Article:
Mandrik, O., Thomas, C., Bessey, A. et al. (2026) Ten recommendations for modelling cost effectiveness of screening: perspectives of an international stakeholder group. PharmacoEconomics. ISSN: 1170-7690

https://doi.org/10.1007/s40273-026-01602-7

10 Recommendations for modelling cost effectiveness of screening: perspectives of an international stakeholder group

Mandrik Olena [1]; Thomas Chloe[1]; Bessey Alice[1*]; Brennan Alan[1]; Carvalho L. Andre [2]; Castilla Rodríguez Ivan[3]; Doroshenko Olena[4]; Hill Harry[1]; Kunst Natalia[5,6]; Nagy Balazs[7,8]; Payne Katherine[9]; Pollard Dan[1]; Ramsey D. Scott[10]; Roitberg Felipe [11]; Shinkins Bethany[12]; Smith A. Robert [13]; Thom Howard[14], Whyte Sophie[1]

Type of submission: "Practical Application"

A running heading: 10 Recommendations for modelling screening

*All authors after the first two authors are in alphabetical order

[1] Sheffield Centre for Health and Related Research (SCHARR), School of Medicine and Population Health, The University of Sheffield, Sheffield, UK

[2] International Agency for Research on Cancer, Lyon, France

[3] University of La Laguna, Canary Islands, Spain

[4] The World Bank Group, Sarajevo, Bosnia and Herzegovina

[5] Centre for Health Economics, the University of York, UK

[6] Yale School of Medicine, New Haven, Connecticut, US

[7] Syreon Research Institute, Budapest, Hungary

[8] Center for Health Technology Assessment, Semmelweis University, Budapest, Hungary.

[9] Division of Population Health, Health Services Research and Primary Care, The University of Manchester, Manchester, UK

[10] Hutchinson Institute for Cancer Outcomes Research (HICOR), Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

[11] Hospital Sirio Libanes, São Paulo, Brazil

[12] Warwick Screening, University of Warwick, Coventry, UK

[13] American Cancer Society, New York, US

[14] University of Bristol, Bristol, UK

Corresponding author: Lena (Olena) Mandrik; Senior Research Fellow, Sheffield Centre for Health and Related Research (SCHARR), School of Medicine and Population Health, The University of Sheffield

Regent Court, 30 Regent Street, Sheffield, S1 4DA

Tel: +44 (0) 7 487 29 22 40; Email: o.mandrik@sheffield.ac.uk

ORCID: https://orcid.org/0000-0003-3755-3031

Word count total: 4,262

Conflict of interests: All authors have professional, personal, or intellectual conflicts of interest arising from non-commercial work related to the evaluation and/or implementation of screening programmes. B. Shinkins and K. Payne are members of the UK National Screening Committee. L. Mandrik is a member of the WHO SAGE IVD group. The views expressed in this manuscript are solely those of the authors and do not necessarily reflect the views of their affiliations. This includes authors affiliated with the UK National Screening Committee, the World Bank Group and with the International Agency for Research on Cancer/World Health Organization, who are solely responsible for the content of this article. The views expressed do not necessarily represent the decisions, policies, or views of the UK National Screening Committee, the World Bank Group or the International Agency for Research on Cancer/World Health Organization. No commercial conflicts of interest were declared.

Acknowledgement:

The authors acknowledge contributions from S. Lombardo and J. Marshall from the National Screening Committee UK and A. Zauber from the Memorial Sloan Kettering Cancer Center, US. The authors are grateful to Dr S. Davis and A. Rayner for external validation of the checklist.

This project has been supported by the Higher Education Innovation Fund (HEIF) #189350, the University of Sheffield. Olena Mandrik holds the individual fellowship from the National Institute for Health and Care Research (NIHR), NIHR306094. Katherine Payne is supported by the NIHR Manchester Biomedical Research Centre (BRC) (NIHR203308). Katherine Payne also holds a NIHR Senior Investigator award.

Abstract

Modelling the cost-effectiveness of screening interventions presents unique challenges. These relate to lack of knowledge about underlying health states and disease progression in the absence of screening, added costs arising from incidental findings, screening recall and follow-up diagnostics, imperfect uptake, potential harms to otherwise healthy people, and impacts on resource capacity and equity. No specific but generalisable advice currently exists to help guide health economic modellers working in this area. There is a need for tailored recommendations beyond the widely used health economic modelling frameworks.

We aimed to develop a set of recommendations for modelling the cost-effectiveness of screening programmes. In our iterative process we first drafted a conceptual document outlining key issues requiring recommendations. This framework was then expanded based on additional themes identified through a survey of screening modelling experts. Next, the draft recommendations were shared with a broader international expert group, which included modellers, health economists, and policy specialists. Finally, the core concepts were refined and agreed upon during a virtual stakeholder meeting.

A set of ten recommendations and a checklist are presented. The document provides guidance on critical methodological requirements for modelling screening interventions. These guidelines are intended to help health economic modellers and screening policy makers working to evaluate screening interventions across a wide range of diseases and jurisdictions with clarity, rigour and consistency.

Key Messages

- Cost-effectiveness modelling for screening interventions presents unique challenges, but no guidelines exist beyond standard health economic modelling frameworks.
- A set of ten recommendations for modelling screening interventions were developed, based on iterative expert input through surveys, comments and meetings.
- These recommendations and a checklist provide guidance for screening modellers and policy makers around important methodological requirements and should help improve consistency between screening models developed across different diseases.

Key Words

Screening; health economic modelling; cost-effectiveness modelling; guidelines; recommendations; stakeholder input; expert advice.

Introduction

Screening involves testing usually asymptomatic individuals, with the goal of detecting disease earlier, when treatment has a higher impact on survival and/or quality of life (Table 1). Screening may shift the diagnostic timeline by months or even years, thereby impacting existing diagnostic pathways (Fig 1). Screening interventions for chronic diseases differ substantially from diagnostic interventions used in symptomatic diagnosis in their implementation. A screening programme can take many forms—it may be organised or opportunistic; population-based, risk-stratified, or targeted. There is also considerable variation in who is screened (e.g. age and individual characteristics), how often screening occurs (e.g. annually, biennially, one-off), and which screening tests and decision criteria are applied. These decisions create a wide array of potential programme configurations.

Implementation decisions for screening programmes are informed not only by comparative screening trials, but also by decision-analytic models. These models estimate the impact of screening beyond the trial period and assess whether the health benefits justify the associated costs by comparing the programme's cost-effectiveness with alternative interventions or with current practice.

Due to complexities of screening programmes configurations, decision-analytic models used to evaluate screening interventions in terms of their impact on health outcomes and costs (hereafter screening models) are expected to differ in structure and scope from those used for treatment or diagnostic purposes. These screening models often need to simulate long-term outcomes and population-level effects while capturing high degrees of heterogeneity and uncertainty. In many cases, natural history disease (NHD) modelling is required, which further compounds model uncertainty due to the need to incorporate multiple risk factors, estimate unobservable health states, and make assumptions about disease progression in the absence of diagnosis, commonly with limited empirical data to inform these assumptions. Together, these factors increase the complexity of conceptualising, developing, validating, analysing, and interpreting screening models.

