiveness of health interventions.

The economic evaluation of health interventions is typically based on the outcomes and costs of the interventions. Depending on how the outcomes and interventions are evaluated, one of the main methodologies used is the cost-effectiveness analysis (CEA). This is an economic evaluation technique in which two or more health interventions are compared in terms of incremental costs and incremental effects relative to the standard of care, with the cost-effectiveness expressed using the incremental cost-effectiveness ratio (ICER) - a measure that divides the incremental costs by the incremental effects. Most countries that regularly use CEA to guide policy decisions around the implementation and reimbursement of interventions specify in their health-economic guidelines that cost-utility analyses should be used, where the denominator of the ICER is expressed in quality-adjusted life-years (QALYs) or disability-adjusted life-years (DALYs). The latter is more frequently used in low- and middle-income countries (LMICs). Given the focus on LMICs in this paper, DALYs will be used, from here onwards.

The interpretation of the ICER depends on where it lies on the cost-effectiveness plane (refer to Fig. 1 in Griffin et al., 2020). If the new intervention is more effective and saves money compared to the standard of care (south-east quadrant), or if the new intervention is less effective and more costly compared to the standard of care (north-west quadrant), the ICER is negative and interpretation is simple. In the former case, the new intervention should clearly be adopted from a cost-effectiveness point of view, whereas in the latter case, the intervention is clearly worse and should not be adopted.

To determine whether an intervention is likely to improve overall population health within healthcare budget constraints, the ICER can be compared with a cost-effectiveness threshold in situations where the intervention is more costly and more effective (north-east quadrant) or less costly and less effective (south-west quadrant). Assuming the decision-maker indeed wants to maximise overall population health, a symmetrical threshold should be applied to both quadrants, whereby interventions that are more effective and more costly should remain below the threshold, and interventions that are less effective and less costly should remain above the threshold.

If a new intervention costs more per DALY avoided than the healthcare it displaces, then health opportunity costs exceed health benefits, and implementing the new intervention would be expected to lead to an overall reduction in population health, measured in DALYs. Theoretically, the cost-effectiveness threshold (Chi et al., 2020; Turner et al., 2021; Ochalek et al., 2018) should reflect the point at which this occurs. Thus, given the available budget, interventions that are more costly and more effective with an ICER below the threshold are expected to improve overall population health, while similar interventions with an ICER above the threshold are expected to worsen overall population health.

Characterising uncertainty is crucial in CEA, particularly when evaluating the need for additional evidence. Value of Information (VoI) analysis (Jackson et al., 2022, 2019) enhances CEA by quantifying the benefit of reducing uncertainty in decision-making. In health decision-analytic models, VoI assesses the potential benefit of obtaining additional data aimed at reducing uncertainty in key parameters influencing decision uncertainty. Two key uncertainties are model input values and model structure. VoI analyses in the literature typically focus only on parameter uncertainty and completely ignore model structure uncertainty. These models are typically law-driven due to a lack of long-term data. To quantify input uncertainty, a probability distribution for true input values is propagated through the model using Monte Carlo sampling, known as probabilistic sensitivity analysis (PSA) (Claxton et al., 2005; Strong et al., 2015). However, PSA only addresses input uncertainty, not structural uncertainty, which is harder to quantify and requires judgments about the model's real-life representation.

Despite its potential, VoI analysis (Wilson, 2015) is constrained by structural uncertainties, which are rarely quantified in model-based analyses. Not quantifying structural uncertainty implies that the model is a perfect representation of real-world processes and relationships. While VoI analysis for structural uncertainty using model selection and model averaging has been explored previously (Strong and Oakley, 2014; Bojke et al., 2009), methods in this area are still underdeveloped. Addressing these limitations is essential to fully leverage VoI analysis in making informed and effective healthcare policy decisions.

1.2. Theoretical background on lymphatic filariasis

Lymphatic filariasis (LF), a debilitating neglected tropical disease caused by parasitic worms transmitted through mosquitoes, affects about 882 million people across 44 countries (World Health Organization, 2011). In 2000, the World Health Organization (WHO) launched the Global Program to Eliminate Lymphatic Filariasis (GPELF), aiming to eradicate LF as a public health problem (EPHP) in 73 endemic nations by 2020 (World Health Organization, 2023). By 2025, 21 countries, including Brazil and Timor-Leste, were validated as having achieved EPHP, with 14 others under surveillance after halting large-scale treatment (World Health Organization, 2023, 2011).

The primary intervention involves annual mass drug administration (MDA) for at least five years in affected areas, employing drug combinations such as diethylcarbamazine (DEC) + albendazole (DA) or albendazole + ivermectin (IA) (World Health Organization, 2011). Some areas also utilize a triple combination ivermectin + DEC + albendazole (IDA) (King et al., 2018; Irvine et al., 2017). To assess MDA impact and determine if infection levels have dropped below stopping thresholds, the WHO recommends epidemiological monitoring surveys and transmission assessment surveys (TAS). The TAS uses blood smears samples, typically surveying children aged 5 years and above for microfilariae (mf) prevalence (World Health Organization, 2011). The current MDA guidelines advise a minimum of 5 rounds of treatment before a pre-TAS is used to determine whether a first full TAS should be conducted, known as TAS-1. MDA can be stopped if TAS-1 is passed. Two subsequent surveys must also be passed before EPHP can be validated, TAS-2 and TAS-3, each within 2–3 years of the previous assessment (see Fig. 1(A) in Antony Oliver et al., 2024).

However, focusing solely on children may underestimate mf prevalence, potentially missing ongoing transmission as adults may have higher mf prevalence. This paper proposes to improve the sensitivity of TAS to evaluate mf prevalence in adults, targeting $< 0.5\,\%$ mf prevalence. The TAS sample size typically involves randomly sampling approximately 30 sites with 40–60 adults per site to replicate the characteristics of an evaluation unit (EU). Achieving and sustaining WHO goals necessitates effective surveillance, identifying new cases post-EPHP target attainment. Intensive surveillance thresholds ($< 2\,\%$ antigenemia (Ag), $< 1\,\%$ mf) may still be inadequate, especially in areas with *Culex* transmission vectors (Antony Oliver et al., 2024; Davis et al., 2019). Mathematical and biological theories (May, 1991) propose a transmission breakpoint influenced by local transmission conditions and biological factors, in helminth infections such as LF, which depend on sexual reproduction of the parasites, where low worm burdens diminish onward transmission, potentially leading to disease extinction in deterministic scenarios. Studies, have suggested that the breakpoint might be substantially lower than 1 % mf prevalence (Gambhir and Michael,

2008; Michael and Singh, 2015). Stochastic extinction can still occur above this breakpoint but with a lower probability (Davis et al., 2019). If MDA is halted after reaching the breakpoint, the low-level remaining transmission will diminish gradually taking a longer time for LF extinction.

In this work, we aim to provide the first detailed model simulations of reducing the TAS stopping threshold in LF from $< 1\,\%$ to $< 0.5\,\%$ mf prevalence in a sample of adults aged 20 years and above. This facilitates the understanding of the different trade-offs between additional rounds of MDA treatment and rebounds that apply to the design of surveillance strategies. In this context, modelling can help us to understand how adjusting the threshold used in TAS impacts decisions about the stop of interventions and at what cost. For many settings, a reduction in the threshold increases the probability of elimination, decreases the number of treatment rounds required, and reduces costs. Importantly, however, in certain circumstances (e.g., when coverage is lower), lower thresholds can imply an increase in the number of rounds of treatment required to reach that threshold (with increased costs) but help mitigate chronic conditions (such as lymphoedema and hydrocele) and result in longer sustained elimination with fewer future rebounds.

To investigate the issues outlined above, here we use mathematical models of the transmission dynamics of LF as a case study to assess the potential implications of modifying the threshold for TAS. The paper addresses a key question: What are the potential trade-offs encountered in uncertainty quantification of cost-effectiveness analysis on lowering the stopping threshold for TAS in adults aged 20 years and above from an economic, epidemiological and mathematical perspective? In this paper, we restrict to lowering the stopping threshold from $< 1\,\%$ mf prevalence to $< 0.5\,\%$ mf prevalence for a sample of adults motivated by the work in Antony Oliver et al. (2024) and Davis et al. (2019). Importantly, we focus on areas with *Culex* mosquitoes as the major transmission vector using IA drug combinations for potential comparisons.

1.3. Contributions

The primary contribution of this work revolves around the characterisation of uncertainty within the health-economic decision model designed to understand the transmission dynamics of LF. This is achieved through the integration of the Linear Wasserstein Framework with Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). This methodology offers better performance in high-dimensional scenarios while maintaining computational efficiency. To address the fundamental question of interest, the study investigates the following three specific sub-questions, each of which contributes to advancing our understanding:

1. What is the interplay between the dynamics of infection on DALY burden and elimination? In Section 5 we show the monotonic behaviour of the DALY burden and probability of elimination for different stopping thresholds, baseline mf prevalences and MDA coverages.
2. What are the dynamics of the costs both pre and post-MDA surveillance when the stopping threshold is lowered? In Section 5 we explain the trade-off illustrated in the observed non-monotonic behaviour of the costs for different stopping thresholds, baseline prevalences and MDA coverages.
3. If lower stopping thresholds are required for elimination of transmission, then are we realistically able to measure them using current tools? In order to circumvent issues related to the ICERs to address this question using the CEA framework in Section 2, we instead adopt a linear formulation of the Expected Incremental Net-Monetary Benefit (EINMB) metric for fixed country-level WTP values as recommended by several studies (Pichon-Riviere et al., 2023; Vallejo-Torres et al., 2023) for DALYs averted and approximate the range of WTP for probability of elimination (due to lack of data) in order to align with the goals of GPELF. We also extend the analysis to quantify uncertainty with every additional sample size using Value of Information Analysis (VoI) with the help of the Expected Value of Sample Size (EVSI) metric for the optimum WTP values per DALY averted and unit increase in the probability of elimination for different baseline prevalences of lower stopping thresholds. Finally, we propose an alternate paradigm, the Linear Wasserstein Framework in Section 3 that might help us resolve some of the proposed limitations, particularly around structural uncertainty of the CEA framework.

Addressing these questions will help to assess whether lower thresholds have the potential to assist programmes in achieving LF local elimination goals and how such decisions impact programme costs aligning with the GPELF objectives.

1.4. Outline of the paper

We begin in Section 2 by summarizing the theoretical framework for CEA. In addition, we prescribe an alternative paradigm that circumvents structural uncertainties of CEA using the Linear Wasserstein framework in Section 3. The numerical implementation is then described in Section 4 and tested in Section 5. Finally, we discuss our findings and present our conclusions in Section 6. For the reader's convenience we have provided a list of key terminologies used in the manuscript in Table C.17 (refer to Appendix C for more details, including additional results on the necessary outcomes and the transmission model).

2. Value of information (VoI) analysis

Decision-making under uncertainty involves selecting from various strategies based on incomplete information. A risk-neutral decision-maker would typically choose the strategy that maximizes the expected payoff. However, uncertainty introduces the potential for adverse outcomes, as the expected payoff may not always materialize in practice. In this context, the VoI analysis is framed from the perspective of a risk-neutral decision-maker, using the outcomes to identify the strategy with the highest expected payoff given current knowledge. VoI analysis helps determine whether further evidence is necessary to reduce uncertainty and improve decision making.

In this section, we outline key concepts of uncertainty analysis within the CEA framework. We extend the net monetary benefit (NMB) approach to account for resource implications, while aligning with GPELF goals by incorporating probability of elimination (Antillon et al., 2021).

2.1. Notation and basic concepts

Health economic decision-making aims to determine the optimal intervention (policy or strategy) considering the costs and health impacts of various clinical outcomes with respect to the k uncertain parameters $\boldsymbol{\theta} = \{\theta_1, \ldots, \theta_k\}$. A key metric is the Incremental Cost-Effectiveness Ratio (ICER), defined as the ratio of the difference in costs ($\Delta C(\boldsymbol{\theta})$) to the difference in health impacts ($\Delta E(\boldsymbol{\theta})$) between two strategies:

$$\text{ICER} = \frac{\Delta C(\boldsymbol{\theta})}{\Delta E(\boldsymbol{\theta})}, \tag{1}$$

Here, we use a cost-utility analysis, where health impacts ($\Delta E(\boldsymbol{\theta})$) are expressed in disability-adjusted life years ($\Delta\text{DALYs}(\boldsymbol{\theta})$) averted by quantifying overall disease burden due to morbidity and mortality. In the current analyses, we only include effects on morbidity (e.g., lymphoedema, hydrocele) as we assume the intervention has no impact on mortality. For our analysis, the health impacts will consider both DALYs averted and the probability of elimination. Importantly, within the context of this paper, we note that the uncertain parameters $\boldsymbol{\theta}$ (as listed in Table C.16 in the Appendix) include MDA coverages (equivalently, relating to MDA and TAS), vector-to-host ratio, bite risk aggregation parameter and other population characteristics. These parameters are the main drivers of the uncertainty induced in the costs, DALYs averted and probability of elimination. They are not selected arbitrarily as they are informed by

law-based principles and prior knowledge of the factors that contribute to the epidemiological dynamics (see Irvine et al., 2016; Davis et al., 2019 for more details).

