



Evaluating the Validation Process: Embracing Complexity and Transparency in Health Economic Modelling

Isaac Corro Ramos¹ · Talitha Feenstra^{2,3} · Salah Ghabri⁴ · Maiwenn Al⁵

Accepted: 18 February 2024
© The Author(s) 2024

Reimbursement decisions and price negotiation of healthcare interventions often rely on health economic model results. Such decisions affect resource allocation, patient outcomes and future healthcare choices. To ensure optimal decisions, assessing the validity of health economic models may be crucial. Validation involves much more than identifying (and hopefully correcting) errors in the model implementation. It also includes assessing the conceptual validity of the model and validation of the model input data, and checking whether the model's predictions align sufficiently well with real-world data [1, 2]. In the context of health economics, validation can be defined as “*the act of evaluating whether a model is a proper and sufficient representation of the system it is intended to represent in view of an application*” [3], meaning that the model complies with what is known about the system and its outcomes provide a robust basis for decision making.

In recent years, recognition of the importance of validation as a fundamental step in the modelling process seems to arise among researchers and decision makers [1, 2, 4–6]. Despite this, validation efforts on health economic models remain unreported. A quick PubMed search for “cost effectiveness” and “model” returned 1126 hits, but when

“validation” was added, it dropped to 27 hits (2.4%). This contrasts with searches for “sensitivity analysis” (48%) and “uncertainty” (18%). Recent reviews in cost effectiveness, with a focus on the models used, point out that validation is a missing element in many model-based studies [7, 8]. Even though model developers will probably have validated their models, this lack of reporting might raise questions about the rigor of model validation in daily practice. It is difficult to assess whether health economic model validation is unreported or unperformed [9], given that validation efforts are not reported in detail and that models that were subject to extensive validation (in theory) are still found to have errors [10]. Guidance and reporting guidelines exist, but their implementation is lagging behind.

Terminology in health economic model validation can often be confusing because of different interpretations and a lack of clear definitions. The term “internal validation” can be used to describe the act of comparing model outcomes to empirical data that were used to build the model [11–14]. The same definition is referred to as “dependent validation” elsewhere [15, 16]. However, other studies use “internal validation” to refer to model “verification” [1, 17, 18], or even to double programming [19]. In line with the first use of “internal validation” previously mentioned, the concept of “external validation” requires comparing model outcomes to empirical data that were *not* used to build the model [11–13, 18]. However, the same definition is referred to as “independent validation” as well [15, 16], whereas “independent validation” has also been employed to indicate validation undertaken by a third party [20]. “External validation” has been used for the comparison between model outcomes against outcomes produced by other models [14], but this is often referred to as “cross-validation” [1, 15, 21]. Finally, some publications include a sensitivity analysis as part of model validation [22, 23], although a sensitivity analysis aims to explore uncertainty, not validate models: a model full of errors may still produce robust results. These are just a few examples to highlight that efforts to establish

✉ Isaac Corro Ramos
corroram@imta.eur.nl

¹ Institute for Medical Technology Assessment, Erasmus University Rotterdam, Rotterdam, The Netherlands

² Groningen Research Institute of Pharmacy, Faculty of Science and Engineering, University of Groningen, Groningen, The Netherlands

³ Center for Public Health, Health Services and Society, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

⁴ Department of Medical Evaluation, Direction of Evaluation and Access to Innovation, French National Authority for Health, HAS, Saint-Denis, France

⁵ Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

standardised definitions and guidelines are essential for clarity and consistency in the field.

In addition to this lack of standardisation in the terminology used, the disparity regarding validation requirements among different guidelines around the world does not help model developers when deciding if, and to what extent, validation efforts should be reported. Many guidelines still address the subject of model validation in a brief and unspecific way, or not at all. Exceptions to this are for example the Dutch or the Australian pharmacoeconomic guidelines [6, 24]. We believe that this should be a subject of concern and debate, especially because reporting and good practice guidance is available. The report of the International Society Modeling Good Research Practices Task Force-7 in 2012 established a framework for conducting validation of health economic models [1]. The Assessment of the Validation Status of Health-Economic decision models (AdViSHE) tool was developed in 2016 for the purpose of documenting validation efforts and offered a selection of items to balance feasibility and rigor [15]. We believe these are the minimum validation standards that all model-based studies should meet. Since then, several validation-specific tools have been developed to provide structured approaches to different aspects of model validation [25–28]. Additionally, some studies have addressed the challenge of metrics: how to judge results of validation tests and when it is good enough [29, 30]. As a result, model developers (and users such as health technology assessment agencies) are, in principle, well equipped to conduct validations of health economic models, report and interpret their results. Regarding health economic decision models in general, organisations such as the International Society for Pharmacoeconomics and Outcomes Research and the Society for Medical Decision Making have developed and refined through the years guidelines and reporting standards [1, 31–37]. Many scientific journals in the field of health economics require authors to adhere to specific reporting guidelines, such as the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist [38]. These standards help ensure that health economic evaluations are reported transparently in research articles. Given the available tools, it is strongly recommended to encourage similar standards for model validations. This would entail researchers, public institutions or private health stakeholders to provide sufficiently detailed descriptions of their validation tests, and their results, in publications as well.

