Am J Transl Res 2026;18(9):8227-8253
www.ajtr.org /ISSN:1943-8141/AJTR0175347

Review Article

Microbiota-assisted screening for colorectal neoplasia: advanced adenoma as the key translational endpoint-a narrative review

Xuanzhi Wang, Jiahao Li, Yali Cheng, Yang Liu, Han Li, Haoqian Tan, Yongjian Hu, Ming Zhao, Junying Liu

Department of Gastroenterology, Zhoukou Central Hospital Affiliated to Henan Medical University, Zhoukou 466000, Henan, China

Received July 14, 2026; Accepted August 12, 2026; Epub September 15, 2026; Published September 30, 2026

Abstract: Microbial differences between established colorectal cancer (CRC) and cancer-free controls are reproducible, but screening translation depends on detecting advanced adenoma and early CRC in asymptomatic populations. We examined the gut microbiota as an adjunct to fecal immunochemical testing (FIT), stool deoxyribonucleic acid (DNA) testing, and colonoscopy, with advanced adenoma as the prevention-relevant endpoint. This narrative review used a structured, purposive search of PubMed, Embase, Web of Science, Scopus, and the Cochrane Library through April 10, 2026, prioritizing human neoplasia cohorts, stage-specific biomarkers, multi-omics studies, FIT comparators, and reporting standards. Across studies, microbial signals were strongest for established CRC and weaker and less consistently reported for advanced adenoma. Quantitative evidence supports selected adjunctive approaches, including fecal *Fusobacterium nucleatum* and multitaxon panels combined with FIT, but external validation remains sparse. The 2025 pooled analysis of 3,741 metagenomes demonstrated reproducible CRC-related features across 18 cohorts but did not report advanced-adenoma-specific accuracy. Clinical utility cannot be inferred from sensitivity, specificity, or the area under the receiver operating characteristic curve alone: prevalence-dependent positive predictive value (PPV), negative predictive value (NPV), calibration, a prespecified action threshold, decision-curve net benefit, and colonoscopy burden are decisive. Translation therefore requires a locked assay and model tested prospectively before bowel preparation in a representative screening cohort, with advanced adenoma reported separately and performance compared directly with FIT or stool DNA testing. Microbiota-based markers currently merit evaluation as adjuncts, not replacements, within established screening pathways.

Keywords: Gut microbiota, colorectal neoplasia, advanced adenoma, fecal microbiome, microbiota-assisted screening, FIT, stool DNA testing

Introduction

Screening for colorectal cancer (CRC) is most valuable when it prevents cancer rather than merely identifies an established tumor. Advanced adenomas and advanced serrated lesions occupy that prevention window: they can be removed before invasion, whereas early CRC can often be treated with curative intent. Global Cancer Observatory (GLOBOCAN) 2022 estimates confirm that CRC remains a major cause of cancer incidence and mortality worldwide [1]. The biological rationale for screening is equally durable. Most sporadic cancers arise through an accumulation of molecular and histological changes over years, creating an interval in which high-risk precursor lesions can be found and removed [2-8].

Existing tests divide that task imperfectly. Colonoscopy is both diagnostic and therapeutic, but bowel preparation, procedural risk, cost, endoscopy capacity, and unequal uptake limit its use as a universal first-line test. The fecal immunochemical test (FIT) is inexpensive and scalable and performs well for bleeding cancers, yet a single round detects only a minority of advanced adenomas. Multitarget stool deoxyribonucleic acid (DNA) testing adds host molecular markers and improves detection of CRC and some advanced precursors, but it is more expensive and less specific than FIT, so

https://doi.org/10.62347/QWXP3480

more participants are referred for colonoscopy [4-7]. A new stool marker is therefore useful only if it changes a referral decision at an acceptable false-positive and resource cost.

The gut microbiota is a plausible source of such complementary information because stool contains organisms, genes, and metabolites that reflect interactions among the lumen, mucus, epithelium, immune system, and tumor-adjacent niche [9, 10]. Hemoglobin reports bleeding, and stool DNA reports selected host alterations; a microbial assay could capture strain-level virulence, ecological displacement, metabolic function, or inflammation that is not measured by either. That complementarity, rather than the performance of any single taxon, is the strongest argument for a microbiota-assisted strategy.

Evidence for a CRC-associated microbiome is now broad, and several reviews have catalogued recurrent taxa, mechanistic pathways, and diagnostic models [11-13]. The translational evidence is narrower. Many studies compare clinically detected cancers with healthy controls, enrich the dataset with disease, tune a classifier and its cut-off in the same cohort, and report an area under the curve without showing what happens to advanced adenoma detection or colonoscopy demand. Such a design can establish biological separation but cannot establish screening utility in an average-risk population.

