Improving the efficacy and patient tolerability of colorectal cancer screening

Reducing the occurrence of preventable colorectal cancer

Milou van Riswijk

RADBOUD UNIVERSITY PRESS

Radboud Dissertation Series

Improving the efficacy and patient tolerability of colorectal cancer screening

Reducing the occurrence of preventable colorectal cancer

Milou Lisa Marit van Riswijk

The work presented in this thesis was carried out at the department of Gastroenterology and Hepatology of the Radboudumc.

The printing of this thesis was financially supported by the department of Gastroenterology and Hepatology of the Radboudumc and the NVGE (Nederlandse Vereniging voor Gastroenterologie).

Improving the efficacy and patient tolerability of colorectal cancer screening
Reducing the occurrence of preventable colorectal cancer
Milou Lisa Marit van Riswijk

Radboud Dissertation Series
ISSN: 2950-2772 (Online); 2950-2780 (Print)

Published by RADBOUD UNIVERSITY PRESS
Postbus 9100, 6500 HA Nijmegen, The Netherlands
www.radbouduniversitypress.nl

Design: Proefschrift AIO | Guus Gijben
Cover: Proefschrift AIO | Guntra Laivacuma
Printing: DPN Rikken/Pumbo

ISBN: 9789465150901
DOI: 10.54195/9789465150901
Free download at: https://doi.org/10.54195/9789465150901

RADBOUD
UNIVERSITY
PRESS

Improving the efficacy and patient tolerability of colorectal cancer screening

Reducing the occurrence of preventable colorectal cancer

Proefschrift ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.M. Sanders,
volgens besluit van het college voor promoties
in het openbaar te verdedigen op

woensdag 19 november 2025
om 12.30 uur precies

door

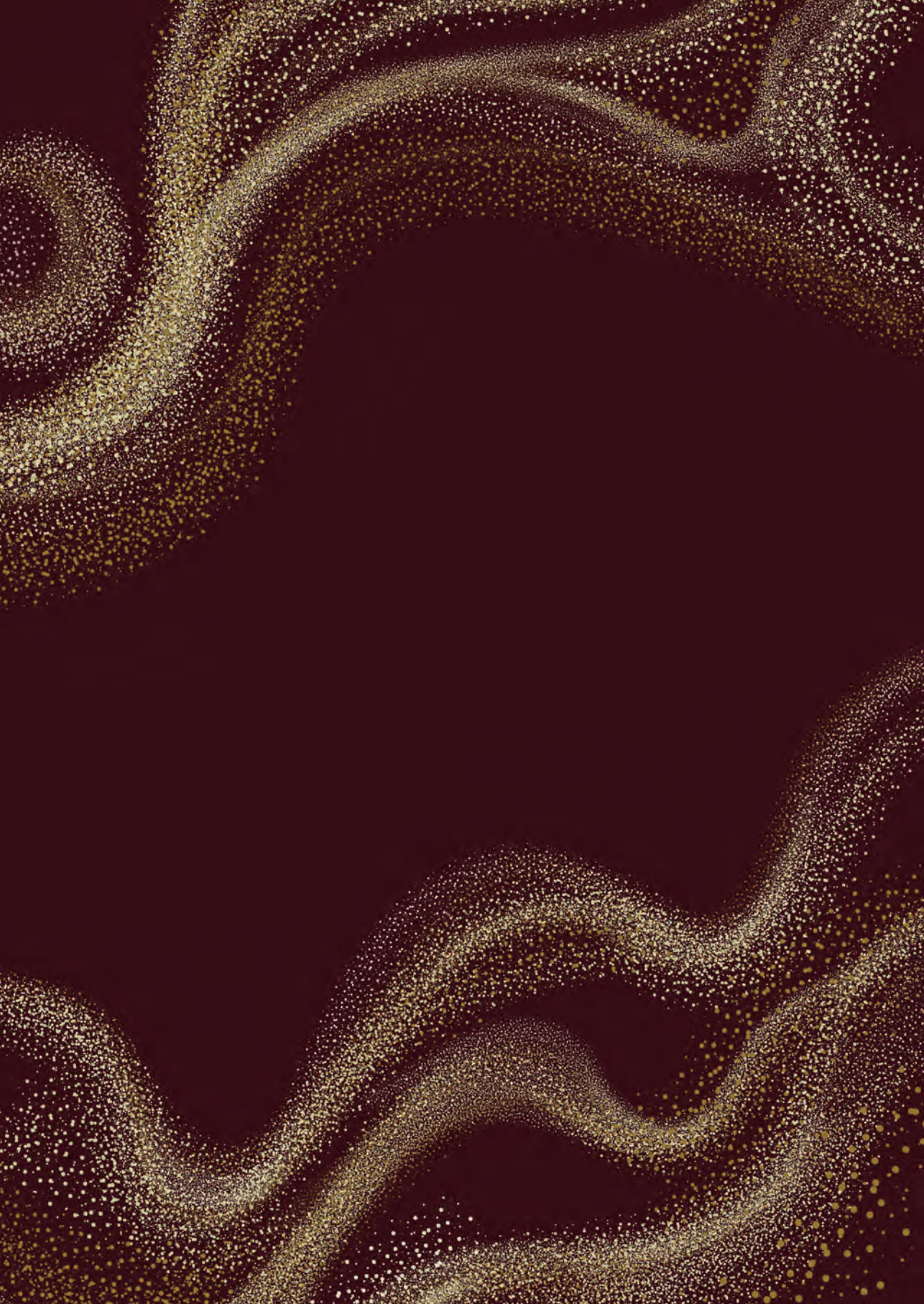
Milou Lisa Marit van Riswijk

Promotor:
Prof. dr. P.D. Siersema

Copromotor:
Dr. A.C.I.T.L. Tan (CWZ)

Manuscriptcommissie:
Prof. dr. M.J.M. Broeders
prof. dr. H.J. Schers
prof. dr. E. Dekker (Amsterdam UMC)

Table of contents



Chapter 1

General introduction and thesis outline

You have only scratched the surface
Of what you're capable of
There are decades
Of victories ahead of you

-Rupi Kaur

Colorectal cancer is a highly prevalent disease with high mortality if detected at a late stage. Screening is effective for both early detection and prevention by removing precancerous lesions. However, several challenges have arisen that reduce the full impact of screening programs. First, screening accessibility is affected by high colonoscopy loads and experienced barriers by patients due to fear of the procedure and the necessary bowel preparation before colonoscopy. Second, a sub-optimally performed colonoscopy may result in missed lesions. This thesis aims to reduce preventable cases of colorectal cancer by exploring alternative screening tests and improving bowel preparation quality.

Colorectal cancer

Colorectal cancer (CRC) ranks third in cancer incidences and is the fourth leading cause of cancer mortality globally (1, 2), with over 1.8 million new cases and approximately 881,000 deaths annually (3). Under the age of 50, it has become the second most common cause of cancer-related death in women and the most common cause in men (4). While the burden of CRC is more pronounced in high-income countries, its prevalence is also rapidly rising in low- and middle-income regions. This rise can be attributed to lifestyle changes, including diet, physical inactivity, and increased life expectancy, which collectively contribute to the growing burden of CRC worldwide. Modifiable risk factors for CRC are associated with a Western lifestyle and include a diet high in red and processed meats, low fiber intake, physical inactivity, obesity, smoking, and high alcohol consumption (5-9). A particularly concerning illustration of this is the rise of early-onset CRC, which is increasing among younger populations under the age of 50 (10) and often presents at more advanced stages (11). The majority of early-onset CRC is not caused by genetic predisposition or hereditary conditions but driven by lifestyle and environmental factors (11-13).

CRC develops from premalignant adenomas and serrated lesions, which gradually progress to malignancy in 10-15 years if left undetected and untreated (14-16). Although a majority of CRC develops from adenomas, fewer than 5% of them develop into cancer. Additionally, the serrated-neoplasia pathway accounts for 10-20% of CRCs (14). A distinct carcinoma pathway is observed in cases linked to Lynch syndrome, where defects in mismatch repair genes result in microsatellite unstable cancers (17). Early detection of CRC is crucial as it significantly improves the chances of successful treatment and survival through downstaging. Removing premalignant lesions during screening is a key preventive measure against CRC development.

Screening for colorectal cancer

Given the importance of early detection, several countries have implemented a population-based screening program for CRC (18). The main benefits of CRC screening are the early detection and resection of premalignant lesions, thereby preventing invasive growth, and tumor detection at an earlier stage (downstaging), so less invasive treatment is required (19, 20). This is illustrated by the decline in CRC incidence after the implementation of the population based screening program in the Netherlands (21-23) and elsewhere (24, 25), ultimately resulting in averted deaths from CRC (26).

Several screening modalities are available, including direct visualization methods such as colonoscopy or sigmoidoscopy (20, 27), as well as non-invasive tests like CT-colonography, fecal occult blood test (FOBT), fecal immunochemical test (FIT), and multitarget stool DNA test (mtDNA), and blood serum tests analyzing circulating tumor DNA or SEPT9 methylation (28-33) (Figure 1). As the goal for CRC screening is not only early detection but also removal of precursor lesions, all non-endoscopic screening modalities, including FIT, mtDNA, and blood tests, require a follow-up colonoscopy in case of a positive result. The accuracy of the different screening modalities varies considerably, particularly for premalignant lesions. For example, for a FIT test from different manufacturers, the sensitivity and specificity of a single FIT is 10.1-37.7% and 85.5-96.6%, respectively (34). However, FIT performance significantly increases when performed in a repeated fashion, especially for advanced adenoma (35, 36). In contrast, the SEPT9 blood test and gFOBT stool test only have an 11.2% and 15% sensitivity for detecting advanced premalignant lesions (37, 38), whereas the mtDNA stool test has higher test characteristics (sensitivity 43.4% specificity 90.6%) (29). Lastly, the recently introduced cell-free DNA blood test reported a 13% sensitivity and 90% specificity for advanced neoplasia (30). Current guidelines do not uniformly recommend one modality (Table 1), but in general favor colonoscopy screening (39-43). In the Netherlands, adults aged 55-75 are invited for CRC screening by FIT with a positivity threshold of 47ug/g feces and colonoscopy in case of a positive FIT (23). Other countries use an opportunistic approach, offering direct colonoscopy screening.

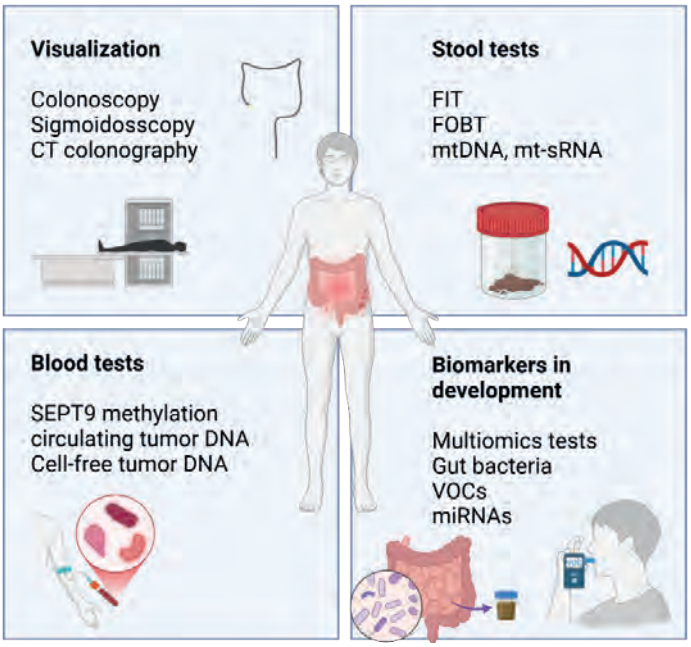


Figure 1. Established and developmental screening modalities for colorectal cancer screening. **CT** - Computed Tomography, **FIT** - Fecal Immunochemical Test, **FOBT** - Fecal Occult Blood Test, **miRNA** - micro RNA, **mtDNA** – multitarget stool DNA test, **mt-sRNA** – multitarget stool RNA test, **SEPT9** - septin 9 gene methylation, **VOC** – volatile organic compounds.

The gold standard for CRC screening is colonoscopy, as it offers both diagnostic and therapeutic qualities and has a high sensitivity for detecting malignant and premalignant lesions (40, 43, 44). However, colonoscopy is not perfect as it may miss lesions. The adenoma miss rate is estimated to be 15% to 37% in repeated colonoscopies (45-49). Precancerous lesions that are missed could result in post-colonoscopy CRC (PCCRC). PCCRC is defined as CRC detected shortly after screening colonoscopy, usually between 6 months and 3 years. PCCRC rates are estimated at 2.7% in the Netherlands (50) and 2.2%-7.0% elsewhere (51-54). One of the primary causes of PCCRC is the failure to detect polyps or early-stage cancer during colonoscopy. This is usually due to one or more of the following reasons: inadequate bowel preparation, endoscopist skill, and location and characteristics of the polyps (e.g., flat lesions are harder to detect) (55). Other causes include incomplete polyp removal or, less commonly, rapidly developing cancer. Therefore, the occurrence of PCCRC is a critical quality indicator for colonoscopy procedures. High rates of PCCRC may indicate issues with the quality of the colonoscopy procedure, such as poor technique, inadequate bowel preparation, or insufficient follow-up care. To minimize the risk of PCCRC, it is essential to ensure high-quality colonoscopy procedures (50).

Table 1. Overview of screening methods by different society guidelines.

	ESGE	ASGE	USPSTF	BSG
Age recommendations	50 years, with adjustments based on risk factors and local conditions.	45 years	50-75 (grade A*) 45-75 (grade B**)	50 years
Screening methods	FIT, colonoscopy, CT colonography, sigmoidoscopy	FOBT, FIT, stool DNA test, CT colonography, sigmoidoscopy	FIT, sigmoidoscopy, colonoscopy, CT colonography, stool DNA test	FIT, colonoscopy, sigmoidoscopy, CT colonography
Ending age for screening	Age 75 years or earlier if life expectancy is less than 10 years.	Based on individual health and prior screening results.	Age 75 years but may continue to age 85 based on individual health status and prior screening history.	Age 74 years
Follow-up	Every 10 years for colonoscopy, every 1-2 years for FIT	Every 10 years for colonoscopy, every 1 year for FIT	Every 10 years for colonoscopy, every 1-3 years for FIT	Every 10 years for colonoscopy, every 1-2 years for FIT

ASGE - American Society for Gastrointestinal Endoscopy, **BSG** - British Society of Gastroenterology, **CT** - Computed Tomography, **ESGE** - European Society of Gastrointestinal Endoscopy, **FIT** - Fecal Immunochemical Test, **FOBT** - Fecal Occult Blood Test, **USPSTF** - United States Preventive Services Task Force. * Grade A level evidence: high evidence of substantial net benefit. ** Grade B level evidence: High certainty of moderate net benefit or moderate certainty of moderate to substantial net benefit.

Challenges in colorectal cancer screening

Building on the established efficacy of current CRC screening methods, this thesis will explore innovative approaches to overcome existing challenges. Despite the demonstrated effectiveness of CRC screening in reducing mortality, several significant barriers persist (Figure 2) (56).

The current two-step screening procedure using a FIT and, in case of a positive FIT result, colonoscopy has become the CRC-screening program in many European countries (40). However, the effect of screening on CRC incidence and mortality hinges on both high participation rates, high accuracy of the FIT, and high-quality colonoscopy completion (57). As FIT serves as a selection tool for colonoscopy, the optimal selection should lead to no missed advanced precancerous lesions and cancers, but also no unnecessary (negative) procedures as the latter may result in high colonoscopy workloads and waiting times. However, barriers to screening participation, often driven by fear or discomfort and the invasive nature of traditional screening methods like colonoscopy, ultimately result in reduced overall screening uptake (58). These barriers could risk reducing the potential impact of CRC prevention programs as the positive effect of CRC screening can only be effectuated

at sufficiently high adherence rates (59). An international trial investigating the effect of invitation for colonoscopy screening on CRC mortality demonstrated that the benefit of colonoscopy screening is severely lowered if screening uptake is low (60). Therefore, screening should be accessible, and care should be taken to address patient groups known to have a lower uptake, such as those with a lower socio-economic status or cultural differences (61-63). Addressing these challenges requires innovative approaches that can improve both accessibility and patient experience without compromising the efficacy of detection.

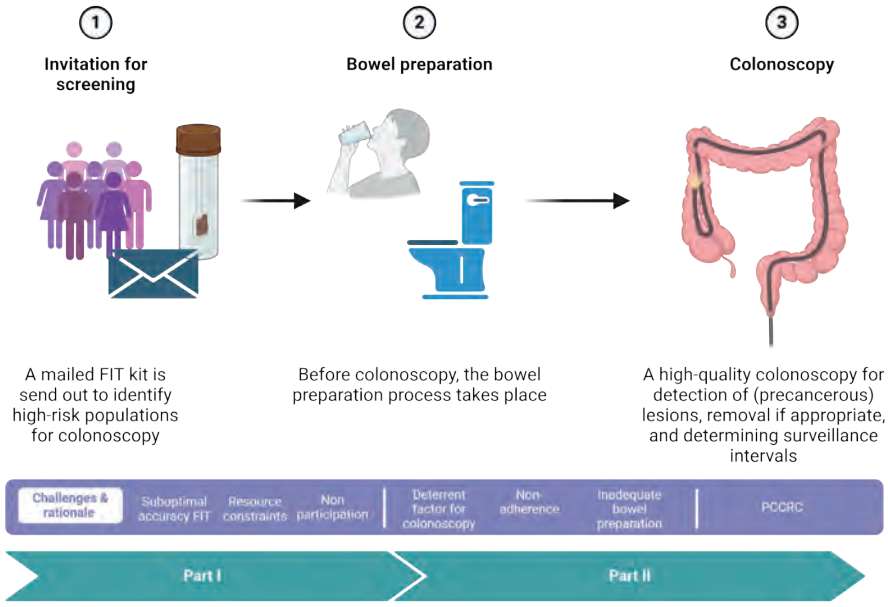


Figure 2. Challenges in the colorectal cancer (CRC) screening process with rationale behind this thesis. **FIT** – fecal immunochemical test; **PCCRC** – post-colonoscopy colorectal cancer.

In addition to adherence issues, the diagnostic accuracy of FIT remains limited. According to a recent Cochrane meta-analysis, FIT demonstrates a sensitivity of 76-89% and a specificity of 94% for detection of CRC (38), resulting in approximately 11-24% of cancers being missed while also leading to avoidable colonoscopies (38, 64, 65). For every 10,000 colonoscopies, this translates to nearly 600 unnecessary procedures (38). These excess procedures impose significant burdens on patients, including anxiety, potential adverse events, and the inconvenience of bowel preparation. Moreover, they strain healthcare systems by escalating costs and contributing to longer waiting times.

Furthermore, FIT mainly identifies cancer by detecting bleeding from lesions, but its accuracy for detecting premalignant polyps that may not bleed is limited. FIT demonstrates a sensitivity of 33% and a specificity of 93% for adenomas, while the sensitivity for advanced serrated lesions is only 5.1%-18.4% (38, 66, 67). This lack of precision can result in delays in diagnosis and treatment.

Several initiatives to improve selection for CRC screening colonoscopy have been considered, such as risk stratification of FIT with hemoglobin concentrations, sex, age, or family history (68-70), but this did not result in improved use and yield of colonoscopy resources. Other biomarkers, including multitarget DNA stool test or Septin 9, have also not reached wide adoption in screening programs (43, 71). Very recently, a new blood test measuring cell-free DNA has the promise to be an easy-to-apply screening test to use nested in routine blood draws (30). However, in modeling studies, (repeated) FIT was still more cost-effective and resulted in lower CRC incidence and mortality (72, 73).

Altogether, an improved selection procedure for screening colonoscopies would be welcomed. Volatile organic compounds (VOCs) have shown potential as non-invasive biomarkers for CRC detection and can be identified in various bodily secretions, including breath, urine, and stool (74, 75). VOCs are gaseous substances produced during normal and altered metabolic processes in the body (76-79). Carcinogenesis is thought to induce specific changes in metabolism and the microbiome, resulting in distinctive VOC patterns in affected individuals (80-82). These can be detected by chemical analysis methods such as gas-chromatography-mass-spectrometry (GCMS) or ion mobility spectrometry (IMS) that are able to detect the specific substances or using sensor-based pattern recognition methods with machine learning (83). Although GCMS and IMS provide detailed analysis, they are resource intensive. On the other hand, sensor-based electronic nose (e-nose) devices integrated with machine learning algorithms may offer the possibility of a low-cost, point-of-care diagnostic tool, which could make screening more accessible.

Systematic reviews and meta-analyses have highlighted the potential of VOCs as diagnostic tools for CRC (74, 84, 85). In the most recent meta-analysis, Wang et al. (2024) included 32 studies that evaluated VOC-based diagnostic tests for CRC (84). The analysis reported a pooled area under the receiver operating characteristics curve (AUC) of 0.93, reflecting a strong predictive capability. Nonetheless, to date, limited large-scale studies investigating VOC analysis for CRC detection have been published, while studies vary significantly in both sample collection- and analysis

methods (85). This lack of standardization limits comparison between studies, leaving the true value unclear. Moreover, external validation is needed before VOC analysis can be adopted in a clinical setting (86).

In the first part of this thesis, we therefore focus on VOCs as a diagnostic biomarker to be used in CRC screening.

Quality of colorectal cancer screening

The effectiveness of the two-step CRC screening is not only dependent on the FIT (first step) but also relies on colonoscopy to effectively diagnose and treat CRC and precursor lesions.

The effectiveness of colonoscopy for CRC screening is inherently tied to the quality of the procedure. As the main aims of CRC screening are early detection and prevention of CRC, the previously discussed measure PCCRC is a critical quality parameter (87, 88). However, as this is only measurable after several years (i.e., it is a lag-time parameter instead of a lead parameter), the most important surrogate quality parameter of colonoscopy is the adenoma detection rate (ADR) (87, 89), as it has shown an inverse correlation with a higher risk of PCCRC both in primary screening colonoscopies and in FIT-positive screening colonoscopies (89, 90). For example, this means that for every 1,000 FIT-positive screening colonoscopies, endoscopists with an ADR of 70% are expected to diagnose approximately 2 PCCRCs within five years. In contrast, endoscopists with ADRs of 65%, 60%, and 55% are expected to diagnose approximately 2.5, 3.5, and 4.5 PCCRCs, respectively, over the same period (90). In addition to ADR, other quality parameters for colonoscopy include withdrawal time (91), cecal intubation rate (92), appropriate polypectomy technique, polyp retrieval rate, complication rate monitoring, patient experience, and appropriate post-polypectomy surveillance intervals (43, 93). The recently renewed quality guidelines by the ASGE have also introduced Sessile Serrated Lesion Detection Rate (SSLDR) (performance threshold 6% or higher) and adenomas per colonoscopy (APC) (performance threshold 0.6) as a quality indicator (94). Maintaining and monitoring quality parameters have been shown to lead to a demonstrable increase in colonoscopy quality (95).

These quality parameters, and thereby the diagnostic effectiveness of colonoscopy, also depend on the quality of bowel preparation. As some lesions may be relatively subtle and flat, suboptimal bowel preparation significantly may impact the ability to detect polyps (96, 97). Furthermore, inadequate bowel preparation is associated with lower cecal intubation rates (93, 96), higher complication rates (98, 99), longer

procedure times (100), and contributes to higher costs, inconvenience and lower patient satisfaction due to the need to reschedule procedures (100). Thus, high-quality bowel preparation is paramount for colonoscopy and CRC screening.

Achieving effective bowel cleansing presents several challenges. Firstly, the bowel preparation process is often intense and burdensome for patients, frequently cited as a significant deterrent to undergoing colonoscopy (101, 102). The need to consume large volumes of preparation fluid, along with strict dietary modifications, increases the challenge for patients. To address these issues, a low-residue diet instead of a clear liquid diet (103, 104) and decreasing volumes from 4L to 2L bowel preparation fluid have been shown to improve tolerability while still ensuring adequate bowel preparation (105). Recently, even lower volume bowel preparations of 1L or less have been developed. However, evidence on the prerequisite efficacy of those solutions is not ubiquitously reported (106). Additionally, the extent of the burden of intermediate or low-volume bowel preparations for patients has been poorly investigated, which is even more the case for low-volume preparations of 1L or less.

The second challenge is that inadequate bowel preparation occurs in up to 20% of colonoscopies (107). Several factors may contribute to this, related to clinical factors that interfere with bowel motility or non-compliance with taking the bowel preparation laxatives, such as a lower education level and poor health literacy (108, 109). Risk factors that have been identified include advanced age, male sex, high body mass index, high American Society of Anesthesiology physical status classification system score (ASA score), polypharmacy, tricyclic antidepressant (TCA) use, opioid use, diabetes, liver cirrhosis, chronic constipation, history of neurologic disease (stroke, spina bifida, dementia, paraplegia, or Parkinson's disease), history of intra-abdominal and/or pelvic surgery, current hospitalization, previous colonoscopies, and history of inadequate bowel preparation (107, 109-114).

By identifying patients who are at risk for insufficient bowel preparation, timely additional measures can be taken to optimize the preparation. This may either be an intensified regime for patients with biological risk factors, or extra attention and education for patient related non-compliance risk factors. Unfortunately, some patients remain repeatedly inadequately prepared for colonoscopy. Evidence on how to proceed with those patients is scarce due to the heterogeneity of the group (97, 115). In routine practice, patients are frequently prescribed an intensified regimen that demands an even greater fluid intake or are admitted to the hospital for following a supervised bowel preparation schedule. This adds to an even greater challenge to patients, which often contributes to inadequate preparation risk. Intraprocedural bowel

cleansing with a specialized device eliminates the reliance on proper preprocedural bowel preparation (116) but has not been assessed in this complex patient group.

Therefore, in the second part of this thesis, we focus on bowel preparation methods that reduce patient discomfort while maintaining or improving the effectiveness of colon cleansing.

Aims of this thesis

The ultimate objective of CRC screening is to reduce the incidence and mortality of this malignancy. This thesis aims to contribute to this goal by exploring the potential of VOC analysis via e-nose technology for CRC detection, and by assessing whether bowel preparation for colonoscopy can be optimized to reduce patient burden while maintaining efficacy. An overview of the research questions and methodologies is presented in Table 2.

Part I of this thesis investigates the feasibility of using VOC analysis through e-nose technology for CRC detection within screening settings. **Chapter 2** presents a multicenter cross-sectional cohort study involving patients invited for screening colonoscopy following a positive FIT result. We assessed the accuracy and reproducibility of exhaled breath analysis by an e-nose by calculating the area under the receiver operating characteristics curve, and sensitivity and specificity. Reproducibility was evaluated in a subset of patients by calculating Cohens' Kappa. Additionally, we assessed patient experience with using the breath test. **Chapter 3** discusses methodological problems encountered during e-nose clinical studies by designing a simulated dataset based on real breath test signals. We assessed the effect of varying neoplasia prevalences, augmented data, and the use of different e-nose devices throughout a study on the performance of the breath test.

In **part II**, we assess if bowel preparation for colonoscopy can be improved by lowering the burden to patients. To this end, in **Chapter** 4, we first performed a systematic review and meta-analysis of the bowel cleansing efficacy of different bowel preparation solutions of 1L or less. We pooled the rate of adequately cleansed patients per type of laxative and provided subgroup analyses for different dosing and diet regimes, and use of additive laxatives. **Chapter 5** presents an open-label multicenter randomized controlled trial comparing a 2L bowel preparation to a specialized preparation of 1L. First, we determined the non-inferiority of a 1L bowel preparation laxative compared to the 2L laxative. Additionally, we assessed the impact of the amount of bowel preparation on patients using validated questionnaires to determine the tolerability and willingness to repeat the bowel

preparation, health-related quality of life differences, and provide insight into work absenteeism and impaired productivity. Finally, in **Chapter 6**, we report the results of a multicenter feasibility study to assess if adequate bowel preparation can be achieved by using an intraprocedural bowel cleansing device in patient who are known with recurrent inadequate bowel preparation.

In **Chapter 8**, the findings of this thesis will be contextualized within the context of literature on colorectal cancer screening. We will explore future perspectives and the implications of these findings for optimizing screening strategies, ultimately aiming to enhance both the efficacy and patient experience of CRC prevention.

Table 2. Summary of aims and methodology used in this thesis.

Chapter	Research question	Methodology
Part I: VOC analysis for CRC screening		
2	What is the accuracy and reproducibility of VOC analysis by electronic nose to diagnose CRC?	Multicenter cross-sectional prospective cohort study including 3469 patients
3	How can methodological issues in e-nose research be overcome?	Simulated data of e-nose breath tests
Part II: Improvement of bowel preparation for colonoscopy		
4	What is the efficacy of bowel preparation solutions of 1L or less?	Systematic review and meta-analysis
5	What is the impact of 1L and 2L bowel preparation on tolerability, health-related quality of life, and working live?	Multicenter open-label, non-inferiority randomized controlled trial including questionnaires before and after bowel preparation
6	Can a bowel cleansing system be used in patients who are difficult to prepare?	Multicenter feasibility study

References

1. Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, et al. Colorectal cancer statistics, 2020. CA Cancer J Clin. 2020;70(3):145-64.
2. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. Gut. 2017;66(4):683-91.
3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
4. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA: A Cancer Journal for Clinicians. 2024;74(1):12-49.
5. Mertens E, Keuchkarian M, Vasquez MS, Vandevijvere S, Peñalvo JL. Lifestyle predictors of colorectal cancer in European populations: a systematic review. BMJ Nutr Prev Health. 2024;7(1):183-90.
6. Bolwell J, Butler R, Burke CA, Liska D, Macaron C. Risk Factors Associated with Advanced Colorectal Neoplasia in Adults Younger than Age 45. J Clin Gastroenterol. 2024.
7. Murphy CC, Zaki TA. Changing epidemiology of colorectal cancer - birth cohort effects and emerging risk factors. Nat Rev Gastroenterol Hepatol. 2023.
8. Mandic M, Safizadeh F, Niedermaier T, Hoffmeister M, Brenner H. Association of Overweight, Obesity, and Recent Weight Loss With Colorectal Cancer Risk. JAMA Network Open. 2023;6(4):e239556-e.
9. Wang K, Lo C-H, Mehta RS, Nguyen LH, Wang Y, Ma W, et al. An Empirical Dietary Pattern Associated With the Gut Microbial Features in Relation to Colorectal Cancer Risk. Gastroenterology. 2024;167(7):1371-83.e4.
10. Sung H, Siegel RL, Laversanne M, Jiang C, Morgan E, Zahwe M, et al. Colorectal cancer incidence trends in younger versus older adults: an analysis of population-based cancer registry data. Lancet Oncol. 2024.
11. Patel SG, Karlitz JJ, Yen T, Lieu CH, Boland CR. The rising tide of early-onset colorectal cancer: a comprehensive review of epidemiology, clinical features, biology, risk factors, prevention, and early detection. The Lancet Gastroenterology & Hepatology. 2022;7(3):262-74.
12. Kanth P, Inadomi JM. Screening and prevention of colorectal cancer. BMJ. 2021;374:n1855.
13. Song J, Han T, Qian L, Zhu J, Qiao Y, Liu S, et al. A decade-long study on pathological distinctions of resectable early versus late onset colorectal cancer and optimal screening age determination. Sci Rep. 2024;14(1):27335.
14. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. Lancet. 2019;394(10207):1467-80.
15. Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, et al. Genetic alterations during colorectal-tumor development. N Engl J Med. 1988;319(9):525-32.
16. Snover DC. Update on the serrated pathway to colorectal carcinoma. Hum Pathol. 2011;42(1):1-10.
17. Nguyen LH, Goel A, Chung DC. Pathways of Colorectal Carcinogenesis. Gastroenterology. 2020;158(2):291-302.
18. Schreuders EH, Ruco A, Rabeneck L, Schoen RE, Sung JJ, Young GP, Kuipers EJ. Colorectal cancer screening: a global overview of existing programmes. Gut. 2015;64(10):1637-49.

19. Winawer SJ, Zauber AG, Ho MN, O'Brien MJ, Gottlieb LS, Sternberg SS, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. N Engl J Med. 1993;329(27):1977-81.
20. Zauber AG, Winawer SJ, O'Brien MJ, Lansdorp-Vogelaar I, van Ballegooijen M, Hankey BF, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. N Engl J Med. 2012;366(8):687-96.
21. IKNL R. Landelijke Monitoring Bevolkingsonderzoek darmkanker 2019. https://www.rivm.nl/documenten/landelijke-monitoring-bevolkingsonderzoek-darmkanker-20192020.
22. Nederlandse Kankerregistratie (NKR) IKNI. Incidentie darmkanker 2019 [updated 16-04-2019; cited 2020 24 march]. Available from: https://www.iknl.nl/kankersoorten/darmkanker/registratie/incidentie.
23. Breekveldt ECH, Lansdorp-Vogelaar I, Toes-Zoutendijk E, Spaander MCW, van Vuuren AJ, van Kemenade FJ, et al. Colorectal cancer incidence, mortality, tumour characteristics, and treatment before and after introduction of the faecal immunochemical testing-based screening programme in the Netherlands: a population-based study. Lancet Gastroenterol Hepatol. 2022;7(1):60-8.
24. Rex DK. The Case for High-Quality Colonoscopy Remaining a Premier Colorectal Cancer Screening Strategy in the United States. Gastrointest Endosc Clin N Am. 2020;30(3):527-40.
25. Levin TR, Corley DA, Jensen CD, Schottinger JE, Quinn VP, Zauber AG, et al. Effects of Organized Colorectal Cancer Screening on Cancer Incidence and Mortality in a Large Community-Based Population. Gastroenterology. 2018;155(5):1383-91.e5.
26. Goddard KAB, Feuer EJ, Mandelblatt JS, Meza R, Holford TR, Jeon J, et al. Estimation of Cancer Deaths Averted From Prevention, Screening, and Treatment Efforts, 1975-2020. JAMA Oncol. 2024.
27. Juul FE, Cross AJ, Schoen RE, Senore C, Pinsky PF, Miller EA, et al. Effectiveness of Colonoscopy Screening vs Sigmoidoscopy Screening in Colorectal Cancer. JAMA Netw Open. 2024;7(2):e240007.
28. Shaukat A, Levin TR. Current and future colorectal cancer screening strategies. Nat Rev Gastroenterol Hepatol. 2022;19(8):521-31.
29. Imperiale TF, Porter K, Zella J, Gagrat ZD, Olson MC, Statz S, et al. Next-Generation Multitarget Stool DNA Test for Colorectal Cancer Screening. New England Journal of Medicine. 2024;390(11):984-93.
30. Chung DC, Gray DM, Singh H, Issaka RB, Raymond VM, Eagle C, et al. A Cell-free DNA Blood-Based Test for Colorectal Cancer Screening. New England Journal of Medicine. 2024;390(11):973-83.
31. Hardcastle JD, Chamberlain JO, Robinson MH, Moss SM, Amar SS, Balfour TW, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet. 1996;348(9040):1472-7.
32. Robertson DJ, Lee JK, Boland CR, Dominitz JA, Giardiello FM, Johnson DA, et al. Recommendations on Fecal Immunochemical Testing to Screen for Colorectal Neoplasia: A Consensus Statement by the US Multi-Society Task Force on Colorectal Cancer. Gastroenterology. 2017;152(5):1217-37.e3.
33. Obaro AE, Plumb AA, Halligan S, Mallett S, Bassett P, McCoubrie P, et al. Colorectal Cancer: Performance and Evaluation for CT Colonography Screening- A Multicenter Cluster-randomized Controlled Trial. Radiology. 2022;303(2):361-70.
34. Levy BT, Xu Y, Daly JM, Hoffman RM, Dawson JD, Shokar NK, et al. Comparative Performance of Common Fecal Immunochemical Tests. Annals of Internal Medicine. 2024;177(10):1350-60.

35. Gerrard AD, Maeda Y, Miller J, Gunn F, Theodoratou E, Noble C, et al. Double faecal immunochemical testing in patients with symptoms suspicious of colorectal cancer. Br J Surg. 2023;110(4):471-80.

36. Farkas NG, Palyvos L, O'Brien JW, Yu KS, Pigott C, Whyte M, et al. The repeat FIT (RFIT) study: Does repeating faecal immunochemical tests provide reassurance and improve colorectal cancer detection? Colorectal Dis. 2024.

37. Church TR, Wandell M, Lofton-Day C, Mongin SJ, Burger M, Payne SR, et al. Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer. Gut. 2014;63(2):317-25.

38. Grobbee EJ, Wisse PH, Schreuders EH, van Roon A, van Dam L, Zauber AG, et al. Guaiac-based faecal occult blood tests versus faecal immunochemical tests for colorectal cancer screening in average-risk individuals. Cochrane Database Syst Rev. 2022;6(6):Cd009276.

39. Davidson KW, Barry MJ, Mangione CM, Cabana M, Caughey AB, Davis EM, et al. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. Jama. 2021;325(19):1965-77.

40. Saftoiu A, Hassan C, Areia M, Bhutani MS, Bisschops R, Bories E, et al. Role of gastrointestinal endoscopy in the screening of digestive tract cancers in Europe: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. Endoscopy. 2020;52(4):293-304.

41. Patel SG, May FP, Anderson JC, Burke CA, Dominitz JA, Gross SA, et al. Updates on Age to Start and Stop Colorectal Cancer Screening: Recommendations From the U.S. Multi-Society Task Force on Colorectal Cancer. Am J Gastroenterol. 2022;117(1):57-69.

42. Monahan KJ, Davies MM, Abulafi M, Banerjea A, Nicholson BD, Arasaradnam R, et al. Faecal immunochemical testing (FIT) in patients with signs or symptoms of suspected colorectal cancer (CRC): a joint guideline from the Association of Coloproctology of Great Britain and Ireland (ACPGBI) and the British Society of Gastroenterology (BSG). Gut. 2022.

43. Shaukat A, Kahi CJ, Burke CA, Rabeneck L, Sauer BG, Rex DK. ACG Clinical Guidelines: Colorectal Cancer Screening 2021. Am J Gastroenterol. 2021;116(3):458-79.

44. Bisschops R, Rutter MD, Areia M, Cristiano S, Domagk D, Kaminski MF, et al. Overcoming the barriers to dissemination and implementation of quality measures for gastrointestinal endoscopy: European Society of Gastrointestinal Endoscopy (ESGE) and United European Gastroenterology (UEG) position statement. United European Gastroenterol J. 2021;9(1):120-6.

45. van Rijn JC, Reitsma JB, Stoker J, Bossuyt PM, van Deventer SJ, Dekker E. Polyp miss rate determined by tandem colonoscopy: a systematic review. Am J Gastroenterol. 2006;101(2):343-50.

46. Wallace MB, Sharma P, Bhandari P, East J, Antonelli G, Lorenzetti R, et al. Impact of Artificial Intelligence on Miss Rate of Colorectal Neoplasia. Gastroenterology. 2022.

47. Jiang W, Xin L, Zhu S, Liu Z, Wu J, Ji F, et al. Risk Factors Related to Polyp Miss Rate of Short-Term Repeated Colonoscopy. Dig Dis Sci. 2023;68(5):2040-9.

48. Shah S, Park N, Chehade NEH, Chahine A, Monachese M, Tiritilli A, et al. Effect of computer-aided colonoscopy on adenoma miss rates and polyp detection: A systematic review and meta-analysis. J Gastroenterol Hepatol. 2023;38(2):162-76.

49. Zhao S, Song Y, Wang S, Wang R, Feng Z, Gong A, et al. Reduced Adenoma Miss Rate With 9-Minute vs 6-Minute Withdrawal Times for Screening Colonoscopy: A Multicenter Randomized Tandem Trial. Official journal of the American College of Gastroenterology | ACG. 2023;118(5):802-11.

50. Wisse PHA, de Boer SY, Oudkerk Pool M, Terhaar sive Droste JS, Verveer C, Meijer GA, et al. Post-colonoscopy colorectal cancers in a national fecal immunochemical test-based colorectal cancer screening program. Endoscopy. 2024;56(05):364-72.

51. Lee JK, Kang JH, Merchant SA, Jensen CD, Burr NE, Corley DA. Overall and Annual Postcolonoscopy Colorectal Cancer Rates in a Large Integrated Healthcare Setting: A Cross-Sectional Study. Clin Gastroenterol Hepatol. 2024;22(1):188-90.e5.
52. Baile-Maxía S, Mangas-Sanjuan C, Sala-Miquel N, Barquero C, Belda G, García-Del-Castillo G, et al. Incidence, characteristics, and predictive factors of post-colonoscopy colorectal cancer. United European Gastroenterol J. 2024.
53. Lamba M, Khaing MM, Ma X, Ryan K, Appleyard M, Leggett B, Grimpen F. Post-colonoscopy cancer rate at a tertiary referral hospital in Australia: A data linkage analysis. J Gastroenterol Hepatol. 2023;38(5):740-6.
54. Sekiguchi M, Kishida Y, Ikematsu H, Konno M, Mizuguchi Y, Hotta K, et al. Proportions and characteristics of interval cancer in annual fecal immunochemical test screening and postcolonoscopy colorectal cancer: Results from a Japanese multicenter prospective study using questionnaires, the C-DETECT study. Dig Endosc. 2024.
55. Ahmad OF, Mazomenos E, Chadebecq F, Kader R, Hussein M, Haidry RJ, et al. Identifying key mechanisms leading to visual recognition errors for missed colorectal polyps using eye-tracking technology. J Gastroenterol Hepatol. 2023;38(5):768-74.
56. Chuang MP-C, Chiu H-M. Does Colonoscopy as a First Screening Test Still Make Sense?—Counterpoint. Digestive Diseases and Sciences. 2024.
57. Lieberman D, Dekker E. How Good Is Good Enough? What Should Be the Target for CRC Screening? Digestive Diseases and Sciences. 2024.
58. Hamer-Kiwacz S, Berntsson H, Galloway G, Tran AM, Tan JY, Hind D, Kurien M. Barriers and facilitators to faecal immunochemical testing in symptomatic populations: A rapid systematic scoping review and gap analysis. J Eval Clin Pract. 2024.
59. Lansdorp-Vogelaar I, Kuntz KM, Knudsen AB, van Ballegooijen M, Zauber AG, Jemal A. Contribution of Screening and Survival Differences to Racial Disparities in Colorectal Cancer Rates. Cancer Epidemiology, Biomarkers & Prevention. 2012;21(5):728-36.
60. Bretthauer M, Løberg M, Wieszczy P, Kalager M, Emilsson L, Garborg K, et al. Effect of Colonoscopy Screening on Risks of Colorectal Cancer and Related Death. New England Journal of Medicine. 2022.
61. Ola I, Cardoso R, Hoffmeister M, Brenner H. Utilization of colorectal cancer screening tests across European countries: a cross-sectional analysis of the European health interview survey 2018-2020. Lancet Reg Health Eur. 2024;41:100920.
62. Belon AP, McKenzie E, Teare G, Nykiforuk CIJ, Nieuwendyk L, Kim MO, et al. Effective strategies for Fecal Immunochemical Tests (FIT) programs to improve colorectal cancer screening uptake among populations with limited access to the healthcare system: a rapid review. BMC Health Serv Res. 2024;24(1):128.
63. van de Schootbrugge-Vandermeer HJ, Lansdorp-Vogelaar I, de Jonge L, van Vuuren AJ, Dekker E, Spaander MCW, et al. Socio-demographic and cultural factors related to non-participation in the Dutch colorectal cancer screening programme. Eur J Cancer. 2023;190:112942.
64. MacDonald S, MacDonald L, Godwin J, Macdonald A, Thornton M. The diagnostic accuracy of the faecal immunohistochemical test in identifying significant bowel disease in a symptomatic population. Colorectal Dis. 2022;24(3):257-63.
65. Wieten E, Schreuders EH, Grobbee EJ, Nieboer D, Bramer WM, Lansdorp-Vogelaar I, et al. Incidence of faecal occult blood test interval cancers in population-based colorectal cancer screening: a systematic review and meta-analysis. Gut. 2019;68(5):873-81.
66. Chang LC, Shun CT, Hsu WF, Tu CH, Tsai PY, Lin BR, et al. Fecal Immunochemical Test Detects Sessile Serrated Adenomas and Polyps With a Low Level of Sensitivity. Clin Gastroenterol Hepatol. 2017;15(6):872-9.e1.

67. Imperiale TF, Ransohoff DF, Itzkowitz SH, Levin TR, Lavin P, Lidgard GP, et al. Multitarget stool DNA testing for colorectal-cancer screening. N Engl J Med. 2014;370(14):1287-97.

68. Auge JM, Pellise M, Escudero JM, Hernandez C, Andreu M, Grau J, et al. Risk stratification for advanced colorectal neoplasia according to fecal hemoglobin concentration in a colorectal cancer screening program. Gastroenterology. 2014;147(3):628-36.e1.

69. Stegeman I, de Wijkerslooth TR, Stoop EM, van Leerdam ME, Dekker E, van Ballegooijen M, et al. Combining risk factors with faecal immunochemical test outcome for selecting CRC screenees for colonoscopy. Gut. 2014;63(3):466-71.

70. Yeoh KG, Ho KY, Chiu HM, Zhu F, Ching JY, Wu DC, et al. The Asia-Pacific Colorectal Screening score: a validated tool that stratifies risk for colorectal advanced neoplasia in asymptomatic Asian subjects. Gut. 2011;60(9):1236-41.

71. Zygulska AL, Pierzchalski P. Novel Diagnostic Biomarkers in Colorectal Cancer. Int J Mol Sci. 2022;23(2).

72. Ladabaum U, Mannalithara A, Weng Y, Schoen RE, Dominitz JA, Desai M, Lieberman D. Comparative Effectiveness and Cost-Effectiveness of Colorectal Cancer Screening With Blood-Based Biomarkers (Liquid Biopsy) vs Fecal Tests or Colonoscopy. Gastroenterology.

73. van den Puttelaar R, Nascimento de Lima P, Knudsen AB, Rutter CM, Kuntz KM, de Jonge L, et al. Effectiveness and Cost-Effectiveness of Colorectal Cancer Screening With a Blood Test That Meets the Centers for Medicare & Medicaid Services Coverage Decision. Gastroenterology.

74. Chandrapalan S, Bosch S, Cubiella J, Guardiola J, Kimani P, Mulder C, et al. Systematic review with meta-analysis: volatile organic compound analysis to improve faecal immunochemical testing in the detection of colorectal cancer. Aliment Pharmacol Ther. 2021;54(1):14-23.

75. de Boer NK, de Meij TG, Oort FA, Ben Larbi I, Mulder CJ, van Bodegraven AA, van der Schee MP. The scent of colorectal cancer: detection by volatile organic compound analysis. Clin Gastroenterol Hepatol. 2014;12(7):1085-9.

76. Markar SR, Chin ST, Romano A, Wiggins T, Antonowicz S, Paraskeva P, et al. Breath Volatile Organic Compound Profiling of Colorectal Cancer Using Selected Ion Flow-tube Mass Spectrometry. Ann Surg. 2019;269(5):903-10.

77. Di Lena M, Porcelli F, Altomare DF. Volatile organic compounds as new biomarkers for colorectal cancer: a review. Colorectal Dis. 2016;18(7):654-63.

78. Bosch S, Berkhout DJ, Ben Larbi I, de Meij TG, de Boer NK. Fecal volatile organic compounds for early detection of colorectal cancer: where are we now? J Cancer Res Clin Oncol. 2019;145(1):223-34.

79. Tyagi H, Daulton E, Bannaga AS, Arasaradnam RP, Covington JA. Non-Invasive Detection and Staging of Colorectal Cancer Using a Portable Electronic Nose. Sensors (Basel). 2021;21(16).

80. van Vorstenbosch R, Cheng HR, Jonkers D, Penders J, Schoon E, Masclee A, et al. Systematic Review: Contribution of the Gut Microbiome to the Volatile Metabolic Fingerprint of Colorectal Neoplasia. Metabolites. 2022;13(1).

81. Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science. 2009;324(5930):1029-33.

82. Tennant DA, Duran RV, Gottlieb E. Targeting metabolic transformation for cancer therapy. Nat Rev Cancer. 2010;10(4):267-77.

83. P H, Rangarajan M, Pandya HJ. Breath VOC analysis and machine learning approaches for disease screening: a review. J Breath Res. 2023;17(2).

84. Wang Q, Fang Y, Tan S, Li Z, Zheng R, Ren Y, et al. Diagnostic performance of volatile organic compounds analysis and electronic noses for detecting colorectal cancer: a systematic review and meta-analysis. Front Oncol. 2024;14:1397259.

85. Scheepers M, Al-Difaie Z, Brandts L, Peeters A, van Grinsven B, Bouvy ND. Diagnostic Performance of Electronic Noses in Cancer Diagnoses Using Exhaled Breath: A Systematic Review and Meta-analysis. JAMA Netw Open. 2022;5(6):e2219372.
86. Leopold JH, Bos LDJ, Sterk PJ, Schultz MJ, Fens N, Horvath I, et al. Comparison of classification methods in breath analysis by electronic nose. Journal of Breath Research. 2015;9(4):046002.
87. Kaminski MF, Regula J, Kraszewska E, Polkowski M, Wojciechowska U, Didkowska J, et al. Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med. 2010;362(19):1795-803.
88. Rutter MD, Beintaris I, Valori R, Chiu HM, Corley DA, Cuatrecasas M, et al. World Endoscopy Organization consensus statements on post-colonoscopy and post-imaging colorectal cancer. Gastroenterology. 2018;155(3):909-25. e3.
89. Corley DA, Jensen CD, Marks AR, Zhao WK, Lee JK, Doubeni CA, et al. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med. 2014;370(14):1298-306.
90. Pieter H.A. Wisse M, MSc, Nicole S. Erler, PhD, Sybrand Y. de Boer, MD, PhD, Bert den Hartog, MD, PhD, Marco Oudkerk Pool, MD, PhD, Jochim S. Terhaar sive Droste, MD, PhD, Claudia Verveer, MD, Gerrit A. Meijer, MD, PhD, Iris Lansdorp-Vogelaar, PhD, Ernst J. Kuipers, MD, PhD, Evelien Dekker, MD, PhD*, and Manon C.W. Spaander, MD, PhD*. Adenoma Detection Rate and Risk for Interval Postcolonoscopy Colorectal Cancer in Fecal Immunochemical Test–Based Screening. Annals of Internal Medicine. 2022;175(10):1366-73.
91. Shaukat A, Rector TS, Church TR, Lederle FA, Kim AS, Rank JM, Allen JI. Longer Withdrawal Time Is Associated With a Reduced Incidence of Interval Cancer After Screening Colonoscopy. Gastroenterology. 2015;149(4):952-7.
92. Thoufeeq MH, Rembacken BJ. Meticulous cecal image documentation at colonoscopy is associated with improved polyp detection. Endosc Int Open. 2015;3(6):E629-33.
93. Kaminski MF, Thomas-Gibson S, Bugajski M, Bretthauer M, Rees CJ, Dekker E, et al. Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. Endoscopy. 2017;49(4):378-97.
94. Rex DK, Anderson JC, Butterly LF, Day LW, Dominitz JA, Kaltenbach T, et al. Quality indicators for colonoscopy. Gastrointest Endosc. 2024;100(3):352-81.
95. Theunissen F, Ter Borg PCJ, Ouwendijk RJT, Bruno MJ, Siersema PD. Overview of a national endoscopy database: The Trans.IT database and its impact on data registration quality. United European Gastroenterol J. 2024.
96. Froehlich F, Wietlisbach V, Gonvers JJ, Burnand B, Vader JP. Impact of colonic cleansing on quality and diagnostic yield of colonoscopy: the European Panel of Appropriateness of Gastrointestinal Endoscopy European multicenter study. Gastrointest Endosc. 2005;61(3):378-84.
97. Hassan C, East J, Radaelli F, Spada C, Benamouzig R, Bisschops R, et al. Bowel preparation for colonoscopy: European Society of Gastrointestinal Endoscopy (ESGE) Guideline - Update 2019. Endoscopy. 2019;51(8):775-94.
98. Chan AO, Lee LN, Chan AC, Ho WN, Chan QW, Lau S, Chan JW. Predictive factors for colonoscopy complications. Hong Kong Med J. 2015;21(1):23-9.
99. Kothari ST, Huang RJ, Shaukat A, Agrawal D, Buxbaum JL, Abbas Fehmi SM, et al. ASGE review of adverse events in colonoscopy. Gastrointest Endosc. 2019;90(6):863-76 e33.
100. Rex DK, Imperiale TF, Latinovich DR, Bratcher LL. Impact of bowel preparation on efficiency and cost of colonoscopy. Am J Gastroenterol. 2002;97(7):1696-700.
101. McLachlan SA, Clements A, Austoker J. Patients' experiences and reported barriers to colonoscopy in the screening context--a systematic review of the literature. Patient Educ Couns. 2012;86(2):137-46.

102. Sharara AI, El Reda ZD, Harb AH, Abou Fadel CG, Sarkis FS, Chalhoub JM, Abou Mrad R. The burden of bowel preparations in patients undergoing elective colonoscopy. United European Gastroenterol J. 2016;4(2):314-8.

103. Avalos DJ, Sussman DA, Lara LF, Sarkis FS, Castro FJ. Effect of Diet Liberalization on Bowel Preparation. South Med J. 2017;110(6):399-407.

104. Nguyen DL, Jamal MM, Nguyen ET, Puli SR, Bechtold ML. Low-residue versus clear liquid diet before colonoscopy: a meta-analysis of randomized, controlled trials. Gastrointest Endosc. 2016;83(3):499-507 e1.

105. Spadaccini M, Frazzoni L, Vanella G, East J, Radaelli F, Spada C, et al. Efficacy and Tolerability of High- vs Low-Volume Split-Dose Bowel Cleansing Regimens for Colonoscopy: A Systematic Review and Meta-analysis. Clin Gastroenterol Hepatol. 2020;18(7):1454-65 e14.

106. Pisera M, Franczyk R, Wieszczy P, Polkowski M, Rupinski M, Chaber-Ciopinska A, et al. The impact of low- versus standard-volume bowel preparation on participation in primary screening colonoscopy: a randomized health services study. Endoscopy. 2019;51(3):227-36.

107. Mahmood S, Farooqui SM, Madhoun MF. Predictors of inadequate bowel preparation for colonoscopy: a systematic review and meta-analysis. Eur J Gastroenterol Hepatol. 2018;30(8):819-26.

108. Nguyen DL, Wieland M. Risk factors predictive of poor quality preparation during average risk colonoscopy screening: the importance of health literacy. J Gastrointestin Liver Dis. 2010;19(4):369-72.

109. Kunnackal John G, Thuluvath AJ, Carrier H, Ahuja NK, Gupta E, Stein E. Poor Health Literacy and Medication Burden Are Significant Predictors for Inadequate Bowel Preparation in an Urban Tertiary Care Setting. J Clin Gastroenterol. 2019;53(9):e382-e6.

110. Dik VK, Moons LM, Huyuk M, van der Schaar P, de Vos Tot Nederveen Cappel WH, Ter Borg PC, et al. Predicting inadequate bowel preparation for colonoscopy in participants receiving split-dose bowel preparation: development and validation of a prediction score. Gastrointest Endosc. 2015;81(3):665-72.

111. Zad M, Do CN, Heffernan A, Johnston L, Al-Ansari M. Factors affecting bowel preparation adequacy and procedural time. JGH Open. 2020;4(2):206-14.

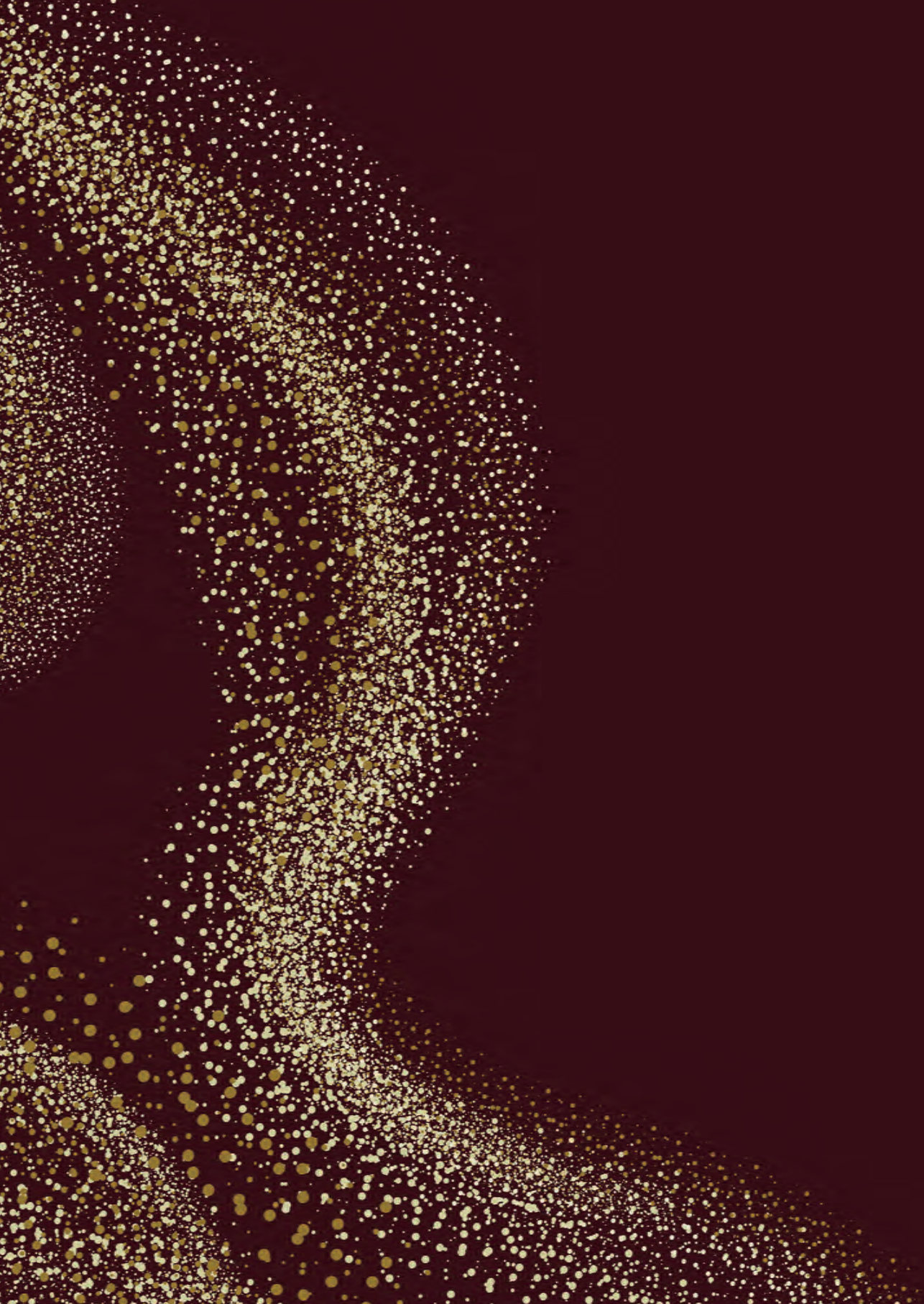
112. Gandhi K, Tofani C, Sokach C, Patel D, Kastenberg D, Daskalakis C. Patient Characteristics Associated With Quality of Colonoscopy Preparation: A Systematic Review and Meta-analysis. Clin Gastroenterol Hepatol. 2018;16(3):357-69 e10.

113. Yee R, Manoharan S, Hall C, Hayashi A. Optimizing bowel preparation for colonoscopy: what are the predictors of an inadequate preparation? Am J Surg. 2015;209(5):787-92; discussion 92.

114. Rex DK. Bowel preparation for colonoscopy: entering an era of increased expectations for efficacy. Clin Gastroenterol Hepatol. 2014;12(3):458-62.

115. Committee ASoP, Saltzman JR, Cash BD, Pasha SF, Early DS, Muthusamy VR, et al. Bowel preparation before colonoscopy. Gastrointest Endosc. 2015;81(4):781-94.

116. van Keulen KE, Neumann H, Schattenberg JM, van Esch AAJ, Kievit W, Spaander MCW, Siersema PD. A novel device for intracolonoscopy cleansing of inadequately prepared colonoscopy patients: a feasibility study. Endoscopy. 2019;51(1):85-92.



Part I: VOC analysis for CRC screening

Test Results Are Inconclusive
Jenna Rindo
JAMA. Published online December 18, 2024.

Test results are available—but first you must
remember your login name and valid
password—to view your recent results
you must triple authenticate with cornea scan
fingerprint and second device.

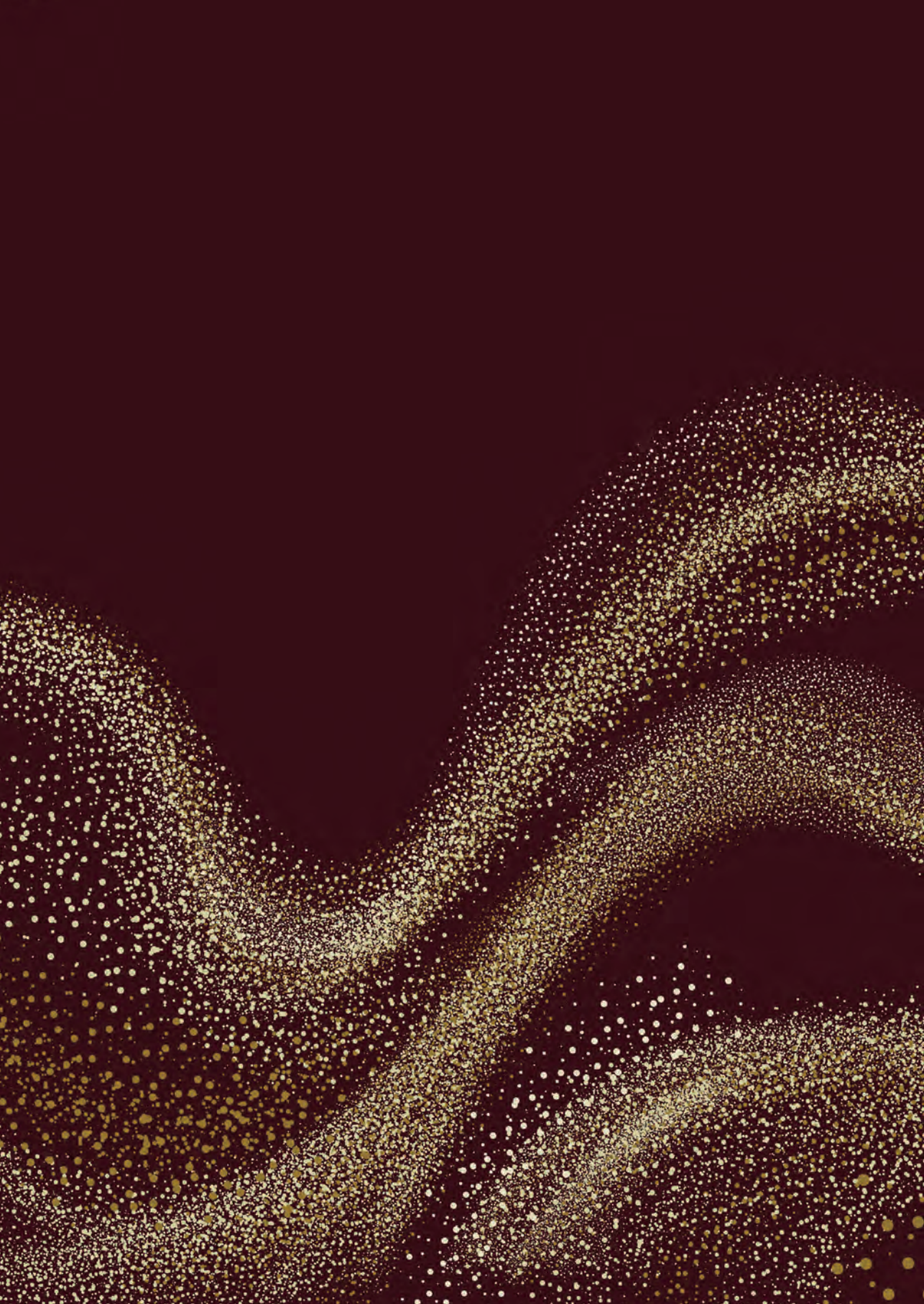
Test results reveal private organs and generational
trauma—IV contrast creates hot spots and viral
neon spirographs from your childhood fixation of
the multicolored pens and unlimited scrap paper.

Test results are official—only when read
by the radiologist. Ultrasound techs are not permitted
to inhale deeply, whistle, or show facial grimaces
as they move their wand in a widening oval.

Test results may vary based on biorhythms, the pollution
from airborne political positions, and the hidden
privilege from conflicting sources within the dark
office of the senior radiologist.

Test results may cause confusion—it's understandable
to request a second opinion—but keep in mind 3D images
seldom outline falsehoods or grant fair reparations.

Test results may grant you a kinesthetic experience.
You may find beauty and abstract art in the digital
breast tomosynthesis—slices from different angles
are both dense and difficult—how bittersweet—
how cognitively dissonant—how nuanced are
the depth and texture of each opacity.



Chapter 2

Breath testing for colorectal cancer detection in patients with a positive fecal immunochemical test – a multicenter prospective cross-sectional study with external validation

Milou L.M. van Riswijk, MD*, Kelly E. van Keulen, MD*, Adriaan C.I.T.L Tan, Ruud W.M. Schrauwen, Wouter H. de Vos tot Nederveen Cappel, Peter D. Siersema, MD, PhD, on behalf of the eNose CRC study group.

* Authors contributed equally

eNose CRC study group

Milou L.M. van Riswijk, MD*, Kelly E. van Keulen, MD*, Adriaan C.I.T.L Tan, Ruud W. Schrauwen, Wouter H. de Vos tot Nederveen Cappel, Peter D. Siersema, MD, PhD, J.M. Jansen, Marc A.M.T. Verhagen, Manon C.W. Spaander, M. Wiegerinck, Monica E.W. Derks, Hui Jun Huang, Renee Visserman, Demi Tenschert, Jip Klerks, Julia Hanevelt, Kim van der Velden, Meike Dooijes, Geanne Welbergen, J.W.A Straathof, MD, PhD

Aliment Pharmacol Ther. 2025 Jul;62(2):208-213. doi: 10.1111/apt.70207.

Abstract

Background: The accuracy of fecal immunochemical testing (FIT) for colorectal cancer (CRC) screening is suboptimal, leading to missed cancers and unnecessary colonoscopies. Although various studies have shown high diagnostic accuracy of volatile organic compound (VOC) analysis as an alternative biomarker for detecting CRC, large-scale validation studies are lacking, leaving the true clinical performance unclear. We aimed to evaluate the diagnostic accuracy and to validate a breath test for CRC using an electronic nose (e-nose) in FIT-positive patients.

Methods: This multicenter cross-sectional study was conducted in 11 Dutch hospitals and consisted of two sub-studies aiming: (1) to acquire sufficient breath samples to build a robust CRC prediction model, and (2) to perform external validation of this model and assess reproducibility. We included FIT-positive patients aged 55-75 years referred for screening colonoscopy and built a cross-validated CRC model using machine learning algorithms. Main outcomes were accuracy in terms of area under the receiver operating characteristics curve (AUC) of the external validation set, sensitivity (Sn), and specificity (Sp). Secondary outcomes included reproducibility and patient experience.

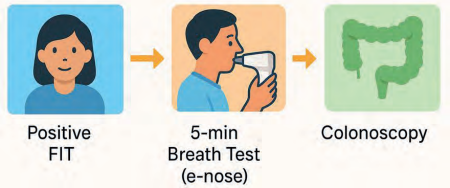
Results: Breath testing was performed in 3,959 patients (40.1% female, median age 64.25 years), 490 of them were withdrawn due to incomplete colonoscopy (n=170), failed/suboptimal breath test (n=255), withdrawal criteria (n=50) or lost to follow-up (n=15)). CRC was diagnosed in 175 (5.0%) patients 1,183 (34.1%) had advanced adenoma, 1,105 (31.9%) non-advanced adenoma, 137 (3.9%) hyperplastic polyps, and 869 (25.1%) a normal colonoscopy. The test characteristics of the breath test were suboptimal, with an AUC of 0.542 (95% CI 0.495-0.589), Sn of 39.5% and Sp of 68.3%. The intraclass correlation coefficient was 0.22, indicating poor reproducibility. Willingness-to-repeat the breath test was 95.3%.

Conclusions: Exhaled breath VOC testing by electronic nose is currently not ready to detect CRC or its precursor stages. The external validation results demonstrate the need for further development, standardization, and reproducibility assessment of e-nose technology before it can replace or supplement current CRC screening methods.

ClinicalTrials.gov Identifier: Clinicaltrials.gov Identifier NCT03346005 and NCT04357158

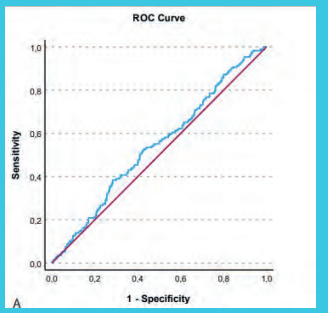
E-nose breath testing for CRC screening in FIT+ patients

Study design & population



- **Multicenter study with external validation of E-nose breath testing**
- **N=3,469 participants with FIT+**
- CRC prevalence 5.0%
- No pathology in colonoscopy 25.1%

Low diagnostic accuracy



AUC 0.542; sensitivity 39.5%; specificity 68.3%
low reproducibility (ICC 0.22)

Discussion

E-nose breath-based VOC analysis did not adequately differentiate CRC from non-CRC patients

Future research should focus on:
- refining VOC detection
- optimizing sensor technology
- conducting larger validation studies before clinical implementation is considered.

