ABSTRACT

Introduction: Integrating AI into pharmacy practice can improve patient outcomes by allowing healthcare professionals to dedicate more time to complex tasks and maintain a patient-centered approach. Despite AI's potential benefits, its accuracy and reliability in answering clinical pharmacy questions remain areas of ongoing research.

Objective: This study evaluates the accuracy of ChatGPT (version 3.5) in answering pharmacy-related clinical questions about to type 2 diabetes mellitus (T2DM). The objective was to assess AI's reliability in providing guideline-based recommendations and its impact on patient care, pharmacy education, and practice.

Methods: ChatGPT 3.5 was instructed to act as a pharmacist to answer 20 questions specific to type 2 diabetes and citing its sources. The questions were categorized into four groups: drug use (n=5), indications (n=5), side effects (n=5), and interactions (n=5). Responses were evaluated based on answers and guideline accuracy using two separate 3-point scales.

Results: A Shapiro-Wilk test suggested slight deviations from normality (W = 0.772, p = 0.061; reference mean, W = 0.791, p = 0.086; response mean). A one-sample t-test showed that ChatGPT's response and reference mean accuracy was not statistically significant compared to a benchmark mean score of 10 (p = 0.278 and p = 0.183, respectively). While ChatGPT provided mostly correct responses and references, minor inconsistencies were observed.

Conclusion: ChatGPT displays moderate to high accuracy in responding to T2DM related clinical questions and aligns with guideline-based standards. While it can be a useful tool, its responses should be reviewed for accuracy with human judgment.