



Augmenting Expertise: A Classifier Algorithm’s Ability to Identify and Categorize Health Economics and Outcomes

Rosie Morland¹, Lynne Cairns¹, Juliet Bell¹, Eilish McBurnie¹, Alissa Epworth², Romney Adams², Yunyu Huang³, and Remon van den Broek⁴
¹Excerpta Medica, London, UK | ²DistillerSR, Ottawa, ON, Canada | ³Excerpta Medica, Amstelveen, Netherlands | ⁴Adelphi Group, Amstelveen, Netherlands

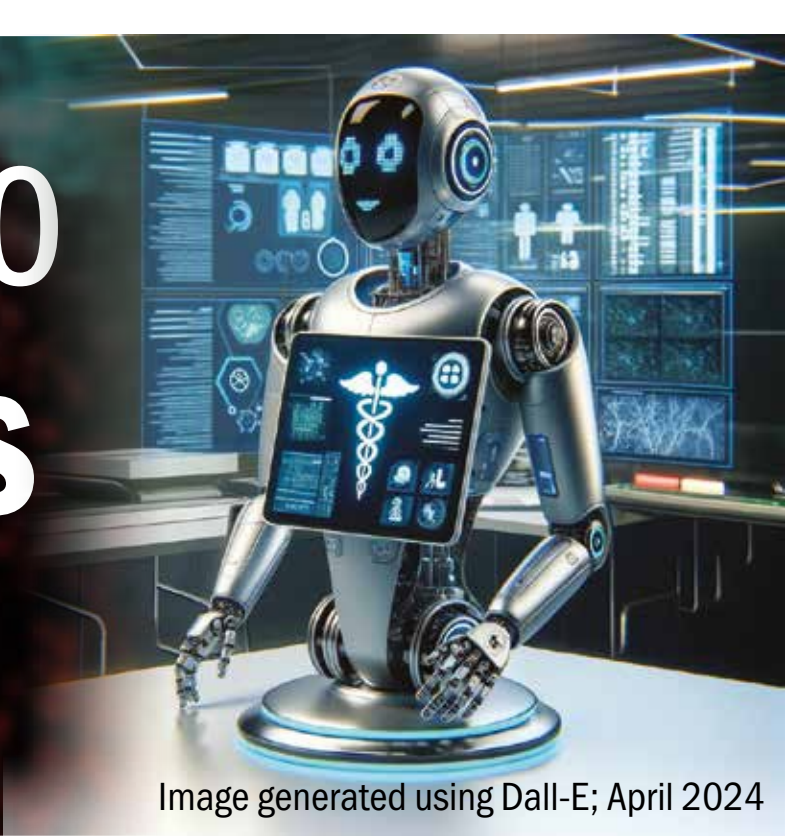


Image generated using Dall-E; April 2024

Objective

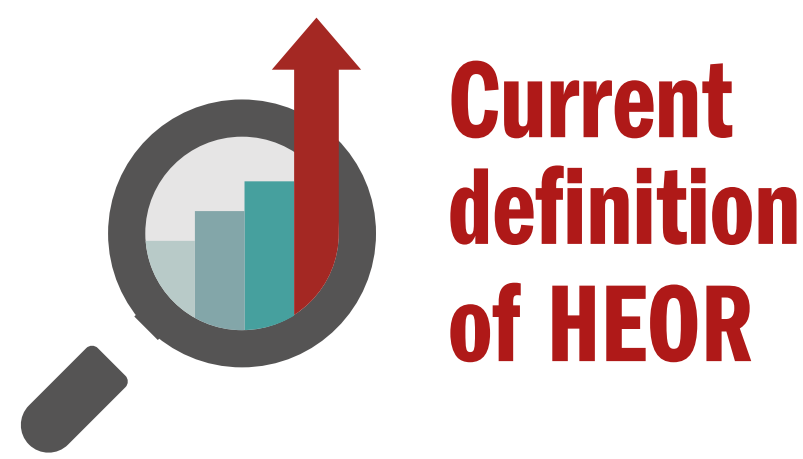
To test a classifier algorithm trained to identify HEOR publications based on title and abstract, using the DistillerSR literature review platform

Background

- ✓ **Reviewing literature can be time-consuming**
This can negatively impact the accurate inclusion of Health Economics and Outcomes Research (HEOR) publications
- ✓ **Human error rate can be 21%**
Studies have shown that human error can be common due to reviewer fatigue¹ and inter-reviewer variability²
- ✓ **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

Conclusions

- **Low false-negatives are desirable** – it’s better to include more hits at screening than risk excluding relevant publications
- Classifier algorithms can **improve the efficiency and reliability of literature review** development improve the efficiency and timelines of literature review development
- Future analyses will work on improving the lexicon of “HEOR” search terms for higher specificity