Van Riswijk, et al. *Aliment. Pharmacol. Ther.*

AP&T

Introduction

Colorectal cancer (CRC) screening with fecal immunochemical tests (FIT), followed by colonoscopy in case of a positive result, has been implemented as a screening program in many countries (1). However, the accuracy of FIT is suboptimal, with a sensitivity of 76-89% and specificity of 94% (2), leading to 11-24% missed cancers on the one hand and unnecessary colonoscopies on the other (2-4). Additionally, FIT primarily detects cancer by identifying bleeding from lesions and has limited accuracy for premalignant polyps that do not bleed, with sensitivity and specificity for advanced adenomas and serrated lesions at 33% and 93%, and an even lower sensitivity of serrated lesions (5.1-18.4%) (2, 5). This suboptimal accuracy could lead to diagnostic and treatment delays.

A better selection procedure for screening colonoscopies is warranted. Recent studies have suggested that volatile organic compounds (VOCs) may serve as a non-invasive biomarker for CRC screening (6-8). VOCs are gaseous compounds that result from metabolic processes and can be found in exhaled breath, blood, urine, or stool samples, amongst others (9-12). Metabolic processes and microbiome changes due to carcinogenesis and this can subsequently lead to a different VOC profile in patients (13). Meta-analyses of the predictive performance of VOCs using various sampling techniques show a promising performance (8, 14, 15). In a recent systematic review, Wang et al. performed a meta-analysis including 32 diagnostic test studies of VOCs in diagnosing CRC (14). This resulted in a pooled area under the receiver operating characteristics curve (AUC) of 0.93. However, only early-stage studies with a modest sample sizes and without external validation were included.

Since large-scale validation studies are not available, the true clinical performance of VOCs as diagnostic biomarkers for CRC is unclear. Validation performed in an external cohort is indispensable because of confounding VOCs or environmental influences and the risk of overfitting and spurious correlations (16, 17). Despite its importance (16-19), only few studies report external validation, all resulting in a lower diagnostic performance (16).

We previously investigated VOC analysis in exhaled breath to predict CRC for screening purposes, using a handheld electronic nose (e-nose) device (7). We included 447 breath tests over five sites, including 77 CRC patients, and established an AUC of 0.84 to predict CRC from a breath sample (7). Given the importance of external validation, the aim of this study is to (1) increase the robustness, and (2) to perform a multicenter validation of our previously developed model.

Methods

Study design

We performed a prospective multicenter cohort study in 11 hospitals in the Netherlands consisting of two sub-studies. The first sub-study was designed to acquire breath samples from patients with and without CRC or premalignant polyps in a large cohort of patients from several hospitals to further improve the robustness of our prediction model for CRC. The second sub-study was designed to validate this prediction model and to assess reproducibility of the breath test. In both studies, patients aged 55-75 years who underwent screening colonoscopy after a positive fecal immunochemical test (FIT) were invited for study participation. Patients who had a history of malignancy, except for non-melanoma skin cancer, history of (partial) bowel resection, active inflammatory bowel disease or polyposis syndromes, and patients who were not able to provide informed consent, were excluded. Patients with missing data of the breath test or the reference test colonoscopy were withdrawn from the final analysis.

Both studies were performed in compliance with the declaration of Helsinki and registered at clinicaltrials.gov (NCT03346005 and NCT04357158). All participants provided informed consent. The studies were approved by the Medical Ethical Committee of Radboudumc (CMO Arnhem-Nijmegen), and thereafter by all participating sites. All data was independently checked by at least two research team members. There was no patient or public involvement during the design, conduct, or reporting of the study. The study protocol, individual patient data, and analytical codes can be provided upon reasonable request.

Study procedures (breath test)

The studies were performed parallel to the Dutch bowel cancer screening program for individuals aged 55-75 years who performed a FIT test. If positive, patients received an invitation for colonoscopy and were consecutively invited for study participation. All study personnel were trained in the study procedures and in performing the breath test. After obtaining baseline characteristics, patients were invited for a 5-minute breath test. Only successful breath tests were included. Causes of failed breath tests and patient experience were documented.

The screening colonoscopies were scheduled within three weeks after the breath test, conform standard of care. Bowel preparation was performed according to each hospital's protocol. Since type of bowel preparation may have impact on the VOC profile of patients, this was documented (20). Patients in the second

sub-study performed a second breath test just prior to the colonoscopy to assess reproducibility.

Study device

All breath tests were performed with Aeonose e-nose devices (the eNose company, Zutphen, the Netherlands), a handheld electronic nose device containing three hotplate metal-oxide sensors (AS-MLV sensors; Applied Sensors GmbH) (21). VOC interaction with these sensors causes redox reactions resulting in slight current changes. The sensors are heated in a sinusoidal pattern, resulting in temperature-specific unique patterns.

The patient breathes through the e-nose via a disposable mouthpiece with a high-efficiency particulate arrestance (HEPA) filter to protect the device against contamination by, e.g., bacteria and viruses, and with a nose clip to prevent inspiration of unfiltered environmental VOCs. During the first two minutes, exhaled breath is lead through the HEPA filters and one-way valves to rinse environmental VOCs from the patients' breath and remove dead air space. In minutes 3-5, exhaled breath is led directly to the sensors for analysis and to a Tenax tube. The latter is used to detect VOCs in low concentrations after breath collection is finished. The following 10 minutes are used to regenerate the sensors before a new breath test can be performed. Breath print data is then transferred to a cloud-based storage platform with a mobile device. Failed breath tests due to incorrect breathing, resulting in environmental VOC pollution of the breath signal, or sensor or connectivity problems were excluded from the final analysis.

Quality control is ensured by the ISO13485 certified manufacturer, who performs yearly calibration cycles with each device. A total of 21 devices were used throughout the study.

Reference test

Breath test results were compared to the colonoscopy outcome and corresponding pathology results. All participating endoscopists and pathologists were certified to perform screening colonoscopies in the Dutch bowel cancer screening program (22). The colonoscopy procedure was performed according to standard of care. Bowel preparation quality was assessed with the Boston Bowel preparation scale (BBPS), and cecal intubation rate (CIR) was recorded. Patients with an incomplete colonoscopy were excluded from the final analysis.

Patients were categorized based on the most advanced lesion into the following categories: Control (no lesions), hyperplastic polyp (HP), non-advanced adenoma (NA), advanced adenoma (AA) (adenomas with one of the following: size ≥ 10mm, >25% villous aspect, or HGD), and CRC. This was based on the ESGE guideline on post-polypectomy colonoscopy surveillance 2013 (23), which was the guideline used in clinical practice during the time the study was performed.

Blinding

Patients, endoscopists, pathologists, and study personnel conducting the breath test were blinded for the breath test outcomes. Two-thirds of the data on final classification was unblinded for model development. The other third remained blinded to ensure adequate validation.

Endpoints

Our main outcome was the accuracy of the breath test to predict CRC in terms of AUC, sensitivity (Sn), specificity (Sp), and positive and negative predictive value (PPV and NPV, respectively). Our secondary endpoints included the accuracy of the breath test to predict AA, reproducibility of the breath test in terms of a Cohens' Kappa coefficient, patient acceptance rating including discomfort on a visual analog scale (VAS), and willingness-to-repeat the breath test.

VOC analysis

Several steps were required from an exhaled breath sample to final prediction of CRC presence or absence. First, the raw sensor data was augmented and then preprocessed to balance out differences between different e-nose devices. Preprocessing included normalization to ensure that the features are on a comparable scale. Depending on the machine learning classifier, the data was also compressed using Singular Value Decomposition (SVD). Preprocessing was done similarly in all groups. The preprocessed data was then used in a machine-learning classifier.

We used previous study data (7) as a base to build a cross-validated model separately for CRC and AA. Various classifiers were evaluated in this study, including a convolutional neural network (CNN), multiplayer perceptron (MLP), and several tree-based classifiers. The data was divided into a training set (80%) and a test set (20%) to validate the training performance (Figure 1). This validation process was used during algorithm development and is separate from the blind predictions from the second sub-study. We took a balanced number of samples from all devices in this study. To avoid class imbalance, the test set was balanced to ensure fair comparison and to evaluate the classifier's ability to distinguish effectively

between binary classes. Hyperparameter tuning for these models was performed by optimizing their performance on the training set. The test set was used for internal validation of the trained algorithm. Performances were considered across different hospitals and devices. The best model was chosen based on test set performances and was not modified after final development. The output of the model consisted of a binary classification. The positivity threshold was set on the test set aiming for high sensitivity and acceptable specificity. This threshold was used as a hard cut-off to avoid indeterminate breath test results. Detailed descriptions of the model development are published elsewhere (7, 24).

The prediction model was externally validated using data from the second sub-study, including devices not used in prior studies (Figure 1).

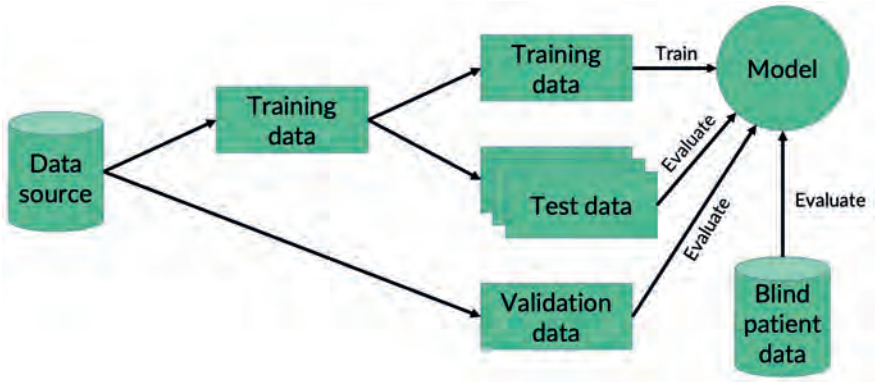


Figure 1. Visual representation of data use and distribution to design the prediction model.

Statistical analysis

Baseline characteristics were summarized as percentages for categorical data, means and standard deviations (SD) for normally distributed data, and medians/interquartile ranges (IQR) for non-normally distributed data. Where appropriate, we used an unpaired t-test, Mann-Whitney-U test, or chi-square test. A two-sided p-value of ≤0.05 was considered statistically significant. No data was omitted from the analyses.

For our primary outcome, the continuous predictive value of the breath test between 0 and 1 was compared to the presence or absence of CRC to assess the performance of the prediction using an AUC, with an AUC of 0.5 meaning random prediction (i.e., flipping a coin), and an AUC of 1.0 perfect prediction. The same

method was used for the accuracy of the prediction of AAs. Additionally, sensitivity, specificity, and positive- and negative predictive values were assessed using the same cut-off of the breath test value.

To assess reproducibility, the predictions of the first and second breath tests in de same patient were compared to the presence of CRC. A Cohens' Kappa statistic was derived to quantify inter-observer agreement.

For VOC analysis, we used Aethena (proprietary software, version 3.0, the eNose Company, Zutphen, the Netherlands), Python 3.10 (Python Software Foundation, Delaware, United States), and Tensorflow 2.13 (Google Brain, Alphabet Inc., California, United States). For statistical analysis, SPSS statistics version 29 (IBM. Corp., Armonk, New York, USA) was used.

Sample size calculation

Sample size calculations assumed a minimum of hundred patients per study group for the model development of the cohort and a 5% CRC prevalence. The sample size for external validation of the model was based on a two-sided alpha of 0.05, a maximum marginal error of 7.5%, an expected sensitivity of 90%, and a 10% drop-out rate based on our pilot study(7, 25).

Results

Baseline characteristics

Between January 2018 and July 2022, a total of 4,479 patients enrolled in the national Dutch CRC screening program were invited for study participation at eleven sites (Figure 2). Bowel preparation was adequate in 97.2% (BBPS ≥6), and CIR was in 98.5%, resulting in 3,469 patients for final analyses. A total of 975 patients were invited for a second breath test for reproducibility assessment prior to the colonoscopy, of which 709 were successfully performed.

Among the included patients, 869 (25.1%) were in the control group, 137 (3.9%) had hyperplastic polyps, 1,105 (31.9%) NAs, 1,183 (34.1%) AAs, and 175 (5.0%) were diagnosed with CRC (Table 1). Overall, patient characteristics were equally divided between groups and between the developmental cohort and validation cohort; however. However, CRC prevalence was lower in the validation cohort with 3.9% compared to 7.0%, respectively (Table 1, Table S1.). In total, 1,484 (40.1%) were female, with a mean age of 64.3 years (SD 6.4). Patients with CRC were significantly

more frequent smokers (p=0.01), used alcohol (p=0.03), had a family history of CRC (p=<0.001), and a previous history of polyps (p=0.01). In the control group, additional findings of colonoscopy, such as hemorrhoids and angiodysplasia, were significantly more frequent. Overall, the adenoma- polyp- and serrated polyp detection rates were 71.0%, 75.2%, and 18.9%, respectively.

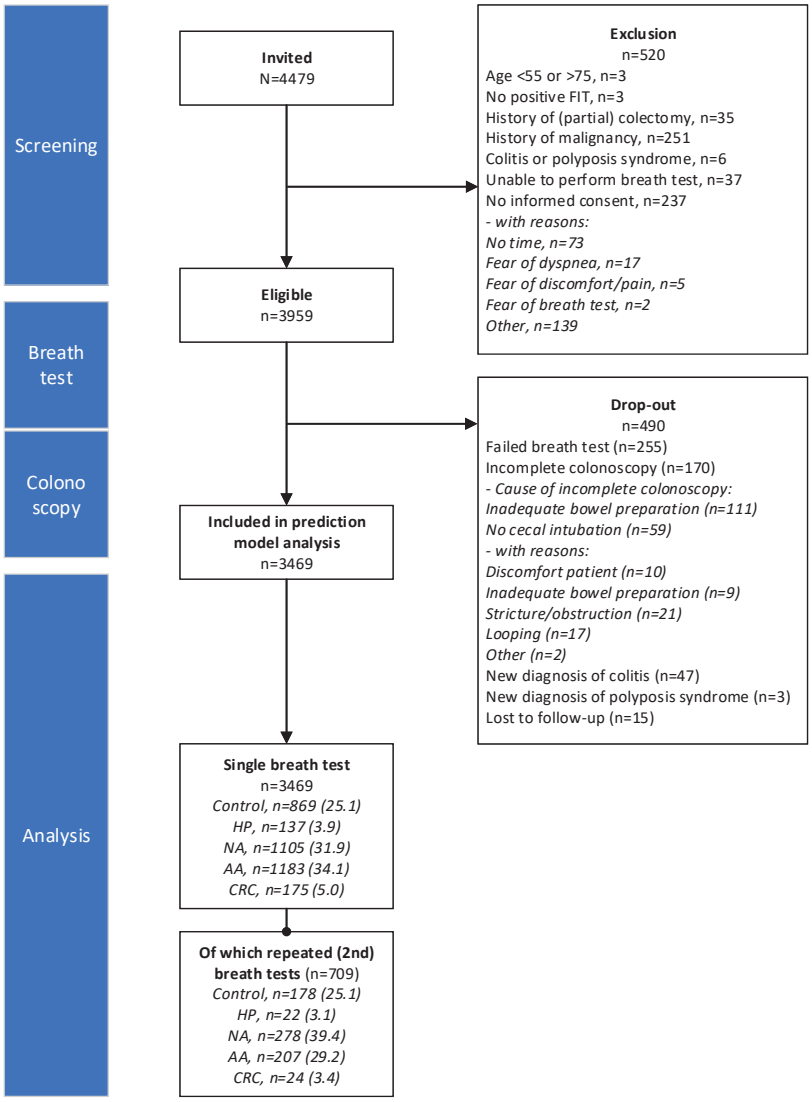


Figure 2. Patient flowchart. After exclusion and dropout, 3,469 breath samples were included in the final analysis. A subset of patients was invited to perform an additional, second breath test to assess reproducibility. The prevalence of (advanced) polyps and colorectal carcinoma within the total cohort of breath samples and the cohort of repeated breath samples is shown. *AA, Advanced adenoma; CRC, colorectal carcinoma; FIT, fecal immunochemical test; HP, hyperplastic polyp; NA, non-advanced adenoma.*

Table 1. Baseline characteristics.

	Control	HP	NA	AA	CRC	P between CRC and non-CRC patients
N (%)	869 (25.1)	137 (3.9)	1105 (31.9)	1183 (34.1)	175 (5.0)	N.A.
Female sex, n (%)	407 (46.8)	55 (40.1)	401 (36.3)	428 (36.2)	71 (40.6)	0.14
Age (years), median (IQR)	63 (58-69)	63 (57-69)	65 (59-71)	63 (58-71)	65 (59-71)	0.06
BMI, median (IQR)	26.5 (24.0-29.4)	27.0 (24.4 -30.1)	26.9 (24.2-29.7)	26.6 (24.4-29.4)	26.3 (24.1-28.9)	0.67
ASA, median (IQR)	2 (1-2)	2 (1-2)	2 (2-2)	2 (1-2)	2 (1-2)	0.30
Smoking						
Never, n (%)						**0.01**
Former, n (%)	439 (50.5)	39 (28.5)	411 (37.2)	465 (39.3)	78 (44.6)	
Current, n (%)	217 (25.0)	38 (27.7)	375 (33.9)	357 (30.2)	35 (20.0)	
Packyears for former/current	213 (24.5)	60 (43.8)	319 (28.9)	361 (30.5)	62 (35.4)	
smokers, median (IQR)	15 (5.2-30)	20 (7.5-33)	17.5 (8.3-30)	18.8 (8-33)	20 (7.5-36)	0.12
Alcohol use						**0.03**
no, n (%)	308 (25.4)	40 (29.2)	331 (30.0)	364 (30.8)	47 (26.9)	
yes, n (%)	561 (64.6)	97 (70.8)	772 (70.0)	819 (69.2)	128 (73.1)	
median units per week (IQR)	5 (2-10)	6 (2-14)	5 (2-14)	6 (2-14)	5 (2-14)	0.28
Dietary restrictions, n (%)						0.85
Vegetarian	15 (1.7)	4 (2.9)	4 (0.4)	14 (1.2)	2 (1.1)	
Vegan	2 (0.2)	0	0	3 (0.3)	0	
Gluten free	2 (0.2)	0	4 (0.4)	2 (0.2)	1 (0.6)	
Other	8 (0.9)	3 (2.2)	9 (0.8)	9 (0.8)	2 (1.1)	
Time fasted before breath test, hours, median (IQR)	0 (0-1.5)	0 (0-1)	1 (0-2)	0 (0-1)	0 (0-1)	0.19

Table 1. Continued

	Control	HP	NA	AA	CRC	P between CRC and non-CRC patients
Past medical history, n (%)						
Abdominal/pelvic surgery	267 (30.7)	49 (35.8)	337 (30.5)	345 (29.2)	57 (32.6)	0.63
Family Hx of CRC	109 (12.5)	23 (16.8)	147 (13.3)	188 (15.9)	39 (22.3)	**<0.001**
Cardiovascular disease	319 (36.7)	49 (35.8)	416 (37.6)	425 (35.9)	58 (33.1)	0.79
Pulmonary disease	92 (10.6)	14 (10.2)	140 (12.7)	146 (12.3)	26 (14.9)	0.41
Renal disease (eGFR <30ml/min)	5 (0.6)	2 (1.5)	18 (1.6)	9 (0.8)	4 (2.3)	**0.03**
Gastro-intestinal disease	124 (14.3)	22 (16.1)	134 (12.1)	96 (8.1)	12 (6.9)	**0.01**
Liver disease	6 (0.7)	0	11 (1.0)	13 (1.1)	1 (0.6)	0.86
Endocrine disease	81 (9.3)	12 (8.8)	108 (9.8)	117 (9.9)	19 (10.9)	0.53
Neurological disease	29 (3.3)	6 (4.4)	53 (4.8)	66 (5.6)	4 (2.3)	0.47
GI history specified, n (%)						
Chronic constipation	35 (4.0)	6 (4.4)	41 (3.7)	26 (2.2)	5 (2.9)	0.46
Inflammatory bowel disease	2 (0.2)	0	2 (0.2)	1 (0.1)	0	0.53
Diverticulitis	15 (1.7)	7 (5.1)	14 (1.3)	10 (0.8)	1 (0.6)	0.26
Gastroenteritis	3 (0.3)	1 (0.7)	2 (0.2)	2 (0.2)	0	0.44
History of polyps	67 (7.7)	10 (7.3)	64 (5.8)	41 (3.5)	4 (2.3)	**0.01**
Antibiotic use in past 2 months, n(%)	19 (2.2)	6 (4.4)	26 (2.4)	17 (1.4)	3 (1.7)	0.69

Table 1. Continued

	Control	HP	NA	AA	CRC	P between CRC and non-CRC patients
CRC type, n (%)						N.A.
Adenocarcinoma					171 (97.7)	
Other					4 (2.3)	
CRC location, n (%)						
Cecum					21 (11.9)	
Ascending					23 (13.1)	
Transverse					18 (10.2)	
Descending/Sigmoid					64 (34.4)	
Rectum					50 (28.4)	
TNM stage, n (%)						
T1					55 (32.4)	
T2					49 (28.8)	
T3					54 (31.8)	
T4					9 (5.3)	
Tx					3 (1.8)	
N0					118 (69)	
N1					42 (24.6)	
N2					9 (5.3)	
Nx					2 (1.2)	
M0					151 (87.8)	
M1					5 (2.9)	
Mx					16 (9.3)	
Metachronous CRC, n (%)					4 (2.3)	
AA category*, n (%)						N.A.
≥10mm				1054 (89.2)		
HGD				13 (1.1)		
Villous aspect				114 (9.7)		

Table 1. Continued

	Control	HP	NA	AA	CRC	P between CRC and non-CRC patients
Other findings colonoscopy**, n (%)						**<0.001**
None	300 (34.5)	51 (37.2)	396 (35.8)	505 (42.7)	96 (54.9)	
Hemorrhoids	164 (18.9)	24 (17.5)	157 (14.2)	129 (10.9)	15 (8.6)	
Angiodysplasia	28 (3.2)	1 (0.7)	15 (1.4)	14 (1.2)	0	
Diverticulosis/-itis	362 (41.7)	58 (42.3)	530 (48.0)	527 (44.5)	64 (36.6)	
Other	15 (1.7)	3 (2.2	7 (0.6)	8 (0.7)	0	

* Polyps were only classified as villous aspect or high-grade dysplasia when smaller than 10mm. All polyps ≥10mm regardless of other advanced aspects were classified as AA-≥*10mm*. ** Multiple answers were allowed. *Control; normal colonoscopy without significant pathology, AA; Advanced adenoma (i.e., >10mm and/or >25% villous histology and/or high-grade dysplasia), ASA; American Society of Anesthesiologists, CRC; Colorectal carcinoma, IQR; interquartile range, NA; non-advanced adenoma, Tx; T stadium not specified, Nx; N stadium not specified, Mx; M stadium not specified.*

Diagnostic performance CRC vs. controls and non-CRC

After training and testing, the average blind predictions combining the CNN, MLP, and Random Forest Classifier (RFC) models yielded an AUC of 0.542 (95% CI 0.495-0.589) when testing for CRC in a dataset with only CRC and control individuals, which is the ideal situation for the prediction model (Table 2, table S2, Figure 3). When adding other non-CRC individuals (i.e., HP, NA, AA), which is the target population for e-nose use, the AUC was 0.500 (95% CI 0.451-0.550) (Table 2). This indicates that the e-nose was unable to predict CRC in this population. At first glance, the NPV may look promising (range 85.1%-90.7%). However, given the 5% prevalence of CRC in our study population and vice versa 95% non-CRC prevalence, the breath test does not aid sufficiently in discriminating healthy from diseased individuals.

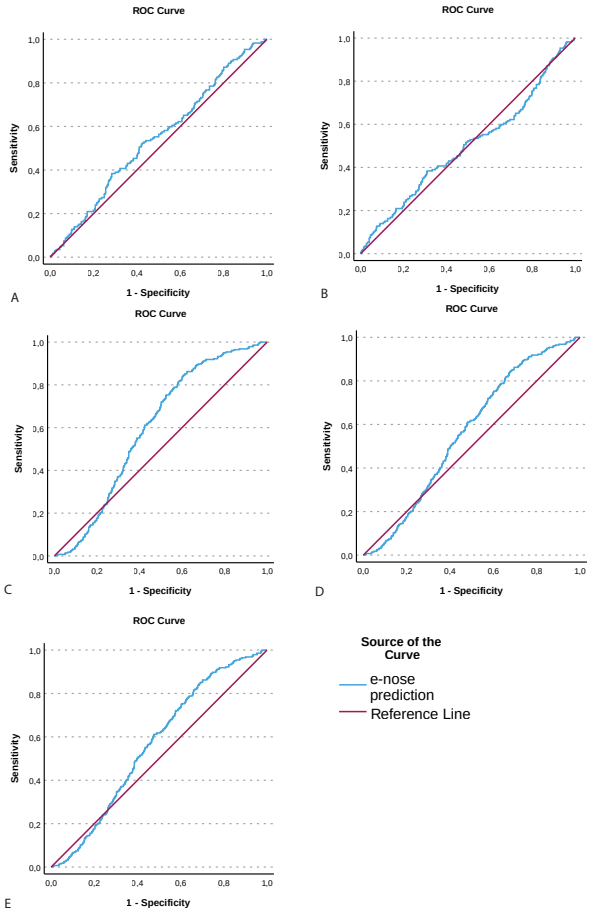


Figure 3. Receiver operating characteristics (ROC) curve of e-nose blind predictions in a dataset containing only CRC and control patients. A – CRC vs control. B – CRC vs non CRC. C – AA vs control. D – AA vs non AA. E – AA vs non AA excluding CRC patients.

Table 2. Diagnostic performance of breath test in various data subsets.

	Cases (n)	Controls (n)	AUC	Sn (%)	Sp (%)	PPV (%)	NPV (%)
PREDICTION OF COLORECTAL CANCER							
CRC-control							
Training model	118	359					
Test	86	86					
Validation	204	445					
Blind predictions	40	236	0.542 (0.495-0.589)	39.5	68.3	19.8	85.1
CRC-non CRC							
Blind predictions	40	930	0.500 (0.451-0.550)	39.5	65.0	11.1	90.7
PREDICTION OF ADVANCED ADENOMA							
AA-control							
Training model	636	440					
Test	98	98					
Validation	734	538					
Blind predictions	287	236					
AA specific model			0.517 (0.492-0.543)	99.2	0.7	57.6	40.0
CRC specific model			0.599 (0.565-0.634)	38.5	68.3	28.4	77.3
AA-nonAA							
Blind predictions	287	723					
AA specific model			0.509 (0.487-0.532)	99.3	1.1	45.0	66.7
CRC specific model			0.563 (0.530-0.596)	38.5	65.1	17.8	84.4
AA-nonAA (ex CRC)							
Blind predictions	287	683					
AA specific model			0.527 (0.504-0.550)	99.2	0.7	48.1	50.0
CRC specific model			0.565 (0.531-0.599)	38.5	65.8	20.0	82.8

AA; advanced adenoma, AUC; area under the receiver operating characteristics curve, CRC; colorectal carcinoma, NPV; negative predictive value, PPV; positive predictive value, pre-test; pre-test likelihood, post-test; post-test likelihood, Sn; sensitivity, Sp; specificity.

Performance AA vs. controls and non-AA

The prediction model trained for CRC-detection was not able to discriminate AA from controls or non-advanced neoplasia, with an AUC of 0.606 (95% CI 0.571-0.641) and 0.570 (95% CI 0.536-0.604), respectively (Table 2, Figure 2 CDE). A predictive model specifically trained for AA-detection showed also no difference (Table 2, Table S2, Figure S2 ABC).

Reproducibility

AUCs of the repeated breath tests were comparable, with an overlapping 95% confidence interval of 0.419-0.708, however were difficult to assess given the poor performance. Therefore, an intraclass correlation coefficient was calculated of 0.22 (95% CI 0.15-0.29), indicating poor correlation in repeated breath tests.

Patient acceptance rate

Patients rated the breath test as non-invasive, with a median NRS of 1 out of 10. Willingness-to-repeat was high, with an overall rate of 95.3%. Patients with a failed breath test reported significantly higher discomfort (median VAS 5 vs 1) and substantially lower willingness-to-repeat (62.6% vs 97.4%). Failed tests were significantly more frequent in females, but other relevant baseline characteristics were not different between groups (Table 3).

Failed breath tests occurred in 255 patients (6.8%). Technical issues were the reason for failure in more than half of the failed breath tests (54.9%) and mainly consisted of connectivity issues, in which the breath data was not adequately transferred from the e-nose to the cloud data storage. Other problems were sensor issues (no signal or 'flatliner') and the disposable mouthpiece or filter getting loose, disrupting the airflow and potentially introducing contaminating environmental VOCs. Non-technical causes for failed breath tests were mostly dyspnea and dyspnea-related anxiety (27.8%) (Table 3).

Table 3. Patient experience of the breath tests and reasons of failed breath tests.

	Successful breath testN=3469 (93.2%)	**Failed breath testN=255 (6.8%)**	**P-value**
Baseline characteristics			
Age, median (IQR)	64.0 (59.0-70.0)	63.0 (59.0-68.3)	0.19
Female sex, n (%)	1452 (39.2)	132 (53.7)	**<0.001**
ASA, median (IQR)	2 (1-2)	2 (1-2)	0.60
BMI, median (IQR)	26.6 (24.2-29.6)	26.5 (24.4-29.3)	0.57
Current or previous smoking, n (%)	2187 (59.0)	145 (58.9)	0.23
Pulmonary disease, n (%)	458 (12.4)	37 (14.9)	0.20
Patient experience			
Breath test 1 VAS, median (IQR)	1 (1-2.5)	5 (1-8)	**<0.001**
1 – Willingness-to-repeat, n (%)	3605 (97.4)	152 (62.6)	**<0.001**
2 – VAS, median (IQR)	1 (0-2)	2 (0-3.25)	0.26
2 – Willingness-to-repeat, n (%)	746 (98.3)	21 (80.8)	**<0.001**
Reason of failed test, n (%)			N.A.
Dyspnea		71 (27.8)	
Coughing		5 (2.0)	
Nausea		6 (2.4)	
e-nose/technical		140 (54.9)	
Not adequately performed by research personnel		21 (8.2)	
Other		12 (4.7)	

ASA, American society of anesthesiologists' classification; BMI, body mass index; IQR, interquartile range; VAS, experienced discomfort on a visual analogue scale.

Discussion

External validation is an essential part for the clinical application of breath-based VOC analysis to predict CRC. We performed a multicenter study with external validation to test whether the e-nose could predict CRC in the exhaled breath of nearly 3,500 FIT-positive patients. While patients regarded the breath test as non-invasive (median NRS 1/10) with a high willingness-to-repeat (95.3%), the e-nose was not able to predict CRC (AUC 0.542) or advanced adenoma (AUC 0.517) from a breath sample. Reproducibility could therefore not be adequately assessed.

The performance of our e-nose is significantly lower than previous studies have reported. This could be attributed to various factors. One limitation of previously reported studies is smaller sample sizes in those studies; however, larger studies have also been conducted. For instance, the COBRA1 study by Woodfield et al. achieved an AUC of 0.87 in a sample of 1,463 patients using the ReCIVA breath sampling device based on gas-chromatography-mass-spectrometry (GCMS), in contrast to our sensor-based approach (26). The discrepancy in performance compared to our study may be due to differences in patient populations and higher prevalence of symptomatic patients that hypothetically lead to higher concentrations of discriminative VOCs in the COBRA1 study or the different VOC sampling method. Another explanation for different results could be confounding factors, such as environmental VOCs, that could introduce noise and variability of the analytical VOC signal, thereby obscuring the detection of disease-specific VOC patterns. This was demonstrated by McFarlane et al., who reported that CRC detection from urinary VOC profiles was unreliable until larger control groups were included (27). In comparison, our study's larger sample size should have provided robust data; however, the presence of CRC cases across multiple devices may have introduced variability, which impacted performance. Lastly, a possible explanation for our discrepant results to other studies is the lack of external validation in other studies. Prediction models tend to perform worse in new datasets, even if internal validation is performed (28). Scheepers et al. conducted a meta-analysis of 52 e-nose studies for various cancers, including 11 Aeonose studies, which was the e-nose device used in the present study (15). Although the pooled sensitivity and specificity were high (90% and 87%, respectively), none of the included studies performed true external validation. Moreover, significant funnel plot asymmetry was observed, suggesting the potential presence of publication bias. These studies collectively show the potential of VOC analysis for CRC detection but also highlight the necessity for larger, multicenter studies to validate these non-invasive diagnostic tools.

There may be several reasons for the current e-nose to perform significantly worse than in our previously published pilot study (7). Firstly, the CRC prevalence of 5.0% compared to the 15.6% prevalence of our pilot study was low. It is well-known that a higher prevalence of the disease of interest often yields higher sensitivities, which could partially explain the discrepancy in our results. For instance, a study evaluating the SEPT9 methylation in plasma for CRC detection observed lower sensitivity in low-prevalence populations compared to a high-prevalence setting (29). In addition, subtle inter-device variations require more data to compensate for, i.e., external validation has been reported to require another 100-200 positive cases (30-32). In the present study, 179 positive (CRC) cases were included, divided over 21 e-nose devices, resulting in a likely too low signal-to noise ratio, which might obscure VOC patterns characteristic of CRC and degrades the quality of the data for prediction models. Moreover, sensor drift may reduce the compatibility of different device signals over time (33). Determining the discriminative ability of an e-nose may therefore require a different study design, which reduces the number of devices used in the study. Hypothetically, prediction model training of e-nose sensor data could be optimized by gaining more insight into the specific sensor data and the VOCs they relate to, so fewer variables will be necessary, resulting in a smaller sample size.

Another explanation for the poor performance in our study might be the breath-based analysis instead of stool-based VOC analysis. This e-nose device has also been tested for lung cancer screening (34). In a multicenter study, Kort et al. included 376 patients with a 43% lung cancer prevalence. After a stepwise validation process, this resulted in an AUC of 0.87. It could be stipulated that in breath analysis, the disease-specific VOC concentration in lung cancer is higher than in CRC as VOCs in CRC have a long route to travel, including absorption into the bloodstream and a first-pass effect through the liver before excretion via the lungs. Consequently, the concentration of disease-specific VOCs in exhaled breath is likely lower in CRC compared to lung cancer. This necessitates more breath samples to develop a predictive model or a higher sensitivity of the device sensors (35). Stool-based VOC analysis for CRC may provide higher concentrations of disease-specific VOCs due to closer proximity to the tumor site, potentially offering better sensitivity and specificity (36).

In contrast to the CRC prevalence, the prevalence of AAs (34.1%) should hypothetically include sufficient data for a discriminative predictive model. However, the accuracy of the e-nose for AAs was also found to be insufficient. It may be that preneoplastic lesions contain less discriminative VOCs. VOC analysis

is based on the metabolic changes excreted through exhaled breath. While some VOCs may originate from the Warburg effect (anaerobic metabolism in cancer cells due to ineffective tumor circulation), it is proposed that several CRC-associated VOCs result from an interplay with gut microbiota changes (37). Microbiome changes are associated with CRC presence, whereby gut dysbiosis may create a pro-oncogenic environment through chronic inflammation and immune response, metabolites, or direct DNA damage (38-40). Gut microbiome variations have shown promise in predicting CRC, with some studies reporting AUCs as high as 0.91 for early-stage CRC detection (41). Recent studies have highlighted specific VOCs linked to gut microbiome alterations in CRC patients, including changes in short-chain fatty acids, amino acids, sulfur compounds, and compounds related to the Warburg metabolism, such as alcohols and ketones (11, 26, 37, 42)**.** It may well be that AAs contain less discriminative VOCs than CRC, as less advanced metabolic and/or microbiome alterations are caused (43). This could explain why AAs could not be predicted despite a sufficiently high prevalence in our study. Integrating microbiota, proteome, and amino acid profiles to VOC analysis may therefore potentially enhance the discriminatory power for CRC and adenomas and needs further study (44).

Although this study did not yield positive results, a future clinical benefit of e-nose use for CRC screening is a potential increase in adherence. The efficacy of colonoscopy screening programs for CRC is strongly dependent on participation (45, 46). Given the high acceptability of patients for a breath test in the present study, a breath test could offer a potential solution to improve participation and the overall effectiveness of CRC screening in the future.

Future initiatives for developing a breath test should consider confounding factors in exhaled breath and the effect of systemic VOC differences between study sites, since this leads to lower signal-to-noise ratios, thereby increasing the required sample size. For this, standardized chemical detection techniques such as GCMS may be necessary, as these are already calibrated and proven functional in current clinical practice. Unfortunately, this means that the attractive point-of-care aspect of the e-nose is not yet feasible. After identification of the CRC-related VOCs, as well as the variability in this profile including environmental influences and what these features look like as e-nose signals, e-nose devices can be tested and calibrated as well. Future research recommendations therefore include standardization of breath-based VOC analysis to advance this field (47, 48). For studies using machine learning techniques specifically, VOC breath test samples must be collected with a prevalence of 50%, including negative controls. Once a robust model is developed, it is essential to validate it in an external cohort before it can be used clinically.

This study has several strengths. First, it was tested in the future target population of the e-nose, i.e., FIT-positive individuals invited for colonoscopy-screening, ensuring future applicability and reducing risk of bias. As our study was performed parallel to the national screening program, which is monitored and audited, our reference standard was of high quality and reliability (49). The large sample size of nearly 3,500 patients reduced the risk of spurious correlations. Moreover, classification data was independently double-checked to avoid misclassification that could impact our results. Finally, external validation using different devices and hospitals adds to the robustness of our results.

There are also limitations. First, while we aimed for a generalizable study population, we excluded subjects with a history of malignancies or active IBD as we hypothesized that this could lead to confounding factors in developing our prediction model. We suggest including these subsets of patients in future studies to address applicability. Second, due to the unfavorable results, we could not perform a full analysis of potential confounders. For example, bowel preparation could have influenced the reproducibility results. Third, our study might not be 1 to 1 comparable to other e-nose devices due to different sensors or other VOC sampling media.

Conclusion

Breath-based VOC analysis by e-nose did not adequately discriminate CRC from non-CRC patients in this study including an external validation. Future e-nose studies should focus on assessing distinctive VOC profiles and their sensor pattern to further study the potential of VOC analysis for CRC detection. Importantly, more external validation studies are warranted before clinical application is feasible.

Acknowledgements

We thank all study sites, research personnel, and volunteers who made this large-scale study possible. We thank all study participants for their engagement.

Authors' contributions

- **Milou L.M. van Riswijk**: Protocol design, coordination, patient enrollment, data extraction, analysis and interpretation, visualization, drafting and revision of manuscript, project administration.
- **Kelly E. van Keulen**: Protocol design, coordination, patient enrollment, data extraction, analysis and interpretation, revision of manuscript, project administration.
- **Adriaan C.I.T.L Tan**: Patient enrollment, data extraction, critical revision of manuscript
- **Ruud W.M. Schrauwen:** Patient enrollment, data extraction, critical revision of manuscript
- **Wouter H. de Vos tot Nederveen Cappel:** Patient enrollment, data extraction, critical revision of manuscript
- **Peter D. Siersema**: Protocol design, discussion on interpretation of results, supervising study execution, critical revision of manuscript

Competing interests

This study is investigator-initiated and was funded by the Dutch Digestive Foundation (MLDS) and European Union (EFRO). The e-nose devices (Aeonose, the eNose company, Zutphen, the Netherlands) were provided by the manufacturer for this study without costs.

Author P.S. receives unrestricted grants from Pentax (Japan), FujiFilm (Japan), Norgine (UK), MicroTech (China) and Magentiq Eye (Israel) and is in the advisory board of Sanofi (The Netherlands). Medical Specialists Company Isala (MSB Isala) invested in the eNose Company, Zutphen, the Netherlands. Author W.d.V.t.N.C. is a member of the MSB but declared to renounce potential future profits of the eNose Company in person. Authors M.R., K.K., A.T., and R.S. declare no competing interests.

References

1. Saftoiu A, Hassan C, Areia M, Bhutani MS, Bisschops R, Bories E, et al. Role of gastrointestinal endoscopy in the screening of digestive tract cancers in Europe: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. Endoscopy. 2020;52(4):293-304.
2. Grobbee EJ, Wisse PH, Schreuders EH, van Roon A, van Dam L, Zauber AG, et al. Guaiac-based faecal occult blood tests versus faecal immunochemical tests for colorectal cancer screening in average-risk individuals. Cochrane Database Syst Rev. 2022;6(6):Cd009276.
3. MacDonald S, MacDonald L, Godwin J, Macdonald A, Thornton M. The diagnostic accuracy of the faecal immunohistochemical test in identifying significant bowel disease in a symptomatic population. Colorectal Dis. 2022;24(3):257-63.
4. Wieten E, Schreuders EH, Grobbee EJ, Nieboer D, Bramer WM, Lansdorp-Vogelaar I, et al. Incidence of faecal occult blood test interval cancers in population-based colorectal cancer screening: a systematic review and meta-analysis. Gut. 2019;68(5):873-81.
5. Chang LC, Shun CT, Hsu WF, Tu CH, Tsai PY, Lin BR, et al. Fecal Immunochemical Test Detects Sessile Serrated Adenomas and Polyps With a Low Level of Sensitivity. Clin Gastroenterol Hepatol. 2017;15(6):872-9.e1.
6. Woodfield G, Belluomo I, Laponogov I, Veselkov K, Cross AJ, Hanna GB, et al. Diagnostic performance of a non-invasive breath test for colorectal cancer: COBRA1 study. Gastroenterology. 2022.
7. van Keulen KE, Jansen ME, Schrauwen RWM, Kolkman JJ, Siersema PD. Volatile organic compounds in breath can serve as a non-invasive diagnostic biomarker for the detection of advanced adenomas and colorectal cancer. Aliment Pharmacol Ther. 2020;51(3):334-46.
8. Chandrapalan S, Bosch S, Cubiella J, Guardiola J, Kimani P, Mulder C, et al. Systematic review with meta-analysis: volatile organic compound analysis to improve faecal immunochemical testing in the detection of colorectal cancer. Aliment Pharmacol Ther. 2021;54(1):14-23.
9. Markar SR, Chin ST, Romano A, Wiggins T, Antonowicz S, Paraskeva P, et al. Breath Volatile Organic Compound Profiling of Colorectal Cancer Using Selected Ion Flow-tube Mass Spectrometry. Ann Surg. 2019;269(5):903-10.
10. Di Lena M, Porcelli F, Altomare DF. Volatile organic compounds as new biomarkers for colorectal cancer: a review. Colorectal Dis. 2016;18(7):654-63.
11. Bosch S, Berkhout DJ, Ben Larbi I, de Meij TG, de Boer NK. Fecal volatile organic compounds for early detection of colorectal cancer: where are we now? J Cancer Res Clin Oncol. 2019;145(1):223-34.
12. Tyagi H, Daulton E, Bannaga AS, Arasaradnam RP, Covington JA. Non-Invasive Detection and Staging of Colorectal Cancer Using a Portable Electronic Nose. Sensors (Basel). 2021;21(16).
13. de Boer NK, de Meij TG, Oort FA, Ben Larbi I, Mulder CJ, van Bodegraven AA, van der Schee MP. The scent of colorectal cancer: detection by volatile organic compound analysis. Clin Gastroenterol Hepatol. 2014;12(7):1085-9.
14. Wang Q, Fang Y, Tan S, Li Z, Zheng R, Ren Y, et al. Diagnostic performance of volatile organic compounds analysis and electronic noses for detecting colorectal cancer: a systematic review and meta-analysis. Front Oncol. 2024;14:1397259.
15. Scheepers M, Al-Difaie Z, Brandts L, Peeters A, van Grinsven B, Bouvy ND. Diagnostic Performance of Electronic Noses in Cancer Diagnoses Using Exhaled Breath: A Systematic Review and Meta-analysis. JAMA Netw Open. 2022;5(6):e2219372.
16. Leopold JH, Bos LD, Sterk PJ, Schultz MJ, Fens N, Horvath I, et al. Comparison of classification methods in breath analysis by electronic nose. J Breath Res. 2015;9(4):046002.

17. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med. 2000;19(4):453-73.
18. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Intern Med. 1999;130(6):515-24.
19. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Ann Intern Med. 2006;144(3):201-9.
20. Krishnamoorthy A, Chandrapalan S, Bosch S, Bannaga A, De Boer NKH, De Meij TGJ, et al. The Influence of Mechanical Bowel Preparation on Volatile Organic Compounds for the Detection of Gastrointestinal Disease-A Systematic Review. Sensors (Basel). 2023;23(3).
21. Bruins M, Gerritsen JW, van de Sande WWJ, van Belkum A, Bos A. Enabling a transferable calibration model for metal-oxide type electronic noses. Sensors and Actuators B-Chemical. 2013;188:1187-95.
22. Bronzwaer MES, Depla A, van Lelyveld N, Spanier BWM, Oosterhout YH, van Leerdam ME, et al. Quality assurance of colonoscopy within the Dutch national colorectal cancer screening program. Gastrointest Endosc. 2019;89(1):1-13.
23. Hassan C, Quintero E, Dumonceau JM, Regula J, Brandão C, Chaussade S, et al. Post-polypectomy colonoscopy surveillance: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. Endoscopy. 2013;45(10):842-51.
24. Kort S, Brusse-Keizer M, Gerritsen JW, van der Palen J. Data analysis of electronic nose technology in lung cancer: generating prediction models by means of Aethena. J Breath Res. 2017;11(2):026006.
25. Hajian-Tilaki K. Sample size estimation in diagnostic test studies of biomedical informatics. Journal of Biomedical Informatics. 2014;48:193-204.
26. Woodfield G, Belluomo I, Laponogov I, Veselkov K, Cross AJ, Hanna GB. Diagnostic Performance of a Noninvasive Breath Test for Colorectal Cancer: COBRA1 Study. Gastroenterology. 2022;163(5):1447-9.e8.
27. McFarlane M, Millard A, Hall H, Savage R, Constantinidou C, Arasaradnam R, Nwokolo C. Urinary volatile organic compounds and faecal microbiome profiles in colorectal cancer. Colorectal Dis. 2019;21(11):1259-69.
28. Bleeker SE, Moll HA, Steyerberg EW, Donders AR, Derksen-Lubsen G, Grobbee DE, Moons KG. External validation is necessary in prediction research: a clinical example. J Clin Epidemiol. 2003;56(9):826-32.
29. Church TR, Wandell M, Lofton-Day C, Mongin SJ, Burger M, Payne SR, et al. Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer. Gut. 2014;63(2):317-25.
30. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. Stat Med. 2016;35(2):214-26.
31. Riley RD, Debray TPA, Collins GS, Archer L, Ensor J, van Smeden M, Snell KIE. Minimum sample size for external validation of a clinical prediction model with a binary outcome. Statistics in Medicine. 2021;40(19):4230-51.
32. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. J Clin Epidemiol. 2005;58(5):475-83.

33. Bosch S, de Menezes RX, Pees S, Wintjens DJ, Seinen M, Bouma G, et al. Electronic Nose Sensor Drift Affects Diagnostic Reliability and Accuracy of Disease-Specific Algorithms. Sensors (Basel). 2022;22(23).

34. Kort S, Brusse-Keizer M, Schouwink H, Citgez E, de Jongh F, van Putten J, et al. Diagnosing Non-Small Cell Lung Cancer by Exhaled-Breath Profiling Using an Electronic Nose: A Multicenter Validation Study. Chest. 2022.

35. Martin J, Falzone C, Romain AC. How well does your E-nose detect cancer? Application of artificial breath analysis for performance assessment. J Breath Res. 2024.

36. Śmiełowska M, Ligor T, Kupczyk W, Szeliga J, Jackowski M, Buszewski B. Screening for volatile biomarkers of colorectal cancer by analyzing breath and fecal samples using thermal desorption combined with GC-MS (TD-GC-MS). J Breath Res. 2023;17(4).

37. van Vorstenbosch R, Cheng HR, Jonkers D, Penders J, Schoon E, Masclee A, et al. Systematic Review: Contribution of the Gut Microbiome to the Volatile Metabolic Fingerprint of Colorectal Neoplasia. Metabolites. 2022;13(1).

38. Pleguezuelos-Manzano C, Puschhof J, Rosendahl Huber A, van Hoeck A, Wood HM, Nomburg J, et al. Mutational signature in colorectal cancer caused by genotoxic pks(+) E. coli. Nature. 2020;580(7802):269-73.

39. Janney A, Powrie F, Mann EH. Host–microbiota maladaptation in colorectal cancer. Nature. 2020;585(7826):509-17.

40. Watanabe D, Murakami H, Ohno H, Tanisawa K, Konishi K, Tsunematsu Y, et al. Association between dietary intake and the prevalence of tumourigenic bacteria in the gut microbiota of middle-aged Japanese adults. Scientific Reports. 2020;10(1):15221.

41. Amini M, Rezasoltani S, Asadzadeh Aghdaei H, Pourhoseingholi MA, Zali MR. Accuracy of the Discriminatory Ability of Combined Fecal Microbiota Panel in the Early Detection of Patients with Colorectal Cancer. J Gastrointest Cancer. 2023.

42. Bhandari MP, Polaka I, Vangravs R, Mezmale L, Veliks V, Kirshners A, et al. Volatile Markers for Cancer in Exhaled Breath-Could They Be the Signature of the Gut Microbiota? Molecules. 2023;28(8).

43. Qu R, Zhang Y, Ma Y, Zhou X, Sun L, Jiang C, et al. Role of the Gut Microbiota and Its Metabolites in Tumorigenesis or Development of Colorectal Cancer. Adv Sci (Weinh). 2023:e2205563.

44. Bosch S, Acharjee A, Quraishi MN, Bijnsdorp IV, Rojas P, Bakkali A, et al. Integration of stool microbiota, proteome and amino acid profiles to discriminate patients with adenomas and colorectal cancer. Gut Microbes. 2022;14(1):2139979.

45. Bretthauer M, Løberg M, Wieszczy P, Kalager M, Emilsson L, Garborg K, et al. Effect of Colonoscopy Screening on Risks of Colorectal Cancer and Related Death. New England Journal of Medicine. 2022.

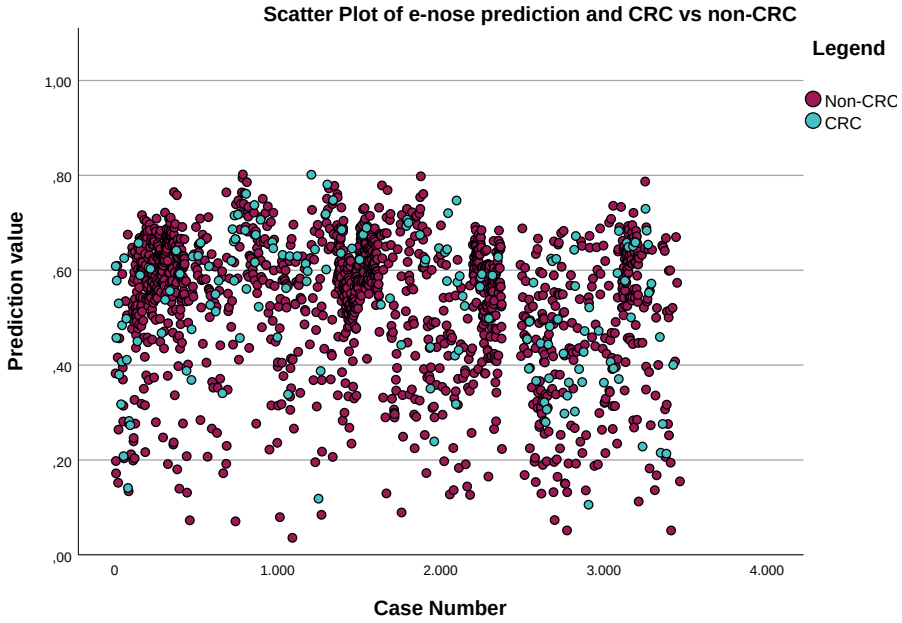
46. Zhu Y, Li X, Hu Y, Chen K, Zheng S, Ding K. Nonadherence to Referral Colonoscopy After Positive Fecal Immunochemical Test Results Increases the Risk of Distal Colorectal Cancer Mortality. Gastroenterology. 2023;165(6):1558-60.e4.

47. Jeerage KM, Berry JL, Murray JA, Goodman CA, Piotrowski PK, Jones CM, et al. The need for multicomponent gas standards for breath biomarker analysis. J Breath Res. 2022;16(4).

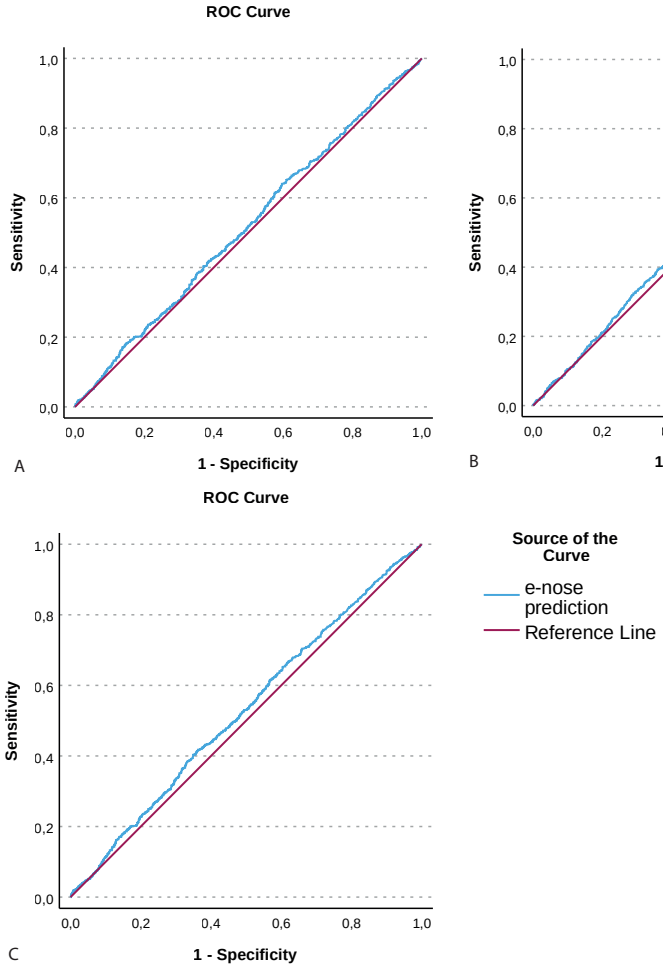
48. Hanna GB, Boshier PR, Markar SR, Romano A. Accuracy and Methodologic Challenges of Volatile Organic Compound-Based Exhaled Breath Tests for Cancer Diagnosis: A Systematic Review and Meta-analysis. JAMA Oncol. 2019;5(1):e182815.

49. (RIVM) tRvVeM. Landelijke monitor bevolkingsonderzoek darmkanker 2022. 2023.

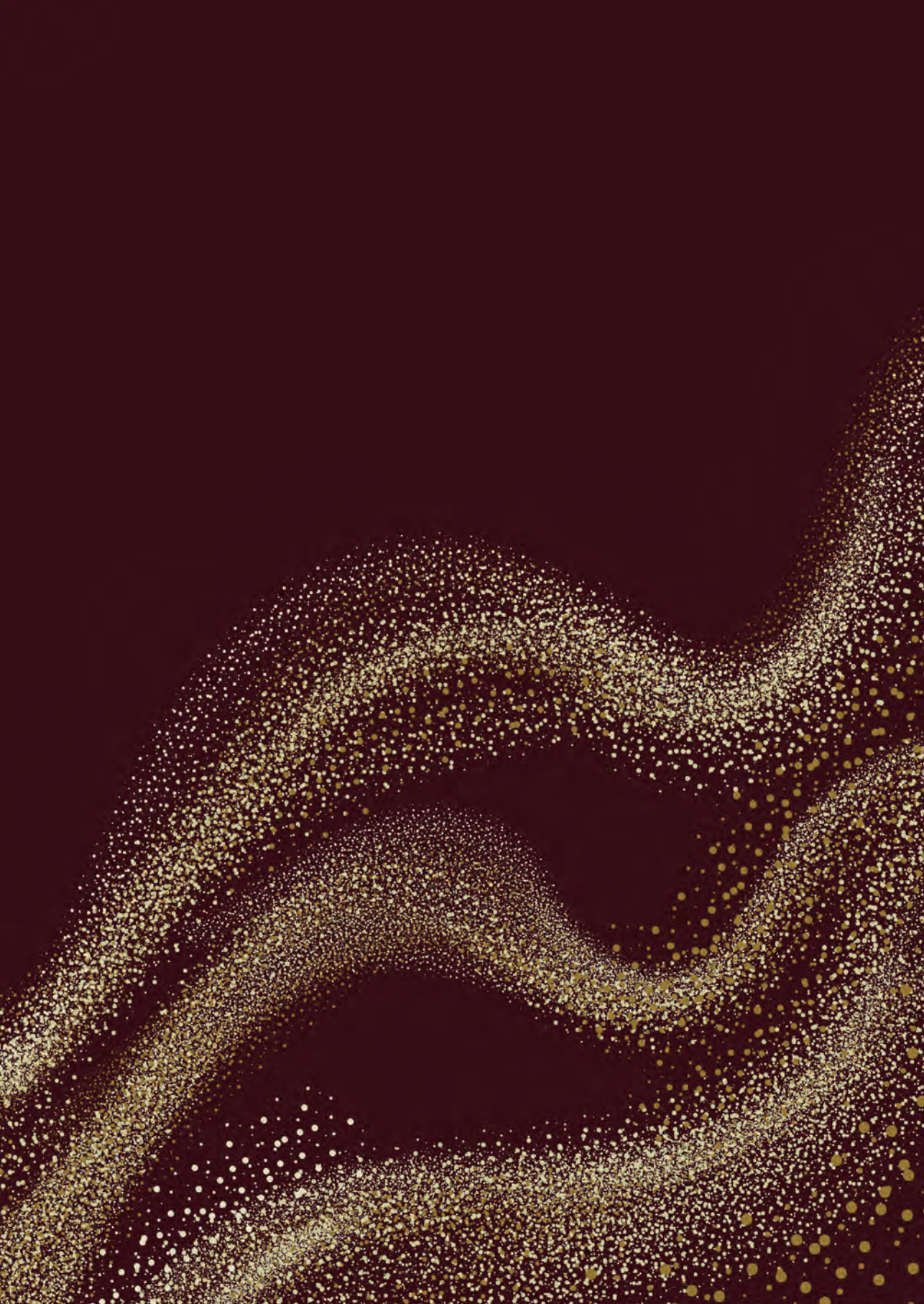
Supporting information



Supplementary figure S1. Scatter plot of breath test predictions. Little to no separation from the e-nose prediction value is seen. CRC; colorectal carcinoma.



Supplementary figure S2. Receiver operating characteristics (ROC) curve of e-nose blind predictions with a predictive model trained for advanced adenoma (AA). A – AA vs control. B – AA vs non AA. C – AA vs non AA excluding CRC subjects.



Chapter 3

Overcoming methodological barriers in electronic nose clinical studies, a simulation data-based approach

Milou L.M. van Riswijk, MD, Bas van Tintelen*, Ruben Lucas*,
Job van der Palen PhD, Peter D. Siersema, MD PhD

* authors contributed equally

J Breath Res. 2025 May 9;19(3). doi: 10.1088/1752-7163/add291.

Abstract

Background and study aim: Analysis of volatile organic compounds by electronic nose (e-nose) may adress gaps in non-invasive screening for neoplasia. Machine learning (ML) impacts study design and sample size requirements, but guidance on clinical study design is limited. This study evaluates how neoplasia prevalence, augmented data, and the number of e-nose devices impact sample size requirements.

Methods: Simulated e-nose breath test data were created using real-world study data. We examined the effect of varying neoplasia prevalence (50% to 5%) and data augmentation on model performance, as well as the impact of using multiple devices. Prediction models were developed using single value decomposition and random forest, and convolutional neural networks. Model performance was displayed as area under the receiver operating characteristics curve (AUC) and F1-score. Stable model performance was defined as the phase where additional data no longer increases model performance.

Results: Lower neoplasia prevalence significantly increased sample size requirements, with low-prevalence settings (5%) requiring up to five times more data than high-prevalence settings (50%) for stable model performance. Model performance varied between devices, and integrating data from multiple devices required larger sample sizes. Approximately 400 datapoints per device at 50% prevalence, and 2100 datapoints at 5% prevalence, were necessary to reach stable model performance.

Conclusions: Sample size requirements for e-nose studies are heavily influenced by disease prevalence and the number of devices used. Limiting device variability and ensuring sufficient case and control samples per device are crucial for achieving reliable predictive performance. Specific requirements will vary based on sensor and disease characteristics.

ClinicalTrials.gov Identifier: Clinicaltrials.gov Identifier NCT03346005 (model study) and NCT04357158 (validation study)

Introduction

Screening and early diagnosis of malignancies potentially improves morbidity and mortality of malignancies (1). Non-invasive methods improve accessibility and thereby effectiveness of screening programs (2). Emerging evidence suggests that volatile organic compounds (VOC) can be used as biomarker for malignancies, such as lung cancer (LC) (3) and colorectal cancer (CRC) (4-6). This could fill the current gap in non-invasive screening methods (7). VOCs can be detected using a low-cost, point-of-care electronic nose (e-nose) using machine-learning (ML) pattern recognition techniques.

Before clinical implementation, validation studies of VOC analysis using e-nose devices are imperative, due to the risk of overfitting and spurious correlations that can lead to overly optimistic claims on the accuracy (8, 9). Ideally, this follows a multi-step approach (10): After confirming the discriminative ability between cancer and healthy subjects in a pilot study, the accuracy can be assessed in larger study samples across the continuity of neoplastic lesions. Preferably, this encompasses larger studies to identify confounders and to provide a provisional positivity threshold.

However, when applying these steps to study designs of diagnostic tests that depend on ML, several methodological difficulties may be encountered (11). Traditionally, studies on new diagnostic tests are designed according to the STARD (12) and TRIPOD (13) (prediction model development and validation, respectively) guidelines. For higher generalizability, large multicenter designs generate better results. However, ML prediction models are generally very data hungry, especially when correcting for background noise (e.g., VOCs in a different hospital) (14), and diversity within cohorts of participants. By using a multicenter design and multiple measuring devices, subtle data differences are introduced, resulting in more data that is required to build a prediction model. Unfortunately, the exact sample size that is needed to cover these subtle differences is difficult to determine. ML prediction models may require ten to over twenty times as many events per variable compared to conventional regression models to develop a stable area under the receiver operating characteristics curve (AUC), due to the vast increase in parameters (14, 15). Furthermore, an increase in number of participating centers leading to higher sample size requirements is counterintuitive to non-ML study designs, where the total study sample size is divided over the participating centers. Illustrative for this, is the abundance of pilot studies in VOC research, while large-scale validation studies are still scarce (6, 16).

The diagnostic accuracy may differ across study populations with varying neoplasia prevalence. A promising accuracy in pilot studies may not hold in follow-up studies with a lower neoplasia prevalence (17), underlining the importance of validation studies. We have previously investigated VOC analysis in exhaled breath for LC and CRC using e-nose technology. For LC, we performed a multicenter diagnostic study including 575 subjects in seven centers (18). After a stepwise validation process (19), an AUC of 0.83 for the training set, and 0.79 in the validation set was derived. For CRC, we performed a similar study but in a CRC screening setting to mimic the setting of the anticipated application (5). In this population, prevalence was significantly lower (5%) compared to the LC prevalence in the aforementioned study of 40-43%, yielding an insufficient AUC of 0.54, despite promising results from a pilot study performed previously (AUC 0.76) (5).

As e-nose technology is relatively new, guidance for clinical diagnostic study designs is hardly available and demands both clinical methodological knowledge as technological knowledge. Therefore, in this study, we aim to provide insight in the effect of varying neoplasia prevalence, augmented data, and the number of measuring devices, on sample size requirements and diagnostic performance, using a series of simulation models.

Methods

E-nose data characteristics

We composed a simulated data set mimicking real study data from our LC- and CRC studies (Table 1). We investigated the diagnostic accuracy of e-nose technology for the detection of CRC in a pilot study (n=447), with a CRC prevalence of 15.7% (5). Breath tests of patients with a histopathological diagnosis of CRC were compared to individuals with a normal colonoscopy (prevalence 28.6%) to develop a predictive model, yielding an AUC of 0.74 in the validation set. Next, we included 3469 patients over 12 study sites referred for screening colonoscopy based on a positive fecal immunochemical test in the Dutch CRC screening program (unpublished results, clinicaltrials.gov identifier NCT03346005). In this study cohort, CRC prevalence was lower compared to the pilot study with 5.0%, and the control group with individuals without polyps or CRC had a prevalence of 25.1%.

The LC study had a sample size of 576 individuals across seven study sites, of which half was suspected of having LC, and the other half was healthy control (18). The

breath test outcomes were compared to the histopathological diagnosis of non-small cell lung carcinoma, which had a prevalence of 41.6% in the study cohort.

All breath tests were performed using the Aeonose (the eNose company, Zutphen, the Netherlands), a handheld e-nose device using three hotplate metal-oxide sensors (AS-MLV sensors; Applied Sensors GmbH, Germany) (20). VOCs interact with these sensors, causing redox reactions on the surface that lead to conductivity changes that can be measured. One breath profile contains over 2000 conductivity values. Unfortunately, the exact breath profile features of a disease as shown in these measurement values are as yet unclear.

Table 1. Overview of lung cancer study and colorectal carcinoma studies.

	LC study (18)	CRC pilot study (5)	CRC study
Study design, sample size (n)	Multicenter cross-sectional, n=575	Multicenter cross-sectional, n=447	Multicenter cross-sectional, n=3469
Neoplasia prevalence, n (%)	239 (41.6)	70 (15.7)	175 (5.0)
(Pre-neoplastic) lesions not included in prediction model	83 (14.4)	234 (52.4)	2425 (69.9)
Controls	253 (44.0)	128 (28.6)	869 (25.1)
Number of study sites, n	7	2	12
Number of used e-nose devices, n	5	5	21
Used ML techniques	Combined result of ANN, logres, RF, RF Extreme, XGBoost	ANN	CNN
AUC (95% CI)	0.86 (0.81-0.91)	0.74	0.54 (0.49-0.58)

ANN, artificial neural network; AUC, area under the receiver operating characteristics curve; CRC, colorectal carcinoma; CI, confidence interval; CNN, convolutional neural network; e-nose, electronic nose; LC, lung cancer; logres, logistic regression; ML, machine learning; RF, random forest; XGBoost, extreme gradient boosting.

Simulation data development

To analyze the effect of varying prevalence, augmented data, and number of e-nose devices on the required sample size, we used simulated data based on e-nose measurements of healthy control subjects from the LC study. On these measurements, we superimposed small deviations in the conductivity profiles, both in the time and temperature domains, to enable a distinction for positive and negative cases. These deviations were designed to mimic real breath samples. Therefore, we first aggregated multiple breath profiles into a single representative data sample by averaging across profiles. This step ensures that the resulting sample captures the overall trend while reducing random variability. Next, we

introduced artificial disease signals at various potential locations within the profile, accounting for the possibility that the disease may manifest inconsistently across different segments of the signal. To further simulate real-world conditions, a small "jitter" in both time and temperature was applied at these locations, mimicking natural variations and measurement noise. Finally, we superimposed Gaussian noise across the entire data sample to simulate ambient fluctuations and external noise, thereby increasing the robustness of the model to real-world environmental variations. The quality of the resemblance of the simulated data was compared with the published LC studies model. To realize this, a classifying ML model was used while reducing the number of cases from the LC study. Afterwards, we compared the performance of the ML classifiers in the LC dataset with an equally sized simulated dataset. Additionally, we compared our results using the *pmsampsize* package in *R* (15), which accounts for outcome prevalence, number of predictors, and expected model performance (AUC). The number of variables was defined as equivalent to the number of features remaining after the preprocessing steps.

Prediction model development

To get from the raw sensor data to a prediction model, several steps are needed (Figure 1). First, since most classifiers in ML prefer training on balanced datasets, i.e. a prevalence of 50%, we used augmented data as training with a prevalence of 5% could easily lead to "negative" classification in all cases, however still resulting in a correct outcome in 95% of all tests. Augmented data are synthetic data based on real data points, to simulate extra events to train the model (21). In our experience, the quality of the augmented data is critical: they should not resemble the original datapoints too much, but they should also not be too deviant to make recognition of profiles difficult. Therefore, the parameters for generating augmented data should be chosen carefully. For our simulation data set, we used Gaussian Noise, which is a statistical noise having a probability density equal to a normal distribution (22), using the Synthetic Minority Oversampling Technique (SMOTE).

