Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Decision-Analytic Models of Detection Strategies for Upper Gastrointestinal Cancers: A Methodological Systematic Review

**Zhezhou He**

wpw518@qmul.ac.uk

Queen Mary University of London https://orcid.org/0009-0009-8955-597X

**Thurkga Moothathamby**

Queen Mary University of London https://orcid.org/0000-0002-1440-2719

**Garth Funston**

Queen Mary University of London https://orcid.org/0000-0002-4156-6401

**Runguo Wu**

Queen Mary University of London https://orcid.org/0000-0002-6523-5754

**Borislava Mihaylova**

Queen Mary University of London https://orcid.org/0000-0002-0951-1304

**Systematic Review**

**Keywords:**

**Posted Date:** March 18th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-9148986/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** The authors declare no competing interests.

# Decision-Analytic Models of Detection Strategies for Upper Gastrointestinal Cancers: A Methodological Systematic Review


## Abstract

**Background:** Early diagnosis of upper gastrointestinal (UGI) cancers is associated with better outcomes. However, decision-analytic models to evaluate UGI cancer detection, and diagnostic and treatment strategies vary widely in structure and underlying assumptions and have not been systematically appraised.

**Purpose:** To develop a structured taxonomy of UGI early detection decision models and identify methodological gaps to guide future modelling.

**Data Sources:** MEDLINE (Ovid), Embase, and the National Health Service Economic Evaluation Database (NHS EED) were searched to November 2024, supplemented by screening of relevant reviews and reference lists of included studies.

**Study Selection:** Modelling studies evaluating UGI cancer early detection through screening, surveillance, or multi-cancer approaches were included.

**Data Extraction:** A structured template captured study characteristics, model design, assumptions, and outcomes.

**Data Analysis:** A four-dimension taxonomy was developed describing simulation level, structural form, time progression, and system interactions.

**Data Synthesis:** Sixty-five studies were included, most focusing on gastric or oesophageal cancers (86%) and conducted in the United States and China (70%). Across studies, early detection strategies were most commonly modelled via cancer stage shift (60%), influencing outcomes through stage at diagnosis and downstream treatment effects. From these studies, 65 model structures were identified; 82% employed a state-transition structure, and none incorporated system-level dynamics or explicit resource constraints.

**Limitations:** This review has several limitations, including restriction to English-language publications and limited applicability of findings to healthcare systems with more specialised or structurally distinct diagnostic pathways.

**Conclusion:** Research on decision modelling of early detection strategies for UGI cancers has been conducted predominantly in gastric and oesophageal cancers. Greater attention is needed

for pancreatic, biliary tract, and gallbladder cancers, alongside more individual characteristics-based simulation approaches and considerations of resource constraints.

## Introduction

Upper gastrointestinal (UGI) cancers, defined here as oesophageal, gastric, pancreatic, gallbladder and biliary tract malignancies, often present with vague or non-specific symptoms that lack clinical specificity and frequently fail to prompt timely investigation [1-8]. By the time symptoms prompt investigation, disease is often advanced, and survival remains poor. Although UGI cancers account for roughly 12% of incident cancers worldwide, they are responsible for nearly 19% of cancer deaths, reflecting their disproportionate lethality [9-12]. The difficulty of recognising UGI cancers early, combined with their substantial mortality burden, provides a clear rationale for research into their earlier detection.

Early detection strategies often involve multiple components, including risk stratification, testing modalities, surveillance intervals and referral pathways, whose combined effects unfold over long time horizons and interact with system capacity and patient flow. These complexities are difficult to capture in observational studies or clinical trials alone [13, 14], particularly for cancers with low incidence or for strategies that require long term follow up. Decision-analytic modelling, therefore, has an increasingly important role in this area. It provides a structured way to synthesise evidence, compare alternative pathways and estimate downstream health and system impacts [15, 16].

However, the existing literature on modelling UGI cancer detection strategies remain fragmented, with no systematic overview or methodological framework guiding how decision-analytic modelling has been applied to study early detection in this cancer group. To address this need, the present review identifies and synthesises decision analytic modelling studies of early detection strategies for UGI cancers, with the aim of describing current methodological approaches, assessing the consistency and transparency of model structures and highlighting opportunities for further methodological development.

# Methods

## Protocol Registration

The review protocol was developed according to the best-practice guidance for health economic and modelling systematic reviews, including PRISMA 2020 and the Centre for Reviews and Dissemination (CRD) framework [17, 18], and was prospectively registered with PROSPERO (CRD42025623416).

## Search Strategy

A comprehensive literature search was undertaken across MEDLINE (via Ovid), Embase and the National Health System Economic Evaluation Database (NHS EED) up to 28 November 2024, supplemented by manual screening of reference lists from included studies and from systematic reviews identified through the database searches. The search strategy was organised around three domains: (1) UGI cancers and associated precursor conditions; (2) health economic evaluation and decision-analytic modelling; and (3) early detection strategies, encompassing screening, surveillance, and emerging diagnostic modalities such as liquid biopsy and circulating tumour DNA. Each domain combined controlled vocabulary (e.g., MeSH, Emtree) with free-text terms, making use of field tags, truncation, Boolean operators, and proximity operators as appropriate. The NHS EED strategy was adapted from those developed for MEDLINE and Embase. To ensure comprehensiveness and reproducibility, all search strategies were piloted and iteratively refined in consultation with a university librarian and benchmarked against a priori identified 10 key studies[19-28]. The references of these key studies, together with the full search strategies are presented in **Supplementary Materials S1 and S2**.

## Study Selection

*Inclusion Criteria*

Studies were eligible if they reported original research on early detection strategies or interventions for UGI cancers in adults, where the study population was predominantly aged 18 years or older. In this review, early detection denotes identifying cancer or precursor conditions, in asymptomatic general or high-risk populations as well as in symptomatic but undiagnosed individuals. Eligible strategies included population or targeted screening, surveillance at varying

intervals, recognition of early symptoms, and emerging approaches such as multi-cancer early detection (MCED), provided that outcomes were specific to UGI cancers. Eligible designs included health economic evaluations and decision-analytic modelling studies that assess detection strategies, model early diagnostic processes, or examine care pathway structures and transitions to diagnosis. Only studies reporting detection-related cancer outcomes (e.g., number of cancers detected or averted, cancer stage at diagnosis distribution, cancer-specific mortality, or survival), with or without associated cost outcomes, were included.

*Exclusion Criteria*

Studies were excluded if they were published as letters, notes, editorials, commentaries, or conference abstracts. Systematic reviews and meta-analyses were not eligible. Studies were also excluded if they focused exclusively on treatment rather than detection strategies, or if they reported only short-term outcomes or examined disease progression without assessing the long-term impact of early detection.

All records retrieved from the database searches were imported into Covidence [29] for deduplication and systematic screening NHS EED was searched manually in parallel, and relevant records were incorporated and deduplicated at the same stage as the main database outputs. Study selection followed a two-stage process: first, titles and abstracts were independently screened by ZH and TH (first and second authors) against the predefined inclusion and exclusion criteria; records not excluded at this stage progressed to full-text review. Second, the same authors independently assessed full texts, with disagreements resolved through discussion or, when necessary, consultation with a senior author. After screening of the primary database searches and NHS EED, the reference lists of included studies and systematic reviews identified through the database searches were manually screened to identify additional records. Identified records were deduplicated and assessed at full text.

*Data Extraction*

A structured, customised data extraction template (available in **Supplementary Material S3**) was used to ensure consistency and transparency [30]. For each study, we recorded bibliographic details (authors, year, country, funding source), study aims, target population, and cancer type. Methodological information included the simulation method, model type, time horizon,

information on patient heterogeneity, model validation approaches, and the specification of treatment effects and their propagation through the modelling framework. Key elements of the economic evaluation were also extracted, including evaluation type, study perspective, cost data, and summary measures [18]. Outcomes of interest comprised cancer diagnosis and detection rates, cancer stage at diagnosis, cancer survival, quality-adjusted life years (QALYs), incremental cost-effectiveness ratios (ICERs), and results from scenario or subgroup analyses. Data extraction was conducted independently by ZH and TH, with disagreements resolved in discussion with a senior author.

## Quality assessment

The quality of reporting and methodological rigour were appraised using the CHEQUE (Criteria for Health Economic Quality Evaluation) tool, which distinguishes between adequacy of methods and clarity of reporting [31]. Full appraisal criteria and scoring procedures are described in **Supplementary Materials S4 and S5**.

## Data synthesis and analysis

Given the heterogeneity of modelling approaches and reporting standards, evidence was synthesised narratively and summarised in structured tables. A taxonomy of modelling methods was developed to classify studies according to salient modelling features, including simulation level, structural form, time progression in model and system interactions. This taxonomy facilitated systematic comparison, allowing the identification of prevailing practices as well as methodological limitations and gaps. To examine variation in modelling approaches and findings, studies were grouped and compared by cancer site. Studies were also compared by setting or geographic region to provide contextual insights.

## Taxonomy Development

The taxonomy was developed based on findings from this review to provide a consistent, theory-informed framework for describing modelling approaches used in early detection. It defines four methodological domains: simulation level, structural form, time progression, and system interaction (**Table 1**).

Simulation level and stochasticity are treated jointly to distinguish cohort-based deterministic models from those generating individual event trajectories through first-order stochastic processes. Structural form captures the mathematical logic by which diagnostic and disease processes are represented. It defined how states, events, and dependencies are structured within a model, determining whether progression is expressed through discrete state transitions or event-driven processes and how history or memory is retained. Temporal representation was classified as non-dynamic, cycle-based, or continuous-time. Finally, system interaction was defined as the endogenous representation of queues, diagnostic delays, capacity constraints, or feedback between competing patient flows.

While grounded in established decision-science and simulation-modelling frameworks (including Brennan et al. [32] and Tsoi et al. [33]), it reflects the specific range of models identified in this review and does not aim to encompass all possible variants used across broader applications.