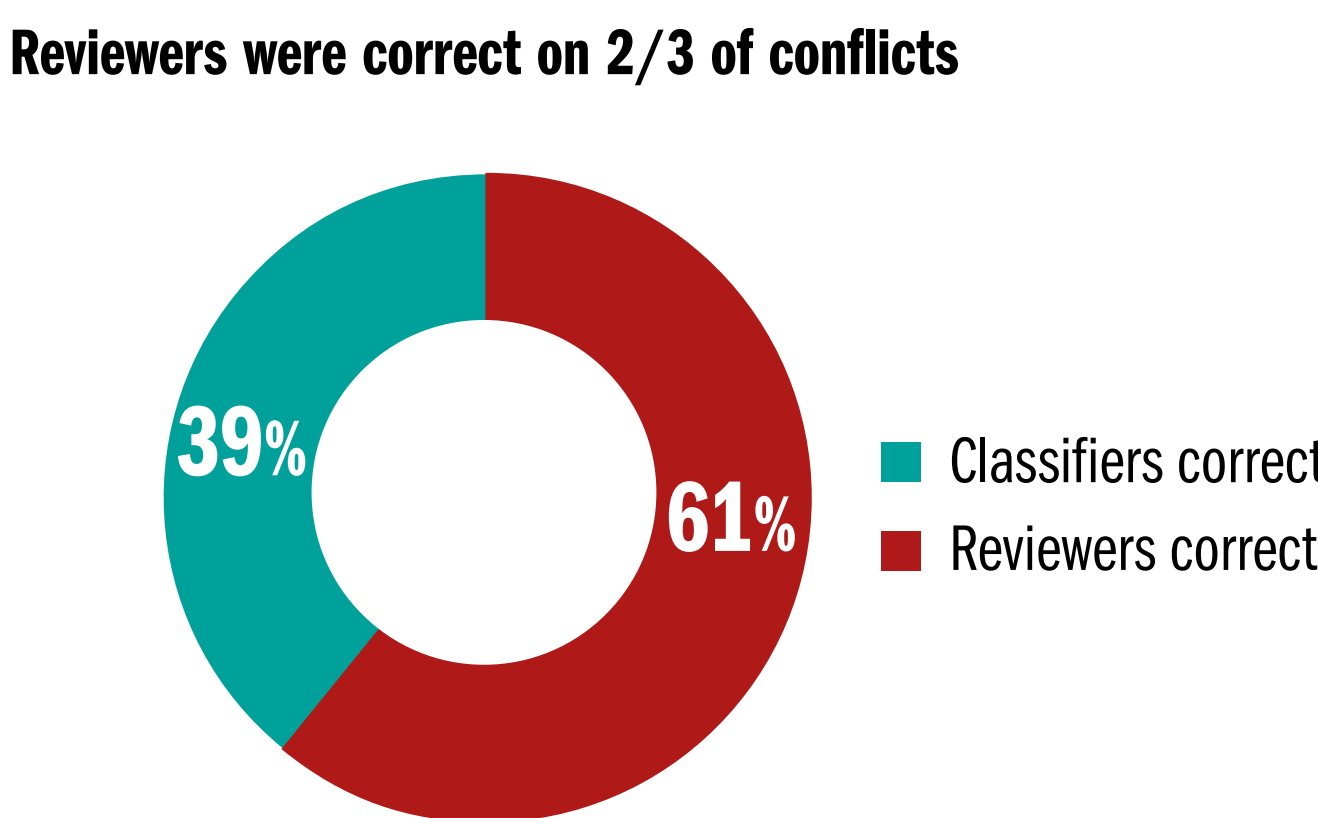