Next, to control for slight variations between multiple e-nose devices before using this data in classifying models, all data is pre-processed. For this, normalization is used by scaling the measurement amplitudes between 0 and 1. Then, as the number of conductivity values per breath profile is usually large compared to the number of participants, it is common practice to apply data compression to prevent overfitting (23). The pre-processed and compressed data is then used in the supervised ML model. We used two different approaches for compressing and classifying both real and simulated breath profiles, using an average performance based on 10 runs of the classifier:

1. Principal Component Analysis (PCA) as compression technique followed by a Random Forest (RF) classifier. This approach worked well for data analysis of the LC study data. Nevertheless, a drawback of the PCA lies in its linear nature, potentially leading to the missing of non-linear features. Additionally, PCA tends to emphasize high-variance features, thereby risking the neglect of more subtle and nuanced features. We used RF with depth of 3 and 10000 trees, proceeded by a SVD that compresses the data to 17 features.
2. Compression and classification using a Convolutional Neural Network (CNN). The CNN is a powerful deep-learning tool, able to handle subtle differences in data sets. A disadvantage is that it requires an even larger number of data points than the PCA + RF approach. However, when using simulated data, this tool can be useful. We used a CNN consisting of the following layers:

    1. Conv2D with 8 filters, kernel size of 7, stride of 2, with padding and a relu activation.
    2. Conv2D with 16 filters, kernel size of 5, stride of 2, with padding and a relu activation.
    3. Batchnormalization
    4. Flatten
    5. Dense to 1 neuron with sigmoid activation.

The classifiers were selected based on experience in prior studies. To prevent overfitting, leave-10%-out cross validation is used, a resampling technique in which 10 rounds of validation are used to increase the likelihood of robust models. After a training and calibration phase, the model is tested on a blinded data set. After satisfactory performance, external validation in a separate cohort is performed.

Analyses

Using the simulated dataset, we aimed to assess three study situations. First, the effect of reducing the prevalence from an ideal 50% towards 5% (i.e., the neoplasia prevalence in the CRC study). Second, the effect of adding augmented data at different neoplasia prevalence values on the required sample size to achieve a similar model performance. Third, the effect of increasing or reducing the number of devices used in the study on the required sample size to achieve similar model performance. This was tested by comparing model performance trained on a single device to model performance trained on multiple devices.

For each situation, we calculated the effect on the model performance and plotted this against the required sample size. Performance was displayed as the F1-score

and AUC of the model (24). The F1 score is the harmonic mean of the precision (positive predicted value) and recall (sensitivity) of the model. A higher F1 score indicates a better accuracy and better representing the model accuracy in case of imbalanced datasets compared to the AUC. With this information, we aimed to get insight into sample size requirements to achieve stable model performance. Stable model was defined as the phase where additional data points no longer lead to an improvement of the AUC and F1-score.

We used Aethena (version 3.0, the eNose Company, Zutphen, the Netherlands) to generate the simulated data and Python 3.10 (Python Software foundation, Delaware, United states) and Tensorflow 2.13 (Google Brain, Alphabet Inc., California, United states) for testing the predictive model performance.

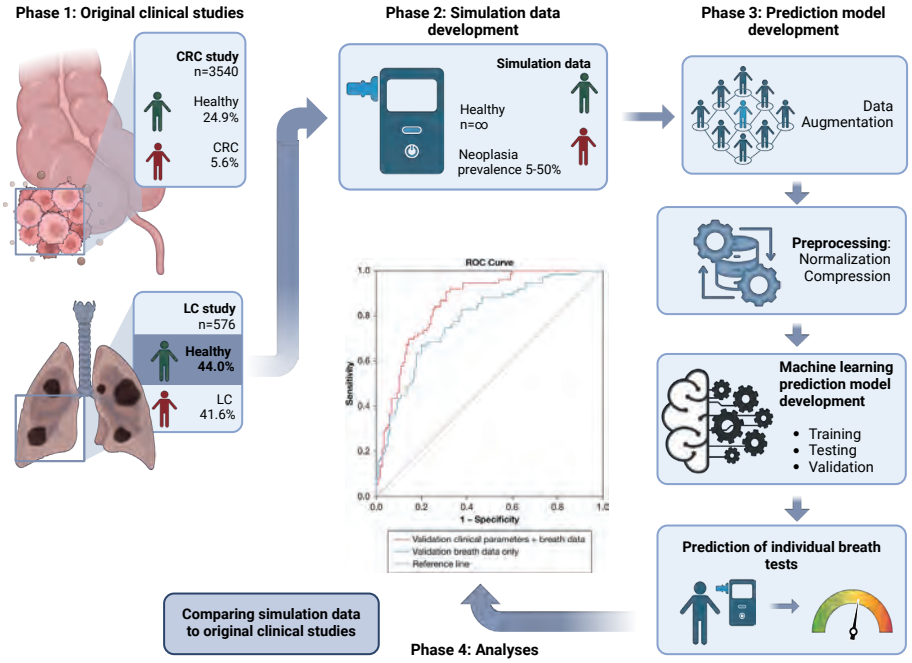


Figure 1. Graphical representation of simulation model development. We derived breath tests from healthy volunteers from our LC study (Phase 1) and superimposed small deviations to make positive and negative breath profiles with varying neoplasia prevalence values (Phase 2). To get from the raw sensor data to the prediction model (Phase 3), we first performed data augmentation using Gaussian Noise, then we preprocessed the data including standardization and compression. The pre-processed and compressed data is then used in the supervised ML model using 10% cross-validation and validation in a separate, resulting in the final prediction model to answer our research questions (Phase 4). Created with BioRender.com

Results

Simulation data characteristics

The resemblance of our simulation data to real data was tested with both CNN and PCA+RF as classifier (Table 2). Using PCA+RF, which were the ML techniques used for the LC study dataset, AUC values of the simulation data in the training set mimicked real data of the LC study. However, in the validation set, performance loss in the simulation dataset was greater than in the LC dataset. This performance loss is likely caused by the tendency of the PCA to prioritise high variance within the data. The artificial disease created for the simulation data is unrelated to the highest variance present in the data and is therefore "overlooked" by PCA compression. However, by using a CNN, which combines classification and compression, we were able to reproduce the LC results (Table 2). In lower prevalence settings, the model performance in the simulation dataset also approached the performance in the CRC dataset, confirming the validity of our simulation data.

Table 2. Overview of model performance in simulation data compared to the performed clinical studies.

	Simulation	LC study (18)	Simulation	CRC pilot study (5)
Neoplasia prevalence	43%	43%	10%	15.6%
Sample size				
Training set	410	410	400	153
Validation set	135	135	135	33
AUCmax				
Training set				
CNN	99.65	84.23	99.94	82.15
PCA+RFC	84.73	84.17	x	98.68
Validation set				
CNN	82.54	73.89	72.85	73.38
PCA+RFC	50.27	76.82	x	70.0

AUC, Area under the receiver operating characteristics curve; CNN, convolutional neural network; CRC, colorectal cancer; LC, lung cancer; RFC, random forest classifier; PCA, Principal components analysis.

Effect of prevalence

To assess the effect of neoplasia prevalence on AUC and F1-score, we used simulated data of five e-nose devices (Figure 2, Figure S1). At decreasing prevalence from 50% (ideal training prevalence) to 5% (comparable to the CRC cohort), increasingly more datapoints are necessary to maintain a similar AUC. For example, to reach an AUC of 0.80 with a disease prevalence of 5%, approximately 2100 datapoints are required, compared to approximately 600 datapoints with a prevalence of 30% and 400 datapoints with a 50% prevalence. Corrected for number of positive cases at lower neoplasia prevalence (i.e., 5% and 10%), the model performance

is lower as well (Figure S2). Furthermore, overfitting risk is demonstrated at lower prevalence models by the loss of high training-model performance to low validation-model performance (Figure S1). The risk of overfitting decreases with increasing datapoints.

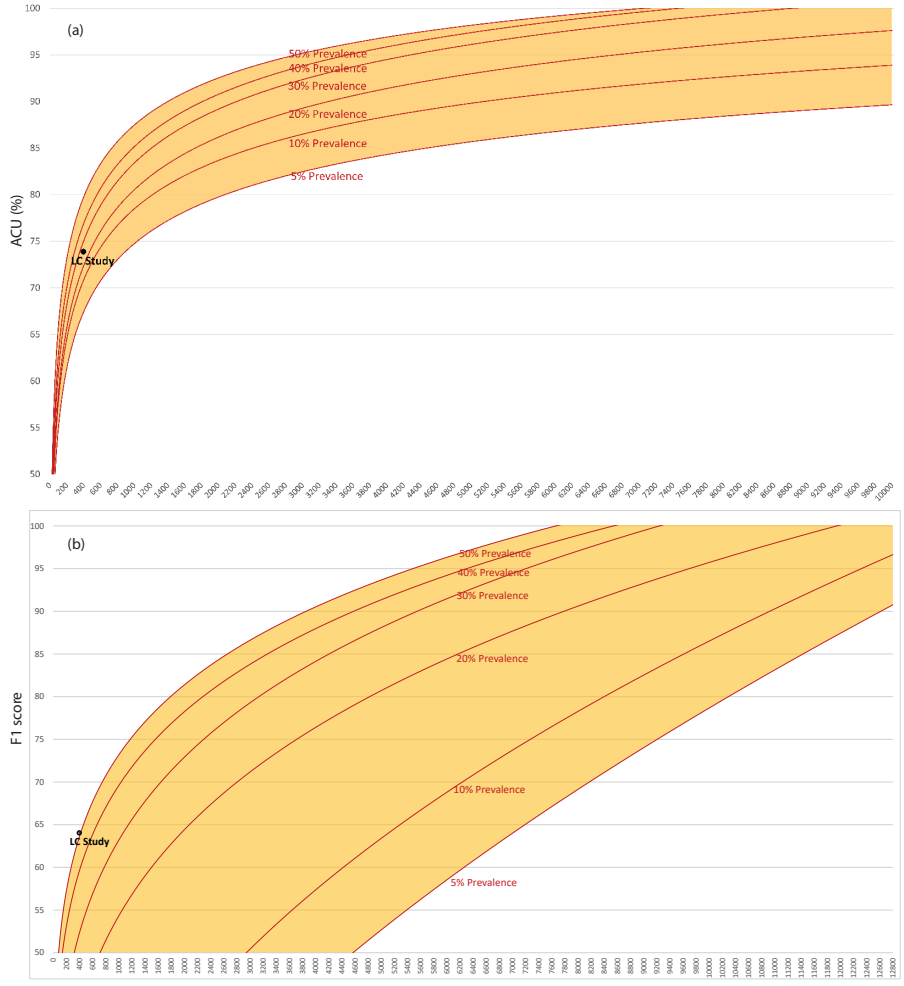


Figure 2. Effect of prevalence on required sample size on area under the curve (AUC) (A) and F1-score (B) calculated on the validation set using a convolutional neural network (CNN). For reference, the results of the LC study have been plotted in the graph. Lines represent polynomials to illustrate patterns since exact sample size requirements are related to disease and device characteristics.

Effect of augmented data

We assessed the effect of adding augmented data to the LC dataset and simulated dataset on model performance at varying sample sizes, using a CNN (Figure S1, dashed lines) and an PCA+RF model (Figure S3) including five e-nose devices. While augmented data may help reducing class-imbalance, the absence of model performance improvement in this figure demonstrates that augmented data is not able to substitute actual data points and provides only marginal improvement. Furthermore, when the used classifier performs insufficient, augmented data will not further improve this (Figure S3).

Effect of number of devices

Model performance varies for each device (Figure 3). Model training on a combination of devices does not translate into an equal increase of predictive performance. Nonetheless, model performance using combined data from three devices (i.e. model training on data from device 1, 2, and 3) was higher than the means of the model performances per individual device. Approximately 300 datapoints per device at a 50% prevalence are required to reach a plateau phase. Further adding devices here does not result in an additionally required total sample size.

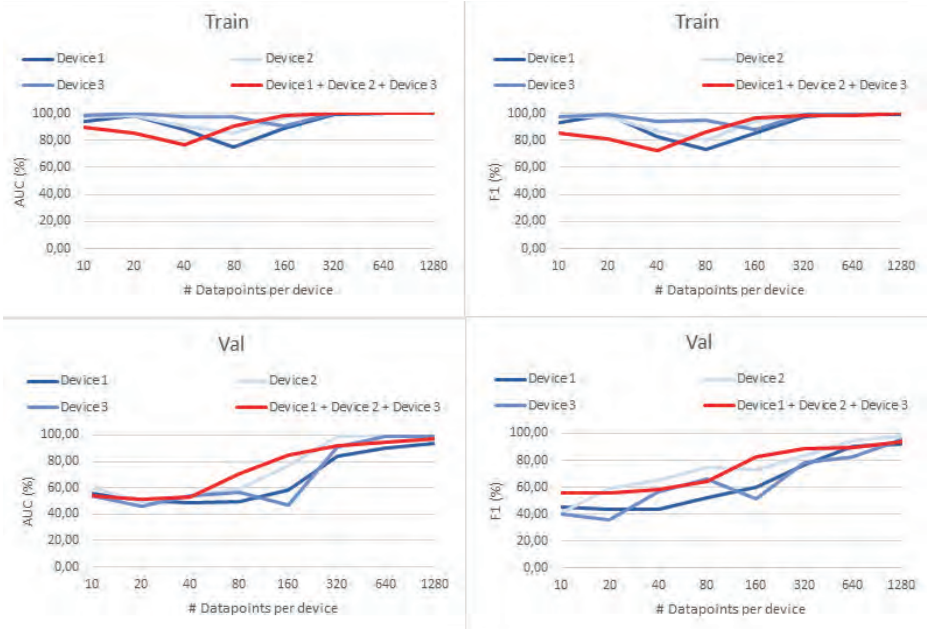


Figure 3. Effect of variation of number of datapoints per device at a 50% prevalence on model performance (AUC and F1-score) of single devices (device 1, device 2, device 3), and combination of devices (Device 1 + device 2 + device 3).

Discussion

The use of VOCs as non-invasive biomarkers for screening purposes seems promising. However, data from large scale studies and external validation are lacking. This could be due to methodological difficulties in e-nose studies, including the unknown impact of background noise, varying neoplasia prevalence, and use of multiple measuring devices on sample size requirement and study design. Our results indicate that in low-prevalent settings, exhaled breath analysis may require up to five times as much data compared to high-prevalent diseases to build a stable predictive model. Augmented data only gives marginal improvement of the predictive performance. Additionally, the use of multiple study devices has a considerable impact on the sample size that is required to build an adequate predictive model.

Previous studies on sample size requirements for diagnostic test studies mainly focus on logistic regression models. Several guidelines for sample size calculation have been proposed, such as the rule of thumb of 10 events per variable (15). However, the total sample size required is also dependent on the total number of participants, the outcome incidence, and expected model performance (15, 25). Riley et al. have proposed a method that includes the precision of the outcome proportion estimate, which is less likely to be overfitted (15). Following the provided *pmsampsize* package in R, where the number of variables is set equal to the number of features that remain after preprocessing of the current studies, this results in similar outcomes as the sample size used in the published LC study (18) and our simulation data. This adds to the validity of our results.

Diagnostic performance in high-prevalent study settings in our simulation were as expected from previous pilot studies, but in low-prevalent settings, predictive performance was subpar. This correlates with findings from our CRC studies. It is known that training of ML models should be performed on a balanced dataset, i.e., number of positive cases and controls should be present in relatively equal parts. In case of unbalanced data, it is general practice to use augmented data (21). In contrast to reports in literature on use of augmented data in imaging techniques (21), our study results demonstrate that augmented data was not able to substitute real datapoints. This may explain the need for extra datapoints in low-prevalent study settings. A second explanation is that our simulated dataset used data of five e-nose devices to mimic actual study settings. However, our results also demonstrate that the number of devices negatively impacts the number of required positive cases, as a low number of positive cases per study device results in a low predictive

performance. Together, this may clarify the effect of a low prevalence of the studied disease on the poor predictive performance in our studies.

Throughout our previous studies, multiple study devices were used, ranging from 5-6 in the LC and CRC pilot studies, to 21 in the CRC study. Due to case mix, local VOCs in ambient air, or sensor characteristics, predictive performances differed per used study device. A frequently occurring problem in e-nose devices is sensor drift over time (26). The micro-hotplate metal-oxide sensors used in the current study device have been developed to enable transferable calibration (20), which minimizes the effect of sensor drift. Other sensor types, such as conducting polymer sensors, may be subject to more inter-device deviations. This impairs robust prediction model development because correction of sensor drift on the signals requires a large dataset of baseline breath samples (26). Therefore, our results demonstrate that required sample size is not only dependent on number of datapoints and prevalence variance but may be more dependent on the number of datapoints per study device. Future study designs should therefore strongly consider the number of devices used, and where possible, restrict the number of devices, instead of using several devices in large multicenter studies to catch as many cases as possible in low-prevalent settings. After establishing a robust model, which is not affected by inter-device differences, external validation in a cohort using a higher number of study devices can be performed to ensure adequate performance in clinical practice.

Despite its vital importance (8, 9, 27, 28), e-nose studies with a truly external validation cohort have yet not been published. Our study demonstrates that for validation purposes in low-prevalence study settings, a substantial number of data points is required. This might explain the discrepancy between the high number of positive pilot studies and the low number of validation studies. Furthermore, this may also delay the use of e-nose technology in clinical practice in the not too far future. Given the high diagnostic performance with pooled sensitivity and specificity of over 80% in meta-analyses of pilot studies (6, 29), there is evidence for VOCs as a diagnostic biomarker, despite the variation in analysis methods. However, for true confirmation, external validation is imperative, preferably in a multicenter setting to guarantee robust results. Analysis methods of VOCs in pilot studies vary greatly, from traditional methods such as gas-chromatography-mass-spectrometry (GCMS) to the use of e-nose technology. Since traditional analysis methods such as GCMS are developed at a high calibrated level, this may result in a lower signal to noise ratio, which positively impacts sample size requirements. Clinical use of VOCs as biomarkers may therefore initially better be done with traditional methods like GCMS, albeit more labor intensive, than with e-nose devices.

Our results implicate that future study designs should consider limiting the number of study devices in a study. In addition, a minimum number of positive cases per study device should be taken into account. Since low-prevalent settings may lead to limited predictive performance, this may warrant a study design with an artificially high prevalence, over development in a setting in which testing will ultimately take place, such as for cancer screening purposes. Following this, external validation could then take place in lower prevalence settings to confirm the accuracy in the intended clinical setting.

This study has several strengths. It is the first to report on data characteristics of e-nose devices and its behavior in machine learning predictive models and may provide directions for future study designs. This may prevent unnecessary endeavors and associated costs for performing a study and expedite reporting of external validation studies. Although different types of e-nose devices may have proprietary components in sensors and analysis methods, the main steps are similar for different devices (23). Our conclusions may apply beyond the specific e-nose we used during the studies.

A limitation of the use of e-nose technology is that the exact sensor signal features that are used for disease recognition remain unknown. Different types of e-nose devices all need a process of pre-processing and feature extraction to evaluate a signal with ML, and the 'black-box' characteristic of ML complicates monitoring of measured signals (23). This inherently limits in depth analysis and more precise calculations. Additionally, it hampers designing of simulation data, which is demonstrated by the loss of predictive performance in ML models that use data compression in a separate step, unlike the CNN model that was used in our final calculations. While our simulated data was designed assuming a normal distribution, VOC distributions in breath research may not always follow a normal distribution. Testing non-normally distributed data might influence the required sample size to achieve stable model performance and should be considered in future studies. Nevertheless, the similarity of the model performance in simulation data with both the LC study and CRC study confirms the validity of our simulation data. Additionally, different diseases, sampling media, and sampling devices may create different profiles that include more explicit features. Hypothetically, as this increases the effect size, and thereby decreases sample size requirements to achieve similar model performance, our results may not directly translate to other clinical contexts. Additionally, the use of different classifiers may be more fitting to be used in data from other e-nose devices and may lead to variations in sample size requirements to obtain stable model performance. In the future, more precise

feature selection and visualization could result in more efficient data analysis, leading to fewer subjects needed to build robust disease prediction models. Therefore, a currently ongoing study aims to explore the exact features of positive and negative signals.

Conclusion

Sample size requirements for e-nose studies strongly depend on disease prevalence and number of study devices used throughout the study. We recommend limiting the number of devices and aim for a minimum number of 150 neoplasia cases per e-nose device for studies aiming at determining discriminative ability and a provisional positivity threshold. This may result in adapting study designs to increase the prevalence in the study cohort to be able to build a robust prediction model.

Acknowledgements

We thank all study staff and collaborators from both clinical studies for their efforts and all subjects for their collaboration.

Ethical statement

The original studies were conducted in accordance with the declaration of Helsinki and approved by the Medical Ethical committee of the participating hospitals before study commencement. All subjects provided written informed consent.

Data accessibility statement

The simulation data and analytical codes can be provided upon reasonable request. For the original breath test data, we refer to the original published studies by Van Keulen et al. and Kort et al.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
2. Lennon AM, Buchanan AH, Kinde I, Warren A, Honushefsky A, Cohain AT, et al. Feasibility of blood testing combined with PET-CT to screen for cancer and guide intervention. Science. 2020;369(6499):eabb9601.
3. Kort S, Tiggeloven MM, Brusse-Keizer M, Gerritsen JW, Schouwink JH, Citgez E, et al. Multi-centre prospective study on diagnosing subtypes of lung cancer by exhaled-breath analysis. Lung Cancer. 2018;125:223-9.
4. Woodfield G, Belluomo I, Laponogov I, Veselkov K, Cross AJ, Hanna GB, et al. Diagnostic performance of a non-invasive breath test for colorectal cancer: COBRA1 study. Gastroenterology. 2022.
5. van Keulen KE, Jansen ME, Schrauwen RWM, Kolkman JJ, Siersema PD. Volatile organic compounds in breath can serve as a non-invasive diagnostic biomarker for the detection of advanced adenomas and colorectal cancer. Aliment Pharmacol Ther. 2020;51(3):334-46.
6. Chandrapalan S, Bosch S, Cubiella J, Guardiola J, Kimani P, Mulder C, et al. Systematic review with meta-analysis: volatile organic compound analysis to improve faecal immunochemical testing in the detection of colorectal cancer. Aliment Pharmacol Ther. 2021;54(1):14-23.
7. de Boer NK, de Meij TG, Oort FA, Ben Larbi I, Mulder CJ, van Bodegraven AA, et al. The scent of colorectal cancer: detection by volatile organic compound analysis. Clin Gastroenterol Hepatol. 2014;12(7):1085-9.
8. Leopold JH, Bos LD, Sterk PJ, Schultz MJ, Fens N, Horvath I, et al. Comparison of classification methods in breath analysis by electronic nose. J Breath Res. 2015;9(4):046002.
9. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med. 2000;19(4):453-73.
10. Bresalier RS, Senore C, Young GP, Allison J, Benamouzig R, Benton S, et al. An efficient strategy for evaluating new non-invasive screening tests for colorectal cancer: the guiding principles. Gut. 2023.
11. Obermeyer Z, Emanuel EJ. Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. New England Journal of Medicine. 2016;375(13):1216-9.
12. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. Bmj. 2015;351:h5527.
13. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Bmj. 2015;350:g7594.
14. van der Ploeg T, Austin PC, Steyerberg EW. Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints. BMC Medical Research Methodology. 2014;14(1):137.
15. Riley RD, Ensor J, Snell KIE, Harrell FE, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. BMJ. 2020;368:m441.
16. van Riswijk MLM, Siersema PD. Letter: meta-analysis of new diagnostic methods for colorectal cancer detection-only as strong as its weakest link. Aliment Pharmacol Ther. 2021;54(4):540.
17. Church TR, Wandell M, Lofton-Day C, Mongin SJ, Burger M, Payne SR, et al. Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer. Gut. 2014;63(2):317-25.

18. Kort S, Brusse-Keizer M, Schouwink H, Citgez E, de Jongh F, van Putten J, et al. Diagnosing Non-Small Cell Lung Cancer by Exhaled-Breath Profiling Using an Electronic Nose: A Multicenter Validation Study. Chest. 2022.
19. Haalboom M, Kort S, van der Palen J. Using a stepwise approach to simultaneously develop and validate machine learning based prediction models. J Clin Epidemiol. 2021.
20. Bruins M, Gerritsen JW, van de Sande WWJ, van Belkum A, Bos A. Enabling a transferable calibration model for metal-oxide type electronic noses. Sensors and Actuators B-Chemical. 2013;188:1187-95.
21. Chlap P, Min H, Vandenberg N, Dowling J, Holloway L, Haworth A. A review of medical image data augmentation techniques for deep learning applications. J Med Imaging Radiat Oncol. 2021;65(5):545-63.
22. Arslan M, Guzel M, Demirci M, Ozdemir S, editors. SMOTE and Gaussian Noise Based Sensor Data Augmentation. 2019 4th International Conference on Computer Science and Engineering (UBMK); 2019 11-15 Sept. 2019.
23. Scott SM, James D, Ali Z. Data analysis for electronic nose systems. Microchimica Acta. 2006;156(3):183-207.
24. Parasa S, Repici A, Berzin T, Leggett C, Gross SA, Sharma P. Framework and metrics for the clinical use and implementation of artificial intelligence algorithms into endoscopy practice: recommendations from the American Society for Gastrointestinal Endoscopy Artificial Intelligence Task Force. Gastrointest Endosc. 2023;97(5):815-24.e1.
25. van Smeden M, Moons KG, de Groot JA, Collins GS, Altman DG, Eijkemans MJ, et al. Sample size for binary logistic prediction models: Beyond events per variable criteria. Statistical Methods in Medical Research. 2019;28(8):2455-74.
26. Bosch S, de Menezes RX, Pees S, Wintjens DJ, Seinen M, Bouma G, et al. Electronic Nose Sensor Drift Affects Diagnostic Reliability and Accuracy of Disease-Specific Algorithms. Sensors (Basel). 2022;22(23).
27. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Intern Med. 1999;130(6):515-24.
28. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Ann Intern Med. 2006;144(3):201-9.
29. Zhou W, Tao J, Li J, Tao S. Volatile organic compounds analysis as a potential novel screening tool for colorectal cancer: A systematic review and meta-analysis. Medicine (Baltimore). 2020;99(27):e20937.

Supporting information

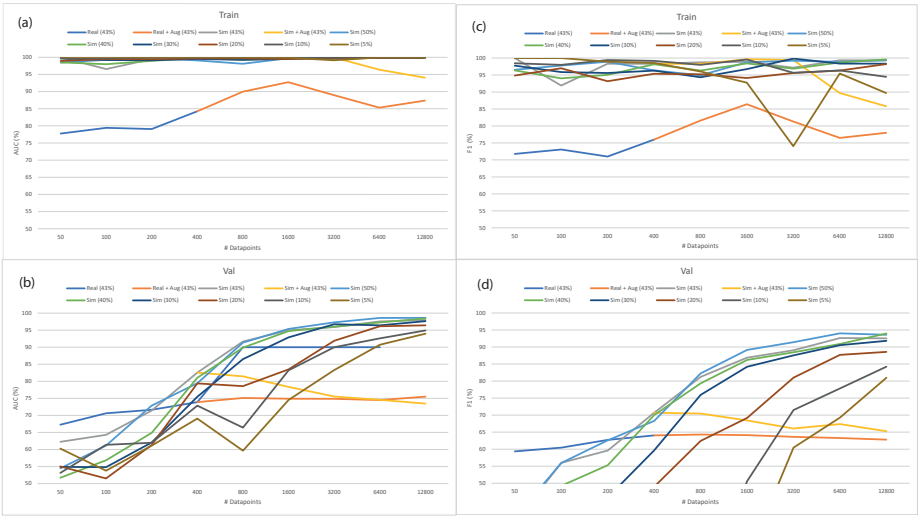


Figure S1 (A-D). Effect of prevalence and addition of augmented data (aug) on required sample size on area under the curve (AUC) (AB) and F1-score (CD) of training and validation cohort. Extra data points increase the models accuracy, while this is only very limited for augmented data. *AUC – area under the receiver operating characteristics curve, Aug – augmented data, Sim – simulation data*.

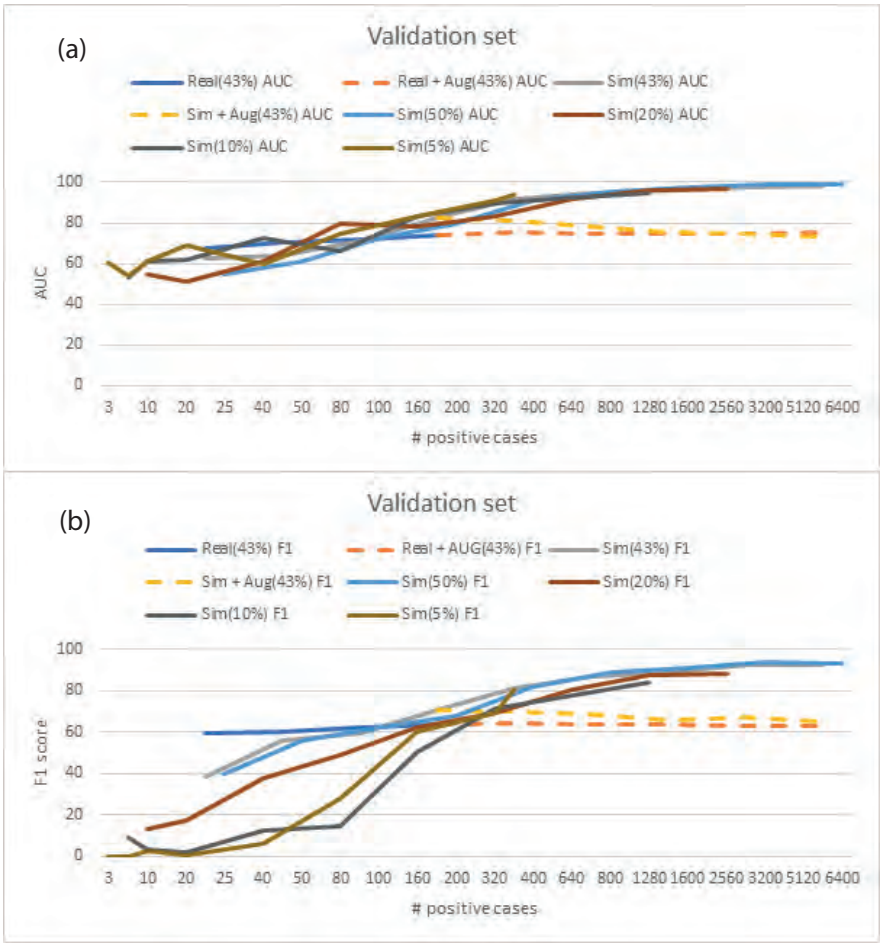


Figure S2 (AB). Effect of prevalence and addition of augmented data (aug) on required positive cases on area under the curve (AUC) (A) and F1-score (B) of validation cohort. *AUC – area under the receiver operating characteristics curve, Aug – augmented data, Sim – simulation data*.

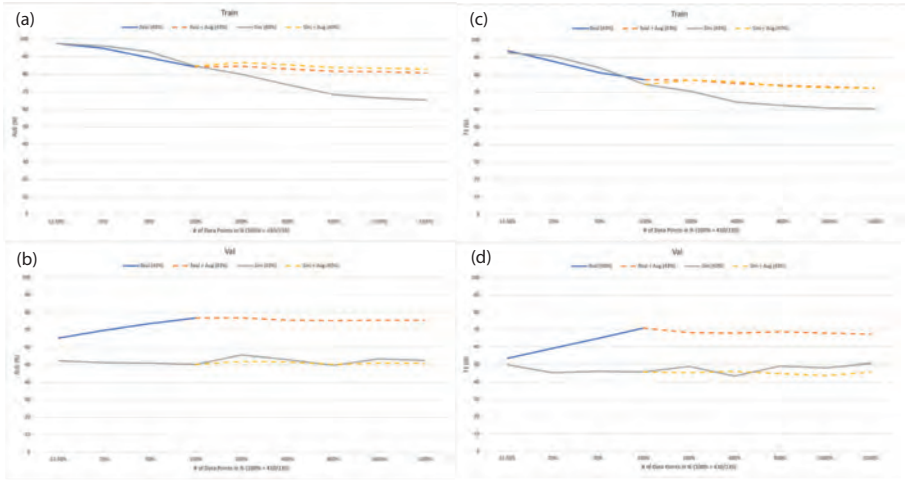
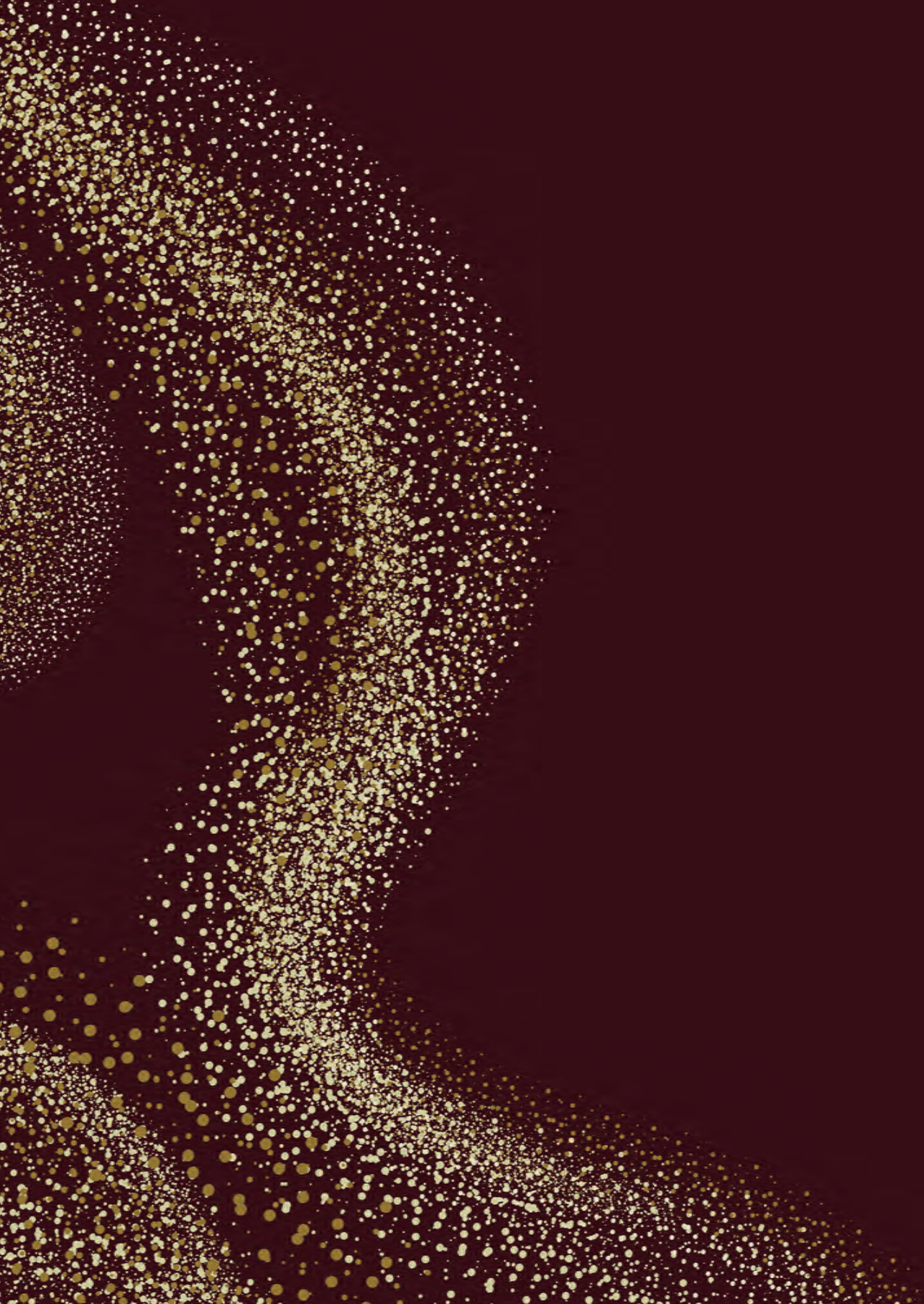
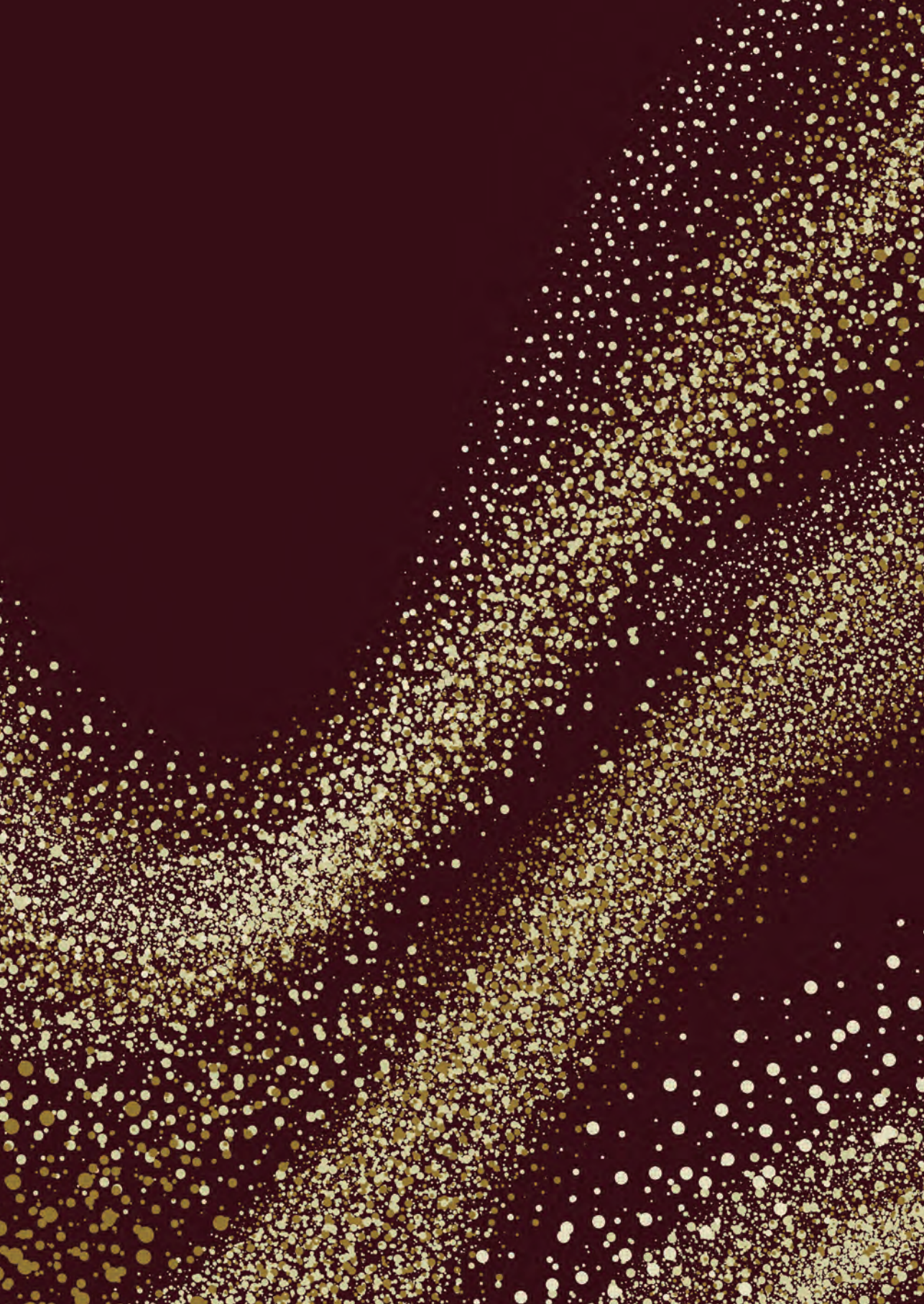


Figure S3 - AB. Effect of addition of augmented data to training (A) and validation (B) cohort on AUC generated with a combined Principal Component Analysis (PCA) random forest (RF) network, versus addition of extra datapoints. CD. Effect of addition of augmented data to training (C) and validation (D) cohort on F1-scores, versus addition of extra datapoints, using a combined PCA-RF network.



Part II: Improvement of bowel preparation for colonoscopy

She is water.
Powerful enough to drown you
Soft enough to cleanse you
Deep enough to save you.



Chapter 4

Efficacy of ultra-low volume (≤1L) bowel preparation fluids - systematic review and meta-analysis

Milou L.M. van Riswijk M.D., Kelly E. van Keulen M.D., Peter D. Siersema M.D. PhD

Dig Endosc. 2022 Jan;34(1):13-32. doi: 10.1111/den.14015.

Digestive Endoscopy 2022 Jan;34(1):13-32

Abstract

Background and aims: High-quality bowel preparation is paramount for the diagnostic accuracy and safety of colonoscopy; however, it is often difficult for patients to adhere to high-volume laxatives, which may contribute to poor bowel preparation. This review aims to assess the efficacy of bowel preparation fluids of 1L or less (≤1L).

Methods: We performed a systematic review including all relevant randomized controlled trials on ultra-low volume (≤1L) bowel preparation fluids for colonoscopy published since 2015. Primary endpoint was the percentage of adequately prepared patients. Secondary endpoints included adenoma detection rate (ADR) and safety.

Results: Bowel preparation with sodium picosulfate/magnesium citrate (SPMC) (19 trials, n=10,287), 1L-polyethylene glycol with ascorbate (PEGA) (10 trials, n=1,717), sodium phosphate (NaP) (2 trials, n=621), and oral sulfate solution (OSS) (3 trials, n=597) was adequate in 75.2%, 82.9%, 81.9%, and 92.1%, respectively, of patients; however, heterogeneity between studies was considerable (I^2 range: 86%-98%). Pooled ADRs were 31.1% with SPMC, 32.3% with 1L-PEGA, 30.4% with NaP, and 40.9% with OSS. Temporary electrolyte changes were seen with all ultra-low volume bowel preparation fluid solutions but without sustained effects in most patients.

Conclusion: Ultra-low volume bowel preparation fluids do not always meet the 90% quality standard for adequate bowel preparation as defined by current guidelines. Nonetheless, they may be considered in patients intolerant for higher-volume laxatives and without risk factors for inadequate bowel preparation or dehydration-related complications.

Definitions: High-volume bowel preparation – more than 2 liters (>2L). Low-volume bowel preparation – 2 liters or less (≤2L). Ultralow-volume bowel preparation – 1 liter or less (≤1L).

Introduction

Colonoscopy is considered the gold standard for screening and surveillance of colorectal cancer (CRC) and its precursor lesions. However, diagnostic accuracy and safety of colonoscopy highly depend on the quality of preprocedural bowel preparation. Inadequate bowel preparation has been reported as frequent as 25% and is associated with a lower adenoma detection rate (ADR), lower procedure completion rate, longer procedure time, higher complication rate, and a higher need for repeat colonoscopy with associated increased healthcare costs (1-7). In light of this, the European Society of Gastrointestinal Endoscopy (ESGE) guidelines advise that at least 90% of the colonoscopy patients should have adequate bowel preparation (1, 8).

Inadequate bowel preparation is often linked to the high volume of laxatives patients need to drink (9, 10). Moreover, the high burden of bowel preparation may be one of the reasons for patients not to undergo colonoscopy (11, 12). In the past few years, several strategies have been developed to ensure adequate bowel cleansing, aiming to improve bowel preparation tolerability while maintaining an adequate cleansing effect. The reference standard for bowel preparation consisted for a long time of 3-4 liters (L) of polyethylene glycol (PEG) electrolyte solution (1, 2) due to its efficacy and favorable safety profile (13). More recent randomized clinical trials (RCT) and meta-analyses comparing lower volumes (2L) of bowel preparation solutions to standard regimes, demonstrated that the former benefit patient compliance and show a higher willingness to repeat colonoscopy while still leading to a high bowel cleansing efficacy (14-16).

Nonetheless, even two liters of poorly tasting laxatives is still less optimal for a subgroup of patients (9, 17). In an effort to further optimize patient experience and compliance, several ultra-low volume bowel preparation fluids of 1L or less have been developed, based on either hyperosmotic solutions or stimulant laxatives (supplementary A). In a recently published observational study including 5000 patients, a 300ml bowel preparation solution consisting of sodium picosulfate with magnesium citrate (SPMC) demonstrated a high willingness to repeat colonoscopy (93.5%) when compared to 4L PEG (69.4%) and 2L PEG with ascorbate (73.2%)(18).

The ESGE recommends both high-volume (>2L) as well as low- (≤2L), or ultralow-volume (≤1L) laxatives in healthy patients, based on a non-inferiority outcome of individual studies. However, not all included studies investigating ultra-low volume laxatives were found to meet the quality standard of a minimum of 90%

adequate bowel preparation. Considering that the cleansing efficacy is even more important than tolerability(3), the scope of this review is to assess the efficacy and safety of ultra-low volume fluids (≤1L) to achieve adequate bowel preparation for colonoscopy.

Materials and methods

Protocol and registration

The protocol was designed in line with the PRISMA guidelines and registered in the PROSPERO database of systematic reviews (CRD42020181630).

Information sources and search strategy

The search was systematically performed on April 17, 2020 in three databases: PubMed, Embase (Ovid interface), and the Cochrane Library (CENTRAL). The search strategy and search terms were developed in collaboration with a medical librarian. Search terms included "colonoscopy", "laxatives", "cathartics", "purgatives", "bowel evacuant", "bowel preparation", "bowel cleansing", "colon cleansing", "visualization", "lavage". Alternative spelling was accounted for. The full search strategy is available in supplementary B.

Eligibility criteria

We included RCTs that investigated bowel preparation fluids with a volume of ≤1L, published between January 1, 2015 and April 17, 2020. We excluded studies that did not report original data, animal studies, studies focusing on a specific study population, and conference abstracts. The search was limited to articles either in English or Dutch, with full-text available through the university library or open access publishing.

Study selection

To deduplicate the records, we used Endnote X9.2 (Clarivate Analytics, Philadelphia, PA, USA), after which all remaining records were transported to the web-based screening program Rayyan QCRI (19). Eligible studies were identified by one researcher (MvR). Uncertainties were resolved through discussion with the senior author (PS).

Outcome measures

Our primary endpoint was the proportion of adequately prepared patients on an intention-to-treat basis. Adequate bowel cleansing was defined as a Boston Bowel

Preparation Scale (BBPS) score ≥6, Aronchick Scale (AS) score ≤2 (good or excellent), Ottawa Bowel Preparation Scale (OBPS) score ≤5, and Harefield cleansing scale (HCS) grade A or B (20).

If the outcome was reported with more than one preparation scale, BBPS and OBPS were preferred over AS, as previous studies have shown better interobserver consistency with a Cohen's kappa coefficient of 0.77, 0.94, and 0.77 for BBPS, OBPS, and AS, respectively. Furthermore, AS is preferred over HCS (kappa 0.457) (20-22). Additionally, BBPS was preferred over OBPS because of more extensive validation and more frequent use in clinical practice (1, 20).

Secondary endpoints included ADR and safety. If the primary outcome was not reported, the study was not included in the meta-analysis for efficacy, but only in the safety analysis.

Statistical analysis

We used a random-effects model to calculate the pooled proportion of adequately prepped patients and ADR per type of fluid, using the restricted maximum likelihood method (23). A Freeman-Tukey double arcsine transformation was used to minimize the effect of extreme proportions (near 1 or 0) in study subsets with small sample sizes and to stabilize variances(24). Additionally, subgroup analyses and meta-regression on predefined subgroups were conducted. We assessed the effects of the use of additives (i.e., adjunctive laxatives drugs prescribed besides the main laxative, e.g., bisacodyl), the dosing protocol (split-dose, same day, or day before), and diet (liquid diet, low-residue diet, or a combination).

Heterogeneity across the pooled studies was assessed using I^2 statistics, with low, moderate, and substantial heterogeneity defined as 25%, 50%, and 75%, respectively (25). To further explore heterogeneity and to detect possible outliers, influence analyses were conducted, including leave-1-out sensitivity analyses and Baujat-plots (26, 27). If more than 10 studies were available, a graphic display of study heterogeneity (GOSH plot) analysis was conducted(28). All analyses were conducted in R3.6.2 (29), using the packages *meta (30), metafor (31),* and *dmetar (32).*

Risk of bias assessment

To assess and visualize risk of bias in the included studies, the Cochrane Collaboration Risk of Bias 2 (RoB2) tool was used for randomized interventional trials(33). Selection bias was assessed using funnel plots.

Results

Search results

Our systematic literature search yielded 5097 citations. After deduplicating, 3029 were screened based on title and abstract. Based on potential relevance, 239 articles were screened full-text, of which 43 were included. The full selection process is shown in the PRISMA flowchart (figure 1) (34).

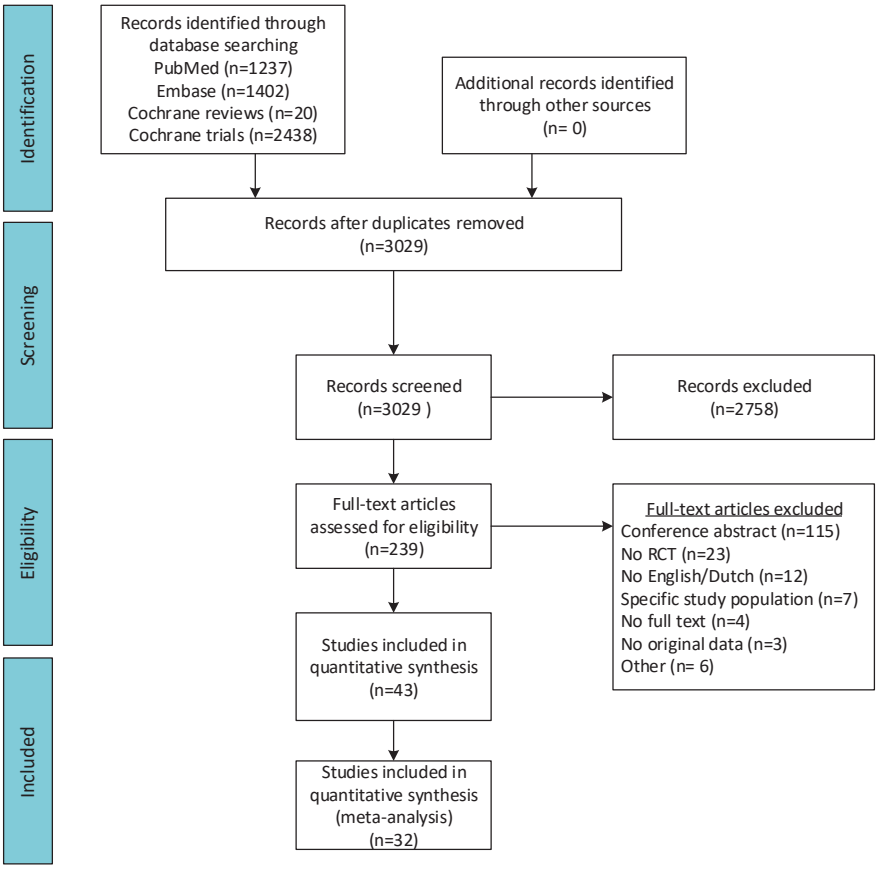


Figure 1. PRISMA flow diagram of study selection. RCT – randomized controlled trial.

Study characteristics

Of the 43 included studies, 26 evaluated SPMC (16, 34-59), 12 1L-PEG with ascorbate (PEGA) (46, 58-68), 4 oral sulphate solution (OSS) (61, 69-71), 4 sodium phosphate solution (NaP) (38, 72-74), 2 sennoside s(72, 75), and 1 magnesium citrate (76) (table 1).

Table 1. Ultralow-volume bowel preparation fluids, study characteristics.

Author (year)	Country	Study design	Reason for colonoscopy	Sample size (ITT)	Age (mean, years)	Intervention (low-volume)	Comparison (high-volume)	Additional fluid intake
Choi 2016[§]	Korea	SC, SB RCT	Diagnostic, therapeutic	1. 102 2. 98	1. 49.9 2. 51.6	1. Split 400ml SPMC (Picolight) + 10mg bisacodyl	2. Split 2L PEGA (Coolprep) + 20ml simethicone	1. 2L 2. 1.5L
Dwyer 2017[§]	Australia	SC non-inferiority RCT	Clinically accepted indications	1. 112 2. 118	54 (full study)	1. split 300ml SPMC (Picosalax), white diet	2. 1L PEG (Glycoprep-C) + split 300ml SPMC (picoprep), LQD	1. 0.2L/h 2. 0.2L/h
Gweon 2015	Korea	SC, SB, non-inferiority RCT	Various	1. 104 2. 105	1. 47.7 2. 50.6	1. SD Split 300ml SPMC + 10mg bisacodyl (no productnames reported)	2. Split 4L PEG	1. Up to 4L 2. NR
Heetun 2016	Ireland	SC RCT	Complaints, screening, surveillance	1. 102 2. 107 3. 122	1. 56.4 2. 58.7 3. 56.8	1. split 300ml SPMC, LQD 3. 240ml NaP split/ DB in morning procedures, LQD	2. 4L PEG, LQD (No productnames reported)	extra fluids encouraged
Hookey 2019[§]	Canada	MC SB non-inferiority RCT	Elective	1. 448 2. 453	57.2 (full study)	1. split 320ml SPMC solution (Clenpiq), LQD 2. split 300ml SPMC powder (Prepopik), LQD		1. 2L 2. 2L
Hung 2020	Taiwan	MC SB non-inferiority RCT	Elective	1. 316 2. 315	1. 47.7 2. 49.4	1. split 300ml SPMC (Bowclean), LRD	2. Split 2L PEG (Kleanprep) + 5mg bisacodyl, LRD	1. 2L 2. None

Table 1. Continued

Author (year)	Country	Study design	Reason for colonoscopy	Sample size (ITT)	Age (mean, years)	Intervention (low-volume)	Comparison (high-volume)	Additional fluid intake
Jun 2017[§]	Korea	SC SB RCT	Screening, diagnostic	1. 99 2. 105	1. 54.6 2. 54.3	1. split 450 ml SPMC (Picolight), LRD 3d 2. Split 450ml SPMC (Picolight) with flexible timing of second dose, LRD 3d.		1. 2.5L 2. 2.5L
Kiesslich 2017[§]	Germany	MC, SB RCT	Elective	1. 131 2. 73	1. 58.3 2. 56.6	1. Split 300ml SPMC (Picoprep) 2. DB 300ml SPMC (Picoprep)		1. 2L 2. 2L
Kim HG 2015	Korea	MC SB non-inferiority RCT	Screening, surveillance, diagnostic, treatment	1. 153 2. 166	1. 53.5 2. 53.8	1. Split 300ml SPMC (Picolight) + 10mg bisacodyl. 3d LRD	2. Split 4L PEG. 3d LRD (Productname NR)	1 .4L 2. None
Kim MJ 2016	Korea	SC SB RCT	diagnostic	1. 59 2. 58 3. 57 4. 55	1. 54 2. 53 3. 53 4. 55	1. DB SPMC (Picolight) + NaP (Clicolon) + 10mg bisacodyl, 1d LRD 2. Split SPMC (Picolight) + NaP (Clicolon)+ 10mg bisacodyl, 1d LRD	3. DB SPMC (Picolight) + 1L-PEGA (Coolprep) +10mg bisacodyl, 1d LRD 4. split SPMC (Picolight) + 1L-PEGA (Coolprep) + 10mg bisacodyl, 1d LRD	1. 1L 2. 1L 3. 1.5L 4. 1.5L

Table 1. Continued

Author (year)	Country	Study design	Reason for colonoscopy	Sample size (ITT)	Age (mean, years)	Intervention (low-volume)	Comparison (high-volume)	Additional fluid intake
Kim SH 2020	Korea	SC SB RCT	Screening, surveillance, diagnostic, treatment	1. 97 2. 99 3. 99	1. 56.5 2. 54.4 3. 58.1	1. split 340ml SPMC (Picosolution) + 10mg bisacodyl, LRD 3d 2. SD 1L-PEGA (Coolprep), 10mg bisacodyl night before, LRD 3d	3. Split 2L PEGA (Coolprep), LRD 3d	1. 2L 2. 1L 3. 1L
Klare 2015	Germany	SC SB RCT	NR	1. 99 2. 101	1. 53.4 2. 56.4	1. DB 300ml SPMC (picoprep), 1d LQD	2. split 2-4L PEG (Oralav) (until clear stool). 1d LQD	1. 250ml/h 2. ≥1L
Kojecky 2017	Czech Republic	MC SB RCT	NR	973‡	NR	1. DB 300ml SPMC (Picoprep), 3d LRD 2. Split 300ml SPMC (Picoprep), 3d LRD	3. DB 2L-PEGA (Moviprep), 3d LRD 4. Split 2L PEGA (Moviprep), 3d LRD 5. 4L PEG (Fortrans) DB, 3d LRD 6. Split 3/1L PEG (Fortrans), 3d LRD	1+2. 2L 3-6. None

Table 1. Continued

Author (year)	Country	Study design	Reason for colonoscopy	Sample size (ITT)	Age (mean, years)	Intervention (low-volume)	Comparison (high-volume)	Additional fluid intake
Kojecky 2018	Czech Republic	MC SB RCT	NR	1+2. 178 3+4. 189 5+6. 181	1+2. 60.2 3+4. 60.5 5+6. 62.4	1. DB 300ml SPMC (Picoprep), LRD 3d 2. Split 300ml SPMC (Picoprep), LRD 3d	3. DB 2L PEGA (Moviprep), LRD 3d 4. Split 2L PEGA (Moviprep), LRD 3d 5. DB 4L PEG (Fortrans), LRD 3d 6. Split 4L PEG (Fortrans), LRD 3d	1+2. 2L 3+4. 1.5L 5+6. None
Munsterman 2015	The Netherlands	SC SB RCT	Surveillance, diagnostic, treatment	1. 85 2. 88	1. 57 2. 55	1. Split 300ml SPMC (Picoprep), LRD 2d, LQD 1d	2. Split 3/1L PEG (Kleanprep), LRD 2d, LQD 1d	1. 4L 2. unspecified
Muñoz-Navas 2015[§]	Spain	SC SB RCT	first-time diagnostic endoscopy	1. 224 2. 213 3. 53	1. 50.7 2. 53.6 3. 55.3	1. DB SPMC, LRD 2d prior, LQD 1d 3. SD 300ml SPMC, LRD 2d prior, LQD 1d	2. DB 4L PEG, LRD 2d prior, LQD 1d (no productnames reported)	1. 250mL/h 2. 250mL/h 3. none
Pisera 2019	Poland	MC SB RCT	screening	1. 6752 2. 6745	55-62 (full study)	1. Mixed DB/split 300ml SPMC (Citrafleet), 1d LRD, 0.5d LQD	2. Mixed DB/ split 4L PEG (Fortrans), 1d LRD, 0.5d LQD	1. 4L 2. None
Pohl 2015[§]	Germany	MC SB RCT	diagnostic, screening, surveillance	1. 193 2. 200	1. 60.1 2. 59.5	1. DB 300ml SPMC (Citrafleet), LRD 1d, 0.5d LQD	2. Split 2L PEGA (Moviprep), LRD 1d, 0.5d LQD	1. 250mL/h 2. 1L
Prieto-Frías 2016	Spain	MC stratified trial	elective outpatient	1. 157 2. 148	1. 53.7 2. 55.7	1. Split SPMC (Citrafleet), LRD 2d 2. DB 300ml SPMC (Citrafleet), LRD 2d		NR

Table 1. Continued

Author (year)	Country	Study design	Reason for colonoscopy	Sample size (ITT)	Age (mean, years)	Intervention (low-volume)	Comparison (high-volume)	Additional fluid intake
Rostom 2019	Canada	SC RCT	Screening, surveillance	1. 33 2. 28 3. 34 4. 36	1. 57.6 2. 56.4	1. DB SPMC, LRD 4d, LQD 1d 2. split SPMC, LRD 4d, LQD 1d	3. DB 2L PEG, LRD 4d, LQD 1d 4. split 2L PEG, LRD 4d, LQD 1d (No productnames reported)	NR
Sahebally 2015	Ireland	SC SB RCT	NR	1. 64 2. 66	1. 57.2 2. 59.3	1. DB 300ml SPMC (Picolax), LQD 0.5d	2. DB 2L PEGA (Moviprep), LQD 0.5d	1. 2L 2. 2L
Schulz 2016[§]	Germany	MC SB RCT	screening, surveillance, diagnostic, IBD	1. 156 2. 159	57.1 (full study)	1. DB 300ml SPMC, LRD 1d, LQD starting preparation 2. Split 300ml SPMC, LRD 1d, LQD starting preparation	No productnames reported	1+2. 250mL/h
Seo 2018	Korea	SC SB RCT	elective outpatient colonoscopy, screening, surveillance, diagnostic	1. 114 2. 109	1. 54.9 2. 56.1	1. split 300/150ml SPMC (Picolight), LRD 3d, LQD 1d	2. Split 2L PEGA (Coolprep), LRD 3d, LQD 1d.	1. 2L 2. 1L
Voiosu 2017	Romania	MC SB RCT	outpatient, first colonoscopy, screening, surveillance, diagnostic	1. 37 2. 37 3. 69 (34/35)	1. 58 2. 57 3. 54	1. Split 300ml SPMC (Picoprep), 1d LRD 3a. Individualized SPMC or PEG based on questionnaire	2. Split 4L PEG (Fortrans), 1d LRD 3b. Individualized SPMC or PEG based on questionnaire.	1. 250mL/h 2. none

Table 1. Continued

Author (year)	Country	Study design	Reason for colonoscopy	Sample size (ITT)	Age (mean, years)	Intervention (low-volume)	Comparison (high-volume)	Additional fluid intake
Yoo 2015	Korea	SC SB RCT	screening, surveillance, diagnostic	1. 100 2. 100	1. 53.3 2. 57.0	1. Split 300ml SPMC (Picolight), LRD 3d, LQD starting preparation	2. Split 2L PEGA (Coolprep), LRD 3d, LQD starting preparation	1. 2L 2. 1L
Schreiber 2019§	Germany	MC SB non-inferiority RCT	screening, surveillance, diagnostic	1. 251 2. 250	1. 52.9 2. 54.6	1. DB 300ml SPMC (Citrafleet), LQD 0.5d 2. DB 1L-PEGA (NER1006, Plenvu)), LQD 0.5d		1. 250mL/h 2. 1L
Bisschops 2019§	Belgium	MC SB non-inferiority RCT	screening, surveillance, diagnostic	1. 275 2. 275 3. 272	1. 56.3 2. 54.9 3. 54.3	1. Split 1L-PEGA (NER1006, Plenvu), 1d LRD, 0.5d LQD 2. SD 1L-PEGA (NER1006, Plenvu), 1d LRD, 0.5d LQD	3. Split 2L PEGA (Moviprep), 1d LRD, 0.5d LQD	ad libitum
DeMicco 2019§	USA	MC SB non-inferiority RCT	out/inpatient, screening, surveillance, diagnostic	1. 276 2. 280	1. 57.5 2. 56.8	1. Split 1L-PEGA (NER1006, Plenvu), 1d LRD, 0.5d LQD 2. Split 1L OSS (Suprep), 1d LRD, 0.5d LQD		1. 1L 2. 2L
Choi 2018	Korea	SC SB non-inferiority RCT	elective outpatient, screening, surveillance, diagnostic	1. 130 2. 130	1. 55.3 2. 58.5	1. split 1L PEGA (Coolprep) + 2mg prucalopride, LRD 3d prior	2. split 2L PEGA (Coolprep), LRD 3d prior	1. 1L 2. 1L
Kamei 2018	Japan	SC non-inferiority trial	outpatients, screening, surveillance, IBD	1. 68 2. 44	45.5 (full study)	1. SD 1L PEGA (Moviprep)+ DB 5mg mosapride citrate hydrate AC, 24 mg sennoside AN	2. SD 2L PEG (NIFLEC) + DB 5mg mosapride citrate hydrate AC, 24 mg sennoside AN	1. 0.5L 2. None

Table 1. Continued

Author (year)	Country	Study design	Reason for colonoscopy	Sample size (ITT)	Age (mean, years)	Intervention (low-volume)	Comparison (high-volume)	Additional fluid intake
Kang 2017	Korea	SC SB RCT	screening, surveillance, diagnostic, previous abdominal surgery	1. 100 2. 100	1. 54.4 2. 56.8	1. SD 1L-PEGA (Coolprep) + 10mg bisacodyl evening before, LRD 3d prior	2. split 2L PEGA (Coolprep), LRD 3d prior	1. 1.5L 2. 1L
Kim SH 2019	Korea	SC SB RCT	screening, surveillance, diagnostic	1. 83 2. 85	1. 52.3 2. 56.8	1. SD 1L-PEGA (Clicool) + 20mg bisacodyl, clear LQD evening prior.	2. split 2L PEGA (Clicool), clear LQD evening prior	1. 1L 2. 1L
Kwon 2016[§]	Korea	MC SB non-inferiority RCT	elective	1. 91 2. 96	1. 59.6 2. 56.0	1. SD 1L PEGA, evening prior 20mg bisacodyl. LRD 3d prior	2. split 2L PEGA, LRD 3d prior (No productnames reported)	1. 1L 2. 1L
Banerjee 2016	India	SC DB non-inferiority trial	Diagnostic, IBD, surveillance	1. 71 2. 75 3. 221 4. 221	45.8 (full study)	1. SD 1L PEG (Peglec) + 24mcg lubiprostone, clear LQD starting preparation	2. SD 1.5L PEG (Peglec) + lubiprostone 3. SD 2L PEG (Peglec) + lubiprostone 4. SD 2L PEG (Peglec)	NRs
Tian 2019	China	SC SB RCT	diagnostic, surveillance	1. 82 2. 84 3. 80	1. 52.3 2. 52.3 3. 49.0	1. split 1L-PEGA + 30ml castor oil + 20ml simethicone, LRD 3d (No productnames reported)	2. split 2L PEG + 30ml castor oil + 20ml simethicone 3. split 1-2L PEG + 30ml castor oil + 20ml simethicone	1. 2L 2. 1L 3. None

Table 1. Continued

Author (year)	Country	Study design	Reason for colonoscopy	Sample size (ITT)	Age (mean, years)	Intervention (low-volume)	Comparison (high-volume)	Additional fluid intake
Poyrazoglu 2015	Turkey	SC SB RCT	elective - screening, surveillance, diagnostic, IBD	1. 66 2. 62	1. 47.7 2. 49.8	1. split 90ml NaP solution, 3d prior LRD, 1d LQD 2. split 500ml 1000mg Sennosides A and B calcium + 66.6g sorbitol	(No productnames reported)	NR
Hung 2019	China	SC RCT	NR	1. 93 2. 97	1. 49.7 2. 48.9	1. split 90ml NaP solution (Fleet Phospho-soda), LRD 2d 0.5d LQD 2. SD 45ml NaP (Fleet Phospho-soda), 2mg prucalopride 8pm DB, LRD 2d, 0.5d LQD		1. 1.5L 2. 1.5L
Kang 2015	Taiwan	SC, non-concurrent control group	Screening	1. 259 2. 172	1. 47.7 2. 47.8	1. DB 90ml NaP (Fleet Phospho-soda) 2. Split 90ml NaP (Fleet Phospho-soda)		NR
Yang 2020[§]	Korea	MC SB non-inferiority RCT	diagnostic, screening, surveillance	1. 112 2. 112	1. 48.2 2. 46.4	1. Split 946ml OSS (Suprep), 1d LQD	2. DB 14 NaP tablets (PBK-1701TC, Pharmbio) + 2,5L water, 1d LQD	1. 2L 2. 2.5L
Lee 2019[§]	Korea	SC SB non-inferiority RCT	diagnostic, screening, surveillance, therapy.	1. 93 2. 94	60.2 (full study)	1. split 946ml OSS (Suclear), LRD 3d prior	2. split 2L PEGA (Coolprep), LRD 3d prior	1. 2L 2. 1L
Gerard 2017	USA	SC SB RCT	elective, screening, surveillance, diagnostic	1. 200 2. 200	1. 60.6 2. 60.1	1. split 946ml OSS (Suprep), 1d LQD	2. Split 2L PEGA (Moviprep), 1d LQD	1. 2L 2. 1L

Table 1. Continued

Author (year)	Country	Study design	Reason for colonoscopy	Sample size (ITT)	Age (mean, years)	Intervention (low-volume)	Comparison (high-volume)	Additional fluid intake
Thukral 2019	USA	SC SB RCT	open-access screening and surveillance	1. 108 2. 107	1. 57.0 2. 55.8	1. split 600ml MagC, LRD 2. split 600ml MagC, LQD	(No productnames reported)	1. 1L 2. 1L
Khorasanynejad 2017	Iran	SC DB RCT	Screening, surveillance	1. 151 2. 160	50.0 (full study)	1. split 180 ml C-lax syrup (containing 270mg Sennosides), LRD 1d	2. 5L PEG DB, LRD 1d prior (no productname reported)	NR

AC – ante coenam, AN – ante noctem, CRC – colorectal carcinoma, DB – Day before, DB – Double-blind, IBD – inflammatory bowel disease, ITT – intention to treat, LQD – liquid diet, LRD – low residue diet, MC – magnesium citrate, MagC – Multicenter, NaP – sodium phosphate solution, NR – Not reported, OSS – oral sulfate solution, PEGA – polyethylene glycol with ascorbate, RCT – randomized controlled trial, SB – Single-blind, SC – single-center, SD – Same day, Split – Split-dose, SPMC – sodium picosulfate magnesium citrate. ‡No subgroup information provided, §Industry-funded trial

All studies were single- or multicenter assessor-blinded RCTs. Fourteen studies (32%) were sponsored by pharmaceutical companies (35, 36, 39, 41, 42, 49, 51, 55, 58, 60, 61, 66, 69, 70). Study populations included outpatient patients with various indications for colonoscopy, i.e., screening, surveillance, and diagnostic. Exclusion criteria were commonly accepted contraindications for colonoscopy and contraindications for bowel preparation in general.

