Comparing, Contrasting, and Validating Health Economic Decision Models: Experiences From the Latest Mt. Hood Challenge in Diabetes and Lessons for Other Disease Areas
Comparing, Contrasting, and Validating Health Economic Decision Models: Experiences From the Latest Mt. Hood Challenge in Diabetes and Lessons for Other Disease Areas

Michael Willis, PhD
Talitha Feenstra, PhD
Mark Lamotte, MD
Alan Brennan, PhD

ISPOR 20th Annual European Congress
Glasgow, SCOTLAND

November 8, 2017
Disclosures

• Michael Willis is an employee at The Swedish Institute for Health Economics and a developer of The Economic and Health Outcomes Model of T2DM (ECHO-T2DM). ECHO-T2DM is proprietary software. No funding or consultant fees were involved in this presentation.

• Talitha Feenstra works at the University Medical Centre Groningen and at the Dutch National Institute for Public Health and the Environment (RIVM), and co-developed the MICADO diabetes model. For the current presentation she has no conflicts of interest. The COPD modellers’ meetings have been organized by the steering committee of which she is a member and were sponsored by: GSK; Novartis; BI; Takeda; Pfizer/BI NL; Nycomed.

• Mark Lamotte is an employee at QuintilesIMS and leader of the QuintilesIMS Core Diabetes Model team. QuintilesIMS received license fees and consulting fees for the use of the CDM. The current project was however done independent from any funding.

• Alan Brennan is an employee of University of Sheffield, has been involved in developing the Sheffield Type 2 Diabetes Model, The Sheffield Type 1 Diabetes Model and the SPHR Diabetes prevention model. He reports no conflicts of interest for this workshop.
ISPOR/SMDM
Best Practices for Model Validation

• **Face validity:** Activities establishing extent to which a model, its assumptions, and applications for which it is used reflect accurately current scientific evidence (as judged by experts)

• **Verification:** Activities establishing extent to which model calculations are correctly implemented, including thorough testing, de-bugging, and ‘stress-testing’ with extreme input values to expose errors of logic and programming

• **Cross-Validation:** Simulate same standardized scenarios with different models, comparing and contrasting results and investigating differences
  - A cornerstone of the Mt. Hood Challenges in T2DM

• **External Validation:** Test concordance between model predictions and observed outcomes for real patients (e.g., key RCTs)
  - Dependent outcomes (from data used in model construction) vs. Independent (from studies not used in models)

• **Predictive validation:** Prospective form of external validation in which the study has not yet been conducted, thus ensuring that the external validation is blinded to the analysts
  - While “strongest” form of evidence, this type of validation requires conditions that are relatively rare

---

1 A series of 7 ISPOR Task Force Reports published in *Value in Health* 15 (2012).
Seminal Mount Hood Challenge

• Analysts involved with the IMIB Model and the Global Diabetes Model (GDM) at the Timberline Lodge on the slopes of Mount Hood in Oregon, USA in August 2000

• Methods
  o IMIB and GDM were loaded with 12 sets of identical T2DM patients and simulated for 20 years
  o Survival, MI, stroke, diabetic retinopathy, albuminuria, and amputation rates were extracted and compared (i.e., cross-validation)
  o Differences were explored and explanations sought (and documented)

• Results
  • ”Both models generated realistic results and appropriate responses to changes in risk factors” (Brown et al, 2000)
  • There were important numerical differences (especially costs), however, but could be explained by differences in model architecture and CVD risk engines

• Importantly, there was an agreement to repeat the Challenge and to invite more modeling groups
MH Challenges Have Occurred Roughly Every Two Years Since

- For anyone interested in diabetes health economics or epidemiology
  - Clinical medicine, academia, pharmaceutical industry, reimbursement decision makers, or government agencies
- Theme varies, but recurring activities include simulating standardized scenarios
- Beginning in 2010, scope was broadened to include abstract submissions and presentations as well
- Attendance has ranged between 70 and 90 participants