This review therefore evaluates a specific clinical proposition: whether microbiota-derived information can improve an existing stool-based screening pathway. Advanced adenoma is treated as the key endpoint, and evidence is judged by incremental performance beyond FIT or stool DNA testing, prevalence-aware predictive values, calibration, prespecified action thresholds, external transportability, and the number of colonoscopies generated. The article is reported in accordance with the Narrative Review reporting checklist.

Literature search approach

This article is a narrative review informed by a structured and purposive literature search, rather than a formal systematic or scoping review. The search was designed to identify representative, clinically relevant, and mechanistically important evidence related to microbiota-assisted colorectal neoplasia screening, with particular attention to advanced adenoma as a prevention-relevant endpoint.

Searches were updated on April 10, 2026, and covered PubMed, Embase, Web of Science, Scopus, and the Cochrane Library. Search concepts included colorectal cancer, colorectal adenoma, advanced adenoma, advanced serrated lesions, the adenoma-carcinoma sequence, gut microbiota, fecal microbiome, mucosa-associated microbiota, microbial biomarkers, fecal immunochemical testing, stool DNA testing, metagenomics, metabolomics, 16S ribosomal ribonucleic acid (rRNA) sequencing, machine learning, external validation, calibration, predictive value, and decision-curve analysis.

Priority was given to human colorectal neoplasia cohorts that reported advanced adenoma or early CRC, studies of microbiota or multi-omics biomarkers, direct FIT or stool DNA comparisons, mechanistic work with a clear link to assay design, screening guidelines, and methodological standards. The reporting frameworks considered included Strengthening the Organization and Reporting of Microbiome Studies (STORMS), Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD), and Standards for Reporting Diagnostic Accuracy Studies (STARD). Foundational experimental studies were included when they clarified a measurable link among a microbial strain, virulence factor, metabolite, host response, and colorectal carcinogenesis.

Because the aim was conceptual and translational synthesis rather than exhaustive evidence mapping, duplicate screening, formal risk-of-bias grading, a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram, and quantitative pooling were not performed. Study selection was purposive, and the review may not capture every microbiome report. No new statistical analysis or meta-analysis was conducted, and no statistical software was used. The review should therefore not be read as a source of pooled diagnostic-accuracy estimates; its purpose is to compare endpoint definitions, identify implementation gaps, and specify the evi-

Table 1. Search strategy summary

Items	Specification
Date of search	Searches were updated on April 10, 2026.
Databases and other sources searched	PubMed, Embase, Web of Science, Scopus, and the Cochrane Library. Additional relevant references were identified from key reviews, original studies, screening guidelines, and methodological reporting standards.
Search terms used	Search concepts included colorectal cancer, colorectal adenoma, advanced adenoma, colorectal neoplasia, adenoma-carcinoma sequence, gut microbiota, fecal microbiome, mucosa-associated microbiota, microbial biomarkers, fecal immunochemical test, FIT, stool DNA testing, metagenomics, metabolomics, 16S rRNA sequencing, machine learning, screening, and early detection.
Timeframe	From database inception to April 10, 2026.
Language and publication status	English-language peer-reviewed full-text articles were prioritized. Conference abstracts without sufficient methodological or outcome details were not prioritized. Foundational mechanistic studies, guidelines, and reporting standards were included when relevant to the review question.
Inclusion criteria	Priority was given to human colorectal neoplasia cohorts; studies reporting advanced adenoma, advanced colorectal neoplasia, or early CRC endpoints; microbiota or multi-omics biomarker studies; studies comparing microbial markers with FIT or stool DNA testing; mechanistic studies relevant to assay design; colorectal cancer screening guidelines; and methodological or reporting standards such as STORMS, TRIPOD, and STARD.
Exclusion criteria	Studies were not prioritized if they were unrelated to colorectal neoplasia screening, lacked relevance to microbiota-assisted detection or assay design, were case reports, were conference abstracts without sufficient methodological or outcome details, or were animal-only studies without clear mechanistic relevance to colorectal carcinogenesis or biomarker development.
Selection process	Literature selection was conducted purposively by the authors and was based on conceptual, clinical, and translational relevance to microbiota-assisted colorectal neoplasia screening. Because this was a narrative review, selection was not performed independently in duplicate, and no formal risk-of-bias assessment, PRISMA flow diagram, or quantitative synthesis was performed.
Additional considerations	Particular attention was given to studies that reported advanced adenoma separately, evaluated FIT-adjunctive performance, addressed stool DNA comparator pathways, or provided mechanistic rationale for measurable stool-based biomarkers. Advanced colorectal neoplasia was interpreted cautiously when CRC and advanced adenoma were combined.