Adequately prepared patients per fluid

For SPMC, the percentage adequately cleaned patients was reported in 19 studies comprising 10,287 patients, with a pooled percentage of 75.2% (95% confidence interval [CI] 67.6-81.4; I^2=96%) (figure 2). The pooled estimate was significantly higher in studies that used additives, such as bisacodyl (n=9714), compared to studies that only used the standard dosage of 300ml SPMC (n=573) (P<0.01) (table 2, Supplementary (S) figure 1). Dosing subgroups (same-day, day before, split-dose) performed significantly different in subgroup analysis (P<0.00001), with a negative trend for day-before dosing (table 2, figure S2). Bowel preparation efficacy did not differ significantly between diet subgroups (liquid, low-residue, or combined) (table 2, figure S3).

Ten studies comprising 1717 patients reported the proportion of adequately prepared patients using 1L-PEGA, with a pooled percentage of 82.9% [95%CI 74.4-90.1; I^2=94%] (figure 3). The pooled efficacy of 7 studies (n=641) using additives was comparable to the efficacy of NER1006, a 1L-PEG solution with a higher ascorbate concentration (table 2, figure S4). In addition, subgroup analysis for dosing subgroups showed comparable efficacy of same-day or split dosing (table 2, figure S5). Subgroup analysis for diet effect did not show a significant difference in bowel cleansing efficacy (table 2, figure S6).

Two studies comprising 621 patients reported the efficacy of NaP, with a pooled percentage adequately prepared patients of 81.9% [95%CI 36.7-97.2; I^2=98%] (figure 4). For OSS, 3 studies comprising 597 patients reported on our primary endpoint, with a pooled percentage of 92.1% [95%CI 79.7-97.2, I^2=86%] (figure 5). Due to the small number of studies available, no subgroup analyses could be performed.

Furthermore, the 2 studies investigating the efficacy of sennosides did not report on preparation adequacy in a proportional manner.

Table 2. Subgroup analyses for sodium picosulfate/magnesium citrate and 1L polyethylene glycol with ascorbate.

Subgroup analysis (category)	Nº of studies	Adequately prepared (%, 95%CI)	I^2 (%)
All studies	43		
SPMC	19	75.15 (67.63-81.41)	96
1L-PEGA	10	82.94 (74.39-90.08)	94
OSS	3	92.06 (79.67-97.17)	86
NaP	2	81.91 (36.75-97.24)	98
SPMC			
With additives	5	85.77 (74.05-92.71)	96
Without additives	14	71.18 (62.37-78.63)	79
Same day	1	86.79 (53.13-97.44)	NA
Day before	6	52.71 (36.52-68.36)	91
Split-dose	9	77.96 (68.40-85.25)	93
Liquid diet	3	78.68 (59.84-90.14)	98
Low-residue diet	6	80.42 (66.45-89.50)	92
Combined	9	69.75 (57.46-79.74)	96
1L-PEGA			
With additives	7	80.04 (75.46-83.95)	91
Without additives (NER1006)	3	80.72 (76.59-84.27)	97
Same day	6	88.69 (79.48-94.07)	84
Day before	1	58.40 (22.30-87.29)	NA
Split-dose	4	80.75 (65.10-90.42)	94
Liquid diet	5	78.56 (61.44-89.39)	93
Low-residue diet	1	89.50 (69.89-96.90)	NA
Combined	4	86.82 (71.13-94.62)	94

1L-PEGA – 1L polyethylene glycol with ascorbate, CI – confidence interval, OSS – oral sulfate solution, Nº – Number, NA – not applicable, NaP – Sodium Phosphate solution, SPMC – sodium picosulfate with magnesium citrate.

Secondary endpoints

Adenoma detection rate

ADR was reported in 10 SPMC studies with a pooled ADR of 31.0% [95%CI 25.6-36.7; I^2=83%] and in 8 1L-PEGA studies with a pooled ADR of 32.4% [95%CI 26.6-38.4; I^2=83%]. ADR was reported in 1 study in the NaP group and was 30.4% [95%CI 20.6-41.2], and in 2 studies in the OSS group with a pooled ADR of 40.9% [95%CI 28.3-54.2; I^2=81%]) (figure 6A-D)

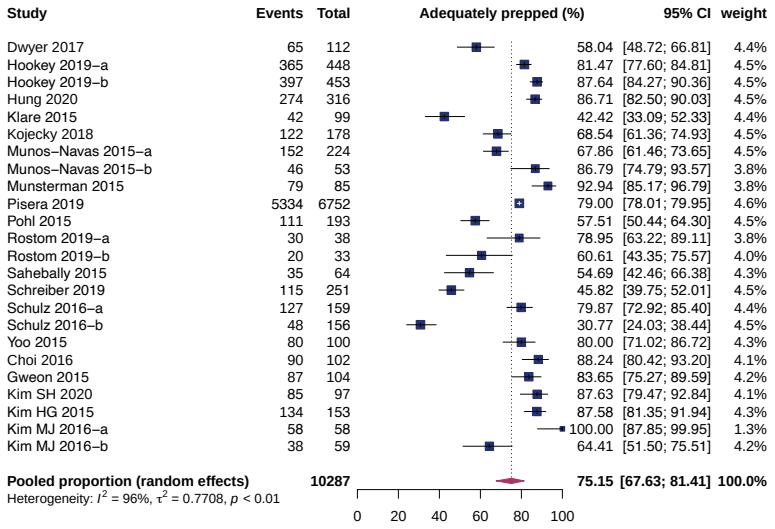


Figure 2. Pooled proportion adequately prepped patients, sodium picosulfate/magnesium citrate (SPMC). CI – confidence interval.

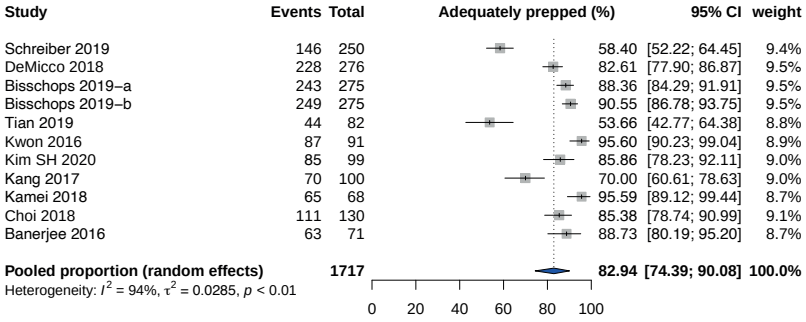


Figure 3. Pooled efficacy for 1L PEG with ascorbate (PEGA).

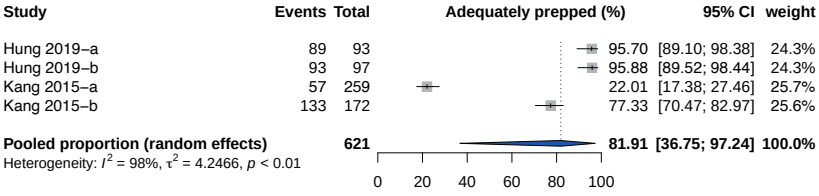


Figure 4. Pooled efficacy Sodium Phosphate solution (NaP). CI – confidence interval.

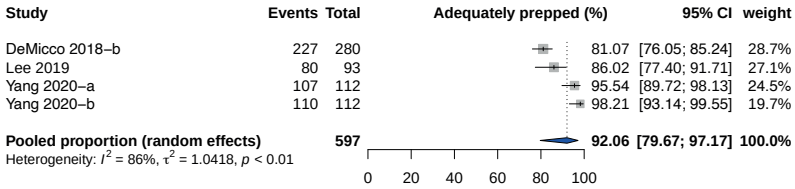


Figure 5. Pooled efficacy oral sulfate solution (OSS) CI – confidence interval.

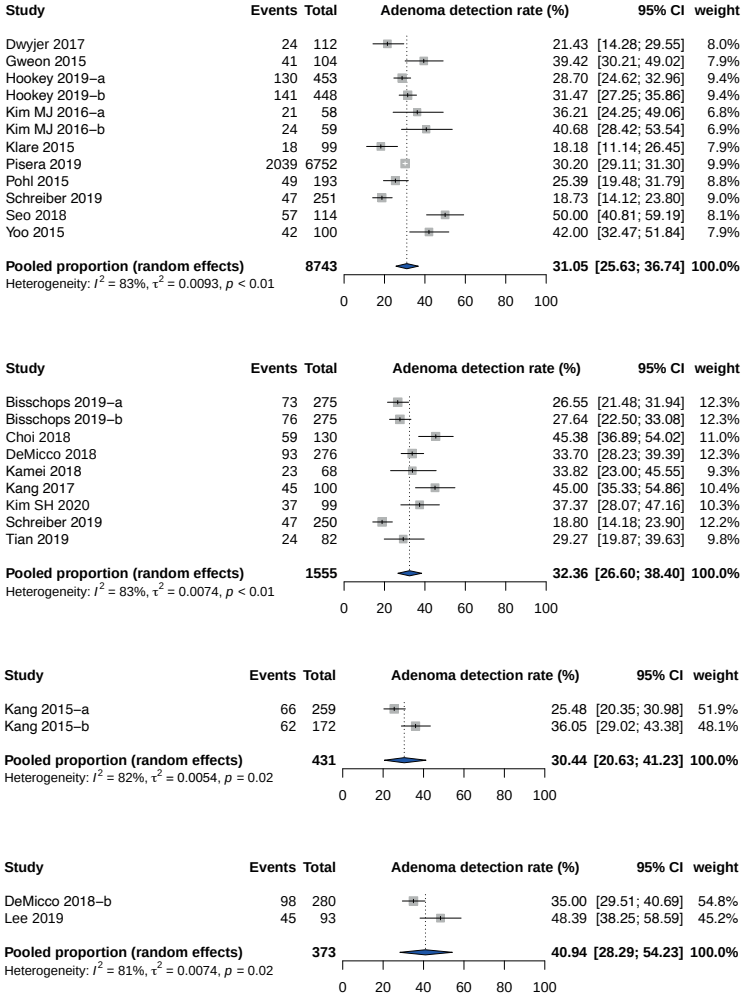


Figure 6. Forrest plots for pooled adenoma detection rate for fluid studies. A – sodium picosulfate/magnesium citrate (SPMC), B – 1L PEG with ascorbate (PEGA), C – sodium phosphate solution (NaP), D – oral sulfate solution (OSS). CI – confidence interval.

Safety

All included studies reported gastrointestinal symptoms such as abdominal pain and distention, anal irritation, nausea, and to a lesser extent vomiting as most frequent adverse events (AEs). Furthermore, headache, dizziness, and general malaise were reported with the use of all fluids.

Of the 26 SPMC studies, 22 reported on AEs, and 6 evaluated laboratory abnormalities. A range of 8.1-85.6% of the patients experienced at least some of the AEs as mentioned above. Moreover, 3 studies (37, 39, 42) reported elevated serum magnesium in 3.6-10.5% of the patients, and decreased serum sodium levels in up to 21.2% of patients (45) with one report of severe hyponatremia (119mmol/L)(55). Schulz et al. reported hyperkalemia (6.2mmol/L and 8.5mmol/L, respectively) in 2 young patients, which resolved without sequelae.

In the 1L-PEGA group, AEs were reported in 11 of 12 studies, occurring in 13.2-43.4% of patients. In the NER1006 studies (58, 60, 61), higher rates of temporary decrease in renal function and hypernatremia (median +4.0 mmol/L from baseline) were reported, compared to other ultra-low volume fluids included in these review.

For NaP, AEs were reported in 2 of the 4 included studies, occurring in 44.3-72.2% of the patients. The type of reported AEs was similar to those reported above. Additionally, an increase in serum inorganic phosphorus levels (from a median of 3.5mEq/L at baseline to 6.3mEq/L at the day of colonoscopy) was noted (72).

AEs were reported in all OSS studies and occurred over a range of 18.5-77.4% of the patients. Laboratory abnormalities included a temporary decrease in renal function (61). Other electrolyte changes were not considered as clinically significant (61, 69).

Heterogeneity

Heterogeneity was considerable for all bowel preparation fluid studies and could only partially be explained by the prespecified subgroups (additive-use and differences in dosing schedule). In subgroup analyses based on type of bowel preparation scale, I^2 remained substantial (BBPS I^2 91%, OBPS I^2 96%, data not shown). No changes in the pooled effect size nor in the extent of heterogeneity were observed in sensitivity analysis (figure S7). Influence analysis, including Baujat- and GOSH plots, identified possible outliers (55, 58, 68, 74) (figure S7-9). Excluding these outliers in the meta-analyses did however not change the pooled effect sizes significantly but reduced the CI. For SPMC, the CI changed from 67.6-81.4 to 73.2-76.0. For 1L-PEGA, after removing the outliers(58, 68), the pooled percentage

changed non-significantly from 82.9% [95%CI 74.4-90.1; I^2=94%] to 77.0% [95%CI 75.7-78.1; I^2=94%].

Risk of bias

Funnel plots showed no evidence for publication bias (figures S11A-D). The overall risk of bias was low in 58.1%, intermediate in 23.3%, and high in 16.3% of the included studies (figure 7, figure S12). The pooled outcome did not change significantly for any of the fluids when excluding the studies classified as high risk of bias (36, 46, 52, 63, 67, 72, 74), but for the 1L-PEGA group, a drop from 83.0% [95%CI 74.4-90.1] to 75.3% [95%CI 73.0-77.3] was found.

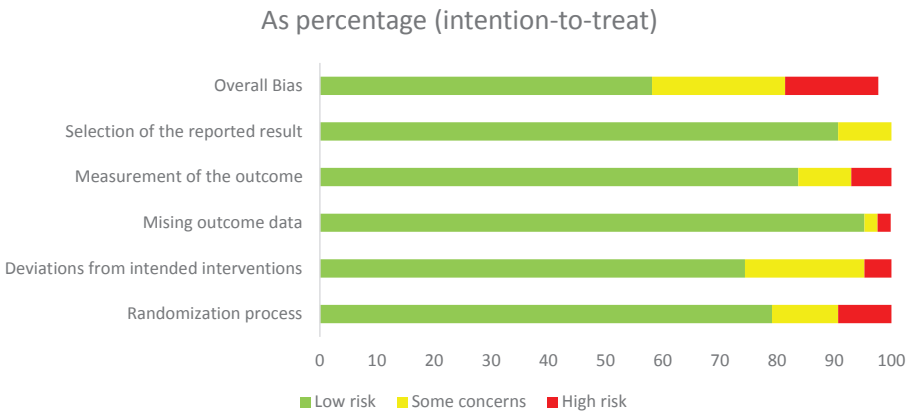


Figure 7. Summary of risk of bias (Cochrane RoB2 tool*). *The Cochrane RoB2 tool assesses the risk of bias across five domains, including randomization process, protocol deviations, missing data, outcome measurement, and selection of the reported result. The overall risk of bias is determined by the highest risk within the subdomains. This figure summarizes the risk of bias within all included studies, as percentage of the total number of studies.

Discussion

This systematic review and meta-analysis show that ultra-low volume (≤1L) bowel preparation with SPMC, 1L-PEGA, NaP, or OSS, was adequate in 75.2%, 82.9%, 81.9%, and 92.1%, of patients, respectively. While ESGE guidelines on bowel preparation and CRC screening recommend an adequate bowel cleansing rate in at least 90% of procedures (1), the majority of these ultra-low volume bowel preparation fluids, with the exception of OSS, do not meet this quality standard as defined by the ESGE in our analysis. It should be noted that only a low number of studies investigating NaP (n=4) and OSS (n=4) were included, which likely mirrors their limited use in

daily clinical practice. The preference for SPMC and 1-PEGA in various international guidelines is motivated by potential side-effects that have been reported with use of NaP and OSS (1, 2, 8).

Evidence on the efficacy of low-volume fluids is contradictory (77). Some meta-analyses comparing high-volume (>3L) PEG with lower-volume fluids (≤2L) have demonstrated a lower efficacy of low-volume fluids (78, 79), whereas others have suggested non-inferiority when comparing these two different volume fluids (14, 80-82). An explanation of the suboptimal efficacy results in our meta-analysis may well be that we limited our analysis to ≤1L (ultra-low volume) fluids, while others classify a volume of ≤2L already as low-volume. High- or intermediate volume laxatives such as 4L-PEG or 2L-PEGA have a well-established efficacy profile (1). Therefore, when 2L preps are included in the same group as the ultra-low-volume fluids, this may improve the overall efficacy results of the low-volume group and give the wrong impression that all low-volume fluids are equally effective or non-inferior to the high-volume counterpart. Another explanation may be that day-before dosing generally performed worse than split-dose protocols in our analysis. This has also been reported in other meta-analyses (78, 80) and might at least partly explain why the pooled efficacy we found is lower than expected as in the included studies both split-dose and day-before dosing was used. It is recommended that a colonoscopy procedure should take place within 2-5 hours after finishing bowel preparation to make sure that the colon is most optimally prepped (1), thereby reducing the risk that neoplastic lesions will be missed due to bowel contamination (83, 84). Furthermore, it is questionable whether split-dosing is feasible with ultra-low volume fluids, especially in isotonic fluids, or that same-day dosing should be standard. Split-dosing may lower the purging effect of the first dose, thereby reducing the final laxative effect of the hypertonic second dose. In our dosing-stratified analyses, the pooled efficacy of same-day and split-dosing were close to the recommended 90%. Further studies on ultra-low volume fluids should focus on the efficacy of dosing on the day of colonoscopy, as in this way ultra-low volume bowel preparation fluids might still be a viable option for bowel cleansing.

The ultra-low volume laxatives presented here may offer a solution for patients having difficulties with drinking high volumes. Additionally, the optimized patient perception as compared to high-volume fluids likely will increase the willingness of patients to repeat colonoscopy and decreases the number of patients avoiding colonoscopy (11, 80, 81). Spadaccini et al. performed a meta-analysis including 17 RCTs (n=7582) in which they showed that the compliance rate, tolerability, and willingness to repeat taking the same preparation were all in favor of low-volume

(<2L) preparations (80). Nevertheless, lowering the volume of bowel preparation fluids does not release patients from drinking large volumes. The stimulant and hyperosmotic pharmacologic mechanism of action that draws water into the gut lumen makes taking extra fluids necessary in addition to the laxatives. Thus, patients should be instructed to maintain hydration to compensate for the large fecal effluent of 2.5L-3L (85). For the 300ml prep SPMC for example, this means drinking at least 2 liters of fluid, which is the same as is recommended in addition to 2L-PEGA.

High-volume bowel preparation fluids such as 4L-PEG may anyhow be preferable in patients with a high risk of dehydration-related complications, such as acute kidney injury, or fluid shifts. The iso-osmotic nature of the PEG electrolyte solution minimizes fluid shifts and thereby reduces the risk of electrolyte disturbances. These electrolyte disturbances such as transient hypermagnesemia for SPMC, hypo- or hypernatremia for 1L-PEGA, and hyperphosphatemia for NaP have been reported in this systematic review, as well as in other publications(86). No serious AEs were reported in our included studies, but there are case reports that report on fatal hyponatremia and hypermagnesemia (87, 88). The rare risk of acute phosphate nephropathy caused by tubular calcium depositions due to NaP use has resulted in a warning by the United States Food and Drug Administration to consider alternative bowel preparations instead of NaP (89). Therefore, hyperosmotic ≤1L laxatives may be less suitable for elderly or patients with renal dysfunction (1, 3, 6).

Nonetheless, it is debatable whether the above-mentioned electrolyte changes are clinically relevant for the majority of patients (90, 91). In a retrospective study of 2.8 million participants, 30- and 90-day hospitalizations for electrolyte changes were <0.1% in patients that used several low-volume bowel preparations, which was not significantly different from patients using high-volume alternatives(92). However, the severe AEs that occur only rarely, are often reported in post-marketing surveillance data in case-series or retrospective studies. For some of the more recently developed bowel preparation fluids, such as NER1006, this data is still limited. Additionally, study populations of included studies in this meta-analysis mostly comprised healthy adult patients, excluding patients at risk for adverse events or unable to drink larger volumes. This should be taken into account when deciding on the most suitable laxative for a particular patient.

Bowel preparation quality can be improved in several ways. On the one hand, lowering the volume of bowel preparation fluids may reduce non-compliance rates in patients (80). On the other hand, diet restrictions may influence the

experienced burden of bowel preparation significantly. Compared to a clear liquid diet (CLD), low residue diets (LRD) are better tolerated (93-95). Two meta-analyses compared LRD to CLD in studies with similar bowel preparation solutions in both arms and found an equal bowel cleansing efficacy but better tolerability with a higher willingness to repeat for LRD (94, 95). Furthermore, too many rules and restrictions for patients can be overwhelming and may undermine understanding the importance of adequate cleansing (96). Enhanced patient education has been shown to improve colonoscopy preparation (97), for example using visual aids or mobile apps in addition to regular counseling (97-99). Two meta-analyses concluded that enhanced instructions benefit bowel preparation quality and ADR (100, 101), although another meta-analysis acknowledging these benefits, pointed to a possible risk of publication bias (102).

An interesting and possibly useful development may be the use of bowel cleansing devices. Using mechanical bowel cleansing before or during colonoscopy is proposed as an alternative to oral laxatives in selected patient groups (103-107). Preprocedural devices work through retrograde bowel lavage using pressurized water one hour before colonoscopy (108-110), while intraprocedural devices can be used during colonoscopy providing water-pressured cleansing (107, 111-113). Feasibility studies have shown a clear potential, with adequate bowel preparation achieved in 97.9-100% and 68.8-91.1% of patients in whom intraprocedural devices (107, 111-113) or preprocedural devices (108-110), respectively, were used. Nonetheless, the use of intraprocedural devices adds significantly to the total procedure time, and preprocedural devices require a specialized nurse to operate the system. The associated costs may prohibit ubiquitous use, but the application of these bowel cleansing devices could be of interest in patients with risk factors for inadequate preparation, in whom a repeat endoscopic procedure often is indicated. Additionally, these devices could reduce admission time for inpatients.

Strengths and limitations

The strengths of our meta-analysis are the large number of patients included (n=13,222) and the robustness of the results in the extensive sensitivity analysis. Moreover, we only included studies in which bowel preparation fluids with a volume of ≤1L were included. Currently, 2L-PEGA is widely prescribed and recommended(1), but some patients still have difficulty with this volume. This makes ≤1L fluids a welcome innovation.

The large heterogeneity in our meta-analysis inevitably limits interpretation of the results. This is illustrated by the reduced efficacy of 1L-PEGA after removing outliers

and studies with a high risk of bias. Pooling the proportion of adequately prepped patients might have introduced heterogeneity in our results, besides the existing heterogeneity due to different study locations (Asia, Europe), dietary instructions, dosing regimens, and use of additives. Through subgroup- and extensive sensitivity analyses, the influence of this heterogeneity could be minimized, and this further endorsed the robustness of our results. Although we could not take into account individual patients' risk factors for poor bowel preparation, such as high age, BMI, history of poor preparation, constipation, or history of neurological disorders (9, 17, 114), the RCTs in this meta-analysis frequently did not include patients who, for example, had serious systemic illnesses or used tricyclic antidepressants. While the use of different bowel preparation scales across studies is a drawback for performing a meta-analysis, our approach is not different from other published meta-analyses (80, 115). The trend that a large proportion of published studies are non-inferiority trials and underpowered to detect superiority, and the great variety of comparative arms led us to only pool the efficacy of the ultra-low volume fluids without the comparative high-volume arms of the included studies. This enabled us to select more studies, giving a more precise direction to the pooled effect.

Conclusion

Large scale use of ultra-low volume bowel preparation is limited by an overall efficacy of these ≤1L fluids below the 90% ESGE quality target. Therefore, their use might mainly be considered in selected patient populations with no risk factors for dehydration-related complications or inadequate preparation, as well as for patients having difficulty drinking large volumes.

Acknowledgments

We thank dr. O.Y. Chan for assistance in performing the systematic search and dr. R. Akkermans for statistical advice.

Conflicts of interest

MLM van Riswijk and KE van Keulen have nothing to declare. PD Siersema receives unrestricted grants from Pentax (Japan), Norgine (UK), Motus GI (USA), MicroTech (China) and The eNose Company (Netherlands) and is in the advisory board of Motus GI (USA) and Boston Scientific (USA).

References

1. Hassan C, East J, Radaelli F, Spada C, Benamouzig R, Bisschops R, et al. Bowel preparation for colonoscopy: European Society of Gastrointestinal Endoscopy (ESGE) Guideline - Update 2019. Endoscopy. 2019;51(8):775-94.
2. Committee ASoP, Saltzman JR, Cash BD, Pasha SF, Early DS, Muthusamy VR, et al. Bowel preparation before colonoscopy. Gastrointest Endosc. 2015;81(4):781-94.
3. Johnson DA, Barkun AN, Cohen LB, Dominitz JA, Kaltenbach T, Martel M, et al. Optimizing adequacy of bowel cleansing for colonoscopy: recommendations from the US Multi-Society Task Force on Colorectal Cancer. Am J Gastroenterol. 2014;109(10):1528-45.
4. Harewood GC, Sharma VK, de Garmo P. Impact of colonoscopy preparation quality on detection of suspected colonic neoplasia. Gastrointest Endosc. 2003;58(1):76-9.
5. Froehlich F, Wietlisbach V, Gonvers JJ, Burnand B, Vader JP. Impact of colonic cleansing on quality and diagnostic yield of colonoscopy: the European Panel of Appropriateness of Gastrointestinal Endoscopy European multicenter study. Gastrointest Endosc. 2005;61(3):378-84.
6. Rex DK, Imperiale TF, Latinovich DR, Bratcher LL. Impact of bowel preparation on efficiency and cost of colonoscopy. Am J Gastroenterol. 2002;97(7):1696-700.
7. Kingsley J, Karanth S, Revere FL, Agrawal D. Cost Effectiveness of Screening Colonoscopy Depends on Adequate Bowel Preparation Rates - A Modeling Study. PLoS One. 2016;11(12):e0167452.
8. Hassan C, Bretthauer M, Kaminski MF, Polkowski M, Rembacken B, Saunders B, et al. Bowel preparation for colonoscopy: European Society of Gastrointestinal Endoscopy (ESGE) guideline. Endoscopy. 2013;45(2):142-50.
9. Dik VK, Moons LM, Huyuk M, van der Schaar P, de Vos Tot Nederveen Cappel WH, Ter Borg PC, et al. Predicting inadequate bowel preparation for colonoscopy in participants receiving split-dose bowel preparation: development and validation of a prediction score. Gastrointest Endosc. 2015;81(3):665-72.
10. Gandhi K, Tofani C, Sokach C, Patel D, Kastenberg D, Daskalakis C. Patient Characteristics Associated With Quality of Colonoscopy Preparation: A Systematic Review and Meta-analysis. Clin Gastroenterol Hepatol. 2018;16(3):357-69 e10.
11. Sharara AI, El Reda ZD, Harb AH, Abou Fadel CG, Sarkis FS, Chalhoub JM, et al. The burden of bowel preparations in patients undergoing elective colonoscopy. United European Gastroenterol J. 2016;4(2):314-8.
12. McLachlan SA, Clements A, Austoker J. Patients' experiences and reported barriers to colonoscopy in the screening context--a systematic review of the literature. Patient Educ Couns. 2012;86(2):137-46.
13. Fordtran JS, Hofmann AF. Seventy Years of Polyethylene Glycols in Gastroenterology: The Journey of PEG 4000 and 3350 From Nonabsorbable Marker to Colonoscopy Preparation to Osmotic Laxative. Gastroenterology. 2017;152(4):675-80.
14. Xie Q, Chen L, Zhao F, Zhou X, Huang P, Zhang L, et al. A meta-analysis of randomized controlled trials of low-volume polyethylene glycol plus ascorbic acid versus standard-volume polyethylene glycol solution as bowel preparations for colonoscopy. PLoS One. 2014;9(6):e99092.
15. Jin Z, Lu Y, Zhou Y, Gong B. Systematic review and meta-analysis: sodium picosulfate/magnesium citrate vs. polyethylene glycol for colonoscopy preparation. Eur J Clin Pharmacol. 2016;72(5):523-32.

16. Seo SI, Kang JG, Kim HS, Jang MK, Kim HY, Shin WG. Efficacy and tolerability of 2-L polyethylene glycol with ascorbic acid versus sodium picosulfate with magnesium citrate: a randomized controlled trial. Int J Colorectal Dis. 2018;33(5):541-8.
17. Hassan C, Fuccio L, Bruno M, Pagano N, Spada C, Carrara S, et al. A predictive model identifies patients most likely to have inadequate bowel preparation for colonoscopy. Clin Gastroenterol Hepatol. 2012;10(5):501-6.
18. Waldmann E, Penz D, Majcher B, Zagata J, Sinkovec H, Heinze G, et al. Impact of high-volume, intermediate-volume and low-volume bowel preparation on colonoscopy quality and patient satisfaction: An observational study. United European Gastroenterol J. 2019;7(1):114-24.
19. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210.
20. Kastenberg D, Bertiger G, Brogadir S. Bowel preparation quality scales for colonoscopy. World J Gastroenterol. 2018;24(26):2833-43.
21. Rostom A, Jolicoeur E. Validation of a new scale for the assessment of bowel preparation quality. Gastrointest Endosc. 2004;59(4):482-6.
22. Lai EJ, Calderwood AH, Doros G, Fix OK, Jacobson BC. The Boston bowel preparation scale: a valid and reliable instrument for colonoscopy-oriented research. Gastrointest Endosc. 2009;69(3 Pt 2):620-5.
23. Veroniki AA, Jackson D, Viechtbauer W, Bender R, Bowden J, Knapp G, et al. Methods to estimate the between-study variance and its uncertainty in meta-analysis. Res Synth Methods. 2016;7(1):55-79.
24. Freeman MF, Tukey JW. Transformations related to the angular and the square root. The Annals of Mathematical Statistics. 1950:607-11.
25. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539-58.
26. Viechtbauer W, Cheung MW. Outlier and influence diagnostics for meta-analysis. Res Synth Methods. 2010;1(2):112-25.
27. Baujat B, Mahe C, Pignon JP, Hill C. A graphical method for exploring heterogeneity in meta-analyses: application to a meta-analysis of 65 trials. Stat Med. 2002;21(18):2641-52.
28. Olkin I, Dahabreh IJ, Trikalinos TA. GOSH - a graphical display of study heterogeneity. Research Synthesis Methods. 2012;3(3):214-23.
29. Team RC, DC R. A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2012 2019 [22-06-2020].
30. Balduzzi S, Rucker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evidence-Based Mental Health. 2019;22(4):153-60.
31. Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. Journal of Statistical Software. 2010;36(3):1-48.
32. Harrer M, Cuijpers P, Furukawa T, Ebert D. dmetar: Companion R Package For The Guide "Doing Meta-Analysis in R". R package version 0.0. 9000. 2019.
33. Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898.
34. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009;339:b2535.

35. Choi HS, Chung JW, Lee JW, Lim MY, Park DK, Kim YJ, et al. Polyethylene glycol plus ascorbic acid is as effective as sodium picosulfate with magnesium citrate for bowel preparation: A randomized trial. J Dig Dis. 2016;17(4):268-73.

36. Dwyer JP, Tan JYC, Paul E, Bunn C, Mangira D, Secomb R, et al. White Diet with split-dose Picosalax is preferred, better tolerated, and non-inferior to day-before clear fluids with polyethylene glycol plus sodium picosulfate-magnesium citrate for morning colonoscopy: A randomized, non-inferiority trial. JGH Open. 2017;1(1):38-43.

37. Gweon TG, Kim SW, Noh YS, Hwang S, Kim NY, Lee Y, et al. Prospective, randomized comparison of same-day dose of 2 different bowel cleanser for afternoon colonoscopy: picosulfate, magnesium oxide, and citric acid versus polyethylene glycol. Medicine (Baltimore). 2015;94(13):e628.

38. Heetun Z, Crowley R, Zeb F, Kearns D, Brennan MH, O'Connor C, et al. Comparison of polyethylene glycol vs sodium picosulphate vs sodium biphosphonate by efficacy in bowel cleansing and patients' tolerability: a randomised trial. Ir J Med Sci. 2016;185(3):629-33.

39. Hookey L, Bertiger G, Lee Johnson K, 2nd, Ayala J, Seifu Y, Brogadir SP. Efficacy and safety of a ready-to-drink bowel preparation for colonoscopy: a randomized, controlled, non-inferiority trial. Therap Adv Gastroenterol. 2019;12:1756284819851510.

40. Hung SY, Chen HC, Chen WT. A Randomized Trial Comparing the Bowel Cleansing Efficacy of Sodium Picosulfate/Magnesium Citrate and Polyethylene Glycol/Bisacodyl (The Bowklean Study). Sci Rep. 2020;10(1):5604.

41. Jun JH, Han KH, Park JK, Seo HI, Kim YD, Lee SJ, et al. Randomized clinical trial comparing fixed-time split dosing and split dosing of oral Picosulfate regimen for bowel preparation. World J Gastroenterol. 2017;23(32):5986-93.

42. Kiesslich R, Schubert S, Mross M, Klugmann T, Klemt-Kropp M, Behnken I, et al. Efficacy and safety of PICOPREP tailored dosing compared with PICOPREP day-before dosing for colon cleansing: a multi-centric randomised study. Endosc Int Open. 2017;5(4):E282-E90.

43. Kim HG, Huh KC, Koo HS, Kim SE, Kim JO, Kim TI, et al. Sodium Picosulfate with Magnesium Citrate (SPMC) Plus Laxative Is a Good Alternative to Conventional Large Volume Polyethylene Glycol in Bowel Preparation: A Multicenter Randomized Single-Blinded Trial. Gut Liver. 2015;9(4):494-501.

44. Kim MJ, Hong CW, Kim BC, Park SC, Han KS, Joo J, et al. Phase II Randomized Controlled Trial of Combined Oral laxatives Medication for BOwel PREParation (COMBO-PREP study). Medicine (Baltimore). 2016;95(7):e2824.

45. Klare P, Poloschek A, Walter B, Rondak IC, Attal S, Weber A, et al. Single-day sodium picosulfate and magnesium citrate versus split-dose polyethylene glycol for bowel cleansing prior to colonoscopy: A prospective randomized endoscopist-blinded trial. J Gastroenterol Hepatol. 2015;30(11):1627-34.

46. Kojecky V, Matous J, Keil R, Dastych M, Kroupa R, Zadorova Z, et al. A head-to-head comparison of 4-L polyethylene glycol and low-volume solutions before colonoscopy: which is the best? A multicentre, randomized trial. Int J Colorectal Dis. 2017;32(12):1763-6.

47. Kojecky V, Matous J, Keil R, Dastych M, Zadorova Z, Varga M, et al. The optimal bowel preparation intervals before colonoscopy: A randomized study comparing polyethylene glycol and low-volume solutions. Dig Liver Dis. 2018;50(3):271-6.

48. Munsterman ID, Cleeren E, van der Ploeg T, Brohet R, van der Hulst R. 'Pico-Bello-Klean study': effectiveness and patient tolerability of bowel preparation agents sodium picosulphate-magnesium citrate and polyethylene glycol before colonoscopy. A single-blinded randomized trial. Eur J Gastroenterol Hepatol. 2015;27(1):29-38.

49. Munoz-Navas M, Calleja JL, Payeras G, Hervas AJ, Abreu LE, Orive V, et al. A randomized trial to compare the efficacy and tolerability of sodium picosulfate-magnesium citrate solution vs.

4 L polyethylene glycol solution as a bowel preparation for colonoscopy. Int J Colorectal Dis. 2015;30(10):1407-16.

50. Pisera M, Franczyk R, Wieszczy P, Polkowski M, Rupinski M, Chaber-Ciopinska A, et al. The impact of low- versus standard-volume bowel preparation on participation in primary screening colonoscopy: a randomized health services study. Endoscopy. 2019;51(3):227-36.
51. Pohl J, Halphen M, Kloess HR, Fischbach W. Impact of the quality of bowel cleansing on the efficacy of colonic cancer screening: a prospective, randomized, blinded study. PLoS One. 2015;10(5):e0126067.
52. Prieto-Frias C, Munoz-Navas M, Betes MT, Angos R, De la Riva S, Carretero C, et al. Split-dose sodium picosulfate-magnesium citrate colonoscopy preparation achieves lower residual gastric volume with higher cleansing effectiveness than a previous-day regimen. Gastrointest Endosc. 2016;83(3):566-73.
53. Rostom A, Dube C, Bishay K, Antonova L, Heitman SJ, Hilsden R. A randomized clinical prospective trial comparing split-dose picosulfate/ magnesium citrate and polyethylene glycol for colonoscopy preparation. PLoS One. 2019;14(3):e0211136.
54. Sahebally SM, Burke JP, Chu S, Mabadeje O, Geoghegan J. A randomized controlled trial comparing polyethylene glycol + ascorbic acid with sodium picosulphate + magnesium citrate solution for bowel cleansing prior to colonoscopy. Ir J Med Sci. 2015;184(4):819-23.
55. Schulz C, Muller J, Sauter J, Miehlke S, Schmocker C, Hartmann D, et al. Superiority of a Split-dose Regimen of Sodium Picosulfate/Magnesium Citrate (SPMC) in Comparison to a Prior-day Schedule (AM/PM) for Colonoscopy Preparation. A Randomized Single-blinded Study. J Gastrointestin Liver Dis. 2016;25(3):295-302.
56. Voiosu T, Tantau A, Voiosu A, Bengus A, Mocanu C, Smarandache B, et al. Preparation regimen is more important than patient-related factors: a randomized trial comparing a standard bowel preparation before colonoscopy with an individualized approach. Rom J Intern Med. 2017;55(1):36-43.
57. Yoo IK, Lee JS, Chun HJ, Jeen YT, Keum B, Kim ES, et al. A randomized, prospective trial on efficacy and tolerability of low-volume bowel preparation methods for colonoscopy. Dig Liver Dis. 2015;47(2):131-7.
58. Schreiber S, Baumgart DC, Drenth JPH, Filip RS, Clayton LB, Hylands K, et al. Colon cleansing efficacy and safety with 1 L NER1006 versus sodium picosulfate with magnesium citrate: a randomized phase 3 trial. Endoscopy. 2019;51(1):73-84.
59. Kim SH, Kim JH, Keum B, Jeon HJ, Jang SH, Choi SJ, et al. A Randomized, Endoscopist-Blinded, Prospective Trial to Compare the Efficacy and Patient Tolerability between Bowel Preparation Protocols Using Sodium Picosulfate Magnesium Citrate and Polyethylene-Glycol (1 L and 2 L) for Colonoscopy. Gastroenterol Res Pract. 2020;2020:9548171.
60. Bisschops R, Manning J, Clayton LB, Ng Kwet Shing R, Alvarez-Gonzalez M, Group MS. Colon cleansing efficacy and safety with 1 L NER1006 versus 2 L polyethylene glycol + ascorbate: a randomized phase 3 trial. Endoscopy. 2019;51(1):60-72.
61. DeMicco MP, Clayton LB, Pilot J, Epstein MS, Group NS. Novel 1 L polyethylene glycol-based bowel preparation NER1006 for overall and right-sided colon cleansing: a randomized controlled phase 3 trial versus trisulfate. Gastrointest Endosc. 2018;87(3):677-87 e3.
62. Choi SJ, Kim ES, Choi BK, Min G, Kim W, Lee JM, et al. A randomized controlled trial comparing the efficacy of 1-L polyethylene glycol solution with ascorbic acid plus prucalopride versus 2-L polyethylene glycol solution with ascorbic acid for bowel preparation. Scand J Gastroenterol. 2018;53(12):1619-24.

63. Kamei M, Shibuya T, Takahashi M, Makino M, Haga K, Nomura O, et al. Efficacy and Acceptability of 1 Liter of Polyethylene Glycol with Ascorbic Acid vs. 2 Liters of Polyethylene Glycol Plus Mosapride and Sennoside for Colonoscopy Preparation. Med Sci Monit. 2018;24:523-30.

64. Kang SH, Jeen YT, Lee JH, Yoo IK, Lee JM, Kim SH, et al. Comparison of a split-dose bowel preparation with 2 liters of polyethylene glycol plus ascorbic acid and 1 liter of polyethylene glycol plus ascorbic acid and bisacodyl before colonoscopy. Gastrointest Endosc. 2017;86(2):343-8.

65. Kim SH, Kim ER, Kim K, Kim TJ, Hong SN, Chang DK, et al. Combination of bisacodyl suppository and 1 L polyethylene glycol plus ascorbic acid is a non-inferior and comfortable regimen compared to 2 L polyethylene glycol plus ascorbic acid. Dig Endosc. 2020;32(4):600-7.

66. Kwon JE, Lee JW, Im JP, Kim JW, Kim SH, Koh SJ, et al. Comparable Efficacy of a 1-L PEG and Ascorbic Acid Solution Administered with Bisacodyl versus a 2-L PEG and Ascorbic Acid Solution for Colonoscopy Preparation: A Prospective, Randomized and Investigator-Blinded Trial. PLoS One. 2016;11(9):e0162051.

67. Banerjee R, Chaudhari H, Shah N, Saravanan A, Tandan M, Reddy DN. Addition of Lubiprostone to polyethylene glycol(PEG) enhances the quality & efficacy of colonoscopy preparation: a randomized, double-blind, placebo controlled trial. BMC Gastroenterol. 2016;16(1):133.

68. Tian X, Shi B, Liu XL, Chen H, Chen WQ. A Randomized Trial of Split Dose 3 L Polyethylene Glycol Lavage Solution, 2 L Polyethylene Glycol Lavage Combined With Castor Oil, and 1 L of Polyethylene Glycol Lavage Solution Combined With Castor Oil and Ascorbic Acid for Preparation for Colonoscopy. Front Med (Lausanne). 2019;6:158.

69. Yang HJ, Park DI, Park SK, Lee CK, Kim HJ, Oh SJ, et al. Novel sulfate tablet PBK-1701TC versus oral sulfate solution for colon cleansing: A randomized phase 3 trial. J Gastroenterol Hepatol. 2020;35(1):29-36.

70. Lee HH, Lim CH, Kim JS, Cho YK, Lee BI, Cho YS, et al. Comparison between an Oral Sulfate Solution and a 2 L of Polyethylene Glycol+Ascorbic Acid as a Split Dose Bowel Preparation for Colonoscopy. Journal of clinical gastroenterology. 2018.

71. Gerard DP, Foster DB, Holden JL, Kang J, Raiser MW. Clinical Trial of 1000 Subjects Randomized to 5 Low-Volume Bowel Preparations for Colonoscopy and Their Acceptance of Split-Dose Bowel Preparations. J Clin Gastroenterol. 2017;51(6):512-21.

72. Poyrazoglu OK, Yalniz M. Two low-dose bowel-cleansing regimens: efficacy and safety of senna and sodium phosphorus solution for colonoscopy. Patient Prefer Adherence. 2015;9:1325-31.

73. Hung JS, Yi CH, Liu TT, Lei WY, Wong MW, Chen CL. Does prucalopride reduce the dose of sodium phosphate in bowel preparation? A single-blind, randomized, and prospective study. J Dig Dis. 2019;20(6):294-300.

74. Kang JW, Chuang CH, Chen CY, Cheng HC, Chang WL, Chen WY, et al. Different schedules of bowel preparation with sodium phosphate lead to different bowel cleansing effects and adenoma detection rates at colonoscopy. Advances in Digestive Medicine. 2015;2(1):12-7.

75. Khorasanynejad R, Norouzi A, Roshandel G, Besharat S. Bowel Preparation for a Better Colonoscopy Using Polyethylene Glycol or C-lax: A Double Blind Randomized Clinical Trial. Middle East J Dig Dis. 2017;9(4):212-7.

76. Thukral C, Tewani SK, Lake AJ, Shiels AJ, Geissler K, Popejoy S, et al. Results of a Community-based, Randomized Study Comparing a Clear Liquid Diet With a Low-residue Diet Using a Magnesium Citrate Preparation for Screening and Surveillance Colonoscopies. J Clin Gastroenterol. 2019;53(1):34-9.

77. Millien VO, Mansour NM. Bowel Preparation for Colonoscopy in 2020: A Look at the Past, Present, and Future. Curr Gastroenterol Rep. 2020;22(6):28.

78. Martel M, Barkun AN, Menard C, Restellini S, Kherad O, Vanasse A. Split-Dose Preparations Are Superior to Day-Before Bowel Cleansing Regimens: A Meta-analysis. Gastroenterology. 2015;149(1):79-88.

79. Enestvedt BK, Tofani C, Laine LA, Tierney A, Fennerty MB. 4-Liter split-dose polyethylene glycol is superior to other bowel preparations, based on systematic review and meta-analysis. Clin Gastroenterol Hepatol. 2012;10(11):1225-31.

80. Spadaccini M, Frazzoni L, Vanella G, East J, Radaelli F, Spada C, et al. Efficacy and Tolerability of High- vs Low-Volume Split-Dose Bowel Cleansing Regimens for Colonoscopy: A Systematic Review and Meta-analysis. Clin Gastroenterol Hepatol. 2020;18(7):1454-65 e14.

81. Rocha RSP, Ribeiro IB, de Moura DTH, Bernardo WM, Minata MK, Morita FHA, et al. Sodium picosulphate or polyethylene glycol before elective colonoscopy in outpatients? A systematic review and meta-analysis. World J Gastrointest Endosc. 2018;10(12):422-41.

82. Yi LJ, Tian X, Shi B, Chen H, Liu XL, Pi YP, et al. Low-Volume Polyethylene Glycol Improved Patient Attendance in Bowel Preparation Before Colonoscopy: A Meta-Analysis With Trial Sequential Analysis. Front Med (Lausanne). 2019;6:92.

83. Seo EH, Kim TO, Park MJ, Joo HR, Heo NY, Park J, et al. Optimal preparation-to-colonoscopy interval in split-dose PEG bowel preparation determines satisfactory bowel preparation quality: an observational prospective study. Gastrointest Endosc. 2012;75(3):583-90.

84. Zawaly K, Rumbolt C, Abou-Setta AM, Neilson C, Rabbani R, Zarychanski R, et al. The Efficacy of Split-Dose Bowel Preparations for Polyp Detection: A Systematic Review and Meta-Analysis. Am J Gastroenterol. 2019;114(6):884-92.

85. Patel V, Nicar M, Emmett M, Asplin J, Maguire JA, Santa Ana CA, et al. Intestinal and renal effects of low-volume phosphate and sulfate cathartic solutions designed for cleansing the colon: pathophysiological studies in five normal subjects. Am J Gastroenterol. 2009;104(4):953-65.

86. Martens P, Bisschops R. Bowel preparation for colonoscopy: efficacy, tolerability and safety. Acta Gastroenterol Belg. 2014;77(2):249-55.

87. Ko SH, Lim CH, Kim JY, Kang SH, Baeg MK, Oh HJ. Case of inappropriate ADH syndrome: hyponatremia due to polyethylene glycol bowel preparation. World J Gastroenterol. 2014;20(34):12350-4.

88. Onishi S, Yoshino S. Cathartic-induced fatal hypermagnesemia in the elderly. Intern Med. 2006;45(4):207-10.

89. Markowitz GS, Stokes MB, Radhakrishnan J, D'Agati VD. Acute phosphate nephropathy following oral sodium phosphate bowel purgative: an underrecognized cause of chronic renal failure. J Am Soc Nephrol. 2005;16(11):3389-96.

90. Mathus-Vliegen EMH, van der Vliet K, Wignand-van der Storm IJ, Stadwijk JS. Split-dose bowel cleansing with picosulphate is safe and better tolerated than 2-l polyethylene glycol solution. Eur J Gastroenterol Hepatol. 2018;30(7):709-17.

91. Lee SP, Park E, Kim HV, Sung IK, Kim JH, Lee SY, et al. Does 2 L Polyethylene Glycol Plus Ascorbic Acid Increase the Risk of Renal Impairment Compared to 4 L Polyethylene Glycol? Dig Dis Sci. 2016;61(11):3207-14.

92. Sacks NC, Sharma A, Cyr PL, Bertiger G, Dahdal DN, Brogadir SP. Real-world comparison of the effectiveness and safety of different bowel preparation agents. Clin Exp Gastroenterol. 2018;11:289-99.

93. Alvarez-Gonzalez MA, Pantaleon MA, Flores-Le Roux JA, Zaffalon D, Amoros J, Bessa X, et al. Randomized Clinical Trial: A Normocaloric Low-Fiber Diet the Day Before Colonoscopy Is the Most Effective Approach to Bowel Preparation in Colorectal Cancer Screening Colonoscopy. Dis Colon Rectum. 2019;62(4):491-7.

94. Nguyen DL, Jamal MM, Nguyen ET, Puli SR, Bechtold ML. Low-residue versus clear liquid diet before colonoscopy: a meta-analysis of randomized, controlled trials. Gastrointest Endosc. 2016;83(3):499-507 e1.

95. Avalos DJ, Sussman DA, Lara LF, Sarkis FS, Castro FJ. Effect of Diet Liberalization on Bowel Preparation. South Med J. 2017;110(6):399-407.

96. Rutherford CC, Calderwood AH. Update on Bowel Preparation for Colonoscopy. Curr Treat Options Gastroenterol. 2018;16(1):165-81.

97. Serper M, Gawron AJ, Smith SG, Pandit AA, Dahlke AR, Bojarski EA, et al. Patient factors that affect quality of colonoscopy preparation. Clin Gastroenterol Hepatol. 2014;12(3):451-7.

98. Liu Z, Zhang MM, Li YY, Li LX, Li YQ. Enhanced education for bowel preparation before colonoscopy: A state-of-the-art review. J Dig Dis. 2017;18(2):84-91.

99. Walter B, Klare P, Strehle K, Aschenbeck J, Ludwig L, Dikopoulos N, et al. Improving the quality and acceptance of colonoscopy preparation by reinforced patient education with short message service: results from a randomized, multicenter study (PERICLES-II). Gastrointest Endosc. 2019;89(3):506-13 e4.

100. Guo X, Yang Z, Zhao L, Leung F, Luo H, Kang X, et al. Enhanced instructions improve the quality of bowel preparation for colonoscopy: a meta-analysis of randomized controlled trials. Gastrointest Endosc. 2017;85(1):90-7 e6.

101. Tian X, Xu LL, Liu XL, Chen WQ. Enhanced Patient Education for Colonic Polyp and Adenoma Detection: Meta-Analysis of Randomized Controlled Trials. JMIR Mhealth Uhealth. 2020;8(6):e17372.

102. Chang CW, Shih SC, Wang HY, Chu CH, Wang TE, Hung CY, et al. Meta-analysis: The effect of patient education on bowel preparation for colonoscopy. Endosc Int Open. 2015;3(6):E646-52.

103. Parekh PJ, Oldfield ECt, Johnson DA. Bowel preparation for colonoscopy: what is best and necessary for quality? Curr Opin Gastroenterol. 2019;35(1):51-7.

104. Obstein KL, Valdastri P. Advanced endoscopic technologies for colorectal cancer screening. World J Gastroenterol. 2013;19(4):431-9.

105. Rigaux J, Juriens I, Deviere J. A novel system for the improvement of colonic cleansing during colonoscopy. Endoscopy. 2012;44(7):703-6.

106. Eliakim R, Yassin K, Lachter J, Chowers Y. A novel device to improve colon cleanliness during colonoscopy. Endoscopy. 2012;44(7):655-9.

107. van Keulen KE, Neumann H, Schattenberg JM, van Esch AAJ, Kievit W, Spaander MCW, et al. A novel device for intracolonoscopy cleansing of inadequately prepared colonoscopy patients: a feasibility study. Endoscopy. 2019;51(1):85-92.

108. Kaiser Junior RL, de Quadros LG, Faria MAG, Kaiser FSL, Campo JCO, Zotarelli Filho IJ. Aquanet Bowel Cleansing Device Versus Oral Sodium Picosulfate for Pre-Endoscopy Bowel Preparation: Propensity Score Analysis for Interventional Effectiveness Evaluation. Gastroenterology Res. 2018;11(1):18-24.

109. Kaiser-Junior RL, LG DE-Q, Flamini-Junior M, Faria MAG, Campo JCO, VL DE-O, et al. New Bowel Preparation Technique for Colonoscopy: Clinical Trial Comparing Aquanet and Mannitol. Arq Bras Cir Dig. 2018;31(3):e1393.

110. Ayub K, Ketwaroo G, Abudayyeh S, Ibrahim A, Cole RA, Brumfield-Brown R, et al. Mechanical colon cleansing for screening colonoscopy: A randomized controlled trial. J Dig Dis. 2017;18(12):691-7.

111. Perez Jimenez J, Diego Bermudez L, Gralnek IM, Martin Herrera L, Libes M. An Intraprocedural Endoscopic Cleansing Device for Achieving Adequate Colon Preparation in Poorly Prepped Patients. J Clin Gastroenterol. 2019;53(7):530-4.

112. Hoffman A, Murthy S, Pompetzki L, Rey JW, Goetz M, Tresch A, et al. Intraprocedural bowel cleansing with the JetPrep cleansing system improves adenoma detection. World J Gastroenterol. 2015;21(26):8184-94.

113. Moshkowitz M, Fokra A, Itzhak Y, Arber N, Santo E. Feasibility study of minimal prepared hydroflush screening colonoscopy. United European Gastroenterol J. 2016;4(1):105-9.

114. Gimeno-Garcia AZ, Baute JL, Hernandez G, Morales D, Gonzalez-Perez CD, Nicolas-Perez D, et al. Risk factors for inadequate bowel preparation: a validated predictive score. Endoscopy. 2017;49(6):536-43.

115. Bucci C, Zingone F, Schettino P, Marmo C, Marmo R. Same-Day Regimen as an Alternative to Split Preparation for Colonoscopy: A Systematic Review with Meta-Analysis. Gastroenterol Res Pract. 2019;2019:7476023.

Supporting information

Supplementary A. Background information bowel preparation fluids
Supplementary B. Search strategy 17 April 2020
Supplementary C. Subgroup and heterogeneity analyses (figures).
Figure S1. Subgroup analysis based on additive use, sodium picosulfate/magnesium citrate (SPMC).
Figure S2. Subgroup analysis based on dosing schedule, sodium picosulfate/magnesium citrate (SPMC).
Figure S3. Subgroup analysis based on diet, sodium picosulfate/magnesium citrate (SPMC)
Figure S4. Pooled efficacy 1L PEG with ascorbate (PEGA), additives subgroups.
Figure S5. Pooled efficacy of 1L PEG with ascorbate (PEGA), dosing subgroups.
Figure S6. Pooled efficacy of 1L PEG with ascorbate (PEGA), diet subgroups.
Figure S7 – leave-1-out analyses for pooled efficacy and heterogeneity. A – Sodium picosulfate with magnesium citrate, B – 1L-PEG with ascorbate, C – Sodium phosphate solution, D – Sodium sulfate solution.
Figure S8 – influence analyses and Baujat-plots for identifying outliers. A – Sodium picosulfate with magnesium citrate, B – 1L-PEG with ascorbate, C – Sodium phosphate solution, D – Sodium sulfate solution.
Figure S9 – Graphic display of study heterogeneity (GOSH) plot analysis for identifying outliers, Sodium Picosulfate with Magnesium Citrate (SPMC)
Figure S10 – Graphic display of study heterogeneity (GOSH) plot analysis for identifying outliers, 1L PEG with ascorbate
Figure S11 – funnel plots of percentage adequately prepped patients per fluid. A – Sodium picosulfate with magnesium citrate, B – 1L-PEG with ascorbate, C – Sodium phosphate solution, D – Sodium sulfate solution.
Figure S12 – Traffic light plot for individual risk of bias of included studies, Cochrane RoB2 tool

Supporting information

A – Background information bowel preparation fluids

Sodium picosulfate magnesium citrate (SPMC) is a 300ml cathartic in 2 doses, combined of 2 active ingredients. Magnesium citrate is an osmotic laxative that stimulates colon motility in a hyperosmolar environment. Sodium picosulfate is a stimulant laxative, decreasing water and salt resorption while increasing water excretion in the bowel lumen.

Other hyperosmolar bowel preparation fluids include sodium phosphate (NaP) and sodium sulfate solution (OSS). NaP is a laxative consisting of 90mL delivered in 2 doses. Although reports demonstrate a better compliance and efficacy compared to PEG[1], NaP is at risk of causing fluid shifts and electrolyte abnormalities, which has led to a black box warning by the United States Food and Drug Administration (FDA) due to the rare risk of acute phosphate nephropathy caused by tubular calcium phosphate depositions[2].

OSS is a phosphate-free alternative to NaP. Gut sulfate resorption is easily saturated, which leaves a high concentration in the bowel lumen, hence its osmotic laxative effect. Sulfate does not cause significant fluid shifts and is delivered in a split dose of a total of 1L solution.

PEG is an iso-osmotic laxative. While its high-volume equivalent exerts bowel cleansing through high-volume lavage, this low-volume counterpart relies more on the use of adjuncts. The resorption of ascorbic acid is easily saturated and thus remains in the gut lumen in high doses, increasing osmolarity[3] and thereby stimulating colon cleansing with a lower volume needed. Other adjuncts include the contact laxatives bisacodyl and lubiprostone, and the prokinetic drug prucalopride.

Less frequently used low-volume cathartics include sennosides. Sennosides are contact stimulant laxative, stimulating colon motility and water and electrolyte excretion. Senna is taken as a syrup of 1mL/kg bodyweight.

B – Search strategy 17 April 2020

Pubmed:

Search	Query	Items found
#3	Search **(#1 NOT (Animals[MeSH] NOT Humans[MeSH]))** Filters: **Publication date from 2015/01/01 to 2020/12/31**	1237
#2	Search **(#1 NOT (Animals[MeSH] NOT Humans[MeSH]))**	3489
#1	Search **((“colonoscopy” [MeSH] OR endoscop*[tiab] OR colonoscop*[tiab]) AND (“Laxatives” [MeSH] OR “Laxatives” [Pharmacological Action] OR “Cathartics” [MeSH] OR “Cathartics” [Pharmacological Action] OR “therapeutic irrigation” [MeSH] OR laxative* [tiab] OR laxan*[tiab] OR laxat*[tiab] OR prepar* [tiab] OR clean* [tiab] OR cathartic* [tiab] OR Bowel Evacuant*[tiab] OR Purgative*[tiab] OR Bowel Preparation Solution*[tiab] OR bowel clean*[tiab] OR bowel preparation*[tiab] OR colon clean*[tiab] OR Low volume *[tiab] OR visualization*[tiab] OR visualisation*[tiab] OR lavage*[tiab] OR prepared colonoscop*[tiab] OR prepared bowel[tiab]) AND (randomized controlled trial[pt] OR controlled clinical trial[pt] OR randomized[tiab] OR randomised[tiab] OR placebo[tiab] OR drug therapy[sh] OR randomly[tiab] OR trial[tiab] OR groups[tiab] OR Aronchick Scale[tiab] OR Aronchick[tiab] OR “Boston Bowel Preparation Scale”[tiab] OR BPPS[tiab] OR Bowel preparation scale*[tiab] OR Boston[tiab] OR “Ottawa Bowel Preparation Scale”[tiab] OR OBPS[tiab] OR scale*[tiab]))**	3656

Embase:

Search		Results
1	exp colonoscopy/ or intestine endoscopy/ or *gastrointestinal endoscopy/ or *rectoscopy/ or *sigmoidoscopy/ or (endoscop* or colonoscop*).ti,ab,kw.	384649
2	exp intestine preparation/ or exp intestine contraction stimulating agent/ or exp intestine lavage/ or (laxative* or laxan* or laxat* or prepar* or clean* or cathartic* or Bowel Evacuant* or Purgative* or Bowel Preparation Solution* or bowel clean* or bowel preparation* or colon clean* or inadequate bowel preparation or inadequate bowel clean* or Low volume* or visualization* or visualisation* or lavage*).ti,ab,kw.	1533557
3	Randomized controlled trial/ or Controlled clinical study/ or ((random$.ti,ab. or randomization/ or intermethod comparison/ or placebo.ti,ab. or (compare or compared or comparison).ti. or ((evaluated or evaluate or evaluating or assessed or assess) and (compare or compared or comparing or comparison)).ab. or (open adj label).ti,ab. or ((double or single or doubly or singly) adj (blind or blinded or blindly)).ti,ab. or double blind procedure/ or parallel group$1.ti,ab. or (crossover or cross over).ti,ab. or ((assign$ or match or matched or allocation) adj5 (alternate or group$1 or intervention$1 or patient$1 or subject$1 or participant$1)).ti,ab. or (assigned or allocated).ti,ab. or (controlled adj7 (study or design or trial)).ti,ab. or (volunteer or volunteers).ti,ab. or human experiment/ or trial.ti.) not (Randomized controlled trial/ or Controlled clinical study/) not (((random$ adj sampl$ adj7 ("cross section$" or questionnaire$1 or survey$ or database$1)).ti,ab. not (comparative study/ or controlled study/ or randomi?ed controlled.ti,ab. or randomly assigned.ti,ab.)) or (Cross-sectional study/ not (randomized controlled trial/ or controlled clinical study/ or controlled study/ or randomi?ed controlled.ti,ab. or control group$1.ti,ab.)) or (((case adj control$) and random$) not randomi?ed controlled).ti,ab. or (Systematic review not (trial or study)).ti. or (nonrandom$ not random$).ti,ab. or "Random field$".ti,ab. or (random cluster adj3 sampl$).ti,ab. or ((review.ab. and review.pt.) not trial.ti.) or ("we searched".ab. and (review.ti. or review.pt.)) or "update review".ab. or (databases adj4 searched).ab. or ((rat or rats or mouse or mice or swine or porcine or murine or sheep or lambs or pigs or piglets or rabbit or rabbits or cat or cats or dog or dogs or cattle or bovine or monkey or monkeys or trout or marmoset$1).ti. and animal experiment/) or (Animal experiment/ not (human experiment/ or human/))))	4442189
4	(Aronchick Scale or Aronchick or Boston Bowel Preparation Scale or BPPS or Bowel preparation scale* or Boston or Ottawa Bowel Preparation Scale or OBPS or scale*).ti,ab,kw.	1094799
5	3 or 4	5196395
6	1 and 2 and 5	8897
7	limit 6 to conference abstract	4099
8	6 not 7	4798
9	8 not ((exp animal/ or exp invertebrate/ or nonhuman/ or animal experiment/ or animal tissue/ or animal model/ or exp plant/ or exp fungus/) not (exp human/ or human tissue/))	4662
10	limit 9 to yr="2015 -Current"	1402

Cochrane – CENTRAL:

(([mh colonoscopy] OR (endoscop* OR colonoscop*):ti,ab,kw) AND ([mh Laxatives] OR [mh Cathartics] OR [mh "therapeutic irrigation"] OR (laxative* OR laxan* OR laxat* OR prepar* OR clean* OR cathartic* OR Bowel Evacuant* OR Purgative* OR Bowel Preparation Solution* OR bowel clean* OR bowel preparation* OR colon clean* OR inadequate bowel preparation OR inadequate bowel clean* OR Low volume* OR visualization* OR visualisation* OR lavage*):ti,ab,kw)) NOT ([mh animals] NOT [mh humans])

Limits: publication date from Jan 2015 to Dec 2020

Results: 20 Cochrane reviews, 2438 trials, 1 Cochrane protocol

C – Subgroup and heterogeneity analyses

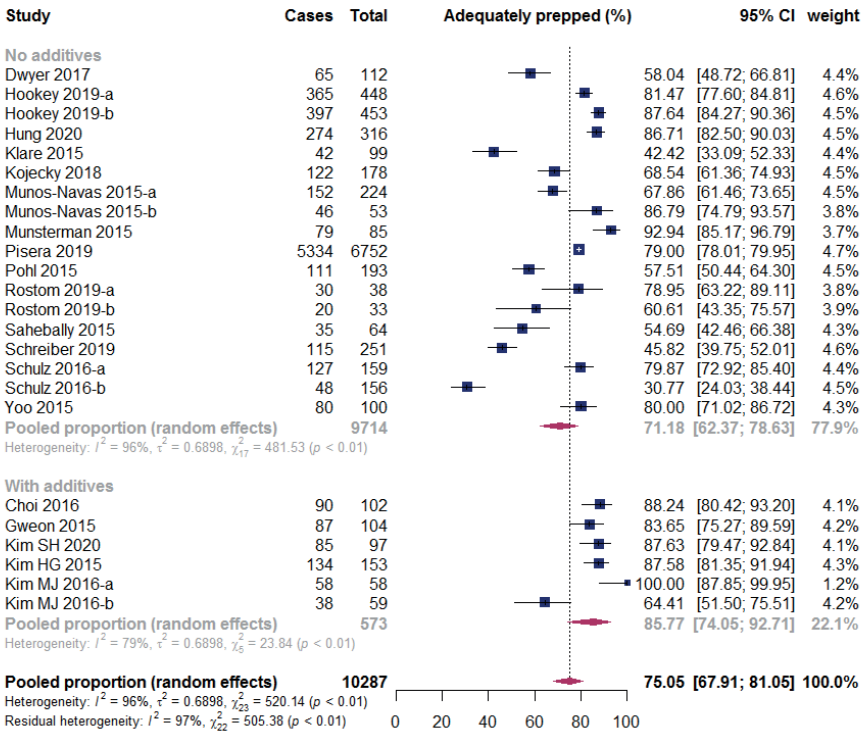


Figure S1. Subgroup analysis based on additive use, sodium picosulfate/magnesium citrate (SPMC).

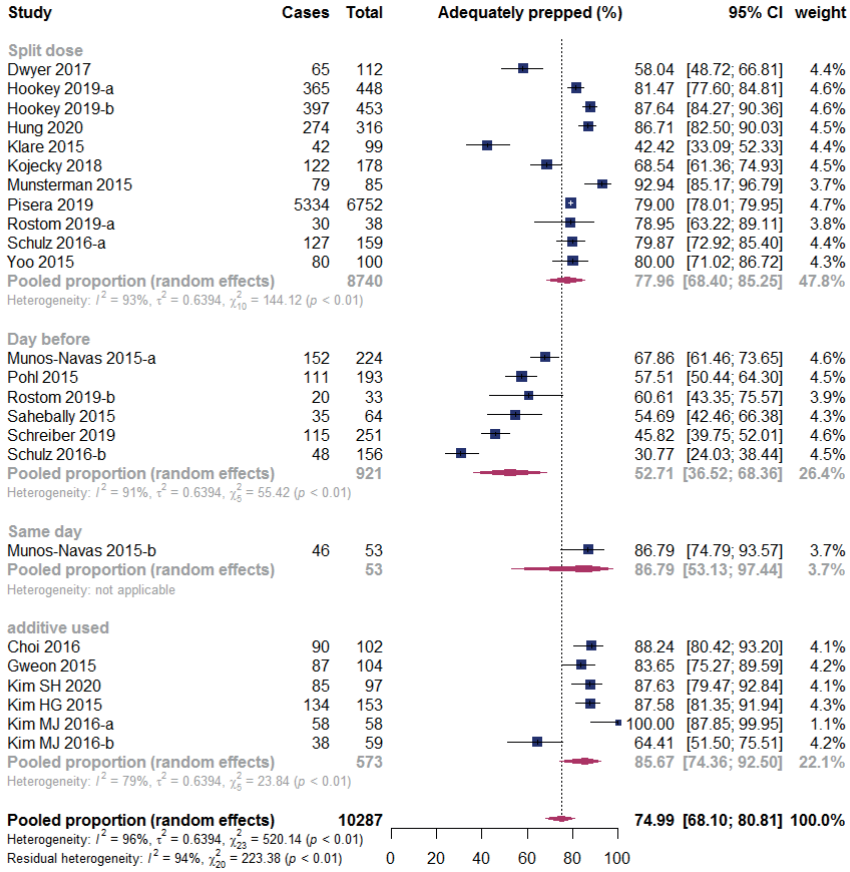


Figure S2. Subgroup analysis based on dosing schedule, sodium picosulfate/magnesium citrate (SPMC).

