EXTERNAL VALIDATION USING REAL WORLD DATA OF A GENERATIVE AI ASSISTED MEDICATION-INDICATION KNOWLEDGE BASE

Saniya Deshpande[1], Florence Ma[1], Alice Beattie[1], Linxuan Zhang[1], Mattia Ficarelli[1], Edward Bonnett[1], Alistair Marsland[1], Luisamanda Selle Arocha[1], Godspower Oboli[1], Alejandra Castanon[1], Benjamin D. Bray[1]

Corresponding author: ben.bray@lcp.uk.com

No funding was provided for this study.

Presented at the international Society for Pharmacoepidemiology (ISPE) Annual Meeting, 24-28 August 2024, Berlin, Germany

1 Health Analytics, Lane Clark & Peacock LLP. London, UK

https://www.lcp.com/en

Summary

We aimed to carry out an external validation of a recently developed medication-indication knowledge base which had been generated using large language AI models. This maps all the potential clinical indications for medicines to enable high throughput pharmacoepidemiology research.

To evaluate how well the knowledge base reflected real-world prescribing patterns, we compared it against a random sample (N = 250) of patient reported medication-indication pairs from the NHANES dataset.

The AI-assisted knowledge base contained 84% of the medication-indication pairs in the NHANES sample. Of the missing indications (N = 40), the majority were judged by clinical experts to be incorrectly-reported indications in NHANES for the specified medication (N = 31), whilst only a small number were off-license (N = 7) or valid indications (N = 2) missed by the knowledge base.

Background

- We have developed and validated a medication-indication knowledge base using the GPT-4 and MedCAT LLMs1, which has 8,270 unique medication-indication pairs of community dispensed medicines.
- We aimed to carry out an external validation of the knowledge base to assess its accuracy compared to indications reported in real world data.
- We used the National Health and Nutrition Examination Survey3 (NHANES) pre-pandemic prescription medications dataset (2017-2020) for external validation, which contains patient reported indications for each medicine that each NHANES participant reported using.

Methods

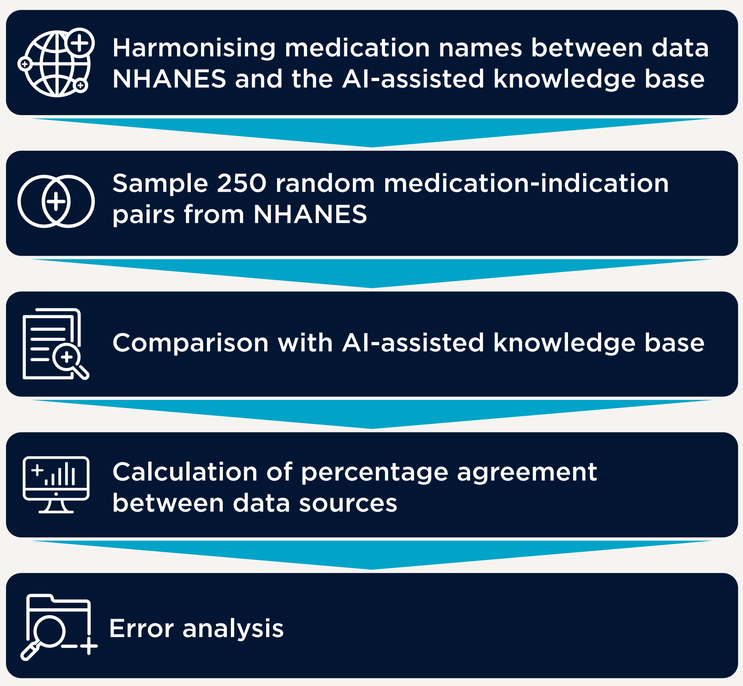

Figure 1: Flowchart of external validation process

- Harmonisation included converting common American names for medications into their respective British versions (e.g. Acetaminophen was converted to Paracetamol).
- Approximate string matching was then conducted using the Levenshtein distance and the R RecordLinkage package4.
- The indications reported in NHANES are patient-reported, so if an indication reported in NHANES was clinically valid (allowing for synonyms and lay language) this was treated as a match.
- Clinicians (two medical doctors and a pharmacist) reviewed all the discrepancies to analyse the reasons behind discrepancies between NHANES and the knowledge base.

