IJHPM
International Journal of Health Policy and Management

Commentary

Characterizing the Validity and Real-World Utility of Health Technology Assessments in Healthcare: Future Directions
Comment on "Problems and Promises of Health Technologies: The Role of Early Health Economic Modelling"

Nadine K. Zawadzki[1*] ⓘ, Joel W. Hay[2,1] ⓘ

Abstract
With their article, Grutters et al raise an important question: What do successful health technology assessments (HTAs) look like, and what is their real-world utility in decision-making? While many HTAs are published in peer-reviewed journals, many are considered proprietary and their attributes remain confidential, limiting researchers' ability to answer these questions. Models for economic evaluations like cost-effectiveness analyses (CEAs) synthesize a wide range of evidence, are often statistically and mathematically sophisticated, and require untestable assumptions. As such, there is nearly universal agreement among researchers that enhancing transparency is an important issue in health economic modeling. However, the definition of transparency and guidelines for its implementation vary. Model registration combined with a linked database of model-based economic evaluations has been proposed as a solution, whereby registered models and their accompanying technical and nontechnical documentation are sourced into a single publicly-available repository, ideally in a standardized format to ensure consistent and complete representation of features, code, data sources, results, validation exercises, and policy recommendations. When such a repository is ultimately created, modelers will not have to reinvent the wheel for every new drug launched or new treatment pathway. These more open and transparent approaches will have substantial implications for model accuracy, reliability, and validity, improving trust and acceptance by healthcare decision-makers.
Keywords: Health Technology Assessment, Economic Modeling, Cost-Effectiveness, Transparency

*Correspondence to:
Nadine K. Zawadzki
Email: nzawadzki13@gmail.com

In seeking to understand the returns to early health economic modeling, Grutters et al[1] found that assessments performed in early development led to further development and research, and that none resulted in a firm 'no-go' recommendation. These results suggest that the benefits of early modeling are potentially very high. With their article, Grutters et al raise an important question: What do successful health technology assessments (HTAs) look like, and what is their real-world utility in decision-making? As stated by Mandelblatt et al, "Models are only as good as their ability to represent reality at the level needed to draw useful conclusions; this, in turn depends upon their structure and on the assumptions that go in to the models."[2] The process of determining how 'good' a model is involves model validation (predictive and structural) and peer review.[3] While there are tools for measuring the general quality of published economic models (eg, the Quality of Health Economic Studies instrument),[4] there are no such standard tools for point-by-point evaluation of validity and applicability in peer review. How, then, can we characterize the systematic factors of HTA success and failure?

One could begin by comparing models and their predictions against available data on market and health outcomes. However, this requires access to detailed information on models. While many HTAs such as cost-effectiveness analyses (CEAs) are published in peer-reviewed journals, many are considered proprietary and their attributes remain confidential. Due to confidentiality concerns, Grutters et al were limited in what they could report; only two assessments reviewed were published in scientific journals. As a result, they were unable to identify specific characteristics contributing to decisions favoring further model development or assess the key points of success and failure.

The trade-offs between confidentiality and model transparency in CEA have been discussed in the literature for over a decade. Because CEAs synthesize a wide range of evidence, are often statistically and mathematically sophisticated, and require untestable assumptions, many key decision-makers may view models as proprietary black boxes designed to give biased results.[5] In the United States,

for example, model results are often discounted by skeptical clinicians and policy-makers.[6] While there is nearly universal agreement among researchers that enhancing transparency is an important issue in health economic modeling,[7-9] the definition of transparency and guidelines for its implementation vary, especially with respect to sharing model code.[10] The Second Panel on Cost-effectiveness in Health and Medicine recommends that modelers should provide enough detail about model structure and parameterization to allow reproducibility, from making entire models available to providing versions that allow users to vary selected inputs, but makes no formal recommendation for sharing model code.[7] The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and the Society of Medical Decision Making recommends that transparent models should provide enough information to enable readers, specialists and non-specialists alike, to understand a model's accuracy, limitations, and potential applications.[8] This requires non-technical documentation detailing model structure and potential applications, though not at the level of detail necessary for replication, in addition to technical documentation more substantially detailing model structure, equations, data input and, at the modeler's discretion, access to model code to permit replication.