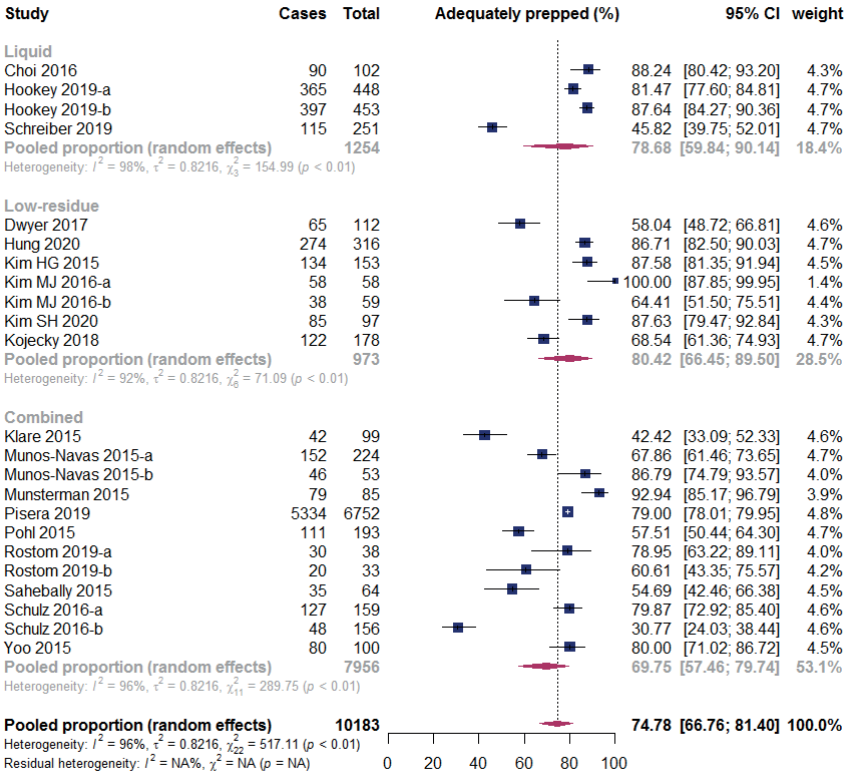


Figure S3. Subgroup analysis based on diet, sodium picosulfate/magnesium citrate (SPMC)

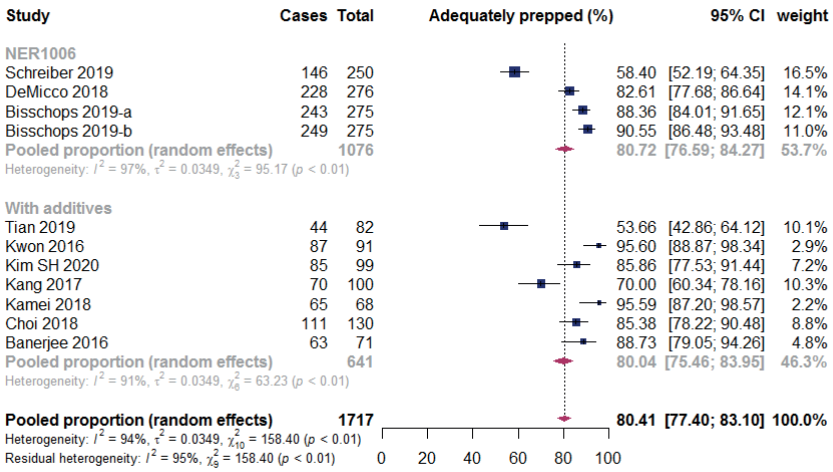


Figure S4. Pooled efficacy 1L PEG with ascorbate (PEGA), additives subgroups.

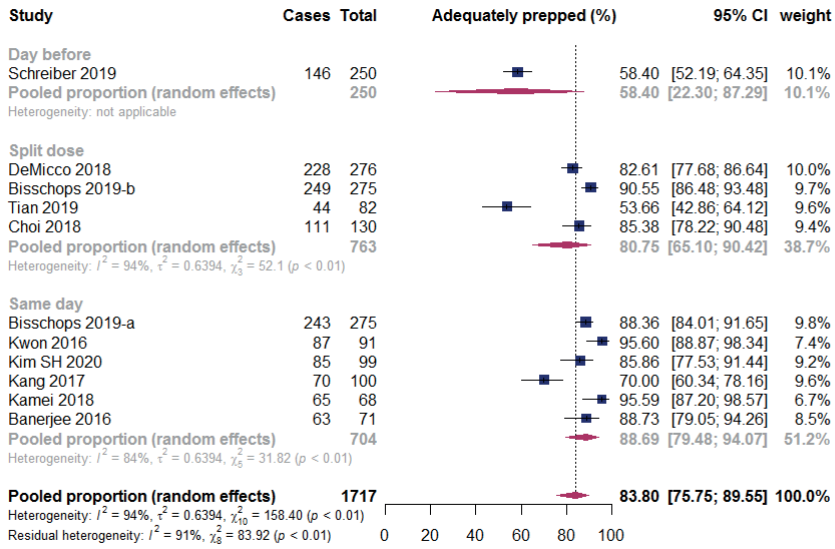


Figure S5. Pooled efficacy of 1L PEG with ascorbate (PEGA), dosing subgroups.

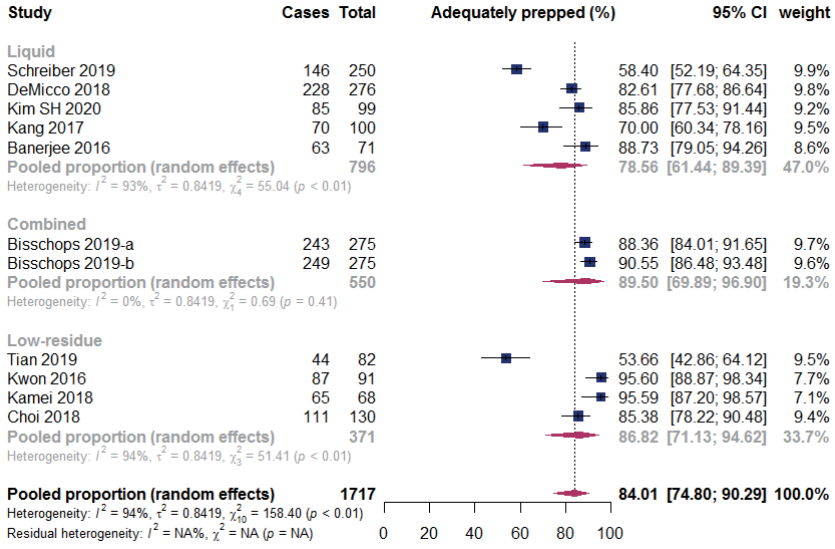


Figure S6. Pooled efficacy of 1L PEG with ascorbate (PEGA), diet subgroups.

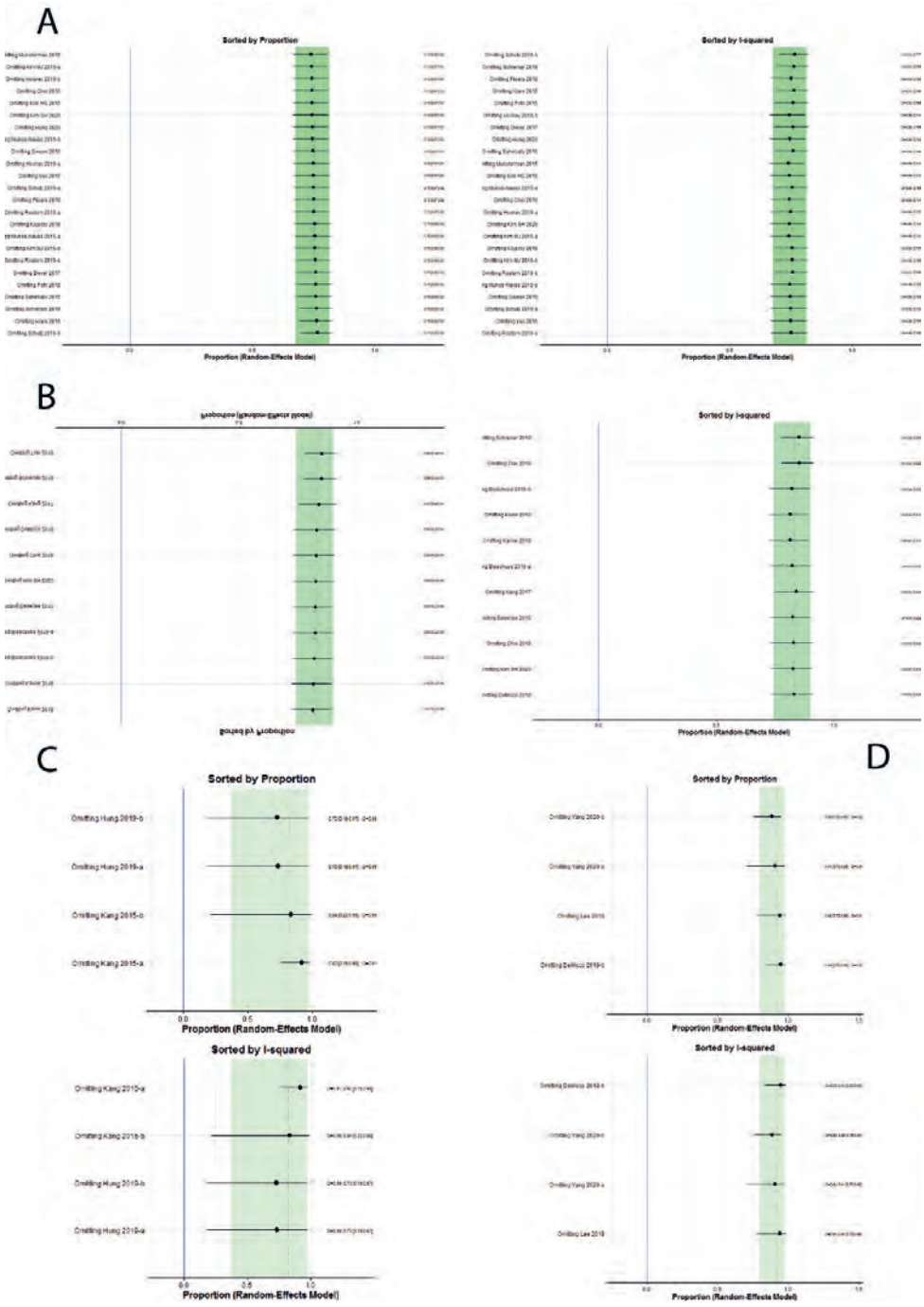


Figure S6 – leave-1-out analyses for pooled efficacy and heterogeneity. A – Sodium picosulfate with magnesium citrate, B – 1L-PEG with ascorbate, C – Sodium phosphate solution, D – Sodium sulfate solution.

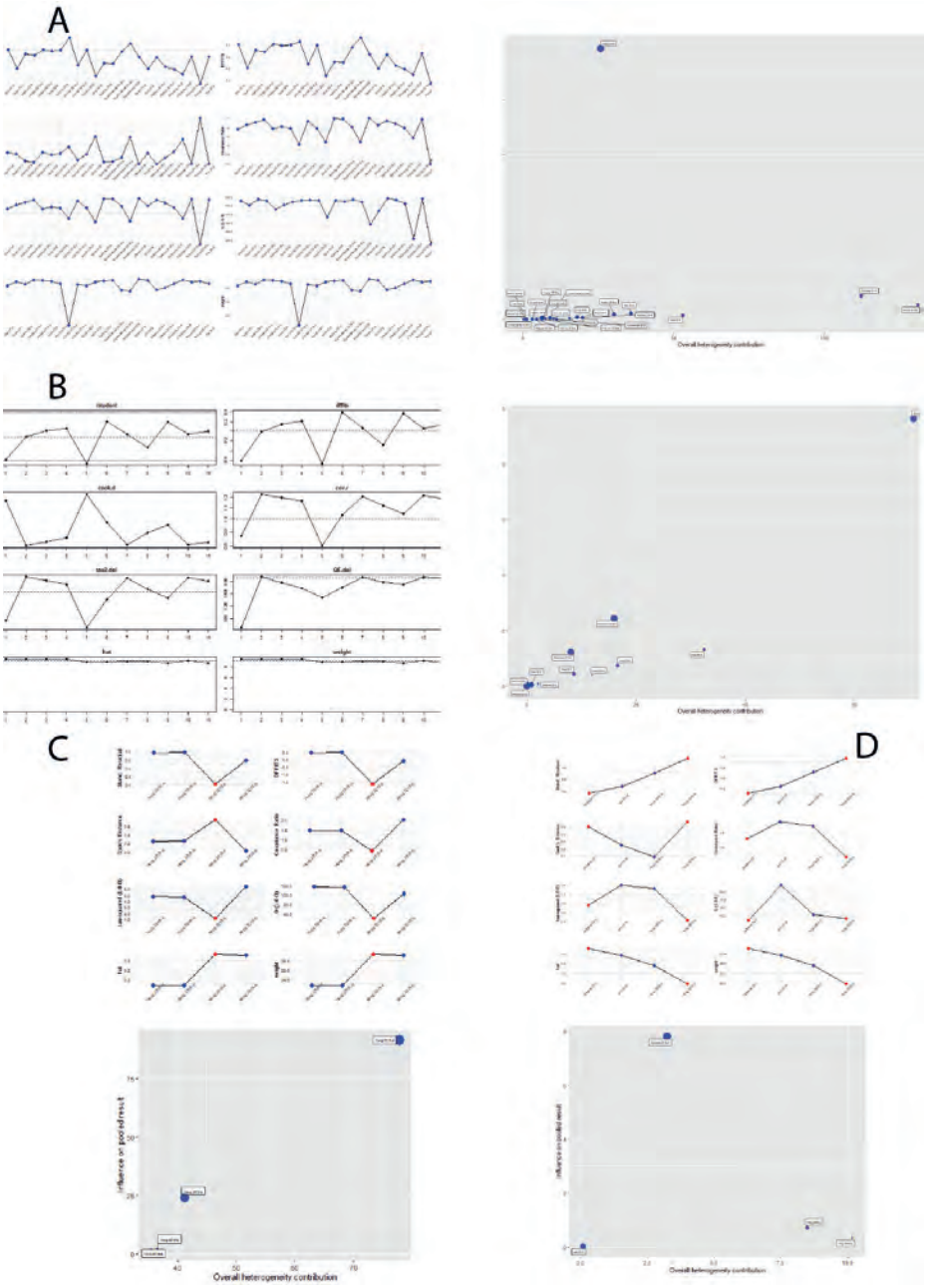


Figure S8 – influence analyses and Baujat-plots for identifying outliers. A – Sodium picosulfate with magnesium citrate, B – 1L-PEG with ascorbate, C – Sodium phosphate solution, D – Sodium sulfate solution.

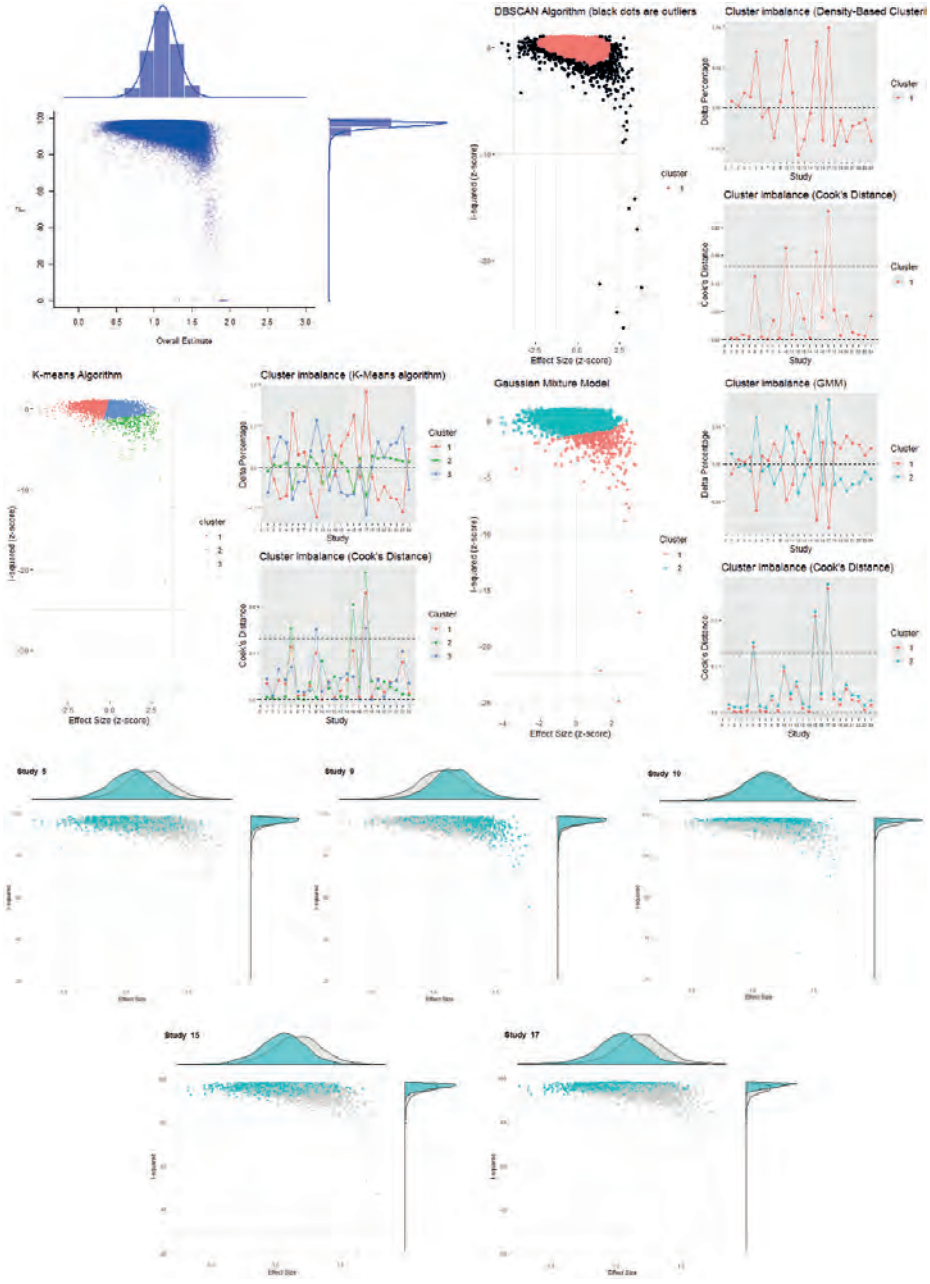


Figure S9 – Graphic display of study heterogeneity (GOSH) plot analysis for identifying outliers, Sodium Picosulfate with Magnesium Citrate (SPMC)

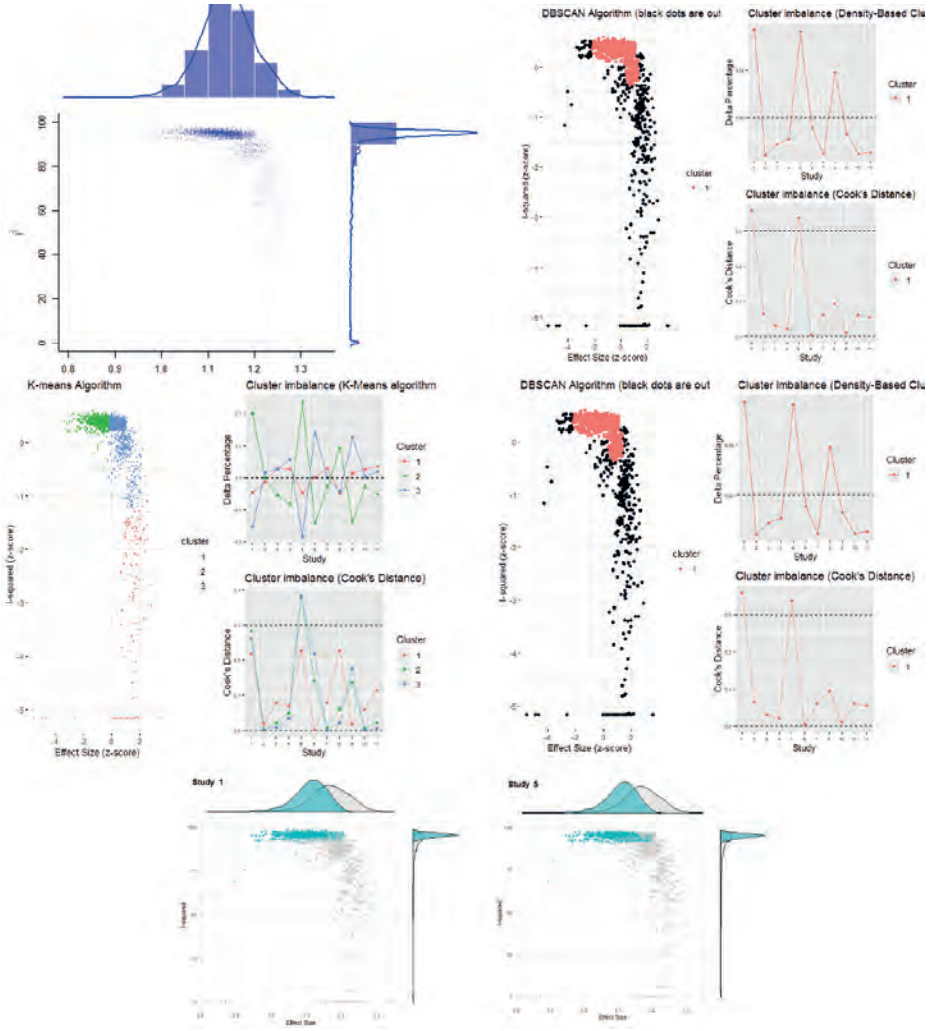


Figure S10 – Graphic display of study heterogeneity (GOSH) plot analysis for identifying outliers, 1L PEG with ascorbate

D – Risk of bias

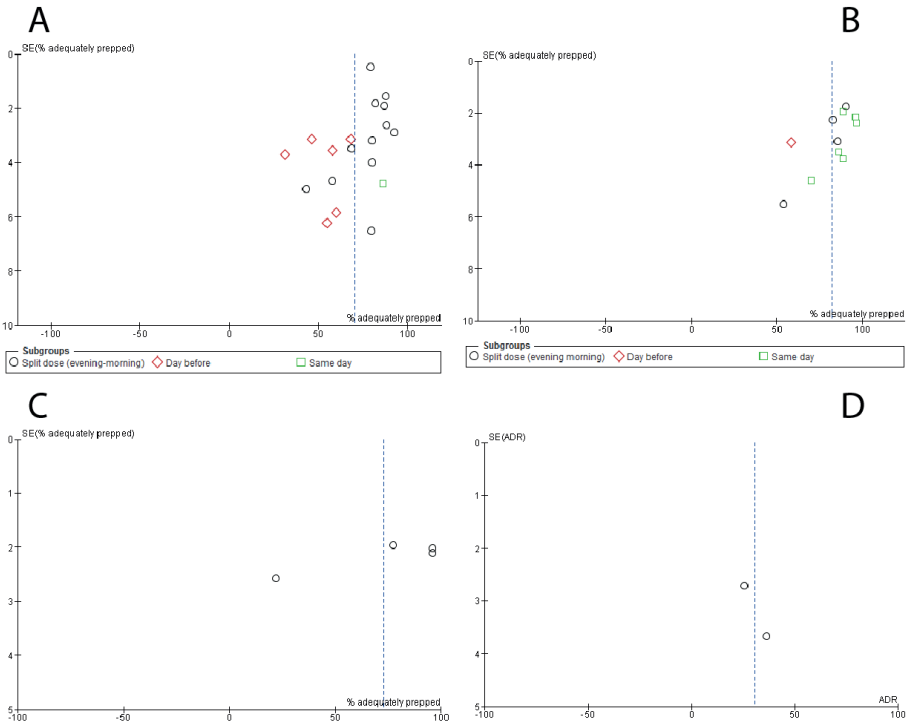


Figure S11 – funnel plots of percentage adequately prepped patients per fluid. A – Sodium picosulfate with magnesium citrate, B – 1L-PEG with ascorbate, C – Sodium phosphate solution, D – Sodium sulfate solution.

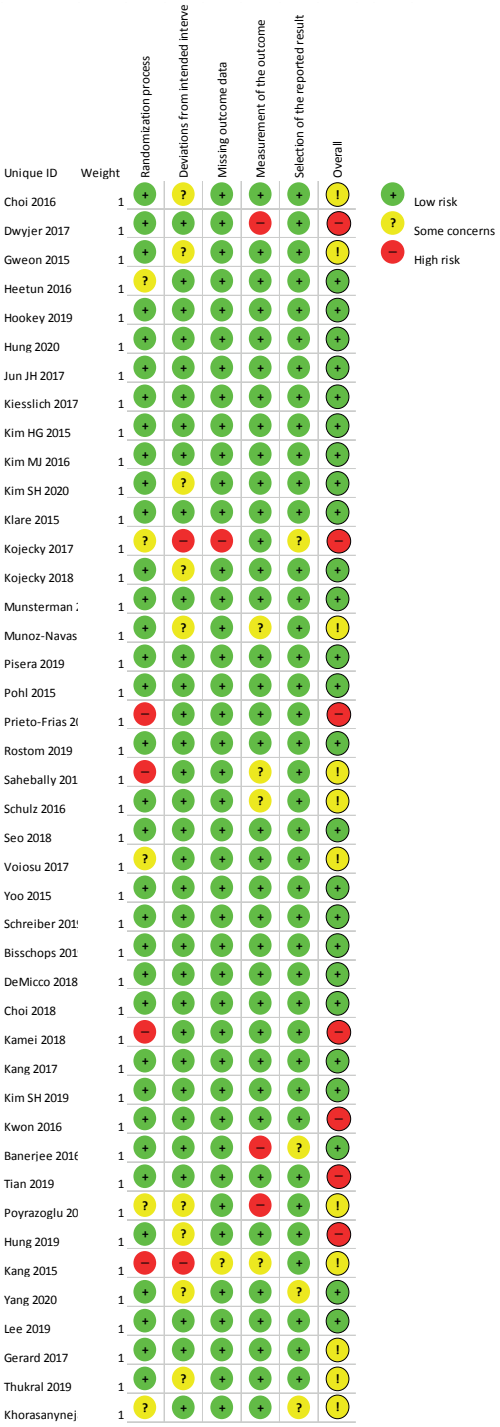
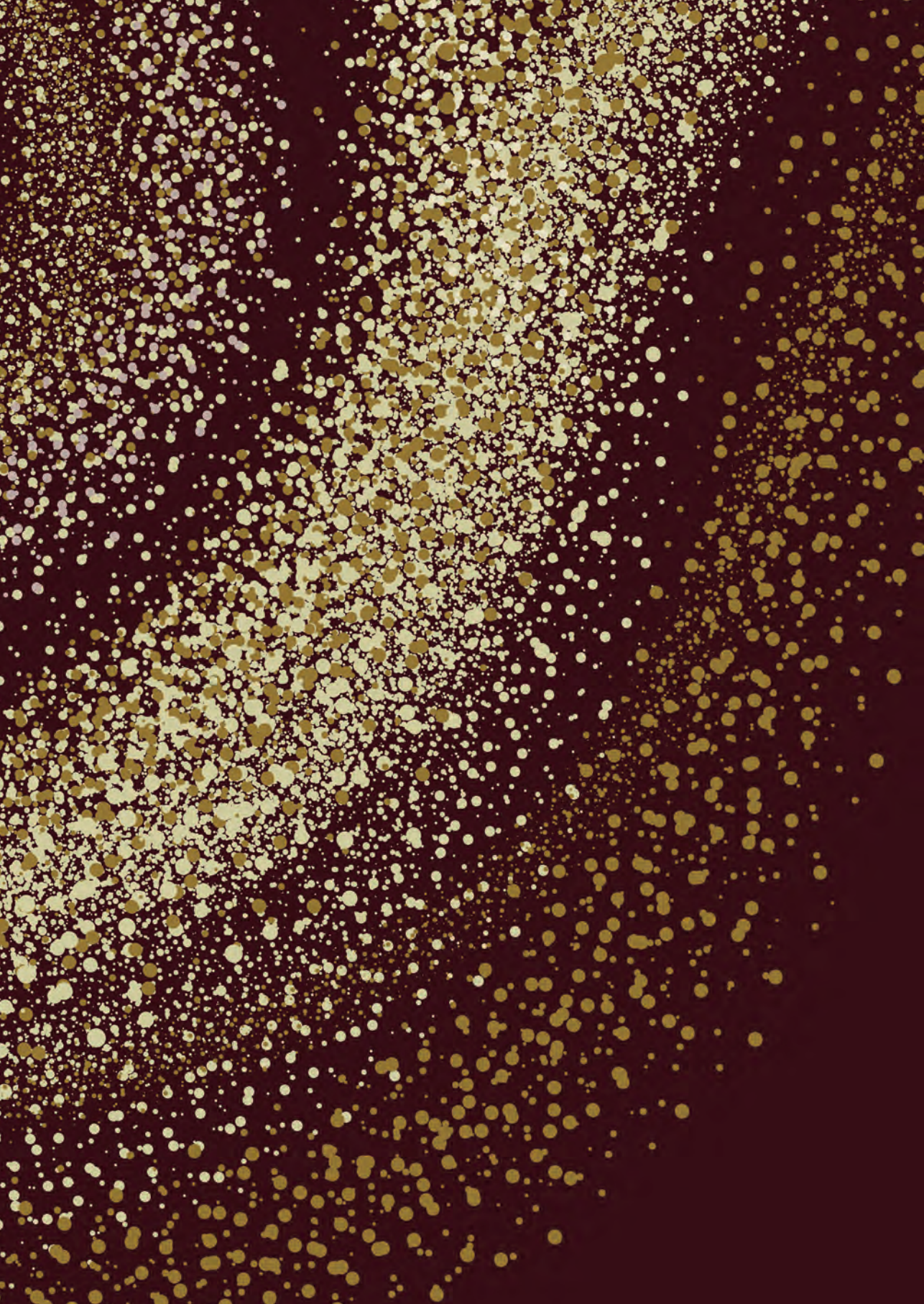


Figure S12 – Traffic light plot for individual risk of bias of included studies, Cochrane RoB2 tool

Supplementary files references

1. Tan JJ, Tjandra JJ. Which is the optimal bowel preparation for colonoscopy - a meta-analysis. *Colorectal Dis*. 2006; **8**: 247-58.
2. Markowitz GS, Stokes MB, Radhakrishnan J, D'Agati VD. Acute phosphate nephropathy following oral sodium phosphate bowel purgative: an underrecognized cause of chronic renal failure. *J Am Soc Nephrol*. 2005; **16**: 3389-96.
3. Fujita I, Akagi Y, Hirano J, *et al.* Distinct mechanisms of transport of ascorbic acid and dehydroascorbic acid in intestinal epithelial cells (IEC-6). *Research communications in molecular pathology and pharmacology*. 2000; **107**: 219-31.



Chapter 5

Comparing low-volume vs. intermediate volume bowel preparation and its impact on work and tolerability: An open-label, non-inferiority, randomized controlled trial

Milou L.M. van Riswijk, MD, F.A. Indemans, MD, K. Hawinkels, MD, R.M. Schreuder, MD, L Wildeman, Adriaan C.I.T.L Tan, MD, PhD, Peter D. Siersema, MD, PhD

The present version represents the submitted manuscript.
The published study can be found at doi.org/10.1055/a-2695-0994.
Endoscopy. 2025 Sep 3.

Abstract

Background: Bowel preparation is essential for colonoscopy, but patients may experience it as a deterrent factor for colonoscopy. Although low-volume laxatives are associated with better tolerability, little is known on their effect on patient reported outcomes. We compared low- and intermediate volume bowel preparation and assessed the impact on tolerability, health-related quality of life (HRQoL) and work.

Methods: We performed an open-label, non-inferiority, randomized trial in 4 Dutch hospitals. Patients were randomized between 2L polyethylene glycol with ascorbate (2LPEG+Asc) or 1L PEG+Asc with added sodium sulfate. Before and after bowel preparation, patients completed validated questionnaires on productivity costs, tolerability, and QoL. Primary outcome was non-inferiority of 1L to 2L PEG+Asc. Secondary outcomes included change in QoL scores, tolerability, and impact on working productivity.

Results: We included 467 patients (2LPEG+Asc, n=229 and 1L PEG+Asc, n=238). 1L PEG+Asc was non-inferior to 2L PEG+Asc, with adequate cleansing rates of 96.1% (95% CI 92.6-98.0%) and 96.4% (95% CI 92.9-98.3%), respectively (p=0.841). Patients in the 1L PEG+Asc group were more willing to repeat bowel preparation (59.88% vs 48.3%, p=0.044), with experienced tolerability being the most influential factor (OR 0.053-0.225 in case of difficult or fair tolerability). No clinically significant changes were found in the HRQoL scores. Absenteeism and impaired working productivity were seen in 7.9% and 12.3% of patients, respectively, but did not differ significantly between groups.

Conclusions: Bowel preparation with 1L PEG+Asc is non-inferior to 2L PEG+Asc and has a higher willingness-to-repeat. Bowel preparation tolerability is fundamental for effective cleansing and reducing colonoscopy barriers.

ClinicalTrials.gov Identifier: NCT05242562

Keywords: Colonoscopy, bowel preparation, quality of life, randomized controlled trial, work productivity

Introduction

High quality colonoscopy cannot be performed without adequate bowel preparation [1]. Inadequate bowel preparation is associated with lower lesion detection rates, and a major predictor of failed cecal intubation. Moreover, the need for repeat procedures increase healthcare costs. Importantly, patients with poor bowel preparation are less likely to be satisfied about their colonoscopy and have a lower willingness-to-repeat colonoscopy [2]. This is detrimental to the benefit of CRC screening, as lower participation rates significantly decrease colonoscopy screening efficiency [3].

Efforts have been made to increase both bowel preparation efficiency and tolerability. Meta-analyses have demonstrated a higher tolerability of lower volume bowel preparation fluids [4]. A 2L poly-ethylene glycol with added ascorbate (PEG+Asc) solution has been shown to be non-inferior to the "gold standard" of 4L PEG, and has a higher willingness-to-repeat [1]. Recently, even lower volume bowel preparations were introduced, such as 1L PEG+Asc with added sodium sulfate.

Nevertheless, the impact of bo wel preparation on patients' working life and health related quality of life (HRQoL) may be significant, leading to postponement and sometimes patients refraining from colonoscopy [5, 6]. Despite its negative correlation with both bowel preparation efficacy and colonoscopy uptake, little is known on the effect of bowel preparation on patient reported outcomes, and differences between various volume laxatives. In a prospective cohort of 1100 patients, Collatuzzo et al. reported a correlation between patient-related factors and symptoms caused by bowel preparation [7]. Fuccio et al. demonstrated that bowel preparation and colonoscopy significantly impacted work, with up to one-third of patients reporting work absenteeism or reduced performance, particularly when using same-day full-dose regimens or experiencing procedure-related symptoms [8]. However, they did not include 1L PEG+Asc, and the studies were not randomized, which may lead to selection bias of patients choosing the laxative of their choice. Additionally, unemployed patients were excluded, while a large proportion of the patient population undergoing colonoscopy also consists of patients who are retired or not working due to illness. Other studies on impact of bowel preparation on patient reported outcomes and working productivity are scarce and mainly involve larger volume preparations (2L and higher).

Insight on impact of experienced symptoms, work, and HRQoL may help patients and caregivers in selecting a bowel preparation and could reduce the experienced

barriers for undergoing colonoscopy. Therefore, this study aims to compare low-volume bowel preparation to intermediate-volume bowel preparation on tolerability, impact on work, and HRQoL.

Methods

Study design

We performed an open label, non-inferiority, randomized controlled trial in four Dutch hospitals. We included adult patients referred for diagnostic, screening, or surveillance colonoscopy. Exclusion criteria included therapeutic procedures, inpatient status, emergency colonoscopy, inflammatory bowel disease, American Society of Anesthesiologists (ASA) score ≥4, (partial) colectomy, inability to provide informed consent, inability to complete Dutch questionnaires via e-mail (due to language barrier, lack of e-mail, visual impairment, or dementia), common contraindications for bowel preparation or its ingredients, and indications for intensified bowel preparation.

The study was approved by the Radboudumc Medical Ethics Committee (Medisch Ethische Toetsingscommissie Oost-Nederland, registration number NL79014.091). It was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines and registered on clinicaltrials.gov (NCT05242562). All patients provided written informed consent. An independent monitor reviewed the study data and informed consent process, and two research team members independently verified all study data.

Randomization

After informed consent, eligible patients were randomized 1:1 using a block design with variable block sizes (4, 6, 8) via a secure web-based system (Castor Electronic Data Capture, Amsterdam, the Netherlands). Randomization was stratified by study site and prior experience with bowel preparation. The randomization sequence was blinded to study team members, but patients and healthcare providers were not blinded to allocation to minimize interference with routine medical care. Blinding of patients was not possible due to the differing volumes. We included additional questions on patient experiences and satisfaction in the questionnaires to assess the impact of potential modifying factors without disrupting standard care.

Study procedures

Bowel preparation and colonoscopy

Patients received either 1L PEG+Asc (Pleinvue, Norgine) or 2L PEG+Asc (Moviprep, Norgine) [9] in an overnight split-dose regimen, each dose followed by 0.5L clear liquids (additional fluids allowed). The last dose needed to be completed 2 hours before travel to the hospital. Patients were also instructed to follow a low-residue diet (LRD) starting two days before colonoscopy, followed by a liquid diet (CLD) upon start taking the laxatives. Instructions were provided by specialized nurses supplemented with a detailed leaflet.

Colonoscopies were performed as part of standard medical care [9], under conscious sedation or propofol, by experienced endoscopists. Bowel preparation quality was assessed using the Boston Bowel Preparation Scale (BBPS), with adequate preparation defined as a minimum score of 2 per segment.

Questionnaires

Participants completed online questionnaires at two timepoints: before starting preparation (baseline) and within one week after colonoscopy. The first questionnaire collected baseline socio-demographic and lifestyle characteristics, clinical details including bowel symptoms, prior bowel preparation experience, risk factors for inadequate preparation, and HRQoL data. The second questionnaire assessed patient experiences with the bowel preparation and colonoscopy, impact on work productivity and costs, and HRQoL.

To assess work productivity impact, we used the validated *Institute for Medical Technology Assessment Productivity Cost Questionnaire* (IPCQ). Unlike the commonly used *Work Productivity and Activity Impairment questionnaire*, the IPCQ also considers the impact of unpaid labor [10]. HRQoL was assessed using the EuroQol Group five dimensions five levels (EQ-5D-5L) and the Short Form-36 (SF-36) version 2 questionnaires. SF-36 version 2 includes improvements in instructions, wording, and response options, enhancing validity, reliability, and precision [11]. The results are directly comparable to version 1 results. Previous studies indicate that the endoscopic procedure minimally affects HRQoL outcomes, with bowel preparation being the primary factor influencing HRQoL [12, 13].

The second questionnaire also included the *Mayo Florida Bowel Preparation Tolerability Questionnaire* (MBTQ), validated in outpatient colonoscopy settings in the United States [14]. For our study, the MBTQ was translated using a forward-

backward translation process by a professional translator. In addition to the MBTQ, we assessed patient satisfaction using the *Dutch Gastrointestinal Endoscopy Satisfaction Questionnaire* (D-GESQ) [15].

Outcomes

Our primary outcome was non-inferiority in bowel cleansing of 1L PEG+Asc to 2L PEG+Asc. Rates of adequate preparation were compared using a pre-defined non-inferiority margin of 5%. For the secondary outcomes we aimed to multidimensionally assess patient experience for a broader insight in factors that drive tolerability of and willingness-to-repeat bowel preparation, next to impact of bowel preparation on HRQoL and working life of patients.

Statistical analysis

Baseline characteristics were summarized using means and standard deviations (SD), medians and interquartile ranges (IQR), or proportions, as appropriate. Comparisons were made using Student's t-test, Mann-Whitney-U test, Wilcoxon's rank test, chi-square test, or Fisher's exact test. Primary outcome analysis was conducted on both an intention-to-treat and per-protocol basis. Non-inferiority was assessed by calculating the confidence interval (CI) of the proportion of adequately prepared patients, with a 5% non-inferiority margin. If non-inferiority was confirmed, superiority was tested using Fisher's exact test. Questionnaire outcomes were analyzed per their manuals. The SF-36 mental and physical component scores (MCS and PCS) and EQ-index of the EQ-5D-5L were compared to the general population [16]. A minimal clinical important change was defined as 0.5SD of the change [17]. Costs of absenteeism, presenteeism, and unpaid labor were calculated using Dutch government-set reference costs (Zorginstituut Nederland), with hourly costs of €34.75 for paid work and €14 for unpaid labor. A multiple logistic regression was used to evaluate associations between willingness-to-repeat the bowel preparation, with variables selected based on literature and outcome events. After assessing linearity and multicollinearity, variables with a P-value <0.2 in univariate analysis were included in multivariate analysis, using backward propagation for model selection. Odds ratios (OR) and 95% CI were reported. Missing data were handled through multiple imputation after confirming data were missing at random, using predictive mean matching with 50 iterations to create 50 imputed datasets. All analyses were performed using SPSS Statistics (IBM, version 25.0), with a two-sided significance level of 5%. Secondary outcomes were assessed in an exploratory fashion, therefore multiple-testing corrections were not applied.

Sample size

Assuming a 95% adequate preparation rate in both groups, a sample of 470 patients (235 per group) was calculated to be required for demonstrating non-inferiority with a 5% one-sided alpha and 80% power. For the secondary endpoint, the study was powered to detect a minimally clinically relevant 5-point difference on the SF-36, with a mean score of 70 and a standard deviation of 20. This required 137 per arm with a 60% response rate.

Results

Baseline characteristics

From May 2022 to February 2023, we enrolled 509 patients across four centers. After excluding 42 patients for consent withdrawal or exclusion criteria, 238 were randomized to the 1L PEG+Asc group and 229 to the 2L PEG+Asc group. The intention-to-treat population included 467 patients. In the per protocol analysis, 35 patients were excluded due to non-compliance or logistical treatment arm switches, resulting in 432 patients (Figure 1).

Baseline characteristics were well balanced (Table 1). The questionnaire response rate was 87.8% (n=410) and was not significantly different between groups (p=0.321). Half of the study population (49.3%) had previous experience with bowel preparation, of which 2L PEG+Asc was the most common prior laxative (24.9%-29.8%). As for pre-existent risk factors for inadequate bowel preparation, most patients had a fair physical activity level, whereas 88.2%-89.9% did not have any risk factors for poor bowel preparation (Table 1). The background of patients was predominantly Dutch (96.2-96.6%), with intermediate to high education levels. Nearly half were employed (51.7-54.0%) (Table 1).

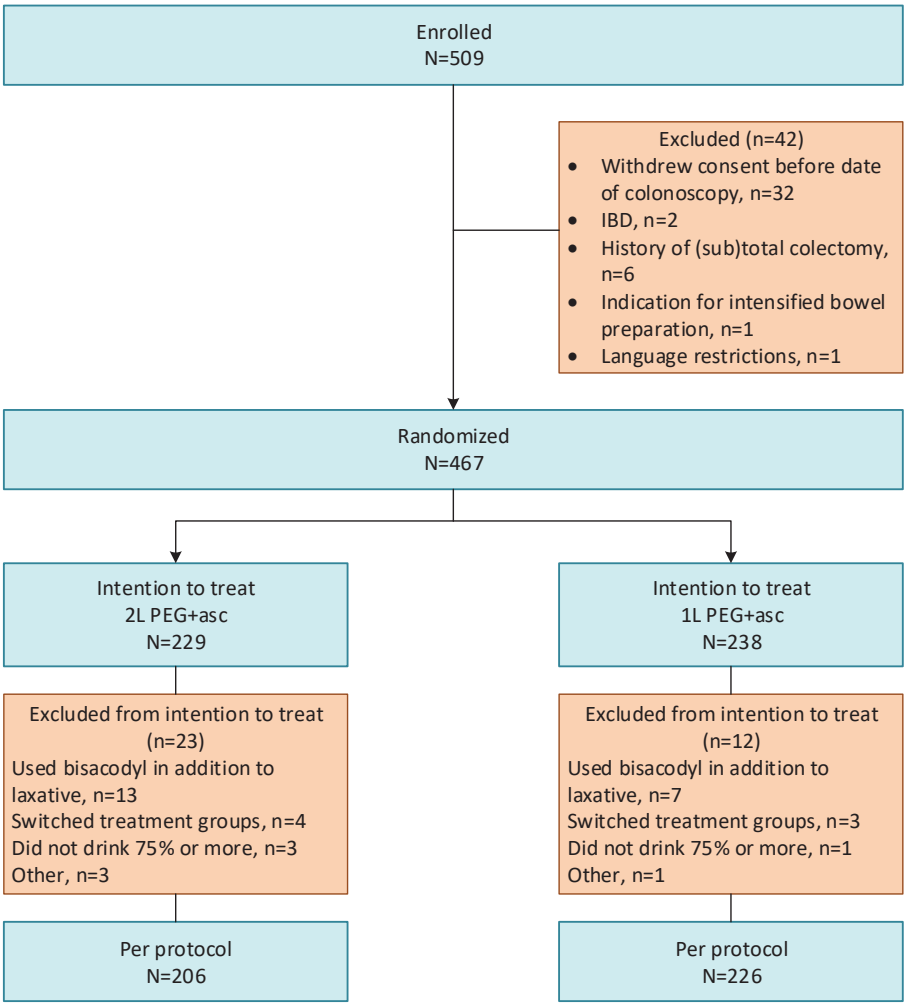


Figure 1. Patient flow chart. *IBD; inflammatory bowel disease, PEG+Asc; polyethylene glycol with added ascorbate*

Table 1. Baseline characteristics.

	1L PEG+Asc (n=238, ITT)*	2L PEG+Asc (n=229, ITT)**
Response rate, n (%)		
Baseline		
After colonoscopy	223 (93.7)	206 (90.0)
Both questionnaires	218 (91.6)	204 (89.1)
median time between	212 (89.1)	198 (86.5)
questionnaires in days (IQR)	8 (4-13)	9 (5-14)
Age, median (IQR)	63 (55-71)	64 (55-70)
Female sex, n (%)	101 (42.6)	98 (42.8)
ASA score, median (IQR)	2 (1-2)	2 (1-2)
BMI, median (IQR)	25.9 (23.6-29.7)	26.3 (23.4-29.4)
Prior experience with bowel preparation, n (%)	113 (49.3)	113 (47.5)
Number of prior colonoscopies, median (IQR)	2 (1-3)	2 (1-3)
Prior laxative used, n (%)		
2L PEG+Asc	71 (29.8)	57 (24.9)
3L PEG+Asc or more	7 (2.9)	14 (6.1)
1L PEG+Asc	3 (1.3)	3 (1.3)
300mL sodium picosulfate magnesium citrate	2 (0.8)	4 (1.7)
4L PEG	9 (3.8)	7 (3.1)
Other	3 (1.3)	2 (0.9)
Do not remember	34 (14.3)	34 (14.8)
Willingness-to-repeat prior laxative, n (%)		
2L PEG+Asc	65 (91.5)	49 (86.0)
3L PEG+Asc or more	6 (85.7)	13 (92.9)
1L PEG+Asc	3 (100.0)	3 (100.0)
300mL sodium picosulfate magnesium citrate	2 (100.0)	4 (100.0)
4L PEG	7 (77.8)	6 (85.7)
Other	2 (66.7)	2 (100.0)
Do not remember	28 (82.4)	26 (76.5)
Colonoscopy indication, n (%)		
Screening	75 (31.5)	89 (38.9)
Surveillance	77 (32.4)	71 (31.9)
Diagnostic	84 (35.3)	67 (29.3)
Smoking, n (%)		
Active smoker	26 (11.1)	25 (11.2)
Former smoker	113 (48.3)	107 (48.0)
Never	94 (40.2)	91 (40.8)
Migration background, n (%)		
Dutch	217 (96.9)	200 (96.2)
Western migration background	3 (1.3)	5 (2.4)
Non-western migration background	2 (0.9)	0
Other	0	2 (0.9)
Prefer not to say	2 (0.9)	1 (0.5)

Table 1. Continued

	1L PEG+Asc (n=238, ITT)*	2L PEG+Asc (n=229, ITT)**
Education level, n (%)		
None	3 (1.3)	0
Primary school	4 (1.8)	3 (1.4)
Secondary school or vocational college	126 (58.0)	119 (57.5)
University or college	89 (39.8)	82 (39.6)
Other	2 (0.9)	3 (1.4)
Marital status, n (%)		
Single	21 (9.4)	20 (9.6)
Married	149 (66.5)	150 (72.1)
Partner, living together	28 (12.5)	25 (12.0)
Partner, not living together	10 (4.5)	4 (1.9)
Widow(er)	14 (6.3)	7 (3.4)
Other	2 (0.9)	2 (1.0)
Physical Activity level, n (%)		
Bedridden	0 (0)	2 (1.0)
Sedentary work	9 (4.0)	3 (1.4)
Work involving walking, no heavy lifting	118 (52.7)	100 (48.3)
Work involving walking and heavy lifting	43 (19.2)	33 (15.9)
Particularly strenuous physical work	54 (24.1)	69 (33.3)
Paid job, n (%)		
Yes	121 (54.0)	107 (51.7)
No	103 (46.0)	100 (48.3)
Work activity per week, mean (SD)		
Hours	31.8 (11.7)	32.6 (11.4)
Days	4.2 (1.1)	4.3 (1.1)
Occupational level, n (%)		
Student	1 (0.4)	2 (1.0)
Employed	90 (40.2)	79 (38.2)
Entrepreneur	19 (8.5)	19 (9.2)
Homemaker	17 (7.6)	6 (2.9)
Unemployed	2 (0.9)	1 (0.5)
Incapacitated for work	7 (3.1)	7 (3.4)
Retired	82 (36.6)	85 (41.1)
Other	6 (2.7)	8 (3.9)
Risk factors for poor bowel preparation, n (%)		
Related conditions		
none	214 (89.9)	202 (88.2)
Constipation	28 (11.8)	26 (11.4)
Abdominal surgery	27 (11.3)	27 (11.8)
Diabetes mellitus	22 (9.2)	18 (7.9)
Kidney disease	3 (1.3)	1 (0.4)
Liver cirrhosis	3 (1.3)	0 (0)
Parkinson's disease	1 (0.4)	1 (0.4)
CVA	4 (1.7)	4 (1.7)
Medication		
Polypharmacy	191 (80.3)	189 (79.9)
Opioid use	6 (2.5)	3 (1.3)
TCA use	3 (1.2)	2 (0.9)

Table 1. Continued

		1L PEG+Asc (n=238, ITT)*	2L PEG+Asc (n=229, ITT)**
Bristol stool scale score, n (%)			
	1	5 (2.2)	4 (1.9)
	2	14 (6.3)	8 (3.9)
	3	39 (17.4)	39 (18.8)
	4	91 (40.6)	87 (42.0)
	5	31 (13.8)	29 (14.0)
	6	39 (17.4)	36 (17.4)
	7	5 (2.2)	4 (1.9)
Median (IQR)		4 (3-5)	4 (4-5)
History of inadequate bowel preparation, n (%)		7 (6.5)	6 (6.1)

* missing n=14 (5.9%), ** missing n=22 (9.6%). *Asc; ascorbate, ASA; American Society of Anesthesiologists, BMI; body mass index, CVA; cerebrovascular accident, IQR; interquartile range, ITT; intention-to-treat, PEG; polyethylene glycol, SD; standard deviation, TCA; tricyclic antidepressant.*

Bowel cleansing efficacy

In the 1L PEG+Asc group, 96.1% (95%CI 92.6-98.0%) and 96.4% (95%CI 92.9-98.3%) of colonoscopies were adequately prepared in the ITT and PP analysis compared to 96.4% (95%CI 93.0-98.3%) and 96.6% (95%CI 92.9-98.5%), respectively, in the 2L PEG+Asc group (p_{ITT}=0.841, p_{PP}=0.917) (Figure 2, Table 2). This met our predefined non-inferiority margin of 5%. Median BBPS scores were significantly higher in the 1L PEG+Asc group for the right (p=0.012) and transverse (p=0.004) colon, but not significantly different in the left colon. High-quality cleansing (BBPS 3-3-3) was more common in the 1L PEG+Asc group, with 72.7% (95% CI 67.0-78.0%) versus 63.8% (95% CI 57.5-70.0%) in the 2L PEG+Asc group for the ITT population (p=0.038). Due to high adequate cleansing rates, in-depth analysis of reasons of inadequate bowel preparation was not performed. Other colonoscopy quality parameters are in Table 2.

Table 1. Baseline characteristics.

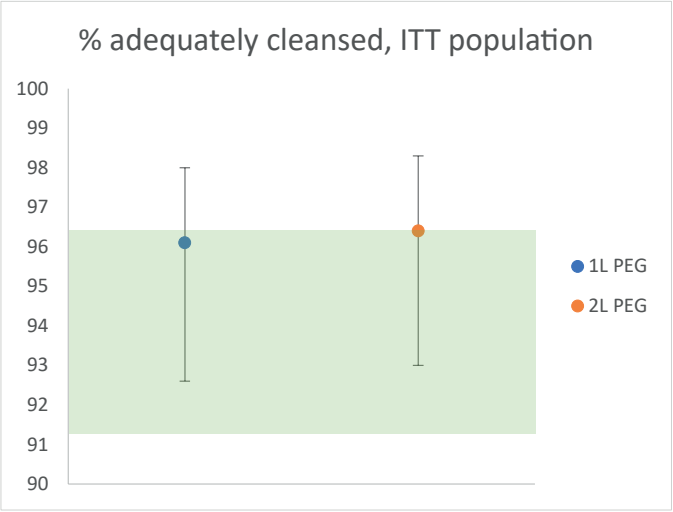


Figure 2. Proportion of adequately cleaned patients using 1L or 2L polyethylene glycol (PEG) bowel preparation with 95% confidence interval. The difference falls in our predefined non-inferiority margin of 5%, establishing non-inferiority.

Table 2. Colonoscopy and bowel preparation outcomes.

		1L PEG+Asc (n=238, ITT)*	**2L PEG+Asc (n=229, ITT)****	**p-value**
Adequate cleansing, n (%)		221 (96.1)	217 (96.4)	0.841
95% CI		92.6-98.0	93.0-98.3	
High quality cleansing rate, n (%)		173 (72.7)	146 (63.8)	**0.038**
95% CI		67.0-78.0	57.5-70.0	
PER PROTOCOL, n		206	226	0.917
Adequate cleansing n (%)		212 (96.4)	196 (96.6)	
95% CI		92.9-98.3	92.9-98.5	
High quality cleansing n (%)		166 (73.5)	131 (63.6)	**0.027**
95% CI		67.7-78.8	57.0-70.2	
BBPS score, median (IQR)				
	Right	3 (3-3)	3 (2-3)	**0.012**
	Transverse	3 (3-3)	3 (3-3)	**0.004**
	Left	3 (3-3)	3 (3-3)	0.077
Experienced difficulty with compliance to low-residue diet, n (%)				0.246
	Very easy	50 (22.6)	56 (27.3)	
	Easy	122 (55.2)	101 (49.3)	
	Fair	37 (16.7)	42 (21.0)	
	Difficult	11 (5.0)	5 (2.4)	
	Very difficult	1 (0.5)	0 (0)	

Table 2. Continued

	1L PEG+Asc (n=238, ITT)*	2L PEG+Asc (n=229, ITT)**	p-value
Experienced difficulty with compliance to liquid diet, n (%)			0.747
Very easy	39 (17.6)	40 (19.5)	
Easy	103 (46.6)	97 (47.3)	
Fair	58 (26.2)	48 (23.4)	
Difficult	16 (7.2)	18 (8.8)	
Very difficult	5 (2.3)	2 (1.0)	
Additional liquids taken, liters, median (IQR)	3 (2-4)	4 (2.4-5.1)	**<0.001**
≥ 75% compliance to laxative, n (%)	218 (99.5)	202 (98.5)	0.356
Physical activity level during bowel preparation, n (%)			0.810
less than usual	85 (38.5)	85 (41.5)	
not changed	131 (59.3)	116 (56.6)	
more than usual	5 (2.3)	4 (2.0)	
Cecal intubation, n (%)	219 (92.8)	219 (96.1)	0.127
Reason for failed cecal intubation, n (%)			0.095
Inadequate bowel preparation	3 (17.6)	0 (0)	
Obstruction/stricture	5 (29.4)	0 (0)	
Pain/discomfort	5 (29.4)	6 (66.7)	
Looping	4 (23.5)	3 (33.3)	
Gloucester comfort scale score, median (IQR)	2 (1-2)	2 (1-2)	0.597
Sedation, n (%)			0.181
No sedation	16 (6.8)	16 (7.0)	
Conscious sedation	207 (87.7)	207 (90.8)	
Deep sedation	13 (5.5)	5 (2.2)	
Need for repeat colonoscopy due to inadequate bowel preparation, n (%)	8 (3.4)	4 (1.8)	0.382
Adenoma detection rate (%)	48.7	51.5	0.396
Adenomas per colonoscopy, median (IQR)	0 (0-1)	0 (0-1)	0.640
Serrated polyp detection rate (%)	17.6	16.2	0.668
Colorectal cancer, n (%)	10 (4.2)	4 (1.8)	0.120
Polyp detection rate (%)	55.9	62.0	0.179
Polyps per colonoscopy, median (IQR)	1 (0-2)	1 (0-2)	0.545
Withdrawal time (minutes), median (IQR)	12 (8-17)	13 (9-19)	**0.047**

* Missing from questionnaires, n=17 (7.1%). ** Missing from questionnaires, n= 24 (10.5%). *Asc; ascorbate, BBPS; Boston Bowel Preparation Scale, CI; confidence interval, IQR; interquartile range, ITT; intention-to-treat, PEG; polyethylene glycol, SD; standard deviation.*

Impact on work and related costs

Among working patients, mean baseline absenteeism was 16.2%, decreasing to 7.9% after colonoscopy (p=0.008), with no significant difference between the groups (Table S1). Additionally, 12.3% of patients in both groups reported reduced productivity (presenteeism) in the past four weeks, with median associated costs of €1390 (Figure 3AB, Table S1).

Unpaid labor loss rates remained stable before and after the procedure in both groups (9.0% before vs. 8.5% after colonoscopy), with mean costs of €1012.3 and €1065.2, respectively. Caregiver involvement was common, with 82.5% of patients requiring assistance, resulting in caregiver absenteeism in 30.6% of cases.

The out-of-pocket costs of patients were modest, with median out-of-pocket costs of €15 (IQR €8.5–€49.5), and no significant difference between study groups (p=0.585). No significant differences between study arms were observed in all categories.

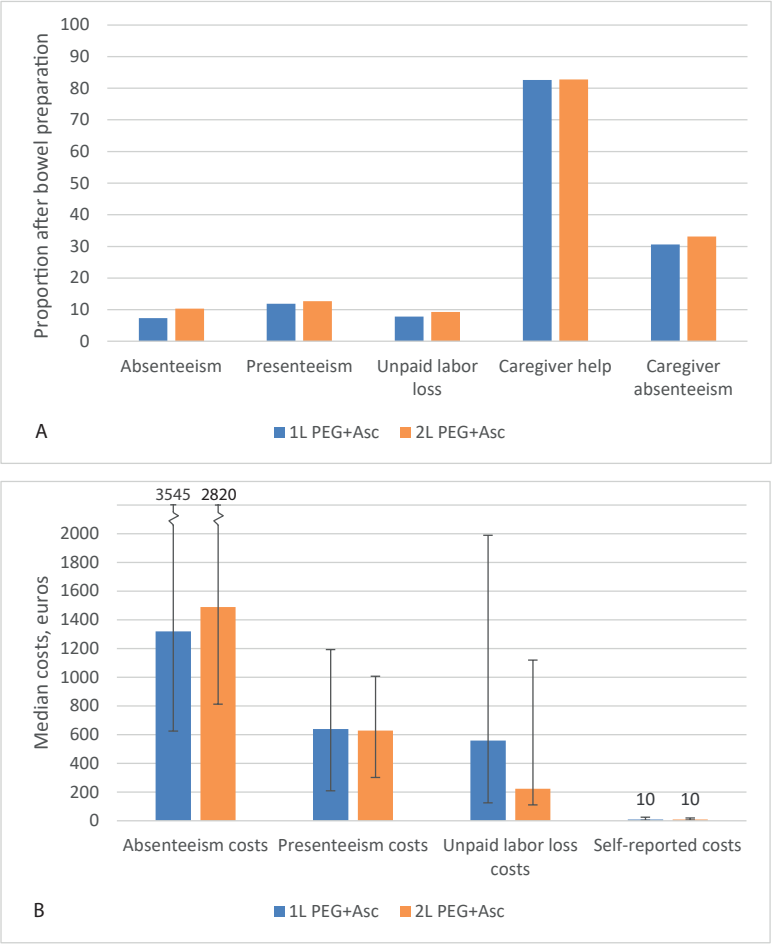


Figure 3. A – impact on work and working productivity in the 1L PEG+Asc and 2L PEG+Asc group using rates of absenteeism and presenteeism in the working subpopulation, and rates of unpaid labor loss, caregiver help, and caregiver absenteeism in the entire study population measured after bowel preparation and colonoscopy. B – Costs of absenteeism, presenteeism, unpaid labor loss, and median self-reported costs in euros.

Tolerability and impact on quality of life

Tolerability

Most respondents drank >75% of the bowel preparation fluid (99.5% in the 1L PEG+Asc group and 98.5% in the 2L PEG+Asc group, p=0.183) (Table S2). Tolerability was rated as "easy" by 51.6% in the 1L PEG+Asc group and 57.1% in the 2L PEG+Asc group (p=0.219). Willingness-to-repeat taking the bowel preparation was higher with 1L PEG+Asc (59.8% vs. 48.3%, p=0.044). Among patients with previous experience with bowel preparation, more patients in the 1L PEG+Asc group found the tolerability better than previous experience (26.9% vs. 9.3%, p<0.001). The total symptom score due to bowel preparation was not significantly different between groups, but patients in the 1L PEG+Asc group reported more often a bad taste (p=0.007) and nausea/vomiting (p=0.006) (Table S2).

In univariable regression analysis, married or living together (OR 1.858), higher endoscopy satisfaction scores (OR 1.067), higher EQ-VAS (OR 1.031), and EQ-Index (OR 1.040) were significantly associated with a higher willingness-to-repeat the bowel preparation. In contrast, 2L PEG+Asc compared to 1L PEG+Asc (OR 0.622), female sex (OR 0.488), fair tolerability (OR 0.178) and difficult compared to good tolerability (OR 0.029) were significantly associated with a lower willingness-to-repeat the bowel preparation. In multivariable analysis, lower tolerability (fair OR 0.225, difficult OR 0.053) and higher symptom score due to bowel preparation (OR 0.852) had a negative impact on willingness-to-repeat the bowel preparation, while patients with a higher endoscopy satisfaction score (OR 1.042), and patients who used 1L PEG+Asc (OR 2.6066 (95% CI 1.592-4.265)) compared to 2L PEG+Asc (OR 0.400) had a higher willingness-to-repeat the bowel preparation (p<0.001) (Table S3).

Health related quality of life

HRQoL measured by the SF-36 was not different before and after bowel preparation and colonoscopy (Figure 4, Table S4). The median PCS and MCS were 51.5 and 53.4, respectively, with no significant difference between study arms. In the 1L PEG+Asc group, physical functioning increased and pain decreased significantly comparing before vs. after colonoscopy (p<0.001 and p=0.002), but this was below the minimal clinically important difference. In the 2L PEG+Asc group, the vitality domain significantly decreased comparing before vs. after colonoscopy (p=0.044), also below the minimal clinically important difference.

In the EQ-5D-5L, patients scored a median baseline EQ-index of 0.89 (1L PEG+Asc) and 0.92 (2L PEG+Asc), both improving to 1.0 after colonoscopy (p<0.001). The

EQ-VAS decreased from 80.41 to 79.95 in the 1L PEG+Asc group and from 81.85 to 81.0 in the 2L PEG+Asc group (p<0.001). These changes were below the minimally clinically important difference. No differences were observed between study arms in subdomains (Figure 4, Table S5), although pain and anxiety in 1L PEG+Asc group and mobility in the 2L PEG+Asc group significantly decreased below the minimally clinically important difference.

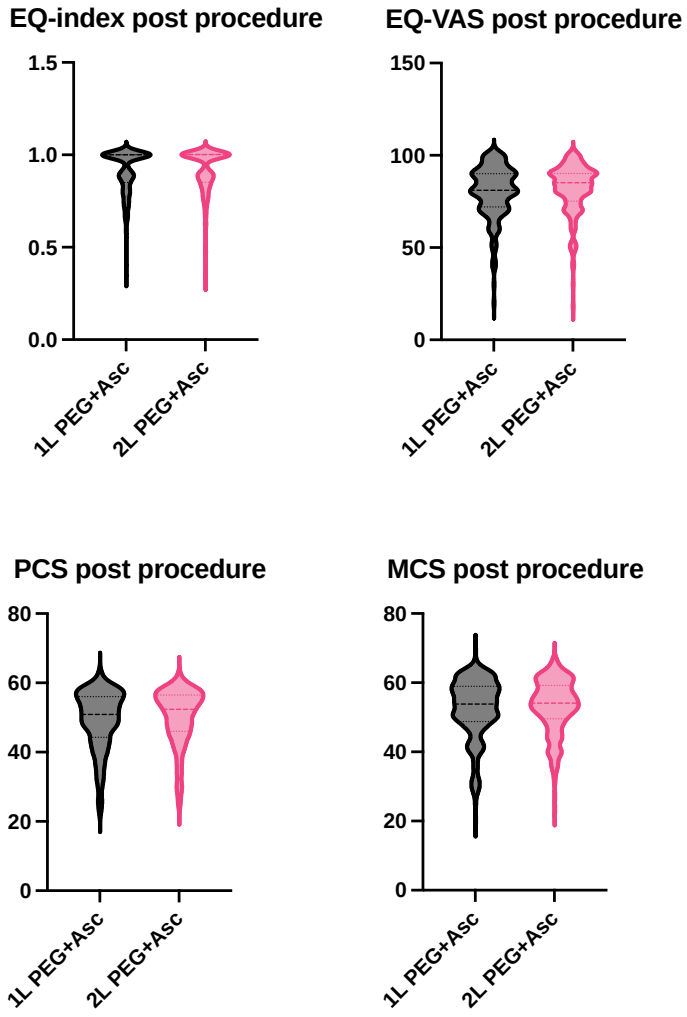


Figure 4. Violin plots of health related quality of life outcomes in the 1L PEG+Asc and 2L PEG+Asc group measured by the EQ index score, EQ VAS score, SF-36 physical component score (PCS), and mental component score (MCS). Higher scores indicate higher quality of life. A wider plot indicates more frequently occurring of the score. Scores were not significantly different between treatment arms.

Endoscopy satisfaction

In general, patients were satisfied about the endoscopic care with a mean score of 85.1 (SD 8.3) (Table S6). Patients were especially satisfied about the endoscopist skills (mean 92.4, SD 8.8) and overall hospital impression (mean 90.5, SD 11.5). Scores were not significantly different between treatment groups.

Adverse events

Sixteen adverse events (AEs) were reported, equally split between the 1L PEG+Asc and 2L PEG+Asc groups. In the 1L PEG+Asc group, all AEs were related to the colonoscopy. In the 2L PEG+Asc group, five AEs were colonoscopy-related, while two involved bowel preparation (both atrial fibrillation). These cases resolved within one day, but the colonoscopies were postponed, and alternative laxatives were used. One serious adverse event, a head injury 30 days post-procedure, was unrelated to the colonoscopy or bowel preparation.

Discussion

While tolerability of bowel preparation is negatively associated with both bowel preparation efficacy and colonoscopy participation, little is known on the impact of bowel preparations with presumed higher tolerability due to lower volumes on patient reported outcomes. Our findings demonstrate that the low-volume preparation (1L PEG+Asc) was non-inferior to the intermediate-volume preparation (2L PEG+Asc) in achieving adequate bowel cleansing (96.1% vs. 96.4%), had higher high-quality cleansing rates (72.7% vs 63.8%), and was associated with greater patient willingness-to-repeat the preparation (59.8% vs. 48.3%). Patient-reported outcomes, including HRQoL and productivity loss, were not significantly different between treatment arms.

Adequate efficacy is a prerequisite for research on patient reported outcomes of bowel preparation. The non-inferiority of 1L PEG+Asc to 2L PEG+Asc in our study is in line with previous studies [18-20]. Our rate of 97% adequately clean colonoscopies is on the high end of the spectrum. This could be due to the exclusion of patients at risk for inadequate bowel preparation in our study; however, in other studies this patient group was also excluded [20-22]. The high compliance in both groups could also be a contributing factor. Furthermore, we included a low-residue diet and split-dose protocol, combined with enhanced instructions, all known to improve efficacy and adherence [1].

Our results demonstrated that tolerability is a critical factor in bowel preparation because it directly impacts patient compliance and willingness-to-repeat, with patients who found the preparation 'difficult' being 94.7% less likely (OR 0.053) to be willing to repeat the preparation compared to those who found it 'easy'. Repici et al. demonstrated that 1L PEG+Asc had significantly higher compliance compared to 4L PEG, resulting in more effective bowel cleansing [20]. Similarly, Bednarska et al. concluded that patients were more willing to repeat a regimen of 1L PEG+Asc compared to 2L PEG+Asc [23], consistent with our findings (59.8% vs 48.3%, P=0.044). Despite a more frequent bad taste (p=0.007) and nausea/vomiting (p=0.006) in the 1L PEG+Asc group, overall tolerability rates were similar between groups**.** Patients who had a prior experience with taking bowel preparation found the 1L PEG+Asc regimen significantly more tolerable than the bowel preparation taken previously, as also reported by Bednarska et al [23]. This may well suggest that volume is an important factor in negatively affecting patient experience, even more than taste or side effects. Therefore, the superior experience of 1L PEG+Asc contributes to the efficacy of the bowel preparation. Furthermore, management of discomfort and anxiety by effective communication likely also plays an important role in improving compliance and overall patient experience [24].

Due to the several instructions patients have to adhere to during bowel preparation, and possible discomfort and anxiety [2], we hypothesized that these could also have an impact on HRQoL. We found comparable HRQoL scores to the general Dutch population, with no clinically relevant difference in both groups. Andronis et al. reported that patients undergoing colonoscopy experienced a temporary decrease in HRQoL due to bowel preparation discomfort [25]. Contrarily, Niv et al. observed that while patients had some discomfort during bowel preparation, the overall HRQoL scores remained relatively stable after colonoscopy, in line with our results [12]. As patients report some discomfort during bowel preparation, traditional HRQoL measures unlikely are able to fully capture impact on patients [26]. However, given the short duration of the bowel preparation process, long-term HRQoL impact by bowel preparation is likely negligible.