Systematic reviews summarising methods used to develop screening models have highlighted inconsistencies in their quality and a frequent failure to adequately address the inherent complexities of screening interventions (1-4). Multiple general frameworks for cost-effectiveness modelling have been previously developed and summarised by the ISPOR task force on reviews with cost-effectiveness outcomes (5). While these frameworks have been recommended for guiding cost effectiveness model development, none of the existing recommendations are specifically tailored to the unique requirements of modelling screening interventions. Some highly intricate modelling frameworks for screening interventions have been developed in the past, such as the CISNET models of cancer screening(6, 7). However, replicating such complex approaches is often impractical due to the significant resources, data, and financial investment required. For the same reason updating complex models may lag behind new evidence (underlying risk, disease progression, treatments, etc). Published reviews have examined cancer model structures, but these have largely been descriptive—summarising model designs, inputs, and outputs—without offering specific recommendations for modelling screening programmes(4, 7, 8).

Here we present a set of recommendations developed and coauthored by a multidisciplinary, multi-national expert group, that aim to provide guidance on critical requirements for screening models and enable consistency in model development.

Methods

Recommendations were developed through an iterative process (see online appendix). First, a conceptual document outlining the key points where screening modelling differs from modelling other types of interventions was drafted based on a scoping review of existing literature about modelling screening interventions. This was followed by the development of an online survey, which incorporated questions based on the main domains of existing checklists for cost-effectiveness studies (4). The combined framework was shared for feedback with a multidisciplinary, multi-national expert group, including modellers, implementation scientists, health economists, and policy specialists. The refined concepts were subsequently discussed during a virtual stakeholders' meeting including all stakeholders, where an effort was made to reach consensus on all recommendations. Any areas where consensus could not be achieved were documented. The final recommendations were unanimously supported by all but one expert (HT, who disagreed with three recommendations). This level of agreement meets our predefined threshold for consensus (<15% disagreement) (9).

The terminology used in this manuscript is reported in Table 1 and the externally validated methodological checklist for modellers - in the Table 2.

Table 1. Terminology and definitions referred to in this publication

Terminology	Definition
Screening implementation related definitions	
Screening for asymptomatic diseases	Testing individuals who do not report any symptoms, with the goal of detecting diseases earlier, when treatment of disease has higher impact on survival or quality of life.
Opportunistic screening **(10)**	A screening test that may be offered by a clinician to an individual during a health care encounter for other reasons, or that could be potentially requested by the patient.
Organised screening **(10)**	A screening programme where individuals are invited to screening proactively.
Population or population - wide screening **(10)**	A nationally or regionally delivered screening programme in which an eligible group to be screened is identified only by such factors as age and sex.
Steady state or fully rolled out programme	A screening programme that has completed its initial implementation phase and is operating at its intended capacity. In a model this implies that all individuals targeted by a specific intervention are invited for screening from the same age.
One-time screening	Process of conducting screening tests as a single event, usually at a specific age (e.g. Abdominal Aortic Aneurysm screening in the UK) or state (e.g. pregnancy or newborn screening).
Repeat screening	Process of conducting screening tests at regular intervals (screening rounds).
Risk-stratified screening **(10)**	A screening programme in which key elements—such as screening interval, type of test, age of initiation or stopping, or follow-up care—are tailored according to an individual's risk level. This may involve variation in one or more of these factors based on risk stratification.
Targeted screening **(10)**	A screening programme which aims to improve health outcomes among groups of people identified as being at elevated/above average risk of a specific disease with the risk factor(s) not related solely to age and sex.
Screening Outcomes	
Screening uptake **(11)**	Proportion of individuals among those invited to participate in a screening programme who adequately participate.
Overdiagnosis	Different definitions of overdiagnosis are used in the literature, including diagnosis of the targeted disease through screening that would not be diagnosed without screening, and (2) have gone on to cause harm in a person's lifetime without screening, or (3) ever result in disease-specific mortality. Overdiagnosis

	may occur due to the slow-growing or non-life-threatening nature of the disease, or because of high competing mortality risks (e.g., limited life expectancy), i.e., the patient would have died from another cause before they would have been diagnosed in the absence of screening. Consequently, the disease does not pose any practical threat to the individual's health but involves the use of healthcare resources (screening and treatment resources) and may compromise patients' safety and health outcomes, and quality of life.
Incidental findings	Set of diseases not specifically targeted by screening but discovered during screening.
Test sensitivity **(12)**	For individual tests, test sensitivity is the proportion of people with disease who will have a positive result.
Test specificity **(12)**	For individual tests, test sensitivity is the proportion of people without the disease who will have a negative result.
False - positive screening result	Because screening includes a sequence of diagnostic procedures, different definitions of false positives are used in the literature, including (1) A screening test result that recalls a person for a further diagnostic follow up when they do not actually have the suspected condition (but may not result in positive diagnosis); (2) A screening test that indicates the possible presence of a disease, prompting further diagnostic follow-up, which ultimately leads to a confirmation that an individual has the suspected disease, despite actually not having it.
False - negative screening result	Because screening includes a sequence of diagnostic procedures, different definitions of false positives are used in the literature, including (1) A screening test result that incorrectly identifies a person as disease-free when they actually have the disease; (2) A screening test result that incorrectly identifies a person as disease-free during any step of the screening or diagnostic pathway, despite the individual actually having the disease.
Spill-over effect	Unintended or indirect consequences of screening that extend beyond the initial test result. For example, impact of newborn screening outcomes on family quality of life or life of their siblings or identification of one's relative as a genetic disease carrier, which makes them change their health behaviour, decrease quality of life, or stimulates them to request additional health checks or health care.
Cascade effect	Chain of systematic medical actions triggered by the initial screening. For example, identification of one's relative as genetic disease carrier what makes them eligible for early or more intensive screening.
Other definitions	
Natural History of Disease **(13)**	The course of a disease from its onset (inception) to its resolution, in the absence of treatment or prevention.
Sojourn time	The period during which the disease is present, and so can be detected by screening, but has not yet become clinically symptomatic.

153

154 **Table 2. Checklist for Modelling Screening Interventions**

Recommendation	Checklist main questions	Clarification points
Recommendation 1. Engage Stakeholders Early and Continuously	Have relevant stakeholders been identified and engaged from the conceptualisation stage?	☐ Does the stakeholder group include wide group of contributors, including policymakers and patients/public representatives? ☐ Has engagement been planned to continue throughout the modelling process?
	Is the model design aligned with stakeholder goals?	☐ Have model aims and objectives been validated with stakeholders? ☐ Do stakeholders agree with the structure, assumptions, and data sources used in the model? ☐ Are implementation challenges highlighted by stakeholders (e.g. feasibility, capacity) captured?
Recommendation 2. Evaluate a Wide Range of Interventions and Comparators	Are a wide range of interventions and comparators evaluated?	☐ Have all relevant screening programme options (e.g. ages, intervals, thresholds, etc.) been considered? ☐ Are all of the comparators that the modelling analysis included relevant and feasible to model? ☐ Have social and ethical acceptability issues been considered?
Recommendation 3. Choose the Modelled Population Carefully	Is the modelled population relevant, representative, and adequate to evaluate the impact of screening?	☐ Will the chosen population enable decision questions to be answered (e.g. equity analyses, sub-groups)? ☐ Is the sample size sufficient (for individual-level simulation)? ☐ Have spill-over/cascade effects been considered?
Recommendation 4. Incorporate Key Screening Characteristics with Precise Definitions	Does the model define key screening characteristics and incorporate them if necessary?	☐ Uptake in screening and follow up ☐ False positives ☐ False negatives ☐ Incidental findings ☐ Overdiagnosis
	If some screening characteristics are omitted, is this impact justified and discussed?	☐ Impact on resource use ☐ Impact on costs ☐ Impact on health benefits
	If broader impacts beyond clinical outcomes considered, are they balanced?	☐ Both broader benefits and harms are considered ☐ Unclear balance between broader benefits and harms
Recommendation 5. Use Conservative Mortality Assumptions	Are mortality assumptions reflecting the best available data, relevant for modelled populations?	☐ Disease-specific mortality ☐ Other cause mortality
Recommendation 6. Identify and Address High Uncertainty	Has uncertainty in screening models been acknowledged?	☐ Natural history disease ☐ Long-term outcomes ☐ Surrogate endpoints ☐ Screening characteristics
	Has uncertainty in screening models been explored and addressed properly?	☐ Methodological ☐ Stochastic ☐ Parameter ☐ Structural
Recommendation 7. Consider Detailed Reporting of Methods and Outcomes	Have methods and outcomes been reported in sufficient detail for screening models?	☐ Standard health economic outcomes (QALYs, NMB, ICER) ☐ Clinical outcomes (e.g. incidence, mortality) ☐ Screening characteristics (see recommendation 4) ☐ Resource utilisation ☐ Subgroup outcomes
Recommendation 8. Consider	Does the model adequately account for heterogeneity	☐ Within-group and between-group differences (e.g., risk, mortality, progression, uptake) ☐ Subgroup-specific outcomes