Formally, a strategy is considered cost-effective if the ICER does not exceed the health planner's WTP per DALY averted (WTP_{DALY}),

$$\text{ICER} = \frac{\Delta C(\boldsymbol{\theta})}{\Delta \text{DALYs}(\boldsymbol{\theta})} \leq \text{WTP}_{\text{DALY}}. \tag{2}$$

The net-benefit framework circumvents issues with ICERs by not having to deal with extended dominance (when one intervention is less cost-effective than a combination of two or more interventions) by transforming the ICER into a linear additive form, known as the net-monetary benefit (NMB). In particular,

$$\begin{aligned} &\Delta C(\boldsymbol{\theta}) \leq \text{WTP}_{\text{DALY}} \cdot \Delta \text{DALYs}(\boldsymbol{\theta}) \implies \\ &\quad 0 \leq \text{WTP}_{\text{DALY}} \cdot \Delta \text{DALYs}(\boldsymbol{\theta}) - \Delta C(\boldsymbol{\theta}) = \text{NMB}(\boldsymbol{\theta}, \text{WTP}_{\text{DALY}}). \end{aligned} \tag{3}$$

By using NMB, which relies on single monetary values rather than ratios, the framework simplifies the evaluation of multiple interventions, regardless of which quadrant of the cost-effectiveness plane the ICER lies in. The uncertainty about the "true" unknown values of $\boldsymbol{\theta}$ is represented by the joint probability distribution, $\mathbb{P}$. Given a Monte Carlo sample of N iterates of $\boldsymbol{\theta}$, a strategy is preferred over the comparator if the expected incremental NMB (EINMB) exceeds zero:

$$0 \leq \mathbb{E}_{\boldsymbol{\theta}}(\text{NMB}(\boldsymbol{\theta}, \text{WTP}_{\text{DALY}})) \approx \frac{1}{N}\sum_{i=1}^{N} \text{NMB}(\boldsymbol{\theta}_i, \text{WTP}_{\text{DALY}}). \tag{4}$$

where for $i = \{1, 2, \dots, N\}$ we have that $\boldsymbol{\theta}_i \overset{\text{iid}}{\sim} \mathbb{P}$. We now extend this framework to multi-strategy decision analysis between S strategies where we recall that a decision maker is faced with a set of mutually exclusive decision options, indexed $s = \{1, \dots, S\}$. Next, it is assumed that a decision model, denoted NMB predicts the utility for strategy s given N iterates of the uncertain parameters $\boldsymbol{\theta}_i^{(s)}$. With current knowledge, the best that a risk-neutral decision maker can do is to choose the decision option that gives the highest expected utility. The utility associated with this option is:

$$\begin{aligned} &\max_{s=\{1,2,\dots,S\}} \mathbb{E}_{\boldsymbol{\theta}^{(s)}}\left(\text{NMB}\left(\boldsymbol{\theta}^{(s)}, \text{WTP}_{\text{DALY}}\right)\right) \\ &\approx \max_{s=\{1,2,\dots,S\}} \frac{1}{N}\sum_{i=1}^{N}\left(\text{NMB}\left(\boldsymbol{\theta}_i^{(s)}, \text{WTP}_{\text{DALY}}\right)\right). \end{aligned} \tag{5}$$

where the $\mathbb{E}_{\boldsymbol{\theta}^{(s)}}(\cdot)$ represents the expectation (mean) taken with respect to $\mathbb{P}_s$, the distribution of ith iteration of $\boldsymbol{\theta}_i^{(s)}$ for each strategy s. We also include benefits related to the probability of elimination of LF, aligning with GPELF goals using WTP per unit increase in probability of elimination ($\text{WTP}_{\text{Elimination}}$). The NMB is reformulated as:

$$\begin{aligned} &\text{NMB}\left(\boldsymbol{\theta}^{(s)}, \text{WTP}_{\text{DALY}}, \text{WTP}_{\text{Elimination}}\right) \\ &\quad = 100 \cdot \text{WTP}_{\text{Elimination}} \times \Delta\mathbb{I}_{\text{Elimination}}(\boldsymbol{\theta}^{(s)}) \\ &\qquad + \text{WTP}_{\text{DALY}} \times \Delta \text{DALYs}(\boldsymbol{\theta}^{(s)}) - \Delta C(\boldsymbol{\theta}^{(s)}), \end{aligned} \tag{6}$$

where we use the symbol Δ to denote the difference in costs, DALYs averted, and probability of elimination between the strategy being evaluated and the comparator. In the context of this work, the comparator chosen is the outcomes related to the dynamics of the stopping threshold < 1% mf prevalence for 5–10 % baseline prevalences in a sample of children. The choice of this comparator is motivated from the current stopping threshold adopted by the WHO to declare LF as EPHP. Here, $\Delta\mathbb{I}_{\text{Elimination}}$ is 1 if only one strategy achieves elimination, and 0 otherwise. Analogous to the traditional NMB, the strategy that ought to be implemented is indicated by,

$$\max_{s=\{1,2,\dots,S\}} \mathbb{E}_{\boldsymbol{\theta}^{(s)}}\left(\text{NMB}\left(\boldsymbol{\theta}^{(s)}, \text{WTP}_{\text{DALY}}, \text{WTP}_{\text{Elimination}}\right)\right). \tag{7}$$

Simultaneously, the framework also allows for a probabilistic interpretation of cost effectiveness. The probability that a strategy $s \in \{1, \dots, S\}$ is cost-effective (CE), given the uncertain parameters $\boldsymbol{\theta}_i^{(s)}$, is expressed by:

$$\mathbb{P}(s \text{ is CE}) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{M}(\boldsymbol{\theta}_i^{(s)}), \tag{8}$$

where

$$\mathbb{M}(\boldsymbol{\theta}_i^{(s)}) = \begin{cases} 1 & \text{if } s = \arg\max_{s \in 1:S} \text{NMB}\left(\boldsymbol{\theta}_i^{(s)}, \text{WTP}_{\text{DALY}}, \text{WTP}_{\text{Elimination}}\right) \\ 0 & \text{Otherwise} \end{cases} \tag{9}$$

The framework therefore presents a measure of uncertainty by the proportion of samples where the strategy has the highest NMB of all strategies.

In general terms, the Expected Value of Perfect Information (EVPI) (Brennan et al., 2007) is the expected value of learning, with certainty, the "true" values of all model parameters $\boldsymbol{\theta}^{(s)}$ (i.e., eliminating all parameter uncertainty). It represents the maximum amount a decision-maker would be willing to pay for perfect information to avoid the potential losses associated with uncertainty. EVPI is defined as,

$$\begin{aligned} \text{EVPI} = &\mathbb{E}_{\boldsymbol{\theta}^{(s)}}\left(\max_{s=\{1,2,\dots,S\}} \text{NMB}\left(\boldsymbol{\theta}^{(s)}, \text{WTP}_{\text{DALY}}, \text{WTP}_{\text{Elimination}}\right)\right) \\ &- \max_{s=\{1,2,\dots,S\}} \mathbb{E}_{\boldsymbol{\theta}^{(s)}}\left(\text{NMB}\left(\boldsymbol{\theta}^{(s)}, \text{WTP}_{\text{DALY}}, \text{WTP}_{\text{Elimination}}\right)\right). \end{aligned} \tag{10}$$

The expected value of acquiring new information about a subset of parameters of interest is used to identify the parameters that are important in driving the decision uncertainty. Here, the vector of parameters can be split in two components $\boldsymbol{\theta}^{(s)} = (\boldsymbol{\phi}^{(s)}, \boldsymbol{\psi}^{(s)})$, where $\boldsymbol{\phi}^{(s)}$ is the subvector of parameters of interest (i.e., those that could be investigated further) and $\boldsymbol{\psi}^{(s)}$ are the remaining "nuisance" parameters. Therefore, the decision option is selected based on the one that maximizes expected utility, conditional on the values $\boldsymbol{\phi}^{(s)}$. This has utility:

$$\max_{s=\{1,2,\dots,S\}} \mathbb{E}_{\boldsymbol{\psi}^{(s)}|\boldsymbol{\phi}^{(s)}}\left(\text{NMB}\left(\boldsymbol{\theta}^{(s)}, \text{WTP}_{\text{DALY}}, \text{WTP}_{\text{Elimination}}\right)\right). \tag{11}$$

where $\mathbb{E}_{\boldsymbol{\psi}^{(s)}|\boldsymbol{\phi}^{(s)}}$ denote the expectation taken with respect to the distribution of $\boldsymbol{\psi}^{(s)}|\boldsymbol{\phi}^{(s)}$. This gives the value of learning $\boldsymbol{\phi}^{(s)}$ with no uncertainty. Of course, we will never be in the position to completely eliminate the uncertainty on $\boldsymbol{\phi}^{(s)}$, so we then average over its current probability distribution while also subtracting the value of the current optimal decision to calculate the Expected Value of Partial Perfect Information (EVPPI) (Brennan et al., 2007; Strong et al., 2015; Heath et al., 2016). The economic value of eliminating all uncertainty about $\boldsymbol{\phi}^{(s)}$ (assuming risk neutrality) is equal to the EVPPI which is given by:

$$\begin{aligned} \text{EVPPI} = &\mathbb{E}_{\boldsymbol{\phi}^{(s)}}\left(\max_{s=\{1,2,\dots,S\}} \mathbb{E}_{\boldsymbol{\psi}^{(s)}|\boldsymbol{\phi}^{(s)}}\left(\text{NMB}\left(\boldsymbol{\theta}^{(s)}, \text{WTP}_{\text{DALY}}, \text{WTP}_{\text{Elimination}}\right)\right)\right) \\ &- \max_{s=\{1,2,\dots,S\}} \mathbb{E}_{\boldsymbol{\theta}^{(s)}}\left(\text{NMB}\left(\boldsymbol{\theta}^{(s)}, \text{WTP}_{\text{DALY}}, \text{WTP}_{\text{Elimination}}\right)\right). \end{aligned} \tag{12}$$

Expected Value of Sample Information (EVSI) (Heath et al., 2019) measures the value of collecting additional data $\tilde{\boldsymbol{\theta}}^{(s)}$ to inform $\boldsymbol{\phi}^{(s)}$, assuming $\tilde{\boldsymbol{\theta}}^{(s)}$ directly updates $\boldsymbol{\phi}^{(s)}$ and is independent of $\boldsymbol{\psi}^{(s)}|\boldsymbol{\phi}^{(s)}$. More formally, it is the value of acquiring information through sampling or additional data collection to reduce uncertainty in decision-making. EVSI is bounded above by EVPPI. If data $\tilde{\boldsymbol{\theta}}^{(s)}$ were observed as $\tilde{x}^{(s)}$, it would update distribution of $\boldsymbol{\phi}^{(s)}$ for each strategy s, impacting the net benefit distribution for each treatment. EVSI is the average value over all possible data sets:

$$\begin{aligned} \text{EVSI} = &\mathbb{E}_{\tilde{\boldsymbol{\theta}}^{(s)}}\left(\max_{s=\{1,2,\dots,S\}} \mathbb{E}_{\boldsymbol{\theta}^{(s)}|\tilde{\boldsymbol{\theta}}^{(s)}}\left(\text{NMB}\left(\boldsymbol{\theta}^{(s)}, \text{WTP}_{\text{DALY}}, \text{WTP}_{\text{Elimination}}\right)\right)\right) \\ &- \max_{s=\{1,2,\dots,S\}} \mathbb{E}_{\boldsymbol{\theta}^{(s)}}\left(\text{NMB}\left(\boldsymbol{\theta}^{(s)}, \text{WTP}_{\text{DALY}}, \text{WTP}_{\text{Elimination}}\right)\right). \end{aligned} \tag{13}$$

where $\mathbb{E}_{\boldsymbol{\theta}^{(s)}|\tilde{\boldsymbol{\theta}}^{(s)}}$ denotes the expectation over the posterior distribution of $\boldsymbol{\theta}^{(s)}$ given data $\tilde{\boldsymbol{\theta}}^{(s)}$.

In this paper, we estimate EVSI computationally using the nested Monte Carlo method based on "Moment Matching" (Heath et al., 2019).

This approach enhances computational efficiency by reducing nested Monte Carlo error and improving the approximation of the posterior sample distribution through Moment Matching (see Algorithm 1 in Appendix A for details). Several alternative methods in literature exist, including Importance Sampling (IS), regression techniques, the Gaussian approximation method, and Integrated Nested Laplace Approximation (INLA) (see Table 3 in Kunst et al., 2020 for comparison of the different methods). However, we choose Moment Matching (Heath et al., 2016) for its better computational efficiency in estimating EVSI (see Table 2 in Heath et al., 2020 which reports the computational time for EVSI evaluation in three case studies). It is particularly beneficial when the data does not have an easily identifiable sufficient statistic, but the model itself remains computationally efficient. Although INLA is highly efficient for Bayesian inference—particularly in latent Gaussian models where it leverages dimensionality reduction by treating PSA simulations as a 'spatial problem'. This method could be more efficient than the Moment Matching method in health economic models with a small number of underlying parameters but with a longer computational time. Therefore, Moment Matching provides a more practical balance between accuracy and efficiency for our EVSI estimation needs.

2.2. Fundamental issues implementing EVSI using moment matching

There are several challenges that arise when evaluating the EVSI metric defined in (13) using Moment Matching:

1. Assumptions on Distributions: Implementing EVSI using the moment-matching method (see Algorithm 1 in Appendix A) involves approximating the distribution of the sample information using moments (mean, variance, etc.). This can introduce errors, especially if the true distribution of the sample information is not well approximated by these moments.
2. Dependence on Prior Information: The quality of EVSI estimates using Moment Matching depends heavily on the prior information available. Poor or inaccurate priors can lead to misleading EVSI estimates.
3. Implementation Challenges: Moment Matching is accurate and efficient when the health economic model has a low computation time but becomes increasingly infeasible as the model runtime increases and can be inaccurate when the sample size is less than 10.
4. Requires Accurate EVPPI Estimation: Moment Matching is more accurate for studies that have significant impact on the underlying uncertainty in the decision-analytic model, i.e., the EVPPI of $\boldsymbol{\phi}^{(s)}$ for each strategy $s \in \{1, \dots, S\}$ needs to be high compared to the value of reducing all model uncertainty (i.e., EVPI), ideally greater than 40 % (Heath et al., 2019).

