

ABSTRACT

Objectives: Abusive head trauma (AHT) is a leading cause of death in young children. Analyses of patient characteristics presenting to Emergency Medical Services (EMS) are often limited to structured data fields. Artificial Intelligence (AI) and Large Language Models (LLM) may identify rare presentations like AHT through factors not found in structured data. Our goal was to apply AI and LLM to EMS narrative documentation of young children to detect AHT.

Methods: This is a retrospective cohort study of EMS transports of children <36 months of age with a diagnosis of head injury from the 2018–2019 ESO Research Data Collaborative. Non-abusive closed head injury (NA-CHI) was distinguished from AHT and child maltreatment (AHT-CAN) through 2 expert reviewers; kappa statistic (κ) assessed inter-rater reliability. A Natural Language Processing (NLP) framework using an LLM augmented with expert derived n-grams was developed to identify AHT-CAN. We compared test characteristics (sensitivity, specificity, negative predictive value (NPV)) between this NLP framework to a Generative Pretrained Transformer (GPT) or n-grams only models to detect AHT-CAN. Association of specific word tokens with AHT-CAN was analyzed using Pearson's chi-square. Area Under the Receiver Operator Curve (AUROC) and Area Under the Precision-Recall Curve (AUPRC) are also reported.

Results: There were 1082 encounters in our cohort; 1030 (95.2%) NA-CHI and 52 (4.8%) AHT-CAN. Inter-rater agreement was substantial ($\kappa = 0.71$). The augmented NLP framework had a specificity and sensitivity of 72.4% and 92.3%, respectively with a NPV of 99.5%. In comparison, the GPT model had a sensitivity of 69.2%, specificity of 97.1% and NPV of 98.4% and n-grams alone had a sensitivity of 53.8%, specificity of 62.0%, NPV of 96.4%. AUROC was 0.91 and AUPRC was 0.52. A total of 44 n-grams and bi-grams were positively associated with AHT-CAN including “domestic,” “various,” “bruise,” “cheek,” “multiple,” “do,” “not respond,” “see EMS.”

Conclusions: AI and LLMs have high sensitivity and specificity to detect AHT-CAN in EMS free-text narratives. Words associated with physical signs of trauma are strongly associated with AHT-CAN. LLMs augmented with a list of n-grams may help EMS identify signs of trauma that aid in the detection of AHT in young children.