

International Journal of Population Data Science

Journal Website: www.ijpds.org



Free Text Analysis: Identification of adverse drug events in clinical notes

Gauri Kamat¹, Mingyang Shan², and Roei Gutman¹

¹Brown University

²Eli Lilly and Company

Objective

To explore the prevalence of adverse drug events (ADE) leading to drug changes using Artificial Intelligence (AI) based free text analysis of primary care encounter notes.

Approach

We used electronic medical record I encounter notes linked to population-based drug dispensation records. Dispensation records were used to identify drug switching to another agent in the same therapeutic class, which could suggest an undesirable side-effect requiring discontinuation of the drug. We annotated clinical notes by identifying ADEs after retrieving the last encounter note written for that patient on or before the identified date. We ran different AI models, including BERT and Large Language Models (LLMs), to identify ADEs.

Results

We annotated 1085 notes with 362 ADEs. These were used to train BERT models (precision or PPV: 0.584) and prompt-based LLM (0.563) but with improved sensitivity (0.725) and specificity (0.739).

Conclusion

Identifying ADEs requires deep medical knowledge and simple Natural Language Processing models trying to identify ADEs only on the basis of language of encounter notes was not successful based on the BERT model performance. Prompt based models have deep knowledge but were trained with the objective of producing grammatically correct, coherent sentences, not with the medical knowledge required. LLMs also require significant computational power.

Implications

While AI does have significant potential to advance our capacity for complex analytic tasks the current LLMs are not yet able to adequately identify ADEs in a corpus of electronic clinical notes.

