WHICH PROPERTY OF ARTIFICIAL INTELLIGENCE AFFECTS HUMAN DIAGNOSTIC PERFORMANCE MORE, INTERPRETABILITY OR PRECISION? A MULTI-READER MULTI-CASE STUDY

Lianlian Wu

UEG Week Copenhagen 2023 | Poster

Oesophagus

October 15, 2023

A multi-reader multi-case study comparing explainable and black-box AI for gastric neoplasm diagnosis found that AI precision, not interpretability, significantly affected endoscopist diagnostic performance.

KEY POINTS

- Both the explainable AI system ENDOANGEL-ED and a traditional sole deep-learning model significantly improved endoscopist accuracy when diagnosing gastric neoplasms from 127 video clips
- Twenty-five endoscopists completed four reading rounds with different AI assistance configurations, including no AI, explainable AI showing diagnostic indexes, and black-box AI showing only results
- When comparing systems with identical precision but different interpretability, no significant differences were found in accuracy (79.21% vs 78.24%, p=0.792), sensitivity, or specificity
- Comparing AI systems with different precision levels showed no significant accuracy difference (78.24% vs 76.91%, p=0.142) but significantly different sensitivity (83.93% vs 88.29%, p=0.030) and specificity (73.89% vs 68.22%, p=0.035)
- The authors conclude that AI precision, rather than interpretability through visible diagnostic indexes, is the critical factor affecting endoscopist performance and should be prioritized in AI development

VIEW THIS CONTENT ON GUTFLIX

https://gutflix.eu/search/d/c02ecdac-743c-11ee-aa04-0242ac140004

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.