## Further Model Features

In addition to the methodological domains above, further key model features were also summarised. First, the ability of models to reflect patient heterogeneity through use of patient characteristics in determinising model parameters and outcomes was summarised. Second, validation of model performance was reviewed through use of (a) external validation, defined as empirical comparison of model outputs with independent data sources not used for model development or parametrisation; (b) internal validation, defined as empirical comparison of model outputs with observed data derived from the same underlying study population or data source used for model development or parameterisation; or (c) face validity, defined as external review of model data sources and parameters. Third, the approaches used to assess the impact of uncertainty in model parameters on model findings were reviews with focus on use of probabilistic or deterministic sensitivity analyses. Finally, the mechanisms to treatment effect propagation in the models were categorised according to how interventions altered the natural history of disease and how these changes were propagated through downstream outcomes. A mechanism refers to the pathway through which an intervention influences treatment eligibility, cancer incidence, cancer stage distribution, or cancer survival. Mechanisms were classified as

principal or concurrent based on their functional role within the intervention strategy. The principal mechanism represented the core pathway targeted to achieve improvements in outcomes. Concurrent mechanisms denoted distinct pathways that operated alongside the principal mechanism. Depending on the intervention design, concurrent mechanism functioned either as subsidiary factors modifying downstream outcomes or as substantive, independent drivers of benefit. This category frequently included multi-component interventions such as screening combined with precursor eradication or stratified surveillance.

***Table 1: Model Taxonomy***

| Model | Simulation level & Stochasticity | Structural form | Time progression | System interaction |
|---|---|---|---|---|
| Decision tree (DT) | Cohort-level; deterministic. | Static branching structure representing short-term or one-off choices. | Non-dynamic; no explicit time evolution. | No system interaction possible. Decision trees lack system dynamics: they represent only branching pathways without evolving states, making them unable to capture queues, capacity limits, feedback, or resource competition. |
| State-transition (cohort) | Cohort-level deterministic (expected-value outputs, no first-order uncertainty). Transition probabilities may be aggregate or individualised. | Markov or semi-Markov state-transition structure with predefined health states. Heterogeneity or history dependence may be incorporated through attributes, tunnel states, or stratified or individualised transition probabilities (e.g., individualised transition probabilities and outcomes. See *Heart Protection Study Collaborative Group* [34]). | Cycle-based. May be extended to semi-Markov form, where transition probabilities depend on time spent in the current state (sojourn-time). | Minimal system interaction possible: adding states or time elements increases structural detail but not true dynamics, as transition probabilities remain unresponsive to environmental factors such as congestion, capacity, or system pressures. |
| State-transition (microsimulation) | Individual-level; stochastic (first-order uncertainty). Each individual experiences a simulated event trajectory. | Markov or semi-Markov structure applied at the individual level; supports heterogeneity and history dependence through attributes and tunnel states. | Cycle-based or event-time based progression. | Partial system interaction possible: individual-level stochastic trajectories introduce within-model dynamics and history dependence, but transition probabilities unresponsive to environmental factors such as such as congestion, capacity, or system pressures. |
| Discrete-event simulation (DES) | Individual-level, stochastic. | Event-driven representation of entities, events, and shared resources; the system state evolves through scheduled events. | Continuous-time; explicit time-to-event scheduling and competing risks. | Full system interaction possible: DES provides a framework for modelling an evolving system state in which, conditional on model structure and assumptions, interactions such as queues, scheduling, resource competition, capacity constraints, congestion, and feedback loops may be generated endogenously through event-driven dynamics. |

# Results

## Search Results

The database searches identified 601 records (Embase n = 264; MEDLINE n = 183; NHS Economic Evaluation Database (NHS EED) n = 154). After duplicate removal, 503 records were screened by title and abstract, and 138 underwent full-text assessment.

Supplementary searches of reference lists from included studies (n = 56) and relevant systematic reviews identified an additional 97 records. After duplicate removal, 62 supplementary records were assessed at full text using the same eligibility criteria.

In total, 200 full-text articles were assessed for eligibility, and 65 studies met the inclusion criteria and were included in the final review. For clarity of reporting, records identified through database and supplementary searches are presented together in the PRISMA flow diagram (**Figure 1**).

## Study Characteristics

To provide an overview of the evidence base, this section summarises the key characteristics of the included studies. Detailed study information (study title, authors, and publication year), as well as trends in publication year over time, are reported in the **Supplementary Materials S6–S8**.

*Disease Focus of Included Models*

Among the 65 included studies, gastric cancer was the most frequently investigated UGI malignancy, reported in 33 studies (51%) [19-23, 25, 27, 35-60], followed by oesophageal cancer in 23 studies (35%) [26, 61-82] and pancreatic cancer in 5 studies (8%) [24, 28, 83-85]. Further four studies (6%) [86-89] examined multi-cancer or cross-site early detection strategies, including multi-cancer early detection (MCED) blood tests. No studies focused specifically on gallbladder or biliary tract cancers. Moreover, no study explicitly defined the healthcare setting, such as primary, secondary, or tertiary care, within which screening, diagnostic evaluation, or surveillance was modelled.

*Geographic Distribution of Studies*

Studies were classified into 4 major geographic regions, as depicted in **Figure 2**. Study activity was most strongly concentrated in North America, primarily reflecting studies undertaken in the United States (n = 30, 46%). East Asia contributed a substantial proportion of studies, dominated by China, including Hong Kong and Taiwan (n = 15, 23%), with additional contributions from Japan (n = 5, 8%) and Singapore (n = 4, 6%). Europe contributed fewer studies overall, most notably the United Kingdom (n = 4) and the Netherlands (n = 3). A small number of studies were conducted in Oceania and Latin America or spanned multiple regions; these were few in numbers and classified collectively as "other".

*Population Data and Targeting Approaches Used in Studies*

Across the included studies, approaches to population definition and screening and early detection varied.

Most studies relied on simulated populations, with only a small minority incorporating real population data to inform baseline population structure or key epidemiological parameters. This pattern was consistent across cancer sites. For oesophageal cancer, 22 of 23 studies (96%) [26, 61-69, 71-82] used simulated populations, with only one study (4%) [70] incorporating real population data. Similarly, 32 of 33 gastric cancer studies (97%) [19, 20, 22, 23, 25, 27, 35-55, 57-60] relied on simulated populations, with one exception [21] using real population data. No pancreatic or multiple UGI cancer studies used real population data.

Variation was also observed in the ways populations were targeted and incorporated into detection pathways. Nearly half of studies evaluated screening-only strategies (31/65, 48%), while others examined combined screening and surveillance pathways (22/65, 34%) or surveillance-only strategies in asymptomatic high-risk populations (12/65, 18%), see **Figure 3**. General-population approaches typically applied screening to unselected populations defined primarily by age and occasionally sex [45, 48, 55], whereas risk-based strategies restricted screening or surveillance to predefined higher-risk subgroups based on established clinical or demographic risk factors, most commonly gastro-oesophageal reflux disease, Helicobacter pylori infection, Barrett's oesophagus, dysplasia, or a family history of UGI cancer.

Finally, screening and early detection strategies differed across cancer sites. To avoid repetition in the main text, these strategies are summarised in **Supplementary Material S9**.

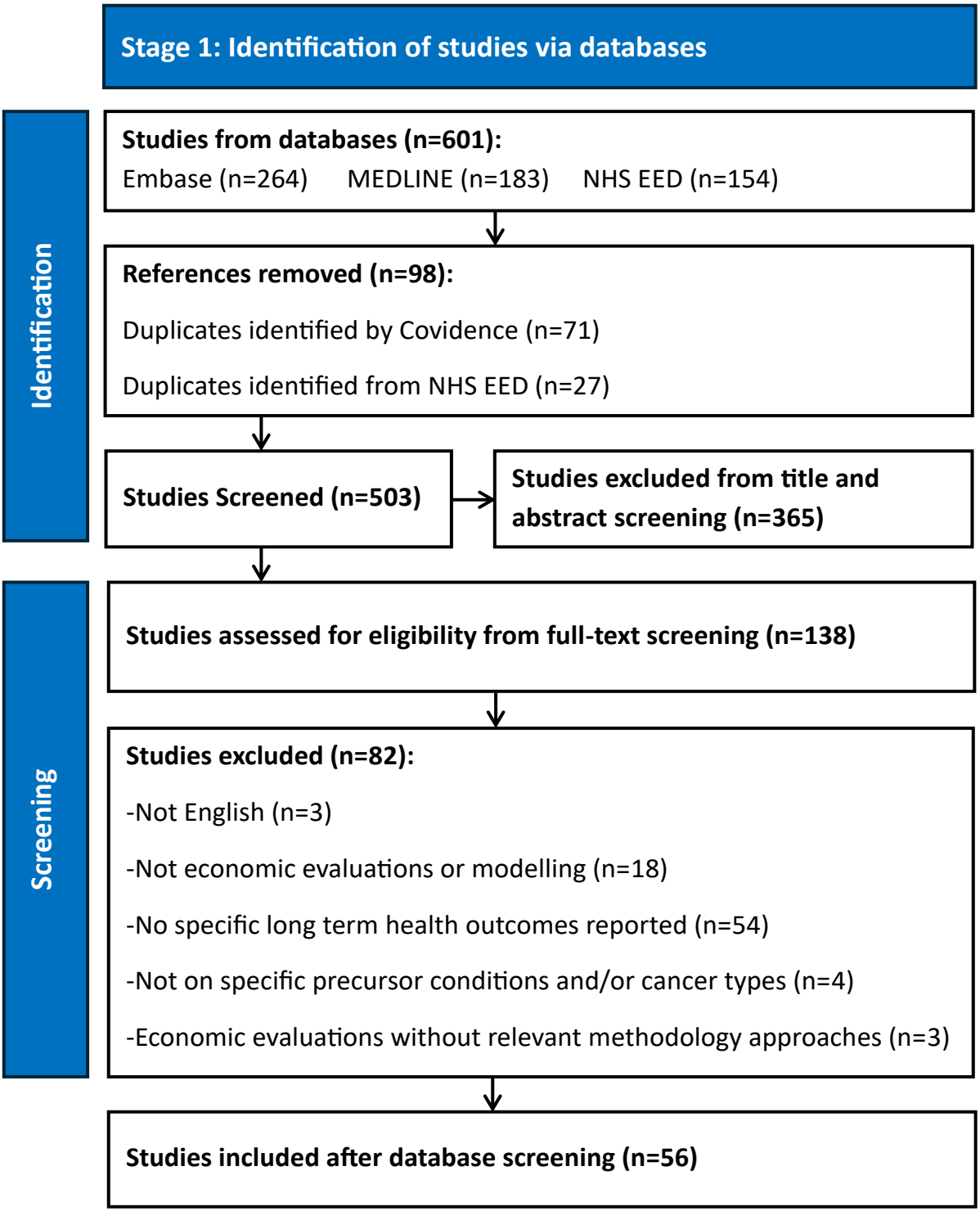
## Quality Assessment Results

The results of the reporting and methodological quality assessments are presented in **Figure 4** and **Figure 5**. These show that elements defining the decision problem, including objectives, populations, comparators, outcomes, and decision relevance, were generally reported well across studies. In contrast, items related to data quality, model validation and analysis of structural uncertainty were frequently absent or only partially addressed, both in reporting and methodological assessments. In particular, probabilistic sensitivity analysis, assessment of structural uncertainty, data quality appraisal, and model validation were often not reported. As a result, the review is able to characterise the prevailing modelling approaches and reported cost effectiveness results but provides limited information on how uncertainty was handled or how results were validated within individual studies.

