Electronic Supplementary Material 2

Comparing the Cohort and Micro-Simulation Modeling Approaches in Cost-Effectiveness Modelling of Type 2 Diabetes Mellitus (T2DM): A Case Study of the IHE Diabetes Cohort Model and the Economics and Health Outcomes Model of T2DM

Journal: Pharmacoeconomics

Michael Willis[1]

Adam Fridhammar[1]

Jens Gundgaard[2]

Andreas Nilsson[1]

Pierre Johansen[2]

[1]The Swedish Institute for Health Economics, Lund, Sweden

[2]Novo Nordisk A/S, Søborg, Denmark

Corresponding author:

Michael Willis, PhD

The Swedish Institute for Health Economics

Box 2017

SE-220 02 Lund

Email: mw@ihe.se

AdViSHE Validation Assessment Tool:
The IHE Diabetes Cohort Model (IHE-DCM)

Adam Fridhammar, MSc
Kristoffer Nilsson, MSc
Sofie Persson, PhD
The Swedish Institute for Health Economics
Lund, Sweden
www.ihe.se

December 19, 2019

Part A: Validation of the Conceptual Model

- **A1:** Have experts been asked to judge the appropriateness of the conceptual model?

  Yes, IHE-DCM has been continuously evaluated for face validity since the first model version was completed in 2013, including feedback regarding design and programming from clinical experts. Additionally, the conceptual model has been peer-reviewed by scientific journal editors and independent reviewers in the course of six published model applications [1-6] and one published model validity study [7].

  IHE-DCM has also been approved by Health Technology Assessment (HTA) reviewers, including the Dental and Pharmaceutical Benefits Agency (TLV) in Sweden and the Norwegian Medicines Agency (NoMA) in Norway. These objective decision-makers in each country judged IHE-DCM to be suitable for estimating cost-effectiveness for the study problems considered, implicitly (and in some cases explicitly) approving the conceptual model of IHE-DCM. NoMa writes that "Markov cohort modeling is a common and well-established method of health economic analysis. For T2DM microsimulation (simulation at the individual level) is considered more suitable because of the complexity of the disease course. Such models are often heavy to run and less transparent than cohort-level simulation. The cohort model loses some accuracy, but this is compensated for a much faster model. A shorter run time facilitates work on the validation of the model and running sensitivity analyzes. The Norwegian Medicines Agency considers that this may be appropriate cohort modeling of T2DM (faster running time), although some accuracy is lost. Lundqvist et al. demonstrated that results from the cohort model were comparable to other established models. It is also positive that all baseline patient characteristics and treatment effects (change in biomarkers such as HbA1c) can be easily changed in the model." [8, 9].

- **A2:** Has this model been compared to other conceptual models found in the literature or clinical textbooks?

  Yes. In the current manuscript.

Part B: Input Data Validation

- **B1:** Have experts been asked to judge the appropriateness of the input data?

  Yes. Although it is unclear where to draw the line between the conceptual model and input data, experts have regularly weighed in on the appropriateness of inputs used in IHE-DCM. In a broad sense, any risk equation used in IHE-DCM is "input data". In a narrow sense, "input data" can be interpreted as the values that are entered into IHE-DCM by a user to parameterize a given application (e.g., mean age of the cohort at study baseline), which differs from programmed features of the model which cannot be modified by the user without changing the source code (e.g., the macrovascular risk equations supported in the model).

  While some of the choices made in selecting risk equations may be unique to IHE-DCM, there is broad consensus among diabetes modelers about others, such as the UKPDS equations used to estimate risks of cardiovascular events and mortality.

  In the narrow sense of "input", that is, the data required to populate the simulations for this manuscript, not all inputs have been judged by independent experts.

- **B2:** When input parameters are based on regression models, have statistical tests been performed?

  Yes. Regression models as a source of risk predictions are hard-coded into model programing, but regressions models for QALY disutility values, such as the CODE-2 [10], can be customized by the model as "user inputs". Statistical assessments of the validity and adequacy of these regression models are largely available in their respective publications. The UKPDS equations used in IHE-DCM for the prediction of macrovascular risks and mortality are well-established in T2DM modeling and appropriate statistical tests were performed.

Part C: Validation of the Computerized Model

- **C1:** Has the computerized model been examined by modeling experts?