<table>
<thead>
<tr>
<th>Place</th>
<th>Date</th>
<th>Participating Models</th>
<th>Theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>MH1 Mt. Hood, Oregon</td>
<td>August 2000</td>
<td>2</td>
<td>Original MH Challenge</td>
</tr>
<tr>
<td>MH2 San Francisco, California</td>
<td>June 12, 2002</td>
<td>6</td>
<td>Improving reliability, validity and usefulness of computer simulation models of diabetes</td>
</tr>
<tr>
<td>MH3 Oxford, England</td>
<td>August 30-31, 2003</td>
<td>6</td>
<td>Predicting future complications, costs, and lifespan for five pre-specified patients and five standardized treatments</td>
</tr>
<tr>
<td>MH4 Basel, Switzerland</td>
<td>September 2-4, 2004</td>
<td>7</td>
<td>(1) Introduce external validation using CARDS study data, (2) simulate DCCT data for T1DM modeling, and (3) estimate outcomes for a precisely defined hypothetical person with type 2 diabetes, with and without glycemic control</td>
</tr>
<tr>
<td>MH5 Malmö, Sweden</td>
<td>September 19-20, 2010</td>
<td>8</td>
<td>(1) Validation against recent clinical trial outcomes and (2) capturing uncertainty</td>
</tr>
<tr>
<td>MH6 Baltimore, Maryland</td>
<td>June 7-8, 2012</td>
<td>8</td>
<td>(1) Validation against new clinical trial and observational data outcomes with emphasis on “blinding” and (2) exploration of 2nd order uncertainty in modeling standardized scenarios</td>
</tr>
<tr>
<td>MH7 Stanford, California</td>
<td>June 17-19, 2014</td>
<td>10</td>
<td>(1) Simulating the new Look AHEAD results, (2) predicting mortality after major events, and (3) exploring ethnicity-related variability in the models (suitability of models geographically)</td>
</tr>
<tr>
<td>MH8 St. Gallen, Switzerland</td>
<td>September 16-18, 2016</td>
<td>10</td>
<td>(1) Transparency of simulations and results and (2) communicating outcomes</td>
</tr>
</tbody>
</table>

https://www.mthooddiabeteschallenge.com/
The Next Mount Hood Challenge

Mt Hood 2018 Diabetes Challenge

5-7 October 2018
Düsseldorf, Germany

https://www.mthooddiabeteschallenge.com/
Mount Hood Challenges Mission

• Knowledge sharing and improved communication:
  o Between modeling groups
  o Between model developers, model users, and consumers of modeling results

• Improve quality of research, set (voluntary) minimum standards
  o Publication of conference proceedings (Brown et al [2000], The Mount Hood 4 Modeling Group [2007], Palmer et al [2013])
  o A web page that contains historical information on previous Challenges, an information repository for diabetes models, user-submitted publication lists, and more (https://www.mthooddiabeteschallenge.com/)
  o Mt. Hood participants worked on ADA "Guidelines for Computer Modeling of Diabetes and Its Complications" (Diabetes Care 27 [Summer 2004])
  o Recommendations for minimum reporting standards under submission (Dr. Lamotte will discuss shortly)

• Platform that promotes external auditing of diabetes modeling
  o Lift perceived credibility to outside actors who often see diabetes models as ”black boxes”
How Do MH Challenges Work in Practice?

• Organizing committee chooses theme and venue
  o Preferably with support of a local university or health economics consultantcy
• Simulation Challenges are defined and distributed to participants (and placed on line to aid recruitment of new modeling groups)
• Modeling groups perform simulations and submit results in advance of congress date
• Abstract submission for short presentations open to all meeting participants (volunteers review abstracts and set up an afternoon of parallel presentation sessions)
• The Challenges (1.5 days):
  o Each participating modeling group presents key model details (briefly)
  o Challenge results are presented globally (were previously presented by each model group individually, but led to time-consuming duplication)
  o Considerable time is reserved for discussion of the results, debate, and consensus building
• Invited speakers/submitted abstracts and presentations:
  o Invited speakers address key issues related to the congress theme
  o Accepted abstracts are presented to share insight in areas of active research
Management Challenges

• Administration/Finances (always a challenge!!):
  o Not for profit
  o Funding via participation fees (with some corporate donations in early conferences)
  o Universities have been leveraged to handle financial aspects

• Human resource requirements are extensive:
  o Time to organize the meeting (both content and logistics)
  o Time for modelers to run simulations and submit documentation
  o Time for someone to organize results across modeling teams and present to group/moderate subsequent discussion

• Don’t overemphasize ”Challenge”
  o Competitions (e.g., for best fit) de-emphasize cooperation
  o Blinded vs. unblinded challenges

• Intellectual Property
  o Recognize where cooperation/sharing begins and ends
  o Focus on common goals: code-sharing/full transparency may work in some areas, but in many settings it doesn’t
  o Ensure that groups feel comfortable with sharing and that submitted results are not used without consent elsewhere

• Creating interest
  o Need minimum number of participants to cover fixed costs (and keep conference fees reasonable)
  o Sites/timing generally linked to big diabetes meetings (ADA, EASD) to reduce travel costs
Some Features of Diabetes That Perhaps Helped Precipitate the MH Challenges

• Complicated pathophysiology
  o DM models must capture disease progression, in which multiple risk factors (including blood glucose and blood pressure) can impact on a wide variety of co-morbid and interdependent health outcomes like cardiovascular disease, renal failure, amputation, and blindness
  o Need for these models to be multi-application, otherwise prohibitively expensive
  o Many consider the models to be black boxes; engendering trust is crucial and the DM field was early to realize the importance (necessity) of model validation

• Big disease prevalence
  o Relative abundance of data
  o Critical mass of interested and knowledgable researchers
  o Mix of different actors with different role; perhaps making cooperation easier?

