**Official Title: Evaluation of Artificial Intelligence for Adenoma Detection in Water Exchange Colonoscopy: the WEAID Randomized Controlled Trial (Water Exchange With Artificial Intelligence-assisted Detection)**

**NCT number: 06173258**

**Date of Document: Oct 1, 2023.**

## Introduction

Adenoma detection rate (ADR) and adenoma per colonoscopy (APC) are crucial predictors for postcolonoscopy colorectal cancer (CRC). Water Exchange (WE) colonoscopy maximizes bowel cleanliness during insertion and has demonstrated superiority in increasing ADR and APC as compared to air insufflation.

Artificial intelligence (AI) lesion computer-aided detection (CADe) devices present promising tools for colonoscopies. Meta-Analysis reveals that AI generates an average ADR gain of 8% and an average APC gain of 0.20 in air-insufflated colonoscopy.

Recent research confirmed that the addition of AI to WE elevated ADR, and the application of WE to AI reduced false positive (FP) instances per colonoscopy, as evidenced by the analysis of pre-recorded video in a randomized controlled trial (RCT) comparing WE and air insufflation. Pilot retrospective data involving 715 patients (367 in the AI group, 348 in the control group) from Italy and Taiwan, with a mean age of 59.6 years (51.5% men) showed that WE with AI demonstrated numerical APC enhancement (1.31 ± 1.80 vs. 1.14 ± 1.38, $P$ = 0.152) and significantly increased the detection rate of proximal sessile serrated lesion (SSL) (5.2% vs. 2.3%, $P$ = 0.044) in comparison to WE alone (Tables 1 and 2) (Cheng CL et al. Gastrointest Endosc 2024;99:AB17–18).

**Table 1. Per-patient ADR and APC in water exchange (WE) with CADe pilot study**

| Detection of adenomas | WE without CADe (n=348) | WE with CADe (n=367) | *P* value |
|---|---|---|---|
| Overall ADR, n (%) | 203 (58.3) | 207 (56.4) | *0.602 |
| Right ADR, n (%) | 106 (30.5) | 107 (29.2) | *0.703 |
| Proximal ADR, n (%) | 132 (37.9) | 149 (40.6) | *0.465 |
| Overall APC, mean (SD) | 1.14 (1.38) | 1.31 (1.80) | †0.152 |

*Fisher exact test. †Student's *t* test

**Table 2. Detection of sessile serrated lesions in WE with CADe pilot study**

| Detection of SSLs | WE without CADe (n=348) | WE with CADe (n=367) | *P* value |
|---|---|---|---|
| **SSL detection rate** | | | |
| Overall SSL detection rate, n (%) | 10/348 (2.9) | 21/367 (5.7) | *0.062 |
| Right SSL detection rate, n (%) | 7 (2.0) | 14 (3.8) | *0.153 |
| Proximal SSL detection rate, n (%) | 8 (2.3) | 19 (5.2) | *0.044 |
| **Detected SSL per colonoscopy** | | | |
| Right SSL per colonoscopy, mean (SD) | 0.02 (0.17) | 0.06 (0.36) | †0.062 |
| Proximal SSL per colonoscopy, mean (SD) | 0.03 (0.22) | 0.10 (0.58) | †0.036 |

*Fisher exact test. †Student's *t* test

**Hypotheses**

We hypothesize that the combination of WE and AI will significantly increase APC as opposed to WE alone, in patients undergoing screening or surveillance colonoscopy. The primary outcome of this study is APC. We aim to confirm whether the combination of WE with AI achieves a significantly higher APC than WE alone.

## Methods

**Study design:** The WEAID (WE-AI-assisted Detection) trial is a two-arm parallel multicenter RCT designed to compare APC between WE with AI and WE colonoscopy

alone. Patient recruitment will be conducted at three hospitals in Italy and Taiwan (Digestive Endoscopy Unit, CTO Hospital, Iglesias, Italy [primary investigator: Sergio Cadoni]; Digestive Endoscopy and Gastroenterology Unit, Manzoni Hospital, Lecco, Italy [primary investigator: Arnaldo Amato]; and Evergreen General Hospital, Taoyuan, Taiwan [primary investigator: Chi-Liang Cheng]). Ethical approval will be sought from the Institutional Review Boards of the participating centers. Informed written consent will be obtained from all enrolled patients. The trial has been registered with ClinicalTrials.gov and assigned the identifier NCT06173258 before patient enrollment. The Sepulveda Ambulatory Care Center, VAGLAHS, UCLA in the USA will be a non-recruiting participating site.

**Study population:** The study population include male and female patients aged 45–75 years at average risk for CRC who plan to undergo colonoscopy for primary screening, postpolypectomy surveillance, and individuals with positive fecal immunochemical test (FIT) or guaiac fecal occult blood test (gFOBT) results. The exclusion criteria incorporate (1) patients with a history of inflammatory bowel disease, hereditary CRC syndrome, serrated polyposis syndrome, CRC, colorectal resection, colonic stricture, or severe comorbid illnesses rendering polypectomy unsafe; (2) patients with colonoscopy contraindications (e.g., acute diverticulitis or toxic megacolon); (3) therapeutic colonoscopy (e.g., hemostasis, removal of a large polyp); (4) emergent colonoscopy; (5) pregnant women or those planning pregnancy; and (6) patients unwilling to participate in the study.