***Figure 1: PRISMA Flow Chart***

Panels (a) and (b) illustrate two sequential stages of the same study selection process: (a) identification and screening via database searches; and (b) supplementary identification via reference list and systematic review screening.

(a)



(b)

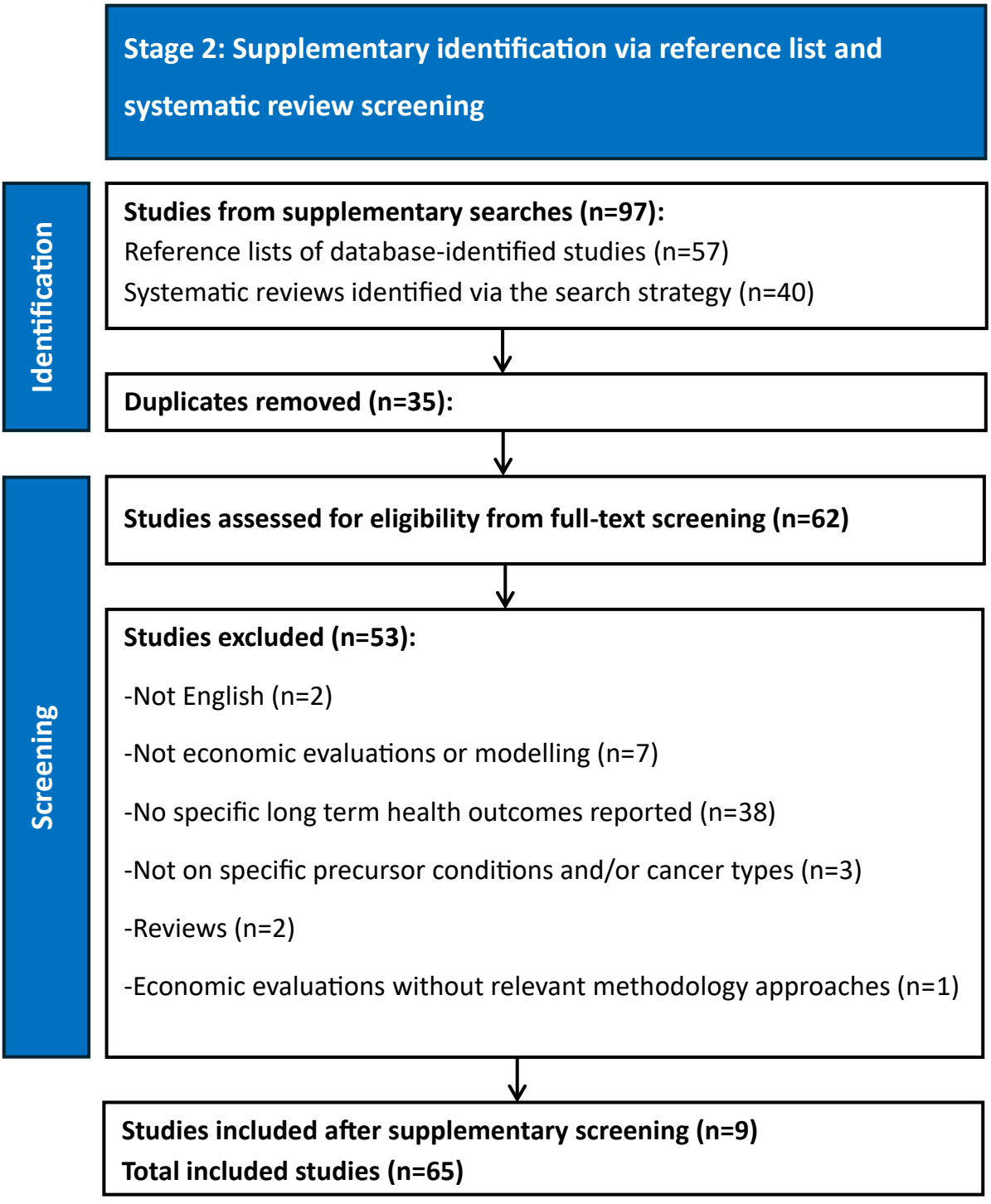

Stage 2: Supplementary identification via reference list and systematic review screening
Identification
Studies from supplementary searches (n=97):
Reference lists of database-identified studies (n=57)
Systematic reviews identified via the search strategy (n=40)
Duplicates removed (n=35):
Screening
Studies assessed for eligibility from full-text screening (n=62)
Studies excluded (n=53):
-Not English (n=2)
-Not economic evaluations or modelling (n=7)
-No specific long term health outcomes reported (n=38)
-Not on specific precursor conditions and/or cancer types (n=3)
-Reviews (n=2)
-Economic evaluations without relevant methodology approaches (n=1)
Studies included after supplementary screening (n=9)
Total included studies (n=65)

