Maastricht University | Maastricht UMC+ | CAPHRI Care and Public Health Research Institute

Faculty of Health, Medicine and Life Sciences

EE149

Uncertainty Assessment in Economic Evaluations of Artificial Intelligence-based Health Technologies: Pitfalls and Recommendations

Mabel Wieman[A], Bram Ramaekers[A], Laure Wynants[A], Inigo Bermejo[B], Andrea Gabrio[A], Nigel Armstrong[C], Marie Westwood[D], Manuela Joore[A] & Sabine Grimm[A]

[A]Maastricht University, Maastricht, Netherlands, [B]University of Hasselt, Diepenbeek, Belgium, [C]Kleijnen Systematic Reviews Ltd., York, UK, [D]University of Bristol, Bristol, UK

Introduction

Despite **Artificial Intelligence** (AI)'s growing potential in clinical care, **inadequate evidence** on (cost-)effectiveness often hinders adoption. [1,2] For successful clinical implementation and societal impact, **robust economic evaluations** (EEs) with **uncertainty assessment**, are essential. [3]

Aim

To **identify common uncertainties** in EEs of AI-based health technologies used in clinical care. Explore how these uncertainties are **currently assessed** in existing model-based EEs. **Formulate recommendations** for practice and research.

Understanding Uncertainty in EEs of AI

Three uncertainties within EEs of AI were defined: (1) Transportability, (2) Human-AI collaboration, and (3) Performance dynamics. All three are caused by **unavailability** and/or **indirectness** of the evidence and can manifest in multiple **model aspects.** [3-6]

Existing EEs **occasionally addressed** transportability and human-AI collaboration, but not performance dynamics. [7,8]

TRANSPORTABILITY = AI performance can be affected by differences between the target setting and the development setting

HUMAN-AI COLLABORATION = AI performance is often directly compared with that of humans, but in real-life human-in-the-loop systems are more common

PERFORMANCE DYNAMICS = AI performance can change over time due to model drift and updates

