EE684

Technical validation of health economic models – is there a gold standard? Assessment and refinement of current checklist tools.

Anna Tokarz MSc[1], Agota Szende PhD[2], Emilija Veljanoska MSc[3]

[1]Fortrea Market Access Consulting & HEOR, Warsaw, Poland, [2]Fortrea Market Access Consulting & HEOR, Leeds, United Kingdom, [3]Fortrea Market Access Consulting & HEOR, Munich, Germany

Introduction

- Model validation is essential in cost-effectiveness (CE) evaluation of new therapies. While internal, external, and cross-validation are standard approaches to assess credibility, tools for verifying the technical correctness of model calculations remain limited.
- Technical validation refers to the process of ensuring that the model's structure, logic, and calculations function as intended. This includes verifying formulas, algorithms, and data flows within the model, independent of clinical or epidemiological assumptions.
- In the context of Health Technology Agency (HTA), technical and/or clerical errors can undermine the reliability of CE and budget impact (BI) analyses, potentially leading to incorrectly informed decision-making.
- Despite its importance, technical verification is often underreported or inconsistently applied. This may reflect the absence of a universally accepted framework or checklist for conducting and documenting technical validation in health economic modelling.

Methods

- A structured review using PubMed and AI-assisted searches was conducted in May 2025 to identify technical validation checklists applied in economic models.
- Identified checklists were classified into four domains: technical (T) – designed to validate model functionalities and logic, methodological (M) – validating methodological assumptions (i.e.: model structure, population etc.), reporting (R) – advising on proper reporting standards for economic evaluations, and country-specific (C) – assessing alignment with specific HTA requirements. The C, M and R checklists were excluded due to their focus on methodological and reporting aspects.
- Identified T checklists were compared using the CHEC-list, a framework of 19 validation test categories proposed by Evers et al., 2005 selected for its broad coverage of model domains.[1]

Results

- Of 42 identified publications, 12 were excluded as country-specific HTA guidelines, leaving 30 articles for analysis. Of the 30 publications, the majority (n=22) primarily focused on reporting standards or methodological validation and were therefore excluded from detailed comparison. Only the remaining eight publications addressed some of the technical aspects of model development and implementation, and were included in the in-depth analysis.
- Table 1 presents comparison of the 8 identified T (n=1) [8] and M/T (n=7) checklists.

Table 1. Presence of technical validation elements in checklists by methodological area

Criterion	Adams 1992 [2]	Sacristan 1993 [3]	Gonzalez-Perez 2002 [4]	Weinstein 2003 [5]	Evers 2005 [1]	Philips 2006 [6]	Inotai 2012 [7]	Büyükkaramikli 2019* [8]
Description of the target population		X			X		X	
Choice of alternatives	X	X	X		X	X	X	
Economic study question		X	X		X	X	X	
Economic study design	X	X		X	X	X	X	
Follow up period		X		X	X	X	X	
Perspective	X	X		X	X	X	X	
Economic identification	X		X	X	X	X	X	
Economic measurement	X	X	X	X	X	X	X	
Economic valuation	X		X	X	X	X	X	X
Outcome identification	X			X	X	X	X	
Outcome measurement	X	X		X	X	X	X	
Outcome valuation				X	X	X	X	X
Adjustement for time variation	X	X	X	X	X		X	X
Sensitivity and uncertanities	X	X	X		X	X	X	X
Presentation of results		X	X		X		X	X
Generalisability of results					X	X		
Incremental analysis	X	X	X		X		X	
Independence of investigators					X			
Ethics and distribution of effects	X	X			X			

* Most of the tests did not match with selected CHEC-list evaluation criterias.

- In Table 1, an "X" was assigned to a checklist if it included at least one item or test explicitly addressing the technical verification of model calculations under the corresponding criterion.
- Despite applying generous inclusion criteria, the majority of reviewed checklists did not comprehensively address all technical validation domains outlined in the CHEC-list, indicating persistent gaps in current validation practices. The observed heterogeneity in checklist scope poses challenges for generalisability efforts.
- Of the eight checklists evaluated, seven included only limited provisions for technical verification, primarily focusing on basic checks related to calculation accuracy, logical consistency of results, and input estimation. Only the most recent checklist, Büyükkaramikli et al., 2019 [8], was more complete as it included tests for calculation verification, sensitivity, and credibility.
- None of the reviewed verification checklists explicitly addressed critical technical aspects such as data flow integrity, validity of model adaptations, computational efficiency, or the use of volatile functions. These omissions represent significant gaps in current validation practices. Additionally, none of the checklists included targeted validation procedures for Visual Basic for Applications (VBA) scripts or Budget Impact Models (BIMs).
- Table 2 presents a summary of key technical validation gaps that were not addressed by the assessed checklists.

