VALUE IN HEALTH ▮ (2018) ▮▮▮–▮▮▮

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Computer Modeling of Diabetes and Its Transparency: A Report on the Eighth Mount Hood Challenge

Andrew J. Palmer, MBBS[1,], Lei Si, PhD[1], Michelle Tew, MPharm, MPH[2], Xinyang Hua, MS[2], Michael S. Willis, PhD[3], Christian Asseburg, PhD[3], Phil McEwan, PhD[4], José Leal, PhD[5], Alastair Gray, PhD[5], Volker Foos, MSc[6], Mark Lamotte, MD[6], Talitha Feenstra, PhD[7,8], Patrick J. O'Connor, MD, MPH[9], Michael Brandle, MD[10], Harry J. Smolen, PhD[11], James C. Gahn, PhD[11], William J. Valentine, PhD[12], Richard F. Pollock, MA, MSci[12], Penny Breeze, PhD[13], Alan Brennan, PhD[13], Daniel Pollard, MSc[13], Wen Ye, PhD[14], William H. Herman, MD[15], Deanna J. Isaman, PhD[14], Shihchen Kuo, RPh, PhD[16], Neda Laiteerapong, MD[17], An Tran-Duy, PhD[2], Philip M. Clarke, PhD[2]*

[1]*Menzies Institute for Medical Research, University of Tasmania, Hobart, Tasmania, Australia;* [2]*Centre for Health Policy, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Victoria, Australia;* [3]*The Swedish Institute for Health Economics, Lund, Sweden;* [4]*Health Economics and Outcomes Research Ltd., Cardiff, UK;* [5]*Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, UK;* [6]*IQVIA, Real-World Evidence Solutions, Zaventem, Belgium;* [7]*National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands;* [8]*Groningen University, University Medical Center Groningen, Groningen, The Netherlands;* [9]*HealthPartners Institute and HealthPartners Center for Chronic Care Innovation, Minneapolis, MN, USA;* [10]*Department of Internal Medicine, Kantonsspital St. Gallen, St. Gallen, Switzerland;* [11]*Medical Decision Modeling Inc., Indianapolis, IN, USA;* [12]*Ossian Health Economics and Communications, Basel, Switzerland;* [13]*School of Health and Related Research, University of Sheffield, Sheffield, UK;* [14]*Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA;* [15]*Departments of Internal Medicine and Epidemiology, University of Michigan, Ann Arbor, MI, USA;* [16]*Departments of Internal Medicine, University of Michigan, Ann Arbor, MI, USA;* [17]*Department of Medicine, University of Chicago, Chicago, IL, USA*

ABSTRACT

Objectives: The Eighth Mount Hood Challenge (held in St. Gallen, Switzerland, in September 2016) evaluated the transparency of model input documentation from two published health economics studies and developed guidelines for improving transparency in the reporting of input data underlying model-based economic analyses in diabetes. **Methods:** Participating modeling groups were asked to reproduce the results of two published studies using the input data described in those articles. Gaps in input data were filled with assumptions reported by the modeling groups. Goodness of fit between the results reported in the target studies and the groups' replicated outputs was evaluated using the slope of linear regression line and the coefficient of determination (R^2). After a general discussion of the results, a diabetes-specific checklist for the transparency of model input was developed. **Results:** Seven groups participated in the transparency challenge. The reporting of key model input parameters in the two studies, including the baseline characteristics of simulated patients, treatment effect and treatment intensification threshold assumptions, treatment effect evolution, prediction of complications and costs data, was inadequately transparent (and often missing altogether). Not surprisingly, goodness of fit was better for the study that reported its input data with more transparency. To improve the transparency in diabetes modeling, the Diabetes Modeling Input Checklist listing the minimal input data required for reproducibility in most diabetes modeling applications was developed. **Conclusions:** Transparency of diabetes model inputs is important to the reproducibility and credibility of simulation results. In the Eighth Mount Hood Challenge, the Diabetes Modeling Input Checklist was developed with the goal of improving the transparency of input data reporting and reproducibility of diabetes simulation model results.

Keywords: computer modeling, diabetes, Mount Hood Challenge, transparency.

Introduction

The use of economic simulation modeling tools to support decision making in the health care setting is widespread and necessary [1,2]. This is especially true for chronic and progressive diseases such as diabetes mellitus (DM), for which the time horizon of interest for decision making is lifetime and thus beyond the time and resource constraints of clinical trials. Health economic modeling provides a unique opportunity to capture the health and cost consequences of new interventions over the

* *Address correspondence to:* Andrew J. Palmer, Menzies Institute for Medical Research, University of Tasmania, Medical Science 1 Building, 17 Liverpool Street (Private Bag 23), Hobart, Tasmania 7000, Australia.

E-mail: Andrew.Palmer@utas.edu.au.

1098-3015$36.00 – see front matter

https://doi.org/10.1016/j.jval.2018.02.002

relevant time horizon as well as across all comparators of interest to decision makers.