Heterogeneity and Equity Impacts	and potential equity impacts of screening?	☐ Equity impact or acknowledgement of implications for access and fairness
Recommendation 9. Consider Healthcare System Capacity	Does the model consider healthcare system capacity and opportunity costs of implementing screening?	☐ Has unconstrained modelling been used to first identify efficient strategies? ☐ Are potential constraints relevant to consider? ☐ Are opportunity costs (e.g., resources diverted from symptomatic patients) acknowledged? ☐ Are resource limitations explicitly considered (e.g. for low-resource settings)?
Recommendation 10. Assess Transferability and Need for Updates	If model adaptation is considered, has model transferability been evaluated?	☐ Are setting-specific differences in healthcare pathways and population risk accounted for? ☐ Have transferable and non-transferable components been defined?
	When model results become outdated?	☐ Has the potential impact of downstream care changes on screening outcomes been considered?

Legend: Each question can be answered with "Yes," "No," or "Not applicable / not relevant" (and with "Can't tell" when the checklist is used for appraisal purposes).

Abbreviations: ICER – Incremental Cost-Effectiveness Ratio; NMB – Net Monetary Benefit; QALYs – Quality-Adjusted Life Years.

Recommendation 1. Engage Stakeholders Early and Continuously to Shape Model Aims and Design

Screening interventions involve multiple interacting components—such as defining the target population, selecting the screening modality, determining the screening interval, and managing follow-up. These components give rise to a wide range of possible programme design scenarios that influence funders, implementors, and beneficiaries. Consequently, the aims and design of a screening model - while remaining within the general aims of cost-effectiveness analysis such as maximising health benefits within fixed budgets - should be guided by the needs and priorities of relevant stakeholders.

To ensure that the model remains flexible, robust, and feasible to implement, stakeholder engagement should begin early in the conceptualisation phase and continue throughout the modelling process. Although stakeholder involvement in conceptual model development is now standard practice (14), ongoing engagement provides additional value by ensuring that emerging evidence relevant to model structure, analysis, or interpretation is not overlooked, and by supporting the interpretation of model validation results and modelling outputs in relation to existing empirical evidence. This is particularly important for screening modelling projects, which are typically longer in duration and more structurally complex and data-intensive than treatment models.

The stakeholder group that contributes to model development (Box 1) should be broad and inclusive, encompassing not only clinicians, statisticians, epidemiologists, and policymakers, but also peer health economists to validate the modelling assumptions and patients and public representatives (i.e. target screening population). This inclusive approach ensures that lived experiences are embedded in the model and aligns the modelling research question with the PICO (Population, Intervention, Comparator, Outcome) framework (15).

Box 1: Engage stakeholders to contribute to:

- Key interventions and priority questions
- Clinical pathway and the role of screening within it
- Constraints and implementation challenges
- Assumptions related to the NHD, data sources, parameter plausibility (e.g. test sensitivity, screening uptake)
- Data transferability when local data are lacking
- Model validation
- Interpretation of the modelling outcomes.

Stakeholders also influence model design indirectly, as their goals shape the model's purpose and structure (16, 17). For instance:

- If the goal is to assess cost-effectiveness for an established national screening programme, the model if possible should simulate the real-world delivery of a fully rolled-out intervention (18). This includes individuals initiating screening at a defined age and continuing until an upper age limit.
- If the goal is to support implementation planning or inform trial design (e.g. to determine optimal sample size, screening intervals, or age ranges), it may be more appropriate to simulate the initiation phase of a screening programme (19). Such models involve offering screening to the entire eligible population at once, resulting in variation in screening histories, costs, and benefits across different age groups.

The complexity of the model should also reflect the nature of the intervention being evaluated and be no more complex than necessary (Table 3).

Table 3. Considerations in the choice of screening model design

Type of Screening Intervention	Model Complexity	Typical Model Structure	Key Modelling Considerations	Examples
One-time screening	Low to moderate	Decision tree (possibly combined with a Markov model)	• Resembles diagnostic models • Focuses on immediate outcomes of a single screening event • Limited data requirement • Efficient computation • Computationally easy probabilistic sensitivity analysis	Newborn screening
Repeated screening	Moderate to high	Markov model, microsimulation, discrete event simulation, or hybrid model	• Accounts for missed cases in earlier rounds • Incorporates de novo disease development • Reflects natural history of disease between screening rounds • Require more data • Includes individual-level correlations such as repeated participation behaviour • Computationally challenging probabilistic sensitivity analysis • May support multiple decision questions	Cancer screening programmes

Recommendation 2. Evaluate a Wide Range of Intervention and Comparators where Feasible

Given the multiple decisions involved in designing a screening programme —including age of initiation and cessation, screening interval, test modality, risk thresholds, and population subgroups —screening models often involve comparison of multiple scenarios. Beyond the clinical and economic impacts, modelled strategies may also need to consider social and ethical aspects to improve public acceptability (20, 21). For instance, if a risk-stratified screening approach leads to exclusion of lower-risk individuals from routine currently available screening, modelling alternative approaches—such as allowing opportunistic screening upon request—may be useful for exploring more publicly acceptable options.

While comparing new interventions with "standard care" is common, other relevant comparators could be included where possible (i.e., where data exist to inform the evaluation, and research resource as well as implementation resources permit such evaluation). These may include:

- A "no screening" scenario, even in contexts where a screening programme is already in place, or
- Comparators involving alternative technologies or strategies likely to appear in the near future (e.g. those that were already approved elsewhere, approved but have not been implemented yet).

Previous methodological work identified a number of questions to be considered when selecting comparators for cost-effectiveness analysis (22). However, feasibility also must be considered when including such comparators (see Fig 2). For example, parametrisation of a "no screening" scenario may be difficult if screening is long established (the example of breast cancer screening (7)) or if opportunistic (often poorly recorded) screening is widespread, as it is the case with use of prostate-specific antigen test in cancer diagnostics and screening (23). In such cases, limiting comparisons to "standard care" may be more realistic (3). Conversely, if the future screening programme is compared to a recently implemented programme that is not yet at steady-state, reliable representation of “standard care” may also be challenging.

Recommendation 3. Choose the Modelled Population Carefully

Although decision-makers often prioritise specific sub-groups (e.g., high-risk populations), developing a model for the entire at-risk population (e.g., the national population), alongside targeted analyses for specific policy questions, offers several advantages—provided sufficient data are available. This approach allows reconstruction of populations with varying characteristics (e.g., age, risk factors, care exposure), enabling flexible and efficient exploration of additional research questions (e.g., broader populations, new screening methods), easier model validation across different screened populations (e.g., trial cohorts), and assessment of screening’s impact on health equity (e.g., socioeconomic disparities).