Note: CRC, colorectal cancer; DNA, deoxyribonucleic acid; FIT, fecal immunochemical test; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; rRNA, ribosomal ribonucleic acid; STARD, Standards for Reporting Diagnostic Accuracy Studies; STORMS, Strengthening The Organization and Reporting of Microbiome Studies; TRIPOD, Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis.

dence needed for a microbiota-assisted screening decision.

A summary of the search strategy is provided in **Table 1**. A representative detailed PubMed search strategy is provided in Supplementary Table 3.

The study selection process was conducted in a purposive manner. Records were first assessed for relevance to microbiota-assisted colorectal neoplasia screening, with emphasis on human stool-based studies, advanced adenoma or advanced colorectal neoplasia endpoints, FIT or stool DNA comparators, external validation, and translational relevance to clinical screening pathways. When multiple publications addressed similar topics, priority was given to studies with larger or screening-based populations, clearer endpoint definitions, external or cross-cohort validation, quantitative performance reporting, and direct relevance to assay design or implementation. Mechanistic studies were selected when they clarified biologically plausible links between microbial taxa, virulence factors, functional pathways, meta-

bolites, host responses, and colorectal carcinogenesis. Because this was a narrative review rather than a systematic review, a formal PRISMA-style record count was not maintained. To avoid implying systematic-review methodology, exact screening counts are not reported. However, the selection principles and prioritization criteria are described to improve transparency.

Microbial alterations across the colorectal adenoma-carcinoma sequence

Microbial changes from healthy controls to adenoma

Fecal microbiome studies suggest that microbial changes can be detected before invasive cancer, but the adenoma signal is generally weaker and less consistent than the CRC signal. Early stool sequencing studies found that microbial community data could differentiate healthy controls, adenoma, and carcinoma groups, particularly when combined with clinical risk factors [14]. Later evidence syntheses reported recurring associations but emphasized variation in sample handling, sequencing platform, taxonomy pipelines, and case definitions [11, 12]. In adenoma, enriched taxa have included members of oral-associated anaerobic communities and pro-inflammatory genera, while depleted taxa often include butyrate producers. However, adenoma definitions vary substantially, and small non-advanced adenomas may not create a strong or stable fecal signal.

Fusobacterium nucleatum (F. nucleatum) has been reported in adenoma and CRC tissues, but its enrichment is more consistently observed in carcinoma than in early adenoma [15, 16]. FadA expression was reported to be higher in adenomas and adenocarcinomas than in normal tissue in mechanistic work [17], suggesting biomarker relevance at a premalignant stage. Nevertheless, fecal *F. nucleatum* alone has insufficient performance and can be influenced by oral health, geography, diet, and analytical method; quantitative fecal *F. nucleatum* has therefore been studied mainly as an adjunct to existing stool-based screening rather than as a stand-alone test [18]. Enterotoxigenic *Bacteroides fragilis* (ETBF) is mechanistically compelling because *B. fragilis* toxin (BFT) can affect epithelial integrity [19, 20]. The polyketide synthase-positive (pks+) *E. coli* lineage is relevant because colibactin can induce DNA damage and characteristic mutational signatures [21, 22]. Multi-cohort fecal metagenomic studies support community-level signatures [23, 24], but marker-panel performance for early lesions remains under-validated [25, 26].

The adenoma stage also illustrates why a simple list of bacteria is insufficient. A small distal tubular adenoma may have little impact on stool microbial composition, whereas a larger villous or high-grade lesion may alter local inflammation, mucus, and luminal shedding. Lesion location matters because right-sided lesions may interact with different microbial gradients, bile acid pools, transit times, and biofilm structures than left-sided lesions. Serrated lesions add another layer of complexity because they may be flat, proximal, mucus-rich, and molecularly distinct from conventional adenomas. Many microbiome studies do not separate these lesion subtypes in sufficient detail, so stage-specific conclusions should be treated as provisional unless the extracted study explicitly reports advanced adenoma or serrated lesion performance.