• Presence of engaged individual researchers
  o Among others, Philip Clarke (University of Melbourne) and Andrew Palmer (University of Tasmania)

• Where similar factors exist for other disease areas, they should be leveraged
• Where differences exist, alternative solutions might be warranted
Lessons from other disease areas

COPD Modelling meetings
other initiatives

Talitha Feenstra,
Short intro to Chronic Obstructive Pulmonary Disease

Chronic Obstructive Pulmonary Disease (COPD) is a common, preventable and treatable disease that is characterized by **persistent respiratory symptoms and airflow limitation** that is due to airway and/or alveolar abnormalities usually caused by significant exposure to noxious particles or gases.

Chronic disorder
- Lung function decline
- Respiratory symptoms
- Exacerbations
- Comorbidities
- Increased mortality
COPD modellers meetings

• Almost annual meetings since 2011 (now 5 in total)
• Most recent was February 2017
• Inspired by Mount Hood Challenges
• Wish to validate Dutch COPD model
• Main organizer: Martine Hoogendoorn (EUR/iMTA)
• Contact: t.l.feenstra@umcg.nl or hoogendoorn@imta.eur.nl
12 models have participated at least once

- NL: Dynamic population COPD Progression model (Hoogendoorn et al)
- US: Dynamic Cohort COPD model (Hansen et al)
- S: Swedish generic model of disease history and economic impact of COPD (Borg et al)
- DE: The German comprehensive care COPD model (Wacker et al)
- It: Italian COPD population model (Dal Negro et al)
- US Pharmacometric-pharmacoeconomic model (represented by Slejko)
- Takeda global COPD model (Samyshkin et al)
- BI bronchodilator therapy COPD Model (Rutten-van Molken et al)
- IMS/Novartis COPD Markov model (Price et al)
- IMS/Novartis COPD patient simulation model (Asukai et al)
- GSK Galaxy COPD Disease Progression model (Briggs et al)
- GSK ICS/LABA model 2005 (Briggs et al)
Meeting Topics

- 2011: “COPD, towards comprehensive, valid and transparent models to support future decision making”
  - Presenting structure of models
  - Hypothetical scenarios changing model-parameters
- 2012: “COPD, towards comprehensive, valid and transparent models to support future decision making”
  - Hypothetical treatment scenarios
- 2014: “Modelling Personalized COPD Care: economic, societal and regulatory implications”
  - Scenarios based on trial-data
  - Scenarios for subgroups
- 2015: “Personalized treatment of COPD in relation to economic modelling”
  - Prediction models for exacerbations
- 2017: “Treatment adherence and meta-modelling”
  - Meta-models
Typical Meeting Activities

• Model structure
  o Present models
  o Analyze heterogeneity
  o Prediction modeling (exacerbations)

• Investigate essential parameters
  o Scenario analyses
  o Meta-modeling

• Validate against external sources
  o scenario analyses

• External speakers
  o New perspectives
Meeting Outcomes

• Better understanding of COPD modeling
  o Main drivers of results
  o Different approaches to model same phenomena
• Better model validity
• New methods (external presenters)
• Insights from clinicians
• Great discussions, leading to publications
Prediction models for exacerbations in different COPD patient populations: comparing results of five large data sources

Martine Hoogendoorn¹
Talitha L Feenstra²,³
Melinde Boland¹
Andrew H Briggs⁴
Sixten Borg⁵
Sven-Arne Jansson⁶
Nancy A Risebrough⁷
Julia F Sleijko⁸
Maureen PMH Rutten-van Mölken¹

ABSTRACT

Background and objectives: Exacerbations are important outcomes in COPD both from a clinical and an economic perspective. Most studies investigating predictors of exacerbations were performed in COPD patients participating in pharmacological clinical trials who usually have moderate to severe airflow obstruction. This study was aimed to investigate whether predictors of COPD exacerbations depend on the COPD population studied.

Methods: A network of COPD health economic modelers used data from five COPD data sources—two population-based studies (COPDGene® and The Obstructive Lung Disease in Norrbotten), one primary care study (RECODE), and two studies in secondary care (Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoint and UPLIFT) —to estimate and validate several prediction models for total and severe exacerbations (~hospitalization). The...
CANCER: The CISNET Initiative

- Not personally involved
- Consortium, sponsored by NCI from 2000 onwards (https://cisnet.cancer.gov/)
- Range of
  - Cancer sites (BC; LC; Cervical; CRC; Esophageal; Prostate)
  - Interventions (Prevention, screening, treatment, new:diagnosis, biomarkers, palliative care)
  - Countries: seems limited, most US-based models
- Number of participating models varies by site from 3 to 8
- Activities:
  - Modeling same problem with various models: comparative modeling
  - Methodological and technical issues (programmers group)
  - Model registry, allowing selection based on site, model type, etc.
    https://resources.cisnet.cancer.gov/registry/home/
  - Tools based on the models, for policy makers
Checking the criteria for cancer and COPD

