VALUE IN HEALTH 16 (2013) 670–685

Available online at www.sciencedirect.com

SciVerse ScienceDirect

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

Computer Modeling of Diabetes and Its Complications: A Report on the Fifth Mount Hood Challenge Meeting

Andrew J. Palmer, BmedSci, MBBS[*], *The Mount Hood 5 Modeling Group*

Menzies Research Institute Tasmania, University of Tasmania, Hobart, TAS, Australia

ABSTRACT

Objectives: The Mount Hood Challenge meetings provide a forum for computer modelers of diabetes to discuss and compare models, to assess predictions against data from clinical trials and other studies, and to identify key future developments in the field. This article reports the proceedings of the Fifth Mount Hood Challenge in 2010. **Methods:** Eight modeling groups participated. Each group was given four modeling challenges to perform (in type 2 diabetes): to simulate a trial of a lipid-lowering intervention (The Atorvastatin Study for Prevention of Coronary Heart Disease Endpoints in Non-Insulin-Dependent Diabetes Mellitus [ASPEN]), to simulate a trial of a blood glucose–lowering intervention (Action in Diabetes and Vascular Disease: Preterax and Diamicron Modified Release Controlled Evaluation [ADVANCE]), to simulate a trial of a blood pressure–lowering intervention (Cardiovascular Risk in Diabetes [ACCORD]), and (optional) to simulate a second trial of blood glucose–lowering therapy (ACCORD). Model outcomes for each challenge were compared with the published findings of the respective trials. **Results:** The results of the models varied from each other and, in some cases, from the published trial data in important ways. In general, the models performed well in terms of predicting the relative benefit of interventions, but performed less well in terms of quantifying the absolute risk of complications in patients with type 2 diabetes. Methodological challenges were highlighted including matching trial end-point definitions, the importance of assumptions concerning the progression of risk factors over time, and accurately matching the patient characteristics from each trial. **Conclusions:** The Fifth Mount Hood Challenge allowed modelers, through systematic comparison and validation exercises, to identify important differences between models, address key methodological challenges, and discuss avenues of research to improve future diabetes models.

Keywords: computer simulation, cost-effectiveness analysis, diabetes, health economics, modeling.

Introduction

A decade after Jonathan Brown and Andrew Palmer met to compare the 20-year predictions of two computer simulation models of type 2 diabetes in the Timberline Lodge, high on the side of Mount Hood near Portland, Oregon, the fifth Mount Hood Challenge meeting was held in Malmö, Sweden, in September 2010 [1]. A total of eight modeling groups participated in the 2010 challenge, which followed a similar format to previous Mount Hood meetings whereby modelers were asked to use their prediction models to simulate the outcomes of clinical studies to inform debate on the challenges facing groups working in this area.

Computer simulation models, in essence a series of mathematical equations combined in a structured framework, have many uses such as allowing data from clinical trials to be extrapolated over longer time periods and to other populations. By providing information for health care decision makers on long-term clinical outcomes and costs, such models allow informed choices to be made between available interventions. As the issue of cost containment becomes ever more pertinent for many health care decision makers, the reliance on computer simulation modeling is increasing. This is particularly true of chronic diseases such as type 2 diabetes, which develop over a long period of time and are associated with significant morbidity and mortality and a substantial economic burden [2].

Although cost-of-illness studies have taught us a great deal about the scale of the economic burden associated with diabetes, as well as the identity of the main cost drivers, they do little to help us understand the incremental value of new interventions in a given population. Clinical trials provide essential information on new interventions, but their limitations in terms of time frame (typically 1–3 years), tightly controlled designs, and often (highly) selected populations can make their findings difficult to generalize to other care settings or populations. Key parameters such as demographics, life expectancy, patient management/medical technology, treatment costs, and health budgets can vary widely between regions and between countries. Flexible computer models have the potential to overcome these problems and provide valuable information, such as assessments of long-term cost-effectiveness, for policymakers and reimbursement decision makers. To fulfill this role, models must be based on

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

E-mail: andrew.palmer@utas.edu.au.

1098-3015/$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2013.01.002

the best available evidence, and validated against clinical data (internal and external validation) as well as each other, and they must also be transparent, documented in detail, and open about their mechanisms and assumptions.

The aim of this article was to report the proceedings of the Fifth Mount Hood Challenge held in Malmö, Sweden, in 2010, with a view to providing a summary of how eight current diabetes models match up to data from published clinical studies as well as to each other, to highlight differences between models, and to offer an insight into the challenges facing diabetes models a decade after the First Mount Hood Challenge.

Research Design and Methods

The Fifth Mount Hood Challenge was the first meeting for 6 years (since the Fourth Mount Hood Challenge in 2004 [3]), and participating modelers were asked to perform simulations based on four published clinical trial data sets, thereby allowing comparison of all eight participating models against clinical data. Treatments and interventions, management of patients, and cohort characteristics were defined in advance to minimize the number of potentially disparate assumptions required to make reliable forecasts over the duration of follow-up reported in the trial publications. The working hypothesis for the Mount Hood Challenge was that this process of standardized comparison is the best method to identify differences between models as well as assess the models' reliability in predicting the consequences of changes in risk factors brought about by an intervention in a clinical trial situation. Readers are referred to the Mount Hood web site for details of instructions given to modeling groups, and for contact details of modeling groups for further clarifications or detailed information regarding individual model structures and assumptions [4]. All groups with a published simulation model of type 2 diabetes were invited to take part in the challenge. In total, eight groups accepted this year's challenge. They were joined at the meeting by 85 participants from 10 countries.

