EVALUATING THE POTENTIAL OF CHATGPT TO SIMULATE CLINICAL TRIAL DATA: IMPLICATIONS FOR THE INTEGRITY OF SCIENTIFIC RESEARCH

AUTHORS: H. RODRÍGUEZ RAMALLO[1], N. BÁEZ GUTIÉRREZ[1], C.C. CABAÑAS[2], R. LÓPEZ RANCHAL[2], E. ALEGRE RUEDA[3], I. GARCÍA CONTRERAS[3]

1-HOSPITAL UNIVERSITARIO NUESTRA SEÑORA DE VALME, PHARMACY, SEVILLA, SPAIN.
2-HOSPITAL UNIVERSITARIO REINA SOFÍA, PHARMACY, CÓRDOBA, SPAIN.
3-HOSPITAL UNIVERSITARIO VIRGEN MACARENA, PHARMACY, SEVILLA, SPAIN.

Poster number 6ER-011

Scan the QR-code to see the online version

Background and importance

Large Language Models (LLMs) represent a significant opportunity for healthcare. However, they may also facilitate fraudulent scientific practices, generating plausible clinical trial (CT)

Aim and objectives

To assess the capability of GPT-4 ADA (OpenAI) to generate a dataset resembling a real CT, comparing two drugs without previous direct comparisons

Material and methods

1. Instructions to the LLM:

- Adapted from Taloni et al. (2023).
- Through the ChatGPT-4 interface, a request was made to generate a database for 500 patients with advanced clear cell renal carcinoma.

2. Trial simulation details

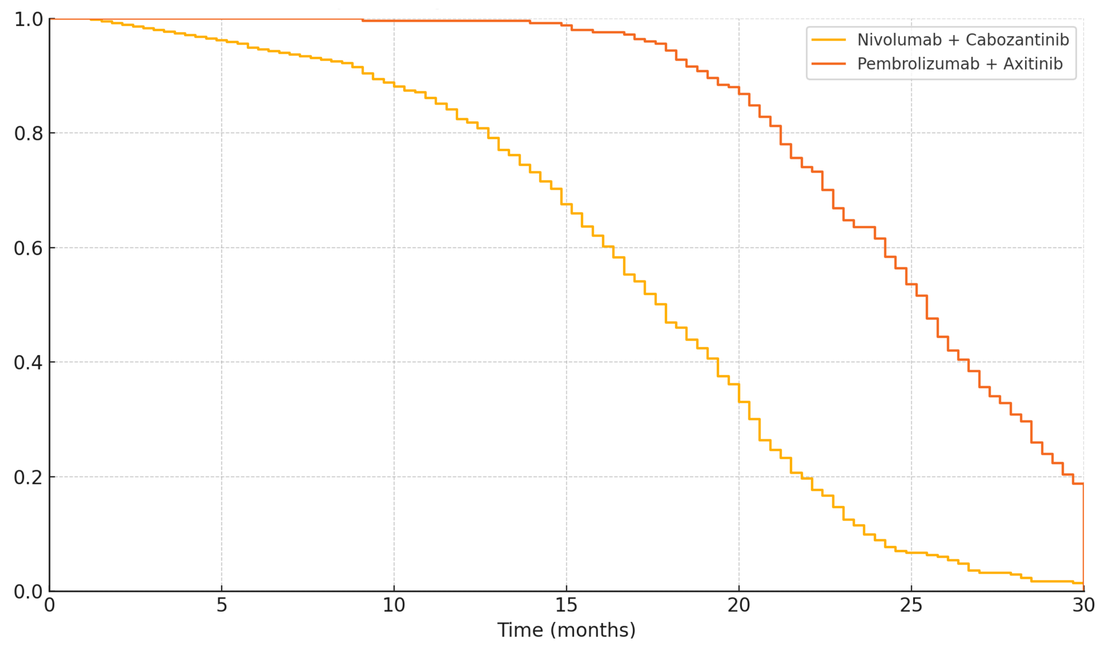
- Patient code
- Sex: 39% male, 61% female
- Date of birth
- Treatment: Nivolumab + Cabozantinib (50%), Pembrolizumab + Axitinib (50%)
- Recruitment region: North America (24%), Western Europe (26%), Others (50%)
- Combined Positive Score for PD-L1: ≥1(59%), <1 (41%)
- Nº of organs with metastasis: 1 (26%), ≥2 (74%)
- Prior radiotherapy: 10%
- Prior nephrectomy: 82%
- Time to death and PFS: measured in months

3. Data analysis

- A Cox proportional hazards model was applied to estimate hazard ratio (HR) and 95% confidence interval (CI).

Results

- ✓ Excel database with 500 pseudo-anonymized patients
- ✓ Two minutes
- ✓ Data met all the predefined criteria
- ✓ PFS: 18.78 months for Nivolumab + Cabozantinib and 25.25 months for Pembrolizumab + Axitinib, with a HR of 0.4 (95% CI: 0.32–0.5, P<0.001).



Conclusion and relevance

- ChatGPT's ability to generate CT-like datasets could facilitate fraudulent publications.
- Greater transparency in trial registration and data management is essential to safeguard research integrity.
- Hospital pharmacists should remain vigilant regarding the potential misuse of LLMs.

References and/or acknowledgements

Taloni A, Scorcia V, Giannaccare G. Large Language Model Advanced Data Analysis Abuse to Create a Fake Data Set in Medical Research. *JAMA Ophthalmol.* 2023;141(12):1174–1175. doi:10.1001/jamaophthalmol.2023.5162