Definition and endpoint heterogeneity of advanced adenoma

Advanced adenoma is commonly defined using histopathological and size-based criteria, including adenoma size ≥10 mm, villous or tubulovillous histology, or high-grade dysplasia [4-7]. Some studies also consider large serrated lesions or advanced serrated lesions within broader advanced precursor categories, although definitions vary across screening programs and research cohorts. Advanced colorectal neoplasia is a broader endpoint and usually includes advanced adenoma and CRC. These distinctions are important because adenoma, advanced adenoma, advanced colorectal neoplasia, early CRC, and established CRC are not interchangeable endpoints for screening translation.

Endpoint heterogeneity can directly influence reported microbiome performance metrics [14, 16, 18, 27, 28, 34]. When advanced adenoma is combined with CRC, sensitivity and area under the receiver operating characteristic curve (AUC) may be inflated by the stronger

microbial signal of established cancer. When non-advanced adenomas are combined with advanced adenomas, the weaker signal from small low-risk lesions may dilute the apparent performance for prevention-relevant lesions. Similarly, differences in size thresholds, villous histology, high-grade dysplasia, serrated lesion inclusion, and lesion location can affect microbial shedding, bleeding, inflammation, and stool-detectable signals. For this reason, microbiota-assisted screening studies should report non-advanced adenoma, advanced adenoma, advanced serrated lesions when available, early CRC, and advanced CRC separately. Advanced adenoma-specific performance should be prioritized when evaluating whether a microbial marker has credible prevention-oriented clinical utility.

Microbial features of advanced adenoma

Advanced adenoma is the key prevention endpoint because removal can interrupt progression before invasive cancer develops. For microbiota-assisted screening, this stage is also the hardest test. Established CRC often bleeds, disrupts tissue architecture, alters oxygen gradients, and creates inflammatory or necrotic niches; advanced adenomas may generate a weaker and more localized fecal signal. A CRC-versus-control AUC therefore cannot be extrapolated to screening for advanced adenoma. Representative studies evaluating microbiota-assisted detection of advanced adenoma and advanced colorectal neoplasia are summarized in **Table 2**.

The reporting problem is as important as the biological signal. Studies that merge non-advanced adenoma, advanced adenoma, early CRC, and advanced CRC obscure the endpoint on which prevention depends. Additional study-level details are provided in Supplementary Table 1 [29-33]. Zhang et al. directly targeted advanced adenoma in a population-based screening cohort and showed that a 13-marker bacterial panel added discrimination to FIT, although the bootstrap-adjusted area under the curve remained modest [34]. Wong et al. demonstrated a FIT-adjunct signal for advanced colorectal neoplasia, but that composite included CRC and cannot be treated as isolated advanced-adenoma evidence [18].

Evidence maturity and translational relevance

From a translational perspective, the studies summarized in **Table 2** should not be interpreted as equivalent levels of evidence [14, 16, 18, 27, 28, 34]. Discovery-oriented adenoma-CRC sequence studies mainly support biological plausibility and biomarker generation, whereas internally validated models provide early evidence of discriminatory potential. Studies with external validation, FIT-adjunctive assessment, advanced adenoma-specific reporting, or screening-based populations are more directly relevant to clinical implementation. Therefore, microbiota-assisted screening should be judged primarily by incremental value in real screening pathways rather than by CRC-enriched case-control performance alone.

The evidence base can be viewed as a maturity continuum. At the earliest level are mechanistic and microbiome-discovery studies that identify taxa, virulence features, functional pathways, or metabolites associated with colorectal neoplasia. A second level includes internally validated classifiers, which are useful for hypothesis generation but remain vulnerable to overfitting and cohort-specific effects. A third level includes externally validated models and cross-cohort analyses, which better address transportability. The most implementation-relevant level consists of prospective or screening-based investigations that report advanced adenoma separately, evaluate performance beyond FIT or stool DNA testing, use prespecified thresholds, and estimate downstream colonoscopy burden. This hierarchy reinforces the central premise of this review: not all microbiome studies contribute equally to screening implementation, and diagnostic-model translation requires prespecified models, external validation, calibration, and clinically relevant thresholds [48-50].

