Transparency in Health Economic Modeling: Options, Issues and Potential Solutions

Eric Wu[1], Zheng-Yi Zhou[2], Jipan Xie[3], Cinzia Metallo[1], Praveen Thokala[4]

[1] Analysis Group, Boston, 111 Huntington Avenue, Boston, MA 02199, USA
[2] Analysis Group, Angel Court, London, England EC2R7HJ, United Kingdom
[3] Analysis Group, 333 South Hope Street, Los Angeles, CA 90071, USA
[4] University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA, United Kingdom

Article Type: Practical Application

Reference Number: PECA-D-18-00366

Corresponding Author:

Eric Q. Wu
111 Huntington Avenue, 14th Floor
Boston, MA 02199
USA

Email: eric.wu@analysisgroup.com
Phone: 617 425 8254

Abstract

Economic models are being increasingly used by health economists to assess the value of health technologies and inform healthcare decision-making. However, most published economic models represent a kind of black box, with known inputs and outputs but undisclosed internal calculations and assumptions. This lack of transparency makes the evaluation of the model results challenging, complicates comparisons between models, and limits the reproducibility of the models. Here, we aim to provide an overview of the possible steps that could be undertaken to make economic models more transparent and encourage model developers to share more detailed calculations and assumptions with their peers. Scenarios with different levels of transparency (i.e. how much information is disclosed) and reach of transparency (i.e. who has access to the disclosed information) are discussed and five key concerns (copyrights, model misuse, confidential data, software, and time/resources) pertaining to model transparency are presented along with possible solutions. While a shift toward open-source models is underway in health economics, as it happened before in other research fields, the challenges ahead should not be underestimated. Importantly, there is a pressing need to find an acceptable tradeoff between the added value of model transparency and the time and resources needed to achieve such transparency. To this end, it will be crucial to set incentives at different stakeholder levels. Despite the many challenges, the many benefits of publicly sharing economic models make increased transparency a goal worth pursuing.

KEY POINTS

- Health economics models are often challenging to understand, interpret, and reproduce as they lack or report limited information regarding underlying calculations and assumptions

- We provide an overview of the ongoing debate on economic model transparency as well as discuss the key challenges facing the implementation of fully transparent models along with possible options and solutions

- A concerted effort involving the research community, modelers, academics, and scholarly journals is needed to improve the level of transparency (i.e. how much information is disclosed) and reach of transparency (i.e. who has access to the disclosed information) of economic models

1. Introduction

Economic models are being increasingly used in conjunction with efficacy and safety data to assess the value of health technologies [1, 2]. However, published economic models generally lack transparency, with transparency broadly defined as full accessibility to any information related to how the model was built, and thus full reproducibility [3]. Indeed, economic models are typically described in terms of key assumptions, inputs, data sources, and outputs but the internal components and assumptions are often not detailed enough to allow other researchers to reproduce the model or fully interpret the results [1, 4]. In particular, the underlying code or calculations are rarely, if ever, shared. As a result, economic models represent a kind of black box with known inputs and outputs but undisclosed internal structure and calculations. To assess the robustness and validity of a model, sensitivity analyses are typically conducted to evaluate whether, and to which extent, alternative assumptions may alter the model results. Nonetheless, sensitivity analyses do not completely solve the issue of limited transparency as authors may consider only the most convenient scenarios [4].

The lack of detailed descriptions of internal model structures and equations complicates comparisons and makes the evaluation of the model results challenging [4, 5]. More transparency would assuage concerns that payers and other stakeholders have expressed regarding the reliability of economic models as well as allay the perception that commercially sponsored models are inherently biased [5]. Making a more detailed description of all the steps used to build an economic model publicly available would allow researchers to more easily validate the model results, compare the model with other available models, update the model inputs based on the most recent evidence, adapt the model to local health systems or different therapeutic areas, and reduce duplication of efforts [6].