Furthermore, notable changes have been occurring in the field of health economic modelling, and data science in general, in recent years [39]. The development of more complex methods (e.g., in the analysis of survival data or evidence synthesis) [40–42], a change in modelling habits (e.g., the use of R in economic modelling) [43–46], the increasing use of open-source models [47], real-world data [48] and

artificial intelligence/machine learning methods are a few examples of these changes [49]. Health economic models are growing in complexity, owing to for example the importance of personalised medicine [50, 51], advanced therapeutic medicinal products [52], modelling of vaccines and immunisation frameworks [53], and the increasing interest for multiple-use models such as whole disease or pathway models [54–61]. Complex models require more extensive validation efforts than straightforward cohort-level state-transition models or decision trees to ensure their accuracy and reliability. The use of open-source modelling software has increased, promoting transparency and collaboration among different stakeholders [44, 47]. However, this does not replace the need for model validation. Integrating real-world data in health economic models has to be combined with adequate validation of model outcomes against such data, for instance by replicating observed data [62]. We believe all these changes, together with a general lack of reporting, call for updated guidance on model validation.

Model validation can play a crucial role in healthcare decision making, at least in theory. Health economic models can be complex, often relying on intricate assumptions and multiple data sources. Ensuring that these models meet high-quality standards is vital to their utility and credibility [2]. This can be achieved by systematically checking all elements of model validation: validating the conceptual model, verifying data sources, testing the plausibility of modelling assumptions, conducting extensive model verification, comparing model outcomes to independent real-world data when available, and addressing potential conflicts of interest when experts are involved in any aspect of model validation. One way to enhance this is through *standardisation*. As part of this, consistent use of terminology is helpful. The development of guidelines and best practices for health economic model validation can help ensure that critical steps are not overlooked and that models consistently adhere to standards. Furthermore, transparency and credibility should be cornerstones of the validation process, making detailed documentation of validation efforts available. More consensus in the field, or clarity by model users, is needed on whether independent third parties should conduct validations, or it is sufficient when model developers clearly report their validation steps and results. The use of external independent validators without conflicts of interest, such as the National Institute for Health and Care Excellence External Assessment Groups, can provide an impartial assessment of the model's strengths and weaknesses, helping to identify potential pitfalls and areas for improvement, enhancing not only the credibility of the model but also improving the transparency of the decision-making process. Nevertheless, such procedures are costly and do not solve all issues raised previously.

Validation of health economic models should be seen as a critical component of evidence-based decision making in

healthcare. However, as of today, it still faces several important challenges, including the lack of consensus guidance and standardised procedures, the need for greater rigour or the question of who should oversee the validation process. To address these challenges, we encourage model developers, agencies requiring models for their decision making and editors of journals that publish models to recommend the use of state-of-the-art tools for reporting (and conducting) validations of health economic models, such as those mentioned in this editorial.

Funding The writing of this editorial was not financially supported.

Data Availability Not applicable.

Declarations

Conflict of interest Isaac Corro Ramos, Talitha Feenstra, Salah Ghabri and Maiwenn Al have no conflicts of interest that are directly relevant to the content of this editorial.