  Yes. As part of HTA submissions, independent and objective HTA boards such as NOMA (Norway) and TLV (Sweden) have reviewed the structure of IHE-DCM and its cost-effectiveness applications (see the entry for **A1**)

- **C2:** Has the model been run for specific, extreme sets of parameter values in order to detect any coding errors?

  Yes, every time a programing change is made to the model, the implementation is tested systematically and debugged by both the programmer and other staff at the Swedish Institute for Health Economics. A formal verification test of IHE-DCM has been conducted recently, where 74 "stress tests" were specified and then conducted. It was conducted by staff at the Swedish Institute for Health Economics. The results of the stress tests were consistent with expectations. An expanded formal verification including several hundred "stress test" is performed after each new model update.

- **C3:** Have patients been tracked through the model to determine whether its logic is correct?

  Yes. While IHE-DCM is a cohort model, and does not produce any output on individual patients, the code is regularly inspected, and traces of health states and biomarkers are examined for the cohort to assess correct implementation in the computerized model and plausibility of assumptions.

- **C4:** Have individual sub-modules of the computerized model been tested?

  Yes, see response to **C2**, above. Furthermore, when model improvements are made to any sub-module of IHE-DCM, the modified sub-module is inspected by running IHE-DCM with inputs that are specifically designed to isolate and investigate the effects of this sub-module on other parts of the model simulation. While these tests are being carried out regularly to verify that the IHE-DCM model calculations perform in line with expected outcomes, we do not follow a set protocol nor are these test results reported.

  Additionally, prior to model execution, IHE-DCM evaluates the model inputs for nonsensical, impermissible values (e.g., negative ages or biomarker values), and the user is prompted to adjust any offending inputs to ensure consistency with model requirements.

Part D: Operational Validation

- **D1:** Have experts been asked to judge the appropriateness of the model outcomes?

  Yes. HTA bodies have judged that the model outcomes are appropriate. NOMA (Norway), TLV (Sweden) have all reviewed the model itself and applications of the model as part of reimbursement submissions.

  Modeling guidelines were followed in the construction of IHE-DCM, including the decision of what outcomes to report. IHE-DCM reports a comprehensive set of outcomes that includes the standard set expected of DM models plus many more to enable a full examination of model results. By providing this level of transparency, we enable all concerned experts to assess appropriateness of the model for a broad range of outcomes, including for example,

  - Summary measures of health economic endpoints, including, costs, LYs, survival, QALYs, ICERs, CEACs, scatterplots
  - Cumulative incidence of individual micro- and macrovascular complications, survival
  - Simulated biomarker evolution over time (e.g., HbA1c, SBP, BMI, cholesterol, eGFR, heart rate)
  - Adverse event rates

- **D2:** Have the model outcomes been compared to the outcomes of other models that address similar problems?

  Yes. In the current manuscript.

- **D3:** Have the model outcomes been compared to the outcomes obtained when using alternative input data?

  Yes. As part of a recent verification (See C2), results were generated for each of the four CVD risk equations supported in IHE-DCM, UKPDS 82 [11],UKPDS [12] , Swedish NDR [13], and Australian FDS [14]), and the results were compared and contrasted.

- **D4:** Have the model outcomes been compared to empirical data?

  Yes, the external validation of the model was formally tested by replicating and simulating 12 clinical trials (both interventional and non-interventional) and comparing 167 predicted microvascular, macrovascular, and mortality outcomes with those observed in the actual trials. Concordance was examined using visual and regression-based analysis of the resulting scatterplots (where an intercept of 0 and a slope of 1 indicate perfect concordance). Additional analyses were conducted on sub-groups of outcomes, including 'dependent' vs. 'independent' endpoints and microvascular vs. macrovascular vs. mortality endpoints to assess the performance of individual model components. The validation study also evaluated the impact of using alternative macrovascular and mortality risk equations.

  Visual inspection indicated that the model predicted outcomes well using each of the three competing sets of macrovascular risk equations. The regression coefficients indicated that the best-fitting regression line for UKPDS 68 was almost perfectly concordant with observed values (slope 0.996), whereas there was a slight tendency for the Swedish NDR risk equations (0.952) and the UKPDS 82 equations to underestimate (0.899) actual outcomes. The R2 values were uniformly high (>0.96). There were no major differences between 'dependent' and 'independent' outcomes, nor were there major differences for microvascular and mortality outcomes. [7]. The external validation indicates that the IHE Diabetes Cohort Model is associated with predictive accuracy in line with other models of T2DM, indicating that the trade-off in accuracy for using the cohort approach vs. micro-simulation might not be that large.