***Figure 2: Studies by Cancer Type and Geographical Region***

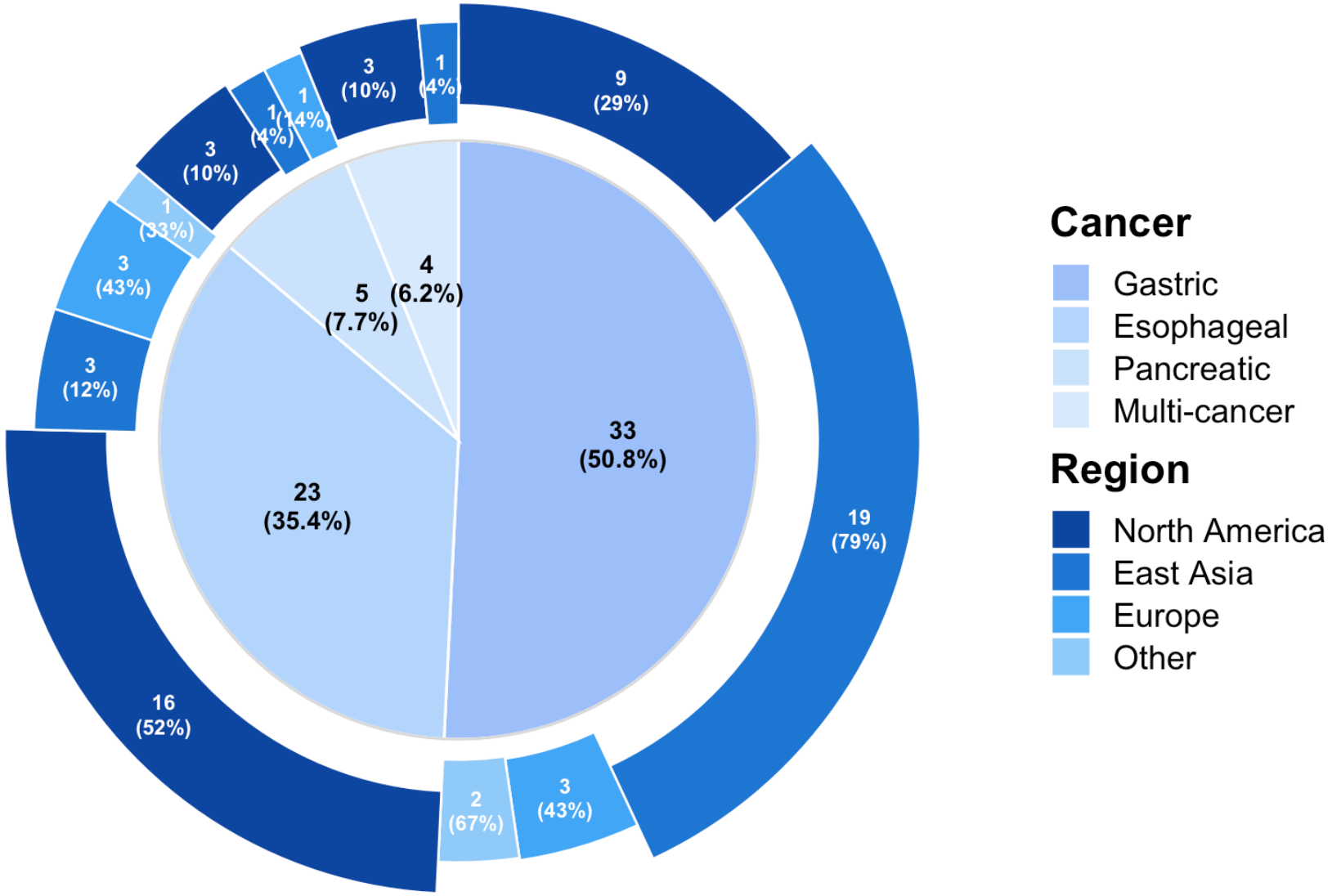


***Figure 3: Distribution of Early Detection Strategy Types and Target Populations***

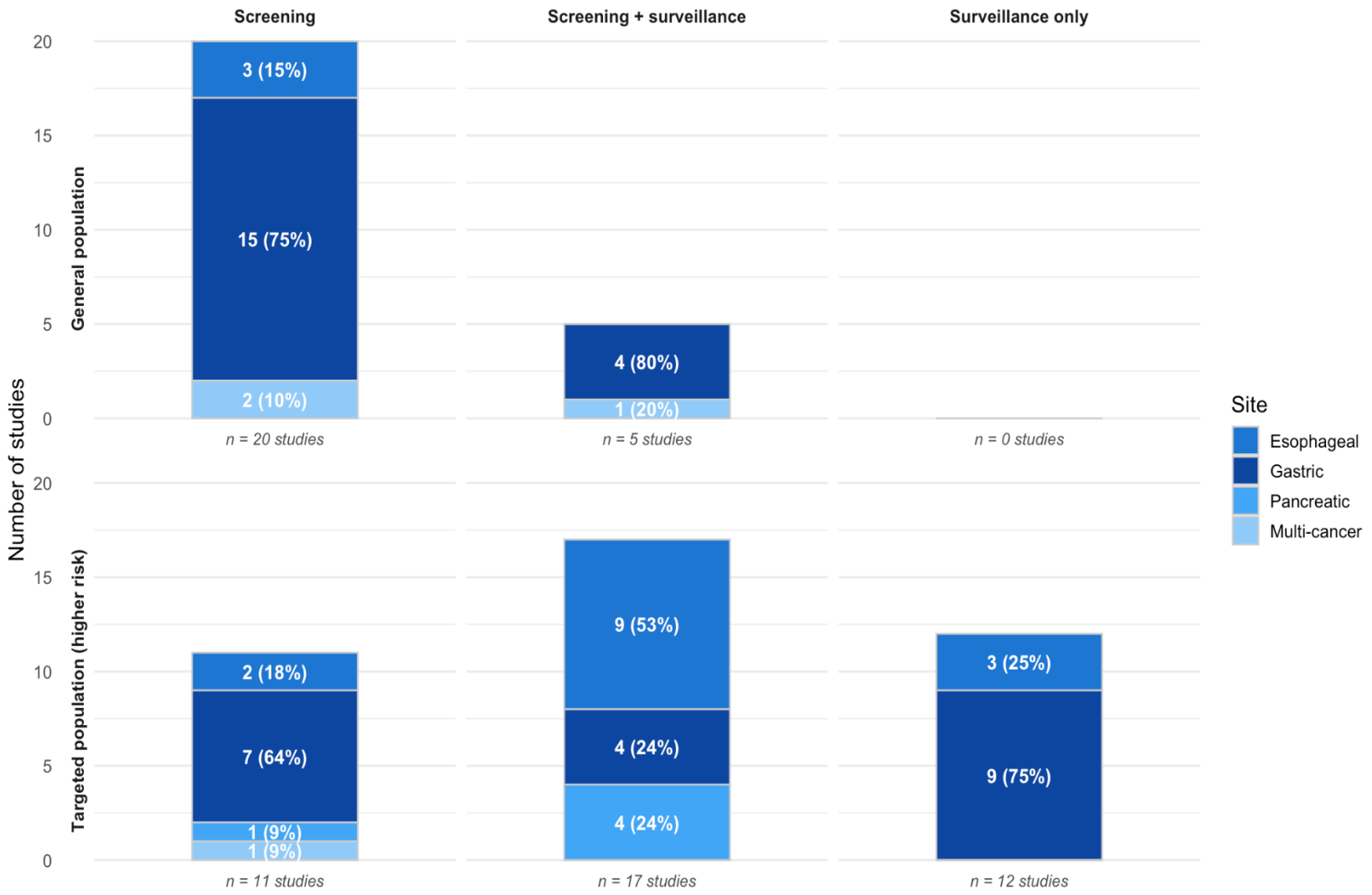
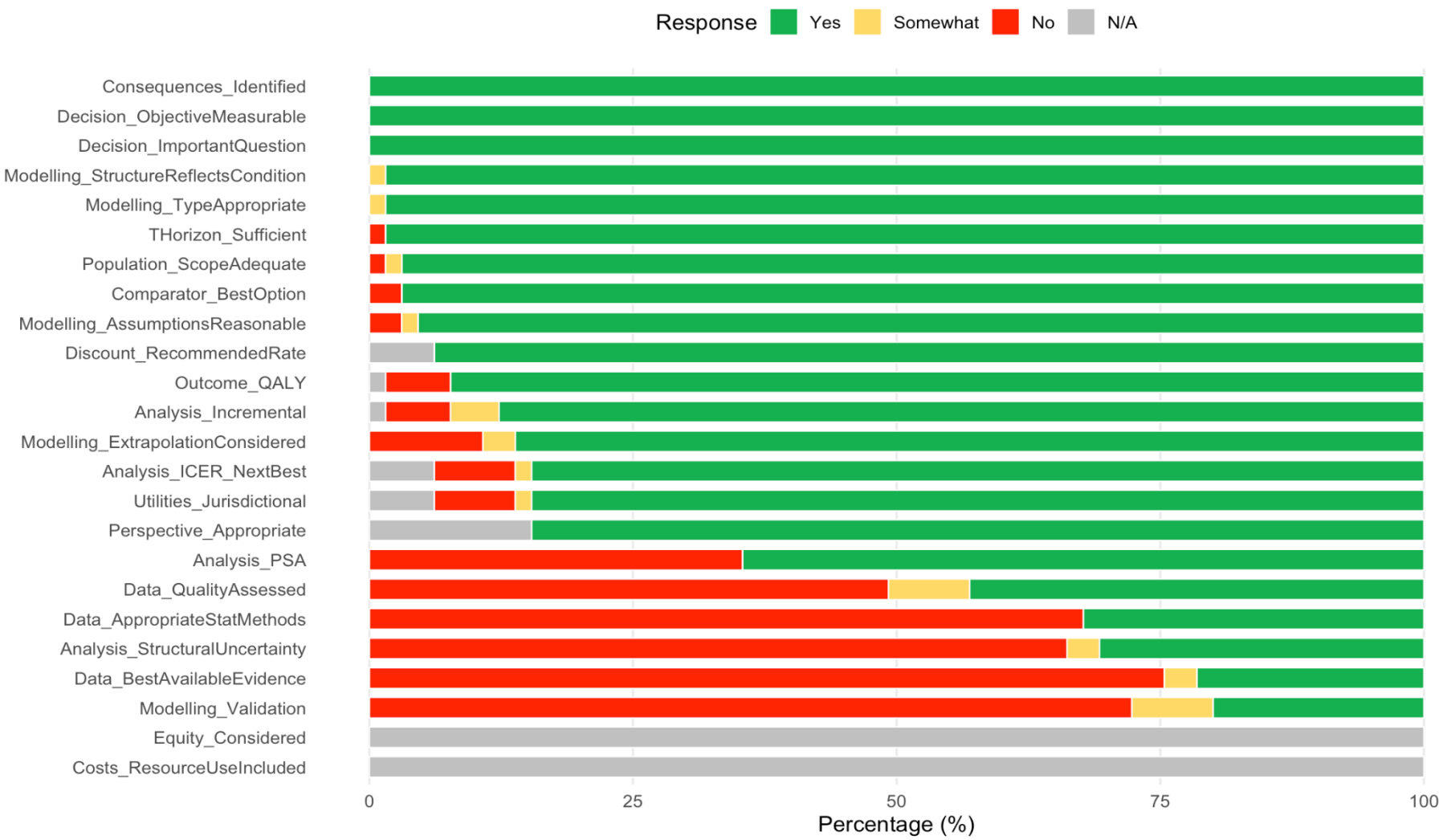


***Figure 4: Reporting Quality Results***



***Figure 5: Method Quality Results***



## Key Methodological Findings and Model Taxonomy

As individual studies frequently employed multiple modelling components or distinct modelling frameworks, model taxonomy was conducted at the level of the model structure, defined as the complete set of analytically linked components addressing a single prespecified decision problem. This approach distinguished structurally integrated modelling components that jointly formed one analytical framework from independently implemented model frameworks addressing the same decision context. Accordingly, structurally integrated components (e.g. a decision tree connected to a state-transition Markov model) were treated as a single model structure. Using this approach, 65 model structures were identified across the 65 included studies.

By contrast, three studies [65, 73, 76] applied the same externally developed and validated model frameworks to their respective decision problems, without analytical integration, and were therefore not counted as distinct models but treated as repeated applications of the same underlying models for the purposes of taxonomy. Model characteristics synthesised across the four methodological domains are summarised in **Figure 6**.

*Model Simulation Level and Stochasticity*

Most models operated at the cohort level with deterministic outputs. Of the 65 model structures, 51 (78%) were deterministic cohort models, representing populations as homogeneous groups and applying fixed transition probabilities such that identical inputs produced identical outputs. Thirteen models (20%) operated entirely at the individual level using microsimulation approaches, encompassing both health economic and biological or mechanistic individual based simulations, and generated first order variability across simulation runs. One model (2%) incorporated both cohort level and individual level components, for example a decision tree combined with an individual level semi-Markov microsimulation and was therefore classified separately as a standalone combination at the simulation level.

*Structural Form*

Across the 65 model structures identified, 54 (83%) employed state transition frameworks, representing the predominant modelling approach. Three models (5%) used discrete event simulation (DES), while one model (2%) adopted a fully decision tree (DT) structure. A further seven models (11%) employed a combined decision tree and state transition configuration, in which a decision tree was used to represent short term decision processes and a state transition model to capture longer term outcomes. Within these classifications, the majority of models were original modelling frameworks developed for the included studies, while three models, comprising two state transition models and one DES model, relied on externally validated modelling frameworks at the structural level [90, 91].