Results

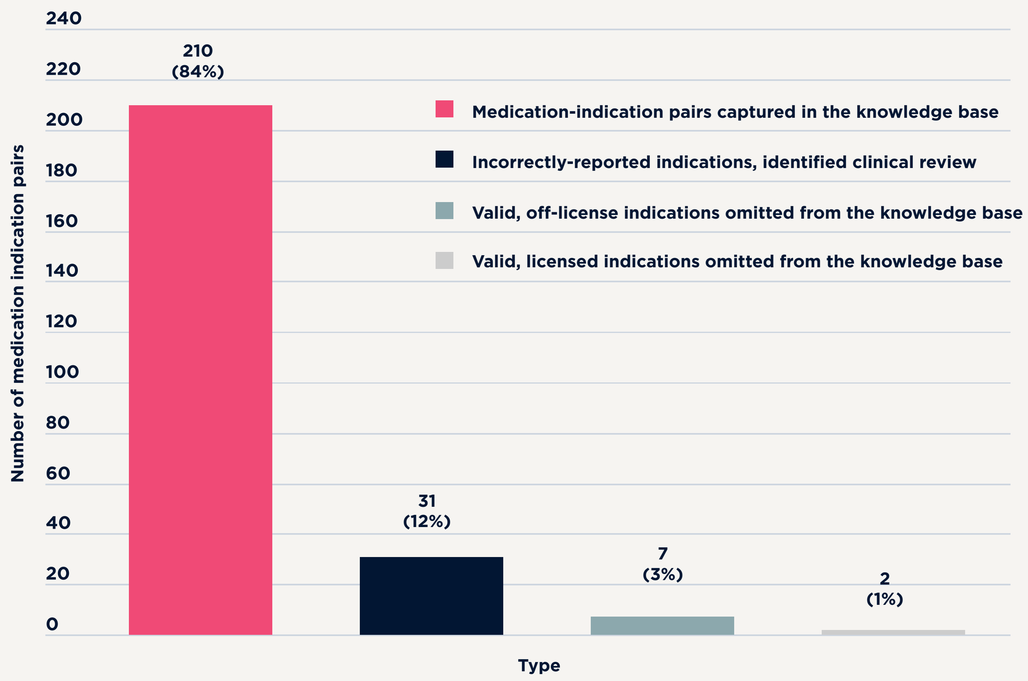

Figure 2: Types of medication-indication pairs in NHANES versus the AI-assisted knowledge base

- There was a high percentage agreement (84%) between the medication-indication pairs appearing in NHANES and in the medication-indication knowledge base.
- Of those missed by the knowledge base, most indications were not valid indications (licensed or otherwise) for the medication reported- this is likely to be a consequence of patient self-report (examples in Table 1). Figure 2 provides a detailed breakdown of omissions.

Medication	Indication in NHANES	Indications in knowledge-base	Evaluation
Metformin hydrochloride/ vildagliptin	Type 2 diabetes mellitus	Type 2 diabetes melitus	Correct - exact match
Finasteride	Enlarged prostate	Benign prostatic hyperplasia	Correct - alternative wording
Alprazolam	Heart disease	Panic disorder, agoraphobia, generalised anxiety disorder, anxiety associated with depression	NHANES incorrect - clinically unrealted indication
Ofloxacin	Unspecified disorder of eye and adnexa	Mixed infections of the urethra and cervix, uncomplicated skin and skin structure infections, uncomplicated cystitis, prostatitis	Correct indication missed by medication - indication knowledge base

Table 1: Examples of comparisons between NHANES and the AI-assisted knowledge base, and the final evaluations for each.

Conclusions

The LLM AI generated knowledge base captures most licensed and off-license indications used in NHANES. Discrepancies were largely due to patients reporting indications which were not clinically valid.

The knowledge base could be used to support pharmaco-epidemiological analytics and research, reducing the requirement for manual indication mapping and codelist creation.

Further reading:

1. Achiam J, Adler S, Agarwal S, et al. GPT-4 Technical Report. doi:https://doi.org/10.48550/arXiv.2303.08774
2. Kraljevic Z, Searle T, Shek A, et al. Multi-domain clinical natural language processing with MedCAT: The Medical Concept Annotation Toolkit. Artificial Intelligence in Medicine. 2021;117:102083. doi:https://doi.org/10.1016/j.artmed.2021.102083
3. NHANES: Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, [2017-March 2020][https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?Cycle=2017-2020]
4. Sariyar M, Borg A. The RecordLinkage Package: Detecting Errors in Data. The R Journal. 2010;2(2):61-67. doi:10.32614/RJ-2010-017
5. Wei WQ, Cronin RM, Xu H, Lasko TA, Bastarache L, Denny JC. Development and evaluation of an ensemble resource linking medications to their indications. J Am Med Inform Assoc. 2013;20(5):954-961. doi:10.1136/amiajnl-2012-001431
6. Zheng NS, Kerchberger VE, Borza VA, Eken HN, Smith JC, Wei WQ. An updated, computable MEDication-Indication resource for biomedical research. Sci Rep. 2021;11(1):18953. doi:10.1038/s41598-021-98579-4