While some argue that open-source publication of models would compromise intellectual property rights and disincentivize their development,[11] economic models themselves often have minimal proprietary value, and in many situations, publication in a prestigious peer-reviewed journal is worth more than the ability to sell the model privately.[6,12] Several top scientific journals such as *Nature*, *Cell*, and *Science* now require data and/or code sharing with readers upon publication.[13] Since most applied modeling techniques and software packages are standard among researchers in the field, model scripts and code contain little innovative intellectual property.[12] Healthcare organizations are more likely to be concerned with the release of proprietary data from clinical trials. However, when clinical trials are biased or not adequately generalizable, those CEAs that use their results will suffer from these same limitations.

Take, for example, a CEA comparing various strategies to reduce cholesterol involving diet, niacin, and lovastatin, published in Appendix C of *Cost Effectiveness in Health and Medicine* by Gold et al,[14] a landmark instructional book for economic modeling in healthcare. At the time of the book's publication in 1996, clinical trials on niacin showed favorable changes in lipid profiles.[15-18] Although it had not been established in clinical research that niacin's effect on lipid levels resulted in lower rates of cardiovascular events, the authors assumed that such a relationship existed between lipid levels and cardiovascular risk for niacin users. Using the biomarker results, the authors of the CEA concluded that care with niacin was cost-effective for prevention of cardiovascular disease, and in 1997, the US Food and Drug Administration approved a prescription extended-release version of niacin. However, nearly two decades later, the results of two large prospective trials showed that niacin did not improve prevention or reduce mortality.[19,20] Although the original clinical trials

were presumed to show clinical effectiveness, most were inadequately powered and relied on surrogate measures that had not been formally validated. Very few reported information on cardiovascular outcomes, and even fewer were designed to detect changes in these clinical outcomes.[21] Ultimately, the CEA published in Appendix C relied on data that were poor quality and its recommendations were invalidated when better outcome data became available. Since CEAs require synthesis of estimates from various sources, it is important that modelers communicate their assumptions and their reasons for selecting certain data sources over others. Justification for model inputs and parameters is integral to model transparency, especially when data is confidential or otherwise not publicly available. If the authors were required to publish technical documentation detailing the assumed relationship between lipid levels and cardiovascular risk, perhaps a reviewer or reader would have questioned their decision to use an unvalidated surrogate endpoint.

It is possible to build accurate and useful economic models without relying solely on clinical data. Indirect and unintended outcome measures from more pragmatic settings and registries that take into account patient heterogeneity and real-life experiences should also be considered to answer a HTA question.[22] Real-world evidence (RWE) is used by several HTA bodies to confirm or supplement findings from clinical trials,[23] and in certain cases could be used to demonstrate treatment effects when clinical trials are not feasible or unethical, or when there is significant unmet need. Advanced statistical and data science techniques such as Deep IV, a deep neural net machine learning algorithm for counterfactual prediction,[24] and E-value analysis of unmeasured confounding[25] help address bias in observational data and permit causal inferences using RWE. RWE has utility in each stage of product lifecycles, from development to market access and post-launch.[22] Of the assessments reviewed by Grutters et al, 23 were constructed without clinical data in early stages of development, all of which led to further development, research, or both.[1] As with clinical trial data, transparency in analyses of RWE would further strengthen HTA insights gained.