While Moment Matching can be a useful tool for approximating EVSI, these limitations must be carefully considered and addressed to ensure accurate and reliable health economic decision-making.

3. Linear wasserstein framework

This section introduces an alternative methodology to evaluate EINMB and EVSI defined in Section 2 by assessing the cost-effectiveness of different strategies using the Linear Wasserstein Framework, also referred to as "Linear Optimal Transport" (LOT), originally formulated in Wang et al. (2013). Wasserstein-like distances serve as metrics on probability measures, quantifying the minimum cost required to transport mass between two distributions, typically using the pth power of the Euclidean distance. While Wasserstein distances are commonly used to model translational shifts, the framework can be extended to incorporate more complex deformations, such as rotations and scalings, by adjusting the cost function. A key challenge with Wasserstein distances is their high computational cost and the limited availability of off-the-shelf data analysis tools. The linearization of optimal transport addresses these issues by mapping $P : \mathcal{P}(X) \to \mathbb{R}^k$ (for some k) such that the Wasserstein distance in $\mathcal{P}(X)$ is well approximated by the Euclidean distance in $\mathbb{R}^k$. This transformation enables the use of standard data analysis techniques, including dimensionality reduction, classification, and modelling, within the Euclidean space.

3.1. Notations

We define some of the commonly used notations in the Linear Wasserstein Framework. We denote the space of probability measures on X as $\mathcal{P}(X)$. Let $X, Y \subseteq \mathbb{R}^d$ with $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$.

Definition 1. Let $\mu \in \mathcal{P}(X)$ and $T : X \to Y$ be a measurable map, the pushforward of μ by T, denoted as $T_{\#}\mu$ is the measure ν defined by,

$$\nu(B) = \mu(T^{-1}(B)). \tag{14}$$

for all measurable set $B \subseteq Y$.

Definition 2. We define $\Pi(\mu, \nu)$ between measures μ and ν to be the set of probability measures (couplings) on the product space $\mathcal{P}(X \times Y)$ whose first marginal is μ and the second marginal is ν. For any transport map $T: X \to Y$, there exists an associated transport plan π such that,

$$\pi = (\mathrm{Id} \times T)_{\#}\mu. \tag{15}$$

where Id denotes the identity map. We recall that if $P^X : X \times Y \to X$ and $P^Y : X \times Y \to Y$ are the canonical projections, then the marginals are $P^X_{\#}\pi = \mu$ and $P^Y_{\#}\pi = \nu$.

3.2. Optimal transport formulations

The *Monge formulation* (Villani, 2021) would be to find the transport map T satisfying (14), given the probability measures $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$. The objective function would be minimising $\mathbb{M}(\mu, \nu)$, where

$$\mathbb{M}(\mu, \nu) := \inf_{T : T_{\#}\mu = \nu} \int_X |x - T(x)|^2 \mathrm{d}\mu(x). \tag{16}$$

We call any T which satisfies $T_{\#}\mu = \nu$ a transport map and the minimizer of the optimisation problem in (16) as the optimal transport map T^*. It is often difficult to handle the non-convex optimisation problem in (16) due to its non-linearity in T.

The *Kantorovich formulation* (Villani, 2021) would be to minimise the objective function $\mathbb{K}(\mu, \nu)$, given the probability measures $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, where

$$\mathbb{K}(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \int_{X \times Y} |x - y|^2 \mathrm{d}\pi(x, y). \tag{17}$$

The minimizer of (17) is the optimal transport plan π^*. In this sense, the *Kantorovich formulation* in (17) can viewed as a relaxation of the *Monge formulation*. Moreover, in cases where the probability measures are discrete we find that the *Monge formulation* is ill-posed as the transport maps may not exist. This motivates the modified formulation in (17) which describes the amount of mass $\pi(x, y)$ that can be transported from x to multiple positions at y.

3.3. Wasserstein distances

Let $X = Y = \mathbb{R}^d$. This allows us to define the 2-Wasserstein distance (Villani, 2021), which is the minimum transportation cost between the probability measures $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, as

$$d_{\mathrm{W}^2}(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \left(\int_{X \times Y} |x - y|^2 \mathrm{d}\pi(x, y) \right)^{\frac{1}{2}}. \tag{18}$$

We note that the construction can be generalized to any power $p \in [1, \infty)$ where an additional bound on the pth moments of the probability measures needs to be imposed, so as to guarantee a finite transport cost.

The d_{W^2} distances are advantageous for Lagrangian modeling due to their simplicity, metric properties (like symmetry), existence of geodesics, Riemannian structure and theoretical benefits like existence

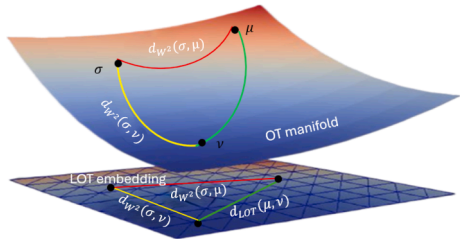

Fig. 1. The Linear Wasserstein framework embeds measures in the tangent space of a fixed reference σ. As a consequence, the Euclidean distance between the non-negative measures μ and ν is an approximation for the 2-Wasserstein distance $d_{W^2}(\mu,\nu)$. This figure is computed using ParaView.

of optimal transport maps and plans. However, they require the inputs to be probability measures, are computationally expensive, are distances defined on the same metric space and lack off-the-shelf data analysis tools. We therefore opt for the Linear Wasserstein Framework, which reduces the computational expense and allows for off-the-shelf Euclidean data analysis tools.

3.4. Linear wasserstein framework

The Linear Wassertein framework introduced by Wang et al. (2013), illustrated in Fig. 1, has several applications in biomedical imaging, analysis of 2-D point cloud data (Basu et al., 2014; Ozolek et al., 2014), telescopic and facial expressions (Kolouri and Rohde, 2015; Kolouri et al., 2017). The term "linear" refers to the (Euclidean) vector space structure that one gains after approximation. Heuristically, instead of computing the geodesic distance on the manifold we compute a 'projection' of the manifold to the tangent plane at a (fixed) reference measure and then compute the distances on the tangent plane. This is the main reason we use the term linear when naming the distance.

To discuss the Linear Wasserstein framework in the continuous setting we consider a domain $X \subseteq \mathbb{R}^d$ that is a bounded, convex and closed subset of $\mathbb{R}^d$ with a non-empty interior, alongside the probability measures $\mu_i \in \mathcal{P}(X)$ for all $i \in \{1,2,\dots,N\}$ and a fixed reference $\sigma \in \mathcal{P}(X)$. The optimal transport map T_i^* between σ and μ_i satisfies,

$$d_{W^2}(\mu_i,\sigma) = \left(\int_X |x - T_i^*(x)|^2 \mathrm{d}\sigma(x)\right)^{\frac{1}{2}} \tag{19}$$

where ${T_i^*}_{\#}\sigma = \mu_i$. This provides the basis to formally introduce the linear Wasserstein distance for two measures say μ_1 and μ_2.

$$d_{LW^2}(\mu_1,\mu_2;\sigma) = \left(\int_X |T_1^*(x) - T_2^*(x)|^2 \mathrm{d}\sigma(x)\right)^{\frac{1}{2}}. \tag{20}$$

This enables us to compute the linear embeddings in the form of projections $P : \mathcal{P}(X) \to T_\sigma\mathcal{P}(X)$ onto tangent space. This can be expressed as,

$$P(\mu_i) = T_i^* - \mathrm{Id}, \tag{21}$$

Equivalently, we relate the Linear Wasserstein distance d_{LW^2} to 2-Wasserstein distance d_{W^2} in Eq. (19) between any two measures μ_1, μ_2 relative to σ. This can be expressed as follows,

$$d_{W^2}(\mu_1,\mu_2;\sigma) = ||P(\mu_1) - P(\mu_2)||_{L^2(\sigma)} = d_{LW^2}(\mu_1,\mu_2;\sigma). \tag{22}$$

Remark 1. This implies that the maps $P(\mu_i)$ form the linear embeddings in the form of projections from the 2-Wasserstein space to a L^2 (Euclidean) space, thereby preserving the optimal transport distance between μ_i and σ. It is assumed that $d_{LW^2}(\mu_1,\mu_2;\sigma) \approx d_{W^2}(\mu_1,\mu_2)$ and the approximation depends on the curvature of the Wasserstein space and in general the linear Wasserstein distance is not equivalent (in terms of metric equivalence) to the Wasserstein space. However, when there is some special structure, such as when the measures are all translations or shearings then one gets established bounds (Moosmüller and Cloninger, 2023) like $cd_{LW^2}(\mu_1,\mu_2;\sigma) \le d_{W^2}(\mu_1,\mu_2) \le Cd_{LW^2}(\mu_1,\mu_2;\sigma)$ where c, C are some positive constants, independent of μ_1 and μ_2.

3.5. Advantages of using the linear wasserstein framework

The linear optimal transport (LOT) framework offers an alternative approach to the calculation of the Expected Value of Sample Information (EVSI), addressing several limitations associated with the Moment Matching method as outlined below:

1. Approximation Accuracy: Unlike Moment Matching, which approximates the distribution using moments, LOT directly operates on the full distribution of the data by leveraging probability measures. This results in more accurate representations of the underlying distributions, thereby reducing approximation errors. According to ISPOR's recommendations (Rothery et al., 2020), uncertainty in input parameters should be represented using probability distributions, with any dependencies between parameters captured by a joint, correlated distribution. LOT facilitates the propagation of uncertainty from the transmission model input parameters to the transmission model outputs by representing them as a point cloud (a discrete set of points in space), thereby quantifying the structural uncertainty more comprehensively.
2. Handling High Dimensions: LOT distance is often more robust in high dimension compared to other well-known metrics on measure spaces, such as Jensen-Shannon and Kullback-Leibler divergence. The reason for this is due to the fact there is no underlying assumption that requires measures to be absolutely continuous (often rare in high-dimensions). In particular, it is often informative to project to a lower dimensional subspace where off-the-shelf data analysis tools can be employed (see 4 below; technically, this follows from the well-known manifold hypothesis (Fefferman et al., 2016)).
3. Computational Efficiency: Evaluating pairwise approximate distances between N samples using LOT requires solving only N optimal transport problems (to compute maps $P(\mu_i)$ for all samples), compared to the $O(N^2)$ complexity required for computing exact pairwise Wasserstein distances.
4. Off-the-shelf Data Analysis Tools: LOT distance benefits from the Euclidean space where standard (off-the-shelf) techniques for clustering, classification, regression, and dimensionality reduction such as PCA can be easily performed.

By directly addressing the distribution of sample information and leveraging efficient optimization techniques, the LOT framework provides a robust and scalable uncertainty quantification tool to calculate cost-effectiveness analysis. This circumvents many of the limitations associated with Moment Matching, leading to more accurate and reliable decision-making.

3.6. Decision making with uncertainty

Uncertainty quantification (UQ) is integral to decision-making but remains challenging due to the inherent trade-off between expressivity and computational tractability. Traditional methods, such as Monte Carlo (MC) simulations and Gaussian Process Regression (GPR), exhibit fundamental limitations when addressing distributional uncertainty, where the underlying probability distribution is unknown or subject to shifts. MC-based techniques rely on random sampling, incurring significant computational costs, especially in high-dimensional settings where the curse of dimensionality leads to slow convergence (Wahba, 1990; Williams and Rasmussen, 1996). Meanwhile, GPR (Stevenson et al., 2004) assumes smooth function priors, restricting its ability to model abrupt changes or heavy-tailed distributions, which are prevalent in real-world applications. These limitations hinder the accurate quantification of uncertainty when probability distributions evolve dynamically over time.

The Linear Wasserstein Framework (also called LOT) provides a nuanced approach that extends beyond conventional UQ techniques by leveraging the underlying geometry of the probability distributions. Unlike Latin Hypercube Sampling (LHS) (McKay, 1992), which relies on uniform stratification and may struggle with multimodal or heavy-tailed distributions, LOT directly captures the geometric structure of probability distributions, making it particularly suitable for uncertainty propagation in systems with non-linear interactions. In contrast to Polynomial Chaos Expansion (PCE) (Xiu and Karniadakis, 2002), which approximates uncertainty through polynomial bases that deteriorate in high dimensions, LOT circumvents such approximations by preserving full distributional information through transport-based embeddings. This distinction enables LOT to remain robust even in scenarios characterised by non-Gaussian distributions.

Unlike Copula-based models (Nelsen, 2006), which capture dependency structures but fail to quantify distributional distances, LOT explicitly preserves the full statistical geometry of uncertainty. This also sets LOT apart from Linear Noise Approximation (LNA) (Fintzi et al., 2022; Swallow et al., 2024), which relies on local linearization and fails to capture non-Gaussian dynamics, as well as quantile emulation (Semochkina et al., 2025), which estimates specific quantiles rather than the complete probabilistic structure. Furthermore, while Bayesian Network Models (BNMs) (Pearl, 1988) enable probabilistic reasoning within structured graphical models, their inference becomes computationally intractable in high dimensions. In contrast, LOT facilitates efficient computation in high-dimensional settings by leveraging its linear structure and Euclidean geometry.