Simulation Challenges in Type 2 Diabetes

To expand on the validation exercises from the Fourth Mount Hood Challenge Meeting in 2004, the modelers performed four external validation analyses against three recent clinical trials that reported the results of interventions attempting to modify key risk factors for the complications of type 2 diabetes. For each of these trials, the modeling groups attempted to predict the event rates of the primary end points and of as many secondary outcomes as possible at the end of the study by using only the size of the risk factor change achieved where possible within the individual models' frameworks, or using relative effects of treatments if this was the only option with a model. All modeling groups were restricted from using any information that they may have had access to (e.g., patient-level data) that was not in the public domain to produce the primary results set for presentation at the meeting. Such data, however, could be used to produce an additional set of results to examine whether model performance could be improved. The end points reported (by default) by the models did not always identically match those reported in the clinical trials. While modeling groups were asked to report as many outcomes as they could for each trial, some models could not generate results for every end point. As a result, the results tables are incomplete, with empty cells where no appropriate end point for comparison was available from that model. Notable differences in end-point definitions have been cited where relevant in this article. Results were to be reported as the proportion of patients experiencing each type of event (so as to match the outcomes reported in the trial publications). The following challenges were set.

Lipid-lowering intervention based on the Atorvastatin Study for Prevention of Coronary Heart Disease Endpoints in Non-Insulin-Dependent Diabetes Mellitus trial [5]

The Atorvastatin Study for Prevention of Coronary Heart Disease Endpoints in Non-Insulin-Dependent Diabetes Mellitus (ASPEN) was a 4-year, double-blind, parallel group trial of 10 mg of atorvastatin versus placebo in patients with type 2 diabetes and low-density lipoprotein cholesterol levels below contemporary guideline targets ($\leq$160 mg/dl [4.1 mmol/l], or $\leq$140 mg/dl [3.6 mmol/l] for subjects with myocardial infarction [MI] or coronary intervention >3 months before screening). The composite primary end point comprised cardiovascular death, nonfatal MI, nonfatal stroke, recanalization, coronary artery bypass surgery, resuscitated cardiac arrest, and worsening or unstable angina requiring hospitalization. Exclusion criteria included hemoglobin A_{1c} (Hb A_{1c}) value of over 10% (86 mmol/mol), blood pressure over 160/100 mm Hg, body mass index (BMI) over 35 kg/m^2, preexisting liver disease, kidney disease, or heart failure treated with digoxin. Patients were advised to adopt a National Cholesterol Education Program diet (which is low in saturated and *trans* fats, and rich in fruits, vegetables, whole grains, fat-free and low-fat dairy products, and lean meat, fish, and poultry). A total of 2410 patients were randomly allocated to receive atorvastatin or placebo. Mean patient age was 60 years, and approximately two-thirds of the study population was male (Table 1). At the end of the study, composite primary end point rates for atorvastatin and placebo were 13.7% and 15.0%, respectively (hazard ratio 0.90; 95% confidence interval [CI] 0.73–1.12). Subgroup analysis in patients with a history of MI or interventional procedure showed a hazard ratio of 0.82 (95% CI 0.59 –1.15). In patients with no prior MI or interventional procedure, the hazard ratio was estimated to be 0.97 (95% CI 0.74–1.28) for atorvastatin versus placebo.

Blood glucose–lowering intervention from the Action in Diabetes and Vascular Disease: Preterax and Diamicron Modified Release Controlled Evaluation trial [6]

In the Action in Diabetes and Vascular Disease: Preterax and Diamicron Modified Release Controlled Evaluation (ADVANCE) trial, a total of 11,140 patients with type 2 diabetes were randomly assigned to undergo either standard glucose control or intensive glucose control, defined as the use of gliclazide (30–120 mg modified release) plus other drugs (metformin, thiazolidinediones, acarbose, and/or insulin) as required to achieve an Hb A_{1c} value of 6.5% or less (Table 1). Primary end points were composites of major macrovascular events (death from cardiovascular causes, nonfatal MI, or nonfatal stroke) and major microvascular events (new or worsening nephropathy or retinopathy), assessed both jointly and separately. Inclusion criteria included a history of macro- and microvascular disease and age 55 years or more. Exclusion criteria included a requirement for insulin at the time of study initiation. After a median 5 years of follow-up, mean Hb A_{1c} value was lower in the intensive-control group (6.5%) than in the standard-control group (7.3%). Intensive control reduced the incidence of combined major macrovascular and microvascular events (18.1%, vs. 20.0% with standard control; hazard ratio 0.90; $P = 0.01$), as well as that of major microvascular events (9.4% vs. 10.9%; $P = 0.01$), primarily because of a reduction in the incidence of nephropathy. The hazard ratio for macrovascular events was 0.94 for intensive treatment versus standard care ($P = 0.32$, not significant).