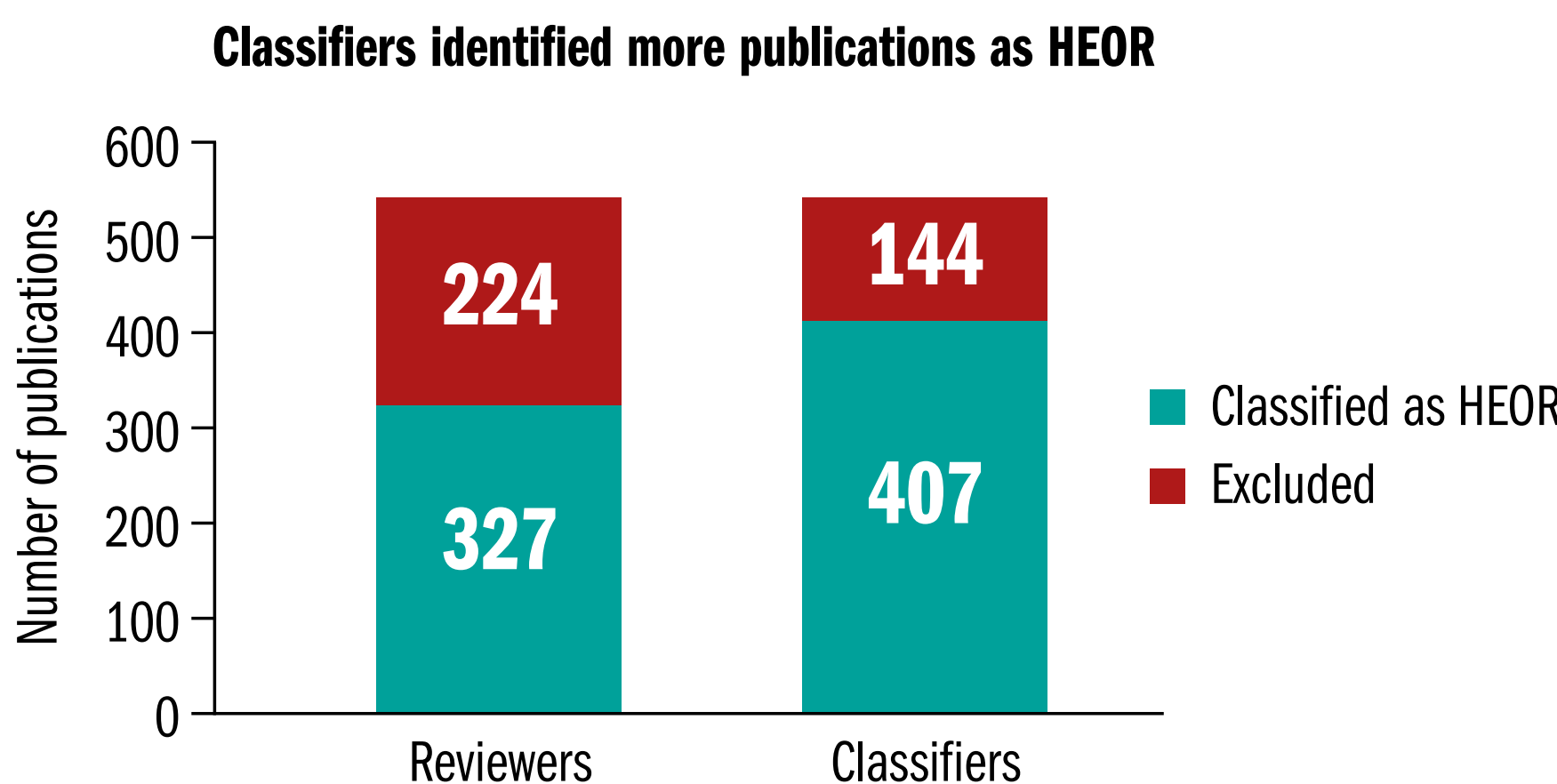
Current definition of HEOR