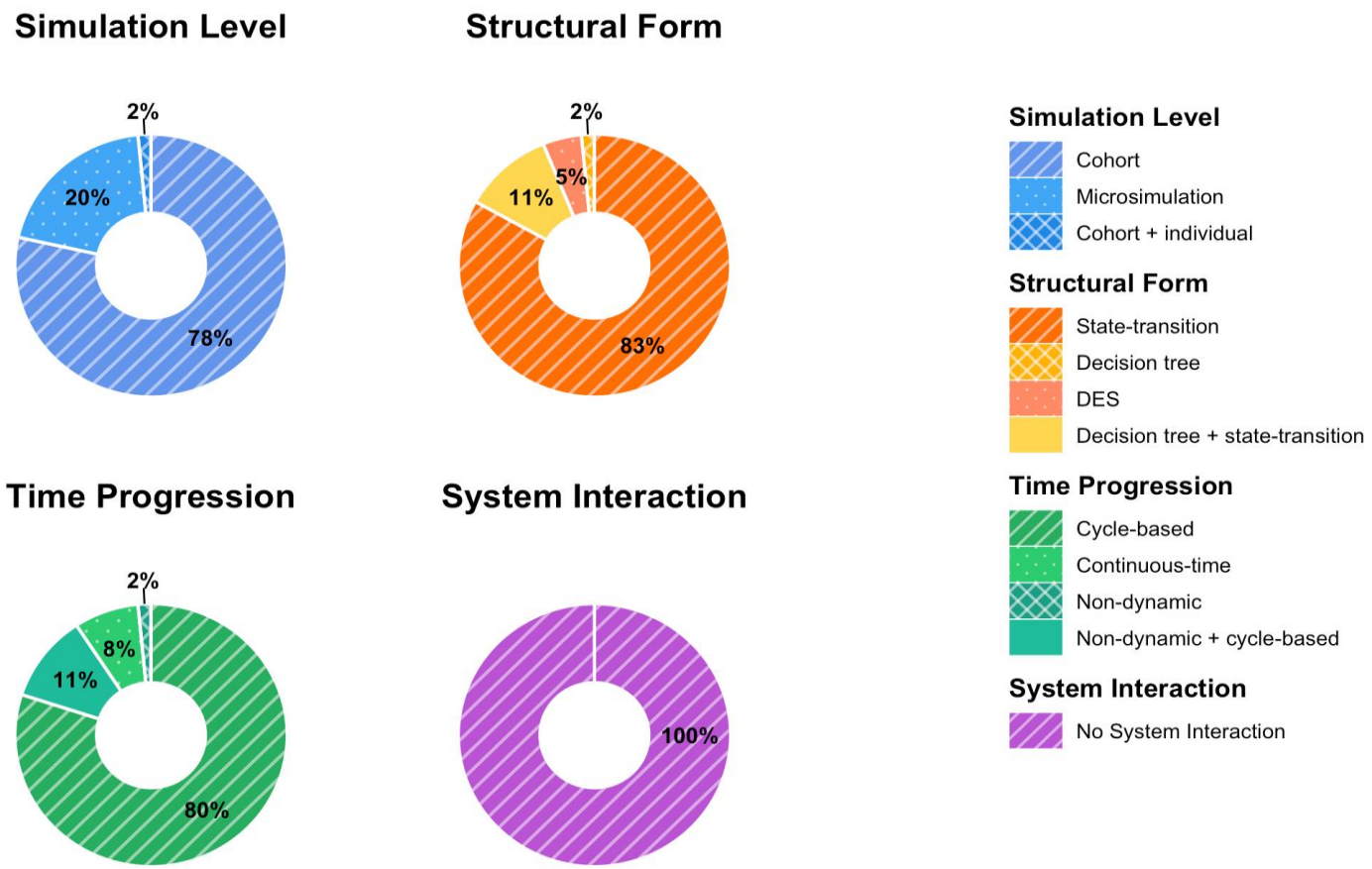
*Time Progression*

Of the 65 model structures identified, 52 (80%) employed cycle-based progression, with state occupancy updated at fixed intervals, typically monthly or annually. Five models (8%) used continuous time representations, including discrete event simulation, in which events occurred at individually timed intervals rather than at fixed cycle boundaries. One model (2%) adopted a non-dynamic structure, while a further seven models (11%) combined a non-dynamic component with a cycle-based state-transition framework.

*System Interaction*

None of the included model structures incorporated system interaction mechanisms. All evaluations assumed unconstrained access to diagnostic and treatment services, with no dynamic system dependencies.

***Figure 6: Model Taxonomy Results (Model Structure)***




Percentages are rounded to the nearest whole number and may not sum to 100%.

## Patient Heterogeneity

Patient characteristics, including population type, sex, ethnicity, and age, were commonly used to describe study populations. In this review, patient heterogeneity refers to whether such characteristics were incorporated into model structures in a way that allowed patient pathways or outcomes to differ within the analysis. Across the reviewed model structures, the extent to which patient heterogeneity was incorporated varied considerably. Overall, twenty-five model structures (32%) incorporated patient heterogeneity.

Among these, six model structures represented heterogeneity at the individual level using microsimulation approaches, including models that incorporated individual risk factors such as smoking status and Helicobacter pylori infection at baseline and, in some cases, allowed these exposures to change over time, thereby influencing disease progression and gastric cancer screening outcomes across the life course [27, 57], and comparative microsimulation analyses in which screening strategies and outcomes differed across sex- and race-specific populations within a single modelling framework [76].

A further nineteen model structures incorporated patient heterogeneity at the cohort level, through stratification into predefined subgroups with group-specific risks or transition parameters, while assuming homogeneity within each subgroup. This approach was implemented in several ways. Some studies stratified populations by demographic characteristics such as age, sex, and ethnicity, applying subgroup-specific gastric cancer incidence and screening effects to evaluate Helicobacter pylori screening in populations with elevated background risk, including indigenous or ethnic minority groups [20]. Other models defined risk groups based on clinical or pathological characteristics, such as the presence of precancerous gastric lesions, and evaluated endoscopic surveillance strategies separately for high- and low-risk subgroups without modelling individual-level variation within groups [60]. Additional studies stratified cohorts according to predefined high-risk clinical conditions, such as diabetes status or pancreatic precursor lesions, and assessed screening strategies within each subgroup using group-specific cancer risks and outcomes [24, 43].

The remaining forty-four model structures (68%) did not incorporate patient heterogeneity, applying average risks and screening effects uniformly across the population.

## Model Validation

Validation procedures were reported in 19 studies (19/65, 29%). External validation was the most frequently described approach and was reported in 13 studies (13/65, 20%). These assessments typically compared model outputs with independent epidemiological data or previously published estimates; in United States settings, this commonly included comparison with incidence profiles from the Surveillance, Epidemiology and End Results (SEER) programme.

Internal validation was conducted in three studies (3/65, 5%), and one additional study combined internal and external validation.

Face validity alone was reported in one study through expert appraisal, while a further study applied a broader validation framework incorporating face validity, cross-validation, internal validation, and external validation. Overall, 46 studies (71%) did not report any form of validation (see **Supplementary Material S10** for the extracted information on model validation).

## Parameter Uncertainty

At the study level (n = 65), approaches used to address second order uncertainty, defined as uncertainty in model input parameters, varied across the included studies.

Probabilistic sensitivity analysis (PSA) was reported in 40 studies (62%), while 25 studies (38%) did not report undertaking PSA.

Deterministic sensitivity analysis (DSA) was reported in 60 studies (92%). Among these studies, one way sensitivity analysis was the most frequently reported approach [44 studies (68%)]. A further 12 studies (18%) reported conducting both one way and two-way sensitivity analyses, while 3 studies (5%) reported two-way sensitivity analysis only. One study (2%) reported a combination of one way and three-way sensitivity analyses. In contrast, 5 studies (8%) did not report undertaking any form of deterministic sensitivity analysis (see **Supplementary Material S11** for the extracted information on parameter uncertainty).

## Treatment Effects in Models

*Stage Shift*

Stage shift was the principal mechanism in 39 of the included studies (60%). In these studies, stage shift arose because screening or surveillance detected cancers during the asymptomatic phase. Under natural history assumptions, diagnosis was assumed to occur only following symptom onset and therefore at more advanced stages. Early detection shifted the diagnostic distribution toward earlier, potentially curable stages, with subsequent improvements in survival and quality adjusted life years arising from the application of stage specific survival and utility inputs.

Stage shift functioned as a concurrent mechanism in 6 studies (9%), typically in studies where it operated alongside another principal mechanism, such as precursor eradication or stratified screening (see **Supplementary Material S12** for a summary of treatment effects).

*Risk-factor elimination*

Risk-factor elimination was the principal mechanism in 26 studies (40%). This approach predominantly reflected H.Pylori eradication in gastric cancer models and endoscopic ablation or oesophagectomy for BE or dysplasia in oesophageal cancer evaluations. Functioning fundamentally to prevent cancer onset, propagation occurred through reduced transition probabilities from precursor states to invasive cancer or by reassigning treated individuals to a lower-risk health state. This effectively blocked the onset of invasive disease, thereby decreasing cancer incidence relative to natural history. Risk-factor elimination appeared as a concurrent mechanism in 14 studies (22%), often in models where screening first identified precursor lesions and subsequent curative removal reduced future progression risk.

*Risk stratification*

Risk stratification appeared exclusively as a concurrent mechanism in 11 studies (17%). Stratification allocated screening or surveillance to individuals at higher baseline risk, defined by precursor lesions, genetic susceptibility or family history. It was not classified as a principal mechanism because stratification determines eligibility for the intervention rather than propagating health gains directly. In these studies, improvements in outcomes were driven by

the associated principal mechanism, most often stage shift or risk-factor elimination, while stratification functioned as a concurrent strategy that refined targeting.

## Discussion

This review identified 65 studies describing 65 distinct decision-analytic model structures of early detection for UGI cancers. A taxonomy was developed to classify models according to simulation level, structural form, time progression, and system interactions. Overall, the findings indicate a field with well-established methodological traditions, but with persistent gaps in data integration, representation of real-world detection processes, and coverage across geographic settings and disease sites.

### ***Cancer-Site and Geographic Concentration***

The concentration of modelling activity within UGI cancers and settings highlights disparities in the existing evidence base. Most studies focused on gastric cancer in East Asia and oesophageal cancer in the United States and United Kingdom, while pancreatic and other UGI cancers were seldom evaluated, the results are mostly consistent with patterns reported in previous literature [92-94]. It is not surprising, however, given that pancreatic, biliary tract, and gallbladder cancers are difficult to detect in the asymptomatic phase, highly lethal once diagnosed, and relatively uncommon compared with other more prevalent cancers worldwide [9, 10].

Geographically, modelling efforts have remained concentrated in high-income, data-rich health systems, particularly in North America and East Asia. Although this distribution reflects where data accessibility and research infrastructure are most advanced, it also reinforces an uneven allocation of analytical capacity across regions. This concentration limits the generalisability of existing evidence and leaves substantial gaps in understanding how early-detection strategies might perform in lower-resource settings or within structurally different health systems. Expanding modelling efforts to underrepresented cancers and contexts will be essential to build an evidence base that more accurately reflects the global epidemiology and operational diversity of UGI cancers.

### ***Dominant Analytical Paradigms and Structural Constraints***

Deterministic cohort state-transition models were the predominant analytical approach identified in this review, with relatively small number of studies employing microsimulation or

discrete-event simulation. However, even in these cases, key operational features such as waiting-time dynamics, queueing processes, and capacity constraints were not represented. This pattern is consistent with earlier reviews of economic evaluations in gastric and oesophageal cancers [92, 94, 95], and recent work on patient-level simulation across the broader oncology domain [96]. Although individual-based frameworks offer greater flexibility to represent heterogeneity and system dynamics, their wider use remains confined by data availabilities and technical complexity. Furthermore, methodological guidance argued that model complexity should align with the decision problem [16, 97]. When incorporating system-level dynamics is unlikely to affect cost-effectiveness outcomes, the additional computational and data burdens of discrete-event or more complex simulation may not be warranted [98, 99]. Thus, the balance between structural realism and analytical simplicity depends on context, data quality, and the marginal value of additional complexity in informing policy decisions.