Bowel preparation can have a significant impact on patients' work as they need to take time off to complete the bowel preparation and undergo the colonoscopy. Fuccio et al. reported in a prospective cohort (n=1137) that absenteeism or presenteeism due to bowel preparation or symptoms after colonoscopy was 30.5% in patients undergoing colonoscopy [8]. Our study's lower rate of both absenteeism (7.9%) and presenteeism (12.3%) may be due to the use of a split-dose regimen and a difference in laxatives used (4L and 2L vs. our study's 1L and 2L). Other studies

observed absenteeism in 20-32% of patients due to bowel preparation, with an association between a higher symptom burden and absenteeism [27, 28]. Additionally, our study observed that the loss of unpaid labor was 8.5% after colonoscopy, which has poorly been investigated so far. Both direct (hospital) and indirect, nonmedical costs are needed for a clearer understanding of the total expenses associated with bowel preparation [29, 30]. Our findings may thus offer a perspective on these costs from both patient and societal viewpoints.

Ongoing interest is needed for bowel preparation tolerability, to ensure adequate preparation but also to minimize the possible negative impact on patients. This includes a split-dose regimen, diet liberations such as low-residue diets with shorter durations [31, 32], and also developing other low-volume laxatives. This will likely reduce experienced barriers to colonoscopy and CRC screening [33]. Furthermore, our results on HRQoL and patient costs using validated instruments may inform healthcare policies and allow them to make informed decisions that improve CRC screening participation while minimizing impact on patients.

Our study has several strengths. In contrast to previous studies, our multicenter randomized design reduces the risk of bias. Additionally, we used validated instruments to increase generalizability and comparability. To minimize possible remaining bias, we used multiple imputations to compensate for missing information, although the response rate was already nearly 90%. The use of clinical and socio-economic variables enabled us to correct for potential confounders and provided more optimal information on the impact of bowel preparation on patients and relatives.

We also acknowledge some limitations. Our study population consisted of predominantly Dutch, high-educated patients. This may limit generalizability to patients with other ethnicities or lower socio-economic levels, while these patients are known to be at a slightly higher risk for inadequate bowel preparation [34, 35]. Lastly, overlapping recall periods of the SF-36 and IPCQ limit the ability to accurately capture the true change before and after bowel preparation. However, since other parameters such as the EQ-5D-5L with shorter recall periods confirmed the results of the SF-36, our study still provides valuable insights.

Conclusion

In conclusion, low-volume 1L PEG+Asc bowel preparation is an effective and tolerable alternative to intermediate-volume 2L PEG+Asc preparation, with high patient satisfaction and minimal impact on quality of life and productivity. Prioritizing tolerability of bowel preparation is essential to increase adequate bowel cleansing and reduce barriers to colonoscopy.

References

1. Hassan C, East J, Radaelli F et al. Bowel preparation for colonoscopy: European Society of Gastrointestinal Endoscopy (ESGE) Guideline - Update 2019. Endoscopy 2019; 51: 775-794. doi:10.1055/a-0959-0505
2. McLachlan S-A, Clements A, Austoker J. Patients' experiences and reported barriers to colonoscopy in the screening context—a systematic review of the literature. Patient education and counseling 2012; 86: 137-146.
3. Zhu Y, Li X, Hu Y et al. Nonadherence to Referral Colonoscopy After Positive Fecal Immunochemical Test Results Increases the Risk of Distal Colorectal Cancer Mortality. Gastroenterology 2023; 165: 1558-1560.e1554. doi:10.1053/j.gastro.2023.08.013
4. Spadaccini M, Frazzoni L, Vanella G et al. Efficacy and Tolerability of High- vs Low-Volume Split-Dose Bowel Cleansing Regimens for Colonoscopy: A Systematic Review and Meta-analysis. Clin Gastroenterol Hepatol 2020; 18: 1454-1465 e1414. doi:10.1016/j.cgh.2019.10.044
5. McLachlan SA, Clements A, Austoker J. Patients' experiences and reported barriers to colonoscopy in the screening context--a systematic review of the literature. Patient Educ Couns 2012; 86: 137-146. doi:10.1016/j.pec.2011.04.010
6. Sharara AI, El Reda ZD, Harb AH et al. The burden of bowel preparations in patients undergoing elective colonoscopy. United European Gastroenterol J 2016; 4: 314-318. doi:10.1177/2050640615594550
7. Collatuzzo G, Fuccio L, Frazzoni L et al. Determinants of symptom burden related to bowel preparation for colonoscopy. Dig Liver Dis 2022. doi:10.1016/j.dld.2022.06.005
8. Fuccio L, Collatuzzo G, Frazzoni L et al. Impact of colonoscopy on working productivity: a prospective multicenter observational study. Gastrointest Endosc 2021. doi:10.1016/j.gie.2021.11.039
9. Bronzwaer MES, Depla A, van Lelyveld N et al. Quality assurance of colonoscopy within the Dutch national colorectal cancer screening program. Gastrointest Endosc 2019; 89: 1-13. doi:10.1016/j.gie.2018.09.011
10. Bouwmans C, Krol M, Severens H et al. The iMTA Productivity Cost Questionnaire: A Standardized Instrument for Measuring and Valuing Health-Related Productivity Losses. Value Health 2015; 18: 753-758. doi:10.1016/j.jval.2015.05.009
11. Ware JE, Kosinski M, Dewey JE: How to score version 2 of the SF-36 health survey. Lincoln, RI: QualityMetric Incorporated; 2000
12. Niv Y, Bogolavski I, Ilani S et al. Impact of colonoscopy on quality of life. Eur J Gastroenterol Hepatol 2012; 24: 781-786. doi:10.1097/MEG.0b013e328352deff
13. Mahabaleshwarkar R, Khanna R, West-Strum D et al. Association between health-related quality of life and colorectal cancer screening. Popul Health Manag 2013; 16: 178-189. doi:10.1089/pop.2012.0044
14. Patel M, Staggs E, Thomas CS et al. Development and validation of the Mayo Clinic Bowel Prep Tolerability Questionnaire. Dig Liver Dis 2014; 46: 808-812. doi:10.1016/j.dld.2014.05.020
15. Veldhuijzen G, de Jong MJP, Roosen CM et al. The gastrointestinal endoscopy satisfaction questionnaire captures patient satisfaction as a key quality indicator of gastrointestinal endoscopy. Eur J Gastroenterol Hepatol 2020; 32: 832-837. doi:10.1097/MEG.0000000000001764
16. M MV, K MV, S MAAE et al. Dutch Tariff for the Five-Level Version of EQ-5D. Value Health 2016; 19: 343-352. doi:10.1016/j.jval.2016.01.003

17. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Medical care 2003; 41: 582-592.

18. López-Jamar JME, Gorjão R, Cotter J et al. Bowel cleansing effectiveness and safety of 1L PEG + Asc in the real-world setting: Observational, retrospective, multicenter study of over 13000 patients. Endosc Int Open 2023; 11: E785-e793. doi:10.1055/a-2125-0025

19. Bednarska O, Nyhlin N, Schmidt PT et al. The Effectiveness and Tolerability of a Very Low-Volume Bowel Preparation for Colonoscopy Compared to Low and High-Volume Polyethylene Glycol-Solutions in the Real-Life Setting. Diagnostics 2022; 12: 1155.

20. Repici A, Spada C, Cannizzaro R et al. Novel 1-L polyethylene glycol + ascorbate versus high-volume polyethylene glycol regimen for colonoscopy cleansing: a multicenter, randomized, phase IV study. Gastrointest Endosc 2021; 94: 823-831 e829. doi:10.1016/j.gie.2021.04.020

21. Bisschops R, Manning J, Clayton LB et al. Colon cleansing efficacy and safety with 1 L NER1006 versus 2 L polyethylene glycol + ascorbate: a randomized phase 3 trial. Endoscopy 2019; 51: 60-72. doi:10.1055/a-0638-8125

22. DeMicco MP, Clayton LB, Pilot J et al. Novel 1 L polyethylene glycol-based bowel preparation NER1006 for overall and right-sided colon cleansing: a randomized controlled phase 3 trial versus trisulfate. Gastrointest Endosc 2018; 87: 677-687 e673. doi:10.1016/j.gie.2017.07.047

23. Bednarska O, Nyhlin N, Schmidt PT et al. The Effectiveness and Tolerability of a Very Low-Volume Bowel Preparation for Colonoscopy Compared to Low and High-Volume Polyethylene Glycol-Solutions in the Real-Life Setting. Diagnostics (Basel) 2022; 12. doi:10.3390/diagnostics12051155

24. Shafer LA, Walker JR, Waldman C et al. Predictors of patient reluctance to wake early in the morning for bowel preparation for colonoscopy: a precolonoscopy survey in city-wide practice. Endosc Int Open 2018; 06: E706-E713. doi:10.1055/s-0044-102298

25. Andronis L, Waugh N, Zanganeh M et al. Impact of colonoscopy on health-related quality of life: findings from the RECEDE study. Health Qual Life Outcomes 2024; 22: 49. doi:10.1186/s12955-024-02262-x

26. Swan JS, Hur C, Lee P et al. Responsiveness of the testing morbidities index in colonoscopy. Value Health 2013; 16: 1046-1053. doi:10.1016/j.jval.2013.07.008

27. Ko CW, Riffle S, Shapiro JA et al. Incidence of minor complications and time lost from normal activities after screening or surveillance colonoscopy. Gastrointest Endosc 2007; 65: 648-656. doi:10.1016/j.gie.2006.06.020

28. Dong MH, Kalmaz D, Savides TJ. Missed work related to mid-week screening colonoscopy. Dig Dis Sci 2011; 56: 2114-2119. doi:10.1007/s10620-010-1545-2

29. Heitman SJ, Au F, Manns BJ et al. Nonmedical Costs of Colorectal Cancer Screening With the Fecal Occult Blood Test and Colonoscopy. Clinical Gastroenterology and Hepatology 2008; 6: 912-917. e911. doi:https://doi.org/10.1016/j.cgh.2008.03.006

30. Rex DK, Imperiale TF, Latinovich DR et al. Impact of bowel preparation on efficiency and cost of colonoscopy. Am J Gastroenterol 2002; 97: 1696-1700. doi:10.1111/j.1572-0241.2002.05827.x

31. Nguyen DL, Jamal MM, Nguyen ET et al. Low-residue versus clear liquid diet before colonoscopy: a meta-analysis of randomized, controlled trials. Gastrointest Endosc 2016; 83: 499-507 e491. doi:10.1016/j.gie.2015.09.045

32. Wang F, Huang X, Wang Z et al. One-day versus three-day low-residue diet bowel preparation regimens before colonoscopy: a meta-analysis of randomized controlled trials. J Gastroenterol Hepatol 2024. doi:10.1111/jgh.16466

33. Jain S, Galoosian A, Wilhalme H et al. Individuals With a Family History of Colorectal Cancer Warrant Tailored Interventions to Address Patient-Reported Barriers to Screening. Clinical and Translational Gastroenterology 2023; 14: e00583. doi:10.14309/ctg.0000000000000583

34. Serper M, Gawron AJ, Smith SG et al. Patient factors that affect quality of colonoscopy preparation. Clin Gastroenterol Hepatol 2014; 12: 451-457. doi:10.1016/j.cgh.2013.07.036

35. Nguyen DL, Wieland M. Risk factors predictive of poor quality preparation during average risk colonoscopy screening: the importance of health literacy. J Gastrointestin Liver Dis 2010; 19: 369-372.

Supporting information

Table S1. Patient costs, absenteeism and presenteeism according to the IMTA productivity cost questionnaire, pre and post bowel preparation and colonoscopy.

	1L PEG+Asc (n=238, ITT)Missing n=14 (5.9%)			2L PEG+Asc (n=227, ITT)Missing n=22 (9.6%)		
	Pre	**Post**	**p**	**Pre**	**Post**	**p**
Payed job, n (%)	121 (54.0)			107 (51.7)		
Amount of work per week, mean (SD)		n/a			n/a	
Hours	31.8 (11.7)			32.6 (11.4)		
Days	4.2 (1.1)			4.3 (1.1)		
Absenteeism in past 4 weeks			0.688	22 (20.6)		0.070
n (%)	15 (12.4)	16 (7.3)		5 (3.8-9)	21 (10.3)	
Number of days, median (IQR)	5 (2-16)	9 (3.3-20)		40 (22.2-53.8)	7 (4-18)	
Number of working hours, median (IQR)	40 (16-80)	38 (18-140)		1390 (770-	42.9 (23.4-124)	
Costs in euros, median (IQR)	1390 (556-2780)	1320 (626-4865)		1870)	1489 (813-4309)	
Presenteeism						
n (%)		26 (11.9)	1.0		26 (12.7)	1.0
Number of days, median (IQR)		6 (3-10.5)			(3-10)	
Working efficacy 0-10		5 (4-8)			6 (4-7)	
Number of working hours with reduced efficacy, median (IQR)		-18.4 (-34.4, -6)			-18.1 (-29, -8.7)	
Costs in euros, median (IQR)		-639 (-1193, -209)			628 (-1008, -303)	
Unpaid labor loss						
n (%)	20 (8.9)	17 (7.8)	0.629	19 (9.2)	19 (9.3)	1.0
Number of days, median (IQR)	8.5 (3.3-18.8)	20 (6.5-24)		10 (4-21)	6 (4-16)	
Number of working hours lost, median (IQR)	33.5 (6.8-100)	40 (9-142)		40 (6-100	16 (8-80)	
Costs in euros, median (IQR)	469 (94.5-1400)	560 (126-1988)		560 (84-1400)	224 (112-1120)	
Caregiver help						
n (%)		180 (82.6)			169 (82.8)	0.977
Caregiver absenteeism, n (%)		55 (30.6)			56 (33.1)	0.721

Table S1. Continued

	1L PEG+Asc (n=238, ITT)Missing n=14 (5.9%)			2L PEG+Asc (n=227, ITT)Missing n=22 (9.6%)		
	Pre	**Post**	**p**	**Pre**	**Post**	**p**
Self-reported costs						0.951
None, n (%)		144 (61.2)			126 (55.5)	
Medication, n (%)		8 (3.4)			6 (2.6)	
Amount, mean (SD)		*13 (27.65)*			*18 (21.75)*	
OTC medication, n (%)		3 (1.3)			5 (2.2)	
Amount, mean (SD)		*5.67 (1.16)*			*13.6 (12.97)*	
Travel costs, n (%)		64 (27.0)			63 (27.8)	
Amount, mean (SD)		*15 (12.51)*			*12.92 (11.27)*	
Other, n(%)		16 (6.8)			22 (9.6)	
Amount, median (IQR)		*6 (1.5-36.3)*			*10 (5-15)*	
Total self-reported costs, median (IQR)		10 (5-25)			10 (5.9-20)	
Sum costs of absenteeism, presenteeism, unpaid labor loss, and self-reported costs, median (IQR)		17 (9-65.6)			15 (7-39)	0.585

Tolerability

Table S2. Results of the Mayo Florida Bowel preparation tolerability questionnaire

	1L PEG+Asc (n=238, ITT) missing n=19 (8.4)	**2L PEG+Asc (n=229, ITT) missing n=24 (10.5)**	**p-value**
Number of bowel movements in week prior to prep, n (%)			0.722
3 or less/week	18 (8.2)	20 (8.7)	
4-8/week	129 (58.9)	124 (60.5)	
9 or more/week	72 (32.9)	61 (26.6)	
Laxative left after drinking to best effort, n (%)			0.341
Less than 25%	218 (99.5)	202 (98.5)	
25% - 50% left	1 (0.5)	1 (0.5)	
50% - 75% left	0	2 (1.0)	
75% or more left	0	0	
Tolerability, n (%)			0.219
Easy	113 (51.6)	115 (57.1)	
Acceptable	63 (28.8)	63 (30.7)	
Somewhat difficult	27 (12.3)	15 (7.3)	
Very difficult	16 (7.3)	10 (4.9)	
Willingness-to-repeat this prep, n (%)			**0.044**
not willing at all	17 (7.8)	16 (7.8)	
somewhat willing	71 (32.4)	90 (43.9)	
mostly willing	131 (59.8)	99 (48.3)	
In case of experienced difficulties in tolerability, do you think it was due to your current health issues? n, (%)			0.550 0.304 for trend
Yes	6 (2.7)	2 (1.0)	
No	52 (23.7)	46 (22.4)	
N/A (no difficulties)	117 (53.4)	111 (54.1)	
N/A (no health issues)	44 (20.1)	46 (22.4)	

Table S2. Continued

		1L PEG+Asc (n=238, ITT) missing n=19 (8.4)	2L PEG+Asc (n=229, ITT) missing n=24 (10.5)	p-value
Symptoms, n (%)				
Bad taste				**0.007**
	None	65 (29.7)	85 (41.5)	
	Mild	92 (42.0)	79 (38.5)	
	Moderate	4 (18.7)	35 (17.1)	
	Severe	21 (9.6)	6 (2.6)	
Feeling full				
	None	63 (28.8)	42 (20.5)	0.166
	Mild	87 (39.7)	81 (39.5)	
	Moderate	55 (25.1)	65 (31.7)	
	Severe	14 (6.4)	17 (8.8)	
Lack of sleep				
	None	110 (50.2)	80 (39.0)	0.091
	Mild	59 (26.9)	61 (29.8)	
	Moderate	35 (16.0)	41 (20.0)	
	Severe	15 (6.8)	23 (11.2)	
Nausea/vomiting				
	None	141 (64.4)	163 (79.5)	**0.006**
	Mild	51 (23.3)	30 (14.6)	
	Moderate	15 (6.8)	7 (3.1)	
	Severe	12 (5.5)	5 (2.4)	
Bloating				
	None	67 (30.6)	64 (31.2)	0.749
	Mild	99 (45.2)	87 (42.4)	
	Moderate	44 (20.1)	41 (20.0)	
	Severe	9 (4.1)	13 (6.3)	
Abdominal pain/cramps				
	None	104 (47.5)	98 (47.8)	0.960
	Mild	84 (38.4)	75 (36.6)	
	Moderate	25 (11.4)	25 (12.2)	
	Severe	6 (2.7)	7 (3.4)	
Headache				
	None	155 (70.8)	149 (72.7)	0.212
	Mild	42 (19.2)	43 (21.0)	
	Moderate	15 (6.8)	12 (5.9)	
	Severe	7 (3.2)	1 (0.5)	
Symptom score*, mean (SD)		5.6 (3.8)	5.5 (3.4)	0.743
What is the tolerability of the laxative used compared to the previous experience? N (%)		N=105	N=98	**<0.001**
	Worse	12 (5.5)	13 (6.3)	
	About the same	34 (15.5)	66 (32.2)	
	Better	59 (26.9)	19 (9.3)	
	N/A (first colonoscopy)	114 (52.1)	107 (52.2)	

Table S3. Factors associated with willingness to repeat in univariate and multivariate logistic regression. OR, Odds ratio; CI, confidence interval; MCS, mental component score post-colonoscopy; MBTQ, Mayo bowel preparation tolerability questionnaire;

Predictor	OR univariate (95% CI)	p-value	OR multivariate (95% CI)	p-value	OR final model (95% CI)	p-value
2L PEG+Asc compared to 1L PEG+Asc	0.622 (0.432-0.896)	**0.011**	0.367 (0.212-0.637)	**<0.001**	0.400(0.239-0.669)	<0.001
Age	1.008 (0.992-1.024)	0.343				
Sex female	0.488 (0.336-0.709)	**<0.001**	1.094 (0.610-1.960)	0.763		
No prior experience	0.822 (0.572-1.183)	0.292				
Colonoscopy indication		**0.161**		0.189		
Screening	Reference		Reference			
Surveillance	1.301 (0.834-2.028)		1.544 (0.610-1.960)			
Diagnostic	0.838 (0.537-1.305)		1.724 (0.835-3.560)			
Education level	Reference	0.765				
None	0.661 (0.182-2.401)					
Secondary school	0.724 (0.197-2.656)					
University						
Married/living together	1.858 (1.118-3.088)	**0.017**	1.598 (0.787-3.245)	0.194		
No paid occupation	0.752 (0.515-1.100)	**0.142**	0.691 (0.401-1.192)	0.184		
No excellent prep	0.928 (0.628-1.371)	0.707				
Gloucester comfort scale score	Reference	**0.077**	Reference	0.893		
1	0.829 (0.546-1.260)		0.959 (0.520-1.767)			
2	0.577 (0.359-0.929)		0.837 (0.393-1.786)			
3/4/5						
Polyp detected	1.355 (0.936-1.962)	**0.108**	0.975(0.540-1.760)	0.933		
No absenteeism	1.616 (0.818-3.192)	**0.167**	1.370 (0.457-4.108)	0.574		
No presenteeism	1.441 (0.805-2.579)	0.219				

Table S3. Continued

Predictor	OR univariate (95% CI)	p-value	OR multivariate (95% CI)	p-value	OR final model (95% CI)	p-value
Tolerability (MBTQ) *good* *fair* *difficult*	Reference 0.178 (0.111-0.286) 0.029 (0.012-0.071)	**<0.001**	Reference 0.208 (0.115-0.375) 0.054 (0.019-0.155)	**<0.001**	Reference 0.225 (0.130-0.389) 0.053 (0.020-0.143)	<0.001
Symptom score	0.726 (0.673-0.783)	**<0.001**	0.826 (0.744-0.918)	**<0.001**	0.852 (0.779-0.932)	<0.001
Total endoscopy satisfaction score	1.067 (1.039-1.095	**<0.001**	1.054 (1.015-1.095)	**0.006**	1.042 (1.008-1.076)	0.015
MCS	1.021 (0.998-1.045)	**0.077**	0.968 (0.932-1.004)	0.083		
EQ VAS	1.031 (1.016-1.047)	**<0.001**	1.004 (0.977-1.031)	0.770		
EQ Index	1.040 (1.021-1.059)	**<0.001**	1.009 (0.989-1.029)	0.404		

Health related quality of life

Table S4. Results of quality-of-life questionnaire SF-36v2 before (pre) and after (post) bowel preparation, Norm-based scores

Category Median (IQR)	**1L PEG+Asc (n=238, ITT)Missing, n=15**				**2L PEG+Asc (n=227, ITT)Missing, n=23**				**P-value of change**
	Pre	**Post**	**p-value**	**Change per patient**	**Pre**	**Post**	**p-value**	**Change per patient**	**pooled**
Physical functioning	53.1 (46.5-56.9)	53.1 (44.3-57.5)	**<0.001**	0 (0-2.2)	55.3 (48.7-57.5)	55.3 (46.5-57.5)	0.183	0 (0-2.2)	**0.022**
Role limitation physical	49.6 (44.4-56.5)	50.5 (44.4-56.5)	0.606	0 (-1.7-1.7)	51.3 (46.2-56.5)	51.3 (44.4-56.5)	0.194	0 (-1.7-1.7)	0.383
Pain	49.6 (44.5-60.7)	53.9 (44.5-60.7)	**0.002**	0 (-4.3-0)	53.9 (44.5-60.7)	53.9 (44.5-60.7)	0.204	0 (-0.9-0)	0.626
General health	50.6 (42.4-55.5)	48.2 (43.4-55.5)	0.699	0 (-2.4-3.1)	50.6 (43.4-55.5)	50.6 (43.4-55.4)	0.681	0 (-2.4-2.4)	0.917
Vitality	50.1 (43.6-56.6)	53.3 (46.8-56.6)	0.396	0 (-3.2-3.2)	53.3 (43.6-56.6)	53.3 (46.8-56.6)	**0.044**	0 (-3.2-3.2)	0.373
Social functioning	51.6 (46-57.1)	51.6 (46-57.1)	0.276	0 (0-5.6)	57.1 (46-57.1)	57.1 (46-57.1)	0.084	0 (0-0)	0.699
Role limitation emotional	55.4 (47.8-55.4)	52.8 (45.2-55.4)	0.105	0 (0-2.5)	54.1 (47.8-55.4)	52.8 (47.8-55.4)	0.304	0 (0-2.5)	0.501
Mental health	54.7 (49-57.6)	54.7 (49-57.6)	0.212	0 (-2.9-2.9)	54.7 (49-57.6)	54.7 (49-60.5)	0.846	0 (-2.9-2.9)	0.320
Physical component score	50.4 (45.1-55.3)	50.9 (44.3-56.0)	0.293	-0.3 (-2.2-1.9)	52.5 (46.4-56.3)	52.4 (46.2-56.5)	0.895	0.1 (-2.0-1.7)	0.823
Mental component score	53.2 (46.5-57.7)	52.5 (47.3-57.6)	0.492	0.2 (-2.9-3.1)	53.5 (47.2-57.3)	52.6 (48.1-57.8)	0.828	-0.1 (-2.3-2.4)	0.430

Table S5. EQ-5D-5L results before (pre) and after (post) bowel preparation.

Category median (IQR)	1L PEG+Asc (n=238, ITT)				2L PEG+Asc (n=227, ITT)				P-value of change
	Pre	Post	Change mean (SD)	P pre vs post	Pre	Post	Change mean (SD)	P pre vs post	
Mobility	1 (1-1)	1 (1-1)	0.01 (0.59)	0.969	1 (1-1)	1 (1-1)	0.08 (0.42)	**0.008**	0.171
Selfcare	1 (1-1)	1 (1-1)	-0.02 (0.21)	0.096	1 (1-1)	1 (1-1)	0.01 (0.28)	1.0	0.233
Activity	1 (1-1)	1 (1-1)	-0.08 (0.71)	0.105	1 (1-1)	1 (1-1)	-0.01 (0.53)	0.894	0.257
Pain	1 (1-2)	1 (1-2)	0.17 (0.79)	**0.002**	1 (1-2)	1 (1-2)	0.08 (0.60)	0.075	0.119
Anxiety	1 (1-1)	1 (1-1)	0.14 (0.59)	**0.001**	1 (1-1)	1 (1-1)	0.06 (0.52)	0.067	0.198
VAS, mean (SD)	80.41 (14.0)	79.95 (13.2)	-0.29 (12.6)	**<0.001**	81.85 (13.4)	81.0 (13.2)	-0.60 (9.8)	**<0.001**	0.963
EQ Index	0.89 (0.82-1.0)	1 (0.85-1.0)	-0.03 (0.13)	**<0.001**	0.92 (0.85-1.0)	1 (0.85-1.0)	-0.02 (0.10)	**<0.001**	0.440

Endoscopy satisfaction

Table S6. Results of the Dutch gastrointestinal endoscopy satisfaction questionnaire. Higher scores indicate higher satisfaction, 0=worst and 100=best.

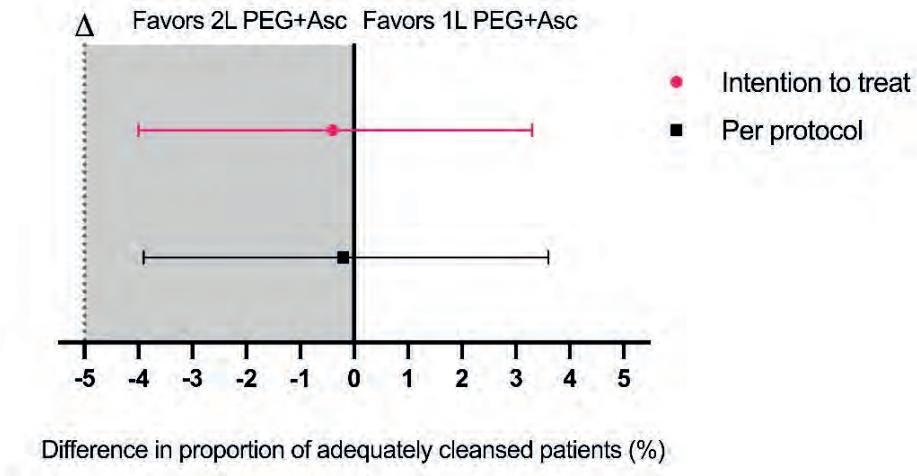
mean scores (SD)	Complete study population, n=42 missing n=77	1L PEG+Asc (n=238, ITT) Missing, n=39 (16.4%)	2L PEG+Asc (n=227, ITT) Missing, n=38 (16.6%)	p-value between study arms
Skills and satisfaction	92.4 (8.8)	92.4 (8.5)	92.5 (9.0)	0.849
Information before endoscopy	85.4 (10.1)	85.3 (10.3)	85.4 (10.0)	0.941
Pain and discomfort	86.3 (12.6)	86.4 (12.6)	86.1 (12.7)	0.912
Information after endoscopy	84.0 (10.3)	84.0 (10.7)	84.0 (10.0)	0.828
Hospital	90.5 (11.5)	89.4 (12.3)	91.6 (10.5)	0.053
Total	85.1 (8.3)	84.9 (8.5)	85.2 (8.2)	0.405

Published article

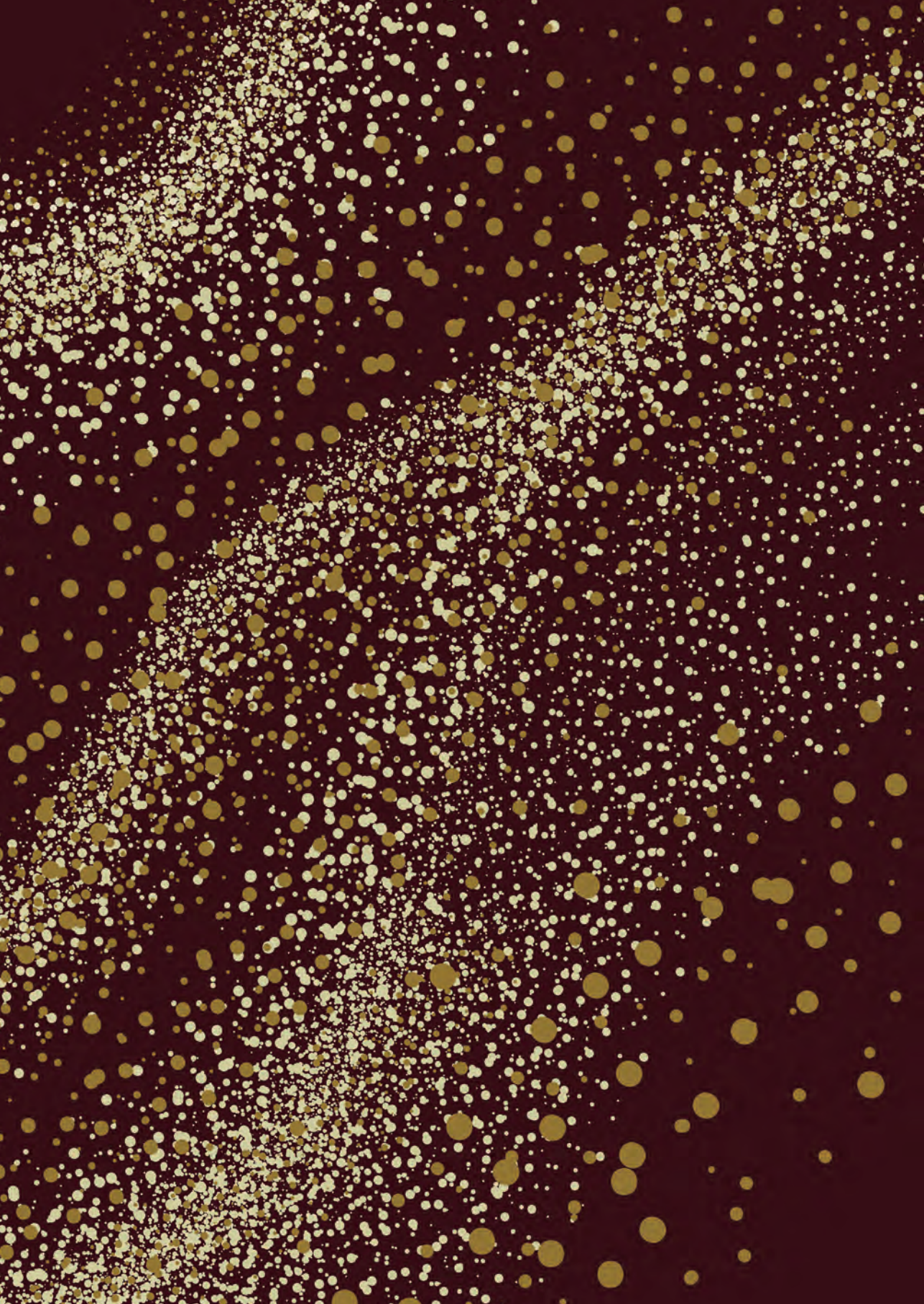
The published version of this manuscript can be found at Endoscopy. 2025 Sep 3. doi: 10.1055/a-2695-0994.

The following major changes have been made to this manuscript in the published version:

- Updated wording in methods to specify why certain procedures and statistical methods were chosen.
- Clarified definition of primary outcome to include bowel cleansing efficacy in terms of non-inferiority in rate of adequate bowel cleansing of 1L PEG+Asc to 2L PEG+Asc in *Abstract* and *Methods*.
- Addition of between group difference of -0.4% (95%CI -4.0-3.3) (ITT) and -0.2% (-3.9-3.6) (PP) in the *Abstract* and *Results* section and updated figure representing the primary outcome (figure 2):



- Omission of SF-36 results due to the plurality of outcomes and overlapping recall period in first and second questionnaire and updated methods and results (including figure 4) accordingly.
- Expansion of the limitations section to include a deviation from ESGE guidelines of the 2-day low residue diet and emphasize the exploratory nature of the secondary analyses performed.



Chapter 6

An intraprocedural bowel cleansing system for difficult-to-prepare patients – a multicenter prospective feasibility study

Milou L.M. van Riswijk, MD, Kelly E. van Keulen, MD, Helmut Neumann, MD, Peter D. Siersema, MD, PhD

United European Gastroenterol J. 2024 Feb;12(1):56-65. doi: 10.1002/ueg2.12501.

Abstract

Background: Adequate bowel preparation is a prerequisite for colonoscopy. However, up to 20% of colonoscopies remain inadequately prepared. Risk factors for inadequate bowel preparation often overlap with those of failed cecal intubation. This study aimed to assess the feasibility of an intraprocedural bowel cleansing system (BCS) in patients with a history of inadequate bowel preparation.

Methods: Patients (n=44) with a history of inadequate bowel preparation in the past two years were included. After a limited preparation with 300mL split-dose sodium picosulfate magnesium citrate, additional cleansing during colonoscopy was performed with the BCS. Primary outcome was adequate bowel preparation using the Boston Bowel Preparation Scale (BBPS). Secondary outcomes included cecal intubation rate (CIR), procedure times, usability, and safety.

Results: Median BBPS increased from 1-2-2 (IQR 1-2) to 3-3-3 (IQR) ($P<0.0001$), with 31.8% and 88.6% of patients adequately prepared before and after using the BCS, respectively ($p<0.0001$). CIR was 88.6%. Reasons for incomplete colonoscopy were looping (n=2), technical failure (n=1), relative stricture (n=1), and residual feces (n=1). In patients with complete colonoscopy, the adequate cleansing rate was 97.5%. Median total procedure time was 26 minutes, of which 5.3 minutes were spent on cleaning. General ease of use was scored with a median of 4 out of 5, representing "as good as conventional colonoscopy". No serious adverse events occurred.

Conclusions: Adequate bowel cleaning can be achieved with an intraprocedural BCS in patients with a history of inadequate bowel preparation which may reduce repeat colonoscopies and clinical admissions for bowel preparation. However, since these patients more frequently have a complicated anatomy (surgical scarring, diverticulosis, etc.), adequate patient selection is advised to avoid incomplete procedures.

ClinicalTrials.gov Identifier: NCT04700410

Study highlights

What is known

- Adequate bowel preparation is essential to perform high-quality colonoscopy.
- Patients known with inadequate bowel preparation are at risk for recurrent inadequate bowel preparation.
- The optimal approach to patients at risk for inadequate bowel preparation is not yet known.

What is new here

- A novel bowel cleansing system could be used during colonoscopy to achieve adequate bowel preparation in patients at risk for inadequate preparation.
- The bowel cleansing system slightly increases the scope diameter and stiffness, which may restrict cecal intubation.
- In selected patients, this bowel cleansing system may reduce the number of repeat colonoscopies and clinical admissions for bowel preparation.

Introduction

High quality colonoscopy reduces the incidence and mortality of colorectal cancer (CRC) (1, 2), but cannot be performed without adequate bowel preparation (3, 4). Inadequate bowel preparation precludes meticulous mucosal inspection and is therefore associated with a lower polyp detection rate (5-8). Other consequences of inadequate bowel preparation include a lower cecal intubation rate, higher complication rate, longer procedure times, higher costs because of repeat procedures, and lower patient satisfaction (3, 4, 9-12).

Unfortunately, inadequate bowel preparation has been reported in up to 20% of colonoscopies (13). Several risk factors have been identified, related to factors interfering with bowel motility (i.e., opioid use, diabetes, constipation), or to non-compliance of laxative use, such as poor health literacy (14). Previous poor preparation is one of the most important and easy identifiable risk factors for inadequate bowel preparation (15, 16). Identifying patients at risk for insufficient preparation enables taking timely additional measures to optimize preparation.

However, the optimal approach to patients at risk for inadequate bowel preparation is unclear (3, 16), because these patients are frequently excluded in clinical trials that assess a cleansing strategy and guideline recommendations are lacking. In daily practice, patients are often prescribed an intensified bowel preparation regime that requires patients to drink an even higher volume. This increases the difficulty for patients, resulting in recurrent inadequate bowel preparation.

Intraprocedural bowel cleansing removes the dependency on adequate preprocedural bowel preparation. In healthy individuals, we previously demonstrated the feasibility of an intraprocedural bowel cleansing system (BCS) (17). However, patients at risk for inadequate bowel preparation would be a more attractive patient population for BCS-assisted colonoscopy. Nevertheless, risk factors for inadequate bowel preparation overlap with risk factors for incomplete procedures (14, 18, 19), and use of an intraprocedural BCS might increase technical difficulty. Patients at risk for inadequate bowel preparation often have a history of abdominal or pelvic surgery, or diverticulosis resulting in narrowing of the lumen or fixation of the sigmoid (14, 18), which makes establishing the feasibility of an BCS in this population of interest. We aimed to assess the feasibility of BCS-assisted colonoscopy in a patient population with a history of inadequate bowel preparation.

Methods

Study design and population

We performed a single-arm multicenter prospective study in a tertiary referral center (Radboud university medical center, Nijmegen, The Netherlands), and an outpatient endoscopy unit (GastroZentrum Lippe, Bad Salzuflen, Germany). The study population consisted of patients aged 18 years or older scheduled for screening or surveillance colonoscopy. All patients had a history of inadequate bowel preparation in the last two years, defined as segmental BBPS of <2, need for repeat colonoscopy, or shortening of surveillance interval due to inadequate bowel preparation. Patients were enrolled consecutively at the date of their scheduled colonoscopy. Patients who had a history of (partial) colectomy, colorectal cancer or colitis (due to a higher risk of per-procedural adverse events), an American Society of Anesthesiologists (ASA) score of ≥ 4, inadequately corrected anticoagulation disorder, bowel obstruction, lower gastrointestinal bleeding with hemodynamic instability, pregnancy or lactating, or inability to provide informed consent, were excluded.

This study was performed in compliance with the Declaration of Helsinki, approved by the ethics committees of both sites, and registered at clinicaltrials.gov (NCT04700410). An independent monitor reviewed all study data. Written informed consent was obtained from all subjects before enrollment.

Study device

The second generation Pure-Vu System, (MotusGI, Tirat Carmel, Israel) is an FDA-approved and CE-certified BCS intended to increase visualization in inadequately prepared colons by offering intraprocedural cleansing. It consists of a disposable oversleeve and a workstation for flushing saline or water through the oversleeve (Figure 1A&B, Video 1), controlled by foot pedals. The oversleeve is available for most standard and slim size colonoscopes, with a maximum outer diameter of 21mm at the tip of the scope and 18mm at the oversleeve when attached on a standard size colonoscope. The workstation provides pulsed irrigation with simultaneous evacuation that breaks up feces faster than standard through-the-scope irrigation. The presence of several jets in the oversleeve and large suctioning tube could increase cleansing efficiency. An auto-purge function may prevent clogging.

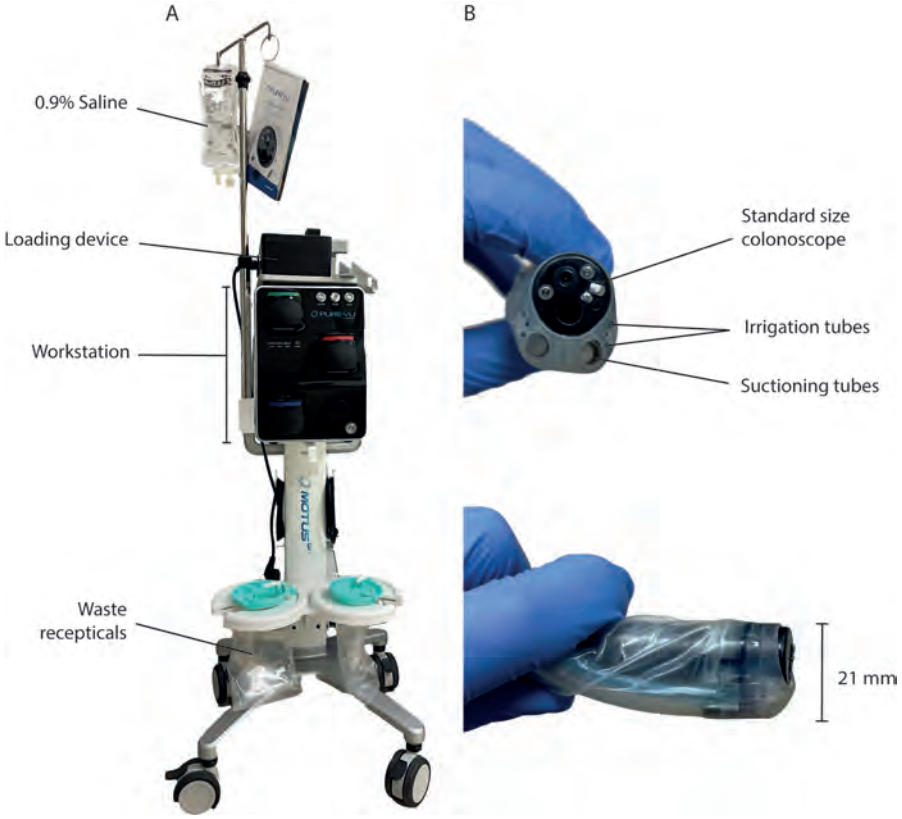


Figure 1. A. System set-up. A 0.9% saline solution or sterile water is used for cleaning and connected to the oversleeve via the workstation. Suction tubes are also connected to the workstation to evacuate debris. The loading fixture is located on the workstation to load the oversleeve onto the colonoscope. B. Details of the oversleeve with a standard size colonoscope.

Study procedures

After informed consent, we obtained participants' baseline characteristics using standardized case report forms. We used a combined approach for BP, using both laxatives and the BCS. All subjects, who would usually have been prescribed an intensified regimen due to previous inadequate preparation, now received a split-dose (2x150mL) sodium picosulfate magnesium citrate (Picoprep, Ferring, Germany), supplemented by a low-fiber diet starting two days before colonoscopy. Instructions were given by specialized endoscopy nurses, and an information leaflet containing detailed diet and laxative instructions was provided to each participant (Supplemental table S1). The laxatives were not intended to provide adequate BP on its own in this patient group. Additional bowel cleansing during colonoscopy using the BCS was performed during insertion and/or withdrawal at the endoscopists' discretion.

Colonoscopy procedure

All colonoscopies were outpatient procedures performed under propofol sedation. At each site, one expert endoscopist (P.S., H.N.), both having performed >2,000 colonoscopies, performed half of the study procedures. Only standard size colonoscopes were used (Pentax EC38-i10L, Fujifilm EC-760R-VL) and corresponding standard size BCS oversleeves.

The time spent for cecal intubation, withdrawal, washing with the BCS, and time for interventions were recorded. Cecal intubation- and withdrawal time did not include time for interventions or cleaning with the BCS. Common colonoscopy techniques for cecal intubation were used, including CO2 insufflation, manual pressure and position changes. All interventions were performed with the BCS attached. Segmental bowel preparation levels were assessed by the endoscopist using BBPS before BCS-use, but after regular flushing, and after full cleaning (20). Each corresponding segment was photo-documented before and after using the BCS. Patient discomfort during the procedure was scored by the attending nurses using the five-point Gloucester Comfort Scale score (GCS).

Follow-up (FU) consisted of a phone call at 48 hours and 1 month after the study procedure, to assess patient-related outcomes and adverse events (AE). Additionally, medical records were evaluated. Participants scored their level of abdominal pain on a numeric rating scale (NRS) from 0 (no pain) to 10 (worst pain).

Outcomes

Our primary endpoint was the proportion of adequately prepared patients before and after using the BCS assessed using BBPS. A second, independent and blinded evaluation was performed by a third expert endoscopist evaluating the photos taken before and after BCS use, blinded to the initial assessment. To minimize interobserver variability, all assessors completed a validated online training for BBPS assessment (21). Adequate bowel cleansing was defined as a BBPS of ≥2 per segment.

Secondary endpoints included cecal intubation rate (CIR), procedure times, device usability, level of patient discomfort and willingness to repeat, procedural findings, and safety. The usability of the BCS was graded by the performing endoscopist in concordance with previous studies on a 5-point scale (from 1 (worst) to 5 (best)) for general ease of use, ease of insertion, ease of angulation, ease of advancement, ease of retroflexion, ease of polyp resection, device stiffness, and device holding forces (Supplemental table S2).

Statistical analysis

The primary outcome was compared using a McNemars' test for the proportion of adequately prepared patients, and a Wilcoxon signed-rank test for segmental median BBPS scores before and after using the BCS. The second assessment of the BBPS scores was compared to the original scores using a Wilcoxon signed-rank test, applying Bonferroni correction for multiple testing.

For descriptive statistics, we used mean and standard deviation (SD) for normally distributed data, and median and interquartile range (IQR) in case of skewed data. Categorical data was represented as counts and percentages. To assess a learning curve, a secondary analysis using the Mann-Whitney-U test was performed for procedure times and usability scores, excluding the first four subjects per endoscopist.

Missing data were not imputed and excluded from analyses. Two-sided p-values of <0.05 were considered statistically significant. All analyses were performed with IBM SPSS 25 (IBM. Corp., Armonk, New York, USA). Graphs were designed using GraphPad Prism 9 (GraphPad Software, San Diego, California, USA).

Sample size calculation

Based on previous studies with this BCS and rates of recurrent inadequate bowel preparation, we assumed that 60% of patients would be adequately prepared before using the BCS, and 95% of patients would be adequately prepared after using the BCS. Using a McNemar's test with a 90% power and two-sided alpha of 5%, we required 36 patients to find a statistically significant difference. Per endoscopist, 4 extra patients were included to become familiar with the second-generation device, making the total sample size 44.

Role of funding source

The study was investigator-initiated and financially sponsored by Motus GI. The financial sponsor did not have a role in data collection, data analysis or interpretation, writing of the manuscript or decision to submit for publication.

Results

Baseline characteristics

Between June 2021 and March 2022, forty-four patients with a history of inadequate bowel preparation were included (Table 1). Prior to inclusion, inadequate BP had

been noted in a median of 70.8% of past colonoscopies. All patients had at least one and some up to six risk factors for inadequate BP (13, 14). The most frequent risk factors were chronic constipation, and history of intra-abdominal or pelvic surgery (19, 22). All patients reported full compliance with the laxatives.

Table 1. Baseline characteristics.

Number of patients, n	44
Male, n (%)	36 (81.8)
Age, median (IQR), years	61 (55-66.8)
BMI, median (IQR), kg/m^2	26.2 (24.4-29.1)
ASA, median (IQR)	2 (1-2)
Medication use, n (%)	
Tricyclic antidepressant	2 (4.5)
Chronic laxative use	6 (13.6)
Patient history, n (%)	
Chronic constipation[1]	5 (11.4)
Diabetes	5 (11.4)
Intra-abdominal or pelvic surgery	7 (15.9)
History of neurological disease[2]	2 (4.5)
Diverticulosis	19 (43.2)
Number of previously inadequately prepped colonoscopies, median (IQR)	1 (1-1)
Number of total colonoscopies, median (IQR)	2 (1-3.75)
Proportion of previous inadequately prepped colonoscopies, median (%)	70.8
BBPS score during previous colonoscopy[3] (total), n (%)	
0	1 (2.3)
3	7 (15.9)
4	7 (15.9)
5	26 (59.1)
6	1 (2.3)
Reason for previous poor BP[4], n (%)	
Non-compliance to diet or laxative	25 (56.8)
Laxative side effects	6 (13.6)
Medical history	9 (20.5)
Medication use	1 (2.3)
Unknown	3 (6.8)
Prior bowel preparation, n (%)	
2L PEG with ascorbate	9 (20.5)
1L PEG with ascorbate	1 (2.3)
300mL Sodium picosulfate/magnesium citrate	3 (6.8)
2L Sodium Phosphate solution	4 (9.1)
4L PEG electrolyte solution	18 (40.9)
Intensified regime[5]	9 (20.5)
Bisacodyl added to bowel preparation, n (%)	16 (36.4)

[1] defined as <3 bowel movements per week. [2] Neurologic disease including stroke, spina bifida, dementia, paraplegia, or Parkinson's disease. [3] Retrospectively collected combined segmental BBPS scores. Separate segmental BBPS scores were not available. [4] Inadequate bowel preparation during last colonoscopy procedure before study inclusion. [5] 4L PEG with ascorbate (n=5, 11.4%), 3L PEG with ascorbate (n=2, 4.6%), clinical preparation with 9L PEG with ascorbate (n=1), clinical preparation with 6.5L PEG with ascorbate (n=1). ASA, American Society of Anesthesiologist classification; BMI, Body mass index; IQR, interquartile range; PEG, polyethylene glycol.

Primary outcome

The median BBPS increased significantly from 1-2-2 (IQR 1-2) before to 3-3-3 (IQR 3-3) after BCS use (P<0.0001), with 31.8% and 88.6% being adequately prepared before and after BCS-use, respectively (P<0.0001). This outcome was confirmed by independent assessment of representative endoscopic photos (Table 2, Figure 2). In five patients (11.4%), cecal intubation was not feasible with the BCS attached and thus the BCS could not perform full colonic cleansing. In patients with successful cecal intubation, adequate cleansing was achieved in 97.5%.

Table 2. Primary outcome.

	Before BCS cleaning (n=44)	After BCS cleaning (n=44)	p-value
Median BBPS scores (IQR)	1-2-2 (1-2)*	3-3-3 (3-3)	P<0.0001
Median BBPS scores (IQR), secondary assessment	1-1-2 (1-2)*	3-3-3 (3-3)	P<0.0001
Adequately cleaned patients, n (%)	14 (31.8%)	39 (88.6%)	P<0.0001

BBPS, Boston Bowel Preparation Scale; A segmental score of ≥2 is defined as adequate; BCS, Bowel cleansing system; IQR, Interquartile range. * Median BBPS before cleaning between primary and secondary assessment P=0.012.

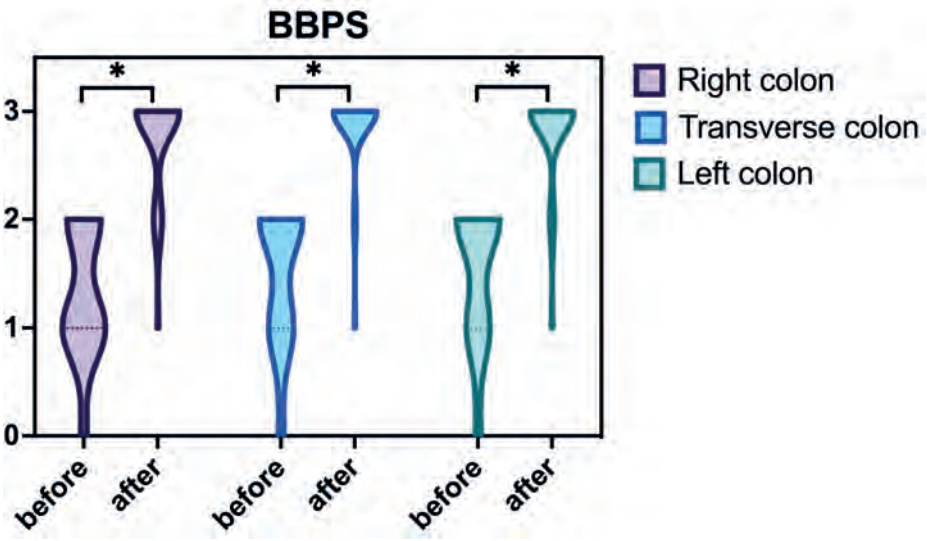


Figure 2. Violin plot of Boston Bowel Preparation Scale (BBPS) scores before and after cleaning with the BCS. The width of the plots indicates the prevalence of that particular score; median scores before BCS-use were 1-2-2, after BCS-use 3-3-3 (P<0.0001).

Secondary outcomes

Cecal intubation

The cecal intubation rate was 88.6%. Causes of failed cecal intubation were scope looping (n=2), technical device failure (n=1), relative stricture (n=1), and residual fecal material (n=1). The technical failure was caused by a protocol breach as the oversleeve was attached to an incompatible colonoscope. In two patients, cecal intubation was still possible with a regular colonoscope without BCS attachment, while this was not attempted for the remaining patients because of time constraints and inability to clean the remaining colon without the BCS. The ascending colon was reached in four of five patients with the BCS attached. In the fifth patient, the procedure was stopped in the transverse colon due to severe diverticulosis resulting in luminal narrowing.

Procedure times

The median total procedure time was 26 minutes, of which a median 5.3 minutes were spent on cleaning with the BCS (Figure 3). Results were not significantly different when the first four patients were left out to determine a learning curve (p=0.267-0.725).

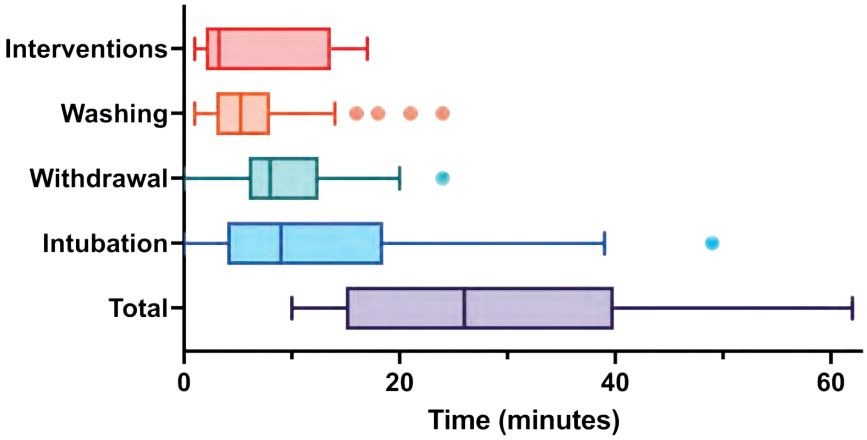


Figure 3. Boxplot of procedure times, median and interquartile range (IQR). Dots indicate times that are >1.5x IQR.

System usability

The performing endoscopist scored the general ease of use of the BCS as "good" or "excellent" in 84.1% of the cases with a median score of 4 out of 5 (IQR 4-5), while in 13.6% and 2.3% it was scored as "acceptable" and "difficult", respectively (Figure 4). The BCS did not interfere with insertion (100%), angulation (97.7%), advancement (97.7%), or polyp resection (100%), but the oversleeve increased the level of device holding forces and stiffness compared to conventional colonoscopy in 30% and 25% of procedures, respectively, which also increased difficulty of retroflexion in 3/11 (27%) patients. Usability scores were not significantly different for the first four included patients, compared to the other subjects (P=0.052 for insertion, P=0.142-0.927 for other categories).

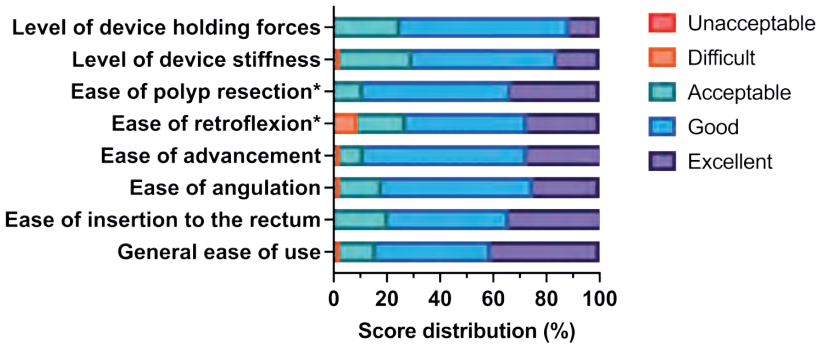


Figure 4. Proportional distribution of usability scores per category (n=44). *Categories "Ease of polyp resection"(n=9) and "ease of retroflex" (n=11) were only performed on a subset of patients, when applicable. A score of "good" represents "as good as conventional colonoscopy".

Patient experience & safety

The median GCS score during colonoscopy was 1 (IQR 1-1), meaning no discomfort. After 48 hours, 11 patients had mild abdominal pain (NRS 3, IQR 3-6). Willingness to repeat BCS-assisted colonoscopy was 92.9%. Three patients indicated that they did not want repeat BCS-assisted colonoscopy in the future, because of incomplete colonoscopy.

Five patients were lost to follow-up. During a median follow-up of 29 days (IQR 28-33), seven AEs in seven patients occurred, all minor, which resolved without sequelae. Two AEs were unrelated to the study device: One patient fell during exercising and one patient experienced transient facial numbness. The other five AEs were possibly related to the procedure or device, i.e., abdominal pain, NRS 5-7 (n=3), and mild rectal blood loss (n=2), both self-limiting.

Procedural findings

Procedural findings included 22 polyps in 10 patients, of which 9 were tubular adenomas, 1 tubulovillous adenoma, 3 sessile serrated adenoma, 6 hyperplastic polyps, and 3 others. This resulted in an adenoma detection rate of 22.7%.

Discussion

In this study, we assessed the feasibility of BCS-assisted colonoscopy for hard-to-prepare patients, to reduce dependency on extensive preprocedural bowel preparation and admission or need for a repeat-colonoscopy. We achieved adequate cleansing in 88.6% of patients, and in 97.5% of patients with successful cecal intubation. Cecal intubation was achieved in 88.6%, and no serious adverse events occurred during the study.

Our results align with previous studies conducted with this BCS on cleansing efficacy, procedure times, and usability (17, 23-25). Two feasibility studies were conducted in patients without a history of inadequate bowel preparation (17, 23). Van Keulen et al. included 47 outpatients from three sites and prescribed 20mg of bisacodyl with additional cleansing by BCS. Baseline adequate cleansing rates improved from 19.1% to 97.9% ($P<0.001$) and BBPS scores increased from median 3 to 9 ($P<0.001$). Perez-Jimenez et al. performed a similar feasibility study including 50 outpatients at two sites. After a 20mg bisacodyl preparation, the proportion of patients with adequate cleansing was 31%, which increased to 98% after BCS use ($P<0.001$). Neumann et al. treated 94 inpatients, who were at higher risk for inadequate bowel preparation, with regular bowel preparation combined with additional cleansing using the BCS (24). Adequate bowel cleaning levels increased from 38% before, to 96% after using the BCS ($P<0.001$).

Our CIR of 88.6% was comparable to the CIR of 84.5% in the study by Neumann et al. (24), which is below the recommended quality target of 90% (4). The BCS oversleeve increases the colonoscope diameter and stiffness, as reflected in the usability scores. Stiffness and holding forces were higher than in conventional colonoscopy with 25% and 30%, respectively. Together, this negatively impacts maneuverability. Additionally, a relatively high number of patients (11/44) with abdominal pain was noticed, although this was comparable to rates of abdominal pain following colonoscopy (26). However, known risk factors for inadequate bowel preparation, such as abdominal and pelvic surgery or diverticulosis, also increase technical difficulty and complicate cecal intubation (19, 22, 27). In the present study,

almost half of the patients (43.2%) had mild to severe diverticulosis. Furthermore, 15.9% had a history of intra-abdominal or pelvic surgery. Therefore, our study's suboptimal CIR should be regarded in light of a population with frequent difficult cecal intubation (27). Nonetheless, careful patient selection is advised to avoid incomplete procedures. For patients with a history of failed cecal intubation, this should be considered.

Existing evidence that recommends a regimen for patients with a history of inadequate bowel preparation is scarce. Two RCTs have been performed in this patient population, comparing 4L PEG to 2L PEG plus ascorbic acid (28), or to 6L PEG (29). Compared to the 6L PEG regime, 4L PEG did not perform significantly different regarding adequate cleansing rate (91.2% vs 87.6%; P =0.44), while it performed superior compared to 2L PEG plus ascorbic acid (81.1% vs. 67.4%, P=0.012). However, in both studies patients with repeated inadequate bowel preparation were excluded, which is the population with the highest burden.

Several potential patients may profit from BCS-assisted colonoscopy. First, in patients with recurrent inadequate bowel preparation, it may prevent repeat colonoscopies. Each repeat colonoscopy doubles healthcare costs and indirectly affects costs by impacting patient's working productivity (30). Arguably, BCS-use may increase costs by adding to the total procedure time. In the present study, median total procedure time was 29 minutes, of which cleaning took 5.3 minutes (18.3% of total procedure time), with 61% of patients only adequately cleaned after cleansing with the BCS. In comparison, in a prospective outpatient colonoscopy cohort, MacPhail et al. described a mean intraprocedural cleansing time of 4.1 minutes (17% of total procedure time) with an adequate bowel preparation rate of 96%, of which 6% was only adequate after intraprocedural cleansing (31). Additionally, BCS use may result in additional costs and CO2 footprint resulting from the single use plastics. However, this may be compensated by reducing repeat colonoscopies. Exact costs of the BCS are not yet established, limiting a cost-effectiveness analysis. We hypothesize that BCS-use would be cost-effective in difficult-to-prepare patients or salvage purposes. Other potential patient populations are inpatients, in which BCS-assisted colonoscopy may limit delays in hospital stays and therefore decreases costs (24, 32). Jacobs et al. demonstrated that inadequate bowel preparation was the most frequent cause of prolonged hospital stays in a retrospective cohort (n=4239). The use of a BCS may also result in fewer clinical admissions for bowel preparation.

Strengths of our study include the multicenter approach, the use of a BBPS-training upfront to decrease interobserver variability, and independent assessment of our primary outcome to minimize the risk of bias. Furthermore, we included patients usually excluded from participating in trials. A limitation of this study is the within-patient comparison of cleansing efficacy instead of a comparative arm. However, no standard effective bowel preparation regimen is available in this patient population. Additionally, the presumed added difficulty of BCS-use combined with the risk of failed cecal intubation in this patient population warranted a feasibility study before performing a randomized controlled trial. However, the true added benefit of BCS-assisted colonoscopy over intensified bowel preparation needs further confirmation. Additionally, our small sample size may hamper the generalizability of our results to new study groups. Finally, we only used standard size colonoscopes, while in patients with difficult cecal intubation frequently slim size colonoscopes are used. Therefore, no robust conclusions regarding our secondary endpoints can be drawn.

Conclusions

Concluding, BCS-assisted colonoscopy seems effective and safe for bowel preparation in patients with a history of recurrent inadequate bowel preparation. However, since these patients frequently have a difficult anatomy to achieve cecal intubation, careful patient selection is advised to avoid incomplete colonoscopies.

Acknowledgements

We thank dr. Geert Bulte for his independent BBPS assessment and acknowledge the nurses and staff at the study sites for their dedication to perform the study.

Competing interests: Author P.S. has received or receives unrestricted grants from Pentax (Japan), Norgine (UK), Motus GI (USA), MicroTech (China) and The eNose Company (Netherlands) and is in the advisory board of Motus GI (USA) and Boston Scientific (USA). Author H.N. is consultant for Fujifilm, Norgine, Medtronic, and 3D-Matrix. Authors M.R. and K.K. have no conflicts to disclose.

Study support: The study was investigator-initiated and financially supported by Motus GI. The financial sponsor did not have a role in data collection, data analysis or interpretation, writing of the manuscript or decision to submit for publication.

Ethics approval statement: This study was performed in compliance with the Declaration of Helsinki, approved by the Ethics committees at the two participating sites (METC Oost-Nederland registration number 2018-4543 and Ethik-Kommission der Ärztekammer Westfalen-Lippe und der Westfälischen Wilhelms-Universität Münster registration number 2021-269-f-S), and registered at clinicaltrials.gov (NCT04700410). An independent monitor reviewed all study data. Written informed consent was obtained from all subjects before enrollment.

Data availability statement: Study protocol, individual patient data, and analyses codes will be shared upon reasonable request.

References

1. Guo F, Chen C, Holleczek B, Schöttker B, Hoffmeister M, Brenner H. Strong Reduction of Colorectal Cancer Incidence and Mortality After Screening Colonoscopy: Prospective Cohort Study From Germany. Am J Gastroenterol. 2021;116(5):967-75.
2. Doubeni CA, Corley DA, Quinn VP, Jensen CD, Zauber AG, Goodman M, et al. Effectiveness of screening colonoscopy in reducing the risk of death from right and left colon cancer: a large community-based study. Gut. 2018;67(2):291-8.
3. Hassan C, East J, Radaelli F, Spada C, Benamouzig R, Bisschops R, et al. Bowel preparation for colonoscopy: European Society of Gastrointestinal Endoscopy (ESGE) Guideline - Update 2019. Endoscopy. 2019;51(8):775-94.
4. Kaminski MF, Thomas-Gibson S, Bugajski M, Bretthauer M, Rees CJ, Dekker E, et al. Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. Endoscopy. 2017;49(4):378-97.
5. Chokshi RV, Hovis CE, Hollander T, Early DS, Wang JS. Prevalence of missed adenomas in patients with inadequate bowel preparation on screening colonoscopy. Gastrointestinal Endoscopy. 2012;75(6):1197-203.
6. Clark BT, Rustagi T, Laine L. What level of bowel prep quality requires early repeat colonoscopy: systematic review and meta-analysis of the impact of preparation quality on adenoma detection rate. Am J Gastroenterol. 2014;109(11):1714-23; quiz 24.
7. Sulz MC, Kröger A, Prakash M, Manser CN, Heinrich H, Misselwitz B. Meta-Analysis of the Effect of Bowel Preparation on Adenoma Detection: Early Adenomas Affected Stronger than Advanced Adenomas. PLoS One. 2016;11(6):e0154149.
8. Clark BT, Laine L. High-quality Bowel Preparation Is Required for Detection of Sessile Serrated Polyps. Clin Gastroenterol Hepatol. 2016;14(8):1155-62.
9. Froehlich F, Wietlisbach V, Gonvers JJ, Burnand B, Vader JP. Impact of colonic cleansing on quality and diagnostic yield of colonoscopy: the European Panel of Appropriateness of Gastrointestinal Endoscopy European multicenter study. Gastrointest Endosc. 2005;61(3):378-84.
10. Kothari ST, Huang RJ, Shaukat A, Agrawal D, Buxbaum JL, Abbas Fehmi SM, et al. ASGE review of adverse events in colonoscopy. Gastrointest Endosc. 2019;90(6):863-76 e33.
11. Rex DK, Imperiale TF, Latinovich DR, Bratcher LL. Impact of bowel preparation on efficiency and cost of colonoscopy. Am J Gastroenterol. 2002;97(7):1696-700.
12. Bugajski M, Wieszczy P, Hoff G, Rupinski M, Regula J, Kaminski MF. Modifiable factors associated with patient-reported pain during and after screening colonoscopy. Gut. 2018;67(11):1958-64.
13. Mahmood S, Farooqui SM, Madhoun MF. Predictors of inadequate bowel preparation for colonoscopy: a systematic review and meta-analysis. Eur J Gastroenterol Hepatol. 2018;30(8):819-26.
14. Dik VK, Moons LM, Huyuk M, van der Schaar P, de Vos Tot Nederveen Cappel WH, Ter Borg PC, et al. Predicting inadequate bowel preparation for colonoscopy in participants receiving split-dose bowel preparation: development and validation of a prediction score. Gastrointest Endosc. 2015;81(3):665-72.
15. Sharara AI, Bou Daher H. Bowel Cleansing Strategies After Suboptimal Bowel Preparation. Clin Gastroenterol Hepatol. 2019;17(7):1239-41.
16. Committee ASoP, Saltzman JR, Cash BD, Pasha SF, Early DS, Muthusamy VR, et al. Bowel preparation before colonoscopy. Gastrointest Endosc. 2015;81(4):781-94.