Author Contributions All authors have contributed to the writing of this editorial and given final approval of the version to be published. The opinions expressed in this editorial are those of the authors and do not necessarily represent the views of their institutions.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

References

- Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. *Value Health*. 2012;15(6):843–50.
- Ghabri S, Stevenson M, Möller J, Caro JJ. Trusting the results of model-based economic analyses: is there a pragmatic validation solution? *Pharmacoeconomics*. 2019;37(1):1–6.
- Vemer P, van Voom GAK, Ramos IC, Krabbe PFM, Al MJ, Feenstra TL. Improving model validation in health technology assessment: comments on guidelines of the ISPOR-SMDM modeling good research practices task force. *Value Health J Int Soc Pharmacoecon Outcomes Res*. 2013;16(6):1106–7.
- Caro JJ, Möller J. Decision-analytic models: current methodological challenges. *Pharmacoeconomics*. 2014;32(10):943–50.
- Karnon J. Model validation: has it's time come? *Pharmacoeconomics*. 2016;34(9):829–31.
- Ministerie van Volksgezondheid, Welzijn en Sport. Guideline for economic evaluations in healthcare (2024 version). <https://english.zorginstituutnederland.nl/about-us/publications/reports/2024/01/16/guideline-for-economic-evaluations-in-healthcare> [Accessed 26 Jan 2024].
- Silva-Illanes N, Espinoza M. Critical analysis of Markov models used for the economic evaluation of colorectal cancer screening: a systematic review. *Value Health J Int Soc Pharmacoecon Outcomes Res*. 2018;21(7):858–73.
- Altunkaya J, Lee JS, Tsiachristas A, Waite F, Freeman D, Leal J. Appraisal of patient-level health economic models of severe mental illness: systematic review. *Br J Psychiatry*. 2021;220(2):1–12.
- De Boer PT, Frederix GWJ, Feenstra TL, Vemer P. Unremarked or unperformed? Systematic review on reporting of validation efforts of health economic decision models in seasonal influenza and early breast cancer. *Pharmacoeconomics*. 2016;34(9):833–45.
- Radeva D, Hopkin G, Mossialos E, Borrill J, Osipenko L, Naci H. Assessment of technical errors and validation processes in economic models submitted by the company for NICE technology appraisals. *Int J Technol Health Care*. 2020. <https://doi.org/10.1017/S0266462320000422>.
- Willis M, Asseburg C, Slee A, Nilsson A, Neslusan C. Development and internal validation of a discrete event simulation model of diabetic kidney disease using CREDENCE trial data. *Diabetes Ther*. 2020;11(11):2657–76.
- Favre-Bulle A, Huang M, Haiderali A, Bhadhuri A. Cost-effectiveness of neoadjuvant pembrolizumab plus chemotherapy followed by adjuvant pembrolizumab in patients with high-risk, early-stage, triple-negative breast cancer in Switzerland. *Pharmacoecon Open*. 2024;8(1):91–101.
- Hoerger TJ, Hilscher R, Neuwahl S, Kaufmann MB, Shao H, Laxy M, et al. A new type 2 diabetes microsimulation model to estimate long-term health outcomes, costs, and cost-effectiveness. *Value Health*. 2023;26(9):1372–80.
- Piena MA, Kroep S, Simons C, Fenwick E, Harty GT, Wong SL, et al. An innovative approach to modelling the optimal treatment sequence for patients with relapsing-remitting multiple sclerosis: implementation, validation, and impact of the decision-making approach. *Adv Ther*. 2022;39(2):892–908.
- Vemer P, Corro Ramos I, van Voorn GAK, Al MJ, Feenstra TL. AdViSHE: a validation-assessment tool of health-economic models for decision makers and model users. *Pharmacoeconomics*. 2016;34(4):349–61.
- Antoniou M, Mateus C, Hollingsworth B, Titman A. A systematic review of methodologies used in models of the treatment of diabetes mellitus. *Pharmacoeconomics*. 2024;42(1):19–40.
- Pollock RF, Norrbacka K, Boye KS, Osumili B, Valentine WJ. The PRIME type 2 diabetes model: a novel, patient-level model for estimating long-term clinical and cost outcomes in patients with type 2 diabetes mellitus. *J Med Econ*. 2022;25(1):393–402.
- Petrou S, Gray A. Economic evaluation using decision analytical modelling: design, conduct, analysis, and reporting. *BMJ*. 2011;342: d1766.
- Dong X, He X, Wu J. Cost effectiveness of the first-in-class ARNI (sacubitril/valsartan) for the treatment of essential hypertension in a Chinese setting. *Pharmacoeconomics*. 2022;40(12):1187–205.
- Tamlyn Anne R, Downes M, Simoncini T, Yu Q, Ren M, Wang Y, et al. Evaluating the cost utility of estradiol plus dydrogesterone for the treatment of menopausal women in China. *J Med Econ*. 2024;27(1):16–26.
- Kent S, Becker F, Feenstra T, Tran-Duy A, Schlackow I, Tew M, et al. The challenge of transparency and validation in health economic decision modelling: a view from Mount Hood. *Pharmacoeconomics*. 2019;37(11):1305.