Microbial shifts from adenoma to colorectal cancer

Across the transition from adenoma to CRC, studies more consistently identify expansion of oral-associated anaerobes, opportunistic pathobionts, and organisms with inflammatory or genotoxic potential. Meta-analyses of fecal metagenomes have found reproducible CRC-associated microbial signatures across populations, including enrichment of Fusobacterium,

Table 2. Representative microbiome studies relevant to advanced adenoma and microbiota-assisted stool-based screening

Study	Study design/population	Sample type and platform	Endpoint reporting	Advanced adenoma/advanced neoplasia evidence	FIT or stool-test comparison	Validation status	Reviewer-level interpretation
Zackular et al. 2014 [14]	Colonoscopy-associated cohort spanning healthy controls, adenoma, and CRC.	Stool; 16S rRNA sequencing.	Adenoma and CRC evaluated across the adenoma-carcinoma sequence.	Advanced adenoma not reported separately.	No FIT comparator.	Discovery-oriented cohort.	Useful for showing that stool microbiota shifts can appear before cancer, but limited direct evidence for advanced adenoma screening.
Feng et al. 2015 [16]	Metagenomic case-control study of CRC, adenoma, and controls.	Stool; shotgun metagenomics.	Adenoma and CRC reported, with emphasis on microbial gene markers.	Advanced adenoma not reported separately.	No FIT comparator.	Cross-cohort validation elements reported.	Important for metagenomic marker discovery; prevention-stage utility cannot be inferred from CRC-enriched performance.
Wong et al. 2017 [18]	Participants evaluated for colorectal neoplasia.	Stool; qPCR quantification of Fusobacterium nucleatum with FIT.	Advanced colorectal neoplasia, including CRC and advanced adenoma; advanced adenoma analyzed separately.	For advanced adenoma, fecal F. nucleatum alone had AUC 0.59 (95% CI 0.51-0.67), sensitivity 32.7%, and specificity 87.0%. FIT alone sensitivity was 15.5% with AUC 0.55 (95% CI 0.48-0.63); FIT + F. nucleatum sensitivity was 38.6%, specificity 89.0%, and AUC 0.63 (95% CI 0.56-0.70).	Direct FIT comparison; validation cohort advanced adenoma AUC improved from 0.55 to 0.63 (P=0.031). For CRC, FIT + F. nucleatum improved sensitivity from 73.1% to 92.3% and AUC from 0.86 to 0.95.	Independent validation cohort reported.	Important proof of concept for FIT-adjunct use in advanced colorectal neoplasia, but ACN includes CRC and should not be treated as isolated advanced adenoma evidence.
Baxter et al. 2016 [27]	Colonoscopy cohort with colorectal lesions and controls.	Stool; 16S rRNA sequencing plus FIT hemoglobin concentration in random-forest model.	Colonic lesions, CRC, and adenoma; endpoint broader than isolated advanced adenoma.	Microbiota/FIT model detected 91.7% of cancers and 45.5% of adenomas versus FIT alone 75.0% and 15.7%; among FIT-missed lesions, the model detected 70.0% of cancers and 37.7% of adenomas.	Direct FIT comparison; combined microbiota/FIT model improved sensitivity for broad colonic lesions.	Leave-one-out and 10-fold cross-validation reported.	Strong evidence that microbiota can complement FIT, but it is not advanced adenoma-specific validation.
Yachida et al. 2019 [28]	Multi-stage colorectal neoplasia cohort.	Stool; shotgun metagenomics and metabolomics.	Multiple stages across colorectal tumorigenesis, including multiple polypoid adenomas, intramucosal carcinoma, and advanced CRC.	Stage-specific shifts were reported in multiple polypoid adenomas and intramucosal carcinomas. F. nucleatum increased from intramucosal carcinoma to more advanced stages; Atopobium parvulum and Actinomyces odontolyticus were increased in multiple polypoid adenomas and/or intramucosal carcinomas.	No FIT or stool DNA comparator and no clinical referral threshold.	Large multi-omics cohort; stage biology emphasized rather than screening implementation.	Supports early lesion biology and metabolomic assay rationale; it does not establish FIT-adjunct screening utility.
Zhang et al. 2021 [34]	Population-based screening setting; 268 advanced adenoma cases and 788 controls; FIT subset 248 advanced adenoma and 732 controls.	Stool; 16S rRNA sequencing with 13 microbial markers plus FIT/APCS models.	Advanced adenoma was the target endpoint and was analyzed separately from non-advanced adenoma and controls.	Thirteen-marker model apparent AUC 0.641 (95% CI 0.601-0.681), adjusted AUC 0.607 (95% CI 0.548-0.660). At 90% specificity, the combined FIT + marker model reached 31.5% apparent and 28.4% adjusted sensitivity.	FIT alone apparent AUC 0.544 and adjusted AUC 0.527; FIT sensitivity 10.5%, specificity 98.4%. FIT + 13 markers improved apparent AUC to 0.671 and adjusted AUC to 0.641 (P<0.001); FIT + markers + APCS apparent AUC 0.727 and adjusted AUC 0.706.	Bootstrap-adjusted estimates reported; no external implementation cohort.	The most directly relevant advanced adenoma study in the table: it shows measurable gain over FIT, but performance remains modest and needs locked external validation before referral use.
Piccinno et al. 2025 [54]	Pooled analysis of 3,741 stool metagenomes from 18 cohorts: 1,471 CRC, 702 adenoma, and 1,568 control samples.	Stool; shotgun metagenomics with species-, pathway-, and strain-level analyses.	CRC stage and location were analyzed; adenomas were pooled as a single category.	No advanced-adenoma-specific subgroup or diagnostic accuracy was reported. Adenoma microbiomes did not differ significantly from controls at the global community level.	No FIT or stool DNA comparator. CRC-versus-control LODO AUC averaged 0.85 (range, 0.71-0.97); this was not an adenoma result.	Multicohort and LODO validation; external transportability was addressed for CRC.	Strong evidence for reproducible CRC and cross-stage microbial features, but no advanced-adenoma detection performance.