In health economics, quantifying uncertainty is crucial for understanding its impact on decision-making. The LOT method offers a refined approach for this process, as illustrated in Algorithms 2 and 3 in Appendix A. This method involves computing the projections defined in (21) and applying Principal Component Analysis (PCA). The rationale behind using PCA is to reduce the dimensionality of the projections, which helps prevent issues like singularity in the within-class scatter matrix when performing LDA. This is especially important because the number of training samples is much smaller than the feature space dimension. By reducing the dimensionality, we ensure that the within-class variance is well-defined, allowing for a meaningful comparison with the dimensionality reduction technique employed in LOT. The PCA components are then used as feature vectors to train the LDA, which classifies the different strategies (e.g., baseline prevalence for $< 0.5\,\%$ stopping threshold) based on the decision boundary (hyperplane) determined by the LDA. We say that a baseline prevalence (for each stopping threshold evaluated) is cost-effective if the test data lies below the decision boundary (hyperplane). We recall that the WTP_{DALY} are obtained from data for 80 % MDA coverage, but no such data is available for $\text{WTP}_{\text{Elimination}}$ at 65 % MDA coverage so we estimate it using the decision boundary.

Next, we extend the same procedure by incorporating bootstrapping to obtain additional projections of the transport maps. This resampling ensures that uncertainty in the input parameters can be propagated to the outcomes and thereby impact the classification performance (utility). Here, utility directly measures the model's performance, and an increase in utility indicates that the added data is enhancing the model's reliability, which is valuable for making optimal decisions. The EVSI is then calculated to assess the added utility of incorporating more data through bootstrapping compared to the current dataset (baseline using a sample size of $N = 500$ independent draws in the transmission model and training the outcomes onto a LDA (20:80 - train:test split ratio)). The goal is to understand how additional data (captured via different bootstrapped sample size $M = 500, 1000, 1500$ draws) improves the classification of various baseline prevalences using the different epidemiological outcomes under $< 0.5\,\%$ stopping threshold with LDA (by increasing the train:test split ratio from 20:80 to 80:20). Unlike traditional methods like Moment Matching, which may overlook complex uncertainties, we expect the LOT projections to capture the inherent structural uncertainties. This offers decision-makers a more comprehensive, distribution-aware view of how uncertainty impacts classification accuracy, leading to more informed and robust policy decisions.

4. Methods

We utilize the stochastic TRANSFIL model (Irvine et al., 2015) with parameters previously estimated (Irvine et al., 2016; Stolk et al., 2018) to represent transmission by *Culex* mosquitoes (for more details on the parameters we refer Table C.16 in Appendix C). The model simulates the health impacts of LF and incorporates MDA effects, based on simulated target coverage, systematic non-adherence, and drug efficacy (Dyson et al., 2017). We excluded other interventions, such as vector control, for this study. We modelled closed populations (i.e., no migration, only births and deaths) of 100,000-50,000 people, reflecting the population size for an EU in standard TAS as per WHO guidelines (World Health Organization, 2011) and an exponential age distribution. The detection parameters were fitted using Bayesian MCMC to data from Malindi, Kenya, Colombo, Gampaha and Sri Lanka (Irvine et al., 2016). MDAs were simulated at 65 % and 80 % coverage. Systematic non-adherence was included by calculating individual treatment probabilities based on coverage and between-round correlation, parameterized with data from Leogane, Haiti, and Egypt (Dyson et al., 2017).

The model also simulates health impacts of lymphoedema, hydrocele, and acute adenolymphangitis (ADL) using published methods. Morbidity due to lymphoedema and hydrocele was modelled using a non-linear functional relationship of infection and morbidity for sub-Saharan Africa (Van der Werf et al., 2003). The model assumes morbidity occurs after accruing a certain cumulative worm burden. ADL incidence was estimated twice per year in 70 % of hydrocele patients and four times annually in 95 % of lymphoedema patients (Chu et al., 2010). Prevalence was converted using published disability weights (Network, 2020). Side-effects of MDA were not considered, despite reports of 13 % feeling unwell post-MDA, as these effects were deemed minor (Willis et al., 2020). Mental illness was also excluded due to lack of accurate data, despite its recognized burden in LF (Ton et al., 2015; Koschorke et al., 2022).

For WHO-prescribed starting and stopping decisions (World Health Organization, 2011), we considered TAS survey samples from 30 sites per EU. Baseline prevalences were sampled from a normal distribution with means of 5–10 %, 10–20 %, or 20–30 %, representing different strategies in our decision analytic model. In each site, we sampled 40–60 adults aged 20 years and above to evaluate TAS. If mf-positive adults were below the stopping threshold MDA was halted until the next survey; otherwise, it continued. We iterated this algorithm 1000 times and reported mean baseline prevalences. Cost simulations considered TAS ($12,494.75 Brady et al., 2017) and MDA rounds ($7640.92 Stolk et al., 2013) over a 20-year horizon, with discounting included. We note that for MDA restarts, the costs of the MDA and TAS are doubled.

For cost-effectiveness analysis, using the Expected Incremental Net Monetary Benefit (EINMB) metric, we used $< 1\,\%$ mf prevalence with 5–10 % baseline prevalence in children (aged 5 years and above) as the comparator. We simulated transmission dynamics and morbidity associated with LF, including DALY burden for 30 sites and a TAS-like survey across those sites. We investigated different MDA coverages (65 % and 80 %) and different baseline LF prevalences. We evaluated WTP_{DALY} for DALYs averted, reflecting opportunity costs and adjusted for purchasing power parity (Turner et al., 2019) using $500 (Ghana), $2500 (Congo) and $5000 (Southern Africa) based on the provided country-specific percentage of GDP per capita estimate that underlies the DALY-4 estimation method by multiplying the total per individual DALY value times a specific proportion of the GDP per capita (Ochalek et al., 2018) for LMIC and $\text{WTP}_{\text{Elimination}}$ per unit increase in local elimination ranging from $0 to $10,000 (Antillon et al., 2021). The evaluation was performed for five different stopping thresholds (0.5 %, 1 %, 2 %, and 5 %) using *Culex* vector, although the primary focus was $< 0.5\,\%$ stopping

threshold. The analysis included samples of children, adults, and the entire eligible population, with prime focus on adults. For each scenario (different combinations of baseline prevalences and MDA coverages), the impact of the stopping threshold was analysed using a model-based approach, which assessed transmission dynamics, health outcomes and economic impacts. To address our key questions we focus on the adult age-group where we utilize the following inputs for the health-economic decision model:

1. Probability of elimination (Stolk et al., 2022), i.e., the probability of achieving local elimination within 20 years post-MDA if mf prevalence in a sample size of <1700 adults aged 20 years and above was below the stopping threshold.
2. Health impact evaluation through DALYs averted for morbidity by quantifying the overall disease burden due to lymphoedema, hydrocele and ADL by the years lived with disability. In this context, we assume that the years of life lost due to premature death is zero as death due to LF is rare.
3. Costs due to MDA rounds and TAS.

We primarily focus on two major tasks - (a) Compute the cost-effectiveness of lowering the stopping threshold to $< 0.5\,\%$ mf prevalence in adults, (b) Evaluate the uncertainty in total costs, DALYs averted and/or probability of elimination for the different strategies (baseline prevalences 5–10 %,10-20 % and 20–30 %) corresponding to the stopping threshold $< 0.5\,\%$ mf prevalence in adults. We note to address (a) from the CEA framework we use the EINMB metric (see Section 2) with fixed country-specific WTP_{DALY} and vary across an approximate range of $\text{WTP}_{\text{Elimination}}$ due to non-availability of data. Analogously, to address (a) using LOT framework in conjunction with PCA and LDA we classify the different baseline prevalences for the stopping threshold $< 0.5\,\%$ mf prevalence in adults. To address (b) from the VoI framework we use the EVSI metric (see Section 2) which is implemented using Moment Matching method (see Algorithm 1 in Appendix A). Due to the above-mentioned challenges in implementing Moment Matching (see Section 2.2) we alternatively use the LOT framework in conjunction with PCA and LDA (see Section 3). The comparator chosen is the parameter values related to stopping threshold $< 1\%$ mf prevalence in children for the 5–10 % baseline prevalences.

We briefly outline the LOT approach used in addressing the above two tasks as follows:

1. **Data Generation (see line 1 in Algorithms 2 and 3)**: Let $N = 1000$ be the number of iterations, $T = 20 \times 12 = 240$ (months) be the number of time steps, and choose the strategy

$$s = \begin{cases} 1 & \text{if baseline prevalence 5–10 \%} \\ 2 & \text{if baseline prevalence 10–20 \%} \\ 3 & \text{if baseline prevalence 20–30 \%} \end{cases}$$

Now, for each $i = \{1, \dots, N\}$ and $s = \{1, 2, 3\}$ we consider the following:

(a) Sample the model parameters $\boldsymbol{\theta}_i^{(s)} = (\theta_{i1}^{(s)}, \dots, \theta_{ik}^{(s)})$ using the transmission model TRANSFIL (Irvine et al., 2015) with parameters previously estimated (Irvine et al., 2016; Stolk et al., 2018) to represent transmission by *Culex* mosquitoes (for more details we refer Table C.16 in Appendix C). The labels are denoted as $\{y_i\}_{i=1}^{N} \in s$.

(b) Run the model for $N = 1000$ iterations with the suitable choice of the parameters $\boldsymbol{\theta}_i^{(s)}$ to obtain a vector of the average (over the population) for the possible outcomes of interest for stopping threshold $< 0.5\,\%$. In short, we denote $c_i^{(s)} \in \mathbb{R}^T$ as an observation of the average cost $C(\boldsymbol{\theta}_i^{(s)}) \in \mathbb{R}^T$, $a_i^{(s)}$ as an observation of the average DALYs$(\boldsymbol{\theta}_i^{(s)}) \in \mathbb{R}^T$, and e_i as an observation of the average probability of elimination $\mathbb{I}_{\text{Elimination}}(\boldsymbol{\theta}_i^{(s)})$.

(c) Construct the probability measure

$$\mu_i^{(s)} = \frac{1}{T}\sum_{j=1}^{T} \delta_{\left(c_{ij}^{(s)}, a_{ij}^{(s)}, e_{ij}^{(s)}\right)} \in \mathcal{P}(\mathbb{R}^3).$$

(d) Similarly, construct the reference measure

$$\sigma = \frac{1}{T}\sum_{j=1}^{T} \delta_{(\hat{c}_{ij}, \hat{a}_{ij}, \hat{e}_{ij})} \in \mathcal{P}(\mathbb{R}^3),$$

such that $\hat{c}_i \in \mathbb{R}^T$ is an observation of the average (over population) cost $C(\hat{\boldsymbol{\theta}}_i)$, $\hat{a}_i \in \mathbb{R}^T$ is an observation of the average (over population) DALYs$(\hat{\boldsymbol{\theta}}_i)$, and $\hat{e}_i \in \mathbb{R}^T$ is an observation of the average probability of elimination $\mathbb{I}_{\text{Elimination}}(\hat{\boldsymbol{\theta}}_i)$ where $\hat{\boldsymbol{\theta}}_i$ represents the uncertain parameter values for the ith point cloud, which are associated with the 5–10 % baseline prevalence and correspond to the stopping threshold $< 1\,\%$ mf prevalence in children.

2. **Computation of LOT embeddings (see Step 1 in Algorithms 2 and 3):** Secondly, we compute the projections $P(\mu_i^{(s)}) \in \mathbb{R}^{3T}$ as defined in (21) for each measure $\mu_i^{(s)}$ relative to the reference σ as defined above.
3. **Dimensionality Reduction using PCA (see Step 2 in Algorithms 2 and 3):** Thirdly, using these projections, we apply PCA to the LOT embeddings (projections) to extract the top $p = 15$ eigenvectors that capture the principal variations in the distributions of the outcomes.
4. **Classification using LDA (see Step 3 in Algorithms 2 and 3):** The projected eigenvectors obtained from PCA serve as training feature vectors, denoted as $\{\mathbf{X}_i\}_{i=1}^{L}$, where $L \in \{0.2N, 0.4N, 0.6N, 0.8N\}$ is the size of the training dataset. This approximately corresponds to the incremental costs and incremental DALYs averted and/or probability of elimination. The three classes of baseline prevalences (5–10 %, 10–20 %, and 20–30 %) are used as training labels (strategies), denoted as $\{y_i\}_{i=1}^{L}$. These feature vectors and labels are then used as inputs for LDA to classify the remaining (unseen) test dataset into the different baseline prevalences for $< 0.5\,\%$ stopping threshold.
5. **Compute EINMB, account for the uncertainty in the outcomes and estimate $\text{WTP}_{\text{Elimination}}$ for 65 % MDA coverage (see Step 4–5 in Algorithm 2 and Step 4 in Algorithm 3):** To compare with the EINMB within the CEA framework we represent the scatter plots recording the mean estimate of test data classified into the different baseline prevalence and provide the confidence ellipses to account for the uncertainty in the classification predictions. We note that for 80 % MDA coverage we use the fixed (plotted as hyperplane using data) WTP_{DALY} as chosen above and for 65 % MDA coverage we estimate the decision boundary from LDA (plotted as hyperplane) as the $\text{WTP}_{\text{Elimination}}$ (see Algorithm 2 in Appendix A for more details).
6. **Compute EVSI (see Step 5–6 in Algorithm 3)**: To define an equivalent notion of the EVSI, we extend the proposed algorithm by incorporating additional samples through bootstrapping. We compute the projections for the bootstrapped samples that are subsequently reduced to a lower-dimensional representation using PCA, followed by classification using LDA. Within this framework, we define a utility function based on model performance, quantifying the number of correctly classified instances across predicted class labels.