Part E: Other Validation Techniques

- **E1:** Have any other validation techniques been performed?

  The formal validation of IHE-DCM included examination of face validity, debugging and stress testing, cross-validation, and dependent and independent validation. Back-of-the-envelope calculations is regularly used to evaluate the face validity of model updates, and double programming is used to double-check the IHE-DCM implementation of risk equations.

References

1. Ericsson, Å. and A. Lundqvist, *Cost Effectiveness of Insulin Degludec Plus Liraglutide (IDegLira) in a Fixed Combination for Uncontrolled Type 2 Diabetes Mellitus in Sweden.* Applied Health Economics and Health Policy, 2017. **15**(2): p. 237-248.
2. Ericsson, A. and A. Fridhammar, *Cost-effectiveness of once-weekly semaglutide versus dulaglutide and lixisenatide in patients with type 2 diabetes with inadequate glycemic control in Sweden.* J Med Econ, 2019. **22**(10): p. 997-1005.
3. Johansen, P., et al., *Cost Effectiveness of Once-Weekly Semaglutide Versus Once-Weekly Dulaglutide in the Treatment of Type 2 Diabetes in Canada.* Pharmacoecon Open, 2019. **3**(4): p. 537-550.
4. Ericsson, A., et al., *Cost-effectiveness of liraglutide versus lixisenatide as add-on therapies to basal insulin in type 2 diabetes.* PLoS One, 2018. **13**(2): p. e0191953.
5. Steen Carlsson, K. and U. Persson, *Cost-effectiveness of add-on treatments to metformin in a Swedish setting: liraglutide vs sulphonylurea or sitagplitin.* J Med Econ, 2014. **17**(9): p. 658-69.
6. Kiadaliri, A.A., et al., *Cost-Utility Analysis of Glucagon-Like Peptide-1 Agonists Compared with Dipeptidyl Peptidase-4 Inhibitors or Neutral Protamine Hagedorn Basal Insulin as Add-On to Metformin in Type 2 Diabetes in Sweden.* Diabetes Ther, 2014.
7. Lundqvist, A., et al., *Validation of the IHE Cohort Model of Type 2 Diabetes and the Impact of Choice of Macrovascular Risk Equations.* PLOS ONE, 2014.
8. The Dental and Pharmaceutical Benefits Agency (TLV). *Luxusia is included in the reimbursement system with limitation (Lyxumia ingår i högkostnadsskyddet med begränsning)*. [2018-01-12]; Available from: https://www.tlv.se/beslut/beslut-lakemedel/begransad-subvention/Lyxumia-ingar-i-hogkostnadsskyddet-med-begransning/.
9. The Norwegian Medicines Agency (NoMA). *Ozempic (semaglutide) for the treatment of type 2 diabetes mellitus*. [2019-12-19]]; Available from: https://legemiddelverket.no/Documents/Offentlig%20finansiering%20og%20pris/Metodevurderinger/O/Ozempic_T2D_2019.pdf.
10. Bagust, A. and S. Beale, *Modelling EuroQol Health-Related Utility Values for Diabetic Complications from CODE-2 Data.* Health Economics, 2005. **14**(3): p. 217-30.
11. Hayes, A.J., et al., *UKPDS outcomes model 2: a new version of a model to simulate lifetime health outcomes of patients with type 2 diabetes mellitus using data from the 30 year United Kingdom Prospective Diabetes Study: UKPDS 82.* Diabetologia, 2013. **56**(9): p. 1925-33.
12. Clarke, P.M., et al., *A model to Estimate the Lifetime Health Outcomes of Patients with Type 2 Diabetes: the United Kingdom Prospective Diabetes Study (UKPDS) Outcomes Model (UKPDS no. 68).* Diabetologia, 2004. **47**(10): p. 1747-59.
13. Ahmad Kiadaliri, A., et al., *Towards renewed health economic simulation of type 2 diabetes: risk equations for first and second cardiovascular events from Swedish register data.* PLOS ONE, 2013. **8**(5): p. e62650.
14. Davis, W.A., M.W. Knuiman, and T.M. Davis, *An Australian cardiovascular risk equation for type 2 diabetes: the Fremantle Diabetes Study.* Intern Med J, 2010. **40**(4): p. 286-92.