Despite agreement on the importance of transparency in health economic modeling, there is no definitive method for assessing transparency. However, all major guidelines emphasize reproducibility in their recommendations. Some argue that replication studies can effectively demonstrate how transparently a model is reported, identify potential calculation errors and inform future reporting practices.[26] The most recent paper to discuss the concept of a successful replication reported on the 8th Mt. Hood Diabetes Challenge Network: Economics, Simulation Modelling and Diabetes Competition,[27] at which researchers compare their health economic diabetes models in terms of their structure and performance through coordinated tests in which they are asked to reproduce a real-world result.[28] Through the competition's process, the researchers present their models in greater detail than typically provided in publications, and may compare data sources, underlying code, and calibration techniques. In the competition, the replication challenge was

used to indicate reporting transparency, and their proposed definition of replication success was described in terms of model transparency.[27] Recently, the Mt. Hood Diabetes Challenge Network has developed formal guidelines and a checklist, published in *Value in Health,*[27] to further improve transparency in reporting of input data and other information underlying model-based analyses for diabetes.

In 2018, Shao et al participated the 9th Mt. Hood Challenge to demonstrate that their BRAVO microsimulation model of diabetes costs and outcomes,[29] based on the publicly available ACCORD patient-level clinical trial data, performs much better on many, if not all, dimensions than other advanced diabetes models, all of which are based on the much older UKPDS data.[30] In development of BRAVO, Shao et al devised a novel approach to estimate regional multipliers for diabetes models that are constructed using data from a single region, thereby improving prediction accuracy by reducing systematic bias and increasing explanation power.[31] Typical methods for external validation comparing model predictions to outcomes from single-region clinical trials may only weakly assess Shao et al's new approach. However, the Mt. Hood Challenge provided Shao et al a unique opportunity to validate their approach against that of their competitors, permitting them to thoroughly identify inconsistencies and determine their source, whether it be the data and/or model calibration.

Of course, conferences and competitions are not feasible for every systematic evaluation of HTAs. Model registration combined with a linked database of model-based economic evaluations has been proposed as a solution, whereby registered models and their accompanying technical and nontechnical documentation are sourced into a single publicly-available repository, ideally in a standardized format to ensure consistent and complete representation of features, code, data sources, results, validation exercises, and policy recommendations.[32] Such a registry would aid researchers and decision-makers who are interested in comparing HTAs, their characteristics, and their relative performance and accuracy to examine and factor all available information into their evaluations. A registry would additionally facilitate collaboration between researchers while avoiding duplicating efforts to answer certain policy questions or validate existing models. ISPOR has already convened a Special Interest Group that is currently undertaking development of such an open-source platform. In the last year, it has developed tools for the ISPOR website, Short Courses, manuscripts in draft for scientific journals, webinars, and presentations in order to promote open-source modeling. The Mount Hood Diabetes Challenge Network has also created an initial registry for diabetes models in response to recent calls for a registry of economic models, located on their website.[33] While their registry does not require model code to be made available and is currently limited to non-technical documentation, with outside links to technical descriptions and published reports, the Network is currently establishing a structured form to house technical and non-technical documentation of contributing models.[32] When a repository is ultimately created, modelers will not have to reinvent the wheel for every new drug launched or new treatment pathway.

Grutters et al were disinclined to report the point estimate results of assessments in terms of expected costs, effects, or cost-effectiveness without proper explanation of the context of analysis and quality of underlying evidence. An open-source repository of models would make this information publicly available through accompanying documentation, permitting Grutters et al to critically examine these results in their study. Assuming that technical documentation would also include probabilistic and scenario analyses beyond that typically reported in published manuscripts, Grutters et al could have additionally commented on existing uncertainty surrounding care. Although Grutters et al have presented many important insights, they caution that their recommendations cannot be interpreted independently as they are strongly related to both the value proposition and positioning of products. If a model repository were available, Grutters et al could have compared their included HTAs to those performed for competitor products at various stages of development, supplementing their sample so that it no longer represents only those innovations for which an early health economic modeling study was commissioned, thereby further informing value, positioning, and commercial viability. These more open and transparent approaches will have substantial implications for model accuracy, reliability, and validity, improving trust and acceptance by healthcare decision-makers. As Grutters et al aimed to demonstrate, evaluating the real-world performance of such models would promote more efficient and effective healthcare HTA.