While population size may be important in any model, it is particularly critical in screening models that rely on individual patient- level simulations. This is because most simulated individuals will incur costs—such as those from screening invitations, procedures, and follow-up of screen-positive cases—while only a small proportion will benefit from screening. For example, for abdominal aortic aneurysm, less than one percent of men screened in England were found to have a positive diagnosis(24). As a result, when the incremental benefits of screening versus a comparator are small

(e.g., comparing different starting ages for screening), using an insufficiently large simulated population may reflect cost differences but fail to detect meaningful differences in outcomes due to random variation alone (25).

The population assigned to screening in screening models should accurately reflect the risk of undiagnosed disease and/or its progression, which for many diseases varies across different population subgroups.

Age is one of the most important demographic factors to include in screening models. Age-dependent risk (i.e., varying prevalence of undiagnosed disease at different screening ages) is key in many conditions, including cardiovascular disease, cancer, and in antenatal screening where risk depends on maternal age (e.g., Down syndrome) (26). Using average instead of age-specific risk can overestimate screening benefits in younger people and underestimate them in older groups.

Other demographic (e.g. sex), environmental (e.g. occupation) and lifestyle (e.g. smoking, diet) factors also matter when modelling populations with varied risks. It is important to consider whether correlated risks or behaviours are relevant (for example alcohol consumption in people who smoke), which may distort estimates of undiagnosed disease prevalence and screening-related behaviour in the simulated population. Including such variables helps ensure accurate modelling of both disease risk and intervention impact.

Unlike many other interventions, screening can have spill-over or cascade effects on other population groups. Therefore, screening models should clearly define the populations being modelled, account for potential spill-over, and consider how excluding these effects might influence cost-effectiveness estimates (27). This effect should for instance be considered in screening for genetically inherited conditions that may prompt testing among relatives if an index case tests positive (28). Another example is cost-effectiveness analysis of HIV screening in pregnant women, where the impact of screening on both the child and the mother should be considered (27).

Recommendation 4. Key Screening Characteristics Should be Incorporated in Models Following Precise Definitions that Suit the Screening Context.

The benefits and harms of screening are likely to be unevenly distributed between individuals and over time. In some inherited genetic conditions, early detection may not change clinical management until symptoms appear (29, 30). In other cases, such as cancer, early detection can trigger aggressive treatment that may reduce quality of life or even life expectancy.

Given these complexities, modern screening models must incorporate a comprehensive set of screening-related characteristics, whose definition may vary according to screening context—some reflecting benefits, others may be arguably considered as harms or unavoidable consequences of screening implementation. These include screening uptake, false positives and false negatives, incidental findings, and overdiagnosis.

To capture screening impacts accurately, as part of conceptual modelling it is necessary to provide a precise definition for how each of these characteristics are operationalised in the model, particularly where more than one valid definition exists (see Table 1).

<u>Screening Uptake</u>

Screening targets asymptomatic individuals, so screening uptake is usually lower than compliance to treatment and may vary by setting, socioeconomic group, or disease risk (19, 31, 32). Uptake is

especially important when modelling resource use or for assessing cost-effectiveness across risk groups. Uptake of diagnostic follow-up should also be considered. If uptake data are not available, initial modelling of perfect uptake scenarios may be suitable and are useful for estimating maximum net benefit, especially if invitation costs are low. If screening appears cost-effective under these assumptions, more realistic scenarios should follow, incorporating observed uptake rates or estimates from similar at-risk populations and interventions (19, 33). This is important since lower uptake of cost-effective interventions may result in less cost-effective screening as was demonstrated in studies of colorectal and breast cancer screening (34, 35)

False Positives and False Negatives

Screening results in recalls including false and true positive cases. False positive cases are often more prevalent than true positives in screening and have implications for cost and patient burden due to follow-up procedures in patients who are determined to not have disease. They should be included in screening models where possible to avoid underestimation of both resource use and harms (4, 33).

In contrast, it may be appropriate to omit false negatives from models where the comparator is 'no screening', but only if screening does not alter time to diagnosis for such patients, disease progression or costs of later diagnosis. For example, omitting false negatives from the first screening round in model analyses of repeat screening would not be appropriate.

Incidental Findings

Incidental findings refer to unrelated clinical conditions detected incidentally during screening. These findings can considerably impact long-term outcomes, costs, and patient experience. Failure to account for them may lead to inaccurate assessments of screening value (36). While not all incidental findings need to be modelled, those with significant health and/or cost implications should be prioritised. Their inclusion should depend on both screen-detectable prevalence and impact per case detected: findings that add benefits may support screening that is otherwise marginally cost-effective, while those adding costs or harms may tip the balance against screening. Examples of incidental findings included in screening models, include detection of multiple conditions during CT scan for abdominal aortic aneurysm, idiopathic T-cell lymphopenia in severe combined immunodeficiency newborn screening, or kidney cancer found during urine dipstick tests for bladder cancer (19, 37, 38).

Overdiagnosis

Screening is applied to generally healthy populations, so tolerance for harm is lower than in diagnostic or therapeutic interventions. Overdiagnosis refers to detection and treatment of disease that would never have become clinically relevant (1, 39). It may cause psychological harm and lead to unnecessary treatment with associated costs and harms, without improvements in survival or quality of life. Unlike incidental findings, overdiagnosis represent diseases the screening aims to detect. Overdiagnosis usually cannot be distinguished on an individual level, but its aggregate impact must be considered. Models should either estimate overdiagnosis directly or run scenario analyses with varying overdiagnosis rates if this is used as an input. Since overdiagnoses can never be measured and only estimated, accuracy of such estimates may vary and have little consensus. For example, the rate of overdiagnosis in CISNET breast cancer models varies from 3% to 66% (7).

Broader Impacts of Screening

From a policymaker perspective, screening is only justified when health benefits outweigh harms and costs to the healthcare system are reasonable. However, some screening programmes provide broader benefits—such as increased certainty, informed reproductive choices (e.g., in genetic or newborn screening), healthier behaviours, or planning for care or early retirement (40). These wider benefits, while important, must be weighed alongside broader harms, including psychological distress or fatalistic responses to risk information, as well as lack of suitable therapies for diagnosed cases. If broader, societal or patient-centred perspectives are adopted, screening-related harms beyond the clinical domain must also be considered.

Recommendation 5. Use Conservative Mortality Assumptions from Relevant Populations, in the Presence of Uncaptured Uncertainty

Mortality in screening models includes two main components: disease-specific and other-cause mortality. Most models assess screening impact by reducing disease-specific mortality through earlier diagnosis. When this is modelled as a shift to earlier disease stages—common in cancer or chronic kidney disease models—it is known as a stage-shift approach (2). However, stage shift may not fully capture screening benefits. It may be appropriate to assign additional survival advantages to screen-detected cases, reflecting that such cases may be less aggressive or progress more slowly (2, 41, 42).

In individual-level stage-shift models, even with a survival benefit from screening, some individuals may die during the lead-time period—before their symptoms would have triggered diagnosis in a no-screening scenario (19). This should not be confused with lead-time bias, which can overstate benefits in survival analyses based on earlier diagnosis and limited follow-up, and rather reflects a probability for some individuals to suffer from early death related to screening (e.g. side effects of radical or systemic therapy) even whilst the population as a whole gains survival benefits. Some models incorporate this nuance, while others assume early treatment always benefits survival. For example, in a bladder cancer screening model, two scenarios were evaluated: one assuming that approximately 0.8% of all cancer deaths in the model are related to aggressive cancer treatment and another one assuming that early cancer detection will not result in any early deaths (19).

