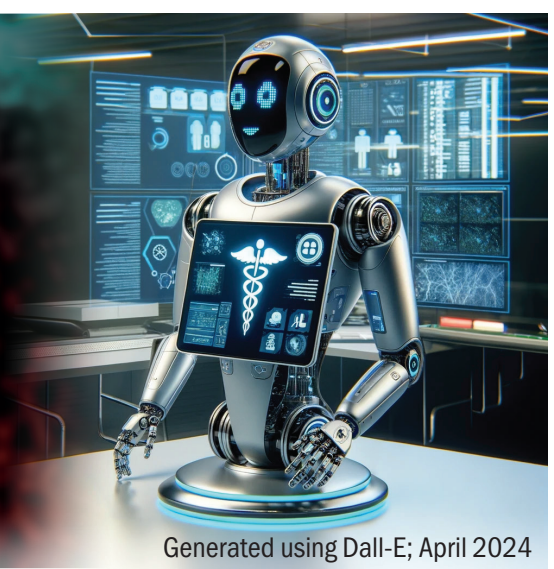


Bridging expertise: A classifier algorithm’s journey through HEOR publication identification

Rosie Morland^{a*}, Lynne Cairns^a, Juliet Bell^{a*}, Eilish McBurnie^a, Alissa Epworth^b, Romney Adams^b, Yunyu Huang^c, and Remon van den Broek^{d*}

^a Excerpta Medica, London, UK | ^b DistillerSR, Ottawa, ON, Canada | ^c Excerpta Medica, Amstelveen, The Netherlands | ^d Adelphi Group, Amstelveen, The Netherlands



Generated using Dall-E; April 2024



Scan the QR code or [click here](#) to learn more

Background:

- Reviewing literature can be time-consuming**
 - This can negatively impact the accurate inclusion of HEOR publications
- Human error rate can be 21%**
 - Studies have shown human error can be common due to review fatigue¹ and inter-reviewer variability².
- Health Economics and Outcomes Research (HEOR) is a wide-ranging, descriptive, term**
 - Definitions and understanding of the term can vary
- Machine-learning classifiers (algorithms that categorize data) could assist**
 - Through successful identification of HEOR publications for inclusion

Aims: To train and test a classifier algorithm that can identify HEOR publications based on title and abstract, using the DistillerSR literature review platform

Developing a definition for Health Economic and Outcomes Research

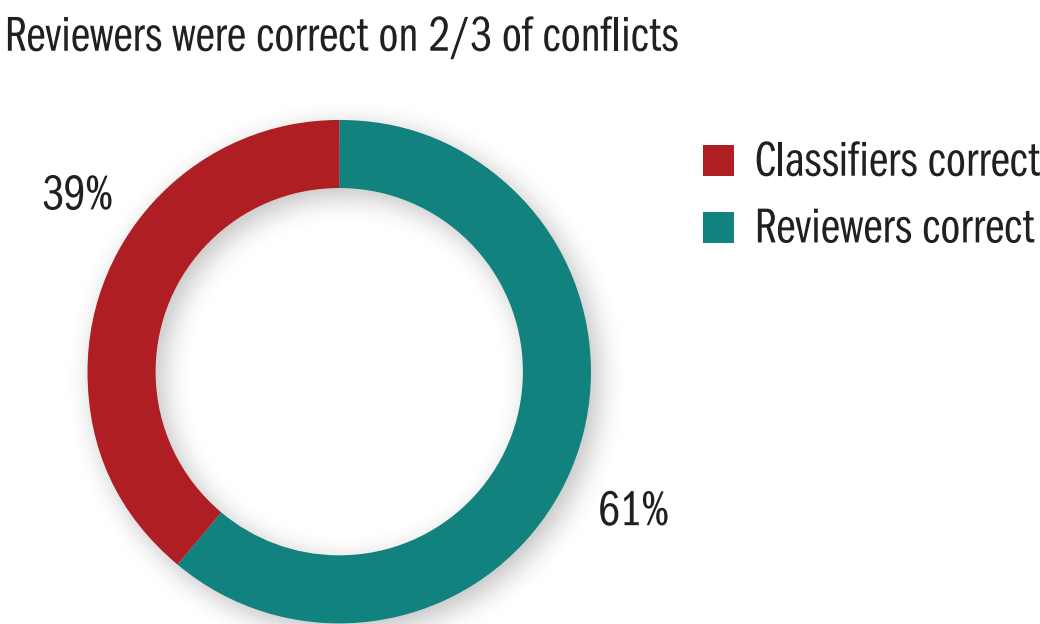
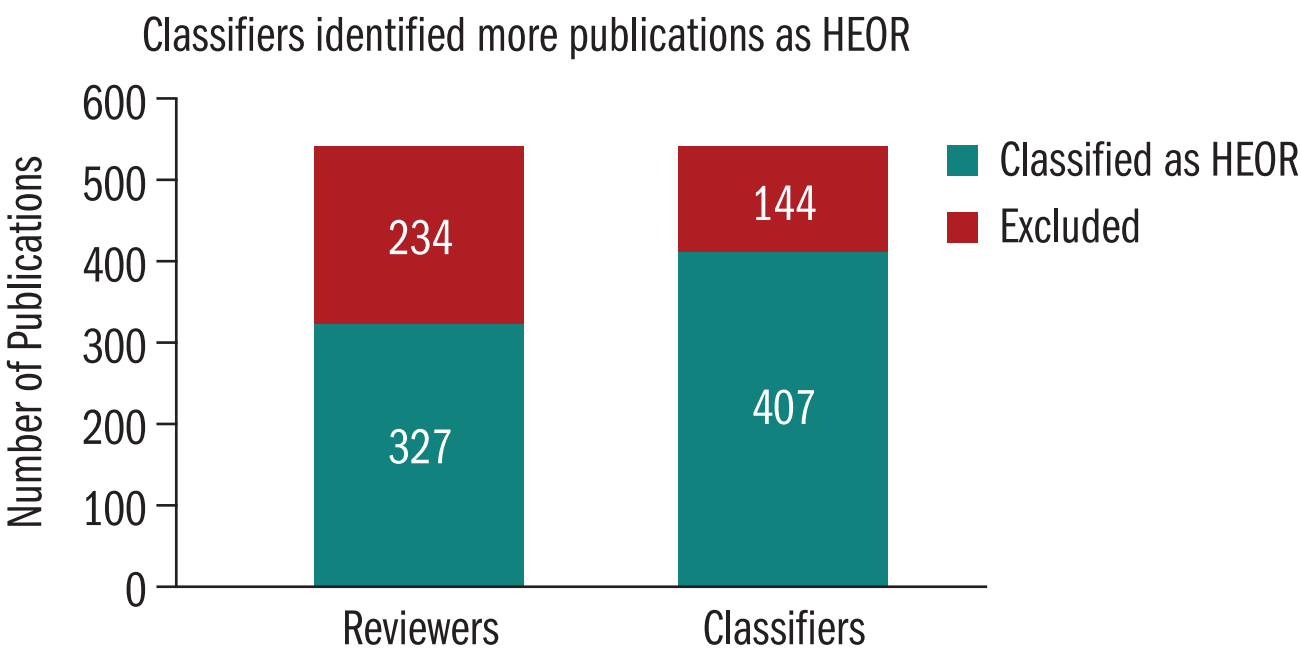


