UNIVERSITAS
GADJAH MADA
Evaluasi Interaksi Obat pada Pasien Rawat Inap dengan Diagnosa Penyakit Stroke Iskemik di Rumah Sakit Akademik Universitas Gadjah Mada Menggunakan Platform Large Language Model Dibandingkan dengan Database UpToDate Lexidrug
Alya Andita Damayanti, Dr. apt. Soni Siswanto, M.Biomed.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

INTISARI

Interaksi obat merupakan salah satu permasalahan dalam pola peresepan yang berpotensi mempengaruhi hasil klinis *(outcome)* pasien. Prevalensi terjadinya interaksi obat pada pasien dengan penyakit kardiovaskuler lebih tinggi dibandingkan dengan kelompok pasien lain dikarenakan jumlah dan penggunaan yang obat yang lebih banyak dan kompleks. Salah satunya yaitu pada penyakit stroke iskemik. Deteksi dan evaluasi interaksi obat sekarang dipermudah dengan adanya perkembangan teknologi berbasis *Artificial Intelligence (AI)* menggunakan model *Large Language Model (LLM)* yang telah menjadi penunjang bagi apoteker dalam memberikan pelayanan farmasi klinis.

Penelitian ini bertujuan untuk mengetahui perbedaan akurasi, sensitivitas, spesifisitas, nilai prediksi positif dan nilai prediksi negatif dari masing-masing platform LLM, yaitu ChatGPT dan DeepSeek, dalam mengidentifikasi kategori interaksi obat pada pasien rawat inap dengan diagnosis stroke iskemik di Rumah Sakit Akademik Universitas Gadjah Mada. Penelitian dilakukan dengan desain observasional dan pendekatan *cross-sectional* menggunakan metode analisis komparatif untuk mengkaji kinerja LLM dalam mendeteksi interaksi obat stroke iskemik. Adapun instrumen penelitian yang digunakan pada penelitian ini adalah data interaksi obat yang diperoleh dari *database* UpToDate Lexidrug dan platform LLM yang terdiri atas ChatGPT dan DeepSeek.

DeepSeek memiliki nilai spesifisitas sempurna (1,0), yang berarti *platform* ini sangat akurat dalam mengabaikan pasangan obat yang tidak berinteraksi secara klinis *(true negative)*. Sementara itu, ChatGPT memiliki spesifisitas yang lebih rendah (0,6) karena menghasilkan lebih banyak *false positive*. Namun, kedua *platform* memiliki nilai sensitivitas yang sama-sama sempurna (1,0), yang berarti keduanya andal dalam mendeteksi seluruh interaksi obat yang signifikan secara klinis *(true positive)*. Nilai akurasi DeepSeek (1,0) juga sedikit lebih tinggi dibandingkan ChatGPT (0,985). DeepSeek memiliki nilai prediksi positif (NPP) yang sempurna (1,0), yang berarti setiap prediksi positifnya dapat dipastikan sebagai interaksi yang benar-benar terjadi. ChatGPT juga memiliki NPP yang sangat baik (0,985). Untuk nilai prediksi negatif (NPN), kedua *platform* mencapai nilai sempurna (1,0), yang mengkonfirmasi bahwa keduanya sangat handal dalam memastikan bahwa prediksi negatif (tidak ada interaksi) adalah benar dan meminimalkan risiko *false negative* yang berbahaya.

Kata kunci: interaksi obat, stroke iskemik, LLM

UNIVERSITAS
GADJAH MADA
Evaluasi Interaksi Obat pada Pasien Rawat Inap dengan Diagnosa Penyakit Stroke Iskemik di Rumah Sakit Akademik Universitas Gadjah Mada Menggunakan Platform Large Language Model Dibandingkan dengan Database UpToDate Lexidrug
Alya Andita Damayanti, Dr. apt. Soni Siswanto, M.Biomed.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

Drug-drug interactions are one of the problems in drug prescribing that can potentially affect patient clinical outcomes. The prevalence of drug-drug interactions in patients with cardiovascular diseases is higher than in other patient groups due to the greater number and more complex use of drugs. One example is ischemic stroke. The detection and evaluation of drug interactions are now facilitated by advancements in Artificial Intelligence (AI)-based technology using Large Language Model (LLM) frameworks, which have become valuable tools for pharmacists in providing clinical pharmacy services.

This study aims to determine the differences in accuracy, sensitivity, specificity, positive predictive value, and negative predictive value of each LLM platform, namely ChatGPT and DeepSeek, in identifying the categories of drug-drug interactions in hospitalized patients diagnosed with ischemic stroke in Gadjah Mada University Academic Hospital. The study will be conducted using an observational design and cross-sectional approach with comparative analysis to assess the performance of LLM in detecting drug-drug interactions in ischemic stroke. The research instruments that will be used in this study includes drug interaction data obtained from the UpToDate Lexidrug database and the LLM platforms, ChatGPT and DeepSeek.

DeepSeek has a perfect specificity score (1.0), meaning the platform is highly accurate at ruling out drug combinations that do not interact clinically (true negatives). Meanwhile, ChatGPT has a lower specificity (0.6) because it generates more false positives. However, both platforms have equally perfect sensitivity scores (1.0), meaning both are reliable in detecting all clinically significant drug interactions (true positives). DeepSeek's accuracy score (1.0) is also slightly higher than ChatGPT's (0.985). DeepSeek has a perfect positive predictive value (PPV) of 1.0, meaning every positive prediction can be confirmed as an interaction that actually occurs. ChatGPT also has an excellent NPP (0.985). For negative predictive power (NPN), both platforms achieve a perfect score (1.0), confirming that they are highly reliable in ensuring that negative predictions (no interactions) are correct and minimizing the risk of dangerous false negatives.

Keywords: drug-drug interactions, ischemic stroke, LLM