***Real-World Populations, Empirical Data, and Validation***

This review revealed that modelling studies have rarely drawn on real-world patient data to directly inform model development and have infrequently reported either internal or external model validation. This pattern, also observed in a previous review of simulation models in oesophageal cancer [95], highlights a persistent gap between theoretical expectations and empirical practice. Many models continue to rely on simulated patient profiles as practical substitutes for real-world data. Such simplifications facilitate model construction when empirical data are limited, but they inevitably narrow the representation of clinical complexity and patient heterogeneity. It should be noted, however, access to linked registry, administrative, and consultation-level datasets are difficult to obtain, resource intensive to harmonise, and prone to inconsistencies in coding and temporal coverage [100-102]. Consequently, the limited adoption of real-world data largely reflects these practical constraints.

Nonetheless, models informed by real patients are more likely to reflect clinical reality and generate insights relevant to decision making. Validation procedures, in turn, assesses whether model outputs are plausible and consistent with observed practice. Rather than a fixed set of tests, it should be viewed as a flexible process that combines empirical evaluation with expert judgement, balanced according to data quality, model scope, and context. Low rates of model

validation, or inadequate reporting thereof, pose a significant risk to model credibility and limit the interpretability of results for decision makers.

Continued refinement of these empirical and methodological foundations will be critical for advancing the credibility and practical relevance of future modelling research.

***Treatment-Effect Propagation and Intervention Logic***

The identification of three classes of treatment effect propagation (stage shift, risk factor elimination, and risk stratification) illustrates how current modelling practices remain closely aligned with available evidence and established state transition conventions. The predominance of stage shift and risk factor elimination highlights a reliance on mechanisms that can be readily parameterised, whereas risk stratification, though conceptually broader, has been applied only concurrently. However, the persistent emphasis on parameter adjustment rather than explicit process representation limits the capacity to capture how early detection interventions reshape diagnostic behaviour or system dynamics. To address this limitation, future modelling should move beyond simple shifts in stage or risk distribution, integrating empirical pathway data and dynamic structures to represent how early detection exerts its effects within real health system contexts.

***Contribution to the literature***

To our knowledge, this is the first review to examine modelling approaches across the full UGI cancer spectrum and the first to classify how early detection models propagate treatment effects. Earlier health economic reviews in the UGI domain, including those on simulation models, have mainly focused on single cancers and offered limited methodological insight, underscoring the need for broader comparative evaluation across UGI modelling studies [93, 103-105]. By consolidating analytical approaches across oesophageal, gastric, pancreatic, biliary tract, and gallbladder cancers, this review highlights recurring structural constraints and areas where methodological development is most needed. The taxonomy introduced here was designed to capture key dimensions of early detection modelling within the UGI cancer domain and provides a structured basis for reporting and comparison. This framework may support clearer documentation, enable more consistent comparison across studies, and guide the development of future models.

***Limitations***

This review has several limitations. First, incomplete and inconsistent reporting across studies limited the availability of detailed methodological information, with key modelling features often described only briefly, hindering consistent data extraction. Second, restricting the search to English-language publications may have excluded relevant work published in other languages. Finally, limited reporting of care settings reduced the applicability of findings to health systems in which diagnostic services are more specialised or structurally distinct.

## Conclusion

Modelling for early detection of UGI cancers has focused on gastric and oesophageal sites, with minimal work on pancreatic, biliary tract, or gallbladder cancers. More attention to these under-represented sites and use of individual-based models with real-world patient data and healthcare settings, are needed.

## References


1. NCCN. NCCN Clinical Practice Guidelines in Oncology: Pancreatic Adenocarcinoma, https://www.nccn.org/guidelines/guidelines-detail?category=1&id=1455 (2025).
2. NCCN. NCCN Clinical Practice Guidelines in Oncology: Gastric Cancer, https://www.nccn.org/guidelines/guidelines-detail?category=1&id=1434 (2025).
3. NCCN. NCCN Clinical Practice Guidelines in Oncology: Esophageal and Esophagogastric Junction Cancers, https://www.nccn.org/guidelines/guidelines-detail?category=1&id=1433 (2025).
4. NCCN. NCCN Clinical Practice Guidelines in Oncology: Biliary Tract Cancers, https://www.nccn.org/guidelines/guidelines-detail?category=1&id=1517 (2025).
5. Conroy T, Pfeiffer P, Vilgrain V, et al. Pancreatic cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Annals of Oncology* 2023; 34: 987-1002. DOI: 10.1016/j.annonc.2023.08.009.
6. Lordick F, Carneiro F, Cascinu S, et al. Gastric cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Annals of Oncology* 2022; 33: 1005-1020. DOI: 10.1016/j.annonc.2022.07.004.
7. Obermannová R, Alsina M, Cervantes A, et al. Oesophageal cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Annals of Oncology* 2022; 33: 992-1004. DOI: 10.1016/j.annonc.2022.07.003.
8. Vogel A, Bridgewater J, Edeline J, et al. Biliary tract cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Annals of Oncology* 2023; 34: 127-140. DOI: 10.1016/j.annonc.2022.10.506.
9. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021; 71: 209-249. 20210204. DOI: 10.3322/caac.21660.
10. Arnold M, Abnet CC, Neale RE, et al. Global Burden of 5 Major Types of Gastrointestinal Cancer. *Gastroenterology* 2020; 159: 335-349 e315. 20200402. DOI: 10.1053/j.gastro.2020.02.068.
11. Arnold M, Rutherford MJ, Bardot A, et al. Progress in cancer survival, mortality, and incidence in seven high-income countries 1995–2014 (ICBP SURVMARK-2): a population-based study. *The Lancet Oncology* 2019; 20: 1493-1505. DOI: 10.1016/S1470-2045(19)30456-5.
12. Berrill JW, Turner JK, Hurley JJ, et al. Upper gastrointestinal cancer in its early stages is predominantly asymptomatic. *Frontline Gastroenterology* 2012; 3: 47. DOI: 10.1136/flgastro-2011-100026.
13. Fitzgerald RC, di Pietro M, O'Donovan M, et al. Cytosponge-trefoil factor 3 versus usual care to identify Barrett's oesophagus in a primary care setting: a multicentre, pragmatic, randomised controlled trial. *The Lancet* 2020; 396: 333-344. DOI: 10.1016/S0140-6736(20)31099-0.
14. Old O, Jankowski J, Attwood S, et al. Barrett's Oesophagus Surveillance Versus Endoscopy at Need Study (BOSS): A Randomized Controlled Trial. *Gastroenterology* 2025; 169: 1233-1243.e1238. DOI: 10.1053/j.gastro.2025.03.021.
15. Sculpher MJ, Claxton K, Drummond M, et al. Whither trial-based economic evaluation for health care decision making? *Health Economics* 2006; 15: 677-687. DOI: 10.1002/hec.1093.

16. Briggs AH, Claxton K and Sculpher MJ. *Decision Modelling for Health Economic Evaluation*. Oxford: Oxford University Press, 2006.
17. Centre for Reviews and Dissemination. Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care, https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf (2009).
18. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021; 372: n71. DOI: 10.1136/bmj.n71.
19. Davies R, Crabbe D, Roderick P, et al. A simulation to evaluate screening for Helicobacter pylori infection in the prevention of peptic ulcers and gastric cancers. *Health Care Manag Sci* 2002; 5: 249-258. DOI: 10.1023/a:1020326005465.
20. Teng AM, Kvizhinadze G, Nair N, et al. A screening program to test and treat for Helicobacter pylori infection: Cost-utility analysis by age, sex and ethnicity. *BMC Infect Dis* 2017; 17: 156. 20170220. DOI: 10.1186/s12879-017-2259-2.
21. Ma Y, Zhou X, Liu Y, et al. An Economic Evaluation of Family-Based Versus Traditional Helicobacter pylori Screen-and-Treat Strategy: Based on Real-World Data and Microsimulation Model. *Helicobacter* 2024; 29: e13123. DOI: 10.1111/hel.13123.
22. Roderick P, Davies R, Raftery J, et al. Cost-effectiveness of population screening for Helicobacter pylori in preventing gastric cancer and peptic ulcer disease, using simulation. *J Med Screen* 2003; 10: 148-156. DOI: 10.1177/096914130301000310.
23. Lee YC, Lin JT, Wu HM, et al. Cost-effectiveness analysis between primary and secondary preventive strategies for gastric cancer. *Cancer Epidemiol Biomarkers Prev* 2007; 16: 875-885. DOI: 10.1158/1055-9965.EPI-06-0758.
24. Kowada A. Cost-effectiveness of MicroRNA for Pancreatic Cancer Screening in Patients With Diabetes. *Pancreas* 2022; 51: 1019-1028. DOI: 10.1097/MPA.0000000000002130.
25. Wong IO, Schooling CM and Cowling BJ. Cost-effectiveness of Helicobacter pylori screening and treatment for gastric cancer in Hong Kong: a decision analytic approach. *Hong Kong Med J* 2014; 20 Suppl 7: 13-15.
26. Swart N, Maroni R, Muldrew B, et al. Economic evaluation of Cytosponge(R)-trefoil factor 3 for Barrett esophagus: A cost-utility analysis of randomised controlled trial data. *EClinicalMedicine* 2021; 37: 100969. 20210618. DOI: 10.1016/j.eclinm.2021.100969.
27. Huang HL, Leung CY, Saito E, et al. Effect and cost-effectiveness of national gastric cancer screening in Japan: a microsimulation modeling study. *BMC Med* 2020; 18: 257. 20200914. DOI: 10.1186/s12916-020-01729-0.
28. Pandharipande PV, Heberle C, Dowling EC, et al. Targeted screening of individuals at high risk for pancreatic cancer: results of a simulation model. *Radiology* 2015; 275: 177-187. 20141112. DOI: 10.1148/radiol.14141282.
29. Veritas Health Innovation. Covidence systematic review software, https://www.covidence.org.
30. Karagiannidou M, Wittenberg R, Landeiro FIT, et al. Systematic literature review of methodologies and data sources of existing economic models across the full spectrum of Alzheimer's disease and dementia from apparently healthy through disease progression to end of life care: a systematic review protocol. *BMJ Open* 2018; 8: e020638. DOI: 10.1136/bmjopen-2017-020638.
31. Kim DD, Do LA, Synnott PG, et al. Developing Criteria for Health Economic Quality Evaluation Tool. *Value Health* 2023; 26: 1225-1234. 20230415. DOI: 10.1016/j.jval.2023.04.004.