These varying assumptions about disease-specific mortality illustrate how different modelling choices can lead to diverging results. When evidence is limited or uncertain, assumptions about mortality should be incorporated directly into structural and probabilistic analyses.

Unlike for many other model parameters uncertainty around disease-specific mortality is not bidirectional over time. This is because long-term trends in treatment improvements often lead to declining disease-specific mortality, which tends to reduce the cost-effectiveness of screening over time, creating an asymmetric risk if mortality benefits are overstated at the time of evaluation. Because of this, in the event that this is not feasible to reflect uncertainty in sensitivity analyses, conservative assumptions — those least favourable to cost-effectiveness — should be used.

Regarding other-cause mortality, models that ignore competing risks—such as life-limiting comorbidity—underestimate screening harms (e.g., overdiagnosis) and misrepresent cost-effectiveness, particularly in older populations. In targeted screening (e.g., for people who smoke), other-cause mortality should reflect higher risk profiles due to co-morbidities.

Recommendation 6. Identify, Measure, and Address High Uncertainty in Screening Models

Screening models typically involve greater uncertainty than treatment models, especially regarding underlying health states and long-term outcomes. This is particularly true for models relying on NHD, where key parameters—such as the prevalence of asymptomatic disease, pre-disease states, and progression without early detection—are under-researched. The challenge is amplified in newer screening programmes, for example in some cancers, where comprehensive NHD data are often lacking (33, 43), or in newborn screening, where data limitations stem from low disease prevalence (44-46).

Uncertainty also arises from the use of lifetime horizons—common in screening models—which necessitate long-term extrapolations or reliance on surrogate outcomes. While long-term projections often rely on limited, outdated or soon-to-be-outdated data due to evolving treatments, surrogate outcomes are useful but may not always reliably predict long-term benefits (47-49). While this uncertainty is somewhat ameliorated by discounting, it increases with life expectancy, making it especially pronounced in newborn and pregnancy screening.

Even when data are available from trials, observational studies, or autopsy research, they may not fully reflect outcomes in the target population due to biases. Autopsy studies may underrepresent younger age groups (selection bias), trials may use flawed methods to estimate undiagnosed disease (measurement bias), and differences between source and target populations can limit applicability.

Screening accuracy and uptake are also major sources of uncertainty. Test sensitivity of some screening tests is assessed in unrepresentative symptomatic populations and may rely on comparisons with imperfect standards or biased study designs. Accuracy can also vary across healthcare settings and populations, particularly between countries of differing economic development (50) or in groups with varying disease risk and prevalence (51). Data on screening uptake for specific populations and diseases are frequently unavailable, leading to reliance on estimates from other conditions or trial settings, which are often overly optimistic—especially for underserved populations.

As shown, screening models face multiple layers of uncertainty. Modellers should ensure all sources of model uncertainty are identified and addressed where possible (Table 4).

Table 4. Approaches to address uncertainties in screening models

Approach to dealing with uncertainty	Types of uncertainty, for definitions see (17, 52, 53)
Address stakeholders for the reference case to follow in the base case and scenario analyses	Methodological uncertainty
Model sufficient population size in individual level models	Stochastic uncertainty
Conduct sufficient number of probabilistic sensitivity analyses for all key inputs	Parameter uncertainty*
Conduct scenario analyses for highly uncertain parameters	Parameter uncertainty*
Consider multi-parameter evidence synthesis	Parameter uncertainty*
Consider external model validation against available data, prioritising external validation using independent datasets when trial or observational study outputs are available	Parameter uncertainty
Report outcomes with both long-term and short-term horizon	Structural uncertainty**
Explore impact of structural assumptions on cost-effectiveness estimates in individually run scenarios, weighted scenarios based on their likelihood, and adding scenarios with different structural assumptions into the probabilistic sensitivity analysis.	Structural uncertainty**
Incorporate expertise from clinical and implementation experts to inform model scenarios.	Structural and parameter uncertainty
Integrate clinical data collection with modelling in an iterative data-model-data approach	Structural and parameter uncertainty
Use Bayesian methods for model parameterization through calibration and for multiparameter evidence synthesis	NHD – related uncertainty (structural and parameter)
Cross-validation of the model in the absence of natural history data	NHD – related uncertainty (structural and parameter)

*Legend: * Parameters with high uncertainty in screening models include uptake, test sensitivity, long-term clinical outcomes, and NHD parameters; ** Structural uncertainties in screening models include long-term extrapolations, use of surrogate outcomes, assumptions about mortality during the lead-time period, and similar model design choices.*

Recommendation 7. Consider Detailed Reporting of Methods and Outcomes for Screening Models

In screening modelling, outcomes beyond costs, life-years gained, quality-adjusted life years, and cost effectiveness (net monetary benefit and incremental cost effectiveness ratio) are important for stakeholders. Additional outcomes that could be considered for model validation or decision-making include false-positive and overdiagnosis rates, disease incidence, mortality, and intermediate endpoints such as stage distribution at diagnosis; disease characteristics affecting curative treatment potential and long-term quality of life; and metrics like time to treatment initiation or disease-free survival. These outcomes are useful for the whole population screened but also subgroups, for example people of different ages, ethnicities, or risks. In addition, health system outcomes such as resource utilisation are important, particularly when follow-up for screen-detected cases places demands on existing diagnostic services.

Screening models are complex and so require comprehensive reporting. While general health economics checklists, such as Consolidated Health Economic Evaluation Reporting Standards - CHEERS (54), provide guidance, they often lack the detail needed for screening modelling studies. In addition to standard health economic components (such as PICO, time horizon, discounting, and

valuation of costs and outcomes) screening models should also report screening uptake, the approach to modelling non-adherent populations, assumptions about sensitivity and specificity (and their mapping to symptomatic data), potential incidental findings, and other key screening characteristics discussed above (Recommendation 4). Given space constraints in manuscripts, a detailed online technical report should accompany screening models, e.g. in supplementary material.

Recommendation 8. Consider Heterogeneity and Equity Impacts of Screening

Patient heterogeneity plays a more prominent role in screening models than in diagnostic or treatment models, as variations in population characteristics affect both disease risk and health behaviours—such as screening uptake (55, 56). In screening contexts, differences in disease onset and progression influence who is detected and when. For example, in cancer screening, some individuals may have slow-progressing disease that remains undiagnosed for years, while others experience rapid progression. Relying on population averages and ignoring individual variability can result in inaccurate estimates of screening effectiveness in non-linear models.

When screening targets specific subgroups, heterogeneity matters at two levels: within-group variation and differences between subgroups. For example, prostate cancer screening may have different implications for Black men, who face a higher risk of the disease (57).

These considerations are closely linked to health equity. Some population groups may face higher disease risk, lower screening uptake, or both. In some cases, screening programmes risk exacerbating health inequity if attention is not devoted to ensuring equal access. Conversely, targeted implementation in underserved populations may help reduce disparities. Even if formal equity analysis is not included in the model, the potential for the screening programme to impact equity benefits to be discussed.

Recommendation 9. Consider the capacity of the healthcare service for implementing screening

Screening for chronic diseases can significantly increase demand on healthcare services, particularly for diagnostic confirmation and treatment following early detection. In some cases, this added pressure may be temporary and offset over time by a reduced burden of severe disease. However, when it is difficult to distinguish between aggressive and non-aggressive forms of disease, screening may lead to sustained increases in healthcare use—even if it improves survival outcomes.