To further the discussion on model transparency and how to achieve it, we aim to provide an overview of the ongoing debate on model transparency as well as discuss the key challenges facing the implementation of fully transparent models along with possible options and solutions. The paper is structured as follows: Section 2 presents the different options for model transparency; Section 3 presents the key issues with model transparency and Section 4 presents the potential solutions.

2. Options for model transparency

While there is a general consensus that the way economic models are reported should be improved, how to achieve this goal is the subject of intense debate. In particular, the pros and cons of making published economic models openly available to researchers and healthcare stakeholders are being increasingly discussed in the literature [2, 4-7]. Generally speaking, possible scenarios for model transparency rests on

finding an appropriate solution to two main issues (**Figure 1**): level of transparency (i.e. how much information is disclosed) and reach of transparency (i.e. who has access to the disclosed information). The current level of transparency of published economic models can be considered low as the model internal structure and computations are not fully disclosed, most often neither to the journal nor to the readers. To increase the level of transparency, a graphical user interface (GUI) could be set up with a full list of model assumptions; data inputs and model settings so that users can interact with the model by selecting the inputs that are of interest to them and visualizing the outputs [16]. A GUI-based solution that allows users to perform additional analyses other than the ones reported in the original publication would further the users' understanding of the model structure and computations [16]. Given that, in this case, the code or internal calculations would not be visible, this would be considered a medium level of transparency. Full transparency would be achieved only by releasing every single step and process involved in building the model, including the source code.

One of the most touted solutions to the problem of model transparency is that of open source models, suggested by health economists [2, 4, 6, 17] and the ISPOR-SMDM Modeling Good Research Practices Task Force [18, 19]. In the open-source framework, detailed descriptions of the model inputs, outputs, structure, and outcomes are made available to the public along with all the codes and calculations used to build the model, either in a repository or upon request [2]. In particular, releasing the source code is believed by many to be the best possible way to ensure that the model is reproducible and its assumptions valid [6, 20]. Some modelers have already chosen the full open-source route and have made their source code, developed either in R or Excel, downloadable. Examples include the model developed by Sullivan et al. [20] in pain therapy, the IVI Rheumatoid Arthritis (IVI RA) model [12], the IVI NSCLC model in non-small cell lung cancer[13], and the models found in the Global Health Cost Effectiveness Analysis (GHCEA) registry [14] and in the University of Exeter's repository [15]. However, it should be noted that these examples represent the exception rather than the norm.

Besides the level of transparency, the issue of who has access to the model information should also be addressed. Indeed, the modeler could share the full model only with the party that commissioned the model such as funding agency or company (low reach of transparency), only with the health technology assessment (HTA) agencies (e.g. NICE) to whom the model is submitted and/or any other relevant stakeholders such as academic groups (medium reach of transparency), or with the public at large, either depositing the model in a repository or making it available upon request (high reach of transparency).

The scenarios with a high level and reach of transparency (**Figure 1**), commonly associated with open source models, carry some risks that may result in unintended consequences. The potential issues and drawbacks associated with increased model transparency need to be carefully evaluated before making

any determinations. To this end, five main concerns (copyrights, model misuse, confidential data, software, and time/resources) pertaining to model transparency are discussed below, along with possible solutions.

3. Issues with model transparency

Currently, when a model-based manuscript is published in a journal, the authors transfer to the journal the copyrights of the content of the manuscript but not those of the source code, which can therefore be reused by the authors at will. Whether, and the extent to which, the source code should be released is currently a sticking point in the ongoing debate on model transparency. Some researchers like Padula et al. [5, 7] have argued that publication of the source code would compromise the copyrights of the modelers and, as a result, would discourage the creation of economic models. This could be true especially for complex models. Others like Cohen at al. [4] have argued that publication of the full source code would not affect the 'intellectual value' of the model as the most valuable part of the model rests on the internal assumptions — which, theoretically, should already be reported in the methods section of a manuscript — not on the lines of code, which are simply used to implement the model assumptions. The scientific credibility of a model, they reckon, depends on its reproducibility, which in turn depends on the availability of the source code. How to reconcile these opposing views and create sharing standards that take into account a modeler's copyrights?

