

## INTISARI

Interaksi obat perlu menjadi perhatian karena dapat mengubah efek obat dalam tubuh. Banyak interaksi obat yang relevan secara klinis melibatkan obat dengan indeks terapi sempit, seperti digoksin dan warfarin. Salah satu cara untuk mengetahuinya adalah dengan memanfaatkan *database* informasi obat, seperti UpToDate Lexidrug yang selama ini digunakan dalam praktik klinis. Namun, seiring berkembangnya teknologi, *platform* AI muncul sebagai alternatif yang dapat dimanfaatkan untuk mengetahui interaksi obat. Meskipun demikian, validitasnya dalam mengidentifikasi interaksi obat masih perlu dieksplorasi lebih lanjut. Oleh karena itu, diperlukan penelitian mengenai kemampuan AI dalam mendeteksi interaksi obat berdasarkan *database* obat yang umum digunakan. Penelitian ini bertujuan untuk mengetahui perbedaan akurasi, sensitivitas, spesifisitas, nilai prediksi positif dan negatif, tingkat kesepakatan, serta kinerja diagnostik dari masing-masing AI dalam mendeteksi interaksi obat, dengan UpToDate Lexidrug sebagai referensi.

Penelitian ini merupakan penelitian observasional dengan pendekatan *cross-sectional* dan metode analisis komparatif. Subjek penelitian terdiri dari *platform* AI (ChatGPT-4o Mini, Copilot, dan Gemini) serta *database* obat konvensional (UpToDate Lexidrug). Obat yang digunakan yaitu digoksin dan warfarin, serta obat yang berpotensi berinteraksi dengannya berdasarkan data FDA, artikel *systematic review* dan *meta-analysis*, serta formularium nasional. Alat yang digunakan meliputi lembar pengumpul data dan *software* pengolah data (Microsoft Excel, JASP, dan RStudio).

Hasil menunjukkan bahwa ChatGPT-4o Mini memiliki spesifisitas dan nilai prediksi positif (NPP) tertinggi pada semua analisis (interaksi digoksin, warfarin serta gabungan digoksin dan warfarin) dan akurasi tertinggi pada interaksi warfarin; Copilot memiliki sensitivitas tertinggi pada interaksi warfarin; Gemini memiliki nilai prediksi negatif (NPN) tertinggi pada semua analisis, serta akurasi dan sensitivitas tertinggi pada interaksi digoksin maupun gabungan digoksin dan warfarin. Nilai kesepakatan tertinggi antara UpToDate Lexidrug dengan *platform* AI ditemukan pada UpToDate Lexidrug – ChatGPT-4o Mini dengan tingkat keparahan *major* menunjukkan kesepakatan tertinggi pada seluruh analisis. Tingkat keparahan *major* juga memiliki nilai AUC tertinggi pada seluruh analisis. Secara keseluruhan, ChatGPT-4o Mini menunjukkan keunggulan dibandingkan Copilot dan Gemini.

**Kata kunci:** interaksi obat, validitas, *artificial intelligence*, digoksin, warfarin

## ABSTRACT

Drug interactions require careful attention because they can alter the effects of drugs in the body. Many clinically relevant drug interactions involve narrow therapeutic index drugs, such as digoxin and warfarin. One way to identify them is by utilizing drug information databases, like UpToDate Lexidrug, which have long been used in clinical practice. However, with technological advancements, AI platforms have emerged as alternative tools for detecting drug interactions. Nevertheless, the validity of identifying drug interactions still needs further exploration. Therefore, research is needed to evaluate the capabilities of AI in detecting drug interactions based on commonly used drug databases. This study aims to compare the accuracy, sensitivity, specificity, positive and negative predictive values, agreement levels, and diagnostic performance of each AI in detecting drug interactions, using UpToDate Lexidrug as a reference.

This study is observational research with a cross-sectional approach and a comparative analysis method. The subjects included AI (ChatGPT-4o Mini, Copilot, and Gemini) and the conventional drug database (UpToDate Lexidrug). The drugs evaluated were digoxin and warfarin, along with other drugs that may potentially interact with them based on FDA data, a systematic review, and meta-analysis article, and the national formulary. The tools used included data collection sheets and data processing software (Microsoft Excel, JASP, and RStudio).

The results showed that ChatGPT-4o Mini had the highest specificity and predictive positive value (PPV) across all analyses (digoxin, warfarin, and their combination), as well as the highest accuracy in warfarin interactions. Copilot demonstrated the highest sensitivity in warfarin interactions, while Gemini exhibited the highest negative predictive value (NPV) in all analyses, along with the highest accuracy and sensitivity in digoxin and combined interactions. The highest agreement was observed between UpToDate Lexidrug – ChatGPT-4o Mini and major severity, which also had the highest AUC values. Overall, ChatGPT-4o Mini demonstrated superior performance compared to Copilot and Gemini.

**Keywords:** drug interactions, validity, artificial intelligence, digoxin, warfarin