Initial cost-effectiveness analyses should generally adopt an unconstrained modelling approach to identify the most efficient strategies without assuming limits on system capacity. This helps assess whether expanding capacity is a worthwhile investment. However, if screening interventions risk diverting resources from other critical services—such as diagnosis of symptomatic cases—additional analysis is needed. This includes considering opportunity costs and how they may be distributed across population groups.

Resource-constrained models can serve as valuable supplements in such cases. For instance, the evaluation of bowel cancer screening explicitly incorporated limitations in diagnostic colonoscopy capacity (58). These types of models are relevant across all health systems but are particularly important in low-resource settings, where scaling up diagnostic and treatment services poses major challenges (59).

Recommendation 10. Assess Transferability and the Need for Model Updates

Screening models are generally less transferable across jurisdictions than treatment or diagnostic models, due to differences in healthcare pathways, levels of trust in healthcare systems, population risk profiles, and disease heterogeneity. Transferability is especially limited when models are adapted between settings with significant healthcare system differences—such as from high- to low-income countries. When attempting such adaptations, differences in diagnostic pathways, disease management, and screening test accuracy must be considered.

That said, if baseline disease risk is comparable, the NHD component may be transferable—provided adjustments are made for population-specific risk factors. When applying a model from another setting to assess cost-effectiveness, it is essential to identify which components require adaptation. While NHD might remain constant, risk distributions, diagnostic practices, and treatment pathways are more likely to differ.

Because screening occurs early in the care pathway and is influenced by downstream changes, screening models require more frequent updates than treatment models. Updates should be considered when significant changes occur in risk factor distribution, diagnostics, treatment strategies or costs, survival outcomes, or when new natural history data become available. Update of treatment pathways is especially important in disease areas where there is quick and radical change in care pathways and disease progression trajectories. Even if the screening method remains the same such treatment changes can strongly influence the conclusion of the analysis.

Limitations

These recommendations were developed through a consensus process involving an international, multidisciplinary group of experts. While we aimed to make the guidance practical and accessible, fully benefiting from it may require some background knowledge in health economics, modelling, and screening. Certain methodological aspects of screening models were not addressed where the group felt that further empirical evidence is needed to support robust recommendations. These include the use of discounting in screening modelling studies and considerations regarding the decision threshold for cost per QALY.

Conclusion

This set of ten recommendations and a checklist provide guidance on critical methodological requirements for modelling screening interventions. The recommendations are intended to help health economic modellers and screening policy makers working to evaluate screening interventions across a wide range of diseases and jurisdictions with clarity, rigour and consistency.

Fig 1. Conceptualisation of screening models

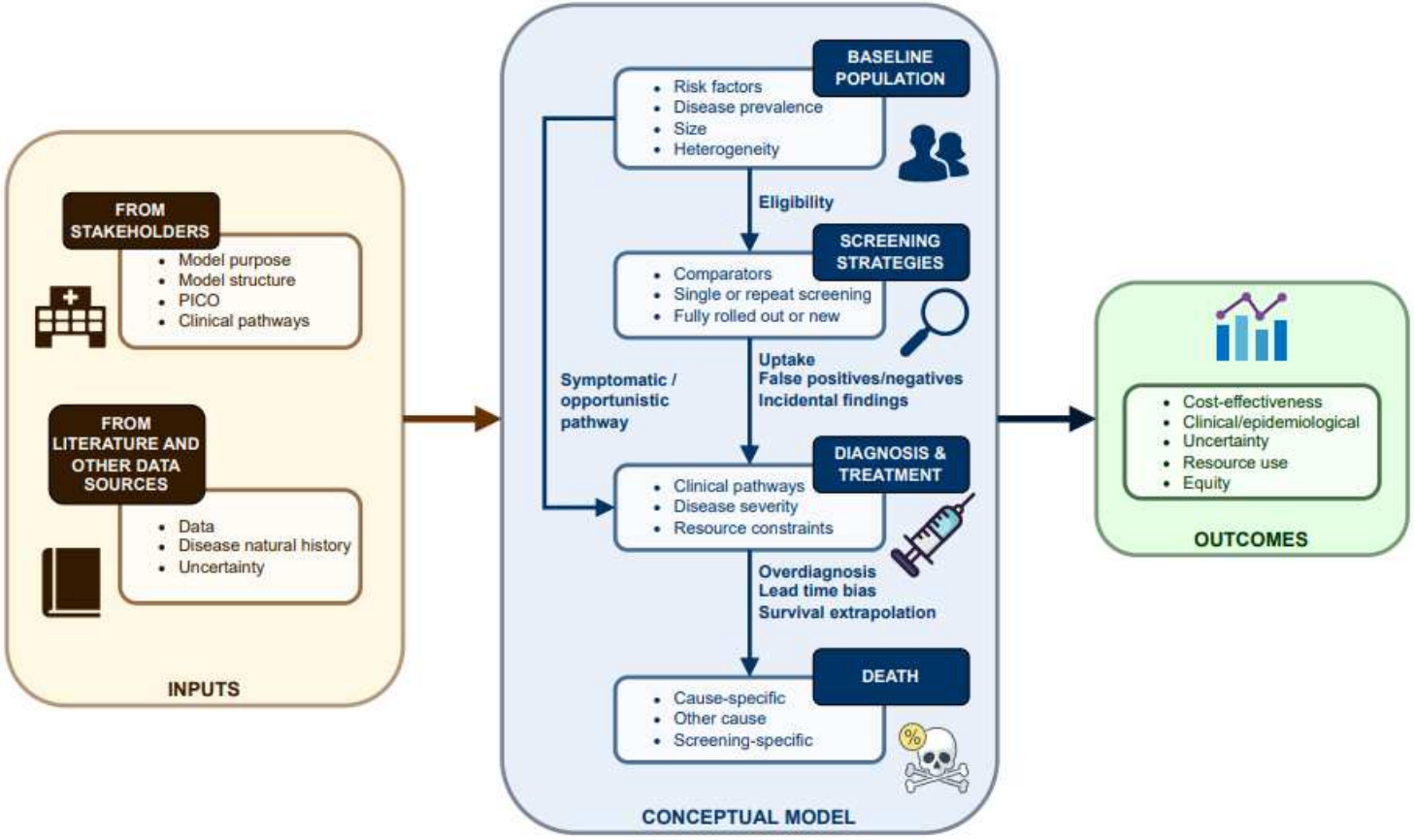
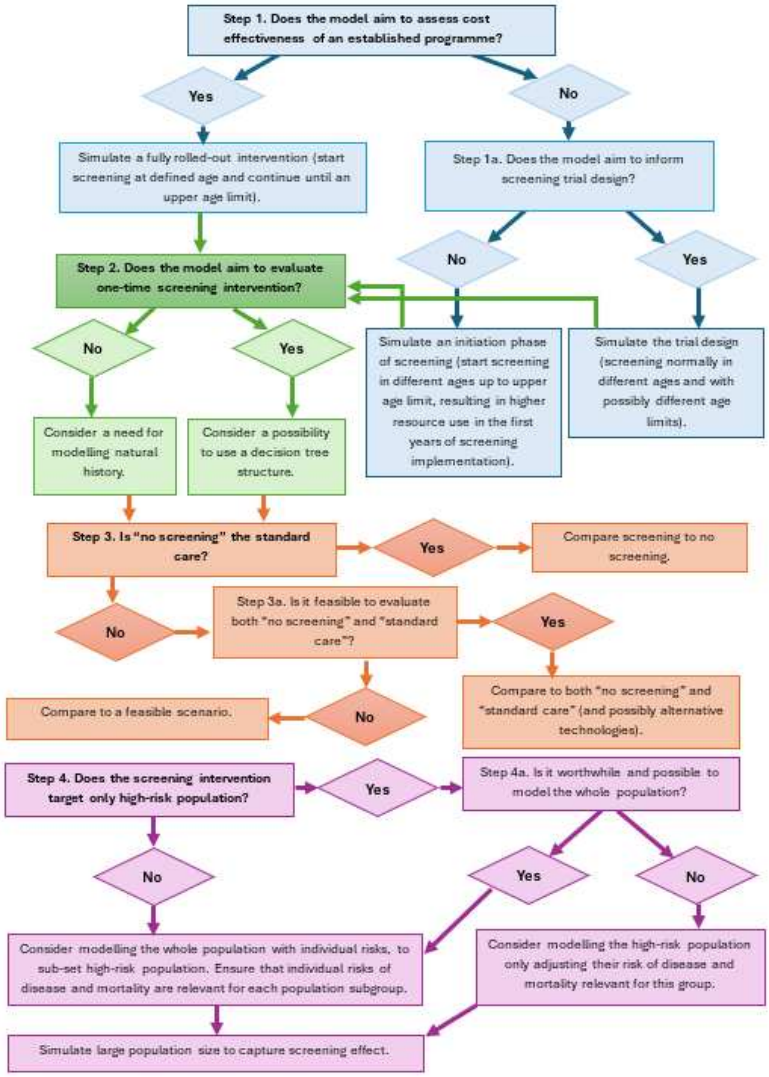