Authors' affiliations
¹Schaeffer Center for Health Policy and Economics, Department of Pharmaceutical and Health Economics, School of Pharmacy, University of Southern California, Los Angeles, CA, USA. ²USC Clinical Economics Research and Education Program (CEREP), Los Angeles, CA, USA.

References
1. Grutters JPC, Govers T, Nijboer J, Tummers M, van der Wilt GJ, Rovers MM. Problems and promises of health technologies: the role of early health economic modeling. *Int J Health Policy Manag.* 2019;8(10):575-582. doi:10.15171/ijhpm.2019.36
2. Mandelblatt JS, Fryback DG, Weinstein MC, Russell LB, Gold MR. Assessing the effectiveness of health interventions for cost-effectiveness analysis. In: Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. *Cost-Effectiveness Analysis in Health and Medicine*. New York (NY): Oxford University Press; 1996:135-164.
3. McCabe C, Dixon S. Testing the validity of cost-effectiveness models. *Pharmacoeconomics.* 2000;17(5):501-513. doi:10.2165/00019053-200017050-00007
4. Ofman JJ, Sullivan SD, Neumann PJ, et al. Examining the value and quality of health economic analyses: implications of utilizing the QHES. *J Manag Care Pharm.* 2003;9(1):53-61. doi:10.18553/jmcp.2003.9.1.53
5. Goodacre S, McCabe C. Being economical with the truth: how to make your idea appear cost effective. *Emerg Med J.* 2002;19(4):301-304. doi:10.1136/emj.19.4.301

6. Hay JW. Now Is the Time for Transparency in Value-Based Healthcare Decision Modeling. *Value Health*. 2019;22(5):564-569. doi:10.1016/j.jval.2019.01.009

7. Neumann PJ, Kim DD, Trikalinos TA, et al. Future directions for cost-effectiveness analyses in health and medicine. *Med Decis Making*. 2018;38(7):767-777. doi:10.1177/0272989x18798833

8. 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. *Med Decis Making*. 2012;32(5):733-743. doi:10.1177/0272989x12454579

9. Husereau D, Drummond M, Petrou S, et al. Consolidated health economic evaluation reporting standards (CHEERS)--explanation and elaboration: a report of the ISPOR health economic evaluation publication guidelines good reporting practices task force. *Value Health*. 2013;16(2):231-250. doi:10.1016/j.jval.2013.02.002

10. Cohen AB. Point-counterpoint: cost-effectiveness analysis in medical care and the issue of economic model transparency. *Med Care*. 2017;55(11):907-908. doi:10.1097/mlr.0000000000000812

11. Padula WV, McQueen RB, Pronovost PJ. Can economic model transparency improve provider interpretation of cost-effectiveness analysis? evaluating tradeoffs presented by the second panel on cost-effectiveness in health and medicine. *Med Care*. 2017;55(11):909-911. doi:10.1097/mlr.0000000000000810

12. Cohen JT, Wong JB. Can economic model transparency improve provider interpretation of cost-effectiveness analysis? a response. *Med Care*. 2017;55(11):912-914. doi:10.1097/mlr.0000000000000811

13. Morin A, Urban J, Adams PD, et al. Research priorities. Shining light into black boxes. *Science*. 2012;336(6078):159-160. doi:10.1126/science.1218263

14. Shepard DS. Cost-effectiveness in Health and Medicine. By M.R. Gold, J.E Siegel, L.B. Russell, and M.C. Weinstein (eds). New York: Oxford University Press, 1996. *J Ment Health Policy Econ*. 1999;2(2):91-92. doi:10.1002/(sici)1099-176x(199906)2:2<91::aid-mhp46>3.0.co;2-i

15. Kane JP, Malloy MJ, Ports TA, Phillips NR, Diehl JC, Havel RJ. Regression of coronary atherosclerosis during treatment of familial hypercholesterolemia with combined drug regimens. *JAMA*. 1990;264(23):3007-3012.