22. Bu DD, Schwam ZG, Wanna GB, Perez E, Cosetti MK. Cost-effectiveness of diffusion-weighted magnetic resonance imaging versus second-look surgery in treating cholesteatoma: a modeling study. *Otol Neurotol*. 2024;45(2):163.
23. Naved N, Umer F, Khowaja AR. Cost-effectiveness analysis of regenerative endodontics versus MTA apexification. *JDR Clin Transl Res*. 2023. <https://doi.org/10.1177/23800844231191515>.
24. Australian Government Department of Health and Aged Care. PBAC guidelines. <https://pbac.pbs.gov.au/information/about-the-guidelines.html> [Accessed 26 Jan 2024].
25. Grimm SE, Pouwels X, Ramaekers BLT, Wijnen B, Knies S, Grutters J, et al. Development and Validation of the TRansparent Uncertainty ASsessmentT (TRUST) tool for assessing uncertainties in health economic decision models. *Pharmacoeconomics*. 2020;38(2):205–16.
26. Tremblay G, Humphries B. A practical guide for clinical and key opinion leader validation of health economic models. *J Med Econ*. 2023;26(1):473–6.
27. Büyükkaramikli NC, Rutten-van Mülken MPMH, Severens JL, Al M. TECH-VER: a verification checklist to reduce errors in models and improve their credibility. *Pharmacoeconomics*. 2019;37(11):1391–408.
28. Pouwels X, Kroeze K, van der Linden N, Kip M, Koffijberg H. MSR35 Systematically validating health economic models using the Probabilistic Analysis Check Dashboard (PACBOARD). *Value Health*. 2022;25(12 Suppl.):S356.
29. Corro Ramos I, van Voorn GAK, Vemer P, Feenstra TL, Al MJ. A new statistical method to determine the degree of validity of health economic model outcomes against empirical data. *Value Health*. 2017;20(8):1041–7.
30. Leunis A, Redekop WK, van Montfort KAGM, Löwenberg B, Uyl-de Groot CA. The development and validation of a decision-analytic model representing the full disease course of acute myeloid leukemia. *Pharmacoeconomics*. 2013;31(7):605–21.
31. Caro JJ, Briggs AH, Siebert U, Kuntz KM, ISPOR-SMDM Modeling Good Research Practices Task Force. Modeling good research practices—overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. *Med Decis Making*. 2012;32(5):667–77.
32. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M, et al. Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-2. *Med Decis Making*. 2012;32(5):678–89.
33. Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. *Value Health*. 2012;15(6):812–20.
34. Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Möller J, et al. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. *Value Health*. 2012;15(6):821–7.
35. Pitman R, Fisman D, Zaric GS, Postma M, Kretzschmar M, Edmunds J, et al. Dynamic transmission modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-5. *Value Health*. 2012;15(6):828–34.
36. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD, et al. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. *Value Health*. 2012;15(6):835–42.
37. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johanneson M, McCabe C, 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):9–17.
38. 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. *BMJ*. 2022;376: e067975.
39. Feenstra T, Corro-Ramos I, Hamerlijnck D, van Voorn G, Ghabri S. Four aspects affecting health economic decision models and their validation. *Pharmacoeconomics*. 2022;40(3):241–8.
40. Palmer S, Borget I, Friede T, Husereau D, Karnon J, Kearns B, et al. A guide to selecting flexible survival models to inform economic evaluations of cancer immunotherapies. *Value Health*. 2023;26(2):185–92.
41. Phillippo DM, Dias S, Ades AE, Belger M, Brnabic A, Schacht A, et al. Multilevel network meta-regression for population-adjusted treatment comparisons. *J R Stat Soc Ser A Stat Soc*. 2020;183(3):1189–210.
42. University of Sheffield. Flexible methods for survival analysis TSD. 2022. <https://www.sheffield.ac.uk/nice-dsu/tsds/flexible-methods-survival-analysis> [Accessed 28 Jan 2024].
43. Alarid-Escudero F, Krijkamp EM, Pechlivanoglou P, Jalal H, Kao SYZ, Yang A, et al. A need for change! A coding framework for improving transparency in decision modeling. *Pharmacoeconomics*. 2019;37(11):1329–39.
44. Jalal H, Pechlivanoglou P, Krijkamp E, Alarid-Escudero F, Enns E, Hunink MGM. An overview of R in health decision sciences. *Med Decis Making*. 2017;37(7):735–46.
45. Alarid-Escudero F, Krijkamp E, Enns EA, Yang A, Hunink MGM, Pechlivanoglou P, et al. A tutorial on time-dependent cohort state-transition models in R using a cost-effectiveness analysis example. *Med Decis Making*. 2023;43(1):21–41.
46. Krijkamp EM, Alarid-Escudero F, Enns EA, Jalal HJ, Hunink MGM, Pechlivanoglou P. Microsimulation modeling for health decision sciences using R: a tutorial. *Med Decis Making*. 2018;38(3):400–22.
47. Pouwels XGLV, Sampson CJ, Arnold RJG, Open Source Models Special Interest Group. Opportunities and barriers to the development and use of Open Source Health economic models: a survey. *Value Health*. 2022;25(4):473–9.
48. Garrison LP, Neumann PJ, Erickson P, Marshall D, Mullins CD. Using real-world data for coverage and payment decisions: the ISPOR Real-World Data Task Force Report. *Value Health*. 2007;10(5):326–35.
49. Padula WV, Kreif N, Vanness DJ, Adamson B, Rueda JD, Felizzi F, et al. Machine learning methods in health economics and outcomes research: the PALISADE checklist: a Good Practices Report of an ISPOR Task Force. *Value Health*. 2022;25(7):1063–80.
50. Rutten-van Mülken M, Versteegh M, Nagy B, Wordsworth S. HEcoPerMed, personalized medicine from a health economic perspective: lessons learned and potential opportunities ahead. *Pers Med*. 2023;20(4):299–303.
51. Vellekoop H, Huygens S, Versteegh M, Szilberhorn L, Zelei T, Nagy B, et al. Guidance for the harmonisation and improvement of economic evaluations of personalised medicine. *Pharmacoeconomics*. 2021;39(7):771–88.
52. Olry de Labry-Lima A, Ponce-Polo A, García-Mochón L, Ortega-Ortega M, Pérez-Troncosa D, Epstein D. Challenges for economic evaluations of advanced therapy medicinal products: a systematic review. *Value Health*. 2023;26(1):138–50.
53. Ultsch B, Damm O, Beutels P, Bilcke J, Brüggjenjürgen B, Gerber-Grote A, et al. Methods for health economic evaluation of vaccines and immunization decision frameworks: a consensus framework from a European vaccine economics community. *Pharmacoeconomics*. 2016;34(3):227–44.