To inform the allocation of resources, models informing such decisions must be clinically credible and valid for the populations and jurisdictions of interest. This can be achieved by reporting models in a transparent manner and testing their internal and external validity. This was emphasized in the International Society for Pharmacoeconomics and Outcomes Research and the Society for Medical Decision Making (ISPOR-SMDM) Modeling Good Research Practices [2], which advocated for "sufficient information to enable the full spectrum of readers to understand a model's accuracy, limitations, and potential applications at a level appropriate to their expertise and needs" [3], and in the DM-specific American Diabetes Association (ADA) guidelines for computer modeling [4], which encouraged reporting "in sufficient detail to reproduce the model and its results" [4]. The Second Panel on Cost-Effectiveness in Health and Medicine [5] similarly advocated transparency, although in a more limited manner.

The main focus of these guidelines is on the transparency of model structure, rather than on the assumptions and data used in simulating an individual application (e.g., population characteristics at baseline and the assumed nature and duration of treatment effects). A model with a fully transparent (and internally and externally valid) structure is not sufficient to reproduce the results of any individual simulation. To achieve this, one must also know what assumptions and input data were included. In the spirit of the Turing test [6] of a machine's ability to exhibit intelligent behavior, we have constructed a hypothetical thought experiment in which two isolated users have access to the same computer simulation model. A simulation would be regarded as transparent if one of the users was able to produce a set of instructions of the simulation they undertook that was sufficiently detailed and comprehensive to allow the other user to implement them and produce identical results using the same model. The ISPOR Consolidated Health Economic Evaluation Reporting Standards checklist [7] outlines many of the items that should routinely be in an economic evaluation, and the Philips checklist is a best practice guideline in model reporting [8]. Both include a range of items concerning application-specific input data. They may, however, be overly general to satisfy the needs in complicated multifactorial disease areas such as DM, and so we have attempted to address this gap in the literature using the Mount Hood diabetes simulation modeling network.

Initiated in 2000 by Andrew Palmer and Jonathan Brown at Timberline Lodge, Mount Hood, OR [9–11], the Mount Hood Challenge is a biennial congress in which as many as 10 DM modeling groups have met to compare and contrast models, methods, and data in the context of simulating standardized treatment scenarios and discussing the results. In September 2016, DM modeling groups gathered in St. Gallen, Switzerland, for the Eighth Mount Hood Challenge, with the aim of standardizing the recording and documentation of simulation inputs and communication of outputs in DM simulation modeling and thereby promoting transparency.

Specifically, the aims of the 2016 Mount Hood Challenge were twofold:

1. to evaluate transparency of key model inputs using two published studies as examples; and
2. to develop a DM-specific checklist for transparency of input data that can be used alongside general health economic modeling guidelines to improve reproducibility of health economic analyses in DM.

The present article summarizes the findings from the first objective and how modelers built on these to develop a series of DM-specific transparency recommendations addressing the second aim. The resulting checklist can serve as a means of improving consistency and transparency in diabetes simulation models and provide a framework for developing similar standards in other disease areas.

Methods

The Eighth Mount Hood Challenge was advertised on the Mount Hood Challenge Web site (https://www.mthooddiabeteschallenge.com/) and all known published diabetes modeling groups were invited to participate. The meeting featured two exercises using instructions provided before the meeting: a transparency challenge on day 1 and a communicating outcomes challenge on day 2. Modeling groups were encouraged to submit results for both challenges. Over the course of 2 days, results were presented and discussed, and paths to improvement were debated. A representative from each of the modeling groups was invited to participate on the third day to choose a topic for a meeting proceedings article. The group chose to focus this article on the transparency challenge only. For details of methods and results of the communicating outcomes challenge, interested readers are referred to the Mount Hood Challenge Web site [12].

The Transparency Challenge

Model transparency, "the extent to which interested parties can review a model's structure, equations, parameter values, and assumptions" [3], is often poor in published economic evaluations, particularly for complex diseases such as DM [4]. More than 10 years after the ADA guidelines promoted increased transparency, this is the first time diabetes modeling groups have attempted to answer the questions "How reproducible are published simulation modeling studies?" and "What is the best way to describe a simulation so that it can be reproduced?"

The modeling groups were assigned two preselected published economic modeling studies in DM [13,14] (see instructions in Appendix 1 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2018.02.002). The first transparency challenge was to replicate the Baxter et al. [13] study, which used the IQVIA-Core Diabetes Model (IQVIA-CDM) to estimate the impact of modest and achievable improvements in glycemic controls on cumulative incidences of microvascular and macrovascular complications and the costs in adults with type 1 (T1DM) or type 2 DM (T2DM) in the UK system [13]. This transparency challenge focused on simulating the T2DM results.

The second transparency challenge was to replicate the UK Prospective Diabetes Study 72 (UKPDS 72), which used the UKPDS Outcomes Model (UKPDS-OM) version 1 to evaluate the cost utility of intensive blood glucose (conventional vs. intensive blood glucose control with insulin or sulphonylureas, and conventional vs. intensive blood glucose control with metformin in overweight patients) and blood pressure control (less tight blood pressure control vs. tight blood pressure control with angiotensin-converting enzyme inhibitors or β-blockers in hypertensive patients) in T2DM [14]. This transparency challenge focused on the comparison of intensive versus conventional blood glucose control in the main randomization.

Modeling groups were asked to use data provided in the study publications including supplementary appendices [13] as inputs into their models and replicate the study analyses. When critical data could not be found in the study publication, they were asked to record assumptions required to fill those data gaps. Simulation results were not blinded. Each group submitted results in advance of the congress.

The data gaps reported by each group were summarized in a tabular format and compared and contrasted during meeting