Scoping Review Protocol

Katharina Leipfinger, Krankenhaus Barmherzige Brüder Regensburg

Prof. Dr. Piso – Klinik für Allgemein- und Viszeralchirurgie, Krankenhaus Barmherzige Brüder Regensburg
Prof. Dr. Raptis – Institut für eHealth & Datensicherheit, OTH Regensburg
Prof. Dr. Baumann – Fakultät Informatik, OTH Regensburg
Dr. Helge Knüttel – Universitätsbibliothek Regensburg
23.02.2026

Title

Performance and Trustworthiness of Large Language Models as Clinical Decision Support Systems in Acute Appendicitis: Protocol for a Scoping Review

Background and Rationale

Large Language Models (LLMs) are rapidly emerging as potential Clinical Decision Support Systems (CDSS). While artificial intelligence is already established in imaging and documentation, the use of LLMs for clinical decision-making in acute surgical conditions remains limited. Acute appendicitis represents one of the most frequent and time-critical surgical emergencies, where incorrect decisions may lead to severe complications. At the same time, increasing workload, staff shortages, and limited night-time coverage especially in peripheral hospitals intensify the need for decision support. Given the fast-growing and heterogeneous body of literature, a scoping review is required to systematically map current evidence on the performance, limitations, and trust-related aspects of LLM-based CDSS in the management of acute appendicitis.

Objectives

The objective of this scoping review is to systematically map the existing evidence of Large Language Models as Clinical Decision Support Systems in the management of acute appendicitis.

Review Questions

How are Large Language Model–based clinical decision support systems used in the management of acute appendicitis in adult and pediatric patients, what performance

outcomes have been reported, and do differences exist in the level of development and clinical application between pediatric and adult surgery?

Which aspects of trustworthiness, transparency, and explainability are addressed?

What ethical, legal, and data protection challenges are described?

Eligibility Criteria

Inclusion Criteria

Studies investigating Large Language Models as Clinical Decision Support Systems

Clinical focus explicitly on acute appendicitis

Adult and/or pediatric patient populations

All study designs (experimental, observational, qualitative, mixed methods, reviews)

Publications in English or German

Publication period from November 2022 to current.

Exclusion Criteria

Studies not related to acute appendicitis

Studies focusing exclusively on image-based artificial intelligence without LLM components

Non-clinical applications of LLMs

Editorials, commentaries, or opinion pieces without empirical or methodological content

Information Sources

Searches will be run in the following databases:

- MEDLINE (Ovid)
- Embase (Ovid)
- Science Citation Index Expanded, Social Sciences Citation Index (Web of Science Core Collection, Clarivate)

- Cochrane Library (Wiley)
- IEEE Xplore
- ACM Digital Library
- ClinicalTrials.gov

Search Strategy and Time Frame

Search strategies will be developed by a medical librarian in cooperation with the domain experts. Initial strategies will be developed for MEDLINE and Embase and the adapted to the other databases. Appropriate controlled vocabulary (e.g., MeSH and Emtree terms) and free-text terms will be used for the search blocks Large Language Models, and acute appendicitis.

The search will be limited to publication dates starting with the public release of functional ChatGPT (November 2022) to current.

Search strategies will be fully documented in a reproducible format in accordance with PRISMA-S.

Study Selection

All retrieved records will be imported into a reference management system and deduplicated. A two-stage screening process (title/abstract screening followed by full-text screening) will be performed using predefined eligibility criteria. The study selection process will be documented using a PRISMA-ScR flow diagram.

Data Extraction

Data will be extracted using a standardized data charting form, including:

Bibliographic characteristics (year, country, study design)

Clinical setting and type of appendicitis

Patient population: adult, pediatric, or mixed

Type of LLM and intended clinical use

Diagnostic and/or therapeutic decision-making (e.g., surgery vs. conservative treatment)

Performance outcomes (e.g., accuracy, concordance, sensitivity, specificity)

Trustworthiness, transparency, and explainability aspects

Ethical, legal, and data protection considerations

Recommendations for clinical implementation

Data Synthesis

Results will be synthesized descriptively and presented in tabular and narrative form. Where data allow, findings will be stratified by adult and pediatric populations.

Ethical Considerations

This scoping review is based exclusively on published literature and does not involve primary patient data. Therefore, no additional ethical approval is required.

Reporting

The review will be conducted and reported in accordance with the PRISMA-ScR guidelines.

Data Availability

Data supporting this review will be made available in a public repository according to FAIR principles, including the search documentation and data extraction forms and screening decisions.

Registration

This scoping review protocol is registered at the University of Regensburg Publication Server (https://epub.uni-regensburg.de/).