Besides copyrights, once the underlying lines of code are made publicly available, a system that guarantees that each published model is appropriately cited when fully or partially used to build a new model should be in place. How can such a system be built?

Another concern regarding open-source models is that people not trained in, or with limited knowledge of, economic models may use an openly available model and generate misleading results regarding the value of different treatment options [5]. Given the potential harm to patients, this issue should not be dismissed and safeguards should be put into place to minimize the risk of improper or unethical use of open-source models. Cohen et al. [4] argued that making models open source would protect against misuse in itself as researchers would eventually uncover any improper use or alteration of a model and report it. However, it should also be noted that economic models are often complex, depending on the therapeutic area, data availability, or treatment regimens. The more complex the model, the easier it is to misuse or misinterpret it as only a few parties have the skills, experience, and resources to fully understand and critically review it [4, 6]. What measures should be implemented to avoid model misuse and dissemination of misleading results?

The use of confidential information that is often needed to build an economic model (e.g. individual patient data) should also be weighted when considering the open-source path. Journals have specific guidelines and requirements in place for when confidential data are used in a publication, whether model-based or not. Precautions should be taken when publicly sharing economic models as well. How to reconcile the use of confidential information with the need for more transparency?

Another important issue that needs to be addressed when discussing open-source models is that of proprietary software. To be usable, open-source models require software that is freely or widely available and easy to share. Currently, most economic models are developed in Excel (e.g. Snowsill et al. [15]), which is part of Microsoft Office, or R (e.g. Sullivan et al. [20] and IVI RA Model), which is freely available, but proprietary software (such as TreeAge, Simul8, and Arena) could also be used. Some vendors make the source code available to customers, who can then modify it but, typically, cannot redistribute it. How to reconcile the use of proprietary software with the need for more publicly available and accessible models?

The time and resources needed to make models more transparent and accessible should not be overlooked, particularly because most of the onus will rest with the researchers creating the models. Copies will need to be created and uploaded to repositories or shared with other researchers upon request. In some cases, GUIs will need to be created as well, as mentioned above. Countermeasures to avoid accidentally disclosing confidential information may also need to be implemented. Training modelers on what constitutes confidential, and thus not disclosable, information may be a first step in the right direction. Built-in countermeasures embedded in the code itself may be another option. Furthermore, since an open model should be readily available but also easily interpretable, modelers may need to spend time adding comments and annotations, based on common standards or repositories' policies, that they may otherwise not include in the code. Technical support for researchers using models made publicly available may also be required to clarify a model's structure or assumptions on an ongoing basis. Given the amount of work required to make models publicly available, how to motivate researchers to share their models?

4. Possible solutions

To mitigate the problem of possible infringement of a modeler's copyrights when publicly disclosing the source code, several measures have been proposed. One option would be to request that anybody (whether an individual researcher or a pharmaceutical company) wanting to reuse the model to ask for permission and/or pay a licensing fee [5]. When publishing the results of a study that used a model

developed by other researchers, the model creators should be acknowledged and proof of permission to publish or payment of the licensing fee should be provided to the journal. To facilitate the sharing of the models, including the source code, without compromising the copyrights of the modelers, a central repository, or several repositories based on the type of economic model, could be created. As proposed by Padula et al. [5], the use of unique digital IDs associated with the models deposited in a repository would ensure that modelers, or the repository itself, can track who has downloaded their model(s) and make sure that the terms of use are respected and the licensing fee paid. Assigning digital IDs to individual economic models could lay out the basis for a citation system similar to the DOI (Digital Object Identifier) system currently used for published articles, further motivating researchers to share their models as they would become part of their publication list and academic curriculum. Such system may also motivate academic institutions to relax their copyright and licensing policies, which tend to be very strict, given the potential to further their reputation by increasing the number of citations attributed to their researchers. As suggested by Bierer et al. [21], recognition of data authorship (in this case, recognition of model authorship) via digital IDs may act as an incentive for data/model sharing and foster collaborations among researchers, further advancing the field of health economics. The growing use of repositories in other research fields suggests that this solution is not farfetched. For instance, cognitive neuroscientists have long shared their computational models in repositories like ModelDB [22], which contains over 1000 public models [23]. GitHub [24] is another popular repository across research areas, where users can choose whether to retain or forego their copyrights. Nevertheless, it should be noted that copyrights laws vary across countries, adding an extra layer of complexity to the problem [25]. For instance, Creative Commons licenses – which enable the free distribution of a copyrighted work without requiring authors to relinquish their copyrights – are not uniformly recognized worldwide and enforcing them has been problematic thus far [26].

