

Introduction

Non-expert endoscopists' optical diagnosis (OD) of colorectal polyps falls short of 'resect and discard' and 'diagnose and leave' thresholds. Previous CADx studies are limited to locally deployed hardware AI systems and mostly evaluated in expert endoscopists. CADDIE is a cloud-based CADx system that receives the live image stream from the endoscopy screen via the hospital's internet network, where the AI algorithm processes the images and outputs its diagnosis (adenoma, non-adenoma or uncertain) on the endoscopy screen.

This RCT aimed to assess the performance of CADDIE when colonoscopy is undertaken by non-expert endoscopists (ADR < 30%).

Methods

We enrolled 739 patients between April 2021 - Dec 2022, scoped by 32 endoscopists in 9 UK hospitals. Each endoscopist was limited to a maximum of 60 procedures and required to perform more than 20% of this target. Patients were block randomised in a 1:1 ratio to standard of care (SOC) or intervention (CADDIE). OD accuracy of humans alone (SOC arm) and humans aided by CADDIE (CADDIE arm) were compared to the CADDIE system's stand-alone performance in the CADDIE arm.

Results

There were 615 evaluable patients after exclusions as pre-defined in the study protocol. There were 523 eligible polyps (270 SOC arm, 253 CADDIE arm). The CADDIE system diagnosed 17.8% of polyps (45/253) as "uncertain".

Human alone (SOC) OD accuracy was 75.9%, human OD aided by CADDIE was 80.2%, whilst the stand-alone CADDIE system was significantly higher than both at 87.5% ($p=0.005$ and $p=0.04$). Similarly, in sub-analysis of small/diminutive polyps ($\leq 10\text{mm}$), human alone OD accuracy was 73.8%, human OD aided by CADDIE was 79.0%, whilst the CADDIE system was significantly higher than both at 87.9% ($p=0.001$ and $p=0.02$). Although results were not significant for diminutive polyps ($\leq 5\text{mm}$), the trend was towards higher accuracy in the CADDIE system (86.6%) compared to humans alone (78.1%; $p=0.011$) and human OD aided by CADDIE (77.3%; $p=0.05$).

In sub-analyses of high-confidence (HC) OD, the CADDIE system's accuracy for $\leq 10\text{mm}$ polyps was significantly higher (87.9%) than humans alone (78.4%; $p=0.04$), with no difference to humans aided by CADDIE (83.3%; $p=0.18$). For HC OD of $\leq 5\text{mm}$ polyps, where the sample size is smaller, the CADDIE system's accuracy was higher (86.6%) but not significant to humans alone (82.7%, $p=0.46$) and humans aided by CADDIE (81.4%, $p=0.28$).

The OD-derived colonoscopy surveillance interval accuracy was 87.2% in humans alone, 86.8% in humans aided by CADDIE and 90.9% for the CADDIE system.

Conclusions

We demonstrate robust results for a cloud-based CADx system. Further research is warranted in the human-computer interaction to optimise non-expert endoscopists' OD performance to that of CADx systems.