

INPLASY202660130

doi: 10.37766/inplasy2026.6.0130

Received: 28 June 2026

Published: 28 June 2026

**Corresponding author:**

Hui Liang

lyxybrtejw@126.com

**Author Affiliation:**

Linfen Central Hospital, Linfen, China.

Liang, H; Yang, XP; Sun, MZ.

**ADMINISTRATIVE INFORMATION****Support** - This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.**Review Stage at time of this submission** - Completed but not published.**Conflicts of interest** - None declared.**INPLASY registration number:** INPLASY202660130**Amendments** - This protocol was registered with the International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY) on 28 June 2026 and was last updated on 28 June 2026.**INTRODUCTION**

**Review question / Objective** This systematic review aims to review NLP applications in epilepsy clinical texts, assess methodological quality using an adapted QUADAS-2 tool supplemented by an AI/NLP-specific framework, and provide directions for future clinical research, particularly for developing clinically applicable predictive models for seizure recurrence.

**Rationale** Approximately 70 million people worldwide suffer from epilepsy, a population with a higher mortality risk than the general public. Early identification of patients with epilepsy at a high recurrence risk and timely initiation of antiepileptic treatment may reduce the recurrence risk. Currently, limited research exists on applying NLP to clinical text data for prediction tasks in epilepsy. However, among the existing studies that do apply NLP to epilepsy clinical texts, the methodological quality and reporting completeness remain unclear. Therefore, a systematic review focusing on

methodological quality assessment is urgently needed to identify current weaknesses and guide future research directions.

**Condition being studied** Epilepsy.

**METHODS**

**Search strategy** The search strategy was built around three conceptual domains (NLP, Epilepsy, Clinical Text) using Boolean operators: terms within each domain were combined with OR, and the three domains were combined with AND. In PubMed, we used a combination of MeSH terms and title/abstract fields ([TIAB]); in Europe PMC, searches were limited to title and abstract fields (TITLE/ABSTRACT). Full search strings are available in Supplementary Appendix S1. Searches were limited to English language and publications from January 1, 2016 to May 1, 2026.

**Participant or population** Epilepsy patients.

---

**Intervention** Natural language processing (NLP) techniques applied to epilepsy clinical text data.

**Comparator** Not applicable.

**Study designs to be included** Original research articles applying NLP techniques to epilepsy clinical text, reporting at least one quantifiable performance metric, with full text in English.

**Eligibility criteria** Original research articles and systematic reviews applying NLP, large language models, or machine learning techniques to epilepsy clinical text, reporting at least one quantifiable performance metric, with full text available in English.

**Information sources** A systematic literature search was performed on May 1, 2026, covering publications from January 1, 2016 to May 1, 2026. The following electronic databases were searched: PubMed and Europe PMC. To supplement the primary database search, forward citation tracking was performed using Google Scholar.

**Main outcome(s)** Study characteristics, NLP technical paradigms, task types, and performance metrics. Methodological quality and risk of bias assessed using QUADAS-2 and an AI/NLP-specific framework, with additional evaluations of reproducibility and class imbalance handling.

**Additional outcome(s)** None.

**Data management** Data will be extracted using a standardized Excel form. Two reviewers will independently perform study selection, with disagreements resolved by discussion. Zotero will be used for reference management. Data will be extracted using a standardized Excel form. Two reviewers will independently perform study selection, with disagreements resolved by discussion or a third reviewer. Zotero will be used for reference management.

**Quality assessment / Risk of bias analysis** Methodological quality and risk of bias were assessed using an adapted QUADAS-2 tool, supplemented by an AI/NLP-specific framework covering data leakage, external validation, and generalization capability. A detailed overview of this assessment framework is provided in Supplementary Table S1. Reproducibility was assessed across data openness, code transparency, and model/rules sharing. Class imbalance was evaluated by metric selection and reporting of positive sample distribution. All assessments were performed independently by

two reviewers, with disagreements resolved through discussion. Results are presented narratively and in tables.

**Strategy of data synthesis** A meta-analysis was not performed due to heterogeneity across studies in task types, dataset sources, gold standards, and evaluation metrics, which made direct numerical comparisons inappropriate. A narrative synthesis was therefore adopted. Extracted data were summarized across four domains: study characteristics, NLP technical paradigms, research tasks and performance metrics, and methodological quality. Results are presented in descriptive text and summary tables; frequencies and percentages were calculated for key methodological indicators where applicable.

**Subgroup analysis** Subgroup comparisons will be performed descriptively based on NLP technical paradigms (rule-based, traditional machine learning, deep learning, and hybrid approaches). No statistical hypothesis testing is planned.

**Sensitivity analysis** Not applicable, as no meta-analysis or quantitative pooling was performed.

**Language restriction** English.

**Country(ies) involved** China.

**Other relevant information** None.

**Keywords** Natural Language Processing; Epilepsy; Clinical Text; Systematic Review; Quality Evaluation; External Validation.

**Dissemination plans** The findings of this systematic review will be disseminated through publication in a peer-reviewed scientific journal.

**Contributions of each author**

Author 1 - Hui Liang - H.Liang contributed to study conception and design, literature search, data extraction, quality assessment, manuscript drafting, and writing-review & editing.  
Email: lyxybrtejw@126.com

Author 2 - Xiao Ping Yang - X. P. Yang participated in screening, verification, dispute resolution, and manuscript revision.  
Email: lyxybrtejw@126.com

Author 3 - Mei Zhen Sun - M. Z. Sun contributed to validation of results, investigation, intellectual discussions, and critical review and editing of the manuscript.  
Email: lyxybrtejw@126.com