54. Brooke A, Dunning L, Bell E, Watson I, Dawoud D, Bouvy J. P2 NICE's pathways pilot: incorporating disease-specific reference models into health technology assessment in England. *Value Health*. 2023;26(12):S1.
55. NICE. Project information: renal cell carcinoma pathways pilot [ID6186]. <https://www.nice.org.uk/guidance/indevelopment/gid-ta11186> [Accessed 30 Jan 2024].
56. NICE. Taking a proportionate approach to technology appraisals. <https://www.nice.org.uk/about/what-we-do/proportionate-approach-to-technology-appraisals> [Accessed 30 Jan 2024].
57. NICE. Project information: cabozantinib with nivolumab for untreated advanced renal cell carcinoma [ID6184]. <https://www.nice.org.uk/guidance/indevelopment/gid-ta11158> [Accessed 30 Jan 2024].
58. NICE. Project information: treatments for non-small-cell lung cancer [ID6234]; 2024. <https://www.nice.org.uk/guidance/indevelopment/gid-ta11289> [Accessed 30 Jan 2024].
59. Jin H, Tappenden P, Ling X, Robinson S, Byford S. A systematic review of whole disease models for informing healthcare resource allocation decisions. *PLoS One*. 2023;18(9): e0291366.
60. Choon-Quinones M, Zelei T, Németh B, Tóth M, Jia XY, Barnett M, et al. Systematic literature review of health economic models developed for multiple myeloma to support future analyses. *J Med Econ*. 2023;26(1):110–9.
61. Blommestein HM, Verelst SGR, de Groot S, Huijgens PC, Sonneveld P, Uyl-de Groot CA. A cost-effectiveness analysis of real-world treatment for elderly patients with multiple myeloma using a full disease model. *Eur J Haematol*. 2016;96(2):198–208.
62. Li X, Hoogenveen R, El Alili M, Knies S, Wang J, Beulens JWJ, et al. Cost-effectiveness of SGLT2 inhibitors in a real-world population: a MICADO model-based analysis using routine data from a GP registry. *Pharmacoeconomics*. 2023;41(10):1249–62.