**Study investigators:** All study investigators will be experts with a minimum of 500 completed WE colonoscopies. Additionally, each investigator will exhibit a minimum baseline ADR of 25%.

**Randomization:** Enrolled patients will be randomized in a 1:1 ratio to either the WE with AI-assisted colonoscopy (study group) or WE colonoscopy alone (control group). Randomization will be carried out by computer-generated sequences using a block design (four participants per block). Stratification will be based on study sites, colonoscopists, and colonoscopy indications (primary screening, surveillance, and positive FIT or gFOBT).

**Artificial intelligence:** Commercially available AI system (CAD-EYE, Fujifilm, EU and Taiwan; Endo-AID, Olympus, EU and Taiwan) will be employed. During AI-assisted procedures, the AI will be activated during the withdrawal phase of the procedure, providing a bounding box as output any time a lesion that is suspected to be a polyp is recognized by the CADe device. At each center, expertise in AI was already developed prior to the start of the study.

**Data collection:** To capture all study data comprehensively, a case record form will be employed. This form, available in written format, will document various aspects of the study, including enrollment procedures, randomization, colonoscopy details, histological results, and safety information. Study investigators and their teams are responsible for completing all sections of a pre-built Microsoft Excel worksheet.

**Colonoscopy procedure:** Bowel preparation regimens will align with established standard practices at each individual site. The study procedures employ standard high-definition colonoscopy video processors with integrated AI systems. In both arms of the study, insertion and withdrawal of the colonoscope will be executed by the same colonoscopist. The evaluation and grading of bowel preparation quality will

be carried out using the Boston Bowel Preparation Scale scoring system.

During the insertion phase of WE colonoscopy, the air pump will be turned off, while the colon will be irrigated with water using a flushing pump. The WE approach involves the simultaneous infusion of water to facilitate luminal expansion and suction of unclean water during insertion. Any encountered air pockets will be aspirated to ensure optimal WE maneuver in salvage cleaning. Upon reaching the cecum, where most of the water is suctioned to collapse the cecal lumen, the air pump will be opened. All investigators will be asked to adhere to WE technique confirmation of nearly equivalent volumes of water infused and suctioned upon cecal arrival. During the withdrawal phase, air or CO2 will be insufflated in both arms.

Withdrawal from the cecum will begin in the left lateral position. Similar withdrawal techniques with adequate luminal distention and comprehensive examination will be emphasized. Adhering to prevailing guidelines set by the American and European Societies of Gastrointestinal Endoscopy, a minimum of 6 minutes of clean withdrawal time will be maintained for all colonoscopies across both arms.

During the withdrawal phase, an exhaustive polyp search and subsequent resection will be performed in both study arms. Comprehensive records will be maintained for all detected polyps, capturing size (assessed in comparison with open forceps or snare), precise location, and morphology (classified according to the Paris classification).

**False positives of AI:** False positive (FP) activation is defined as the identification by AI systems of an area during the withdrawal phase that is not deemed to be a colorectal lesion after re-examination by the colonoscopist. Real-time flagging of FP activation is recorded by investigators and study staff. The NOISE classification will be

used to categorize the causes of FP activations.

**Study outcomes:** The primary outcome is APC, calculated as the total number of adenomas detected divided by the total number of colonoscopies performed. Secondary outcomes encompass 1) overall ADR, 2) ADR and APC in the right and proximal colon, 3) SSL detection rates, 4) mean number of non-neoplastic lesions resected per colonoscopy, 5) real-time number of FPs per colonoscopy, and 6) procedural times (e.g., insertion time, withdrawal time, and total procedural time).

**Sample size calculation:** For group comparison, the sample size was determined assuming that the APC of the control group is 0.9. A clinically significant relative increase of 25% in APC was considered. With a two-sided significance level of 5% and 80% power, 376 patients will be required per group (752 randomized patients). Accounting for a 10% overall dropout rate, enrollment of at least 836 patients is planned.

**Statistical analysis:** Categorical variables are presented as frequencies and percentages and continuous variables as means with standard deviations. The analysis will follow the intention-to-treat approach. Student's *t*-test will be used for continuous variables and chi-square test will be used for categorical variables. Poisson regression will be used to calculate incident rate ratios comparing the APC between the two arms, adjusting for age, sex, body mass index, smoking history, CADe system, and bowel preparation quality. A planned interim analysis will be conducted after the recruitment of approximately 70% of the planned enrollment to determine whether early termination of the study is appropriate to minimize the number of patients in the control group given the less effective, or clearly inferior

treatment. All statistical analyses will be performed using Python version 3.10 (SciPy Library). A *P* value < 0.05 is considered significant.