Oral Communication

OC-313

Enhancing Traumatic Brain Injury Patient Pathway Analysis With AI: An Evaluation of LLM-Based Data Extraction

Topic: Miscellaneous

Emmeline Montané[1], Inès Bigarré[2], Emmanuel Gilson[2], Claire Lebely[2], Xavier De Boissezon[2]

[1]*Toulouse University Hospital, Toulouse, France*
[2]*France*

BACKGROUND

Understanding the care pathways of traumatic brain injury (TBI) patients is essential for optimizing management and outcomes. The French SNDS database offers valuable insights into healthcare consumption but lacks data on initial injury severity, a critical factor in assessing patient trajectories. This study aims to develop and evaluate a Large Language Model (LLM)-based system to extract clinical data, particularly TBI severity and factors influencing care continuity, from 23,395 patient records.

METHODS

Before processing the full dataset, we developed and evaluated the LLM-based extraction tool on a subset of 20 patients from the PMSI-coded hospital records (2014-2018) to refine its performance. All data were anonymized before processing. The model was designed to extract Glasgow Coma Scale (GCS) scores (3-15), trauma mechanisms (falls, traffic accidents, fights, suicide attempts, substance abuse), and trauma dates. Performance was evaluated against a reference standard manually annotated by four experts, using 487 correspondences from 20 patients. The model was assessed at two levels: correspondence-level analysis for individual annotations and patient-level analysis aggregating results across correspondences. Metrics included precision (P), recall (R), F1-scores (F1) for categorical variable and accuracy (A) for continuous variable. Performance was categorized as poor (0.95). Annotation time was compared between human experts and LLM extraction.

RESULTS

The model extracted relevant data from 68 of the 487 correspondences. At the correspondence level, GCS extraction (A = 0,64) and trauma date extraction (A = 0,67) showed average accuracy, while trauma mechanism extraction (p = 0.73/R = 0.94/F1 = 0.82) performed well.

At the patient level, after prioritizing data across multiple correspondences per patient, performance improved: GCS extraction (A = 0.85) and trauma date extraction (A = 0,80) showed good accuracy while trauma mechanism extraction (p = 0.91/R = 0.64/F1 = 0.75) performed very well.

The LLM completed 487 annotations in 79 minutes, compared to seven hours and 25 minutes for human annotation.

CONCLUSION

Large Language Models efficiently extract structured clinical data, reducing annotation time. We will refine the model to improve processing speed, apply it to all records, and integrate findings with the SNDS database. This will enhance understanding of TBI care continuity and help identify factors influencing patient trajectories and long-term management.

Statement of Disclosures

None

Doi: 10.5606/archisprm.2026.OC-313

Corresponding Author:
Emmeline Montané
Toulouse University Hospital, Toulouse, France
Email: emmelinemontane@gmail.com