16. Caruzzo C, Liboni W, Bonzano A, et al. Effect of lipid-lowering treatment on progression of atherosclerotic lesions--a duplex ultrasonographic investigation. *Angiology*. 1995;46(4):269-280. doi:10.1177/000331979504600401

17. Guyton JR, Brown BG, Fazio S, Polis A, Tomassini JE, Tershakovec AM. Lipid-altering efficacy and safety of ezetimibe/simvastatin coadministered with extended-release niacin in patients with type IIa or type IIb hyperlipidemia. *J Am Coll Cardiol*. 2008;51(16):1564-1572. doi:10.1016/j.jacc.2008.03.003

18. Sang ZC, Wang F, Zhou Q, et al. Combined use of extended-release niacin and atorvastatin: safety and effects on lipid modification. *Chin Med J (Engl)*. 2009;122(14):1615-1620.

19. Boden WE, Probstfield JL, Anderson T, et al. Niacin in patients with low HDL cholesterol levels receiving intensive statin therapy. *N Engl J Med*. 2011;365(24):2255-2267. doi:10.1056/NEJMoa1107579

20. Landray MJ, Haynes R, Hopewell JC, et al. Effects of extended-release niacin with laropiprant in high-risk patients. *N Engl J Med*. 2014;371(3):203-212. doi:10.1056/NEJMoa1300955

21. D'Andrea E, Hey SP, Ramirez CL, Kesselheim AS. Assessment of the role of niacin in managing cardiovascular disease outcomes: a systematic review and meta-analysis. *JAMA Netw Open*. 2019;2(4):e192224. doi:10.1001/jamanetworkopen.2019.2224

22. Oortwijn W, Sampietro-Colom L, Trowman R. How to deal with the inevitable: generating real-world data and using real-world evidence for HTA purposes - from theory to action. *Int J Technol Assess Health Care*. 2019;35(4):346-350. doi:10.1017/s0266462319000400

23. Makady A, van Veelen A, Jonsson P, et al. Using real-world data in health technology assessment (HTA) practice: a comparative study of five HTA agencies. *Pharmacoeconomics*. 2018;36(3):359-368. doi:10.1007/s40273-017-0596-z

24. Hartford J, Lewis G, Leyton-Brown K, Taddy M. Deep IV: A Flexible Approach for Counterfactual Prediction. Proceedings of the 34th International Conference on Machine Learning; 2017.

25. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. *Ann Intern Med*. 2017;167(4):268-274. doi:10.7326/m16-2607

26. McManus E, Turner D, Sach T. Can you repeat that? exploring the definition of a successful model replication in health economics. *Pharmacoeconomics*. 2019;37(11):1371-1381. doi:10.1007/s40273-019-00836-y

27. Palmer AJ, Si L, Tew M, et al. Computer modeling of diabetes and its transparency: a report on the eighth mount hood challenge. *Value Health*. 2018;21(6):724-731. doi:10.1016/j.jval.2018.02.002

28. The Mt Hood Diabetes Challenge. https://www.mthooddiabeteschallenge.com. Accessed November 26, 2019.

29. Shao H, Fonseca V, Stoecker C, Liu S, Shi L. Novel risk engine for diabetes progression and mortality in USA: building, relating, assessing, and validating outcomes (BRAVO). *Pharmacoeconomics*. 2018;36(9):1125-1134. doi:10.1007/s40273-018-0662-1

30. BRAVO Model. https://www.mthooddiabeteschallenge.com/bravo. Accessed November 26, 2019.

31. Shao H, Yang S, Stoecker C, Fonseca V, Hong D, Shi L. Addressing regional differences in diabetes progression: global calibration for diabetes simulation model. *Value Health*. 2019. doi:10.1016/j.jval.2019.08.007

32. Kent S, Becker F, Feenstra T, et al. The challenge of transparency and validation in health economic decision modelling: a view from Mount Hood. *Pharmacoeconomics*. 2019;37(11):1305-1312. doi:10.1007/s40273-019-00825-1

33. Diabetes simulation modeling database. https://www.mthooddiabeteschallenge.com/registry. Accessed November 26, 2019.