Fig 2. Decision pathway in developing design for the screening models



References

Commented [OM1]: : In the References section, please provide 3 names before et al. For ref. 39, please provide the name of the publisher and the city location. For ref. 48, please provide the full citation details.

1. Carter JL, Coletti RJ, Harris RP. Quantifying and monitoring overdiagnosis in cancer screening: a systematic review of methods. Bmj. 2015;350:g7773.
2. Keeney E, Thom H, Turner E, Martin RM, Morley J, Sanghera S. Systematic Review of Cost-Effectiveness Models in Prostate Cancer: Exploring New Developments in Testing and Diagnosis. Value Health. 2022;25(1):133–46.
3. Hiligsmann M, Wyers CE, Mayer S, Evers SM, Ruwaard D. A systematic review of economic evaluations of screening programmes for cardiometabolic diseases. Eur J Public Health. 2017;27(4):621–31.
4. Mandrik O, Hahn AI, Catto JWF, Zauber AG, Cumberbatch M, Chilcott J. Critical Appraisal of Decision Models Used for the Economic Evaluation of Bladder Cancer Screening and Diagnosis: A Systematic Review. Pharmacoeconomics. 2023;41(6):633–50.
5. Mandrik OL, Severens JLH, Bardach A, Ghabri S, Hamel C, Mathes T, et al. Critical Appraisal of Systematic Reviews With Costs and Cost-Effectiveness Outcomes: An ISPOR Good Practices Task Force Report. Value Health. 2021;24(4):463–72.
6. Pineda-Antunez C, Seguin C, van Duuren LA, Knudsen AB, Davidi B, Nascimento de Lima P, et al. Emulator-Based Bayesian Calibration of the CISNET Colorectal Cancer Models. Med Decis Making. 2024;44(5):543–53.
7. Trentham-Dietz A, Alagoz O, Chapman C, Huang X, Jayasekera J, van Ravesteyn NT, et al. Reflecting on 20 years of breast cancer modeling in CISNET: Recommendations for future cancer systems modeling efforts. PLoS Comput Biol. 2021;17(6):e1009020.
8. Knudsen AB, McMahon PM, Gazelle GS. Use of modeling to evaluate the cost-effectiveness of cancer screening programs. J Clin Oncol. 2007;25(2):203–8.
9. Keeney E, Thom H, Turner E, Martin RM, Sanghera S. Using a Modified Delphi Approach to Gain Consensus on Relevant Comparators in a Cost-Effectiveness Model: Application to Prostate Cancer Screening. Pharmacoeconomics. 2021;39(5):589–600.
10. UK National Screening Committee: screening in healthcare: UK National Screening Committee; 2024 [updated 10 May 2024. Available from: https://www.gov.uk/guidance/principles-of-population-screening/principles-of-screening#:~:text=The%20UK%20NSC%20does%20not,no%20symptoms%20or%20particular%20risk.
11. Bowel cancer screening programme standards: valid for data collected from 1 April 2018: NHS England; [updated 4 June 2025. Available from: https://www.gov.uk/government/publications/bowel-cancer-screening-programme-standards/bowel-cancer-screening-programme-standards-valid-for-data-collected-from-1-april-2018?utm_source=chatgpt.com.
12. Akobeng AK. Understanding diagnostic tests 1: sensitivity, specificity and predictive values. Acta Paediatrica. 2007;96(3):338–41.
13. Last JM, International Epidemiological A. A dictionary of epidemiology / edited by John M. Last. 4th ed. ed. Oxford: Oxford : Oxford University Press, 2001.; 2001.
14. Squires H, Chilcott J, Akehurst R, Burr J, Kelly MP. A Framework for Developing the Structure of Public Health Economic Models. Value Health. 2016;19(5):588–601.
15. Luijendijk HJ. How to create PICO questions about diagnostic tests. BMJ Evid Based Med. 2021;26(4):155–7.
16. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. Health Econ. 2006;15(12):1295–310.
17. Fasseeh A MA, Nagy B., Nemeth B, Szilberhorn L. The role of modelling in economic evaluations in health care: Eötvös University Press; 2019. 72 p.

18. Thomas C, Mandrik O, Whyte S. Modelling cost-effective strategies for minimising socioeconomic inequalities in colorectal cancer screening outcomes in England. Prev Med. 2022;162:107131.
19. Mandrik O, Thomas C, Akpan E, Catto JWF, Chilcott J. Home Urine Dipstick Screening for Bladder and Kidney Cancer in High-Risk Populations in England: A Microsimulation Study of Long-Term Impact and Cost-Effectiveness. Pharmacoeconomics. 2025;43(4):441–52.
20. Dennison RA, Usher-Smith JA, John SD. The ethics of risk-stratified cancer screening. Eur J Cancer. 2023;187:1–6.
21. Dunlop K, Rankin NM, Smit AK, Salgado Z, Newson AJ, Keogh L, et al. Acceptability of risk-stratified population screening across cancer types: Qualitative interviews with the Australian public. Health Expect. 2021;24(4):1326–36.
22. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-Effectiveness in Health and Medicine. New York, UNITED STATES: Oxford University Press, Incorporated; 1996.
23. Karlsson A, Jauhiainen A, Gulati R, Eklund M, Grönberg H, Etzioni R, et al. A natural history model for planning prostate cancer testing: Calibration and validation using Swedish registry data. PLoS One. 2019;14(2):e0211918.
24. NHS screening programmes in England: 2019 to 2020: Office for Health Improvement & Disparities; 2023 [updated 16 February 2023. Available from: https://www.gov.uk/government/publications/nhs-screening-programmes-annual-report/nhs-screening-programmes-in-england-2019-to-2020.
25. Keeling MJ, Ross JV. On methods for studying stochastic disease dynamics. J R Soc Interface. 2008;5(19):171–81.
26. Salisbury A, Pearce A, Howard K, Norris S. Impact of Structural Differences on the Modeled Cost-Effectiveness of Noninvasive Prenatal Testing. Med Decis Making. 2024;44(7):811–27.
27. Postma MJ, Beck EJ, Mandalia S, Sherr L, Walters MDS, Houweling H, et al. Universal HIV screening of pregnant women in England: cost effectiveness analysis. BMJ. 1999;318(7199):1656–60.
28. Roberts MC, Dotson WD, DeVore CS, Bednar EM, Bowen DJ, Ganiats TG, et al. Delivery Of Cascade Screening For Hereditary Conditions: A Scoping Review Of The Literature. Health Aff (Millwood). 2018;37(5):801–8.
29. McKnight D, Morales A, Hatchell KE, Bristow SL, Bonkowsky JL, Perry MS, et al. Genetic Testing to Inform Epilepsy Treatment Management From an International Study of Clinical Practice. JAMA Neurol. 2022;79(12):1267–76.
30. Botkin JR. Ethical issues in pediatric genetic testing and screening. Curr Opin Pediatr. 2016;28(6):700–4.
31. Lund JL, Rivera MP, Su IH, Long JM, Chen X, Pak J, et al. Estimating the Effects of Cancer Screening in Clinical Practice Settings: The Role of Selective Uptake and Suboptimal Adherence along the Cancer Screening Continuum. Cancer Epidemiol Biomarkers Prev. 2024;33(8):984–8.
32. Weller DP, Campbell C. Uptake in cancer screening programmes: a priority in cancer control. Br J Cancer. 2009;101 Suppl 2(Suppl 2):S55–9.
33. Iragorri N, Spackman E. Assessing the value of screening tools: reviewing the challenges and opportunities of cost-effectiveness analysis. Public Health Rev. 2018;39:17.
34. Sekiguchi M, Igarashi A, esaki M, saito Y, Kobayashi N, Matsuda T. Cost-Effectiveness Analysis of Fecal Immunochemical Test- and Colonoscopy-based Colorectal Cancer Screening across Varying Uptake Rates. DEN Open. 2026;6(1):e70236.
35. Masuku SD, Mandrik O, Mdege ND, Mishra G, Muwonge R, Meyer-Rath G, et al. Breast Cancer Screening Using Clinical Breast Examination: A Cost-Effectiveness Analysis for South Africa. Value Health Reg Issues. 2025;49:101127.
36. Davenport MS. Incidental Findings and Low-Value Care. American Journal of Roentgenology. 2023;221(1):117–23.