- Complicated pathophysiology:
  - Need for models to be multi-application, otherwise prohibitively expensive
- Big disease prevalence
  - Relative abundance of data
  - Critical mass of interested and knowledgeable researchers
- Presence of engaged individual researchers
- Ability to solve issues of
  - Finances
  - Time of participants
  - Confidentiality
Conclusion

• Cross model validation worth the effort
  o But it is an effort indeed
• Increases separate models’ validity
• Increases methodological knowledge of modellers as well as users
• Increases insight in models
  o For model developers
  o For model users
How reproducible are published simulation modeling studies?

Mark Lamotte
The modeling groups participating to Mount Hood 8 received the following challenge

- Two published papers were selected
  - Baxter et al. (2016)
  - UKPDS 72
- Modelling teams to attempt to replicate the analyses
  - Extract information from the PDFs and Supplementary Appendices provided and load model to “best of ability”
    - If anything contradictory or unclear, the groups were charged with deciding and documenting
  - In the event of data gaps, groups were charged with filling the gaps and documenting
  - Simulate the decision problems in the PDFs
- Submit in advance
  - Brief summary (<300 words) that “could potentially from the methods section of a published paper”
  - Detailed methods section that would be “fully transparent … (and permit a ‘blinded’ researcher to reproduce … results)”
  - Summary of the data gaps in the PDFs and assumptions required
  - Challenge results
Caveat

• Baxter and UKPDS team were kind enough to act as sacrificial lambs
  o Idea behind challenge was not to criticize the publications, but rather to leverage them to create momentum/direction for standards to promote transparency and replicability in DM modeling
Focus of the challenge was on Type 2 diabetes
Objectives of the Baxter study

- Estimate the potential cost avoidance that may be achieved through reducing complication rates by making achievable, incremental improvements in glycaemic control, when compared with the levels currently delivered in clinical practice
- It is not predicated on any specific therapy, but simply more timely and appropriate interventions to improve care
Savings are reported per HbA1c interval per person and projected to UK population

Table 2: Type 2 diabetes cost reductions per person, and for the total current UK adult population with Type 2 diabetes, from avoided complications for management of HbA1c, treatments levels applied to the National Institute for Health and Care Excellence.

<table>
<thead>
<tr>
<th>HbA1c</th>
<th>5 years</th>
<th>10 years</th>
<th>15 years</th>
<th>20 years</th>
<th>25 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adults 2 diabetes, per-person cost reductions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 7% uncontrolled</td>
<td>154.11</td>
<td>191.24</td>
<td>236.97</td>
<td>290.46</td>
<td>342.88</td>
</tr>
<tr>
<td>7% uncontrolled</td>
<td>154.11</td>
<td>191.24</td>
<td>236.97</td>
<td>290.46</td>
<td>342.88</td>
</tr>
<tr>
<td>&gt; 7% uncontrolled</td>
<td>154.11</td>
<td>191.24</td>
<td>236.97</td>
<td>290.46</td>
<td>342.88</td>
</tr>
<tr>
<td>Adults 2 diabetes, total population cost reductions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 7% uncontrolled</td>
<td>154.11</td>
<td>191.24</td>
<td>236.97</td>
<td>290.46</td>
<td>342.88</td>
</tr>
<tr>
<td>7% uncontrolled</td>
<td>154.11</td>
<td>191.24</td>
<td>236.97</td>
<td>290.46</td>
<td>342.88</td>
</tr>
<tr>
<td>&gt; 7% uncontrolled</td>
<td>154.11</td>
<td>191.24</td>
<td>236.97</td>
<td>290.46</td>
<td>342.88</td>
</tr>
</tbody>
</table>