- Health-economics: The use of resources to efficiently improve the health of a population*^{3,4}
- Outcome research: Any research that attempts to measure the effectiveness, efficiency, equality and patient-centeredness of healthcare*^{5,6}
- Based on the AMCP definition of outcomes research, we provided further HEOR guidance to reviewers based on 3 main types:⁶
 - Humanistic – patient-/caregiver-reported outcomes
 - Clinical – real-world or statistical comparison of treatments, e.g. indirect treatment comparisons
 - Economic – cost-effectiveness evaluations
- The type of study was not important, only whether HEOR outcomes were reported
- Under this definition, any randomized controlled trial that reported PROs was included

Results:

Expert-trained classifiers were able to identify publications reporting HEOR outcomes using the abstract and title alone.

Reviewers tended towards exclusion (higher false negatives), whereas the classifiers tended towards inclusion (lower false negatives).



Methods:

References imported to build database

- Embase and Pubmed
- DistillerSR literature review software was populated with HEOR and non-HEOR publications from two clinical indications
- Reviews and publications without an abstract were excluded

TRAIN database developed

- Experienced reviewers (n = 7) classified publications as HEOR or not
- “Is HEOR data reported in the title/abstract? Answer: “yes” or “no””
- Conflicts were resolved via online/verbal discussion N = 245

Classifier trained

- Binary classifier algorithm ‘learns’ from experienced reviewers
- The classifier was trained with unconflicted data

TEST database developed

- The trained classifier was applied to TEST
- Reviewers without HEOR experience (n = 5) review the same publications as the classifier N = 551

Analyse

- Review conflicts in TEST database
- Calculate false negative and false positive rate
- Who was ‘correct’, based on the definition provided?
 - This was used to examine false negative (e.g. incorrectly excluded) and false positive (e.g. incorrectly included) characteristics
- Descriptive statistics are reported

Conclusions:

- Low false-negatives are desirable** – it’s better to include too many at screening than be too rigorous and exclude relevant publications
- Classifier algorithms can play a role in **improving the efficiency and reliability of literature review** development improve timelines and accuracy during publication development
- Future analyses will refine the classification and subtyping system that defines HEOR publications

Disclosures:

Rosie Morland, Lynne Cairns, Juliet Bell, Eilish McBurnie, Yunyu Huang, and Remon van den Broek report no disclosures. Alissa Epworth and Romney Adams report employment with DistillerSR

Acknowledgements:

Acknowledgements: We would like to thank our colleagues who contributed as ‘Reviewers without HEOR experience’ (Emma Butterworth, Dali Ma, Dale Owens, Eliana Portilla, and Adam Sellers), and our Studio team.

References:

- Wang Z, Nayfeh T, Tetzlaff J, O’Blenis P, Murad MH. Error rates of human reviewers during abstract screening in systematic reviews. PLOS ONE (2020), 0.1371/journal.pone.0227742
- Belur J, Tompson L, Thornton A., Simon M. Reviewer fatigue & inter-reviewer reliability: (2021). Interrater Reliability in Systematic Review Methodology: Exploring Variation in Coder Decision-Making. Sociological Methods & Research, 50(2), 837-865. <https://doi.org/10.1177/0049124118799372>
- CDC Podcast on Economic Evaluation; available from: https://www.cdc.gov/dhdsr/programs/spha/economic_evaluation/docs/podcast_1.pdf
- National Institute for Health and Care Excellence. Methods for the development of NICE public health guidance (third edition). Published 29/09/2012. Available from: <https://www.nice.org.uk/process/pmg4/chapter/incorporating-health-economics>
- Horgan B. Health Care Reform Bill Contains Opportunities for Integrative Medicine. EXPLORE (2010). 6(4):221-224.
- Weeda ER, Plich NA, Annemans L. Chaper 20 – Outcomes Research. Clinical Pharmacy Education, Practice, and Research. (2019)289-301. Academy of Managed Care. Outcomes Research. Last updated 18/7/2019. Available from: <https://www.amcp.org/about/managed-care-pharmacy-101/concepts-managed-care-pharmacy/outcomes-research>