Comment on "External Validation of the Core Obesity Model to Assess the Cost-Effectiveness of Weight Management Interventions"

Ramaekers, Bram L. T.; Wijnen, Ben; Armstrong, Nigel; Petersohn, Svenja; Feenstra, Talitha; Wang, Junfeng; Joore, Manuela A.

Published in:
Pharmacoeconomics

DOI:
10.1007/s40273-020-00985-5

Publication date:
2021

PharmacoEconomics (2021) 39:133–135
https://doi.org/10.1007/s40273-020-00985-5

LETTER TO THE EDITOR

Comment on "External Validation of the Core Obesity Model to Assess the Cost-Effectiveness of Weight Management Interventions"

Bram L. T. Ramaekers[1,2] · Ben Wijnen[1,2,3] · Nigel Armstrong[4] · Svenja Petersohn[1,2,5] · Talitha Feenstra[6,7] · Junfeng Wang[6,8] · Manuela A. Joore[1,2]

Accepted: 10 November 2020 / Published online: 30 November 2020

Lopes et al. [1] compared predictions made by the Core Obesity Model (COM) with data used and not used to develop the COM. This was defined by the authors as dependent and independent external validation. The COM, a cohort state-transition model, was developed to assess the cost-effectiveness of weight management interventions. The validation with external data was performed to ensure that model predictions can be interpreted with confidence by clinicians, budget holders, and other stakeholders. The authors concluded that the COM predictions showed good concordance in both dependent and independent validations, suggesting the model is suitable for decision-making. The predictions of cardiovascular events and mortality were mentioned as key areas for future refinement of the COM.

We would like to compliment Lopes et al. [1] for publishing this relevant paper and believe the development of multi-use disease (or reference/generic) models is an excellent initiative as it ensures consistency between decisions (e.g. for different weight management interventions) and potentially improves model validity and transparency [2]. Additionally, we agree that thorough validation of models should be carried out to enhance confidence in their predictions. Despite validation against external data being an essential part of the validation process, publications of these validation exercises are often lacking, which makes it difficult to trust the model outcomes [3]. Recently, the COM was used in a single technology assessment (STA) by the National Institute for Health and Care Excellence (NICE) [4]. This STA considered the (cost-)effectiveness of liraglutide for managing being overweight or obese. As members of the evidence review group during this STA and/or experienced modellers in this disease area, we believe that commenting on this validation against external data is warranted, highlighting our main concerns and suggestions for improvement.

The publication considered the COM version 6.1, while version 7.0 was used in the liraglutide STA (submitted to NICE in June 2019), and results produced by version 10.0 were recently presented during the second NICE committee meeting (September 2020). Notably, the various updates from version 7.0 to 10.0 included fixing errors related to the verification/internal validity of the computerised model, which were identified during this STA and are unlikely to be only applicable to the liraglutide case (see the evidence review group [ERG] report and technical engagement responses from Novo Nordisk Ltd., Appendix 1 [4]). This

This comment refers to the article available online at https://doi.org/10.1007/s40273-020-00941-3.

The reply to this comment is available online at https://doi.org/10.1007/s40273-020-00982-8.

✉ Bram L. T. Ramaekers
bram.ramaekers@mumc.nl

1 Department of Clinical Epidemiology and Medical Technology Assessment (KEMTA), Maastricht University Medical Center+, Maastricht, The Netherlands

2 Care and Public Health Research Institute, Maastricht University, Maastricht, The Netherlands

3 Center for Economic Evaluation and Machine Learning, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, The Netherlands

4 Kleijnen Systematic Reviews Ltd, York, UK

5 Pharmerit International, Rotterdam, The Netherlands

6 Centre for Nutrition, Prevention and Health Services Research, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

7 Faculty of Science and Engineering, Groningen Research Institute of Pharmacy, University of Groningen, Groningen, The Netherlands

8 Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

included errors related to adjustments for the cycle length and the calculated risks for developing type 2 diabetes (T2DM) as well as cardiovascular events, potentially affecting the degree of correspondence between model predictions and external data. Therefore, it is unclear to what extent the conclusions drawn from the published validation against external data (based on the COM version 6.1) are applicable to the corrected COM version 10. To avoid ambiguity in the interpretation of the validation against external data, the assessment of face validity and verification of model code should preferably be reported in more detail and/or be performed more explicitly, especially since tools such as AdVishE [5] and TECH-VER [6] are available.

Although the validation seemed to be performed very carefully, the interpretation of the assessment might require additional discussion. In particular, what is a "slight" or "acceptable" over- or underestimation of the observed outcomes? Figure 3, plotting the observed versus predicted results from the independent validations indicated that the COM underestimated events. More specifically, the estimated slope coefficient (with a debatable assumption of no intercept) indicated risk reductions of 19% and 16% for cardiovascular plus mortality outcomes and T2DM incidence, respectively. Although potential reasons for the underestimations have been discussed by the authors, it is still concluded that there is a good concordance between observed and predicted outcomes. While it would remain arbitrary to define cut-offs for "acceptable", this uncertainty resulting from the validity assessment should ideally be reflected in the uncertainty analyses when using the COM for decision-making, for instance by recommending scenario analyses with increased occurrence of events or by parameterising this uncertainty [7].