In addition to repositories, sharing standards and regulations could be mandated by funding agencies. Currently, the National Institutes of Health (NIH) and the National Science Foundation (NSF) have specific model sharing requirements and mandates for certain types of grants and funding initiatives [22]; other agencies could follow their example. Journals would play a key role in mandating the use of repositories. Some journals (e.g. Plos ONE, Science, and the Journal of Biological Chemistry) already require models be shared when they are a key to obtaining the published results, but the vast majority of journals only encourage this practice. In the clinical realm, ClinicalTrials.gov [27] is a good example of how mandating the registration of privately or publicly funded trials as condition for publication in peer-reviewed journals can increase the popularity of a database. Recognizing the issue of copyrights infringement when making a model fully open source, the ISPOR-SMDM joint task force suggested making all the technical details of a model, including the source code, available to peer reviewers after a

study is submitted for possible publication to a journal given that reviewers are bound to confidentiality [18]. A similar proposal has been advanced by the International Committee of Medical Journal Editors [28]. However, many journals may not be well equipped to take on the task of mandating full disclosure of the source code and some restructuring within the publishing industry may need to take place for that to happen. One issue that will need to be addressed, for instance, is that peer reviewers may not be inclined to spend time and effort reviewing a model's calculations or lines of code unless appropriately motivated, monetarily or otherwise.

To minimize the risk of model misuse, more training in health economics and economic modeling may be required for all healthcare stakeholders, including researchers and payers. To curtail the dissemination of deceitful or inaccurate results more health economics training among health care professionals may also be needed. Padula et al. [5, 7] advocated for a curriculum reform for medical students given that they represent some of the future healthcare professionals that could benefit the most from a better understanding of economic models. We agree that introducing medical students to the basic concepts of economic modeling, particularly cost-effectiveness analysis, would help translate model results into practice and prevent the spread of misinformation based on the improper use of openly available models. However, we recognize that the medical school curriculum is already burdensome and adding extra courses or modules may be challenging. It would be important for health economists to work with the American Association of Medical Colleges (AAMC) to find possible avenues of collaboration and identify possible training opportunities during or after medical school, as also proposed elsewhere [5]. In addition to training healthcare professionals on the principles of health economics, safeguards should be embedded in the dissemination of open models through repositories, journals, or funding and HTA agencies, as mentioned above. The reference models proposed by Afzali and Karnon [29] may also contribute to reducing model manipulation and misuse by standardizing models by therapeutic area.

Besides model misuse, accidental disclosure of the confidential data that often populate an economic model is a major concern for which a solution that does not compromise the reliability of the model itself is needed. One option would be to use dummy data so that other researchers can reuse the model without having access to the original dataset. Another option would be to add a random error to the confidential data; however, the error may mask a feature of interest. Synthetic, but statistically identical, data could also be used as done for the Synthetic Longitudinal Business Database (SynLBD) based on US census data [30]. Depending on the type of data, standards should be established to safeguard the confidentiality of the data researchers use in their models.

Regarding the issue of proprietary software, when submitting a model to a journal or an HTA agency, a

possible solution would be to engage only reviewers with knowledge of that particular software, though this may slow down the review process as fewer reviewers may be available. Some commercial software (e.g. TreeAge) allows users to convert models to Excel, facilitating model review. Models created with proprietary software are not necessarily flawed or inaccurate, therefore researchers should not be penalized for using it. A delicate balance based on a modeler's available resources and modeling needs should be struck between the right of modelers to use proprietary software and the necessity of more transparency in the field of health economics.