As a baseline, we consider $N = 500$ independent draws in the transmission model and perform steps 1–4, training an LDA classifier (using a 20:80 train:test split ratio). To evaluate the effect of additional information, we generate bootstrapped samples of increasing sizes (500, 1000, and 1500 draws), and repeat steps 1–4, progressively adjusting the train:test split ratio from 20:80 to 80:20. The EVSI is then computed as the difference between the expected utility obtained from the augmented bootstrapped samples and that from the original baseline sample. This formulation provides a principled way to assess the value of acquiring additional information, allowing the evaluation of whether the proposed strategy (baseline prevalence) yields a measurable improvement in predictive

performance. For implementation details, refer to Algorithm 3 in Appendix A.

5. Results

The impact of MDA on the interruption of LF transmission and reduction of the disease burden using DALYs is dependent on the threshold criteria defined for passing the TAS, as illustrated in the example of a setting with a baseline prevalence of 5–10 % and 80 % MDA coverage of a single population size of 1000 for 10 simulations (Fig. 2).

In Fig. 3 (circles) replicating the characteristics of an EU, we find that the probability for local elimination at 5–10 % baseline prevalence with 80 % MDA coverage and a threshold of < 0.5 % mf prevalence was 89.2 % (≥5 years), 91.8 % (≥20 years), and 90.72 % (entire eligible population). For a threshold of < 1 % mf prevalence, it was 80.05 %, 83.8 %, and 81.76 %, respectively. Lowering the threshold increases the probability for local elimination across different age-groups. These trends follow across different baseline prevalences, and MDA coverages (refer B.2, B.3, and B.4 in Appendix B).

Additionally in Fig. 3 (triangles), a lower threshold results in fewer MDA rounds (7 MDA rounds or more) and surveys (4 surveys or more) due to reduced probability of restarting after stopping. However, for the lowest baseline prevalence, restarting MDA is unlikely for any threshold for children and adults, with slightly higher costs for the lower threshold due to extra standard MDA rounds (5 MDA rounds) needed. For the entire eligible population, higher thresholds for low baseline levels result in increased costs, primarily due to the potential need for MDA restarts, where random events have a greater influence on transmission dynamics (Collyer et al., 2020). In general, more restarts occur at higher baseline prevalences and lower MDA coverage for all thresholds due to the stochastic nature of the model dynamics accounting for higher transmission and increased treatment rounds to achieve elimination (refer Tables B.5, B.6 and B.7, in Appendix B).

Fig. 3 (squares) shows mean DALYs averted across different thresholds for 80 % MDA coverage. Lowering the threshold results in more DALYs averted due to a small change in the morbidity prevalence. Trends are similar for 65 % MDA coverage. The limited change in the incidence of DALY burden drops dramatically as average worm burdens drop, so most morbidity prevalence is due to historic infection, before the MDA.

To evaluate costs, health impact, and monetization benefits of local elimination, we use expected incremental net monetary benefit (EINMB). Higher EINMB indicates optimal cost-effectiveness at a given WTP_{DALY}. Our findings (Fig. 4a) show that at 80 % coverage, switching to a lower threshold is cost-effective across all baseline prevalences, keeping costs per DALY averted below national WTP thresholds (positive EINMB). Variability in results is due to demographic factors such as age, treatment strategy, and population growth (Stone et al., 2016). At 65 % coverage (Fig. 4b), more rounds and surveys suggest switching to a lower threshold is cost-effective based on WTP per 1 % increase in local elimination probability, aligning with GPELF goals (refer Tables B.8, B.9 in Appendix B). For WTPs of approximately $4200, $3000, and $1000 per 1 % increase in local elimination for different baseline prevalence, switching is recommended (Fig. 4b, black solid line).

Health economic decision-analytic models are used to estimate the expected net benefits of competing decision options. The true values of the input parameters of such models are rarely known with certainty,

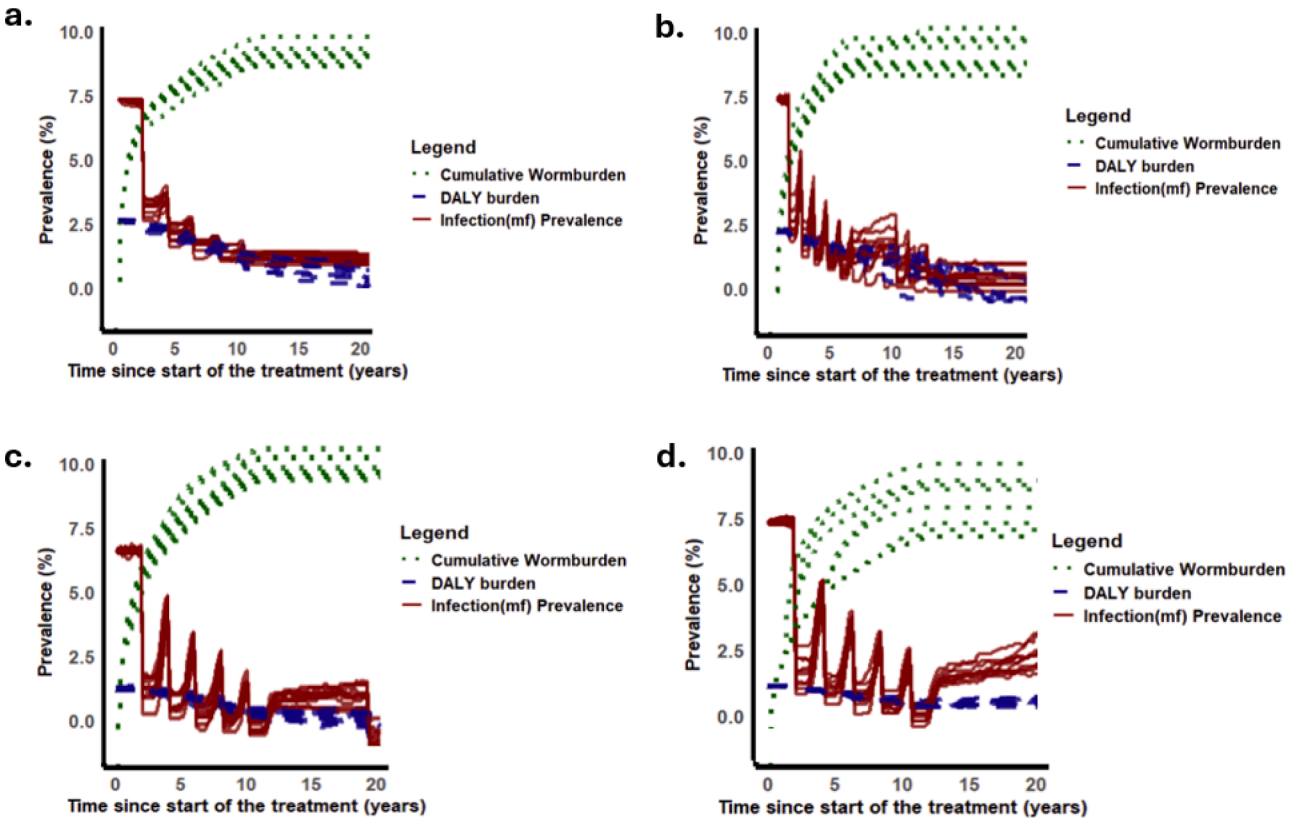

Fig. 2. Simplified timeline plots for a single population of size 1000 for 10 simulations illustrating the model-predicted temporal trends in mf prevalence (solid red lines), DALY burden are computed as the morbidity prevalence of lymphoedema, hydrocele and acute adenolymphangitis (dashed blue lines) times the disability weights (Network, 2020) and cumulative wormburden (dotted green lines) for 5–10 % mf prevalence using (a) < 0.5 %, (b) < 1 %, (c) < 2 % and (d) < 5 % as the stopping threshold criteria for TAS with 80 % MDA coverage for a sample of adults.

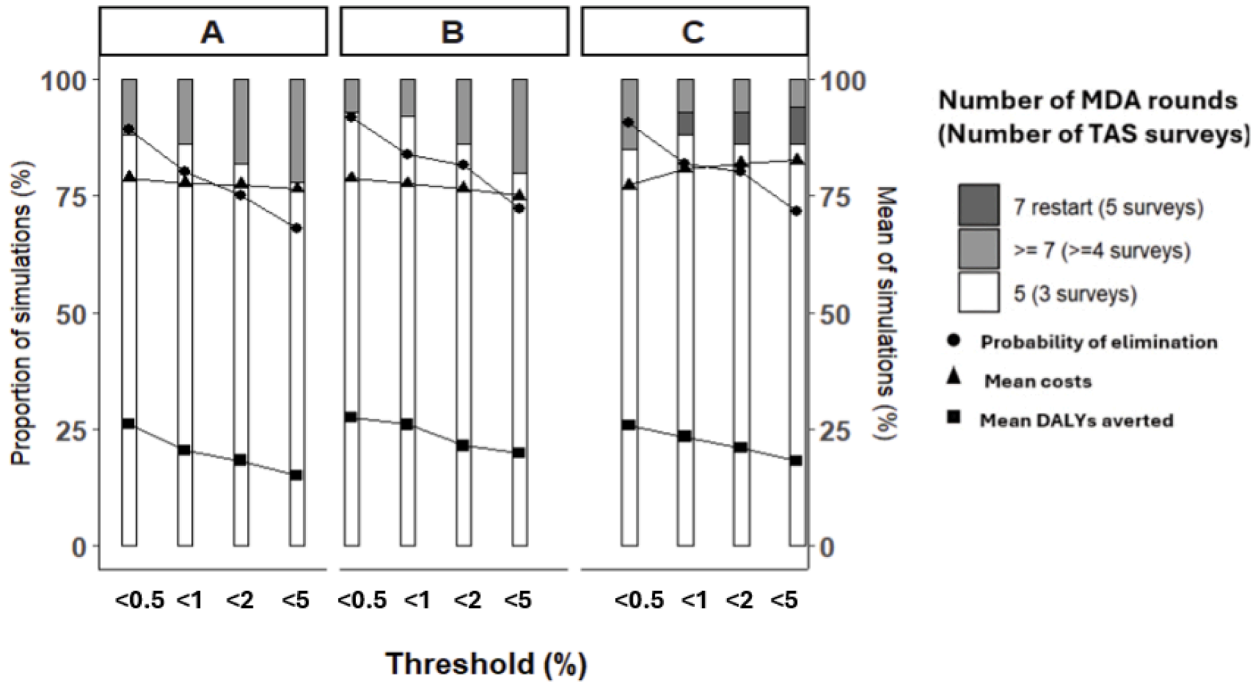

Fig. 3. Epidemiological outcomes of an eligible population post-surveillance for different MDA stopping thresholds and TAS corresponding to 5–10 % baseline prevalence and 80 % MDA coverage. Outcomes are represented by different symbols: probability of elimination (circles), mean cost (triangles), and mean DALYs averted (squares) and sampled for three age-groups: (A) children over 5 years old, (B) adults over 20 years old, and (C) random sampling across all age groups. The minimum number of MDA rounds and TAS required to achieve elimination of infection are represented with varying shades: white for standard treatment (5 MDA rounds and 3 TAS), light grey for moderate treatment (at least 7 MDA rounds and 4 TAS), and dark grey for extensive treatment (7 MDA rounds with restarts and 5 TAS). A restart indicates infection resurgence, requiring additional MDA rounds and TAS to verify elimination. Here, we represent the mean outcomes by randomly sampling approximately 30 sites with 40–60 people stratified by age per site to replicate the characteristics of an evaluation unit (EU, < 500,000 people). Note that costs and DALYs averted are normalized to the same scale for improved visualization in the plots.