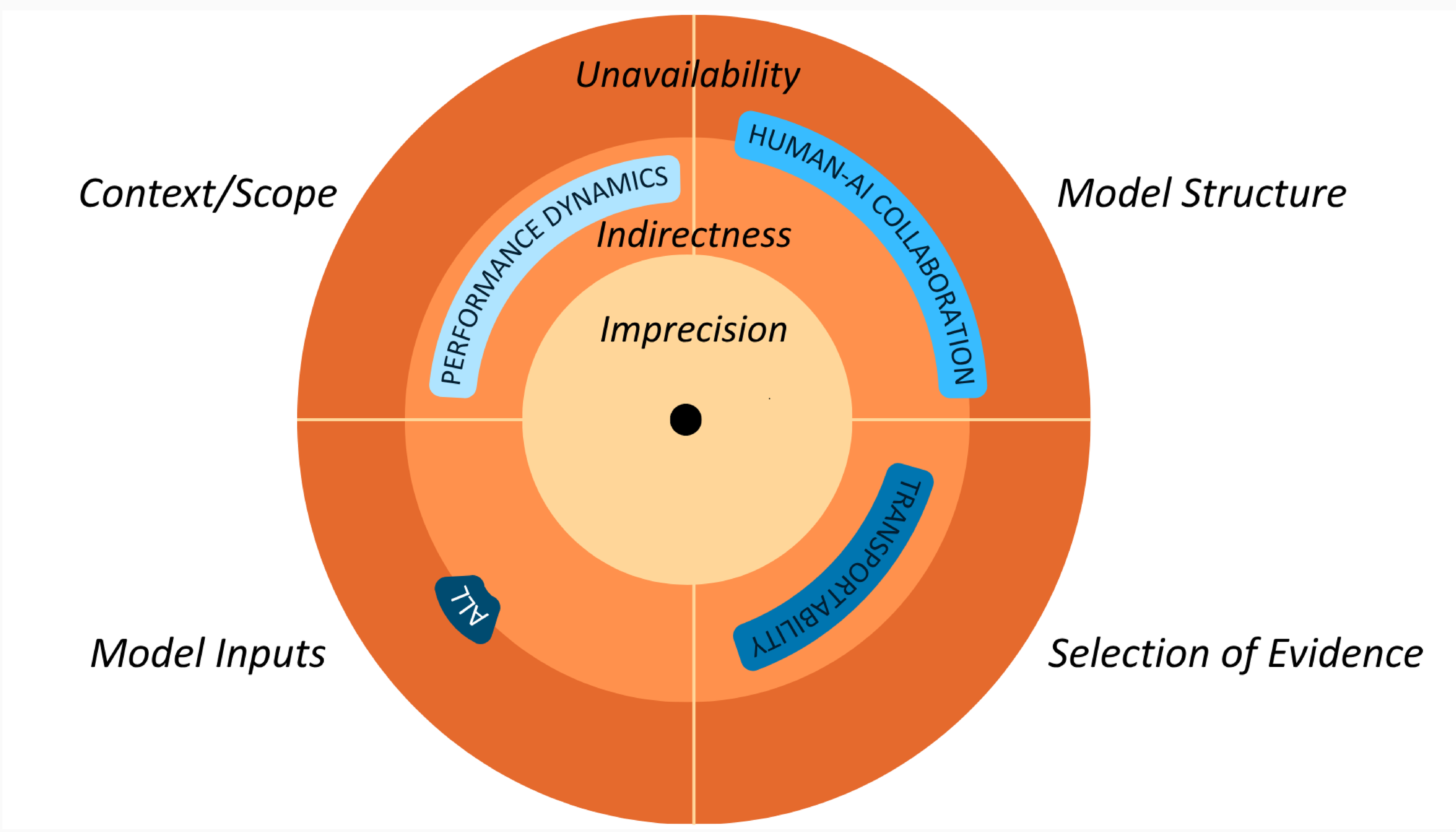

Figure 1: Understanding uncertainty in AI-based health technologies: multiple dimensions

Recommendations for future EEs of AI

We give **general** approaches for **identifying and analyzing** all uncertainties, and **specific** methods tailored to the three common uncertainties, based on literature on existing EEs and uncertainty assessment methods. [7-9]

TRANSPORTABILITY
- Random effect meta-analysis

HUMAN-AI COLLABORATION
- Reliance discrepancy terms

PERFORMANCE DYNAMICS
- Life-cycle assessment methods including VOI analysis

ALL
- Systematic uncertainty identification
  - TRUST
  - GRADE, PROBAST+AI, QUADAS-2 or APPRAISE
- Structured expert elicitation
- Discrepancy approaches
- Scenario analysis & model averaging

Abbreviations: TRUST = Transparent Uncertainty Assessment Tool, GRADE = Grading of Recommendations Assessment, Development and Evaluation, PROBAST+AI = Prediction model Risk Of Bias Assessment Tool, QUADAS-2 = Quality Assessment of Diagnostic Accuracy Studies

Conclusion

- AI-based health technologies have the potential to transform the health care sector, but uncertainties within model-based EEs, often caused by a **lack of context-specific evidence,** need to be appropriately managed.
- We have developed recommendations for uncertainty identification and analysis for assessing **transportability, human-AI collaboration, and performance dynamics.**
- Further research is needed to **apply and further refine these methods** in future EEs of AI-based health technologies.

References

[1]Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med*. Oct 29 2019;17(1):195. doi:10.1186/s12916-019-1426-2

[2]Hendrix N, Veenstra DL, Cheng M, Anderson NC, Verguet S. Assessing the Economic Value of Clinical Artificial Intelligence: Challenges and Opportunities. *Value Health*. Mar 2022;25(3):331-339. doi:10.1016/j.jval.2021.08.015

[3]Grimm SE, Pouwels X, Ramaekers BLT, et al. Development and Validation of the TRansparent Uncertainty ASsessmenT (TRUST) Tool for Assessing Uncertainties in Health Economic Decision Models. *Pharmacoeconomics*. Feb 2020;38(2):205-216. doi:10.1007/s40273-019-00855-9

[4]Van Calster B, Steyerberg EW, Wynants L, van Smeden M. There is no such thing as a validated prediction model. *BMC Medicine*. 2023/02/24 2023;21(1):70. doi:10.1186/s12916-023-02779-w

[5]Reverberi C, Rigon T, Solari A, et al. Experimental evidence of effective human–AI collaboration in medical decision-making. *Scientific Reports*. 2022/09/02 2022;12(1):14952. doi:10.1038/s41598-022-18751-2

[6]Vela D, Sharp A, Zhang R, Nguyen T, Hoang A, Pianykh OS. Temporal quality degradation in AI models. *Scientific Reports*. 2022/07/08 2022;12(1):11654. doi:10.1038/s41598-022-15245-z

[7]Vithlani J, Hawksworth C, Elvidge J, Ayiku L, Dawoud D. Economic evaluations of artificial intelligence-based healthcare interventions: a systematic literature review of best practices in their conduct and reporting. *Front Pharmacol*. 2023;14:1220950. doi:10.3389/fphar.2023.1220950

[8]Voets MM, Veltman J, Slump CH, Siesling S, Koffijberg H. Systematic Review of Health Economic Evaluations Focused on Artificial Intelligence in Healthcare: The Tortoise and the Cheetah. *Value Health*. Mar 2022;25(3):340-349. doi:10.1016/j.jval.2021.11.1362

[9]Otten TM, Grimm SE, Ramaekers B, Joore MA. Comprehensive Review of Methods to Assess Uncertainty in Health Economic Evaluations. *Pharmacoeconomics*. Jun 2023;41(6):619-632. doi:10.1007/s40273-023-01242-1

Correspondence to:
Mabel Wieman, Maastricht University/MUMC+

mabel.wieman@mumc.nl
www.mumc.nl/research/kemta

Dept of Clinical Epidemiology and Medical Technology Assessment

T +3143 387 5559

Gaetano Martinolaan 85, niveau 7/8
Postadres: Postbus 5800 | 6202 AZ Maastricht

CAPHRI Care and Public Health Research Institute