As all the issues discussed above indicate, making a model transparent and accessible to other researchers is a task that requires time and effort. As such, researchers making their models publicly available should be acknowledged and properly credited for providing a useful service to the whole health economics community. However, incentives are needed to improve the transparency and sharing of economic models. Compensating modelers for their time and effort in making models more transparent, for instance in the form of licensing fees, could be a powerful financial incentive. In addition, regulatory and reimbursement agencies could be crucial in providing incentives and motivating researchers to contribute to model transparency. Currently, most HTA agencies (e.g. NICE, CADTH, PBAC) require the model in full, including the source code. Sometimes agencies review the whole model or build their own scenario tests. For example, NICE typically delegates the assessment of a model (and the source code) submitted by a pharmaceutical company to an independent evidence review group. At the end of the review process, only a report, not the full model, is shared with the public. Should a few HTA agencies start to require the full model be made public after appraisal, other may follow and the practice may easily expand to models published in peer review journals. However, this change would require substantial modifications to the overall HTA review process and may receive pushback from pharmaceutical companies and other healthcare stakeholders, particularly if the concerns discussed above are not adequately addressed.

5. Conclusions

A shift toward open-source models is underway in health economics, as it happened before in other research fields such as genomics, protein structure, and cognitive neuroscience. Nonetheless, several key issues still need to be addressed and the challenges ahead should not be underestimated. An important first step would be to find a reasonable solution to the issue of possible infringement of a modeler's copyrights. The potential misuse of publicly shared models and the use of confidential information and proprietary software should also be weighted. Whenever possible, model complexity should be reduced to facilitate the interpretation and dissemination of model results, making sure that a model is as simple as it

can be while representing the disease or phenomenon of interest appropriately – but without redundant complexity. Importantly, there is a pressing need to find an acceptable tradeoff between the added value of model transparency and the time and resources needed to achieve such transparency. Last but not least, besides implementing sharing standards and policies, it will be critical to induce a cultural shift in the health economics community. Despite the challenges, the many benefits of publicly sharing economic models make increased transparency a goal worth pursuing.

Acknowledgments

All authors have contributed to manuscript ideation and development.

Conflict of Interests

No funding was received for the development of this review article. The authors declare no conflicts of interest. However, it is noted that Eric Wu, Zheng-Yi Zhou, and Jipan Xie are employees of Analysis Group, a strategy consulting firm, where they advise clients on matters related to Health Economics and Outcomes Research (HEOR), including economic models. At the time of manuscript development, Cinzia Metallo was also an employee of Analysis Group. Praveen Thokala is a Senior Research Fellow at the University of Sheffield; his research focuses on health economic modeling.