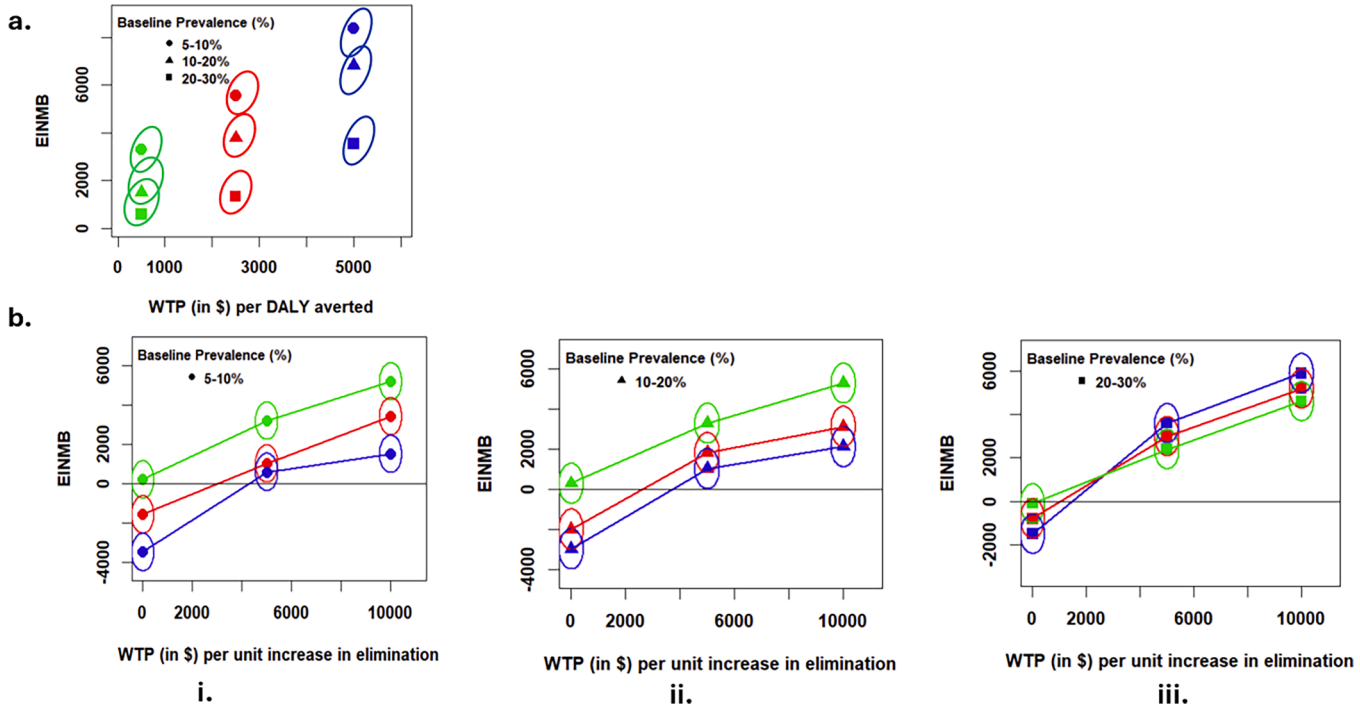

Fig. 4. a. EINMB based on the WTP for range of DALY averted for morbidity: $500 (green), $2500 (red), $5000 (blue) for 5–10 % (circles) , 10–20 % (triangles) and 20–30 % (squares) baseline prevalence for a sample of adults. b. EINMB based on the WTP for 1 % increase in probability of elimination from $0-$10,000 and the WTP_{DALY}: $500 (green), $2500 (red), $5000 (blue) for sample of adults for (i) 5–10 % - circles (ii) 10–20 % - triangles (iii) 20–30 % - squares baseline prevalences comparing < 0.5 % threshold in a sample of adults with respect to < 1 % threshold of mf prevalence in children (comparator). The confidence ellipses account for the 95 % confidence interval obtained from 1000 random draws of costs, DALYs averted and probability of elimination using Monte Carlo simulations. Note: In Figure b. the intersection of the different coloured lines with the black solid lines gives the optimal $WTP_{Elimination}$.

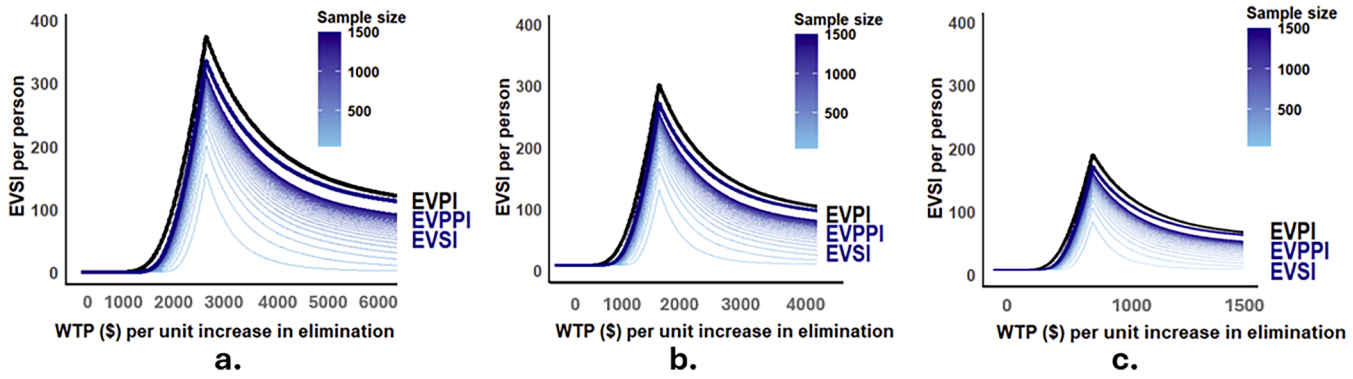

Fig. 5. EVSI estimated using the Moment Matching method, for three baseline prevalence ranges: (a) 5–10 % with WTP_{DALY} =$500, (b) 10–20 % with WTP_{DALY} =$2500, and (c) 20–30 % with WTP_{DALY} =$5000 by varying the number of iterations (sample size) used in the transmission model (equivalently, the epidemiological outcomes used as inputs in the health-economic decision model). The EVSI quantifies the value of reducing uncertainty when lowering the stopping threshold for TAS to < 0.5 % mf prevalence in a sample of adults. Estimates are shown across a range of $WTP_{Elimination}$, assuming 65 % MDA coverage.

Table 1
Comparison of the EVSI per person and computational time for different number of independent iterations (sample sizes) used in the transmission model for post-MDA surveillance. The epidemiological outcomes of the transmission model are evaluated using Moment Matching (MM), Nested Monte Carlo (MC), and LOT + PCA + LDA (80:20 train:test split) methods to obtain EVSI per person. The results are based on the dynamics for < 0.5 % stopping threshold with 65 % MDA coverage across different baseline prevalence.

Baseline prevalence	WTP_{DALY},$WTP_{Elimination}$	Sample size (500)			Sample size (1000)			Sample size (1500)		
		MC	MM	LOT + LDA	MC	MM	LOT + LDA	MC	MM	LOT + LDA
EVSI per Person										
5–10 %	$500, $3000	280	255	260	310	290	295	325	310	320
10–20 %	$2500, $2000	220	190	197	270	245	252	290	267	280
20–30 %	$5000, $500	185	152	160	200	172	180	215	185	198
Computational Time										
5–10 %	$500, $3000	>8 h	≈2 min	≈1 min	>8 h	≈2 min	≈1 min	>8 h	≈2 min	≈1 min
10–20 %	$2500, $2000	>8 h	≈2 min	≈1 min	>8 h	≈2 min	≈1 min	>8 h	≈2 min	≈1 min
20–30 %	$5000, $500	>8 h	≈2 min	≈1 min	>8 h	≈2 min	≈1 min	>8 h	≈2 min	≈1 min

and it is often useful to quantify the value to the decision maker of reducing uncertainty by collecting new data. In the context of understanding how to measure the baseline prevalence for < 0.5 % stopping threshold with precision, we need a handle to quantify the uncertainty arising from the costs due to MDA rounds and surveys alongside the DALY averted and the unit increase in probability of elimination. In this light, the value of the proposed research design for every additional sample size can be quantified by the EVSI metric as defined in Section 2. In Fig. 5(a), the EVSI reaches its maximum at a $WTP_{Elimination}$ of approximately $2500–$3000 for a baseline prevalence of 5–10 %. In contrast, in Fig. 5(b), the EVSI peaks at a lower $WTP_{Elimination}$ of around $1500–$2000 for a baseline prevalence of 10–20 %. Finally, in Fig. 5(c), the peak EVSI occurs at an even lower $WTP_{Elimination}$ of approximately $500–$1000 for baseline prevalence levels of 20–30 %. This pattern indicates that the value of additional information as measured by EVSI is greatest at lower prevalence levels and lower stopping thresholds, where it plays a more critical role in reducing uncertainty in MDA stopping decisions.

Additionally, in Table 1, we find that Moment Matching is much faster than the benchmark nested Monte–Carlo method, which converges to the same EVSI at larger sample sizes (relative difference is much lower for 5–10 % baseline prevalence than 20–30 % baseline prevalence as expected due to additional benefits at lower baseline prevalence). Fig. 5 demonstrates the theoretically expected trend in the relationship between EVSI and sample size. As sample size increases, the EVSI (shown by the dark blue lines) rises and gradually converges toward the EVPPI (represented by the thick dark blue line just below the EVPI), as described in Section 2. This convergence reflects the notion that larger studies yield more valuable information. In the limit, the value of this information approaches the maximum possible gain from resolving the uncertainty: specifically, the uncertainty about achieving elimination for different baseline prevalence at the stopping threshold < 0.5 % mf prevalence.

In order to further test the robustness of the MDA stopping decision based on the cost-effectiveness of lower stopping threshold (< 0.5 % mf prevalence in adults), we rely on the Linear Wasserstein Framework in conjunction with PCA and LDA. In Fig. 6 a, we find that the scattergram of the test data (incremental costs to incremental DALYs averted) for 80 % MDA coverage show that the lower stopping thresholds (< 0.5 %) for different baseline prevalences (represented as circles, triangles and squares) are cost-effective using the fixed country-specific WTP_{DALY} ranging from $500-$5,000. Each symbol represents the mean incremental costs to incremental DALYs averted alongside the confidence ellipses to account for the uncertainty in the distributions to obtain a standardized comparison to the EINMB metric in Fig. 4a. Likewise, in Fig. 6b. we find that the scattergram of the test data (incremental costs with incremental DALYs averted in addition to the probability of elimination) for 65 % MDA coverage shows that the < 0.5 % stopping thresholds for different baseline prevalences are cost-effective with a narrower estimated $WTP_{Elimination}$ per unit increase in the probability of elimination from $500-$3000. Similarly, each symbol here represents the mean incremental costs to the incremental probability of elimination alongside the confidence ellipses accounting for uncertainty to obtain a standardized comparison with the EINMB metric in Fig. 4b. We note that the results illustrated in Fig. 6a and b were calculated using Algorithms 2 and 3 (in Supplementary Appendix A) with the training sample size 80 % and stopping thresholds < 0.5 % mf prevalence in adults for different baseline prevalence. For the trade-off between the different combinations of

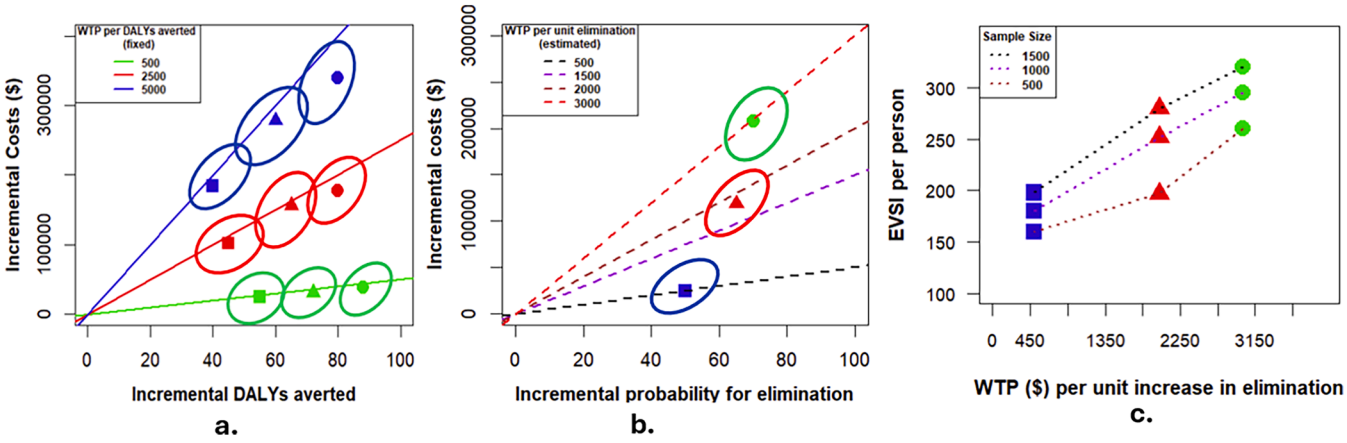

Fig. 6. A summary figure illustrating the variations in the cost-effectiveness of the < 0.5 % stopping threshold for TAS using the Linear Wasserstein framework. The LOT embeddings are projected into a lower-dimensional space using PCA, where the first and second variations correspond to the most important features (namely, incremental total costs and incremental DALYs averted). The feature vector is composed of the ratio of incremental costs to incremental DALYs averted for morbidity with fixed WTP_{DALY} for morbidity given by \$500 (green), \$2500 (red), \$5000 (blue). This reduced feature vector is then classified using LDA, with the plots demonstrated for an 80:20 train-test split. The baseline prevalence represented as circles (5–10 %), triangles (10–20 %), and squares (20–30 %), are used as labels for classification corresponding to the < 0.5 % stopping threshold. a. Scatter plot of incremental costs versus incremental DALYs averted for 80 % MDA coverage using fixed WTP_{DALY} for morbidity. b. Scatter plot of incremental costs versus incremental unit increase in probability of elimination for 65 % MDA coverage using (estimated) $WTP_{Elimination}$ per unit elimination and (fixed) WTP_{DALY}. c. EVSI per person for the (estimated) $WTP_{Elimination}$ per unit elimination and (fixed) WTP_{DALY} for morbidity with varying number of independent iterations (sample size) used in the transmission model. The ellipses estimated from the covariance matrix and the mean vectors of each baseline prevalence (class labels) denote the 95 % confidence intervals accounting for the uncertainty and variability within the distribution of each class. Each symbol represents the mean of the test data (incremental costs to incremental DALYs averted or probability of elimination) classified by their respective baseline prevalences.

baseline prevalence and the values of WTP_{DALY}, $WTP_{Elimination}$, we refer to Table B.10 and Table B.11 in Supplementary Appendix B.