Table 2. Gaps in current technical verification tools

Omitted category	Explanation	Examples of validation tests
Visual Basic for Application	VBA is commonly used to automate calculations and processes in Excel-based health economic models. However, VBA code is often hidden from view and not subject to standard formula auditing tools, making it a potential source of undetected errors or unintended logic, impacting model efficiency.	• Ensure that all VBA codes run without error. • Check if VBA is running efficiently (e.g.: any unused code/part of codes, any not needed loops etc.). • Check if VBA code does not include any hardcoded model inputs/cells.
Adaptations Validity	Health economic models used in HTA submissions are frequently adapted from previously published models or core frameworks. However, without rigorous validation, such adaptations can introduce inconsistencies, misaligned assumptions, or structural discrepancies that compromise the model's accuracy and relevance. These issues may result in divergent outcomes compared to the original model and diminish the model's clinical representativeness, thereby limiting its applicability to real-world patient populations and decision-making contexts.	• Compare outputs from the orginal and adapted models; results should be consistent. • Compare outputs from the original and adapted models using the same inputs. • Do the results match with corresponding trial endpoints? Are there discrepancies? • Do the survival curves estimated (e.g.: from FPNMA) in the model match with statistical analysis output?
Budget Impact Model	Budget impact models are frequently used in HTA submissions to estimate the financial consequences and affordability of adopting a new intervention. Unlike cost-effectiveness models, BIMs often rely on dynamic population inputs, market uptake assumptions, and pricing structures that are prone to implementation errors.	• The same set of logical tests impacting model results as for CEM should be used. • Switch current and new scenarios market shares to each other. • Check if time on treatment is properly included - if duration of treatment is more than 12 months, it is properly distributed between years.
Volatile Functions	Volatile functions in Excel (e.g., RAND(), INDIRECT(), OFFSET()) recalculate every time the worksheet changes, which can lead to inconsistent outputs, unintended variability, and difficulties in reproducing results. Overusage of these functions might result in computing inefficiencies and big model size.	• Are there any volatile functions used (if yes, they should be limited)?
Model logic and conditional logic components	These components are increasingly used in Excel-based health economic models to support the logical representation and structured layout of complex data, such as treatment pathways and resource use. However, they often involve intricate calculations, conditional logic, and automation, which can be difficult to audit and may introduce implementation errors.	• Check if data flow is maintained through the model. • Check if vial sharing/dose optimisation is working as expected. • Is a dosing schedule implemented flexible in the model? • Check if there are no hardcoded values in the engines. • Tornado diagram program to only show meaningful ICERs (from quadrant one).

Conclusions

- Current validation frameworks in health economic modelling predominantly focus on methodological soundness and reporting standards, offering limited assurance regarding the technical correctness of model implementation.
- Technical validation remains an underdeveloped and inconsistently applied domain, despite its critical role in ensuring the reliability of model outputs used in HTA decision-making.
- There is a growing need for structured, scalable, and transparent validation tools that can be integrated into model development workflows, including automated checks, independent audits, and version-controlled documentation to enhance transparency, reproducibility, and stakeholder confidence in model-based evidence.
- Establishing a consensus on best practices for technical validation – potentially through collaboration between HTA bodies, academic institutions, and industry – could enhance the credibility and robustness of health economic evidence.
- Further research is needed to strengthen the methodological basis for technical validation in health economic modelling.

Abbreviations

CE = cost-effectiveness, HTA = Health Technology Agency, T = technical, M = methodological, R = reporting, C = country-specifc, N = number, VBA = Visual Basic for Appliaction, BIM = budget impact model.

References

1. Evers S, et al., Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria, Int J Technol Assess Health Care 2005; 21(2), p. 240-245.
2. Adams ME, et al., Economic analysis in randomized control trials, Med Care 1992; 30(3), p. 231-243.
3. Sacristan JA, et al., Evaluation of pharmacoeconomic studies: utilization of a checklist, Ann Pharmacother 1993; 27(9), p. 1126-1133.
4. Gonzalez-Perez JG, et al., Developing a scoring system to quality assess economic evaluations, Eur J Health Econom 2002; 3(2), p. 131-136.
5. Weinstein MC, et al., Principles of Good Practice for Decision Analytic Modeling in Health-Care Evaluation: Report of the ISPOR Task Force on Good Research Practices – Modeling Studies, Value Health 2003; 6(1), p. 9-17.
6. Philips Z, et al., Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment, Pharmacoeconomics 2006; 24(4), p. 355-371.
7. Inotai A, et al., Attempt to increase the transparency of fourth hurdle implementation in Central-Eastern European middle income countries: publication of the critical appraisal methodology, BMC Health Serv Res 2012; 12, p. 332-339.
8. Büyükkaramikli NC, et al., TECH-VER: A Verification Checklist to Reduce Errors in Models and Improve Their Credibility, Pharmacoeconomics 2019; 37(11), p. 1391-1408.

ISPOR Europe 2025

Fortrea