TOTAL: 154.11 191.24 236.97 290.46 342.88
And also on number of complications

<table>
<thead>
<tr>
<th>Type 2 diabetes</th>
<th>5 years</th>
<th>10 years</th>
<th>15 years</th>
<th>20 years</th>
<th>25 years</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Eye disease</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDR</td>
<td>16,515</td>
<td>39,086</td>
<td>61,222</td>
<td>71,948</td>
<td>68,404</td>
</tr>
<tr>
<td>PDR</td>
<td>3,791</td>
<td>8,836</td>
<td>13,083</td>
<td>14,673</td>
<td>14,830</td>
</tr>
<tr>
<td>ME</td>
<td>22,544</td>
<td>55,338</td>
<td>87,923</td>
<td>100,391</td>
<td>93,842</td>
</tr>
<tr>
<td>SVL</td>
<td>6,215</td>
<td>21,794</td>
<td>37,921</td>
<td>46,768</td>
<td>48,418</td>
</tr>
<tr>
<td>Cataract</td>
<td>7,711</td>
<td>16,738</td>
<td>24,844</td>
<td>27,288</td>
<td>25,275</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>56,777</td>
<td>141,792</td>
<td>224,992</td>
<td>261,069</td>
<td>250,768</td>
</tr>
<tr>
<td><strong>Renal disease</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA</td>
<td>37,844</td>
<td>93,221</td>
<td>143,466</td>
<td>158,051</td>
<td>142,375</td>
</tr>
<tr>
<td>ESRD</td>
<td>276</td>
<td>2,183</td>
<td>6,540</td>
<td>11,663</td>
<td>14,524</td>
</tr>
<tr>
<td>Nephropathy</td>
<td>31</td>
<td>572</td>
<td>2,108</td>
<td>4,887</td>
<td>7,288</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>38,151</td>
<td>95,975</td>
<td>152,114</td>
<td>174,601</td>
<td>164,187</td>
</tr>
<tr>
<td><strong>Foot ulcers and amputations and neuropathy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ulcer</td>
<td>11,088</td>
<td>46,422</td>
<td>87,773</td>
<td>112,120</td>
<td>113,076</td>
</tr>
<tr>
<td>Amputation</td>
<td>872</td>
<td>6,695</td>
<td>16,331</td>
<td>25,601</td>
<td>30,449</td>
</tr>
<tr>
<td>Neuropathy</td>
<td>110,053</td>
<td>221,893</td>
<td>285,619</td>
<td>274,814</td>
<td>230,104</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>122,013</td>
<td>275,011</td>
<td>389,723</td>
<td>412,535</td>
<td>373,629</td>
</tr>
<tr>
<td><strong>Cardiovascular</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHF</td>
<td>14,766</td>
<td>32,569</td>
<td>52,270</td>
<td>59,807</td>
<td>52,241</td>
</tr>
<tr>
<td>PVD</td>
<td>1,837</td>
<td>4,460</td>
<td>7,312</td>
<td>8,666</td>
<td>8,187</td>
</tr>
<tr>
<td>Angina</td>
<td>4,785</td>
<td>10,560</td>
<td>17,048</td>
<td>16,844</td>
<td>16,315</td>
</tr>
<tr>
<td>Stroke</td>
<td>4,750</td>
<td>11,274</td>
<td>19,070</td>
<td>18,821</td>
<td>9,605</td>
</tr>
<tr>
<td>MI</td>
<td>1,852</td>
<td>3,031</td>
<td>2,190</td>
<td>-721</td>
<td>-3,960</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>27,991</td>
<td>61,893</td>
<td>97,890</td>
<td>106,416</td>
<td>82,387</td>
</tr>
</tbody>
</table>
A Brief Look at Aggregate Results for Baxter Replication

• The following groups participated:
  o Cardiff, the Cardiff Model;
  o ECHO-T2DM, the Economics and Health Outcomes Model of T2DM;
  o MDM-TTM, Medical Decision Modeling Inc - Treatment Transitions Model;
  o QI-CDM, Quintiles IMS-Core Diabetes Model
  o MMD: Michigan model (only commented on inputs)
## Overview of data gaps identified

<table>
<thead>
<tr>
<th>Category</th>
<th>Baxter study reported</th>
<th>Model input gaps identified by modelling groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline characteristics of simulated patients</td>
<td>Refer to IMS Disease Analyser (UK database)</td>
<td>Lack of baseline patient characteristics, Sample size not presented; No point estimates for baseline HbA1c provided within the ranges</td>
</tr>
<tr>
<td>Treatment effect / thresholds</td>
<td>Refers to HbA1c treatment intensification levels in Khunti et al. (21) and NICE guidelines (18)</td>
<td>Referred value not present in the paper and count could not be discerned</td>
</tr>
<tr>
<td>Effect evolution</td>
<td>Modelling of modification of treatment at HbA1c thresholds indicated by current NICE guidelines (18)</td>
<td>Risk factor evolution for time-dependent parameters not specified; Unclear if there was a treatment algorithm with rescue treatment</td>
</tr>
<tr>
<td>Prediction of complications</td>
<td>Quintiles IMS Core Diabetes Model</td>
<td>Choice of rates/equations was not reported and should be for the Core Diabetes Model which has the ability to run different risk equations</td>
</tr>
<tr>
<td>Cost</td>
<td>Supplementary table of direct costs of complications and management costs</td>
<td>Cost for some complications missing (fatal MI, ulcers)</td>
</tr>
</tbody>
</table>
## Cost savings per HbA1c and per patient