Furthermore, in Table 1, we compare the EVSI values obtained using LOT alongside the benchmark methods: Moment Matching and nested Monte Carlo for different baseline prevalence at < 0.5 % stopping threshold. We find that the LOT method is much faster than the other two, although both converge to the standard EVSI (evaluated by Monte Carlo) at larger sample sizes for all baseline prevalence. Fig. 6c presents the EVSI per person for varying number of independent iterations (sample size) used in the transmission model, alongside the estimated optimal $WTP_{Elimination}$ (see Algorithm 3 for details). This figure highlights how the additional information obtained through bootstrapping, reflected in different sample sizes, impacts EVSI estimation. A key insight from the figure is that EVSI tends to increase when baseline prevalence is lower for the stopping threshold of < 0.5 % in adults. This indicates that at lower prevalence levels, decision-makers may be willing to allocate a higher $WTP_{Elimination}$, as the value of additional information (represented by larger sample sizes) strengthens the case for further research. This rise in EVSI underscores the increased benefits of elimination, including more DALYs averted for lower baseline prevalence, as illustrated in Fig. 3. In summary, when baseline prevalence is low for stopping threshold < 0.5 %, the potential health and economic advantages of elimination become more significant, making the investment in further data collection even more valuable. Consequently, in Tables B.12 and B.13, B.14, B.15 (refer to Supplementary Appendix B), we present the classification error to accurately predict the baseline prevalences (class labels) for < 0.5 % stopping threshold, which decreases as the training sample sizes increase. This improvement enhances the power of the utility function, which reflects the additive benefits gained with lowering the baseline prevalence for < 0.5 % stopping threshold, as indicated by the EVSI metric. This effect is achieved by training the LDA classifier with different fractions of the sample sizes, demonstrating that larger training datasets for the lower baseline prevalence lead to more accurate classifications and thus greater potential benefits from additional data.

6. Discussion

The probability of local elimination is determined by stopping thresholds, which are crucial for many disease control policies. That being said, it would be worthwhile to examine the effects of a lower threshold on overall program costs, as well as whether it increases the likelihood of local elimination. The application of such a lower threshold in China (World Health Organization Regional Office for the Western Pacific and Control of Lymphatic Filariasis in China Editorial Board, 2003), and its significance in effective LF control, serves as examples of the potential advantages of a lower threshold, which this study highlights. However, the GPELF can use this example to gather crucial data to establish standards for assessing whether MDA has successfully reduced infection prevalence to a point where recrudescence is unlikely to occur.

As we reduce the mf prevalence threshold from < 1 % to < 0.5 %, the likelihood of local elimination increases, according to our analysis of the effects of various stopping thresholds for TAS across 30 sites. A lower threshold reduces both DALY burden and the probability of programmatic restart, despite requiring more rounds. Employing the defined EINMB metric for CEA reveals that switching to a lower threshold is economical at 80 % MDA coverage. However, for 65 % MDA coverage, extra benefits are needed, such as utilizing the $WTP_{Elimination}$ per unit increase for elimination. The limited amount of data, especially on systematic non-adherence, $WTP_{Elimination}$ and wider disease impacts like mental illness, is the reason for the conservative morbidity estimates (Ton et al., 2015; Koschorke et al., 2022).

The expanded use of CEA in healthcare faces several challenges. First, decision-makers must account for social concerns like prioritizing the sick and reducing health disparities by integrating more social concerns into CEA techniques. Second, current CEA practices, which are focused on evaluating new strategies or technologies, often overlook signs of resource misallocation. Third, assessing the broad range of interventions needed for CEA to improve allocative efficiency can be prohibitively expensive and time-consuming. Additionally, many CEA

studies produce context-specific results, limiting their applicability to different populations. Progress towards providing timely, affordable information on the costs and effects of various interventions remains limited, particularly for LMICs (Turner et al., 2021; World Health Organization et al., 2003). A key limitation of uncertainty analysis within the CEA framework is the inherent circularity when we utilize the TRANSFIL model for obtaining our epidemiological outcomes (such as costs, DALYs averted, and probability of elimination). The model incorporates the uncertain parameters θ, and their impact on the outcome is directly dictated by its underlying dynamics. However, the identification of these key parameters θ and the selection of an optimal strategy rely on the same foundational knowledge that informs the model itself, which is usually gauged from prior knowledge.

On the other hand, the Linear Wasserstein Framework, despite being mathematically rigorous, has its own limitations. Firstly, the framework makes several modelling assumptions. Namely, that the distance should be proportional to the cost of translations. This can make the distance sensitive to outliers. Secondly, the Linear Wasserstein distance is also an approximation of the Wasserstein distance, and this approximation may deteriorate depending on the local curvature. Thirdly, the Monge formulation is ill-posed when the probability measures are discrete, and the transport maps may not exist. To avoid this degeneracy, we need to consider an equal number of time steps ($T = 240$ (months)) for each probability measure $\mu_i^{(s)}$ for the different baseline prevalences (classes/labels). We note that if this is not assumed, it may disproportionately affect the robustness of the results, if an imbalance in masses exists. Fourthly, to simplify the computation, we select the reference measure σ as the distribution corresponding to a 5–10 % baseline prevalence at < 1 % stopping threshold in children. While it would be ideal to use $\sigma^{(s)}$ with various combinations of baseline prevalences at the < 1 % stopping threshold, doing so would significantly increase computational complexity. Additionally, the projections computed in this manner would belong to different linear spaces which would not allow for a fair comparison. A further extension to this framework could be to generate future projections of the model simulations for different baseline prevalences using fewer runs to save the computational power of the TRANSFIL model from a Bayesian perspective (Park and Thorpe, 2018) or use state-of-the-art methods such as graph-based semi-supervised methods (Calder et al., 2020) that can leverage the advantages of this geometric embedding when very little information on the data is provided, so that it can learn the geometry of the underlying point cloud data effectively.

Our study assumes constant survey implementation costs, excluding potential out-of-pocket expenses and future cost changes (Sawers and Stillwaggon, 2020). Despite challenges in estimating precise costs for MDA and TAS due to incomplete records and data access issues, simulations help in understanding the TAS threshold's impact on stopping MDA. A limitation in this study is the exclusion of vector control benefits, which remain debated. While some studies suggest combined MDA and vector control benefits in low endemic regions, others find no added advantage over MDA alone, upon which further research is needed. Another major limitation is that our modelling study relies on *Culex* vector due to its increased efficiency in transmission. Although the direct implication of *Culex* species in the transmission of LF in West and Central Africa is still not well documented (Samy et al., 2016; Appawu et al., 2001), in East Africa, *Culex* species particularly *Cx. quinquefasciatus* is known to have a major role in LF transmission (Derua et al., 2017; Mwakitalu et al., 2013). With a changing climate, associated with increased traffic between East and West African countries and rapid expansion of this species in urban settings, it is becoming crucial to assess the role of *Culex* species in the transmission of diseases like LF. We also restrict our analysis to the IA drug, but studies (Turner et al., 2024) for oncho have found that IA may not lead to elimination of transmission (EoT) in all endemic areas, and moxidectin-based strategies could accelerate progress toward EoT and reduce programmatic delivery costs compared with ivermectin-based strategies. We also acknowledge the benefits of the three-drug combination IDA, but it presents specific challenges in implementing the survey design, which could result in reductions in mf prevalence. However, further evidence is needed to confirm this, as noted in (Stolk et al., 2018).

Despite these drawbacks, our research emphasizes how important it is to choose the right framework for uncertainty quantification when making decisions, especially when it comes to disease interventions, particularly LF. It is also essential to comprehend the dynamics of local elimination post-threshold crossing and how it interacts with LF interventions. Our research indicates that although there is a long transient phase involved in the path to LF local elimination post-MDA surveillance, lower thresholds may help programs achieve their objectives. In addition, we also propose the need for a better framework to quantify the uncertainty inherent in the model parameters to analyse the cost-effectiveness of lowering the stopping threshold in LF.

Funding

MCAO, MG and TDH were supported by funding from the Bill & Melinda Gates Foundation (INV-030046), via the NTD Modelling Consortium. TDH is supported by funding from the Li Ka Shing Foundation at the Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford. MT would like to acknowledge the support of the Leverhulme Trust Research through the Project Award "Robust Learning: Uncertainty Quantification, Sensitivity and Stability" (RPG-2024-051) and the EPSRC Mathematical and Foundations of Artificial Intelligence Probabilistic AI Hub (EP/Y007174/1). LP gratefully acknowledges funding from the Wellcome Trust and Royal Society Sir Henry Dale Fellowship (202562/Z/16/Z), the Wellcome Trust Discovery Award "Harnessing epidemiological and genomic data for understanding of respiratory virus transmission at multiple scales" (227438/Z/23/Z) and the UKRI Impact Acceleration Award (IAA 386). KBP is supported by the Medical Research Foundation (MRF-160-0017-ELP-POUW-C0909). MCAO acknowledges the receipt of funding obtained from the Health Data Research UK-The Alan Turing Institute Wellcome (218529/Z/19/Z) and the Cambridge Trust scholarship from the Commonwealth European and International Trust (CCEIT). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

CRediT authorship contribution statement

Mary Chriselda Antony Oliver: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization; **Matthew Graham:** Writing – review & editing, Validation, Software; **Ioanna Manolopoulou:** Writing – review & editing; **Graham F. Medley:** Writing – review & editing; **Lorenzo Pellis:** Writing – review & editing; **Koen B. Pouwels:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization; **Matthew Thorpe:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization; **T. Deirdre Hollingsworth:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary material

Supplementary material associated with this article can be found in the online version at 10.1016/j.jtbi.2025.112197.

References

Antillon, M., Huang, C.-I., Rock, K.S., Tediosi, F., 2021. Economic evaluation of disease elimination: an extension to the net-benefit framework and application to human African trypanosomiasis. Proc. Natl. Acad. Sci. 118 (50), e2026797118.

Antony Oliver, M.C., Graham, M., Gass, K.M., Medley, G.F., Clark, J., Davis, E.L., Reimer, L.J., King, J.D., Pouwels, K.B., Hollingsworth, T.D., 2024. Reducing the antigen prevalence target threshold for stopping and restarting mass drug administration for lymphatic filariasis elimination: a model-based cost-effectiveness simulation in Tanzania, India and Haiti. Clin. Infect. Dis. 78 (Suppl. 2), S160–S168.

Appawu, M.A., Dadzie, S.K., Baffoe-Wilmot, A., Wilson, M.D., 2001. Lymphatic filariasis in Ghana: entomological investigation of transmission dynamics and intensity in communities served by irrigation systems in the upper east region of Ghana. Trop. Med. Int. Health 6 (7), 511–516.

Basu, S., Kolouri, S., Rohde, G.K., 2014. Detecting and visualizing cell phenotype differences from microscopy images using transport-based morphometry. Proc. Natl. Acad. Sci. 111 (9), 3448–3453.

Bojke, L., Claxton, K., Sculpher, M., Palmer, S., 2009. Characterising structural uncertainty in decision analytic models: a review and application of methods. Value Health 12 (5), 739–749.

Brady, M.A., Stelmach, R., Davide-Smith, M., Johnson, J., Pou, B., Koroma, J., Frimpong, K., Weaver, A., 2017. Costs of transmission assessment surveys to provide evidence for the elimination of lymphatic filariasis. PLoS Negl. Trop. Dis. 11 (2), e0005097.

Brennan, A., Kharroubi, S., O'hagan, A., Chilcott, J., 2007. Calculating partial expected value of perfect information via Monte Carlo sampling algorithms. Med. Decis. Mak. 27 (4), 448–470.

Calder, J., Cook, B., Thorpe, M., Slepcev, D., 2020. Poisson learning: graph based semi-supervised learning at very low label rates. In: International Conference on Machine Learning. PMLR, pp. 1306–1316.

Chi, Y.-L., Blecher, M., Chalkidou, K., Culyer, A., Claxton, K., Edoka, I., Glassman, A., Kreif, N., Jones, I., Mirelman, A.J., et al., 2020. What next after GDP-based cost-effectiveness thresholds? Gates Open Res. 4, 176.

Chu, B.K., Hooper, P.J., Bradley, M.H., McFarland, D.A., Ottesen, E.A., 2010. The economic benefits resulting from the first 8 years of the global programme to eliminate lymphatic filariasis (2000–2007). PLoS Negl. Trop. Dis. 4 (6), e708.

Claxton, K., Sculpher, M., McCabe, C., Briggs, A., Akehurst, R., Buxton, M., Brazier, J., O'Hagan, T., 2005. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. Health Econ. 14 (4), 339–347.

Collyer, B.S., Irvine, M.A., Hollingsworth, T.D., Bradley, M., Anderson, R.M., 2020. Defining a prevalence level to describe the elimination of lymphatic filariasis (LF) transmission and designing monitoring & evaluating (M&E) programmes post the cessation of mass drug administration (MDA). PLoS Negl. Trop. Dis. 14 (10), e0008644.

Davis, E.L., Reimer, L.J., Pellis, L., Hollingsworth, T.D., 2019. Evaluating the evidence for lymphatic filariasis elimination. Trends Parasitol. 35 (11), 860–869.

Derua, Y.A., Rumisha, S.F., Batengana, B.M., Max, D.A., Stanley, G., Kisinza, W.N., Mboera, L. E.G., 2017. Lymphatic filariasis transmission on Mafia Islands, Tanzania: evidence from xenomonitoring in mosquito vectors. PLoS Negl. Trop. Dis. 11 (10), e0005938.

Dyson, L., Stolk, W.A., Farrell, S.H., Hollingsworth, T.D., 2017. Measuring and modelling the effects of systematic non-adherence to mass drug administration. Epidemics 18, 56–66.