32. Brennan A, Chick SE and Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ* 2006; 15: 1295-1310. DOI: 10.1002/hec.1148.
33. Tsoi B, O'Reilly D, Jegathisawaran J, et al. Systematic narrative review of decision frameworks to select the appropriate modelling approaches for health economic evaluations. *BMC Research Notes* 2015; 8: 244. DOI: 10.1186/s13104-015-1202-0.
34. Heart Protection Study Collaborative; Borislava Mihaylova AB, Jane Armitage, Sarah Parish, Alastair Gray, Rory Collins. Lifetime cost effectiveness of simvastatin in a range of risk groups and age groups derived from a randomised trial of 20 536 people. *BMJ* 2006; 333: 1145. DOI: 10.1136/bmj.38993.731725.BE.
35. Ascherman B, Oh A and Hur C. International cost-effectiveness analysis evaluating endoscopic screening for gastric cancer for populations with low and high risk. *Gastric Cancer* 2021; 24: 878-887. 20210217. DOI: 10.1007/s10120-021-01162-z.
36. Chen Q, Liang X, Long X, et al. Cost-effectiveness analysis of screen-and-treat strategy in asymptomatic Chinese for preventing Helicobacter pylori-associated diseases. *Helicobacter* 2019; 24: e12563. 20190122. DOI: 10.1111/hel.12563.
37. Dan YY, So JB and Yeoh KG. Endoscopic screening for gastric cancer. *Clin Gastroenterol Hepatol* 2006; 4: 709-716. DOI: 10.1016/j.cgh.2006.03.025.
38. Fendrick AM, Chernew ME, Hirth RA, et al. Clinical and economic effects of population-based Helicobacter pylori screening to prevent gastric cancer. *Arch Intern Med* 1999; 159: 142-148. DOI: 10.1001/archinte.159.2.142.
39. Han Y, Yan T, Ma H, et al. Cost-Effectiveness Analysis of Helicobacter pylori Eradication Therapy for Prevention of Gastric Cancer: A Markov Model. *Dig Dis Sci* 2020; 65: 1679-1688. 20191030. DOI: 10.1007/s10620-019-05910-1.
40. Harris RA, Owens DK, Witherell H, et al. Helicobacter pylori and gastric cancer: what are the benefits of screening only for the CagA phenotype of H. pylori? *Helicobacter* 1999; 4: 69-76. DOI: 10.1046/j.1523-5378.1999.98057.x.
41. Hassan C, Zullo A, Di Giulio E, et al. Cost-effectiveness of endoscopic surveillance for gastric intestinal metaplasia. *Helicobacter* 2010; 15: 221-226. DOI: 10.1111/j.1523-5378.2010.00752.x.
42. Kowada A and Asaka M. Economic and health impacts of introducing Helicobacter pylori eradication strategy into national gastric cancer policy in Japan: A cost-effectiveness analysis. *Helicobacter* 2021; 26: e12837. 20210718. DOI: 10.1111/hel.12837.
43. Kowada A. A Population-Based Helicobacter pylori Eradication Strategy Is More Cost-Effective than Endoscopic Screening. *Dig Dis Sci* 2023; 68: 1735-1746. 20221224. DOI: 10.1007/s10620-022-07795-z.
44. Kowada A. Cost-Effectiveness of Population-Based Helicobacter pylori Screening With Eradication for Optimal Age of Implementation. *Helicobacter* 2024; 29: e13120. DOI: 10.1111/hel.13120.
45. Libanio D, Antonelli G, Marijnissen F, et al. Combined gastric and colorectal cancer endoscopic screening may be cost-effective in Europe with the implementation of artificial intelligence: an economic evaluation. *Eur J Gastroenterol Hepatol* 2024; 36: 155-161. 20231115. DOI: 10.1097/MEG.0000000000002680.
46. Ma J, Yu M, Shao QQ, et al. Both family-based Helicobacter pylori infection control and management strategy and screen-and-treat strategy are cost-effective for gastric cancer prevention. *Helicobacter* 2022; 27: e12911. 20220615. DOI: 10.1111/hel.12911.

47. Oh A, Truong H, Kim J, et al. Cost-effectiveness of screening with polymerase chain reaction for Helicobacter pylori to prevent gastric cancer and peptic ulcers. *J Gastrointest Oncol* 2022; 13: 2186-2196. DOI: 10.21037/jgo-21-911.
48. Parsonnet J, Harris RA, Hack HM, et al. Modelling cost-effectiveness of Helicobacter pylori screening to prevent gastric cancer: a mandate for clinical trials. *Lancet* 1996; 348: 150-154. DOI: 10.1016/s0140-6736(96)01501-2.
49. Qin S, Wang X, Li S, et al. Clinical Benefit and Cost Effectiveness of Risk-Stratified Gastric Cancer Screening Strategies in China: A Modeling Study. *Pharmacoeconomics* 2022; 40: 725-737. 20220615. DOI: 10.1007/s40273-022-01160-8.
50. Qin S, Wang X, Li S, et al. Benefit-to-harm ratio and cost-effectiveness of government-recommended gastric cancer screening in China: A modeling study. *Front Public Health* 2022; 10: 955120. 20220810. DOI: 10.3389/fpubh.2022.955120.
51. Wang Z, Han W, Xue F, et al. Nationwide gastric cancer prevention in China, 2021-2035: a decision analysis on effect, affordability and cost-effectiveness optimisation. *Gut* 2022; 71: 2391-2400. 20220728. DOI: 10.1136/gutjnl-2021-325948.
52. Xie F, Luo N, Blackhouse G, et al. Cost-effectiveness analysis of Helicobacter pylori screening in prevention of gastric cancer in Chinese. *Int J Technol Assess Health Care* 2008; 24: 87-95. DOI: 10.1017/S0266462307080117.
53. Xie F, Luo N and Lee HP. Cost effectiveness analysis of population-based serology screening and (13)C-Urea breath test for Helicobacter pylori to prevent gastric cancer: a markov model. *World J Gastroenterol* 2008; 14: 3021-3027. DOI: 10.3748/wjg.14.3021.
54. Xie F, O'Reilly D, Ferrusi IL, et al. Illustrating economic evaluation of diagnostic technologies: comparing Helicobacter pylori screening strategies in prevention of gastric cancer in Canada. *J Am Coll Radiol* 2009; 6: 317-323. DOI: 10.1016/j.jacr.2009.01.022.
55. Yeh JM, Ho W and Hur C. Cost-effectiveness of endoscopic surveillance of gastric ulcers to improve survival. *Gastrointest Endosc* 2010; 72: 33-43. 20100428. DOI: 10.1016/j.gie.2010.01.047.
56. Yeh JM, Hur C, Kuntz KM, et al. Cost-effectiveness of treatment and endoscopic surveillance of precancerous lesions to prevent gastric cancer. *Cancer* 2010; 116: 2941-2953. DOI: 10.1002/cncr.25030.
57. Yeh JM, Hur C, Ward Z, et al. Gastric adenocarcinoma screening and prevention in the era of new biomarker and endoscopic technologies: a cost-effectiveness analysis. *Gut* 2016; 65: 563-574. 20150316. DOI: 10.1136/gutjnl-2014-308588.
58. Yeh JM, Kuntz KM, Ezzati M, et al. Exploring the cost-effectiveness of Helicobacter pylori screening to prevent gastric cancer in China in anticipation of clinical trial results. *Int J Cancer* 2009; 124: 157-166. DOI: 10.1002/ijc.23864.
59. Zhang C, Qi YB, Hu RB, et al. Family-based Helicobacter pylori infection control and management strategy and screen-and-treat strategy are highly cost-effective in preventing multiple upper gastrointestinal diseases in Chinese population at national level. *Helicobacter* 2024; 29: e13063. DOI: 10.1111/hel.13063.
60. Zhou HJ, Dan YY, Naidoo N, et al. A cost-effectiveness analysis evaluating endoscopic surveillance for gastric cancer for populations with low to intermediate risk. *PLoS One* 2013; 8: e83959. 20131227. DOI: 10.1371/journal.pone.0083959.

