HealthManagement
Promoting Management and Leadership

Text Mining Algorithms Detect Adverse Drug Reaction Detection

Adverse drug reactions (ADRs) are a critical concern in healthcare, often leading to patient harm and increased healthcare costs. Despite the widespread implementation of clinical decision support systems (CDSS) in electronic health records (EHRs), these systems struggle to detect ADRs documented in free-text form. The complexity of unstructured data poses a significant challenge for accurate ADR identification, which is crucial for improving patient safety and treatment outcomes. A recent Jamia Open article explores the development and effectiveness of a text mining algorithm designed to identify ADRs in the free text of Dutch EHRs, highlighting its potential to enhance healthcare practices.

Development of the Text Mining Algorithm

The algorithm was developed in two phases, each focusing on improving its ability to identify ADRs from unstructured EHR data accurately. The previously existing CDSS algorithm was recoded into an R-based text mining (TM) tool in the first phase. This recoding allowed for better handling of free text, resulting in a significant increase in the sensitivity of ADR detection. The algorithm's sensitivity was further refined by implementing a series of Plan-Do-Check-Act (PDCA) cycles, ultimately achieving a sensitivity of 93% and a positive predictive value (PPV) of 13%.

The second step involved integrating Medical Dictionary for Regulatory Activities (MedDRA) terms and Systematized Nomenclature of Medicine Clinical Terms (SNOMED-CT) into the algorithm. These additions were crucial in enhancing the identification of ADRs by linking symptoms described in the EHRs with standardised medical terminologies. Despite some challenges in matching MedDRA terms exactly, the algorithm demonstrated a notable improvement in detecting potential ADRs, achieving an overall sensitivity of 86% and a PPV of 10%.

Enhancement through R-Scripts and Final Algorithm Performance

The second development phase focused on further refining the algorithm by incorporating six R-scripts designed to enhance sensitivity and PPV. Among these, scripts for drug names, MedDRA/SNOMED-CT terms, deduplication, and seriousness of ADRs showed the most significant improvements. The deduplication script, in particular, was instrumental in reducing false positives by removing duplicate ADR entries. The final version of the algorithm achieved a PPV of 70% and a sensitivity of 73%, demonstrating a substantial improvement over previous versions and meeting the target for clinical applicability.

These improvements underline the algorithm's capability to identify ADRs that might be overlooked in manual reviews or traditional CDSS. For example, the algorithm was able to detect variations in the documentation of the same ADR, such as "myopathy due to steroids", described in different ways across the EHR notes. This flexibility is crucial in a clinical setting where terminologies may vary significantly between healthcare professionals.

Challenges and Future Directions

Despite its promising results, the algorithm still faces several challenges. One of the primary limitations is the reliance on a dataset from a single hospital, which may not represent the diversity of documentation practices across different institutions. Additionally, excluding scanned or imported documents from the analysis may have led to incomplete ADR identification. Expanding the algorithm's applicability to include these types of documents could further enhance its accuracy and utility.

Another challenge is the need for internal and external validation of the algorithm. While the current study demonstrated significant improvements in ADR detection, these results must be validated in different clinical environments and with more diverse patient populations. Moreover, integrating the algorithm with existing CDSS would require careful consideration of how to balance the algorithm's outputs with clinical workflows

to avoid alert fatigue among healthcare professionals.

Conclusion

Developing a text-mining algorithm for identifying ADRs in Dutch EHRs represents a significant advancement in pharmacovigilance. By successfully recoding and enhancing a CDSS algorithm, integrating MedDRA and SNOMED-CT terminologies, and employing advanced R-scripts, the study has demonstrated that such an algorithm can significantly improve the detection of ADRs from unstructured data. However, further validation and refinement are necessary to integrate this tool fully into clinical practice. Ultimately, the widespread adoption of such algorithms could lead to improved patient safety, reduced healthcare costs, and better clinical outcomes.

With enhanced sensitivity and PPV, this text mining algorithm holds great promise for future ADR detection in healthcare. Its ability to analyse free-text data and integrate with clinical systems makes it a valuable tool in improving patient safety and healthcare quality.

Source: Jamia Open

Image Credit: iStock

Published on : Mon, 2 Sep 2024