17. van Keulen KE, Neumann H, Schattenberg JM, van Esch AAJ, Kievit W, Spaander MCW, et al. A novel device for intracolonoscopy cleansing of inadequately prepared colonoscopy patients: a feasibility study. Endoscopy. 2019;51(1):85-92.
18. Hsu CM, Lin WP, Su MY, Chiu CT, Ho YP, Chen PC. Factors that influence cecal intubation rate during colonoscopy in deeply sedated patients. J Gastroenterol Hepatol. 2012;27(1):76-80.
19. Shah HA, Paszat LF, Saskin R, Stukel TA, Rabeneck L. Factors associated with incomplete colonoscopy: a population-based study. Gastroenterology. 2007;132(7):2297-303.
20. Lai EJ, Calderwood AH, Doros G, Fix OK, Jacobson BC. The Boston bowel preparation scale: a valid and reliable instrument for colonoscopy-oriented research. Gastrointest Endosc. 2009;69(3 Pt 2):620-5.
21. Calderwood AH, Logan JR, Zurfluh M, Lieberman DA, Jacobson BC, Heeren TC, et al. Validity of a Web-based educational program to disseminate a standardized bowel preparation rating scale. J Clin Gastroenterol. 2014;48(10):856-61.
22. Takahashi Y, Tanaka H, Kinjo M, Sakumoto K. Prospective evaluation of factors predicting difficulty and pain during sedation-free colonoscopy. Dis Colon Rectum. 2005;48(6):1295-300.
23. Perez Jimenez J, Diego Bermudez L, Gralnek IM, Martin Herrera L, Libes M. An Intraprocedural Endoscopic Cleansing Device for Achieving Adequate Colon Preparation in Poorly Prepped Patients. J Clin Gastroenterol. 2019;53(7):530-4.
24. Neumann H, Latorre M, Zimmerman T, Lang G, Samarasena J, Gross S, et al. A multicenter, prospective, inpatient feasibility study to evaluate the use of an intra-colonoscopy cleansing device to optimize colon preparation in hospitalized patients: the REDUCE study. BMC Gastroenterol. 2021;21(1):232.
25. Gross SA, Gerson LB, Lewis BS, Ganz RA. A novel device for improving visualization in an inadequately prepared colon. Gastrointest Endosc. 2018;87(3):883-8.
26. Randel KR, Schult AL, Botteri E, Hoff G, Bretthauer M, Ursin G, et al. Colorectal Cancer Screening With Repeated Fecal Immunochemical Test Versus Sigmoidoscopy: Baseline Results From a Randomized Trial. Gastroenterology. 2021;160(4):1085-96.e5.
27. Rex DK. Achieving cecal intubation in the very difficult colon. Gastrointest Endosc. 2008;67(6):938-44.
28. Gimeno-García AZ, Hernandez G, Aldea A, Nicolás-Pérez D, Jiménez A, Carrillo M, et al. Comparison of Two Intensive Bowel Cleansing Regimens in Patients With Previous Poor Bowel Preparation: A Randomized Controlled Study. Official journal of the American College of Gastroenterology | ACG. 2017;112(6):951-8.
29. Sey MSL, Von Renteln D, Sultanian R, McDonald C, Bouin M, Chande N, et al. A Multicenter Randomized Controlled Trial Comparing Two Bowel Cleansing Regimens for Colonoscopy After Failed Bowel Preparation. Clin Gastroenterol Hepatol. 2021.
30. Fuccio L, Collatuzzo G, Frazzoni L, Cadoni S, Anderloni A, Laterza L, et al. Impact of colonoscopy on working productivity: a prospective multicenter observational study. Gastrointest Endosc. 2021.
31. MacPhail ME, Hardacker KA, Tiwari A, Vemulapalli KC, Rex DK. Intraprocedural cleansing work during colonoscopy and achievable rates of adequate preparation in an open-access endoscopy unit. Gastrointest Endosc. 2015;81(3):525-30.
32. Jacobs CC, Jaber JF, Ladna M, Ponniah S, Madhok I, Ruiz N, et al. Factors Associated With Inpatient Endoscopy Delay and its Impact on Hospital Length-of-Stay and 30-Day Readmission. Clin Gastroenterol Hepatol. 2021;19(12):2648-55.

Video legend

Video 1. *Demonstration of intraprocedural cleansing of an inadequately prepared colon using the bowel cleansing system. To view this video in the full-text HTML version of the article, please visit https://onlinelibrary.wiley.com/doi/10.1002/ueg2.12501.*

Supporting information

Table S1. Detailed bowel preparation instructions.

2 Days before colonoscopy	**Low fiber diet**
	• food low in fat and fiber AND without seeds, hard pieces, peels and skins
	• Examples:
	• Breakfast, lunch, and snacks: White bread, waffles, cracker, rice cake spread (thinly) with butter or margarine
	  • Topped with: Low-fat meats such as smoked meat, chicken breast, turkey breast, ham 30+ cheese, low fat cheese spread, dairy spread, egg prepared in any way, sweet toppings such as jam without seeds, honey, marmite, apple syrup, agave / date syrup, colored sprinkles
	• Beverages: Water, tea, coffee (maximum 2 cups), non-carbonated soft drinks, lemonade, tomato juice, fruit juice without pulp, low-fat or semi-skimmed milk or milk substitutes (oat milk, almond milk and soy milk, low-fat or semi-skimmed (soy) yogurt without seeds and pieces, (soy) custard, low-fat cottage cheese, semolina porridge, rice porridge, cornflakes, clear soup/bouillon.
	• Hot meal:
	  • White rice, plain macaroni, or spaghetti (no high fiber/ whole grain), boiled potatoes, mashed potatoes
	  • A choice of the following cooked vegetables: carrots, cauliflower, broccoli, beets, squash, applesauce (smooth)
	  • Tomato puree from a can
	  • Lean meat, chicken, or fish such as steak, ham steak, chicken breast, turkey breast, white fish, tofu, quorn chopped or quorn pieces, hachee
	• NOTE: legumes, nuts, cabbage, onion, raw vegetables and fresh fruit are NOT allowed

1 Day before colonoscopy	**Liquid diet** On the day before the colonoscopy, you should take liquid food such as custard, yoghurt, defatted broth or soup without pieces. At 5 pm you can take a last liquid meal. After this you are not allowed to eat anymore but you allowed to drink clear liquids. Clear liquids include water, lemonade, non-carbonated soft drinks, clear defatted broth (without vegetables, meat, or noodles), clear fruit juices without pulp such as apple juice and white grape juice, coffee and tea without milk (limit coffee to no more than 2 cups), and clear isotonic sports drinks. **Start sodium picosulfate magnesium citrate** At 6 pm, take the first sachet of sodium picosulfate magnesium citrate. After drinking the dissolved sachet, drink 2 liters of clear liquids. We recommend that you divide the amount of clear liquids until 8 pm and do not drink it all at once. Coffee, black tea, red fruit juices, cloudy fruit juices (such as orange juice with pulp), or milk (products) are not allowed Important: Take the indicated amount of fluids in addition to the amount you are used to drink daily. After 8 pm, you may continue to drink clear fluids. TIP: Do not only drink water but vary as much as possible by drinking different types of clear liquids. Take at least once a defatted broth (such as from a bouillon cube).

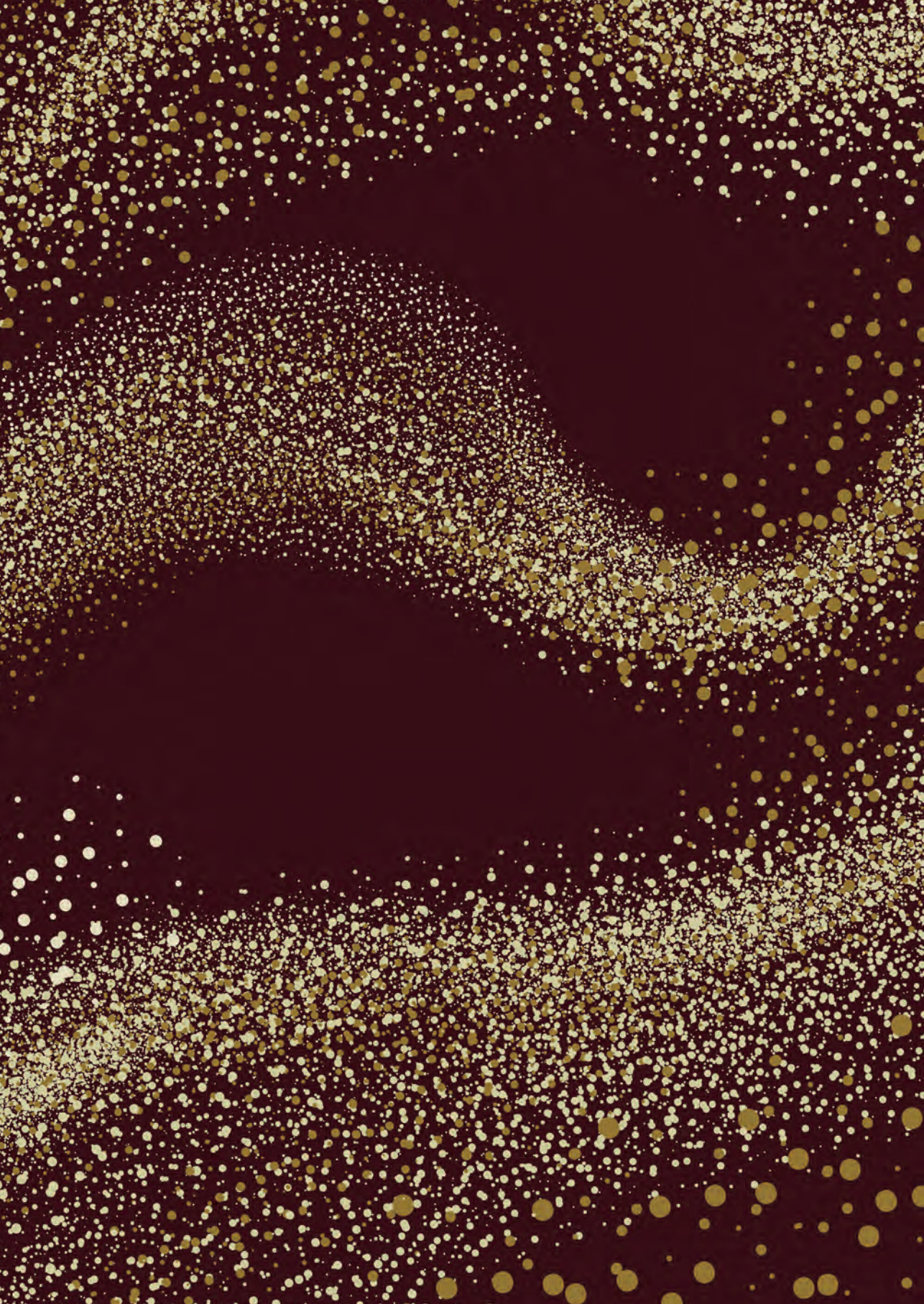
Table S1. Continued

On the day of colonoscopy	Four hours before going to the hospital, you need to drink another sachet of sodium picosulfate magnesium citrate and 2 liters of clear liquid. This should be finished two hours before leaving for the hospital. This is an important part of the preparation to have enough time to empty your bowels and achieving a clean bowel. Your stool after using 2 sachets of sodium picosulfate magnesium citrate should look just like the one you drank - clear and with a limited number of solid particles. You are done when your stool is yellow, light, watery and clear (just like urine).

Table S2. Scoring form used for grading the usability of the bowel cleansing system during colonoscopy, scored by the performing endoscopist.

	Unacceptable	Difficult	Acceptable	Good (as good as conventional colonoscopy)	Excellent	NA/ND
General ease of use						
Ease of insertion to the rectum						
Ease of angulation						
Ease of advancement						
Ease of retroflex						
Ease of polyp resection						
Level of device stiffness						
Level of device holding forces (weight, comfort)						

NA/ND; Not applicable/not done.



Chapter 7

General Discussion

Screening for colorectal cancer (CRC) is effective in early diagnosis and prevention, but faces several challenges that potentially result in cases of CRC that could have been prevented. In this thesis, we assessed the use of volatile organic compound (VOC) analysis by electronic nose (e-nose) as a new diagnostic biomarker and explored bowel preparation methods that may reduce patient discomfort while maintaining efficient cleaning, to improve colonoscopy quality. In this chapter, we reflect and contextualize this research and give suggestions for improvement of CRC screening. Additionally, we delineate future research aims. Figure 1 and Table 2 provide an overview of the main findings and limitations and comments per research chapter. Hereafter, we provide a discussion of the main findings, limitations, clinical implications, and future research recommendations per research chapter.

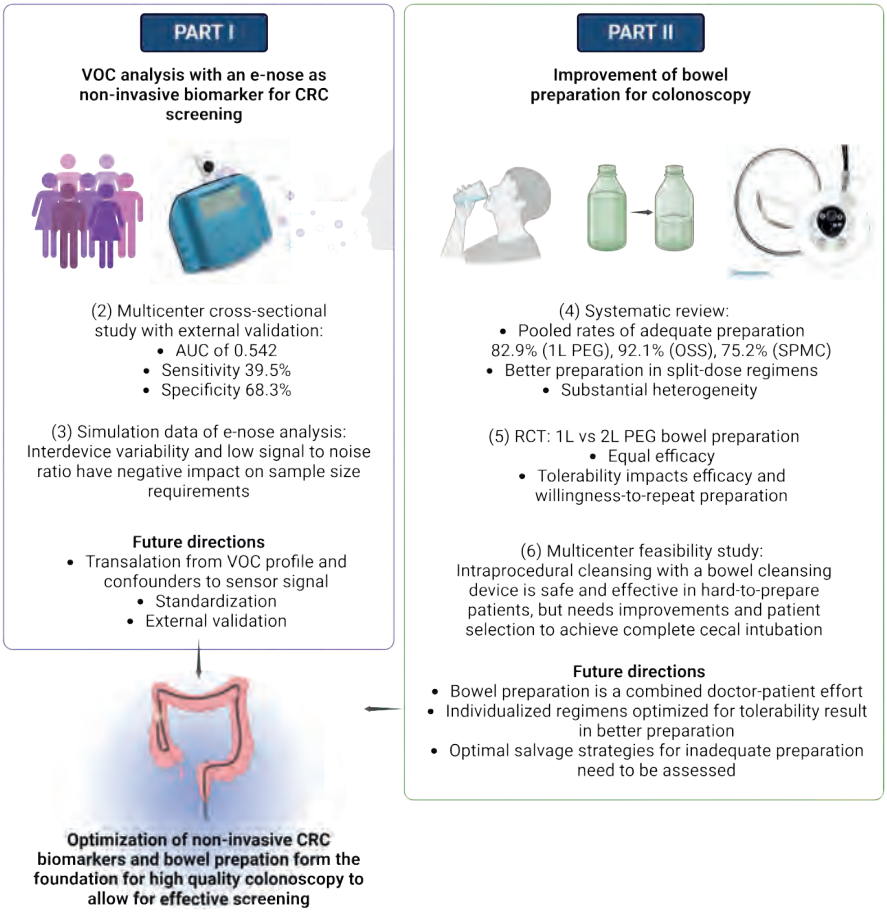


Figure 1. Overview of the most important conclusions in part I and II of this thesis.

Part I - Can CRC screening be improved by a breath test?

Although CRC screening is well-established and effective in the prevention of CRC and reducing CRC-related mortality, some challenges persist that impact the full potential and efficacy of screening. These can be divided into categories that represent phases of screening: In the phase in which individuals receive an invitation for screening, in which a fecal immunochemical test (FIT) kit is send via mail to identify high-risk populations for colonoscopy, key issues are related to non-participation and suboptimal FIT accuracy. In **Part I** of this thesis, the main aim was therefore to assess whether a breath test could be used as screening instrument for CRC to address these issues.

Main research findings

After a successful pilot study (1), in which VOCs in exhaled breath by e-nose were deemed a potential biomarker for CRC detection, we set out to create a robust study that not only could validate the breath test, but also could shed light on its robustness across different devices and study sites. Unfortunately, the findings of the pilot study could not be replicated. In **chapter 2**, we found that VOC analysis of exhaled breath by e-nose reached an AUC 0.542 (95% CI 0.495-0.589), with a sensitivity of 39.5% and specificity of 68.3%. These results were not different when using the prediction model from the initial pilot study, nor when designing a new prediction model based on the newly collected data. Patients had a high willingness to repeat the breath test (95.3%).

To find out why the results of **chapter 2** differed substantially from the initial pilot study, we performed a study of simulated breath test data in **chapter 3**. We hypothesized that the number of CRC cases per study device greatly impacted the adequate prediction of CRC presence. The per-device prevalence of CRC was significantly lower in the current study than in the pilot study due to lower overall CRC prevalence, but also due to a stricter patient selection (in the pilot study some patients provided a breath test after CRC diagnosis) and used more study devices. Because of this, the signal-to-noise ratio significantly decreased, which we presumed to have played a role in the experienced difficulty to predict CRC from breath samples. In **chapter 3**, we confirmed this hypothesis and found that low-prevalent disease prediction requires more positive cases (i.e. positive for neoplasia) than study samples with high neoplasia prevalence. We initially hoped that data augmentation of the positive cases would partially compensate for this, but this was not confirmed in **chapter 3**; compared to an increase in real data points (i.e. a breath test), substitution with augmented data provided only marginal

improvement of the prediction model performance. Lastly, we determined that the use of multiple measuring devices resulted in a significant increase of required data to train the prediction model, i.e., achieving an AUC of 0.80 would require around 2,100 data points when the disease prevalence is 5%, whereas only about 600 data points are needed at a 30% prevalence and 400 data points at a 50% prevalence. This could be due to case mix or small differences in the sensor signals, which lowers the signal-to-noise ratio and thereby increases difficulty to accurately predict CRC from a breath test.

Strengths and limitations

Strengths of **chapter 2** include the large-scale study (3,469 samples) that was nearly nationwide. This allowed us to assess the robustness of the results. In other words, in a positive scenario with a high accuracy of the breath test, this would probably have been established with a high level of certainty. Additionally, we tested the breath test in the future target population, with the intention of ensuring future applicability. Limitations included the exclusion of patients with a history of other malignancies and IBD to avoid confounding factors in the model and to remain close to our pilot study. After future establishment of a successful prediction model, this patient population needs to be addressed. Secondly, the VOC profile likely is subtle, making it prone to environmental disturbances including hospital variations or patient variations such as diet, comorbidities and medication use. Additionally, different sensor types in e-nose technology limit comparison between different types of e-nose devices, while sensor-drift and environmental disturbances introduce limitations to comparisons between devices of the same manufacturer.

Chapter 3 provides insight into inter-device disturbances and gives practical information on the impact of low-prevalent disease, which was the case in the bowel cancer screening population in our study. However, although we are confident that the main principle of low-prevalent disease detection also applies for other VOC detection methods, it not certain that specific device characteristics are directly translatable to other e-nose devices that use different sensors, data-processing, or machine learning techniques. Nonetheless, as all e-nose devices follow a comparable approach to go from the raw sensor data to a prediction model (2, 3), we believe that the main message still holds.

Clinical implications

The adagium '*the best test is the test that gets done*' certainly goes for colorectal cancer screening. While colonoscopy is recommended by European and American endoscopy society guidelines, because its sensitivity and effect in screening and

possibilities for treatment, accessibility and adherence of colonoscopy as primary screening method is subject to improvement. This is to a large extent due to the invasiveness of the procedure that poses a deterrent to patients on the one hand, and increasing resource demands that increase waiting times and costs on the other hand. Individual reasons for non-participation include logistical challenges, personal discomfort, perceived risk of CRC, and socioeconomic factors (4, 5). Addressing these barriers form a means to increase the CRC screening participation rate but often have a small effect (6, 7). A non-invasive screening method that offers selection of patients at risk for advanced neoplasia could improve accessibility of CRC screening.

The most frequently used non-invasive screening method for CRC is FIT screening in stool samples. A recent nested case control study confirmed that FIT screening resulted in a 33% lower risk of CRC-specific mortality (8), which was observed across different ethnic groups. This suggests that FIT screening is also effective in groups with a lower adherence. However, as FIT-based screening is based on the premise of occult blood loss by (pre)cancerous lesions, this potentially results in missed cases as not all lesions bleed enough to cross the threshold. This is reflected by the imperfect sensitivity of FIT (76-89%) (9) and the rate of colonoscopies in which no (advanced) neoplasia is detected in over 60% of patients participating in the Dutch national CRC screening program (10). The sensitivity and specificity of FIT screening is tied to the Hb threshold that is used, which is currently 47 µg/g feces in the Netherlands. Evaluation of this threshold after introduction of the national screening program indicated that a lower threshold may be associated with less missed cases of neoplasia, but also results in an increase of false-positive results and unnecessary colonoscopies with its associated costs (11). Higher cutoffs resulted in an increase of missed adenomas and cancers. This limits the room for further improvement in FIT-based screening and may open the door for other strategies.

Non-invasive biomarkers including proteins, DNA, microbiota, or VOCs in blood, stool, urine, or exhaled breath may offer alternatives. However, current research often includes early phase studies that have not yet been extensively tested in a screening population. **Chapter 2** aimed to build a bridge between proof-of-principle studies and clinical application. It is not uncommon that biomarkers in a controlled experimental setting have a higher accuracy than in sequential clinical settings. In a screening population, the added predictive value would be even lower (please see the box below), rendering these biomarkers unsuitable for population-based screening programs (12). This underscores the importance of external validation studies. This is especially true for studies investigating VOC-analysis, as there is

a risk of overfitting and spurious correlations due to the data-driven nature (13). Additionally, the multifactorial pathogenesis of VOC involvement in metabolomic changes, that also includes dietary habits and environmental influences, could lead to differences in VOC profiles across countries or even hospitals (14). Together, this may lead to overly optimistic claims on model performance and accuracy, that do not hold in new patients (15, 16).

Box 1. Effect of prevalence on diagnostic test performance

Suppose a population of 10,000 individuals, a hypothetical disease with a prevalence of 40% in cohort A and 5% in cohort B. The diagnostic test has a sensitivity of 95% and specificity of 90%.

In population A, this leads to the following test metrics:
Disease Cases: 10,000 ×0.40=4,000
Non-Disease Cases: 10,000-4,000=6,000

Metric	Calculation	Value
True Positives (TP)	4,000 ×0.95	3,800
False Negatives (FN)	4,000-3,800	200
True Negatives (TN)	6,000 ×0.90	5,400
False Positives (FP)	6,000-5,400	600
Positive Predictive Value (PPV)	$\frac{TP}{TP+FP} = \frac{3{,}800}{3{,}800+600}$	86.4%
Negative Predictive Value (NPV)	$\frac{TN}{TN+FN} = \frac{5{,}400}{5{,}400+200}$	96.4%

In this population, the likelihood of disease went from 40% to 86.4% following a positive test, and the likelihood of being healthy went from 60% to 96.4%.

In population B, the test metrics, using the same test but lower disease prevalence, are different:
Disease Cases: 10,000 ×0.05=500
Non-Cases: 10,000-500=9,500

Metric	Calculation	Value
True Positives (TP)	500×0.95	475
False Negatives (FN)	400-475	25
True Negatives (TN)	9,500×0.90	8,550
False Positives (FP)	9,500-8,550	950
Positive Predictive Value (PPV)	$\frac{TP}{TP+FP} = \frac{475}{475+950}$	33.3%
Negative Predictive Value (NPV)	$\frac{TN}{TN+FN} = \frac{8{,}550}{8{,}550+25}$	99.7%

In population B, the likelihood of disease went from 5% pre-test to 33.3% following a positive test, and the likelihood of being healthy went from 95% to 99.7% after a negative test

Thus, a diagnostic test with high sensitivity and specificity applied to a low prevalence population adds less to the post-likelihood chances of having the disease.

The future clinical benefits of e-nose technology in CRC screening are linked to participation rates. The efficacy of nationwide CRC screening programs heavily depends on participation, which varies significantly across European countries (17, 18). Given the high acceptability of the breath test investigated in **Chapter 2**, with a willingness-to-repeat rate exceeding 95%, a breath test could improve participation and the overall effectiveness of CRC screening. Studies have shown increased adherence when alternative tests, such as blood tests, are offered to individuals who previously declined screening (19). Another study reported the highest uptake rates for hypothetical esophageal adenocarcinoma screening scenarios when a breath test was offered (20). Therefore, implementing a breath test could enhance participation rates, addressing the gaps in CRC screening adherence.

The research in **chapter 2 and 3** demonstrated the complexity of VOC analysis. The performance of the current e-nose was significantly lower than in the previous pilot study, likely due to a lower CRC prevalence and issues such as sensor drift and variability in the data collected from different devices. External validation studies suggest that at least 100-200 positive CRC cases are needed for reliable results (21-23), and the variability between devices requires careful consideration in future study designs. Optimizing prediction models through better feature selection could reduce the necessary sample size and improve the reliability of e-nose technology.

This requires more insight into biochemical processes that lead to a CRC VOC profile and the signals to which this translates.

Research agenda

Looking forward, while VOC-based detection of CRC shows potential, further studies are still needed. The identification and validation of specific VOCs associated with CRC, the standardization of detection techniques, and the development of robust machine learning models are crucial steps. Additionally, addressing environmental influences on VOCs and standardizing breath tests will be essential for successful clinical application. Although the e-nose currently faces several challenges, continued research could eventually lead to its adoption as a non-invasive, accessible, and effective CRC screening tool. In Table 1, a comprehensive proposed research agenda for VOC analysis by e-nose is outlined.

Parallels can be drawn between VOC and e-nose research and the *Gartner Hype cycle*, in which initial excitement of technological innovations can lead to disillusionment, but with persistent development, it may find its place in practical applications (24). An equivalent might be the criticized high-profile study on 33 different cancers and their unique cancer microbiome that was published in Nature in 2020 and retracted in 2024 because questions were raised on its robustness (25). Liquid biopsy, the strategy that employs blood-based biomarkers for multi-cancer screening, also shows similarities with e-nose research, with hopeful pilot study results but low sensitivity when tested in larger study populations (26). Like VOCs, the signal of blood-based cancer biomarkers is heavily diluted, particularly in early stage cancers, requiring ultrasensitive tests to detect (26). Understanding these parallels can help in setting realistic expectations and guiding research and development strategies (Figure 2).

To increase the accuracy of VOC analysis in exhaled breath, it is necessary to unravel the metabolic processes that generate the specific VOC profile in CRC patients. These processes may include lipid peroxidation, amino acid metabolism, and fermentation, which are influenced by tumor-specific metabolic alterations (27). Oxidative stress in CRC leads to lipid peroxidation and the generation of hydrocarbons and aldehydes, such as hexanal and pentane, which can be detected in exhaled breath (27). These compounds serve as direct indicators of tumor-induced metabolic dysregulation and oxidative damage. Since VOCs might be linked to gut microbiota, playing a role in tumorigenesis as well (28, 29), information on the interplay of the microbiome and the volatome may aid in understanding the results and select VOCs that are more likely to be CRC specific (30). Gut dysbiosis

can create a pro-oncogenic environment through chronic inflammation, immune responses, and direct DNA damage, for example by pks+ E. Coli (31-33). VOCs related to gut microbiome changes, including short-chain fatty acids, amino acids, sulfur compounds, and metabolites like alcohols and ketones, are emerging as potential biomarkers for early CRC detection. The integration of microbiota, proteome, and amino acid profiles may improve the accuracy of VOC analysis by offering a more precise feature selection (34). Selecting the right VOC profile may potentially result in higher efficacy to train prediction models, which could possibly decrease the need for large samples to build an initial prediction model.

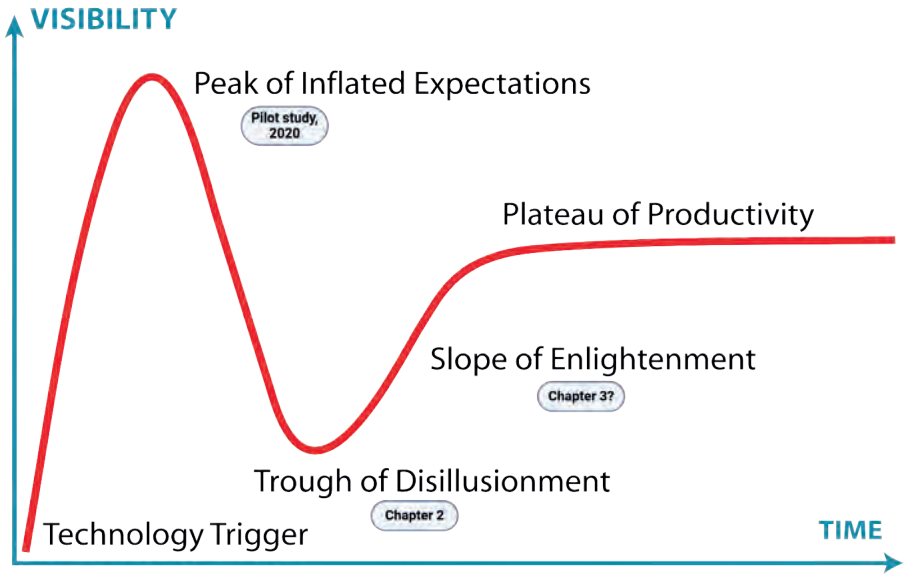


Figure 2. Parallels of Gartner Hype Cycle and e-nose research in this thesis. *Adapted from File:Gartner Hype Cycle.svg - Wikimedia Commons*

Based on the findings in **chapter 2 and 3**, we assume that the VOC profile has a low signal-to-noise ratio. Therefore, small disturbances from the test site (different hospitals), but also from the patient due to diet, smoking, comorbidities, and medication use, potentially disturb the signal. Environmental disturbances can be addressed through standardization of testing, a process that has been stressed to compare different VOC studies (35, 36). Confounding factors in patient variations need to be addressed in future studies (37). Additionally, differences between VOC media (e.g. breath, stool, etc), and between sensors and traditional chemical analysis methods such as GCMS should be assessed. By addressing these various

disturbances that negatively impact the signal-to-noise ratio, better comparisons between study devices and between studies can be made (38). Only then, clinical advancements can be achieved (38). However, developing analysis standards with stability and traceability at the required concentrations is resource-intensive and lack consensus, posing practical barriers to adoption (38, 39). These fundamental challenges and limitations implicate that VOC analysis for CRC screening will not be ready for clinical application in the near-future (36).

Table 1. Proposed research agenda for VOC analysis by e-nose to use in CRC screening.

Step	Study aims
Identification of Relevant VOCs for CRC Detection	Conduct a study to identify which VOCs are most indicative of CRC using GCMS
Characterization of VOC Signals in E-nose Devices	Investigate how the identified VOCs translate into signals on e-nose devices. Analyze the features of these signals to understand how they are represented and detected.
Enhancement of Current Breath Tests Using Improved Feature Selection	Explore whether the improved identification and characterization of VOCs can lead to better feature selection and, consequently, more accurate predictions in current breath tests for CRC detection.
Assessment of Signal Variability in E-nose Devices	Study the variability of e-nose signals generated by VOCs to determine the consistency and reliability of these devices in detecting CRC.
Variability of VOCs in Breath Samples	Investigate the variability of VOCs present in breath samples among different individuals and under various conditions to understand the robustness of VOC-based CRC detection.
Mitigation of Environmental VOCs in Breath Analysis	Investigate methods to reduce the influence of environmental VOCs on breath sample analysis to improve the accuracy of CRC detection using breath-based tests.
Comparison of Breath Tests to Other Media for VOC Detection	Conduct a comparative study to evaluate the differences between VOC detection in breath tests versus other media (e.g., stool, blood, urine) to determine the most effective medium for CRC detection.
Standardization of Breath Tests for VOC Analysis	Develop standardized protocols for breath test analysis, to ensure consistency and reliability in the detection of breath biomarkers for CRC.
External Validation of VOC-Based Breath Tests in a Screening Setting	Perform external validation of VOC-based breath tests for CRC in a large-scale screening setting. This study should involve diverse populations and multiple study sites to assess the effectiveness and applicability of the breath tests in real-world screening scenarios, ensuring the robustness and generalizability of the detection methods.

Based on **chapter 3**, we propose a study design with a CRC prevalence nearing 50% over a maximum of 5 e-nose devices including 100-150 cases per device to assess the discriminative ability. This should lead to robust modeling that potentially results in a transferable model across devices. Next, accuracy in screening settings should be assessed. Finally, external validation is crucial to assess the robustness of the prediction models (40). As the robustness of the model is not only subject to neoplasia prevalence, but also to AI-related issues like spurious correlations and overfitting, clinical application cannot be started until the test is validated in screening cohorts. Lastly, the exact implementation needs information on the added cost-effectiveness and place in the current CRC screening program.

Part II – Can bowel preparation be improved by lowering the burden while maintaining efficacy?

After selecting individuals with higher risk for CRC based on a positive FIT, the second part of the CRC screening process includes a high-quality colonoscopy to a) detect (early stage) CRC, and b) to detect and treat CRC precursor lesions. However, a prerequisite for this is adequate bowel preparation. In clinical practice, bowel preparation may be regarded as difficult, with some individuals even refraining from colonoscopy, thereby significantly reducing the effect of CRC screening. Therefore, in **Part II** of this thesis, the main aims were 1) to systematically assess the efficacy of bowel preparation fluids of 1 liter or less in literature, 2) to compare 1L bowel preparation to 2L bowel preparation and assess impact on work and quality of life, and 3) to investigate whether adequate bowel preparation could be achieved with an intraprocedural bowel cleansing device in patients that have recurrent inadequate bowel preparation.

Main research findings

In **chapter 4**, we conducted a systematic review and meta-analysis to assess the efficacy of one liter or less bowel preparations. We assessed 1L poly-ethylene glycol (PEG), 1L oral sulfate solution, 300mL sodium picosulfate combined with magnesium citrate, and sodium phosphate. The latter is not recommended for bowel preparation anymore due to an increased risk of nephropathy (41, 42). We found that in the included studies, low volume preparations frequently do not meet the 90% quality target set by the European Society of Gastrointestinal Endoscopy (ESGE) (41) (Table 2). Furthermore, we saw a positive trend for adequate bowel preparations in split-dose or same-day regimens and added adjuvant laxatives. The findings of **chapter 4,** with a subpar rate of adequate cleansing for low volume

preparations, were somewhat contradictory to the current bowel preparation guidelines of the ESGE, that state that low volume bowel preparations all could be used interchangeably with higher volume preparations such as 2L PEG+ascorbate or 4L PEG in patients without other contraindications for low-volume bowel preparation. Although direct comparisons cannot be drawn from our meta-analysis, it questions whether intermediate and high-volume bowel preparation perform subpar in studies as well.

In **chapter 5**, we performed a multicenter, open label, randomized controlled trial comparing a specialized 1L bowel preparation with a 2L bowel preparation. Considering our findings from **chapter 4** and the importance of adequate bowel preparation, our primary aim was to assess if the low volume preparation was non-inferior to the intermediate volume. Adequate bowel preparation was found in 96.1% in the low volume group and in 96.4% in the intermediate volume group, respectively. Therefore, we confirmed non-inferiority. The second part of **chapter 5** aimed to assess the impact of the bowel preparation process on patients. Through questionnaires before and after bowel preparation, we found that actual quality of life differences (EQ-5D-5L and SF36v2) are negligible, with no impact on work productivity (presenteeism) and absenteeism. However, we did find a close association between 1L bowel preparation and a higher tolerability and willingness to repeat.

Chapter 5 did not include patients at risk for inadequate bowel preparation. This represents a heterogenous group without uniform guidelines to advise a specific bowel preparation regimen. Therefore, in **chapter 6**, we performed a study investigating if intraprocedural bowel cleansing could serve as a viable option for bowel preparation in this patient group. We found that a bowel cleansing device could efficiently clean inadequately prepared colons, but with an increased difficulty for achieving cecal intubation (cecal intubation rate 88.6%).

Strengths and limitations

Chapters 4, 5, and **6** cover several aspects of bowel preparation. In **chapter 4**, a thorough systematic review was performed, including all studies that investigated bowel preparations of 1L or less. Studies dating from 2015 up to 2020 were included in this meta-analysis, which was performed separately for each type of preparation. Since the publication of this systematic review more post-marketing data has become available on specialized 1L bowel preparation (e.g., Pleinvue, Eziclen), often reporting higher rates of adequate bowel cleansing than reported in the included studies, and increasingly use a split-dose regimen, which is more effective in achieving adequate cleansing than a day before colonoscopy regimen.

A network meta-analysis could have enabled comparison between the laxatives, which could have been more informative to guide decision making in clinical practice. A second limitation is the high heterogeneity of the meta-analyses, which we assessed with sensitivity analyses and graphic display of study heterogeneity (GOSH) plots without substantial changing pooled outcomes.

In **chapter 5**, we conducted a pragmatic open-label multicenter randomized controlled trial, comparing 1L and 2L bowel preparations. The study provides valuable real world data representation, and the randomization reduces the risk of bias in the questionnaires compared to other non-randomized prospective questionnaire studies. After enrollment and data-analysis, the interval between the two questionnaires turned out to be shorter than anticipated when setting up the trial. As the IMTA productivity cost questionnaire is an instrument that uses a four-week recall period, this unfortunately limits comparison before and after bowel preparation. Nevertheless, as this is still an under-investigated topic, the questionnaire that was filled in after colonoscopy still provides interesting information on the impact of bowel preparation and colonoscopy on work absenteeism, presenteeism, and associated costs. Lastly, the study population consisted of predominantly Dutch subjects with a higher education. Since migration background and lower socio-economic levels are known risk factors for inadequate bowel preparation, it would be beneficial to collect information in this subpopulation as well. Additionally, it might be interesting to design a three-questionnaire study, with questionnaires before-, during-, and after bowel preparation, although this might negatively impact response rates and also complicates analysis, because it requires more sophisticated statistical approaches and has power considerations resulting in larger sample sizes.

Chapter 6 is a multicenter feasibility study of a bowel cleansing device in a subpopulation that is frequently inadequately prepared. This is a unique population that is frequently underrepresented in clinical trials assessing bowel preparation efficacy and therefore provides valuable information about this subgroup. Although we could confirm that this device may lead to more adequate cleansing in this challenging subgroup, due to the small sample size, our findings need confirmation in larger studies. New studies should compare intraprocedural cleansing by the study device to A) other salvage options, such as extra laxatives or an enema before performing the colonoscopy, and B) intensified bowel preparation regimes. Additionally, only two experienced endoscopists took part in the study, and since the publication of the study a new generation of the device that is proposed to be slimmer and easier to use as salvage option was introduced (Pure-Vu EVS, MotusGI). This potentially could prevent incomplete cecal intubation, which was seen

in 88.6% in our study, and this needs to be assessed prior to large scale studies. Furthermore, the specific subpopulation in which the added overtube does not lead to reduced cecal intubation still needs to be defined as this is not routinely documented in endoscopy reports in clinical practice.

Clinical implications

What is the most optimal bowel preparation regime?

Bowel preparation is a combined doctor-patient effort that consists of several components: High-quality instructions, diet modification, dosing regimen, and finally, a well-tolerated laxative. For optimal bowel preparation, all these components must be addressed (Figure 2).

Importance of patient education

High-quality instruction involves clear communication between the healthcare provider and the patient, ensuring that the patient fully understands the steps and importance of each element of the preparation process. Patients who receive detailed and easy-to-follow instructions, adhere more to recommended dietary changes, follow the prescribed dosing regimen, and use an effective laxative, are more likely to achieve high-quality bowel preparation. Special attention must be paid to those with lower health literacy and motivation. Various methods of enhanced education have been shown to notably enhance patient adherence and satisfaction to bowel preparation instructions (43-47). For example, a smartphone application to deliver instructions performed similar to paper instructions in an RCT, but with higher patient satisfaction (48). Similarly, use of an educational video lead to a higher rate of adequate preparation and was associated with a higher PDR (49). An educational video offers an easy low-cost option for centers to improve bowel preparation.

Diet modifications

Diet modification typically includes a low-residue diet or clear liquid diet before the procedure to reduce the amount of fecal material in the intestines. Over the years, evidence has however shown that a stricter liquid-only diet does not provide better results than a more liberal diet. In a meta-analysis of 9 RCTs, Nguyen demonstrated higher tolerability (OR 1.92), willingness to repeat (OR 1.86), and compliance (OR 1.04) of low-residue diet (LRD) compared to clear-liquid diet (CLD), with similar rates of adequate bowel preparation (50). Building on this, a more recently published systematic review demonstrated that one day of LRD is non-inferior to three days of diet restrictions (51). Machlab et al. additionally concluded in an RCT (n=582) that no diet restrictions resulted in an equally adequate bowel preparation

rate as a one-day LRD (52). Notably, randomization was stratified by inadequate bowel preparation risk using a prediction score (53).

Timing of the laxative

The dosing regimen refers to the timing and administration of the bowel preparation solution, which may be divided into split doses in the evening prior and on the morning of colonoscopy to enhance effectiveness and improve patient tolerance (54). A runway time of 4-6 hours must be kept to increase bowel preparation quality (41). Although for morning colonoscopies, a split-dose regimen is most practical, a same-day regimen can also be chosen for afternoon colonoscopies, resulting in less bloating (OR 0.68) and higher sleeping quality (OR 0.44) (55). Skipping the dose on the day of the colonoscopy by drinking all laxatives on the day prior to colonoscopy performs significantly worse (56).

Choosing the right laxative

Finally, the choice of laxative plays a crucial role; selecting the appropriate laxative type and dose can significantly impact the quality of the bowel cleansing, ultimately influencing the success of the procedure. Guidelines do not recommend a specific *go-to* laxative (41, 57). Nonetheless, lower volumes of bowel preparation are associated with better patient adherence and tolerability (54). A more tolerable preparation regimen reduces the likelihood of incomplete or inadequate bowel preparation and is even associated with a higher polyp detection rate (58). While in **Chapter 4** not all types of low-volume bowel preparation are recommended, **Chapter 5** demonstrates that 1L PEG+Asc (Pleinvue) provides an equally effective bowel preparation as 2L PEG+Asc (Moviprep), but with higher tolerability and willingness to repeat. In a recent meta-analysis, a 1L OSS (Eziclen) performed equally to 1L PEG+Asc (59). Moreover, a network meta-analysis comparing 14 bowel preparations over 22 RCTs in 7179 patients identified 2L PEG+ simethicone, 2L PEG+lactulose, and 1L PEG + Asc as the laxatives with the highest adequate cleansing rates (60). Sodium picosulfate magnesium citrate (SPMC) is the bowel preparation laxative with the lowest volume (2x150mL) but results in a slightly lower quality bowel preparation (61). A new development is that of bowel preparation tablets. Oral sulfate tablets (OST) could potentially improve tolerability and adherence compared to liquid laxatives, and studies report comparable cleaning efficacy to PEG-based laxatives (62, 63). However, OST do not release patients from drinking fluids as they should be taken with water as well. Additionally, they are associated with a slightly increased risk of erosive gastritis and peptic ulcers (64).

Adjuvant laxatives, such as bisacodyl, may increase bowel cleansing efficacy but do not lead to higher rates of adequate bowel preparation when at baseline effective laxatives are used and are therefore not recommended to use routinely (41, 59). Notably, they were not used in **Chapter 5** and both laxatives still reached a >96% adequate cleansing rate.

Altogether, while the specific choice of laxative may vary, the emphasis on patient comfort and adherence through lower volume preparations is key to achieving successful bowel preparation and, ultimately, a successful procedure. Indeed, based on results presented in this thesis, already four regional hospitals have adopted a more tolerable bowel preparation regimen.

Recommendations for failed bowel preparation

Patient at risk for inadequate preparation should preferably identified before colonoscopy, so timely additional measures can be taken. Risk factors can be divided into non-compliance (65) and reduced bowel motility factors, such as higher age, ASA score, high BMI, abdominal surgery, constipation, opioid use, neurological disease, polypharmacy, and the increasingly used GLP1-agontist or DDP4-inhibitors (66-69). Several prediction models that incorporate these risk factors have been developed, but their performance in effective identification of patients at risk is not perfect and information on real-world performance is lacking (53, 70-72). Nonetheless, recognition of known risk factors for inadequate bowel preparation may aid healthcare providers to select a targeted approach in these patients.

While non-compliance may be addressed through patient education and coaching in addition to a bowel preparation regimen with high tolerability, a uniform approach to the category of reduced bowel motility is not available (41). Where possible, an approach of same-day additional 2L PEG preparation and a second colonoscopy attempt should be considered to further complete the already partially performed bowel preparation (73). **Chapter 6** indicates that intraprocedural bowel cleansing using a device (Pure-Vu) may provide a solution to this, although careful patient selection is advised to not increase the risk of failed cecal intubation. At baseline, it is important to emphasize that salvage techniques are no substitute for a thorough and well-coordinated approach to managing patients with a history of inadequate bowel cleansing.

Our suggested approach for this is tailored to the individual patient and covers several aspects (Figure 2): Compliance, intensified laxative regimens, and monitoring and adjusting during the process. First, compliance should not be underestimated.

Intentional compliance can be addressed through extra education, stressing the importance of bowel preparation. A qualitative study revealed that even in patients with prior experience to bowel preparation and colonoscopy, numerous difficulties resulting from insufficient knowledge are experienced (74). This emphasizes the importance of high-quality instructions to patients. Unintentional compliance, for example due to side effects, taste, or volume difficulties, can be resolved through discussion of alternative laxatives with the patient.

When compliance is ensured, the next step is to add laxatives to the existing regimen. This can be in the form of adjunctive drugs, such as bisacodyl, or an extra dose of the used bowel preparation fluid (75-79). Some individuals may need more bowel preparation, depending on comorbidities, but clear guidance from the literature is lacking. In order to have a sufficiently high dose of laxatives but lower the risk of side effects and lessen the burden of bowel preparation, monitoring of the process with real-time patient feedback including fecal effluent is recommended. A photo taken by patients themselves, analyzed with an app, could be an easy way to guide taking bowel preparation fluids at home (80, 81).

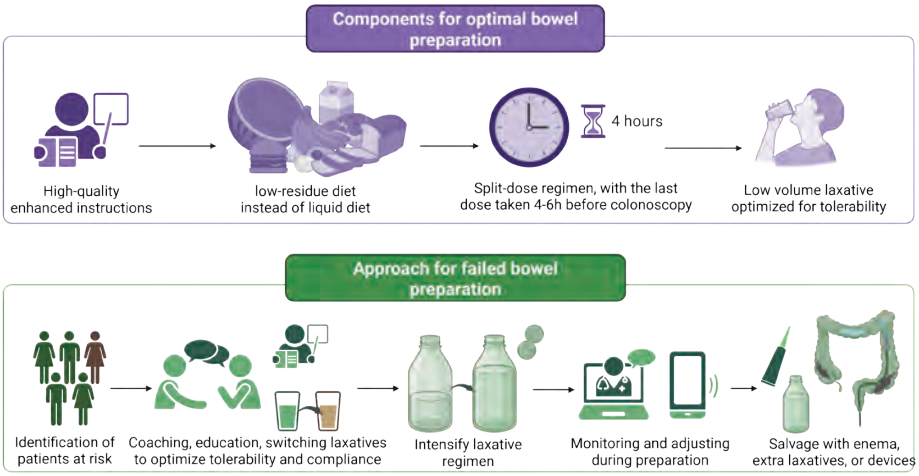


Figure 2. Overview of bowel preparation components that are the base to optimal bowel preparation and suggestions to approach inadequate bowel preparation.

Research agenda

Patient-centered approaches, which consider the individual's preferences and tolerability, may possibly further optimize outcomes of bowel preparation by improving the overall patient experience and compliance. Future research should

assess the application of individualized bowel preparation in hospitals and assess differences between subjects with and without prior bowel preparation experience.

Furthermore, in the current quality parameters for colonoscopy, patient experience using validated measurement instruments is recommended as a standard quality parameter(82). In clinical practice and current research however, this is an underrepresented topic. Our research demonstrates the importance of patient experience and satisfaction, as it is tied to compliance and willingness-to-repeat the bowel preparation. We recommend that colonoscopy experience including bowel preparation as part of this should be continuously monitored, to guide individual choices, but also as a quality measure of the hospital colonoscopy process. In patients who need future colonoscopy, the experience during previous colonoscopy should be considered to optimize the probability of adequate bowel cleansing. Current questionnaire recommendations include the global rating scale (GRS) or gastrointestinal endoscopy satisfaction questionnaire (GESQ) (83, 84). These do however not provide in depth information on bowel preparation. For bowel preparation experience specifically, the Mayo Florida bowel preparation questionnaire has been developed (85). Future studies should assess development, application, and results of a practical instrument that is easy to use and captures both colonoscopy and bowel preparation experience. Ideally, this is a digital survey coupled to a quality monitoring system to allow continuous monitoring.

To further optimize tolerability, additional studies are needed to investigate the optimal dietary guidelines for bowel preparation, particularly the effectiveness of more liberal dietary approaches versus stricter diets, and to explore whether further simplification of dietary restrictions could maintain or even improve bowel cleansing outcomes. Furthermore, optimizing options for bowel preparation to minimize patient burden are warranted. Recently, OSTs were developed (62), but the efficacy and added tolerability compared to other low-volume bowel preparation agents need to be compared.

Lastly, while salvage techniques are not recommended as primary strategies, further studies should investigate the specific patient populations for whom these techniques may be most beneficial and how they can be effectively integrated into a comprehensive bowel preparation protocol without compromising procedure success. With respect to the bowel cleansing device used in **Chapter 6**, future technical development may improve the usability of the device with respect to cecal intubation in more complex colonoscopies. Comparative trials are needed to assess the benefit of device cleaning to other strategies for either salvage cleaning or bowel preparation regimens for patients at risk for inadequate bowel preparation.

Research in context and future perspectives - Quality of colorectal cancer screening

Now that we have discussed the selection for colonoscopy with non-invasive screening methods, aiming to reduce the number of unnecessary procedures and to improve screening uptake, and directions to improve bowel preparation for colonoscopy as a foundation for high-quality colonoscopy screening, we provide an outlook on the impact of quality improvement of screening colonoscopies.

Current guidelines for colonoscopy screening suggest that a colonoscopy result without polyps or only a small number of non-advanced polyps (1-4 <10mm adenomas with low-grade dysplasia) justifies a surveillance interval of 10 years (86-88). Notably, the 10-year interval recommendation is on an empirical basis (89), and may require reassessment due to significant advancements in colonoscopy techniques and quality standards in recent years, which we will address below (90).

After the landmark study of Kaminski et al. that discussed the variations in ADR between endoscopists tied to the risk of PCCRC (91), quality improvement initiatives have ubiquitously improved colonoscopy quality (82). Current recommended quality parameters include a ≥90% adequate bowel preparation, ≥95% cecal intubation, an ADR of ≥35% and ≥50% in FIT-positive patients with a mean number of adenomas per colonoscopy of 0.6, a serrated lesion detection rate of ≥6%, a withdrawal time of ≥8 minutes, an appropriate polypectomy technique in ≥80%, measurement of complication rate and patient experience, and appropriate post-polypectomy surveillance recommendations (82, 92). Organized screening programs have implemented standardized quality monitoring and auditing, which further have improved adherence to quality standards (93). Recent data indicate an increase in ADR, suggesting enhanced detection quality (94-97). The Dutch Trans. IT database, a national endoscopy database aiming to monitor endoscopy quality, has reported improved quality parameters compared to the start of monitoring as well (98). This increased quality was reflected in a decreased PCCRC rate (0.166% vs. 0.027%) in a recent study from New Zealand, including 19 383 individuals who had undergone a screening colonoscopy (99).

In addition to a stricter adherence to quality standards, technological advancements including enhanced imaging quality (high-definition white light (HDWL) instead of standard definition) and increased emphasis on bowel cleanliness, have further contributed to quality improvements. Additional innovations in colonoscopy have shown promise in optimizing mucosal exposure and increasing ADR (100-102). Meta-analyses indicate that computer aided detection (CADe) can yield superior

outcomes compared to traditional methods (103). A network meta-analysis of 94 RCTs including 61,172 patients investigated different interventions to improve ADR for colonoscopy (100). Compared to mucosal exposure devices, chromoendoscopy, and water-based techniques, CADe performed significantly better to improve ADR. Interestingly, also water exchange colonoscopy and virtual chromoendoscopy options like iSCAN, narrow-band imaging (NBI), and linked color imaging (LCI) as part of commonly us HDWL endoscopes, performed better than 'HDWL-only', suggesting that ADR can be further improved using already available techniques. This was confirmed by a different network meta-analysis (104), that noted that dual observation and a 9-minute withdrawal time increased ADR as well.

Studies consistently show that CADe improves adenoma detection rates and reduces the miss rate, particularly for small and flat polyps (105, 106). However, although CADe increases the detection of small lesions, its effectiveness in detecting advanced adenomas (larger or with high-grade dysplasia) is less clear (107), with some concerns about the potential harm of increased detection of non-neoplastic lesions (108). A recent meta-analysis raised concerns about the quality of evidence, highlighting a serious risk of bias and limited applicability to symptomatic patients due to the focus on screening populations (108). Implementation or real-world studies of CADe have not shown the same benefits as randomized trials, which indicates potential gaps between controlled settings and routine practice (109). Our research group has recently published a multicenter randomized controlled trial (Discovery II study) investigating the effectiveness of a CADe (Discovery system, Pentax), and concluded that it did not significantly increase ADR nor APC, but increased the SSL rate by 58% (110). This could be due to a high ADR in the control arm, suggesting a lower additive benefit for high-detecting endoscopists, and is consistent with other real-world publications of CADe use (111, 112).

CADe and mucosal exposure devices such as balloon-aided colonoscopy (G-Eye colonoscope), potentially complement each other by increasing exposure of colonic mucosa, whereas CADe offers detection of lesions on the additionally exposed mucosal surface. Studies suggest that during conventional colonoscopy, a median 19% of mucosal surface is still missed due to colonic folds and curves, and endoscope slippage (113, 114). Therefore, we are currently performing a prospective study in which we compare the combined detection of CADe and balloon-assisted colonoscopy to CADe-only colonoscopy, to assess whether there is a synergistic effect of those techniques (Discovery III study, NCT05220345).

Future research is needed to demonstrate the added benefit of CADe on the long term. It is unsure whether detection of non-advanced lesions will lower CRC incidence and mortality, while increasing recourse demands by the need for additional surveillance colonoscopies (115). Hereto, we suggest a long-term follow-up study of subjects initially participating in CADe studies, which could be done for example using medical record data solutions like CTCue. CTCue is a GDPR-approved program coupled to the hospital electronic health record, enabling direct searches into electronic health record data for research questions. Furthermore, detailed insight into the type of lesions that are detected through CADe is warranted to guide future clinical application. Lastly, improved endoscopy quality also improves the detection of lesions. Moreover, it is suggested that optimization of the endoscopist's visual gaze pattern to include thorough examination of the screen periphery may enhance ADR and APC (116). The added benefit of CADe should therefore be compared to high-quality endoscopy. It would be of interest to compare the effect of a quality-improving training to CADe, which has been demonstrated to improve lesion detection as well (117-119).

Other quality aspects of colonoscopy in which AI may help is withdrawal time monitoring or visualizing the proportion of mucosal exposure. Gong et al. have introduced the ENDOANGEL, which monitors the withdrawal time and concluded that its use increased ADR from 8% to 16% (120). Similarly, Lui et al. designed an AI model that recognizes the cecum, rectum in retroflexion, and polypectomy or biopsy and thereby offers monitoring the effective withdrawal time, which showed a strong correlation with ADR (121). As most quality parameters including ADR and withdrawal time are surrogate markers for meticulous mucosal inspection, it could be useful to introduce a computer aided system that visualizes the proportion of mucosal exposure during colonoscopy and provides a map or even warns the endoscopist during colonoscopy of portions that may have been missed during inspection (113).

Furthermore, adequate assessment of the size and risk of neoplastic transformation largely determines the required surveillance after polypectomy (122). However, visual size assessment by the endoscopist greatly differs, resulting in 10-25% inappropriate surveillance intervals (123, 124). Therefore, we initiated a study using a laser assisted measurement of polyps to compare the accuracy to visual estimation and assess effect on surveillance intervals (Polyp size study, NTC NCT05489380). It is likely that AI may turn out helpful in this as well, as it may be integrated into existing CADe solutions (125). A difficulty in the development of improvement of polyp size measurement is the lack of a gold standard in polyp size. Due to cauterization and

suction effects accompanying polypectomy and shrinkage of polyps in formalin, the polyp size measured by the pathologist may differ from the intraluminal size. Comparing the data measurement variability between measurement methods can provide insight into the preciseness of the new measurement instrument in *in vivo* studies, while in *ex vivo* studies caliper measurements can be performed.

Lastly, computer aided characterization (CADx) has been developed for optical diagnosis of polyps (125, 126). CADx has shown potential in improving the accuracy of polyp characterization during colonoscopy, which could lead to more effective CRC screening and prevention. However, the performance of these systems is highly dependent on the quality of the data and the AI model's training, which has been shown to vary between studies (127). A recent systematic review and meta-analysis including 11 CADx studies did not find an added benefit of CADx for a resect-and-discard strategy, which is proposed to save pathology costs (128). The integration of AI with advanced imaging techniques, such as magnifying endoscopy, could potentially further improve the diagnostic accuracy of CADx systems.

Overall, given the improved quality of colonoscopy and expected further increase in the future with the help of innovations, it is hypothesized that a longer surveillance interval may be feasible following a colonoscopy without polyps or only a small number of non-advanced polyps (1-4 <10mm adenomas with low-grade dysplasia). Extending the interval from 10 to 15 years could reduce the burden of unnecessary colonoscopies even further, while maintaining the effectiveness of CRC prevention strategies. Recent studies using data from Swedish, Canadian, and British national registries have provided evidence supporting this approach, suggesting that the risk of CRC and CRC-specific mortality remains low up to 15 years post-colonoscopy (90, 129, 130). Data from a Swedish registry indicated that both the standardized incidence ratio and mortality ratio up to 15 years are 0.72 and 0.55 compared to the control group, respectively. The authors concluded that increasing the interval from 10 to 15 years results in 2 additional CRC cases and 1 CRC death, but also to a reduction of 1000 unnecessary colonoscopies (90). It should be noted that this is a primary colonoscopy screening cohort without a prior FIT. The Canadian registry confirmed this with a HR of 0.62 for female CRC incidence and HR of 0.57 for male incidence 15 years after a negative screening colonoscopy (129). In the British registry, consisting of data from population-based cohorts (n=195 453), a lower CRC HR of 0.51 was observed even after 20 years following a negative screening colonoscopy (130). Furthermore, sigmoidoscopy-only screening has also been shown to reduce CRC incidence and mortality up to 20 years in the invited-to-screen group (HR 0.76) (131).

In addition, in repeated FIT screening, selection based on prior fecal hemoglobin concentrations may offer less intensive screening. In a study nested in the Dutch CRC screening program including nearly 2.5 million individuals, subjects with undetectable hemoglobin had a substantial 8.71 times higher number needed to scope (132). This could imply less intensive screening in this subgroup. Similarly, an Italian screening program-based study found that repeated undetectable hemoglobin in FIT significantly decreased the probability of advanced neoplasia (1.4% vs 25.5%) or CRC (0.17% vs 4.5%) (133). Shifting the focus from selection of high-risk individuals to lessen the burden from low-risk individuals may be a promising avenue to decrease the harm-to-benefit ratio of screening. Next to VOC as possible biomarker, the addition of an AI-algorithm to FIT may also aid in selecting patients that benefit from colonoscopy (134).

In conclusion, the advancements in colonoscopy quality, driven by both technological innovations and stricter adherence to quality measures, suggest that extending surveillance intervals could be a beneficial strategy. Future guidelines should consider these improvements to optimize CRC screening protocols.

Conclusions

To reduce preventable CRC cases, we addressed several issues of CRC screening. Non-invasive screening tests aid in improving accessibility and lower missed cases. Unfortunately, VOC-analysis in exhaled breath by e-nose currently does not perform sufficiently to endorse application in clinical practice. Future studies should first focus on establishing reproducible VOC profiles of CRC patients before assessing new measuring devices such as electronic noses.

Conversely, improvements in bowel preparation protocols, particularly the use of low-volume solutions and intraprocedural bowel cleansing devices, offer benefits in enhancing patient compliance and colonoscopy quality. Improvement of bowel preparation quality has the potential to improve multiple facets of colonoscopy screening. This thesis underscores the critical role of tolerability in bowel preparation, showing that improved patient comfort enhances the efficacy and that lower volumes of bowel preparation improve tolerability. In patients that have recurrent inadequate bowel preparations, an intraprocedural bowel cleansing device seems safe and effective.

Together with improved quality of screening colonoscopy, these findings pave the way for more patient-centered approaches in colorectal cancer screening, potentially leading to higher participation rates and better overall outcomes.

Table 2. Overview of main findings per research chapter, including limitations and comments.

Chapter	Main findings	Limitations and comments
Part I: VOC analysis for CRC screening		
2	*Breath testing for colorectal cancer detection in patients with a positive fecal immunochemical test – a multicenter prospective cross-sectional study with external validation*	
	• The accuracy of volatile organic compound (VOC) analysis in exhaled breath by e-nose is currently insufficient for clinical application. • Sensitivity was 39.5% and specificity was 68.3%, with an AUC of 0.542 (95% CI 0.495-0.589). • In depth reproducibility assessment could not be performed due to insufficient accuracy. The intraclass correlation coefficient was 0.22, indicating poor reproducibility. • Willingness to repeat the breath test was 95.3%.	• Large scale validation of VOC study is lacking but imperative for clinical application, as this study demonstrates • Future research is needed to focus on the translation of the exact VOC profile to sensor-based analysis, to assess the effect of environmental disturbances and confounders, and comparison between GCMS and sensor-based analysis. • After standardization, new validation studies are required
3	*Overcoming methodological barriers in electronic nose clinical studies, a simulation data-based approach*	
	• Simulation data of VOC analysis can be used to assess the effect of varying neoplasia prevalence and number of used devices throughout the study • Low prevalence and high number of devices negatively impact the ability to build an adequate prediction model • Future e-nose studies need to include a minimum of 150 positive and negative cases per device	• This provides new insights into the results of chapter 2 and helps determining what components in the study design had the most impact on the accuracy. • This study aids new study designs in e-nose research, which may prevent unnecessary study endeavors and costs. • Lack of insight into the *black-box* proprietary sensor data and analysis limits in depth analysis.
Part II: Improvement of bowel preparation for colonoscopy		
4	*Efficacy of ultra-low volume (≤1L) bowel preparation fluids - systematic review and meta-analysis*	
	• Pooled rates of adequate preparation were 82.9% (1L PEG), 92.1% (OSS), 75.2% (SPMC) • Split dose or same day regimens provide highest rates of adequate preparation • Heterogeneity of meta-analyses was considerable (I^2 range: 86%-98%) despite in depth heterogeneity assessment	• The included studies did not all use a specialized low-volume bowel preparation, especially in the 1L PEG group. However, subgroup analyses including only the specialized low-volume laxatives were below 90% as well. • This review confirms that split-dose bowel preparation is preferent to day-before dosing • Direct comparisons between low-volume laxatives could not be drawn based on this systematic review

Table 2. Continued

Chapter	Main findings	Limitations and comments
5	*Comparing low-volume vs. intermediate volume bowel preparation and its impact on work and tolerability: An open-label, non-inferiority, randomized controlled trial*	
	• 1L PEG+Asc is non-inferior to 2L PEG+Asc. Both preparations provide a high rate of adequate cleansing (96.1% and 96.4%, respectively) • Impact on HRQoL is negligible and not different between 1L and 2L • Absenteeism and presenteeism occurred in 7.9% and 12.3%, respectively, but did not differ significantly between groups or before and after bowel preparation. • Tolerability is tied to the efficacy of bowel cleansing. 1L PEG+Asc has a higher tolerability and willingness to repeat.	• This study confirms the non-inferiority of 1L PEG+Asc with high efficacy in patients without indication for an intensified bowel preparation regime • The high questionnaire response rate, randomized design, use of validated questionnaires, and application of multiple imputations enhances generalizability. • Future studies need to address impact of bowel preparation on patients at risk of inadequate bowel preparation, and patients who have a non-Dutch background and/or lower socio-economic levels as these were not included in this study.
6	*An intraprocedural bowel cleansing system for difficult-to-prepare patients – a multicenter prospective feasibility study*	
	• Intraprocedural bowel cleansing using a bowel cleansing device provides adequate cleansing in patients with a history of inadequate bowel preparation, with 31.8% of patients adequately prepared before using the BCS and 88.6% after its use ($p<0.0001$). • Cecal intubation was reached in 88.6%. • In patients with successful cecal intubation, success rate for adequate preparation was 97.7%. • Time needed for cleaning was a median of 5.3 minutes per procedure. • The bowel cleansing device complicates cecal intubation due to increased endoscope diameter and stiffness, in a patient population that is already at risk of difficult cecal intubation	• Future research needs to address cecal intubation using newer generations of the bowel cleansing device including slim colonoscope use. • Comparative studies at a larger scale to other strategies for patients at risk for inadequate bowel preparation are required to understand the full benefit of a bowel cleansing device for both salvage strategies and complete bowel preparation regimens.

References

1. van Keulen KE, Jansen ME, Schrauwen RWM, Kolkman JJ, Siersema PD. Volatile organic compounds in breath can serve as a non-invasive diagnostic biomarker for the detection of advanced adenomas and colorectal cancer. Aliment Pharmacol Ther. 2020;51(3):334-46.
2. Smolinska A, Hauschild AC, Fijten RR, Dallinga JW, Baumbach J, van Schooten FJ. Current breathomics--a review on data pre-processing techniques and machine learning in metabolomics breath analysis. J Breath Res. 2014;8(2):027105.
3. Scott SM, James D, Ali Z. Data analysis for electronic nose systems. Microchimica Acta. 2006;156(3):183-207.
4. Berg-Beckhoff G, Leppin A, Nielsen JB. Reasons for participation and non-participation in colorectal cancer screening. Public Health. 2022;205:83-9.
5. Mohl JT, Ciemins EL, Miller-Wilson L-A, Gillen A, Luo R, Colangelo F. Rates of Follow-up Colonoscopy After a Positive Stool-Based Screening Test Result for Colorectal Cancer Among Health Care Organizations in the US, 2017-2020. JAMA Network Open. 2023;6(1):e2251384-e.
6. Hamer-Kiwacz S, Berntsson H, Galloway G, Tran AM, Tan JY, Hind D, et al. Barriers and facilitators to faecal immunochemical testing in symptomatic populations: A rapid systematic scoping review and gap analysis. J Eval Clin Pract. 2024.
7. Sheeran P, Frisch C, Listrom O, Pei Y, Bermudez A, Rothman AJ, et al. Systematic review of interventions to promote colorectal cancer screening: Benchmarking effect sizes and screening rates. Health Psychol. 2024.
8. Doubeni CA, Corley DA, Jensen CD, Levin TR, Ghai NR, Cannavale K, et al. Fecal Immunochemical Test Screening and Risk of Colorectal Cancer Death. JAMA Netw Open. 2024;7(7):e2423671.
9. Grobbee EJ, Wisse PH, Schreuders EH, van Roon A, van Dam L, Zauber AG, et al. Guaiac-based faecal occult blood tests versus faecal immunochemical tests for colorectal cancer screening in average-risk individuals. Cochrane Database Syst Rev. 2022;6(6):Cd009276.
10. INKL. Belangrijke cijfers over darmkanker 2019 [Available from: https://iknl.nl/getmedia/e27ff657-1605-460c-af01-2c3657c46448/folder_darmkanker_belangrijkste_cijfers_over_maart2019.pdf.
11. Toes-Zoutendijk E, van Leerdam ME, Dekker E, van Hees F, Penning C, Nagtegaal I, et al. Real-Time Monitoring of Results During First Year of Dutch Colorectal Cancer Screening Program and Optimization by Altering Fecal Immunochemical Test Cut-Off Levels. Gastroenterology. 2017;152(4):767-75.e2.
12. Church TR, Wandell M, Lofton-Day C, Mongin SJ, Burger M, Payne SR, et al. Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer. Gut. 2014;63(2):317-25.
13. Leopold JH, Bos LDJ, Sterk PJ, Schultz MJ, Fens N, Horvath I, et al. Comparison of classification methods in breath analysis by electronic nose. Journal of Breath Research. 2015;9(4):046002.
14. O'Keefe SJD. Diet, microorganisms and their metabolites, and colon cancer. Nat Rev Gastroenterol Hepatol. 2016;13:691-706.
15. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Intern Med. 1999;130(6):515-24.
16. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Ann Intern Med. 2006;144(3):201-9.
17. Bretthauer M, Løberg M, Wieszczy P, Kalager M, Emilsson L, Garborg K, et al. Effect of Colonoscopy Screening on Risks of Colorectal Cancer and Related Death. New England Journal of Medicine. 2022.

18. Zhu Y, Li X, Hu Y, Chen K, Zheng S, Ding K. Nonadherence to Referral Colonoscopy After Positive Fecal Immunochemical Test Results Increases the Risk of Distal Colorectal Cancer Mortality. Gastroenterology. 2023;165(6):1558-60.e4.
19. Coronado GD, Jenkins CL, Shuster E, Johnson C, Amy D, Cook J, et al. Blood-based colorectal cancer screening in an integrated health system: a randomised trial of patient adherence. Gut. 2024.
20. Sijben J, Rainey L, Maas F, Broeders MJ, Siersema PD, Peters Y. The public's intended uptake of hypothetical esophageal adenocarcinoma screening scenarios: a nationwide survey. Am J Gastroenterol. 2024.
21. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. Stat Med. 2016;35(2):214-26.
22. Riley RD, Debray TPA, Collins GS, Archer L, Ensor J, van Smeden M, et al. Minimum sample size for external validation of a clinical prediction model with a binary outcome. Statistics in Medicine. 2021;40(19):4230-51.
23. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. J Clin Epidemiol. 2005;58(5):475-83.
24. Campani M, Vaglio R. A simple interpretation of the growth of scientific/technological research impact leading to hype-type evolution curves. Scientometrics. 2015;103(1):75-83.
25. Poore GD, Kopylova E, Zhu Q, Carpenter C, Fraraccio S, Wandro S, et al. Retracted: Microbiome analyses of blood and tissues suggest cancer diagnostic approach. Nature. 2020;579(7800):567-74.
26. Farooq M, Leevan E, Ahmed J, Ko B, Shin S, De Souza A, et al. Blood-based multi-cancer detection: A state-of-the-art update. Curr Probl Cancer. 2024;48:101059.
27. Barbosa JMG, Filho NRA. The human volatilome meets cancer diagnostics: past, present, and future of noninvasive applications. Metabolomics. 2024;20(5):113.
28. van Vorstenbosch R, Cheng HR, Jonkers D, Penders J, Schoon E, Masclee A, et al. Systematic Review: Contribution of the Gut Microbiome to the Volatile Metabolic Fingerprint of Colorectal Neoplasia. Metabolites. 2022;13(1).
29. Qu R, Zhang Y, Ma Y, Zhou X, Sun L, Jiang C, et al. Role of the Gut Microbiota and Its Metabolites in Tumorigenesis or Development of Colorectal Cancer. Adv Sci (Weinh). 2023:e2205563.
30. Hernandez-Leyva AJ, Berna AZ, Liu Y, Rosen AL, Lint MA, Whiteside SA, et al. The breath volatilome is shaped by the gut microbiota. medRxiv. 2024.
31. Pleguezuelos-Manzano C, Puschhof J, Rosendahl Huber A, van Hoeck A, Wood HM, Nomburg J, et al. Mutational signature in colorectal cancer caused by genotoxic pks(+) E. coli. Nature. 2020;580(7802):269-73.
32. Janney A, Powrie F, Mann EH. Host–microbiota maladaptation in colorectal cancer. Nature. 2020;585(7826):509-17.
33. Watanabe D, Murakami H, Ohno H, Tanisawa K, Konishi K, Tsunematsu Y, et al. Association between dietary intake and the prevalence of tumourigenic bacteria in the gut microbiota of middle-aged Japanese adults. Scientific Reports. 2020;10(1):15221.
34. Bosch S, Acharjee A, Quraishi MN, Bijnsdorp IV, Rojas P, Bakkali A, et al. Integration of stool microbiota, proteome and amino acid profiles to discriminate patients with adenomas and colorectal cancer. Gut Microbes. 2022;14(1):2139979.