37. Bessey A, Chilcott J, Leaviss J, de la Cruz C, Wong R. A Cost-Effectiveness Analysis of Newborn Screening for Severe Combined Immunodeficiency in the UK. Int J Neonatal Screen. 2019;5(3):28.
38. Thomas C, Heathcote L, Sun Y, Callister MEJ, Kitt J, Rossi SH, et al. Cost-effectiveness of one-off upper abdominal CT screening as an add-on to lung cancer screening in England. Br J Cancer. 2025;133(2):239–47.
39. Smith RA, Mettlin CJ, Eyre H. Methodologic Issues in the Evaluation of Early Detection Programs. In: Kufe DW PR, Weichselbaum RR, editor. Cancer Medicine 6th edition: Holland-Frei 2003, Hamilton, Ontario, Canada.
40. Lee CL, Chuang CK, Chiu HC, Chang YH, Tu YR, Lo YT, et al. Understanding Genetic Screening: Harnessing Health Information to Prevent Disease Risks. Int J Med Sci. 2025;22(4):903–19.
41. Mandrik O, Thomas C, Whyte S, Chilcott J. Calibrating Natural History of Cancer Models in the Presence of Data Incompatibility: Problems and Solutions. Pharmacoeconomics. 2022;40(4):359–66.
42. Sasieni P, Smittenaar R, Hubbell E, Broggio J, Neal RD, Swanton C. Modelled mortality benefits of multi-cancer early detection screening in England. Br J Cancer. 2023;129(1):72–80.
43. Stout NK, Knudsen AB, Kong CY, McMahon PM, Gazelle GS. Calibration methods used in cancer simulation models and suggested reporting guidelines. Pharmacoeconomics. 2009;27(7):533–45.
44. Prosser LA, Grosse SD, Kemper AR, Tarini BA, Perrin JM. Decision analysis, economic evaluation, and newborn screening: challenges and opportunities. Genet Med. 2013;14(8):703–12.
45. Langer A, Holle R, John J. Specific guidelines for assessing and improving the methodological quality of economic evaluations of newborn screening. BMC Health Serv Res. 2012;12:300.
46. Castilla-Rodríguez I, Vallejo-Torres L, Couce ML, Valcárcel-Nazco C, Mar J, Serrano-Aguilar P. Cost-Effectiveness Methods and Newborn Screening Assessment. Adv Exp Med Biol. 2017;1031:267–81.
47. Owens L, Gulati R, Etzioni R. Stage Shift as an Endpoint in Cancer Screening Trials: Implications for Evaluating Multicancer Early Detection Tests. Cancer Epidemiol Biomarkers Prev. 2022;31(7):1298–304.
48. Rebolj M, Brentnall AR, Geppert J, et al. Late-Stage Outcomes as Surrogates for Mortality in Cancer Screening Trials: A Systematic Review and Meta-analysis. Cancer Epidemiol Biomarkers Prev. 2025 Oct 3;34(10):1694-1709.
49. Dai JY, Georg Luebeck E, Chang ET, Clarke CA, Hubbell EA, Zhang N, et al. Strong association between reduction of late-stage cancers and reduction of cancer-specific mortality in meta-regression of randomized screening trials across multiple cancer types. J Med Screen. 2024;31(4):211–22.
50. Daubner-Bendes R, Kovács S, Niewada M, Huic M, Drummond M, Ciani O, et al. Quo Vadis HTA for Medical Devices in Central and Eastern Europe? Recommendations to Address Methodological Challenges. Front Public Health. 2020;8:612410.
51. Garcia-Zamalloa A, Vicente D, Arnay R, Arrospide A, Taboada J, Castilla-Rodríguez I, et al. Diagnostic accuracy of adenosine deaminase for pleural tuberculosis in a low prevalence setting: A machine learning approach within a 7-year prospective multi-center study. PLoS One. 2021;16(11):e0259203.
52. Jackson CH, Sharples LD, Thompson SG. Structural and Parameter Uncertainty in Bayesian Cost-Effectiveness Models. Journal of the Royal Statistical Society Series C: Applied Statistics. 2009;59(2):233–53.
53. Briggs AH. Handling Uncertainty in Cost-Effectiveness Models. PharmacoEconomics. 2000;17(5):479–500.
54. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C, et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) Statement: Updated Reporting Guidance for Health Economic Evaluations. Value Health. 2022;25(1):3–9.

55. Pignone M, Saha S, Hoerger T, Mandelblatt J. Cost-effectiveness analyses of colorectal cancer screening: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med. 2002;137(2):96–104.
56. Shields GE, Clarkson P, Bullement A, Stevens W, Wilberforce M, Farragher T, et al. Advances in Addressing Patient Heterogeneity in Economic Evaluation: A Review of the Methods Literature. Pharmacoeconomics. 2024;42(7):737–49.
57. Nyame YA, Gulati R, Heijnsdijk EAM, Tsodikov A, Mariotto AB, Gore JL, et al. The Impact of Intensifying Prostate Cancer Screening in Black Men: A Model-Based Analysis. J Natl Cancer Inst. 2021;113(10):1336–42.
58. Whyte S, Thomas C, Chilcott J, Kearns B. Optimizing the Design of a Repeated Fecal Immunochemical Test Bowel Cancer Screening Programme With a Limited Endoscopy Capacity From a Health Economic Perspective. Value Health. 2022;25(6):954–64.
59. Sharma M, John R, Afrin S, Zhang X, Wang T, Tian M, et al. Cost-Effectiveness of Population Screening Programs for Cardiovascular Diseases and Diabetes in Low- and Middle-Income Countries: A Systematic Review. Front Public Health. 2022;10:820750.