

EVALUATION OF THE DIAGNOSTIC ACCURACY OF CHATGPT-4 IN RECOGNITION AND CLASSIFICATION OF ESOPHAGEAL VARICES: A MULTICENTER STUDY

Edoardo Giovanni Giannini

UEG Week Berlin 2025

Poster

Oesophagus

October 6, 2025

ChatGPT-4 achieved diagnostic accuracy comparable to junior and trainee endoscopists in detecting and classifying esophageal varices from endoscopic images, but was less effective at differentiating small from large varices and identifying red signs.

KEY POINTS

- In a study of 279 good-quality esophagogastroduodenoscopy images from six Italian centers, ChatGPT-4 without prior task-specific training was compared to experienced, junior, and trainee endoscopists using experienced endoscopists' evaluations as ground truth.
- ChatGPT-4 detected esophageal varices with AUC 0.77 (sensitivity 86.2%, specificity 68.4%), comparable to most endoscopists but lower than one junior endoscopist.
- ChatGPT-4 distinguished small (F1) from large (F2-F3) varices with AUC 0.75 (sensitivity 67.6%, specificity 81.7%), performing worse than 3/4 junior and 3/5 trainee endoscopists.
- ChatGPT-4 classified F3 varices with AUC 0.71 (sensitivity 52.8%, specificity 89.7%), comparable to all endoscopists and superior to one trainee, but identified red signs poorly (AUC 0.58, sensitivity 20.5%).
- The authors conclude the model shows potential as a supportive tool for training and preliminary triage but cannot replace experienced clinical judgment and requires further refinement for routine clinical use.

VIEW THIS CONTENT ON GUTFLIX

<https://gutflix.eu/search/d/f9629014-a814-11f0-a144-0242ac140006>



AI-generated summary

This summary was generated by an AI large language model based on the content transcript. It is for informational purposes only and should not be considered a substitute for clinical judgment. Always rely on your professional expertise and the full clinical context when making clinical decisions.