35. Hanna GB, Boshier PR, Markar SR, Romano A. Accuracy and Methodologic Challenges of Volatile Organic Compound–Based Exhaled Breath Tests for Cancer Diagnosis: A Systematic Review and Meta-analysis. JAMA Oncology. 2019;5(1):e182815-e.

36. Ma T-t, Chang Z, Zhang N, Xu H. Application of electronic nose technology in the diagnosis of gastrointestinal diseases: a review. Journal of Cancer Research and Clinical Oncology. 2024;150(8):401.

37. Skawinski M, Schooten FJv, Smolinska A. A comprehensive guide to volatolomics data analysis. Journal of Breath Research. 2025;19(1):015001.

38. Jeerage KM, Berry JL, Murray JA, Goodman CA, Piotrowski PK, Jones CM, et al. The need for multicomponent gas standards for breath biomarker analysis. J Breath Res. 2022;16(4).

39. Skawinski M, van Schooten FJJ, Smolinska A. A Comprehensive Guide to Volatolomics Data Analysis. J Breath Res. 2024.

40. Parasa S, Repici A, Berzin T, Leggett C, Gross SA, Sharma P. Framework and metrics for the clinical use and implementation of artificial intelligence algorithms into endoscopy practice: recommendations from the American Society for Gastrointestinal Endoscopy Artificial Intelligence Task Force. Gastrointest Endosc. 2023;97(5):815-24.e1.

41. Hassan C, East J, Radaelli F, Spada C, Benamouzig R, Bisschops R, et al. Bowel preparation for colonoscopy: European Society of Gastrointestinal Endoscopy (ESGE) Guideline - Update 2019. Endoscopy. 2019;51(8):775-94.

42. Markowitz GS, Stokes MB, Radhakrishnan J, D'Agati VD. Acute phosphate nephropathy following oral sodium phosphate bowel purgative: an underrecognized cause of chronic renal failure. J Am Soc Nephrol. 2005;16(11):3389-96.

43. Liu Z, Zhang MM, Li YY, Li LX, Li YQ. Enhanced education for bowel preparation before colonoscopy: A state-of-the-art review. J Dig Dis. 2017;18(2):84-91.

44. Farid N, Sharifian S, Ghafouri R. Evaluation of the effect of reinforced education on the satisfaction of patients undergoing colonoscopy: A randomized controlled trial. PLoS One. 2024;19(1):e0296126.

45. He X, Lei X, Li J, Li P. Telephone instructions improve the quality of bowel preparation for colonoscopy: A meta-analysis of randomized controlled trials. PLoS One. 2023;18(11):e0289063.

46. Guo X, Yang Z, Zhao L, Leung F, Luo H, Kang X, et al. Enhanced instructions improve the quality of bowel preparation for colonoscopy: a meta-analysis of randomized controlled trials. Gastrointest Endosc. 2017;85(1):90-7 e6.

47. Chang CW, Shih SC, Wang HY, Chu CH, Wang TE, Hung CY, et al. Meta-analysis: The effect of patient education on bowel preparation for colonoscopy. Endosc Int Open. 2015;3(6):E646-52.

48. Patel SV, Yu D, Taylor C, McKay J, Hookey L. Smartphone Application Versus Standard Instruction for Colonoscopic Preparation: A Randomized Controlled Trial. J Clin Gastroenterol. 2024.

49. Archer T, Corfe B, Dear K, Cole A, Foley S, Andreyev HJN, et al. Can an educational video improve the adequacy of bowel preparation for patients undergoing their first colonoscopy? Results of the EBOPS RCT. Endosc Int Open. 2024;12(3):E402-e12.

50. Nguyen DL, Jamal MM, Nguyen ET, Puli SR, Bechtold ML. Low-residue versus clear liquid diet before colonoscopy: a meta-analysis of randomized, controlled trials. Gastrointest Endosc. 2016;83(3):499-507 e1.

51. Wang F, Huang X, Wang Z, Yan Z, Wang S, Pan P, et al. One-day versus three-day low-residue diet bowel preparation regimens before colonoscopy: a meta-analysis of randomized controlled trials. J Gastroenterol Hepatol. 2024.

52. Machlab S, Martínez-Bauer E, López P, Ruiz-Ramirez P, Gómez B, Gimeno-Garcia AZ, et al. Restrictive diets are unnecessary for colonoscopy: Non-inferiority randomized trial. Endosc Int Open. 2024;12(3):E352-e60.

53. Dik VK, Moons LM, Huyuk M, van der Schaar P, de Vos Tot Nederveen Cappel WH, Ter Borg PC, et al. Predicting inadequate bowel preparation for colonoscopy in participants receiving split-dose bowel preparation: development and validation of a prediction score. Gastrointest Endosc. 2015;81(3):665-72.

54. Spadaccini M, Frazzoni L, Vanella G, East J, Radaelli F, Spada C, et al. Efficacy and Tolerability of High- vs Low-Volume Split-Dose Bowel Cleansing Regimens for Colonoscopy: A Systematic Review and Meta-analysis. Clin Gastroenterol Hepatol. 2020;18(7):1454-65 e14.

55. Cheng YL, Huang KW, Liao WC, Luo JC, Lan KH, Su CW, et al. Same-day Versus Split-dose Bowel Preparation Before Colonoscopy: A Meta-analysis. J Clin Gastroenterol. 2018;52(5):392-400.

56. Martel M, Barkun AN, Menard C, Restellini S, Kherad O, Vanasse A. Split-Dose Preparations Are Superior to Day-Before Bowel Cleansing Regimens: A Meta-analysis. Gastroenterology. 2015;149(1):79-88.

57. Committee ASoP, Saltzman JR, Cash BD, Pasha SF, Early DS, Muthusamy VR, et al. Bowel preparation before colonoscopy. Gastrointest Endosc. 2015;81(4):781-94.

58. Holt EW, Yimam KK, Ma H, Shaw RE, Sundberg RA, Verhille MS. Patient tolerability of bowel preparation is associated with polyp detection rate during colonoscopy. J Gastrointestin Liver Dis. 2014;23(2):135-40.

59. Liu X, Yu W, Liu J, Liu Q. Oral sulfate solution versus polyethylene glycol for bowel preparation before colonoscopy, meta-analysis and trial sequential analysis of randomized clinical trials. Tech Coloproctol. 2024;28(1):99.

60. Maida M, Marasco G, Fuccio L, Vitello A, Mocciaro F, Amata M, et al. Comparative efficacy of different bowel preparations for colonoscopy: A network meta-analysis. Dig Liver Dis. 2024.

61. Serradesanferm A, Torá-Rocamora I, Pozo À, Ocaña T, Diaz M, Moreira R, et al. Adenoma detection rate and tolerability of two ultra-low volume bowel preparations in screening: a noninferiority randomized controlled trial. Gastrointest Endosc. 2024.

62. Song JH, Bae JH, Yim JY. Efficacy of oral sulfate tablets for bowel preparation and adenoma detection rate. J Gastroenterol Hepatol. 2023;38(3):410-5.

63. Di Palma JA, Bhandari R, Cleveland MV, Mishkin DS, Tesoriero J, Hall S, et al. A Safety and Efficacy Comparison of a New Sulfate-Based Tablet Bowel Preparation Versus a PEG and Ascorbate Comparator in Adult Subjects Undergoing Colonoscopy. Am J Gastroenterol. 2021;116(2):319-28.

64. Haydel JM, Xu AA, Mansour NM. High volume, low volume, or pills, which way should we go? a review of bowel preparation for colonoscopy. Curr Opin Gastroenterol. 2024;40(1):21-6.

65. Serper M, Gawron AJ, Smith SG, Pandit AA, Dahlke AR, Bojarski EA, et al. Patient factors that affect quality of colonoscopy preparation. Clin Gastroenterol Hepatol. 2014;12(3):451-7.

66. Gandhi K, Tofani C, Sokach C, Patel D, Kastenberg D, Daskalakis C. Patient Characteristics Associated With Quality of Colonoscopy Preparation: A Systematic Review and Meta-analysis. Clin Gastroenterol Hepatol. 2018;16(3):357-69 e10.

67. Mahmood S, Farooqui SM, Madhoun MF. Predictors of inadequate bowel preparation for colonoscopy: a systematic review and meta-analysis. Eur J Gastroenterol Hepatol. 2018;30(8):819-26.

68. Abu-Freha N, Yitzhak A, Shirin H, Nevo-Shor A, Abu-Jaffar J, Abu-Rafe S, et al. Glucagon-like peptide-1 receptor agonists significantly affect the quality of bowel preparation for colonoscopy. Endoscopy. 2024.

69. Beran A, Abboursheid T, Ali AH, Albunni H, Mohamed MF, Vargas A, et al. Risk Factors for Inadequate Bowel Preparation in Colonoscopy: A Comprehensive Systematic Review and Meta-Analysis. Am J Gastroenterol. 2024.

70. Hassan C, Fuccio L, Bruno M, Pagano N, Spada C, Carrara S, et al. A predictive model identifies patients most likely to have inadequate bowel preparation for colonoscopy. Clin Gastroenterol Hepatol. 2012;10(5):501-6.

71. Gimeno-Garcia AZ, Baute JL, Hernandez G, Morales D, Gonzalez-Perez CD, Nicolas-Perez D, et al. Risk factors for inadequate bowel preparation: a validated predictive score. Endoscopy. 2017;49(6):536-43.

72. Gkolfakis P, Kapizioni C, Tziatzios G, Facciorusso A, Frazzoni L, Thomopoulos K, et al. Comparative performance and external validation of three different scores in predicting inadequate bowel preparation among Greek inpatients undergoing colonoscopy. Ann Gastroenterol. 2023;36(1):25-31.

73. Sharara AI, Bou Daher H. Bowel Cleansing Strategies After Suboptimal Bowel Preparation. Clin Gastroenterol Hepatol. 2019;17(7):1239-41.

74. He B, Xin Y, Li R, Lin FC, Zhang GM, Zhu HJ. Bowel preparation experiences and needs before follow-up colonoscopy in older adult postoperative colorectal cancer patients: A qualitative study. World J Gastrointest Surg. 2024;16(9):2968-78.

75. Xu H, He Z, Liu Y, Xu H, Liu P. Application of linaclotide in bowel preparation for colonoscopy in patients with constipation: A prospective randomized controlled study. J Gastroenterol Hepatol. 2024.

76. Gerard DP, Gerard SF, Raiser MW. High-dose polyethylene glycol-3350 and gatorade solutions for patients with previous inadequate bowel preparations for colonoscopy are safe and effective. BMC Gastroenterol. 2023;23(1):146.

77. Sey MSL, Von Renteln D, Sultanian R, McDonald C, Bouin M, Chande N, et al. A Multicenter Randomized Controlled Trial Comparing Two Bowel Cleansing Regimens for Colonoscopy After Failed Bowel Preparation. Clin Gastroenterol Hepatol. 2021.

78. Gimeno-García AZ, Hernandez G, Aldea A, Nicolás-Pérez D, Jiménez A, Carrillo M, et al. Comparison of Two Intensive Bowel Cleansing Regimens in Patients With Previous Poor Bowel Preparation: A Randomized Controlled Study. Official journal of the American College of Gastroenterology | ACG. 2017;112(6):951-8.

79. Ibanez M, Parra-Blanco A, Zaballa P, Jimenez A, Fernandez-Velazquez R, Fernandez-Sordo JO, et al. Usefulness of an intensive bowel cleansing strategy for repeat colonoscopy after preparation failure. Dis Colon Rectum. 2011;54(12):1578-84.

80. Zhong H, Hou C, Huang Z, Chen X, Zou Y, Zhang H, et al. A clinical pilot trial of an artificial intelligence-driven smart phone application of bowel preparation for colonoscopy: a randomized clinical trial. Scand J Gastroenterol. 2024:1-6.

81. Inaba A, Shinmura K, Matsuzaki H, Takeshita N, Wakabayashi M, Sunakawa H, et al. Smartphone application for artificial intelligence-based evaluation of stool state during bowel preparation before colonoscopy. Dig Endosc. 2024.

82. Kaminski MF, Thomas-Gibson S, Bugajski M, Bretthauer M, Rees CJ, Dekker E, et al. Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. Endoscopy. 2017;49(4):378-97.

83. Veldhuijzen G, de Jong MJP, Roosen CM, Siersema PD, Drenth JPH, van Esch AAJ. The gastrointestinal endoscopy satisfaction questionnaire captures patient satisfaction as a key quality indicator of gastrointestinal endoscopy. Eur J Gastroenterol Hepatol. 2020;32(7):832-7.

84. Hutchings HA, Cheung WY, Alrubaiy L, Durai D, Russell IT, Williams JG. Development and validation of the Gastrointestinal Endoscopy Satisfaction Questionnaire (GESQ). Endoscopy. 2015;47(12):1137-43.

85. Patel M, Staggs E, Thomas CS, Lukens F, Wallace M, Almansa C. Development and validation of the Mayo Clinic Bowel Prep Tolerability Questionnaire. Dig Liver Dis. 2014;46(9):808-12.

86. Gupta S, Lieberman D, Anderson JC, Burke CA, Dominitz JA, Kaltenbach T, et al. Recommendations for Follow-Up After Colonoscopy and Polypectomy: A Consensus Update by the US Multi-Society Task Force on Colorectal Cancer. Gastrointest Endosc. 2020;91(3):463-85 e5.

87. Shaukat A, Kahi CJ, Burke CA, Rabeneck L, Sauer BG, Rex DK. ACG Clinical Guidelines: Colorectal Cancer Screening 2021. Am J Gastroenterol. 2021;116(3):458-79.

88. Hassan C, Antonelli G, Dumonceau JM, Regula J, Bretthauer M, Chaussade S, et al. Post-polypectomy colonoscopy surveillance: European Society of Gastrointestinal Endoscopy (ESGE) Guideline - Update 2020. Endoscopy. 2020;52(8):687-700.

89. Muto T, Bussey HJ, Morson BC. The evolution of cancer of the colon and rectum. Cancer. 1975;36(6):2251-70.

90. Liang Q, Mukama T, Sundquist K, Sundquist J, Brenner H, Kharazmi E, et al. Longer Interval Between First Colonoscopy With Negative Findings for Colorectal Cancer and Repeat Colonoscopy. JAMA Oncology. 2024.

91. Kaminski MF, Regula J, Kraszewska E, Polkowski M, Wojciechowska U, Didkowska J, et al. Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med. 2010;362(19):1795-803.

92. Rex DK, Anderson JC, Butterly LF, Day LW, Dominitz JA, Kaltenbach T, et al. Quality indicators for colonoscopy. Gastrointest Endosc. 2024;100(3):352-81.

93. Bronzwaer MES, Depla A, van Lelyveld N, Spanier BWM, Oosterhout YH, van Leerdam ME, et al. Quality assurance of colonoscopy within the Dutch national colorectal cancer screening program. Gastrointest Endosc. 2019;89(1):1-13.

94. Zarandi-Nowroozi M, Taghiakbari M, Barkun A, Pohl H, Nauche B, Chagnon M, et al. Effect of fecal immunochemical test cut-off levels on adenoma detection rate: a systematic review and meta-analysis. Scand J Gastroenterol. 2024:1-11.

95. Ammann C, Maqkaj R, Schneider MA, Hehl SJ, Fritsch R, Pohl D, et al. Detection rate of colorectal cancer by routine colonoscopy is comparable in patients aged 45-49 and 50-54 years. Swiss Med Wkly. 2024;154:3769.

96. Wisse PHA, Erler NS, de Boer SY, den Hartog B, Oudkerk Pool M, Terhaar Sive Droste JS, et al. Adenoma Detection Rate and Risk for Interval Postcolonoscopy Colorectal Cancer in Fecal Immunochemical Test-Based Screening : A Population-Based Cohort Study. Ann Intern Med. 2022;175(10):1366-73.

97. Hassan C, Piovani D, Spadaccini M, Parigi T, Khalaf K, Facciorusso A, et al. Variability in Adenoma Detection Rate in Control Groups of Randomized Colonoscopy Trials. Gastrointest Endosc. 2022.

98. Theunissen F, Ter Borg PCJ, Ouwendijk RJT, Bruno MJ, Siersema PD. Overview of a national endoscopy database: The Trans.IT database and its impact on data registration quality. United European Gastroenterol J. 2024.

99. Pakneshan S, Moy N, O'Connor S, Hourigan L, Messmann H, Shah A, et al. Costs and benefits of a formal quality framework for colonoscopy: Economic evaluation. Endosc Int Open. 2024;12(11):E1334-e41.

100. Aziz M, Haghbin H, Sayeh W, Alfatlawi H, Gangwani MK, Sohail AH, et al. Comparison of Artificial Intelligence With Other Interventions to Improve Adenoma Detection Rate for Colonoscopy: A Network Meta-analysis. J Clin Gastroenterol. 2024;58(2):143-55.

101. Weissman S, Mehta TI, Stein DJ, Tripathi K, Rosenwald N, Kolli S, et al. Comparative Efficacy of Endoscopic Assist Devices on Colonic Adenoma Detection: A Systematic Review With Network Meta-analysis. J Clin Gastroenterol. 2022;56(10):889-94.

102. Gkolfakis P, Tziatzios G, Facciorusso A, Muscatiello N, Triantafyllou K. Meta-analysis indicates that add-on devices and new endoscopes reduce colonoscopy adenoma miss rate. Eur J Gastroenterol Hepatol. 2018;30(12):1482-90.
103. Wei MT, Fay S, Yung D, Ladabaum U, Kopylov U. Artificial Intelligence-Assisted Colonoscopy in Real-World Clinical Practice: A Systematic Review and Meta-Analysis. Clin Transl Gastroenterol. 2024;15(3):e00671.
104. Khan R, Ruan Y, Yuan Y, Khalaf K, Sabrie NS, Gimpaya N, et al. Relative Efficacies of Interventions to Improve the Quality of Screening-related Colonoscopy: A Systematic Review and Network Meta-analysis of Randomized Controlled Trials. Gastroenterology. 2024.
105. Repici A, Badalamenti M, Maselli R, Correale L, Radaelli F, Rondonotti E, et al. Efficacy of Real-Time Computer-Aided Detection of Colorectal Neoplasia in a Randomized Trial. Gastroenterology. 2020;159(2):512-20 e7.
106. Jin XF, Ma YH, Shi JW, Cai JT. Efficacy of Artificial Intelligence in Reducing Miss Rates of Gastrointestinal Adenomas, Polyps, and Sessile Serrated Lesions: A Meta-analysis of Randomized Controlled Trials. Gastrointest Endosc. 2024.
107. Soleymanjahi S, Huebner J, Elmansy L, Rajashekar N, Lüdtke N, Paracha R, et al. Artificial Intelligence–Assisted Colonoscopy for Polyp Detection. Annals of Internal Medicine. 2024.
108. Hassan C, Spadaccini M, Mori Y, Foroutan F, Facciorusso A, Gkolfakis P, et al. Real-Time Computer-Aided Detection of Colorectal Neoplasia During Colonoscopy : A Systematic Review and Meta-analysis. Ann Intern Med. 2023;176(9):1209-20.
109. Patel HK, Mori Y, Hassan C, Rizkala T, Radadiya DK, Nathani P, et al. Lack of Effectiveness of Computer Aided Detection for Colorectal Neoplasia: A Systematic Review and Meta-Analysis of Nonrandomized Studies. Clin Gastroenterol Hepatol. 2024;22(5):971-80.e15.
110. Maas MHJ, Rath T, Spada C, Soons E, Forbes N, Kashin S, et al. A computer-aided detection system in the everyday setting of diagnostic, screening and surveillance colonoscopy: an international, randomized trial. Endoscopy. 2024.
111. Ladabaum U, Shepard J, Weng Y, Desai M, Singer SJ, Mannalithara A. Computer-aided Detection of Polyps Does Not Improve Colonoscopist Performance in a Pragmatic Implementation Trial. Gastroenterology. 2023;164(3):481-3.e6.
112. Levy I, Bruckmayer L, Klang E, Ben-Horin S, Kopylov U. Artificial Intelligence-Aided Colonoscopy Does Not Increase Adenoma Detection Rate in Routine Clinical Practice. Official journal of the American College of Gastroenterology | ACG. 2022;117(11):1871-3.
113. McGill SK, Rosenman J, Wang R, Ma R, Frahm JM, Pizer S. Artificial intelligence identifies and quantifies colonoscopy blind spots. Endoscopy. 2021;53(12):1284-6.
114. Moons LM, Gralnek IM, Siersema PD. Techniques and technologies to maximize mucosal exposure. Gastrointest Endosc Clin N Am. 2015;25(2):199-210.
115. Mori Y, Wang P, Løberg M, Misawa M, Repici A, Spadaccini M, et al. Impact of Artificial Intelligence on Colonoscopy Surveillance After Polyp Removal: A Pooled Analysis of Randomized Trials. Clin Gastroenterol Hepatol. 2023;21(4):949-59.e2.
116. Nagai M, Ishibashi F, Okusa K, Mochida K, Ozaki E, Morishita T, et al. Optimal visual gaze pattern of endoscopists for improving adenoma detection during colonoscopy. Gastrointest Endosc. 2024.
117. Schult AL, Hoff G, Holme O, Botteri E, Seip B, Ranheim Randel K, et al. Colonoscopy quality improvement after initial training: A cross-sectional study of intensive short-term training. Endosc Int Open. 2023;11(1):E117-E27.
118. Kim DH, Park SJ, Cheon JH, Kim TI, Kim WH, Hong SP. Does a Pre-Training Program Influence Colonoscopy Proficiency during Fellowship? PLoS One. 2016;11(10):e0164360.

119. Meulen LWT, Haasnoot KJC, Vlug MS, Wolfhagen FHJ, Baven-Pronk M, van der Voorn M, et al. Training in optical diagnosis in community hospitals is associated with improved recognition and treatment of T1 CRC: a prospective multicenter intervention study (OPTICAL II). Endoscopy. 2024.

120. Gong D, Wu L, Zhang J, Mu G, Shen L, Liu J, et al. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study. Lancet Gastroenterol Hepatol. 2020;5(4):352-61.

121. Lui TKL, Ko MKL, Liu JJ, Xiao X, Leung WK. Artificial intelligence-assisted real-time monitoring of effective withdrawal time during colonoscopy: a novel quality marker of colonoscopy. Gastrointest Endosc. 2024;99(3):419-27.e6.

122. Shaukat A, Kaltenbach T, Dominitz JA, Robertson DJ, Anderson JC, Cruise M, et al. Endoscopic Recognition and Management Strategies for Malignant Colorectal Polyps: Recommendations of the US Multi-Society Task Force on Colorectal Cancer. Gastroenterology. 2020;159(5):1916-34 e2.

123. Chaptini L, Chaaya A, Depalma F, Hunter K, Peikin S, Laine L. Variation in polyp size estimation among endoscopists and impact on surveillance intervals. Gastrointest Endosc. 2014;80(4):652-9.

124. Hassan C, Repici A, Rex DK. Addressing bias in polyp size measurement. Endoscopy. 2016;48(10):881-3.

125. van Bokhorst QNE, Houwen B, Hazewinkel Y, Fockens P, Dekker E. Advances in artificial intelligence and computer science for computer-aided diagnosis of colorectal polyps: current status. Endosc Int Open. 2023;11(8):E752-e67.

126. Samarasena J, Yang D, Berzin TM. AGA Clinical Practice Update on the Role of Artificial Intelligence in Colon Polyp Diagnosis and Management: Commentary. Gastroenterology. 2023;165(6):1568-73.

127. Vinsard DG, Mori Y, Misawa M, Kudo S-e, Rastogi A, Bagci U, et al. Quality assurance of computer-aided detection and diagnosis in colonoscopy. Gastrointestinal Endoscopy. 2019;90(1):55-63.

128. Hassan C, Rizkala T, Mori Y, Spadaccini M, Misawa M, Antonelli G, et al. Computer-aided diagnosis for the resect-and-discard strategy for colorectal polyps: a systematic review and meta-analysis. Lancet Gastroenterol Hepatol. 2024.

129. Ruco A, Moineddin R, Sutradhar R, Tinmouth J, Li Q, Rabeneck L, et al. Duration of risk reduction in colorectal cancer incidence and mortality after a complete colonoscopy in Ontario, Canada: a population-based cohort study. The Lancet Gastroenterology & Hepatology. 2024.

130. Knudsen MD, Wang K, Wang L, Polychronidis G, Berstad P, Hjartåker A, et al. Colorectal Cancer Incidence and Mortality After Negative Colonoscopy Screening Results. JAMA Oncol. 2024.

131. Wooldrage K, Robbins EC, Duffy SW, Cross AJ. Long-term effects of once-only flexible sigmoidoscopy screening on colorectal cancer incidence and mortality: 21-year follow-up of the UK Flexible Sigmoidoscopy Screening randomised controlled trial. Lancet Gastroenterol Hepatol. 2024.

132. Toes-Zoutendijk E, van de Schootbrugge-Vandermeer HJ, Katsara MA, de Jonge L, Spaander MCW, van Vuuren AJ, et al. Harm-to-Benefit Ratio of Fecal Immunochemical Test-Based Screening for Colorectal Cancer Given Prior Fecal Hemoglobin Concentrations. Clin Gastroenterol Hepatol. 2024.

133. Senore C, Zappa M, Campari C, Crotta S, Armaroli P, Arrigoni A, et al. Faecal haemoglobin concentration among subjects with negative FIT results is associated with the detection rate of neoplasia at subsequent rounds: a prospective study in the context of population based screening programmes in Italy. Gut. 2020;69(3):523.

134. Ayling RM, Cotter F. Diagnostic application of the ColonFlag AI tool in combination with faecal immunochemical test in patients on an urgent lower gastrointestinal cancer pathway. BMJ Open Gastroenterol. 2024;11(1).

Summary

Colorectal cancer (CRC) remains a significant global health burden, ranking third in cancer incidence and fourth in cancer mortality worldwide. Early detection and prevention through screening programs, have proven effective in reducing CRC-related morbidity and mortality. The two-step screening process consists of a non-invasive screening method to select the population at risk for CRC or precursor lesions, followed by colonoscopy to detect and treat lesions if appropriate. However, limitations at each step persist. This includes, among others, the suboptimal accuracy of current non-invasive screening methods that can lead to both unnecessary colonoscopies and missed cancers and may form barriers to colonoscopy adherence. Additionally, adequate bowel preparation represents the foundation for high-quality colonoscopy screening but faces several issues leading to inadequate bowel preparation. This is primarily due to the burdensome process, which is often cited as a deterrent to colonoscopy. Addressing these challenges could result in in lower post-colonoscopy CRC incidence.

The aims of this thesis were to reduce preventable cases of CRC, by investigating volatile organic compound (VOC) analysis using an electronic nose (e-nose) as alternative non-invasive screening test, and by improving bowel preparation strategies to enhance colonoscopy quality and patient tolerability.

Part I - Can CRC screening be improved by a breath test?

Chapter 2 evaluated the diagnostic potential of VOC analysis via e-nose technology as non-invasive CRC screening tool. A multicenter prospective study with external validation involving 3,469 participants showed an area under the receiver operating characteristics curve (AUC) of 0.542 (95% CI 0.495-0.589), with sensitivity and specificity of 39.5% and 68.3%, respectively. Despite high willingness to repeat the breath test (95.3%), we found that the e-nose lacked sufficient accuracy for CRC detection in its current form. Key challenges included a low signal-to-noise ratio, environmental influences, and inter-device variability. This is representative of the current limitations of VOC analysis with e-nose technology for CRC screening and emphasizes the importance of external validation studies before clinical application. Future efforts should focus on refining e-nose technology, standardizing methodologies, and addressing inter-device variability.

Building on these findings, in **Chapter 3**, a simulation data-based study was conducted to assess methodological barriers in e-nose studies. We used breath test data from healthy subjects in previous studies to design a simulation data

set, enabling us to synthetically increase or decrease the disease prevalence and additionally assess the effect of using multiple e-nose devices for training and validating a model that can robustly predict neoplasia. This revealed that low disease prevalence and the use of multiple e-nose devices significantly increased the required sample size to achieve robust model performance. This study underscores the need for careful consideration of disease prevalence and device consistency in study designs for VOC-based diagnostic test studies. Our findings support the importance of collecting large, high-quality datasets and developing robust prediction models to improve diagnostic performance.

Part II – Can bowel preparation be improved by lowering the burden while maintaining efficacy?

In **Chapter 4**, we present a systematic review and meta-analysis assessing the efficacy of ultra-low volume (≤1L) bowel preparation solutions, such as sodium picosulfate/magnesium citrate and 1L polyethylene glycol with ascorbate (1L-PEG+Asc). Pooled analysis demonstrated adequate bowel preparation rates of 75.2% to 92.1%, with adenoma detection rates (ADR) ranging from 30.4% to 40.9%. While ultra-low volume preparations are suggested to be equally effective to other bowel preparation formulations in literature, they did not consistently meet the 90% adequacy benchmark recommended by current guidelines. The findings in this meta-analysis implicate that ultra-low volume bowel preparations offer an alternative to improve patient tolerability but require further optimization to consistently meet adequacy benchmarks.

In **Chapter 5**, we performed a randomized controlled trial comparing 1L-PEG+Asc to 2L-PEG+Asc preparations in terms of efficacy, tolerability, and patient-reported outcomes. The results showed non-inferior bowel cleansing efficacy with 1L-PEGA (96.1% vs 96.4% adequate preparation in patients not at risk for inadequate bowel preparation), alongside improved willingness to repeat the preparation. Additionally, tolerability was the most influential factor for willingness to repeat (OR 0.053-0.225 in case of difficult or fair tolerability). Other contributing factors were side effects and satisfaction on the general endoscopic procedure. These findings support using lower-volume preparations in clinical practice to enhance adherence and patient experience to bowel preparation protocols.

In **Chapter 6,** an intraprocedural bowel cleansing system (BCS) was evaluated in a multicenter feasibility study involving patients with a history of inadequate bowel preparation. The system increased the Boston Bowel Preparation Scale (BBPS) score from a median of 1-2-2 to 3-3-3, achieving adequate preparation in 88.6%

of patients. This approach may therefore reduce the need for repeat procedures and improved overall colonoscopy completion rates. An important limitation however is the decreased maneuverability that comes with the oversleeve of the BCS, resulting in a cecal intubation rate of 88.6%. As this is below the 90% quality benchmark, adequate patient selection and device optimization is advised to prevent incomplete procedures.

In conclusion, this thesis aimed to reduce preventable CRC cases by exploring breath testing as alternative screening modality and improving bowel preparation methods. VOC-based breath testing using an e-nose requires substantial further refinement before clinical implementation comes in sight. However, the advancements in bowel preparation strategies have immediate potential to enhance CRC screening colonoscopies. This thesis highlights the importance of tolerability in bowel preparation. For patients with recurrent inadequate preparations, using an intraprocedural cleansing devices appears feasible, but needs finetuning for use in clinical practice. These findings support more patient-centered colorectal cancer screening, potentially boosting participation and further improving outcomes.

Nederlandse samenvatting (ook voor niet-ingewijden)

Darmkanker (CRC) blijft wereldwijd een aanzienlijke belasting voor de volksgezondheid en staat op de derde plaats qua incidentie en vierde qua sterfte door kanker. Vroege opsporing en preventie via screeningsprogramma's zijn bewezen effectief in het verminderen van CRC-gerelateerde morbiditeit en mortaliteit. Het twee-staps-screeningsproces bestaat uit een niet-invasieve ontlastingstest om de mensen met verhoogd risico op CRC of de voorloper poliepen te selecteren, gevolgd door een darmonderzoek om deze op te sporen en indien nodig te behandelen. Er blijven echter beperkingen bestaan in beide stappen. De suboptimale nauwkeurigheid van de huidige ontlastingstest uit stap 1 leidt tot zowel onnodige darmonderzoeken, als het missen van kanker door een fout-negatieve ontlastingstest, en vormt mogelijk een barrière voor de deelname aan het bevolkingsonderzoek.

Daarnaast vormt een adequate darmvoorbereiding (de darm moet helemaal schoon zijn van ontlasting) de basis voor een hoogwaardig darmonderzoek, maar dit proces kent diverse moeilijkheden die leiden tot inadequate voorbereiding. Het grootste probleem vormt de grote belasting die de darmvoorbereiding vraagt van patiënten, omdat zij liters aan (veelal vies-smakende) laxeermiddelen moeten drinken. Dit wordt bovendien vaak genoemd als reden om het darmonderzoek uit te stellen of zelfs niet deel te nemen.

De doelstellingen van dit proefschrift waren het verminderen van vermijdbare gevallen van CRC door 1) het onderzoeken van vluchtige organische stoffen (VOC-analyse) in uitgeademde lucht door middel van een elektronische neus (e-nose) als alternatieve niet-invasieve screeningstest, en 2) door het verbeteren van darmvoorbereidingsstrategieën om de kwaliteit van darmonderzoeken en de tolerantie voor patiënten te verhogen.

Deel I - Kan CRC-screening worden verbeterd met een ademtest?

Hoofdstuk 2 evalueerde VOC-analyse via e-nose technologie als niet-invasieve screeningstest voor CRC. Een multicenter prospectieve studie met externe validatie, waarbij 3.469 deelnemers betrokken waren, toonde een *oppervlakte onder de receiver operating characteristics curve* (AUC) van 0,542 (95% betrouwbaarheidsinterval 0,495-0,589), met een sensitiviteit en specificiteit van respectievelijk 39,5% en 68,3%. Dit betekent dat de e-nose in deze studie slechts marginaal beter is dan toeval, vergelijkbaar met de uitkomst van een willekeurige gok. Ondanks een hoge bereidheid om de ademtest te herhalen door de deelnemers (95,3%), bleek

de e-nose dus onvoldoende nauwkeurig voor CRC-detectie in de huidige vorm. Belangrijke uitdagingen waren een beïnvloeding door ruis, omgevingsinvloeden en variabiliteit tussen de gebruikte apparaten. Dit laat de huidige beperkingen van VOC-analyse via e-nose technologie voor CRC-screening zien en benadrukt het belang van externe validatiestudies voordat klinische toepassing mogelijk is. Toekomstig onderzoek zou zich moeten richten op het verfijnen van de e-nose technologie, het standaardiseren van methodologieën en het aanpakken van variabiliteit tussen de gebruikte e-nose apparaten.

Voortbouwend op deze bevindingen werd in **Hoofdstuk 3** een op simulatie-data gebaseerde studie uitgevoerd om methodologische uitdagingen in e-nose studies te beoordelen. We gebruikten ademtestgegevens van gezonde proefpersonen uit eerdere studies om een simulatie dataset te ontwerpen, waarmee we de ziekteprevalentie synthetisch konden verhogen of verlagen. Daarnaast konden we het effect beoordelen van het gebruik van meerdere e-nose apparaten bij het trainen en valideren van een AI-model dat kanker robuust kan voorspellen. Hieruit bleek dat een lage ziekteprevalentie en het gebruik van meerdere apparaten de benodigd aantal deelnemers in een studie aanzienlijk verhoogden om dezelfde resultaten te bereiken. Deze studie benadrukt daarom de noodzaak van zorgvuldige overweging van de ziekteprevalentie en het gebruik van meerdere e-nose apparaten in een studie bij het ontwerp van diagnostische studies die e-nose apparaten gebruiken. Onze bevindingen ondersteunen het belang van het verzamelen van grote, hoogwaardige datasets en het ontwikkelen van robuuste voorspellingsmodellen om de diagnostische prestaties te verbeteren.

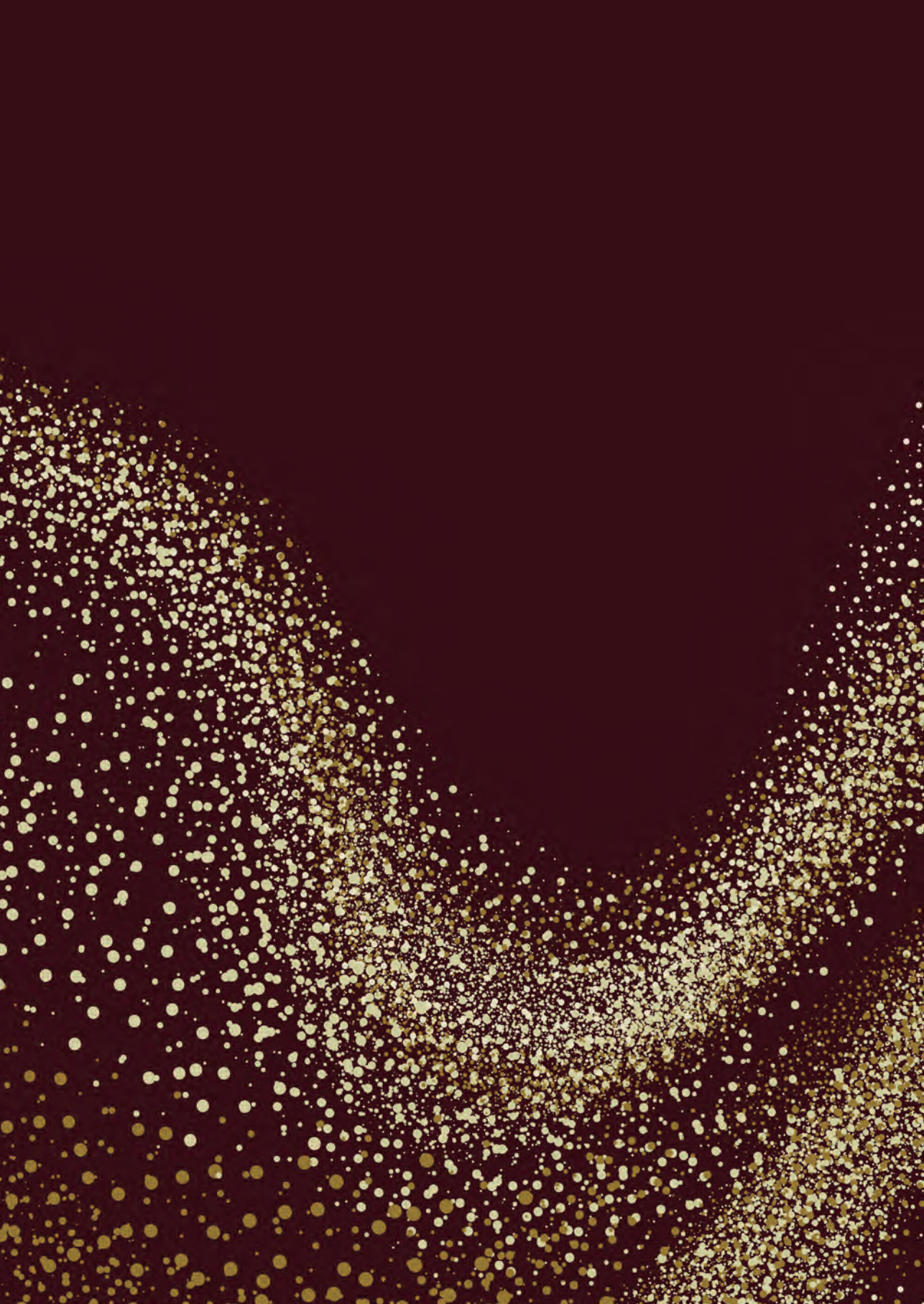
Deel II – Kan darmvoorbereiding worden verbeterd door de belasting te verlagen zonder in te leveren op effectiviteit?

In **Hoofdstuk 4** presenteren we een systematische review en meta-analyse waarin de effectiviteit van ultra-laag volume (≤1L) darmvoorbereidingsoplossingen, zoals natriumpicosulfaat/magnesiumcitraat en 1L polyethyleenglycol met ascorbaat (vitamine C) (1L-PEG+Asc), werd beoordeeld. De gepoolde analyse toonde adequate darmvoorbereidingspercentages van 75,2% tot 92,1%, met adenoomdetectie percentages (ADR) variërend van 30,4% tot 40,9%. Hoewel ultra-laag volume voorbereidingen volgens de wetenschappelijke literatuur even effectief lijken te zijn als andere middelen, voldeden ze in onze studie niet consequent aan de 90%-adequate darmvoorbereiding norm die wordt aanbevolen door de huidige richtlijnen. Deze meta-analyse suggereert dat ultra-laag volume darmvoorbereidingen een alternatief bieden om de tolerantie van patiënten te verbeteren, maar verdere optimalisatie is nodig om consistent aan de normen te voldoen.

In **Hoofdstuk 5** voerden we een gerandomiseerd onderzoek uit waarin 1L-PEG+Asc darmvoorbereiding werd vergeleken met 2L-PEG+Asc darmvoorbereiding op het gebied van effectiviteit, tolerantie en door patiënten gerapporteerde uitkomsten over invloed op werk en kwaliteit van leven. De resultaten toonden dat 1L-PEG+Asc net zo goed werkte als 2L-PEG+Asc (96,1% versus 96,4% adequate voorbereiding bij patiënten zonder risico op inadequate voorbereiding). Daarnaast waren mensen uit de 1L-PEG+Asc groep meer bereid om dit middel nogmaals te gebruiken. Tolerantie bleek de meest invloedrijke factor te zijn voor de bereidheid om het middel nogmaals te gebruiken (OR 0,053-0,225 bij moeilijke of matige tolerantie). Andere bijdragende factoren waren het optreden van bijwerkingen, en tevredenheid over de endoscopische procedure. De bevindingen van deze studie ondersteunen het gebruik van darmvoorbereidingsmiddelen met een lager volume in de klinische praktijk om de therapietrouw en patiëntervaring te verbeteren.

In **Hoofdstuk 6** werd een apparaat geëvalueerd dat de darmen tijdens het darmonderzoek kan schoonspoelen (BCS), in een multicenter haalbaarheidsstudie bij patiënten met een voorgeschiedenis van inadequate voorbereiding. Het systeem verbeterde de Boston Bowel Preparation Scale-score van een mediaan van 1-2-2 naar 3-3-3, waarbij 88,6% van de patiënten een adequate voorbereiding bereikte. Deze benadering kan daardoor de noodzaak tot herhalen van darmonderzoeken verminderen. Een belangrijke beperking is echter de verminderde manoeuvreerbaarheid van de colonoscoop door de *oversleeve* van het BCS, wat resulteerde in een coecum intubatiepercentage van 88,6% (dus 11,4% van de darmonderzoeken waren niet volledig). Aangezien dit onder de kwaliteitsnorm van 90% ligt, wordt adequate patiëntselectie en optimalisatie van het apparaat aanbevolen om incomplete procedures te voorkomen.

Concluderend, dit proefschrift had als doel het aantal vermijdbare CRC-gevallen te verminderen door ademtesten te verkennen als alternatieve screeningsmethode en door darmvoorbereidingsmethoden te verbeteren. Ademtesten met een e-nose vereisen aanzienlijke verdere verbetering voordat klinische implementatie mogelijk is. De verbeteringen in darmvoorbereidingsstrategieën hebben echter directe potentie om CRC-screening darmonderzoeken te verbeteren. Dit proefschrift benadrukt het belang van tolerantie bij darmvoorbereiding. Voor patiënten met herhaaldelijk onvoldoende voorbereidingen lijkt het gebruik van een apparaat om de darm tijdens het darmonderzoek schoon te spoelen haalbaar, maar dit vereist verdere verfijning voor klinisch gebruik. De bevindingen van dit proefschrift ondersteunen meer patiëntgerichte CRC-screening, wat mogelijk de deelname en uitkomsten verbetert.



Appendix

dankwoord
data management
list of publications
phd portfolio
over de auteur

The recipe is simple
Dream big
And when you think
The dream is big enough
Triple it in size
Once you've visualized
The weight of what you're capable of
Stretch it past the horizon of your mind
Until you see into the future
Sink into what's possible
Then come back and
Get to work

-Rupi Kaur

Dankwoord

Promoveren was een avontuur dat ik niet alleen heb afgelegd. Dit proefschrift is niet alleen mijn werk, maar ook het resultaat van alle mensen die mij gedurende dit traject hebben gesteund en geïnspireerd. Graag wil ik daarom een aantal mensen bijzonder bedanken.

Allereerst wil ik de vele patiënten die hebben meegedaan aan de studies in dit proefschrift bedanken. Zonder deze onbaatzuchtige deelname was mijn proefschrift er niet geweest.

Leden van de manuscriptcommissie, bedankt voor het lezen en beoordelen van mijn proefschrift.

Prof. Dr. Siersema, beste Peter, jouw optimisme en de vrijheid en aanmoediging die je me hebt gegeven om stappen te zetten zonder een tot in de puntjes uitgewerkt plan, hebben me stressbestendigheid, flexibiliteit en onafhankelijkheid bijgebracht. Je leerde me dat soms juist het onbekende de meest waardevolle inzichten oplevert. Ook onze samenwerking met bedrijven was een verrijkende ervaring, die deuren hebben geopend.

Dr. Tan, beste Adriaan, vanaf het begin nam je me onder je vleugels en begeleidde je me bij alle keuzes, groot en klein. Je mentorrol maakte mijn promotietraject niet alleen inhoudelijk sterker, maar ook een stuk aangenamer. Je hulp bij de politieke en organisatorische aspecten van onderzoek was onmisbaar.

Verder wil ik graag alle coauteurs bedanken voor hun bijdrage aan dit proefschrift, maar ook alle betrokken endoscopisten, en verpleegkundigen. Zonder jullie bijdrage hadden vele hoofdstukken uit dit proefschrift niet tot stand kunnen komen.

Dames van het stafsecretariaat, **Linda**, **Lionne**, **Rachelle**, ik kon altijd even langskomen voor een kop koffie wanneer ik weer logistieke perikelen of vragen over geldzaken had.

Heren van de eNose Company, jullie doorzettingsvermogen is bewonderingswaardig. Bedankt voor jullie steun en bijdrage, ook ver na het faillissement.

Aleyd en Joris van Norgine, Mark en Annie van MotusGI, dank voor het mede mogelijk maken van hoofdstuk 5 en 6 van mijn proefschrift.

Ook wil ik graag de studenten bedanken die me hebben geholpen bij het includeren voor de e-nose en RESULT-studie: Jip, Geanne, Kim, Meike, Silke, Ilse, Demi, Renée, Femke, Sabine, en Leanne. Zonder jullie was dit nooit zo vlot gelukt.

Daarnaast kunnen mijn PhD Council-collega's niet ontbreken: Yannick, Lara, Toine, Isa, Kim Phi, Lara, Lex, Lisanne, Merel, Milou van den Bemd, Miriam, Daan, prof. Judith Prins en dr. Marieke de Visser. Het was erg fijn en inspirerend om met jullie samen te werken, en om een inkijkje te krijgen in de organisatie van de Graduate school. Het heeft mijn horizon verbreed voorbij de medische promotietrajecten, en we hebben een aantal fantastische events neergezet!

Arts-onderzoekers MDL: Ayla, Britt, Daan, Dorien, Edo, Elsa, Fenna Beeren, Fenna Jansen, Gijs, Jasmijn, Julia, Kelly, Lia, Lieke, Lotte, Lucas, Maarten, Marleen, Melissa, Michiel, Mike, Monica, Naomi, Pepijn, Renée, Romée, Sabien, Veerle, Yonne; dit promotietraject was niet zo leuk geweest zonder jullie. Ik denk met genot terug aan onze lunches en soms second lunches en koffiepauzes bij de befaamde rode bankjes, of de discussies in de even beroemde Mark Rutte kamer. Ook de al dan niet digitale vrijmibo's had ik niet willen missen, om niet te spreken over de legendarische onderzoekers nacht! De congressen in Veldhoven, Kopenhagen, Praag, en San Diego behoorden tot de hoogtepuntjes van mijn PhD.

Ayla, Kelly, Yonne, Lieke, Lotte, Gijs, Michiel, Jasmijn, Mike, Fenna, collega's van **team Siers**, jullie maakten de afgelopen jaren niet alleen leerzaam, maar ook plezierig. Bedankt dat jullie er voor me zijn, en me helpen herinneren wat het verschil tussen ITT en PP is 😉. Kelly, bedankt dat je met jouw proefschrift de fundering voor mijn onderzoek hebt gelegd. Michiel, door onze samenwerking kon onze Discovery III studie niet soepeler verlopen, voor altijd voorstander van gezamenlijke projecten!

Britt, zo knap hoe je staande houdt tussen Nijmegen en Maastricht, en ondertussen ook nog vele kilometers maakt om te trainen. Jaloers op jouw marathontijd hoor! **Daan**, heerlijk hoe jij onder alles toch relaxed blijft, en je ervoor zorgt dat elk feestje blijft doorgaan. Ik krijg ook zin om met een camper door Europa te gaan! **Dorien**, Do! Bedankt voor jouw fijne optimisme en vrolijkheid, jouw aanwezigheid zorgt altijd voor een verbetering van de sfeer. Ook bedankt voor het gedeelde enthousiasme over *Computeren de podcast*. **Edo**, als jij er bent is het altijd gezellig. Ook bedankt dat je ondanks je rugpijn jezelf opofferde om in de woonkamer van onze AirBnB in San Diego te gaan slapen zodat de dames in de slaapkamer konden. **Fenna Beeren,** love de koffietjes en lunches in Arnhem, zo fijn om in het Rijnstate ook iemand te hebben om lekker te kletsen over promotie perikelen (en shoppen).

Fenna Jansen, wat hebben we mooie momenten meegemaakt in *California*. Ik kijk altijd met bewondering naar zowel jouw creativiteit om problemen aan te pakken, maar ook je ongelooflijke sportiviteit (waarmee je met een belachelijk lage hartrate onze monsterhike in Yosemite aflegde). Beach, fietsen, hardlopen, en ook nog moeder zijn, werken, en je PhD afmaken, chapeau! **Gijs**, jouw skills om iemands beweegredenen en gedachten laagje voor laagje af te pellen zijn legendarisch, maar ook je kookkunsten om een perfecte culinaire caprese salade tijdens ons afscheidsetentje te maken zijn ongeëvenaard. **Jasmijn**, schrijfdag-hardloop-birthday buddy, ik bewonder jouw ijzeren discipline die je zonder enige schijnbare afleiding recht op je doel beweegt. Het is inspirerend hoe je durft af te wijken van de standaard routes, maar wel overwogen je eigen weg volgt. **Julia**, jij weet niet alleen van doorpakken, maar je hebt ook altijd oog voor je mede collega als het even minder gaat. Ook bedankt dat je me hebt geïnspireerd om te gaan mealpreppen, volgens mij kent inmiddels iedereen mijn overnight oats. **Lia**, mijn perceptie van competitief spelen veranderde toen ik met jou *Undercover* speelde, meedogenloos, haha! **Lieke**, het begon als mede-Tukker, maar al vanaf onze eerste werkdag wist ik dat ik lief en leed met jou kon delen. Daarnaast was je zorgzame kant duidelijk zichtbaar toen je het me gunde dronken te worden in Praag, met nadien nog spierpijn van het dansen de hele nacht. Dank dat jij samen met Yonne mijn paranimf wil zijn! **Lotte**, jouw doorzettingsvermogen en kracht om je werk zo nauwkeurig mogelijk uit te voeren is prijzenswaardig. **Maarten**, na eerst 3 jaar samen te werken als PhD collega's, was ik erg blij om jouw woordgrappen weer te mogen ontvangen bij de interne. Jammer dat je nu alweer naar de MDL terug bent! **Marleen**, AJAX! Jouw nuchterheid is heerlijk verfrissend. **Melissa**, echt mooi hoe onze plantjes altijd in leven blijven tijdens onze vakanties en dat je presentaties foto's van glitter levercysten bevatten. **Michiel**, professor Maas voor intimi, met jou is een congres, en vooral het feest bij het congres altijd gemaakt. Verder dank voor je vergevingsgezindheid toen Lieke en ik toch nog precies te laat bij je praatje kwamen door de brakheid. **Mike,** jij weet een netwerk samen te brengen en mensen te bewegen als geen ander. **Monica**, jij weet altijd precies het goede te zeggen om mensen op hun gemak te laten voelen. Ook bedankt dat ik door jou de klassieker *Kei blij op de kinderboerderij* heb ontdekt. **Naomi**, ook als niemand meewerkt, krijg jij nog bijzonder grote projecten gedaan, heel knap. **Pepijn**, jij bent de onbetwiste syntax koning, nog maanden nadat jij alweer klaar was kon ik gebruik maken van jouw ppts. **Renée**, ik bewonder jouw vermogen om niet alleen alles altijd glashelder door te hebben, maar ook je lef om hiernaar te handelen. Ik kijk altijd weer uit naar onze koffie/lunch/ontbijt/diner afspraakjes! **Romée**, powervrouw, jij bent altijd de sprankeling op elke bijeenkomst. Ik heb bewondering voor hoe jij zo veel ballen in de lucht kan houden, en dit ook nog super succesvol doet. **Sabien**, jij

hebt altijd zulke verfrissende inzichten! **Veerle**, zo fijn dat jij iedereen kon voorzien van stekjes, maar ook tijdens covid ervoor waakte dat alle plantjes wel genoeg water kregen. **Yonne**, met jou kan je niet alleen lief en leed delen, maar ook heel veel plezier maken en daarbij compleet de tijd vergeten. Dit resulteerde dan ook in de oprichting van de hej t al heurt groepsapp samen met Lieke. Dank dat jij mijn paranimf wilt zijn!

MDL Rijnstate, bedankt voor de begeleiding. Door jullie heb ik kunnen groeien als beginnend arts-assistent.

Interne geneeskunde Rijnstate, enorm bedankt voor het warme welkom, het is alsof ik thuiskom.

Mam, Pap, jullie onvoorwaardelijke steun en vertrouwen waren de basis waarop ik dit onderzoek kon bouwen. Jullie zorgden voor een stabiele achtergrond en een luisterend oor op de juiste momenten. Anny, jij laat iedereen altijd op zijn gemak voelen. Je bent een rots in de branding. Daniel, wat goed dat je je niet door tegenslagen uit het veld laat slaan maar blijft opbouwen en groeien. Quinte (meow), lief zusje, jij weet altijd gelijk hoe ik denk. Ik ben trots op de tweede dokter in het gezin. Stiekem is diergeneeskunde misschien nog wel leuker dan humane geneeskunde.

En dan **Dion** – jouw geduld, liefde en steun zijn van onschatbare waarde geweest. Jij was er altijd, in de pieken en dalen. Zonder jou was dit traject niet alleen een stuk zwaarder geweest, maar ook veel langer. Dank je wel voor alles lieverd.

Research data management

All studies in this thesis were conducted in accordance with the Declaration of Helsinki. Each study received approval before starting from the Medical Ethics Committee (METC Oost-Nederland; file numbers: 2021-13200 (chapter 5), 2018-4543 (chapter 6)) or a waiver of ethical approval from the institutional review board (CMO Radboudumc; file numbers: 2020-6184 (chapter 2-3)), depending on the study's nature. All participants provided written informed consent before any study procedures. Consent was also obtained for reuse of the (pseudonymized) data for future research.

For studies involving human subjects (Chapters 2, 5, and 6), participant privacy was safeguarded through pseudonymization. Data handling complied with the European General Data Protection Regulation (GDPR). The pseudonymization key was stored separately from the study data and was accessible only to the principal investigator and select research team members whose roles required access.

Data for Chapters 2, 5, and 6 were collected via electronic health record searches, clinical experiments, and secure online questionnaires using CastorEDC. Standardized electronic Case Report Forms were utilized for data collection and exported to SPSS or R Studio for statistical analysis. Crude breath samples were processed by the eNose Company (Zutphen, the Netherlands). Data for Chapter 3 were generated from prior studies using the Aeonose device and compared with results from Chapter 2. Chapter 4 data were sourced from published literature.

Data management followed the FAIR principles (Findability, Accessibility, Interoperability, and Reusability). All studies were registered in appropriate trial registries: clinicaltrials.gov (NCT03346005, NCT04357158 for Chapters 2 and 3; NCT05242562 for Chapter 5; NCT04700410 for Chapter 6) and PROSPERO (CRD42020181630 for Chapter 4). To ensure data transparency and reproducibility, filenames, raw and processed data, metadata, descriptive files, and analysis scripts were thoroughly documented and stored together in standardized formats. Data cleaning and analysis were performed in SPSS and R Studio, with syntaxes maintained to facilitate reproducibility. Datasets and analyses are not publicly available for reuse because permission for data sharing was not obtained, but archived for 15 years in a closed-access Data Acquisition Collection (DAC) of the Radboud data repository. The corresponding metadata are publicly accessible via DOI: 10.34973/s30f-xm61 (chapter 2); 10.34973/vc4p-e493 (chapter 4); 10.34973/bwe5-yk64 (chapter 5); 10.34973/s4a8-0m39 (chapter 6).

List of publications

This thesis

M.L.M. van Riswijk*, K.E. van Keulen*, A.C.I.T.L Tan, R.W.M. Schrauwen, W.H. de Vos tot Nederveen Cappel, P.D. Siersema, on behalf of the eNose CRC study group. *Breath testing for colorectal cancer detection in patients with a positive fecal immunochemical test – a multicenter prospective cross-sectional study with external validation*. Aliment Pharmacol Ther. 2025 Jul;62(2):208-213. doi: 10.1111/apt.70207.

M.L.M. van Riswijk, B. van Tintelen*, R. Lucas*, J. van der Palen, P.D. Siersema. *Overcoming methodological barriers in electronic nose clinical studies, a simulation data-based approach.* J Breath Res. 2025 May 9;19(3). doi: 10.1088/1752-7163/add291.

M.L.M. van Riswijk, K.E. van Keulen, P.D. Siersema. *Efficacy of ultra-low volume (≤1 L) bowel preparation fluids: Systematic review and meta-analysis.* Digestive Endoscopy 2022 Jan;34(1):13-32

M.L.M. van Riswijk, F.A. Indemans, K. Hawinkels, R.M. Schreuder, L. Wildeman, A.C.I.T.L Tan, P.D. Siersema. *Comparing low-volume vs. intermediate volume bowel preparation and its impact on work and tolerability: An open-label, non-inferiority, randomized controlled trial.* Endoscopy. 2025 Sep 3. doi: 10.1055/a-2695-0994.

M.L.M. van Riswijk, K.E. van Keulen, H. Neumann, P.D. Siersema. *An intraprocedural bowel cleansing system for difficult-to-prepare patients-A multicenter prospective feasibility study.* United European Gastroenterol J. 2024 Feb;12(1):56-65.

Other publications

Simsek M, Lissenberg-Witte BI, **van Riswijk M.L.M.**, et al. *Off-label prescriptions of drugs used for the treatment of Crohn's disease or ulcerative colitis.* Alimentary pharmacology & therapeutics 2019;49:1293-300.

M.L.M. van Riswijk, J. Grootjans, JJ Bergman. *Curious Endoscopy Corner: Pinpoint stenose van het duodenum en eosinofilie.* Medidact Gastroenterologie, uitgave 21-6-2019.

van Munster S, Nieuwenhuis E, Weusten BLAM, Alvarez Herrero L, Bogte A, Alkhalaf A, Schenk BE, Schoon EJ, Curvers W, Koch AD, van de Ven SEM, de Jonge PJF, Tang TJ, Nagengast WB, Peters FTM, Westerhof J, Houben MHMG, Bergman JJ, Pouw RE; Dutch Barrett Expert Centers. *Long-term outcomes after endoscopic treatment*

for Barrett's neoplasia with radiofrequency ablation ± endoscopic resection: results from the national Dutch database in a 10-year period. Gut. 2022 Feb;71(2):265-276. doi: 10.1136/gutjnl-2020-322615.

van Riswijk, M.L.M. and Siersema, P.D. (2021), *Letter: meta-analysis of new diagnostic methods for colorectal cancer detection–only as strong as its weakest link*. Aliment Pharmacol Ther, 54: 540-540.

Milou van Riswijk, Yonne Peters, Peter Siersema. *Kun je kanker vaststellen in iemands adem? Kanker – Over de laatste ontwikkelingen in het kankeronderzoek*. Stichting Biowetenschappen en Maatschappij, Uitgeverij Lias, 2023

van Riswijk M.L.M., Siersema PD. *Letter: Methodological Gaps Undermine Conclusions on Electronic Nose Use for Colorectal Cancer Detection-Authors' Reply*. Aliment Pharmacol Ther. 2025 Aug;62(3):374-375. doi: 10.1111/apt.70243.

PhD portfolio

Training activities	Hours
Courses	
– Basic life support (2020)	4.00
– RIHS - Introduction course for PhD candidates (2020)	15.00
– Radboudumc - Introduction day (2020)	6.00
– Stanford Medical statistics MOOC (2020), Stanford university	60.00
– Online intensive writing week (2020), Radboud in'to languages	60.00
– Feedback training (2020)	8.00
– RU - Project management for PhD candidates (2020)	56.00
– Radboudumc - eBROK course (for Radboudumc researchers working with human subjects) (2020)	26.00
– UEG course functional dyspepsia (2020)	2.00
– How to sell your science (2020) RIHS	4.00
– How to make a poster presentation (2021) RIMLS	8.00
– Fundamentals of Machine Learning for Healthcare (2021) Stanford university	36.00
– DRE Hands-on course (2021)	2.00
– How to write a rebuttal (2021) Radboud in'to languages	2.00
– Stem en presentatie training (2021) Twentify	8.00
– Workshop: Work is fun, isn't it? (2021) RIHS	2.00
– Radboudumc - Scientific integrity (2021)	20.00
– Cursus perfectionisme (2022)	15.00
– Workshop: Negotiation skills (2022) RIHS	2.00
– Managing your internship (2022) RIMLS	2.00
– RU - Analytic Storytelling (2022)	20.00
– Meet the expert: Visualization in R:from basic to advanced (2022) RIMLS	2.00
– Workshop: prepare your defence, answering questions (2022) RIHS	2.00
– Meet the expert: Adobe illustrator for beginners (2023) RIMLS	2.00
– RU - The Art of Finishing Up (2023)	10.00
– Radboudumc - Re-registration BROK (2023)	5.00
Seminars	
– Meet the expert: LinkedIn (2020)	2.00
– Social media for scientists (2020)	1.00
– Online Recruitment of Study Participants (2020)	2.00
– Workshop van Aula naar Binnenhof (2021) KNAW	3.00
– Mental Health: Breaking the taboo (2021)	1.00
– NVGE PhD Summer school (2022)	6.00
– Workshop how to chair a session (2022)	2.00
– De wetenschap over het bevolkingsonderzoek darmkanker (2022)	6.00
– RIHS PhD retreat 2022 (2022)	12.00
– T-Pensant MDL Oost Nederland (presentation) (2023)	6.00
– GI-HEP meetings with international experts (2023)	10.00
– Research integrity rounds - Patents (2023)	2.00
– Soeterbeeck GI lectures (2024)	20.00

Training activities	Hours
Conferences	
– Online CaRe symposium (2020)	2.00
– Amsterdam Virtual Live Endoscopy (2020)	8.00
– Digital Dutch Digestive Disease days (moderated poster presentation) (2021)	20.00
– United European Gastroenterology week (moderated poster presentation) (2021)	30.00
– RIHS PhD retreat 2021 (2022)	20.00
– ESGE days (Champions' Den) (2022)	30.00
– Digestive Disease Week San Diego (poster presentation) (2022)	40.00
– Dutch Digestive Disease Days 2023 (2x oral presentation (2023)	40.00
– United European Gastroenterology week (poster presentation) (2023)	50.00
Other	
– Grant writing (STIMAG) (2021)	30.00
– Review scientific publication UEG (2021)	2.00
– Grant writing (TURBO) (2021)	20.00
– Organization RIHS PhD retreat 2021 (2021)	40.00
– Grant writing (Norgine) (2021)	20.00
– RIHS PhD council secretary (2021)	120.00
– Organization RIHS PhD retreat 2022 (2022)	40.00
– RIHS PhD Council chair (2022)	120.00
– Journal Club Gastroenterology & Hepatology (weekly) (2023)	252.00
– Intervision sessions PhD students (3x yearly (2023)	22.00
– Coordination GI-Hep lectures with international experts (2023)	50.00
– Coordination Soeterbeeck GI lectures (2023)	40.00
Teaching activities	
Supervision of internships / other	
– Student coaching (Geanne) (2020)	10.00
– Student internship coaching (Demi) (2021)	30.00
– Student coaching (Kim) (2021)	10.00
– Student internship coaching (Renée) (2021)	30.00
– Student coaching (Meike) (2022)	10.00
– Student coaching (Jip) (2022)	10.00
– Student internship coaching (Femke) (2022)	30.00
– Student internship coaching (Sabine) (2022)	30.00
– ROC course for nursing students (2022)	2.00
– Student coaching (Silke) (2022)	15.00
– Gastroenterology nursing teaching course on icterus (2023)	4.00
– Student Coaching (Ilse) (2023)	30.00
– Student internship teaching (Leanne) (2023)	30.00
Total	**1689.00**

Over de auteur

Milou Lisa Marit van Riswijk werd geboren op 15 februari 1994 te Denekamp. Van jongs af aan was zij nieuwsgierig en gefascineerd door de werking van de wereld om haar heen. Al vanaf haar profielwerkstuk op de middelbare school wist zij dat ze een wetenschappelijke carrière ambieerde.

Na haar gymnasiumopleiding aan het Twents Carmel College in Oldenzaal startte zij haar studie Geneeskunde aan de Universiteit van Amsterdam. Tijdens haar studie was zij actief binnen IFMSA-UvA, waar zij cursussen, keuzevakken, en educatieve bijeenkomsten organiseerde.

In 2020 startte Milou haar promotieonderzoek aan de afdeling Maag-, Darm- en Leverziekten van het Radboudumc in Nijmegen, onder begeleiding van prof. dr. P.D. Siersema en dr. A.C.I.T.L. Tan. In 2022 won zij de ESGE Champions' Den Research Grant, waarbij zij haar onderzoeksvoorstel succesvol verdedigde voor een kritische jury. Naast haar onderzoek was zij onder andere voorzitter van de PhD Council van het Radboud Institute for Health Sciences, waar zij promovendi vertegenwoordigde en bijdroeg aan de fusie van twee graduate schools.

Haar ambitie is om arts en onderzoeker te blijven en wetenschap te combineren met patiëntenzorg. Daarnaast vindt zij het belangrijk om oog te hebben voor de persoon achter de patiënt. Door verder te kijken dan de diagnose en te begrijpen wat ziekte betekent in iemands dagelijks leven, wil zij bijdragen aan zorg die niet alleen effectief, maar ook mensgericht is. In haar vrije tijd geniet Milou van hardlopen, reizen en lezen. Zij woont in Arnhem samen met haar partner Dion.

About the author

Milou Lisa Marit van Riswijk was born on February 15, 1994, in Denekamp. From an early age, she was curious and fascinated by how the world around her worked. During her high school research project, she already knew that she aspired to a scientific career.

After completing her pre-university education (gymnasium) at Twents Carmel College in Oldenzaal, she began her medical studies at the University of Amsterdam. During her studies, she was actively involved in IFMSA-UvA, where she organized courses, electives, and educational events.

In 2020, Milou started her PhD research at the Department of Gastroenterology and Hepatology at Radboudumc in Nijmegen, under the supervision of Prof. Dr. P.D. Siersema and Dr. A.C.I.T.L. Tan. In 2022, she won the ESGE Champions' Den Research Grant, successfully defending her research proposal before a critical jury.. In addition to her research, she served as chair of the PhD Council of the Radboud Institute for Health Sciences, representing PhD candidates and contributing to the merger of two graduate schools.

Her ambition is to continue working as both a physician and researcher, combining science with patient care. Beyond medical excellence, she values seeing the person behind the patient. By looking beyond the diagnosis and understanding what illness means in someone's daily life, she aims to contribute to care that is not only effective but also patient-centered. In her free time, Milou enjoys running, traveling, and reading. She lives in Arnhem with her partner, Dion.

I get so lost
In where i want to go
I forget that the place i'm in
Is already quite magical

-Rupi Kaur (home body)

9 789465 150901 >

Radboudumc university medical center

IN·DEI·NOMINE·FELICITER

Radboud University