<table>
<thead>
<tr>
<th>Baxter study</th>
<th>Participating modelling groups</th>
<th>Cardiff</th>
<th>ECHO-T2DM</th>
<th>MDM-TTM</th>
<th>QI-CDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 59 mmol/mol (7.5%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 years</td>
<td>£83</td>
<td>£16</td>
<td>£154</td>
<td>£7</td>
<td>£13</td>
</tr>
<tr>
<td>10 years</td>
<td>£317</td>
<td>£73</td>
<td>£418</td>
<td>£174</td>
<td>£151</td>
</tr>
<tr>
<td>15 years</td>
<td>£682</td>
<td>£179</td>
<td>£644</td>
<td>£353</td>
<td>£605</td>
</tr>
<tr>
<td>20 years</td>
<td>£1,078</td>
<td>£307</td>
<td>£838</td>
<td>£484</td>
<td>£1,283</td>
</tr>
<tr>
<td>25 years</td>
<td>£1,280</td>
<td>£422</td>
<td>£911</td>
<td>£521</td>
<td>£1,799</td>
</tr>
<tr>
<td>&gt; 59 mmol/mol (7.5%) to 64 mmol/mol (8.0%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 years</td>
<td>£132</td>
<td>£26</td>
<td>£170</td>
<td>£60</td>
<td>£9</td>
</tr>
<tr>
<td>10 years</td>
<td>£449</td>
<td>£104</td>
<td>£457</td>
<td>£208</td>
<td>£317</td>
</tr>
<tr>
<td>15 years</td>
<td>£995</td>
<td>£235</td>
<td>£658</td>
<td>£337</td>
<td>£1,069</td>
</tr>
<tr>
<td>20 years</td>
<td>£1,510</td>
<td>£385</td>
<td>£860</td>
<td>£379</td>
<td>£1,906</td>
</tr>
<tr>
<td>25 years</td>
<td>£1,678</td>
<td>£518</td>
<td>£976</td>
<td>£324</td>
<td>£2,503</td>
</tr>
<tr>
<td>&gt; 64 mmol/mol (8.0%) to 75 mmol/mol (9.0%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 years</td>
<td>£138</td>
<td>£68</td>
<td>£157</td>
<td>£83</td>
<td>-£16</td>
</tr>
<tr>
<td>10 years</td>
<td>£607</td>
<td>£201</td>
<td>£412</td>
<td>£218</td>
<td>£294</td>
</tr>
<tr>
<td>15 years</td>
<td>£1,366</td>
<td>£384</td>
<td>£651</td>
<td>£329</td>
<td>£1,198</td>
</tr>
<tr>
<td>20 years</td>
<td>£1,999</td>
<td>£580</td>
<td>£869</td>
<td>£331</td>
<td>£2,440</td>
</tr>
<tr>
<td>25 years</td>
<td>£2,223</td>
<td>£748</td>
<td>£942</td>
<td>£236</td>
<td>£3,810</td>
</tr>
<tr>
<td>&gt; 75 mmol/mol (9.0%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 years</td>
<td>£105</td>
<td>£160</td>
<td>£150</td>
<td>£146</td>
<td>£169</td>
</tr>
<tr>
<td>10 years</td>
<td>£622</td>
<td>£402</td>
<td>£427</td>
<td>£372</td>
<td>£1,019</td>
</tr>
<tr>
<td>15 years</td>
<td>£1,274</td>
<td>£697</td>
<td>£750</td>
<td>£561</td>
<td>£2,442</td>
</tr>
<tr>
<td>20 years</td>
<td>£1,591</td>
<td>£993</td>
<td>£923</td>
<td>£584</td>
<td>£4,255</td>
</tr>
<tr>
<td>25 years</td>
<td>£1,559</td>
<td>£1,231</td>
<td>£1,088</td>
<td>£476</td>
<td>£5,590</td>
</tr>
</tbody>
</table>
Costs Avoided per Patient, by HbA1c

- < 59 mmol/mol (7.5%)
  - 0 500 1,000 1,500 2,000
  - 0 5 10 15 20 25
  - Total Costs Avoided
  - Years

- > 59 mmol/mol (7.5%) to 64 mmol/mol (8.0%)
  - 0 1,000 2,000 3,000
  - 0 5 10 15 20 25
  - Total Costs Avoided
  - Years

- > 64 mmol/mol (8.0%) to 75 mmol/mol (9.0%)
  - 0 1,000 2,000 3,000 4,000 5,000
  - 0 5 10 15 20 25
  - Total Costs Avoided
  - Years

- > 75 mmol/mol (9.0%)
  - 0 1,000 2,000 3,000 4,000 5,000 6,000
  - 0 5 10 15 20 25
  - Total Costs Avoided
  - Years
Complications Avoided, for Full Population

Eye Complications (Total)

Renal Complications (Total)

Cardiovascular Complications (Total)

ECHO-T2DM and IMS-CDM assumed baseline BDR prevalence of 36% (UKPDS 33)
Complications avoided for individual cardiovascular complications

