AI-Assisted Chart Review to Understand Disease Flares in Systemic Lupus Erythematosus

Yupeng Li[1], Lixia Yao[1,2], Ernest Vina[3], Huanmei Wu[2], David Fleece[2], Jay Patel[2]

1. Polygon Health Analytics LLC, Chalfont, PA, USA 2. Temple University, Philadelphia, PA, USA. 3. University of Arizona, College of Medicine, Tucson, AZ, USA

RWD182

Objectives

Systemic lupus erythematosus (SLE) is a chronic autoimmune disorder characterized by a broad spectrum of symptoms and manifestations across organs and tissues. It follows a relapsing-remitting course with varying outcomes and morbidity.

Accurate recognition and definition of disease flares are crucial for evaluating treatment responses, clinical outcomes, and guiding therapeutic decisions. Various flare definitions have been established in the context of clinical trials, including the use of the SLE Disease Activity Index (SLEDAI) where an increase >10 indicates a severe flare and >3 a mild/moderate flare. However, flare reporting in real-world clinical settings is often less systematic, residing in free-text clinical notes without standardized coding, necessitating manual review.

To address this challenge, we employed an advanced Generative Artificial Intelligence (GenAI) model to efficiently and accurately identify SLE flares from clinical notes.

Data and Methods

2,085 de-identified medical records between 1/1/2012 and 05/31/2022 were obtained from Temple University Health System's Epic EHR system.

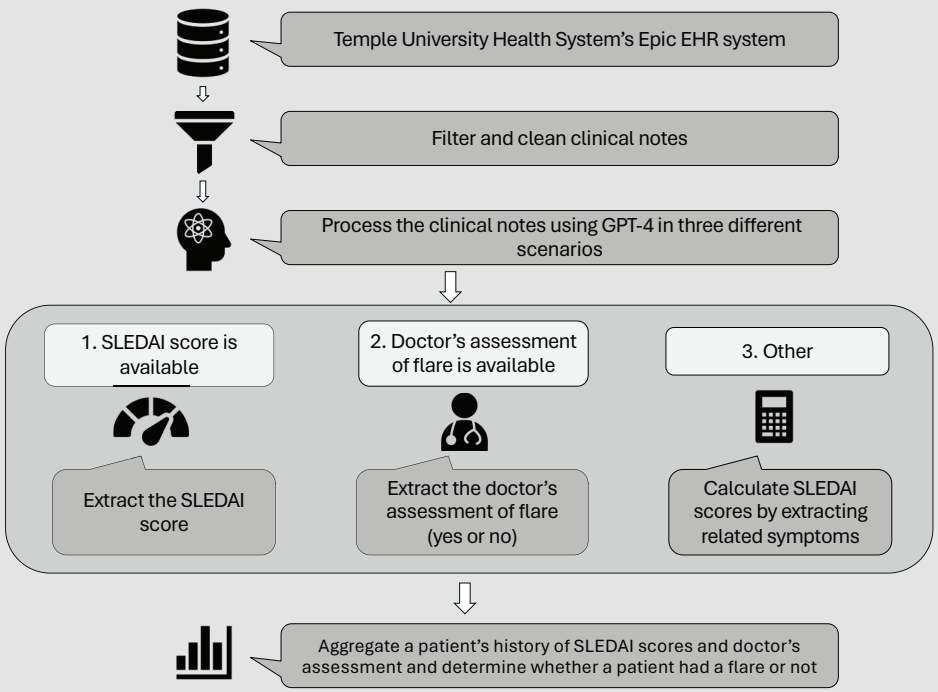
The dataset included 408 patients aged 18 years or above, with at least one rheumatology visit, at least one SLE ICD-10-CM code (M32.*) and SLE mentioned in clinical notes.

AI Model

We utilized the "gpt-4-1106-preview" version of the GPT-4 model developed by OpenAI to process the clinical notes.

Our initial comparison of multiple AI models and versions confirmed "gpt-4-1106-preview" has the highest accuracy.

Workflow



Examples

Clinical note	GPT-4 extraction
"Total (SLEDAI) score : 11 (rash, arthritis, alopecia, leucopenia, increased DNA binding)"	• Scenario = 1 • SLEDAI score = 11*
"currently having mild flare of disease."	• Scenario = 2 • Flare = yes
"Also describes loss of her peripheral vision"	• Scenario = 3 • Symptom = Visual disturbance • Score = 8*
"swelling of her knees and swelling throughout her body"	• Scenario = 3 • Symptom = Arthritis • Score = 4*

* Whether it is a flare was determined by calculating the total score and comparing the score with the previous visit.

Results

- GPT-4 identified total 611 flare events among 408 patients.
  - 7 were based on SLEDAI score changes (3 points increase from previous visit)
  - 71 were based on physicians' assessments mentioned in the notes
  - 533 were based on SLEDAI calculations from symptoms identified in the notes
- A review of 200 randomly selected notes by trained medical professionals showed a 97.5% agreement with GPT-4 results in terms of flare identification.

Scenario	No. of notes	No. of notes correctly processed by GPT-4
1	20	20
2	20	19 • In 1 note, physician's assessment is no flare, but GPT-4 classify it as flare.
3	160	156 • 1 note is not SLE-related, but GPT-4 still tried to calculate the SLEDAI score • In 3 notes, the SLEDAI scores calculated by GPT-4 were incorrect.

Conclusions

- GenAI can substantially enhance traditional manual chart review processes, which are labor-intensive and susceptible to errors due to human fatigue.
- Cutting-edge large language models, exemplified by GPT-4, demonstrate significant promise in accelerating SLE research by harnessing the wealth of information contained within electronic health records data.
- The workflow can be further automated and applied to broader therapeutic areas and scenarios.