

CAN ARTIFICIAL INTELLIGENCE ASSIST IN IDENTIFYING DRUG INTERACTIONS IN CRITICALLY ILL PATIENTS?

ChatGPT vs. STANDARD CLINICAL PRACTICE

P. Martínez Tomas; P. Benito Juez; L. Sánchez Luque; M. de Frutos del Pozo; Z. Rodríguez Fernandez; A. Miguel Domínguez; MT. Esteban Alonso; B. Castaño Rodríguez; N. Revilla Cuesta

Pharmacy Department. Hospital Universitario de Burgos, Spain

✉ pmartinezto@saludcastillayleon.es

WHY WAS DONE?

- Artificial intelligence is progressively being integrated into daily clinical practice, including hospital pharmacy. However, rigorous validation is necessary to guarantee its reliability and clinical applicability.

WHAT WAS DONE?

- To evaluate the capability of a generative artificial intelligence (AI) model to detect drug-drug interactions (DDIs) compared to databases commonly used in standard clinical practice (SCP).

HOW WAS IT DONE?

This study included hospitalized 30 patients in critical care units (ICU and PACU) with ≥ 5 prescribed drugs.

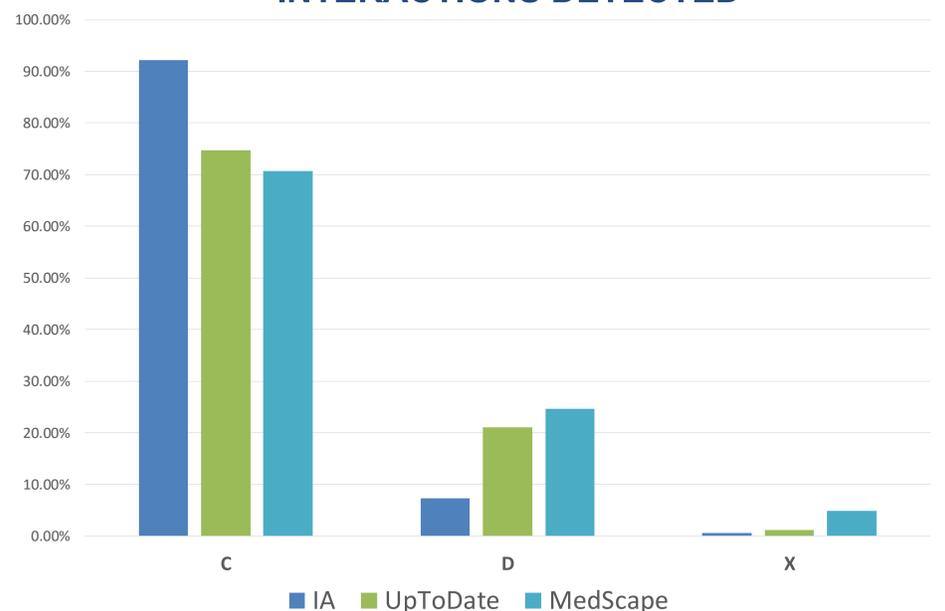
- Sample:** 293 drugs analyzed (mean 9,8 drugs per patient).
- Comparator:** AI model (ChatGPT 4.0) vs. Reference databases: UpToDate, Medscape.
- Classification system:** Lexicomp®
 - X:** Contraindicated
 - D:** Therapy modification required
 - C:** Therapy monitoring required
- Variables:** Patient count, admitting unit, total drugs analyzed, mean drugs and mean DDIs per patient.
- Statistical analysis:** Stata 16.0 with Student's t-test, Z-test, and Fisher's exact test ($p < 0.05$).

WHAT HAS BEEN ARCHIVED?

- Total DDIs detected:** AI (n=218), UpToDate (n=166), Medscape (n=167).
- Mean DDIs per patient (95% IC):**

Method	Mean DDIs per patient
AI Model	7.3 (6.3-8.3)
UpToDate	5.5 (4.2-6.9)
Medscape	5.6 (3.9-7.2)

INTERACTIONS DETECTED



All differences between the reference databases and AI were statistically significant, except for type X interactions between UpToDate and AI.

WHAT IS NEXT?

- No significant difference in total DDIs detected between methods was observed.
- AI overrepresented type C interactions and underreported clinically impactful interactions (types D and X). The low frequency of type X interactions warrants larger studies.
- AI's variability and lack of reproducibility limit widespread clinical application. While AI shows promise as a DDI detection tool, understanding its limitations and validating its outputs remain crucial.