61. Benaglia T, Sharples LD, Fitzgerald RC, et al. Health benefits and cost effectiveness of endoscopic and nonendoscopic cytosponge screening for Barrett's esophagus. *Gastroenterology* 2013; 144: 62-73 e66. 20121003. DOI: 10.1053/j.gastro.2012.09.060.
62. Gerson L and Lin OS. Cost-benefit analysis of capsule endoscopy compared with standard upper endoscopy for the detection of Barrett's esophagus. *Clin Gastroenterol Hepatol* 2007; 5: 319-325. DOI: 10.1016/j.cgh.2006.12.022.
63. Gerson LB, Groeneveld PW and Triadafilopoulos G. Cost-effectiveness model of endoscopic screening and surveillance in patients with gastroesophageal reflux disease. *Clin Gastroenterol Hepatol* 2004; 2: 868-879. DOI: 10.1016/s1542-3565(04)00394-5.
64. Gordon LG, Mayne GC, Hirst NG, et al. Cost-effectiveness of endoscopic surveillance of non-dysplastic Barrett's esophagus. *Gastrointest Endosc* 2014; 79: 242-256 e246. 20130927. DOI: 10.1016/j.gie.2013.07.046.
65. Heberle CR, Omidvari AH, Ali A, et al. Cost Effectiveness of Screening Patients With Gastroesophageal Reflux Disease for Barrett's Esophagus With a Minimally Invasive Cell Sampling Device. *Clin Gastroenterol Hepatol* 2017; 15: 1397-1404 e1397. 20170224. DOI: 10.1016/j.cgh.2017.02.017.
66. Honing J, Kievit W, Bookelaar J, et al. Endosheath ultrathin transnasal endoscopy is a cost-effective method for screening for Barrett's esophagus in patients with GERD symptoms. *Gastrointest Endosc* 2019; 89: 712-722 e713. 20181029. DOI: 10.1016/j.gie.2018.10.024.
67. Hur C, Nishioka NS and Gazelle GS. Cost-effectiveness of aspirin chemoprevention for Barrett's esophagus. *J Natl Cancer Inst* 2004; 96: 316-325. DOI: 10.1093/jnci/djh039.
68. Inadomi JM, Sampliner R, Lagergren J, et al. Screening and surveillance for Barrett esophagus in high-risk groups: a cost-utility analysis. *Ann Intern Med* 2003; 138: 176-186. DOI: 10.7326/0003-4819-138-3-200302040-00009.
69. Inadomi JM, Somsouk M, Madanick RD, et al. A cost-utility analysis of ablative therapy for Barrett's esophagus. *Gastroenterology* 2009; 136: 2101-2114 e2101-2106. 20090306. DOI: 10.1053/j.gastro.2009.02.062.
70. Kastelein F, van Olphen S, Steyerberg EW, et al. Surveillance in patients with long-segment Barrett's oesophagus: a cost-effectiveness analysis. *Gut* 2015; 64: 864-871. 20140718. DOI: 10.1136/gutjnl-2014-307197.
71. Li Y, Du L, Wang Y, et al. Modeling the Cost-effectiveness of Esophageal Cancer Screening in China. *Cost Eff Resour Alloc* 2020; 18: 33. 20200910. DOI: 10.1186/s12962-020-00230-y.
72. Nietert PJ, Silverstein MD, Mokhashi MS, et al. Cost-effectiveness of screening a population with chronic gastroesophageal reflux. *Gastrointest Endosc* 2003; 57: 311-318. DOI: 10.1067/mge.2003.101.
73. Omidvari AH, Ali A, Hazelton WD, et al. Optimizing Management of Patients With Barrett's Esophagus and Low-Grade or No Dysplasia Based on Comparative Modeling. *Clin Gastroenterol Hepatol* 2020; 18: 1961-1969. 20191206. DOI: 10.1016/j.cgh.2019.11.058.
74. Provenzale D, Kemp JA, Arora S, et al. A guide for surveillance of patients with Barrett's esophagus. *Am J Gastroenterol* 1994; 89: 670-680.
75. Rubenstein JH, Inadomi JM, Brill JV, et al. Cost utility of screening for Barrett's esophagus with esophageal capsule endoscopy versus conventional upper endoscopy. *Clin Gastroenterol Hepatol* 2007; 5: 312-318. DOI: 10.1016/j.cgh.2006.12.008.

76. Rubenstein JH, Omidvari AH, Lauren BN, et al. Endoscopic Screening Program for Control of Esophageal Adenocarcinoma in Varied Populations: A Comparative Cost-Effectiveness Analysis. *Gastroenterology* 2022; 163: 163-173. 20220329. DOI: 10.1053/j.gastro.2022.03.037.
77. Rubenstein JH, Vakil N and Inadomi JM. The cost-effectiveness of biomarkers for predicting the development of oesophageal adenocarcinoma. *Aliment Pharmacol Ther* 2005; 22: 135-146. DOI: 10.1111/j.1365-2036.2005.02536.x.
78. Shaheen NJ, Inadomi JM, Overholt BF, et al. What is the best management strategy for high grade dysplasia in Barrett's oesophagus? A cost effectiveness analysis. *Gut* 2004; 53: 1736-1744. DOI: 10.1136/gut.2003.033837.
79. Singer ME and Smith MS. Wide Area Transepithelial Sampling with Computer-Assisted Analysis (WATS(3D)) Is Cost-Effective in Barrett's Esophagus Screening. *Dig Dis Sci* 2021; 66: 1572-1579. 20200623. DOI: 10.1007/s10620-020-06412-1.
80. Vij R, Triadafilopoulos G, Owens DK, et al. Cost-effectiveness of photodynamic therapy for high-grade dysplasia in Barrett's esophagus. *Gastrointest Endosc* 2004; 60: 739-756. DOI: 10.1016/s0016-5107(04)02167-4.
81. Wu B, Wang Z and Zhang Q. Age at Initiation and Frequency of Screening to Prevent Esophageal Squamous Cell Carcinoma in High-risk Regions: an Economic Evaluation. *Cancer Prev Res (Phila)* 2020; 13: 543-550. 20200309. DOI: 10.1158/1940-6207.CAPR-19-0477.
82. Xia R, Li H, Shi J, et al. Cost-effectiveness of risk-stratified endoscopic screening for esophageal cancer in high-risk areas of China: a modeling study. *Gastrointest Endosc* 2022; 95: 225-235 e220. 20210819. DOI: 10.1016/j.gie.2021.08.008.
83. Ibrahim IS, Vasen HFA, Wasser M, et al. Cost-effectiveness of pancreas surveillance: The CDKN2A-p16-Leiden cohort. *United European Gastroenterol J* 2023; 11: 163-170. 20230213. DOI: 10.1002/ueg2.12360.
84. Peters MLB, Eckel A, Seguin CL, et al. Cost-Effectiveness Analysis of Screening for Pancreatic Cancer Among High-Risk Populations. *JCO Oncol Pract* 2024; 20: 278-290. 20231212. DOI: 10.1200/OP.23.00495.
85. Rubenstein JH, Scheiman JM and Anderson MA. A clinical and economic evaluation of endoscopic ultrasound for patients at risk for familial pancreatic adenocarcinoma. *Pancreatology* 2007; 7: 514-525. 20071001. DOI: 10.1159/000108969.
86. Gupta N, Bansal A, Wani SB, et al. Endoscopy for upper GI cancer screening in the general population: a cost-utility analysis. *Gastrointest Endosc* 2011; 74: 610-624 e612. 20110713. DOI: 10.1016/j.gie.2011.05.001.
87. Lipscomb J, Horton S, Kuo A, et al. Evaluating the impact of multicancer early detection testing on health and economic outcomes: Toward a decision modeling strategy. *Cancer* 2022; 128 Suppl 4: 892-908. DOI: 10.1002/cncr.33980.
88. Tafazzoli A, Ramsey SD, Shaul A, et al. The Potential Value-Based Price of a Multi-Cancer Early Detection Genomic Blood Test to Complement Current Single Cancer Screening in the USA. *Pharmacoeconomics* 2022; 40: 1107-1117. 20220830. DOI: 10.1007/s40273-022-01181-3.
89. Xia C, Xu Y, Li H, et al. Benefits and harms of polygenic risk scores in organised cancer screening programmes: a cost-effectiveness analysis. *Lancet Reg Health West Pac* 2024; 44: 101012. 20240122. DOI: 10.1016/j.lanwpc.2024.101012.
90. Kroep S, Heberle CR, Curtius K, et al. Radiofrequency Ablation of Barrett's Esophagus Reduces Esophageal Adenocarcinoma Incidence and Mortality in a Comparative Modeling

Analysis. *Clin Gastroenterol Hepatol* 2017; 15: 1471-1474. 20170112. DOI: 10.1016/j.cgh.2016.12.034.
91. Cancer Intervention and Surveillance Modeling Network (CISNET). Esophageal Cancer Models – CISNET Model Registry, https://cisnet.cancer.gov/resources/model-registry/esophageal-models/ (accessed 17th Nov 2025).
92. Lewis D, Jimenez L, Mansour MH, et al. A Systematic Review of Cost-Effectiveness Studies on Gastric Cancer Screening. *Cancers (Basel)* 2024; 16 20240627. DOI: 10.3390/cancers16132353.
93. Matsuda A, Saika K, Tanaka R, et al. Simulation Models in Gastric Cancer Screening: A Systematic Review. *Asian Pac J Cancer Prev* 2018; 19: 3321-3334. 20181225. DOI: 10.31557/apjcp.2018.19.12.3321.
94. Matchett CL, Sahinkoc M, Chhatwal J, et al. Cost effectiveness analysis of Barrett's esophagus screening: A systematic review. *Best Practice & Research Clinical Gastroenterology* 2025; 75: 101980. DOI: 10.1016/j.bpg.2025.101980.
95. Matsuda A, Saika K, Tanaka R, et al. Simulation Models in Gastric Cancer Screening: A Systematic Review. *Asian Pacific Journal of Cancer Prevention* 2018; 19: 3321-3334. DOI: 10.31557/APJCP.2018.19.12.3321.
96. Busschaert SL, Van Deynse H, De Ridder M, et al. Patient-level simulation models in cancer care: a systematic review. *Front Public Health* 2025; 13: 1335300. 20250509. DOI: 10.3389/fpubh.2025.1335300.
97. Caro JJ, Briggs AH, Siebert U, et al. Modeling good research practices--overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. *Med Decis Making* 2012; 32: 667-677. DOI: 10.1177/0272989x12454577.
98. Karnon J, Stahl J, Brennan A, et al. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. *Med Decis Making* 2012; 32: 701-711. DOI: 10.1177/0272989x12455462.
99. Breeze PR, Squires H, Ennis K, et al. Guidance on the use of complex systems models for economic evaluations of public health interventions. *Health Econ* 2023; 32: 1603-1625. 20230420. DOI: 10.1002/hec.4681.
100. Enewold L, Parsons H, Zhao L, et al. Updated Overview of the SEER-Medicare Data: Enhanced Content and Applications. *J Natl Cancer Inst Monogr* 2020; 2020: 3-13. DOI: 10.1093/jncimonographs/lgz029.
101. Warren JL, Klabunde CN, Schrag D, et al. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. *Med Care* 2002; 40: Iv-3-18. DOI: 10.1097/01.Mlr.0000020942.47004.03.
102. Badrick E, Renehan I and Renehan AG. Linkage of the UK Clinical Practice Research Datalink with the national cancer registry. *Eur J Epidemiol* 2019; 34: 101-102. 20181005. DOI: 10.1007/s10654-018-0441-5.
103. Canakis A, Pani E, Saumoy M, et al. Decision model analyses of upper endoscopy for gastric cancer screening and preneoplasia surveillance: a systematic review. *Therap Adv Gastroenterol* 2020; 13: 1756284820941662. 20200716. DOI: 10.1177/1756284820941662.
104. Lewis D, Jimenez L, Mansour MH, et al. A Systematic Review of Cost-Effectiveness Studies on Gastric Cancer Screening. *Cancers* 2024; 16: 2353.

105. Wittram R, Kreis L, König HH, et al. Economic evaluations of early detection strategies for pancreatic cancer: a systematic review. *Eur J Health Econ* 2025; 26: 1655-1670. 20250605. DOI: 10.1007/s10198-025-01793-4.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterials.docx