CHF

Angina

Stroke

MI

BAXTER  
Cardiff  
IMS-CDM  
MDM  
ECHO-T2DM
Possible reasons why we see differences ....
Hypothesis #1
Differences in assumed baseline patient characteristics may matter

• Baseline patient characteristics were poorly reported
  – Baseline HbA1c was in four brackets (<59, 59-64, 64-75, and >75 mmol/mol or <7.5%, 75-8%, 8-9%, and >9.0%)
  – Mean HbA1c (and SD) were not given,
  – No on other covariates.
  – Assumptions varied across groups:
    • Cardiff assumed fixed initial HbA1c of 7.0%, 7.75%, 8.5% and 9.5%, for each of the four brackets
    • ECHO-T2DM used distributions from NHANES: mean HbA1c 6.35%, 7.68%, 8.43%, 10.60%
    • QI-CDM sources baseline HbA1c from NICE guideline
    • Various assumptions regarding other covariates, which were set to the same values in all HbA1c subgroups (Cardiff, QI-CDM) or to subgroup- varying values (ECHO-T2DM)

• Unclear what effect this had on results
Hypothesis #2
The choice of risk equations may matter

- Unclear what Baxter used
- Cardiff, ECHO-T2DM, QI-CDM all used **UKPDS 82** for the T2DM patients, MDM?
  - Expect reasonably similar incidences of CVD morbidity and mortality?
    - Good for MI, maybe stroke, but complicated by differences in covariate values over time
Hypothesis #3
Assumption about downstream treatment intensification may matter

• QI-CDM modeled treatment intensification with additional efficacy, others (probably) did not

• Cardiff applied common HbA1c intensification threshold (7.5%), QI-CDM and ECHO-T2DM had separate threshold for the two arms, MDM?
Hypothesis #4
Differences in assumptions about unit costs may matter

- QI-CDM assumed costs for events other than found in Baxter, whereas other models applied only Baxter
  - Expect higher costs for IMS-CDM
  - Supported, but does not explain what drives differences between other models
Conclusion

- Detailed reporting of data inputs is needed
- If not, results cannot be reproduced
- Reader has a black box feeling
- HTA agencies will not believe us
  ➔ recommendations!
The Mount Hood Diabetes Modelling Transparency Checklist

Alan Brennan
School of Health and Related Research,
University of Sheffield,
United Kingdom
Background & building upon …

• Transparency of model inputs important to reproducibility & credibility of simulation results.

• ISPOR/SMDM Modeling Good Research Practices - “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” (1)

• The ISPOR CHEERS checklist (2), Philips checklist on best practice guideline in model reporting (3), AdViSHE (4).

• American Diabetes Association (ADA) guidelines for computer modeling - “sufficient detail to reproduce model and results” (5)

---

(4) P. Vemer1,2 • I. Corro Ramos3 • G. A. K. van Voorn• M. J. A. D. • F. L. Feenstra1,5 PharmacoEconomics (2016) 34:349-361
Objective

• Eighth Mount Hood Challenge – an exercise to address this
• Diabetes modelling groups attempted to answer 2 questions
  Q1) “how reproducible are published simulation modelling studies?”
  Q2) “what is the best way to describe a simulation so that it can be reproduced?”

Objective:
To develop a diabetes-specific checklist for transparency of input data that can be used alongside general health economic modelling guidelines to improve reproducibility of health economic analyses and simulation model results in diabetes
Method

• Modelling groups examined 2 cases studies
  o Data gaps reported by each group were summarized in a tabular format and compared and contrasted during meeting proceedings.
  o Documented lack of transparency in reporting model inputs including important deficiencies such as baseline patient characteristics, treatment effects, HbA1c evolution, treatment use over time.
  o Modelling groups generally sourced missing information from literature and made different assumptions

• MONDAY meeting after the Challenge
  o Discussed key issues & reached consensus to start draft guidelines
  o Post-meeting, draft paper proposing and motivating a checklist
  o 2 rounds of revision with all authors,
  o Final refined position paper was created - submitted to ViH journal
Alan, I tried to polish these a bit, can you confirm or change back?
Michael Willis, 3/11/2017
## The Checklist