References

1. Bell CF. The pursuit of transparency and quality improvement in cost-effectiveness analysis - a case study in disease-modifying drugs for the treatment of multiple sclerosis. Journal of managed care pharmacy : JMCP. 2011 Jul-Aug;17(6):463-8.
2. Dunlop WCN, Mason N, Kenworthy J, Akehurst RL. Benefits, Challenges and Potential Strategies of Open Source Health Economic Models. PharmacoEconomics. 2017;35(1):125-8.
3. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB, et al. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--7. Value Health. 2012 Sep-Oct;15(6):843-50.
4. Cohen JT, Wong JB. Can Economic Model Transparency Improve Provider Interpretation of Cost-Effectiveness Analysis? A Response. Medical care. 2017 Nov;55(11):912-4.
5. Padula WV, McQueen RB, Pronovost PJ. Finding Resolution for the Responsible Transparency of Economic Models in Health and Medicine. Medical care. 2017 Nov;55(11):915-7.
6. Cohen JT, Neumann PJ, Wong JB. A Call for Open-Source Cost-Effectiveness Analysis. Annals of internal medicine. 2017 Sep 19;167(6):432-3.
7. 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. Medical care. 2017 Nov;55(11):909-11.
8. !!! INVALID CITATION !!! [2, 4, 6, 8].
9. !!! INVALID CITATION !!! [9, 10].
10. !!! INVALID CITATION !!! [6, 11].
11. !!! INVALID CITATION !!! [11].
12. Incerti D, Curtis JR, Shafrin J, Lakdawalla DN, Jansen JP. A Flexible Open-Source Decision Model for Value Assessment of Biologic Treatment for Rheumatoid Arthritis. Pharmacoeconomics. 2019 Jun;37(6):829-43.
13. Incerti D, Jansen JA. A Description of the IVI-NSCLC Model v1.0; last updated January 31, 2019. [cited 2019 August 1, 2019]; Available from: https://innovationvalueinitiative.github.io/IVI-NSCLC/modeldoc/model-doc.pdf
14. Global Health Cost Effectiveness Analysis (GHCEA) Registry. Open-Source Model Clearinghouse. 2019 February 2019]; Available from: http://healtheconomics.tuftsmedicalcenter.org/orchard/open-source-model-clearinghouse
15. University of Exeter. Immunosuppressive therapy for kidney transplantation in adults: Economic model (dataset). 2019 January 2019]; Available from: https://ore.exeter.ac.uk/repository/handle/10871/25885
16. Dragicevic P, Jansen Y, Sarma A, Kay M, Chevalier F. Increasing the Transparency of Research Papers with Explorable Multiverse Analyses. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. Glasgow, Scotland Uk: ACM; 2019. p. 1-15.
17. Poole C, Agrawal S, Currie CJ. Let cost effectiveness models be open to scrutiny. BMJ (Clinical research ed). 2007 Oct 13;335(7623):735.
18. 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. Medical decision making : an international journal of the Society for Medical Decision Making. 2012 Sep-Oct;32(5):733-43.
19. Caro JJ, Briggs AH, Siebert U, Kuntz KM. Modeling good research practices--overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. Medical decision making : an international journal of the Society for Medical Decision Making. 2012 Sep-Oct;32(5):667-77.

20. Sullivan W, Hirst M, Beard S, Gladwell D, Fagnani F, Lopez Bastida J, et al. Economic evaluation in chronic pain: a systematic review and de novo flexible economic model. The European journal of health economics : HEPAC : health economics in prevention and care. 2016 Jul;17(6):755-70.
21. Bierer BE, Crosas M, Pierce HH. Data Authorship as an Incentive to Data Sharing. The New England journal of medicine. 2017 Apr 27;376(17):1684-7.
22. McDougal RA, Bulanova AS, Lytton WW. Reproducibility in Computational Neuroscience Models and Simulations. IEEE transactions on bio-medical engineering. 2016;63(10):2021-35.
23. SenseLab. ModelDB. 2019 February 2019]; Available from: https://senselab.med.yale.edu/modeldb/
24. GitHub Inc. GitHub 2019 February 2019]; Available from: https://github.com/
25. Introduction of International Copyright Law. [cited 2019 August 1, 2019]; Available from: https://www.copyrightlaws.com/introduction-international-copyright-law/
26. Creative Commons. Available from: https://creativecommons.org/licenses/
27. ClinicalTrials.gov ClinicalTrials.gov Database. 2019 February 2019]; Available from: https://clinicaltrials.gov/
28. Taichman DB, Sahni P, Pinborg A, Peiperl L, Laine C, James A, et al. Data Sharing Statements for Clinical Trials: A Requirement of the International Committee of Medical Journal Editors. Jama. 2017 Jun 27;317(24):2491-2.
29. Haji Ali Afzali H, Karnon J. Addressing the challenge for well informed and consistent reimbursement decisions: the case for reference models. PharmacoEconomics. 2011 Oct;29(10):823-5.
30. US Census Bureau. Synthetic Longitudinal Business Database (SynLBD). 2012 February 2019]; Available from: https://www.census.gov/ces/dataproducts/synlbd/

FIGURES

Fig. 1 Level and reach of transparency: possible scenarios

HTA: Health Technology Assessment; **GUI**: Graphic User Interface