- 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-centredness of healthcare^{5,6}
- Based on the Academy of Managed Care Pharmacy definition of outcomes research, we provided further HEOR guidance to reviewers based on 3 main types:⁶
 - Humanistic – patient-/caregiver-reported outcomes (PROs/CROs)
 - 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

The term “HEOR” has evolved over decades and can be interpreted differently depending on use, e.g. demonstrating cost-effectiveness for payers, or improving quality of life for patients; therefore the definition of HEOR can be vague and can vary depending on its context.

Results

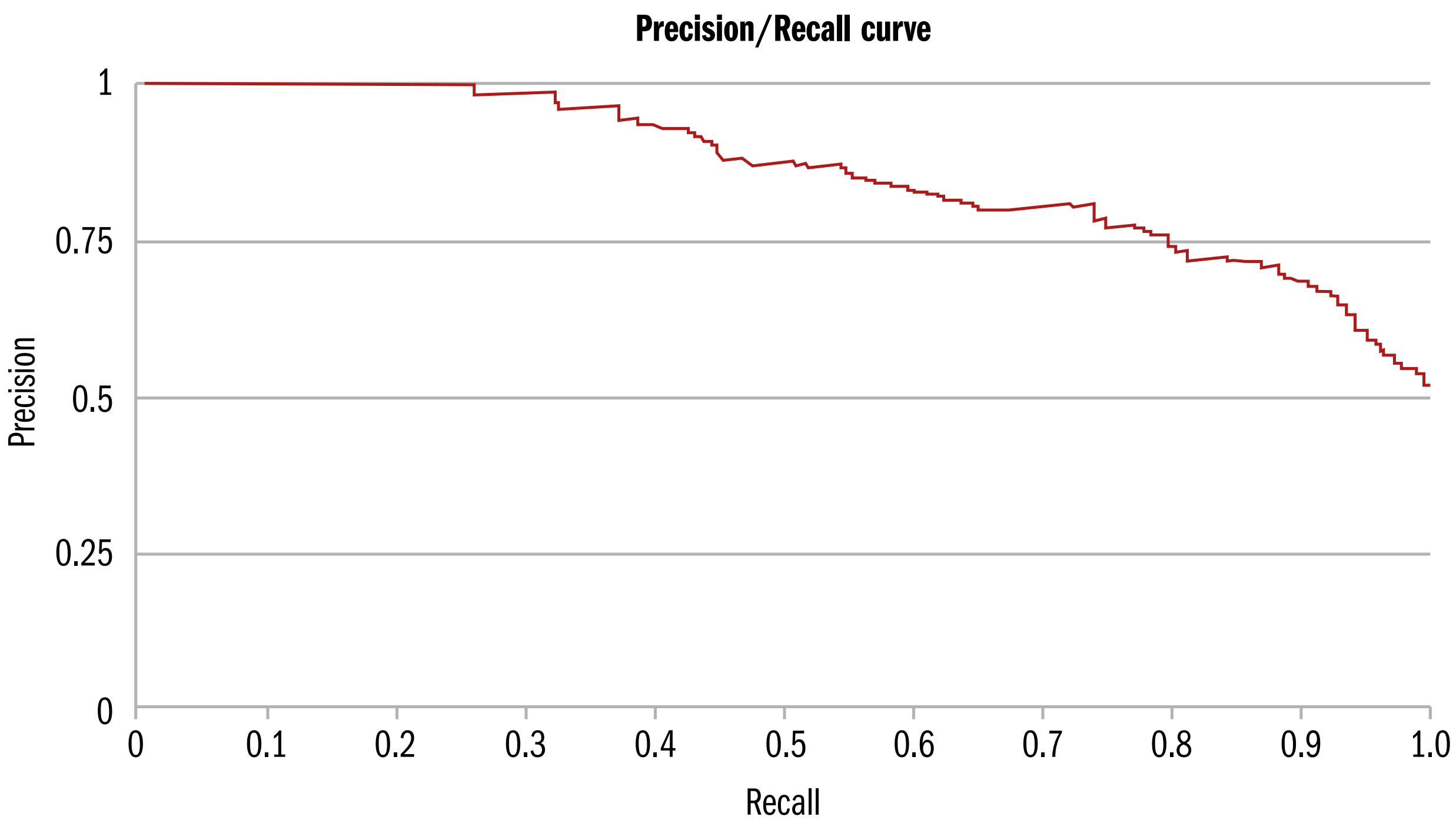
- **Expert-trained classifiers were able to identify publications reporting HEOR outcomes using the abstract and title alone**
- **Reviewers tended towards exclusion (higher false-negatives), whereas the classifiers tended towards inclusion (lower false-negatives)**



Training metrics

Classifier question		Does this publication report HEOR data?	
Statistics			
Total references, n	472	Balanced accuracy score	0.55 (±0.09)
No	249 (52.75%)	Recall score	0.72 (±0.20)
Yes	223 (47.25%)	F1 score	0.75 (±0.08)

Reviewer's classification	DAISY's algorithm classification	
	No	Yes
No	68.67 (± 10.62)	20.67 (± 15.17)
Yes	14.33 (± 10.62)	53.67 (± 14.82)



Limitations

- Only those publications in conflict were re-examined; therefore it is possible that some false-positive/false-negative results occurred via 2 expert reviewers, or both the classifier and non-expert reviewer, selecting the same incorrect option (i.e. erroneously agreeing)
- While 2 disparate indications were selected to improve generalizability, it is possible that our findings would not be applicable to all disease areas and patient populations
- Studies not typically classed as HEOR may have been included if they reported relevant outcomes, e.g. clinical trials that include PROs and measures of quality of life
- During this project, we became aware of the differences in how HEOR is defined, depending on individual experience, e.g.
 - Those with market access experience may see contribution to reimbursement discussions as a defining feature

Methods

References imported to build database

- Embase and PubMed
- DistillerSR literature review software was populated with HEOR and non-HEOR publications from 2 clinical indications
- Reviews and publications without an abstract were excluded

TRAIN database developed

- Experienced reviewers (n = 7) classified publications as HEOR or not
- “Are 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 (n = 122)
- 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

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:

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:

1. Wang Z, et al. PLoS One. 2020;15:e0227742.
2. Belur J, Thompson 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), 837-65. <https://doi.org/10.1177/0049124118799372>
3. CDC Podcast on Economic Evaluation; available from: <https://www.cdc.gov/cardiovascular-resources/php/training/economic-evaluation.html>
4. National Institute for Health and Care Excellence. Methods for the development of NICE public health guidance (third edition). 2012. Available from: <https://www.nice.org.uk/process/pmg4/chapter/incorporating-health-economics>
5. Horigan BJ. Explore (NY). 2010;6:221-4.
6. Weeds ER, et al. In: 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>