Various risk models were used to model long-term outcomes such as cardiovascular disease, mortality, and T2DM based on surrogate outcomes such as body mass index, blood pressure, cholesterol, and glycated haemoglobin. Although extrapolating surrogate to long-term outcomes is a strength of decision-analytic models, the uncertainty associated with the extrapolation should ideally be explored (either in probabilistic sensitivity analyses by parameterising this structural uncertainty or in scenario analyses). Notably, the (uncertainty due to the) selection of the risk models could be debated, given that different risk models were used for estimating cardiovascular events in patients with and without T2DM. This could be considered suboptimal as using different risk models dependent on T2DM status might "introduce bias in terms of rates of disease progression when these are dependent on the study and the population informing the model rather than on the stage of disease" as highlighted in a recent review of prediabetes decision models [8]. Alternatively, the QRISK®3 risk model [9] could be used to estimate primary cardiovascular events and the Framingham risk model [10] could be used for secondary cardiovascular events (independent of T2DM status). In addition to this, as highlighted during the first NICE committee meeting for liraglutide [11], it is unclear whether the risk models used to extrapolate surrogate to long-term outcomes are appropriate for making predictions based on temporary changes in patient characteristics due to, for instance, weight management interventions that the COM aims to assess, particularly given that these risk models are developed to make predictions for patients in a relatively "steady state" (see also appraisal consultation document, Sect. 3.10 [11]). Although there is no clear alternative, this adds to the uncertainty, which should be acknowledged, explored, and preferably supported by additional evidence (potentially on a case-by-case basis) showing that temporary changes in surrogate outcomes translate into long-term outcomes.

In conclusion, we hope that our suggestions will provide additional insights regarding the validation of the COM as described by Lopes et al. [1], which may help the adoption of an otherwise potentially useful tool in the field of obesity.

Declarations

Funding None.

Conflict of interest Bram Ramaekers, Ben Wijnen, Nigel Armstrong, Svenja Petersohn, Talitha Feenstra, Junfeng Wang, and Manuela Joore have no conflicts of interest to declare.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Code availability Not applicable.

Author contributions All authors have commented on the submitted manuscript and have given their approval for the final version to be published.

References

1. Lopes S, Johansen P, Lamotte M, McEwan P, Olivieri AV, Foos V. External validation of the Core Obesity Model to assess the cost-effectiveness of weight management interventions. Pharmacoeconomics. 2020;38(10):1123–33.
2. Sampson CJ, Arnold R, Bryan S, Clarke P, Ekins S, Hatswell A, et al. Transparency in decision modelling: what, why, who and how? Pharmacoeconomics. 2019;37(11):1355–69.
3. Ghabri S, Stevenson M, Moller J, Caro JJ. Trusting the results of model-based economic analyses: is there a pragmatic validation solution? Pharmacoeconomics. 2019;37(1):1–6.
4. Single Technology Appraisal Liraglutide for managing overweight and obesity [ID740]: Committee Papers. 2020 [cited 6-10-2020].

https://www.nice.org.uk/guidance/gid-ta10388/documents/committee-papers.
5. Vemer P, Corro Ramos I, van Voorn GA, Al MJ, Feenstra TL. AdViSHE: a validation-assessment tool of health-economic models for decision makers and model users. Pharmacoeconomics. 2016;34(4):349–61.
6. Buyukkaramikli NC, Rutten-van Molken M, Severens JL, Al M. TECH-VER: a verification checklist to reduce errors in models and improve their credibility. Pharmacoeconomics. 2019;37(11):1391–408.
7. Corro Ramos I, van Voorn GAK, Vemer P, Feenstra TL, Al MJ. A new statistical method to determine the degree of validity of health economic model outcomes against empirical data. Value Health. 2017;20(8):1041–7.
8. Leal J, Morrow LM, Khurshid W, Pagano E, Feenstra T. Decision models of prediabetes populations: a systematic review. Diabetes ObesMetab. 2019;21(7):1558–69.
9. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. BMJ. 2017;23(357):j2099.
10. D'Agostino RB, Russell MW, Huse DM, Ellison RC, Silbershatz H, Wilson PW, et al. Primary and subsequent coronary risk appraisal: new results from the Framingham study. Am Heart J. 2000;139(2 Pt 1):272–81.
11. Single Technology Appraisal Liraglutide for managing overweight and obesity [ID740]: Appraisal consultation document. 2020 [cited 6-10-2020]. https://www.nice.org.uk/guidance/gid-ta10388/documents/129.