Fefferman, C., Mitter, S., Narayanan, H., 2016. Testing the manifold hypothesis. J. Am. Math. Soc. 29 (4), 983–1049.

Fintzi, J., Wakefield, J., Minin, V.N., 2022. A linear noise approximation for stochastic epidemic models fit to partially observed incidence counts. Biometrics 78 (4), 1530–1541. https://doi.org/10.1111/biom.13538

Gambhir, M., Michael, E., 2008. Complex ecological dynamics and eradicability of the vector borne macroparasitic disease, lymphatic filariasis. PLoS One 3 (8), e2874.

Griffin, S., Fusco, F., Naidoo, B., Taylor, M., Walker, S.M., 2020. Does health technology assessment guidance give adequate consideration to decisions about less costly and less effective alternatives?

Heath, A., Kunst, N., Jackson, C., Strong, M., Alarid-Escudero, F., Goldhaber-Fiebert, J.D., Baio, G., Menzies, N.A., Jalal, H., 2020. Calculating the expected value of sample information in practice: considerations from 3 case studies. Med. Decis. Mak. 40 (3), 314–326.

Heath, A., Manolopoulou, I., Baio, G., 2016. Estimating the expected value of partial perfect information in health economic evaluations using integrated nested laplace approximation. Stat. Med. 35 (23), 4264–4280.

Heath, A., Manolopoulou, I., Baio, G., 2019. Estimating the expected value of sample information across different sample sizes using moment matching and non-linear regression. Med. Decis. Mak. 39 (4), 347–359.

Irvine, M.A., Njenga, S.M., Gunawardena, S., Njeri Wamae, C., Cano, J., Brooker, S.J., Deirdre Hollingsworth, T., 2016. Understanding the relationship between prevalence of microfilariae and antigenaemia using a model of lymphatic filariasis infection. Trans. R. Soc. Trop. Med. Hyg. 110 (2), 118–124.

Irvine, M.A., Reimer, L.J., Njenga, S.M., Gunawardena, S., Kelly-Hope, L., Bockarie, M., Hollingsworth, T.D., 2015. Modelling strategies to break transmission of lymphatic filariasis-aggregation, adherence and vector competence greatly alter elimination. Parasit. Vectors 8, 1–19.

Irvine, M.A., Stolk, W.A., Smith, M.E., Subramanian, S., Singh, B.K., Weil, G.J., Michael, E., Hollingsworth, T.D., 2017. Effectiveness of a triple-drug regimen for global elimination of lymphatic filariasis: a modelling study. Lancet Infect. Dis. 17 (4), 451–458.

Jackson, C., Presanis, A., Conti, S., De Angelis, D., 2019. Value of information: sensitivity analysis and research design in Bayesian evidence synthesis. J. Am. Stat. Assoc 114 (528), 1436–1449.

Jackson, C.H., Baio, G., Heath, A., Strong, M., Welton, N.J., Wilson, E.C., 2022. Value of information analysis in models to inform health policy. Annu. Rev. Stat. Appl. 9 (1), 95–118.

King, C.L., Suamani, J., Sanuku, N., Cheng, Y.-C., Satofan, S., Mancuso, B., Goss, C.W., Robinson, L.J., Siba, P.M., Weil, G.J., et al., 2018. A trial of a triple-drug treatment for lymphatic filariasis. N. Engl. J. Med. 379 (19), 1801–1810.

Kolouri, S., Park, S.R., Thorpe, M., Slepcev, D., Rohde, G.K., 2017. Optimal mass transport: signal processing and machine-learning applications. IEEE Signal Process. Mag. 34 (4), 43–59.

Kolouri, S., Rohde, G.K., 2015. Transport-based single frame super resolution of very low resolution face images. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4876–4884.

Koschorke, M., Al-Haboubi, Y.H., Tseng, P.-C., Semrau, M., Eaton, J., 2022. Mental health, stigma, and neglected tropical diseases: a review and systematic mapping of the evidence. Front. Trop. Dis. 3, 808955.

Kunst, N., Wilson, E. C.F., Glynn, D., Alarid-Escudero, F., Baio, G., Brennan, A., Fairley, M., Goldhaber-Fiebert, J.D., Jackson, C., Jalal, H., et al., 2020. Computing the expected value of sample information efficiently: practical guidance and recommendations for four model-based methods. Value Health 23 (6), 734–742.

May, R.M., 1991. Infectious Diseases of Humans: Dynamics and Control. Oxford University Press.

McKay, M.D., 1992. Latin hypercube sampling as a tool in uncertainty analysis of computer models. Proceedings of the 1992 Winter Simulation Conference, pp. 557–564.

Michael, E., Singh, B.K., 2015. Heterogeneous dynamics, parameter sloppiness, robustness/fragility trade-offs, and the eradication of the macroparasitic disease, lymphatic filariasis. BMC Med 14 (1), 1–23.

Moosmüller, C., Cloninger, A., 2023. Linear optimal transport embedding: provable wasserstein classification for certain rigid transformations and perturbations. Inf. Inference 12 (1), 363–389.

Mwakitalu, M.E., Malecela, M.N., Pedersen, E.M., Mosha, F.W., Simonsen, P.E., 2013. Urban lymphatic filariasis in the metropolis of dar es Salaam, Tanzania. Parasit. Vectors 6, 1–13.

Nelsen, R.B., 2006. An Introduction to Copulas. Springer.

Network, G., 2020. Global Burden of Disease Study 2019 (GBD 2019) Disability Weights. Institute for Health Metrics and Evaluation (IHME), Seattle, United States of America.

Ochalek, J., Lomas, J., Claxton, K., 2018. Estimating health opportunity costs in low-income and middle-income countries: a novel approach and evidence from cross-country data. BMJ Glob. Health 3 (6), e000964.

Ozolek, J.A., Tosun, A.B., Wang, W., Chen, C., Kolouri, S., Basu, S., Huang, H., Rohde, G.K., 2014. Accurate diagnosis of thyroid follicular lesions from nuclear morphology using supervised learning. Med. Image Anal. 18 (5), 772–780.

Park, S., Thorpe, M., 2018. Representing and learning high dimensional data with the optimal transport map from a probabilistic viewpoint. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7864–7872.

Pearl, J., 1988. Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann.

Pichon-Riviere, A., Drummond, M., Palacios, A., Garcia-Marti, S., Augustovski, F., 2023. Determining the efficiency path to universal health coverage: cost-effectiveness thresholds for 174 countries based on growth in life expectancy and health expenditures. Lancet Glob. Health 11 (6), e833–e842.

Rothery, C., Strong, M., Koffijberg, H.E., Basu, A., Ghabri, S., Knies, S., Murray, J.F., Schmidler, G. D.S., Steuten, L., Fenwick, E., 2020. Value of information analytical methods: report 2 of the ISPOR value of information analysis emerging good practices task force. Value Health 23 (3), 277–286.

Samy, A.M., Elaagip, A.H., Kenawy, M.A., Ayres, C. F.J., Peterson, A.T., Soliman, D.E., 2016. Climate change influences on the global potential distribution of the mosquito culex quinquefasciatus, vector of west nile virus and lymphatic filariasis. PLoS One 11 (10), e0163863.

Sawers, L., Stillwaggon, E., 2020. Economic costs and benefits of community-based lymphedema-management programs for lymphatic filariasis in India. Am. J. Trop. Med. Hyg. 103 (1), 295.

Semochkina, D., Forrester, A. I.J., Woods, D.C., 2025. Multiobjective optimization using expected quantile improvement for decision making in disease outbreaks. SIAM/ASA J. Uncertain. Quantif. 13 (1), 228–250.

Stevenson, M.D., Oakley, J., Chilcott, J.B., 2004. Gaussian process modelling in conjunction with individual patient simulation modelling: a case study describing the calculation of cost-effectiveness ratios for the treatment of established osteoporosis. Med. Decis. Mak. 24 (1), 89–100.

Stolk, W.A., ten Bosch, Q.A., De Vlas, S.J., Fischer, P.U., Weil, G.J., Goldman, A.S., 2013. Modelling the impact and costs of semiannual mass drug administration for accelerated elimination of lymphatic filariasis. PLoS Negl. Trop. Dis. 7 (1), e1984.

Stolk, W.A., Coffeng, L.E., Bolay, F.K., Eneanya, O.A., Fischer, P.U., Hollingsworth, T.D., Koudou, B.G., Meite, A., Michael, E., Prada, J.M., et al., 2022. Comparing antigenaemia-and microfilaraemia as criteria for stopping decisions in lymphatic filariasis elimination programmes in Africa. PLoS Negl. Trop. Dis. 16 (12), e0010953.

Stolk, W.A., Prada, J.M., Smith, M.E., Kontoroupis, P., De Vos, A.S., Touloupou, P., Irvine, M.A., Brown, P., Subramanian, S., Kloek, M., et al., 2018. Are alternative strategies required to accelerate the global elimination of lymphatic filariasis? Insights from mathematical models. Clin. Infect. Dis. 66 (Suppl. 4), S260–S266.

Stone, C.M., Kastner, R., Steinmann, P., Chitnis, N., Tanner, M., Tediosi, F., 2016. Modelling the health impact and cost-effectiveness of lymphatic filariasis eradication under varying levels of mass drug administration scale-up and geographic coverage. BMJ Glob. Health 1 (1), e000021.

Strong, M., Oakley, J.E., 2014. When is a model good enough? Deriving the expected value of model improvement via specifying internal model discrepancies. SIAM/ASA J. Uncertain. Quantif. 2 (1), 106–125.

Strong, M., Oakley, J.E., Brennan, A., Breeze, P., 2015. Estimating the expected value of sample information using the probabilistic sensitivity analysis sample: a fast, nonparametric regression-based method. Med. Decis. Mak. 35 (5), 570–583.

Swallow, B., Rand, D.A., Minas, G., 2024. Bayesian inference for stochastic oscillatory systems using the phase-corrected linear noise approximation. Bayesian Anal 1–30.

Ton, T.G., Mackenzie, C., Molyneux, D.H., 2015. The burden of mental health in lymphatic filariasis. Infect. Dis. Poverty 4, 34. https://doi.org/10.1186/s40249-015-0068-7

Turner, H.C., Archer, R.A., Downey, L.E., Isaranuwatchai, W., Chalkidou, K., Jit, M., Teerawattananon, Y., 2021. An introduction to the main types of economic evaluations used for informing priority setting and resource allocation in healthcare: key features, uses, and limitations. Front. Public Health 9, 722927.

Turner, H.C., Kura, K., Roth, B., Kuesel, A.C., Kinrade, S., Basáñez, M.-G., 2024. An updated economic assessment of moxidectin treatment strategies for onchocerciasis elimination. Clin. Infect. Dis. 78 (Suppl. 2), S138–S145.

Turner, H.C., Lauer, J.A., Tran, B.X., Teerawattananon, Y., Jit, M., 2019. Adjusting for inflation and currency changes within health economic studies. Value Health 22 (9), 1026–1032.

Vallejo-Torres, L., Claxton, K., Edney, L.C., Karnon, J., Lomas, J., Ochalek, J., Paulden, M., Stadhouders, N., Vanness, D.J., 2023. Challenges of calculating cost-effectiveness thresholds. Lancet Glob. Health 11 (10), e1508.

Villani, C., 2021. Topics in Optimal Transportation. Vol. 58. American Mathematical Soc.

Wahba, G., 1990. Spline Models for Observational Data. SIAM.

Wang, W., Slepcev, D., Basu, S., Ozolek, J.A., Rohde, G.K., 2013. A linear optimal transportation framework for quantifying and visualizing variations in sets of images. Int. J. Comput. Vis. 101 (2), 254–269.

Van der Werf, M.J., De Vlas, S.J., Brooker, S., Looman, C.W.N., Nagelkerke, N.J.D., Habbema, J.D.F., Engels, D., 2003. Quantification of clinical morbidity associated with schistosome infection in Sub-Saharan Africa. Acta Trop. 86 (2–3), 125–139.

Williams, C.K., Rasmussen, C.E., 1996. Gaussian processes for regression. Adv. Neural Inf Process. Syst. 8, 514–520.

Willis, G.A., Mayfield, H.J., Kearns, T., Naseri, T., Thomsen, R., Gass, K., Sheridan, S., Graves, P.M., Lau, C.L., 2020. A community survey of coverage and adverse events following country-wide triple-drug mass drug administration for lymphatic filariasis elimination, Samoa 2018. PLoS Negl. Trop. Dis. 14 (11), e0008854.

Wilson, E., 2015. A practical guide to value of information analysis. PharmacoEconomics 33 (2), 105–121.

World Health Organization, 2011. Lymphatic Filariasis: Monitoring and Epidemiological Assessment of Mass Drug Administration, A Manual for National Elimination Programmes. World Health Organization.

World Health Organization, 2023. Weekly epidemiological record (WER), 13 October 2023, vol. 98, no. 41, pp. 489–502. Weekly Epidemiological Record.

World Health Organization, Baltussen, R. M.P.M., Adam, T., Tan-Torres Edejer, T., Hutubessy, R. C.W., 2003. Making Choices in Health: WHO Guide to Cost-Effectiveness Analysis. World Health Organization.

World Health Organization Regional Office for the Western Pacific and Control of Lymphatic Filariasis in China Editorial Board, 2003. Control of Lymphatic Filariasis in China. World Health Organization Regional Office for the Western Pacific.

Xiu, D., Karniadakis, G.E., 2002. The Wiener–Askey polynomial chaos. SIAM J. Sci. Comput. 24 (2), 619–644.