### Table 6. Checklist of reporting model input in diabetes health economics studies.

<table>
<thead>
<tr>
<th>Diabetes Modelling Input Checklist</th>
<th>Checkbox</th>
<th>Comments (e.g. justification if not reported)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model input</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulation cohort</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnicity/race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI/weight</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of diabetes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline HbA1c, lipids and blood pressure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comorbidities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Treatment intervention</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment algorithm for HbA1c evolution over time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment algorithm for other conditions, e.g.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hypertension, dyslipidaemia, excess weight</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment initial effects on baseline biomarkers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rules for treatment intensification, e.g. the cut-off</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HbA1c level to switch the treatment, the type of new</td>
<td></td>
<td></td>
</tr>
<tr>
<td>treatment and whether the rescue treatment is an</td>
<td></td>
<td></td>
</tr>
<tr>
<td>addition or substitution to the standard treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-term effects, adverse effects, treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>adherence and persistence and residual effects after</td>
<td></td>
<td></td>
</tr>
<tr>
<td>the discontinuation of the treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trajectory of biomarkers, BMI, smoking and any</td>
<td></td>
<td></td>
</tr>
<tr>
<td>other factors that are affected by treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cost</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Differentiated by acute event in first year and</td>
<td></td>
<td></td>
</tr>
<tr>
<td>subsequent years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost of intervention and other costs, e.g.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>managing complications, adverse events, diagnostics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Please report unit prices and resource use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>separately and give information on discount rates</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Health state utilities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operational mechanics of the assignment of utility</td>
<td></td>
<td></td>
</tr>
<tr>
<td>values, i.e. utility- or disutility-oriented</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Management of multi-health conditions</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>General model characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choice of mortality table and any specific event-related mortality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choice and source of risk equations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>If microsimulation: number of Monte Carlo simulations conducted and justification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Components of model uncertainty being simulated (e.g. risk equations, risk factor trajectories, costs, treatment effect), number of simulations and justification</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Diabetes Modelling Input Checklist

Simulation cohort

Baseline patient characteristics of simulated cohort should be clearly stated, incl.

- age, sex, ethnicity/race, smoking status,
- body mass index (BMI)/weight, physical activity
- duration of diabetes, baseline HbA1c, lipids and blood pressure levels,
- comorbidities,
- baseline treatments
  - aspirin, statins, ACE-inhibitors/angiotensin II receptor blockers and/or glucose-lowering treatments.
- Baseline characteristics should be presented in a table as mean with standard deviation or as proportion. Statistical distributions for baseline characteristics should be reported in the table.
Diabetes Modelling Input Checklist

Treatment interventions

1. treatments / algorithms for blood glucose control, hypertension, dyslipidaemia, excess weight etc. for comparator & intervention
2. specify initial impact of treatment(s) on baseline biomarkers
3. rules for treatment intensification and thresholds triggering changes should be specified for HbA1c, blood pressure, lipids, BMI, eGFR
4. specify the set of long-term effects, adverse effects, treatment adherence and persistence, and assumptions on legacy effects i.e. residual treatment effects after the discontinuation of a treatment
5. direct and indirect links from treatment effects on glucose / lipids levels to health outcomes, costs and effectiveness e.g. HbA1c directly affects stroke, MI, retinopathy, nephropathy risks, HbA1c indirectly affects mortality through its impact on CVD
6. include effects on biomarker trajectories over time for HbA1c, lipids, blood pressure, BMI, eGFR, smoking
Diabetes Modelling Input Checklist

Costs & Utilities

Costs of …

• interventions themselves

• being in specific health state and on specific treatments

• complication management should consider timing of events
e.g. macrovascular complications high cost at the time of the event andlower follow-up management costs

• adverse events, diagnostics

• If a societal perspective is used then specify assumptions e.g. 
  productivity losses through absenteeism, presentism, or early retirement.

Health state utilities (HSUs)

• Methodology for utility for multiple co-morbidities should be stated e.g.
  ‘minimum’ (using value of the condition with the lowest utility score),
  ‘additive’ (using the arithmetic sum of utility decrements), or
  ‘multiplicative’ (using the product of utility decrement factors).
Diabetes Modelling Input Checklist
General Model Characteristics

1. choice of country-specific life table for all-cause mortality should be stated in methods, and specific event-related mortality must be stated.
2. document the source and details of risk equations used in the model.
3. if using a microsimulation model, authors should report and justify number of Monte Carlo simulations performed per individual.
4. when performing probabilistic sensitivity analysis it is important to document and justify distributions for components (e.g. risk equations, risk factor trajectories and treatment effect).
Discussion of Usefulness

• **Publications** - Modellers should document simulation inputs via checklist, and submit as supplementary materials with publications. Journal editors/reviewers permit (or require) inclusion of checklist.

• Use for each application - this is a minimal checklist for typical analyses - for some analyses other things will be needed. It is for each specific application of a model - not a general overall model ‘validation’ Issue

• “Costs of Transparency” - full transparency requires considerable resources of modellers and consumers of results. The checklist is a pragmatic solution, focused on influential parameters and assumptions.

• Further Work on …
  - Standardised model outputs to enable cross comparison of results
  - Test if checklist increases transparency at a future Mt Hood.

• Conclusion: - improve credibility and clarity. We hope the checklist will inspire modellers in similarly complex fields to promote transparency of inputs & improve reliability of outputs.