Note: ACN, advanced colorectal neoplasia; APCS, Asia-Pacific Colorectal Screening score; AUC, area under the receiver operating characteristic curve; CI, confidence interval; CRC, colorectal cancer; DNA, deoxyribonucleic acid; FIT, fecal immunochemical test; LODO, leave-one-dataset-out; qPCR, quantitative polymerase chain reaction; rRNA, ribosomal ribonucleic acid.

Parvimonas, Peptostreptococcus, Porphyromonas, and other oral cavity-associated taxa [23, 24]. The driver-passenger model proposes that some organisms may participate early in tumor initiation, whereas others may preferentially colonize the tumor microenvironment created by neoplastic transformation [35]. This framework is useful but should not be treated as proven for every organism.

Bacteroides requires nuanced interpretation. The genus contains common commensals and specific pathobiont strains. ETBF, defined by BFT production, has mechanistic relevance that cannot be inferred from genus-level Bacteroides abundance alone. Similarly, *E. coli* is not uniformly genotoxic; concern is focused on strains carrying the pks island that produce colibactin. This strain-level distinction is one reason shotgun metagenomics or targeted virulence-gene assays may be more informative than 16S rRNA profiling for translation.

Depleted taxa are also important. Reduced abundance of *Faecalibacterium*, *Roseburia*, and other butyrate-producing organisms is frequently interpreted as loss of anti-inflammatory or barrier-supportive functions. *Bifidobacterium* and *Lactobacillus* are sometimes discussed as beneficial genera, but this framing can be misleading because effects are species- and strain-dependent and may vary by host diet, baseline ecology, and disease context. For biomarker development, depletion markers may improve models by capturing loss of health-associated functions, but they are usually less specific than enrichment of pathobionts or virulence genes. Their best role may be as part of multi-dimensional signatures rather than stand-alone diagnostic claims.

Differences between early and advanced colorectal cancer

Early CRC is the screening-relevant stage, but many studies combine stages or include clinically detected advanced cases. Yachida and colleagues reported stage-specific fecal metagenomic and metabolomic phenotypes, supporting the possibility that early lesions have detectable but distinct microbial features [28]. Advanced CRC may show stronger fecal signals because tumors are larger, may bleed, create necrotic or inflammatory niches, and alter luminal metabolism. F. nucleatum and oral-associated anaerobes are often more prominent in established CRC than in controls, while depletion of short-chain fatty acid (SCFA)-producing commensals is frequently reported. However, comparisons between early and advanced CRC may be confounded by tumor location, obstruction, bleeding, bowel habit changes, antibiotic exposure, diet modification, and treatment status.

Fecal microbiota versus mucosa-associated microbiota

Fecal and mucosal microbiota should not be used interchangeably. Fecal samples are practical for screening and can be integrated with FIT or stool DNA collection, but they may dilute localized tumor-associated communities. Mucosa-associated, biofilm-associated, or intratumoral communities may better reflect epithelial interaction. Biofilm organization has been described in proximal colorectal tumors and adjacent mucosa [36]. However, mucosal sampling requires colonoscopy, can be affected by bowel preparation, and is less suitable as a first-line screening test.

Taken together, microbial alterations across the adenoma-carcinoma sequence should be read as candidate measurable signals rather than a substitute screening modality. Key microbial taxa across the colorectal adenoma-carcinoma sequence are summarized in **Table 3**. **Figure 1** maps taxa and strain-level features to functional pathways, host effects, assay formats, and screening implications; **Figure 2** places those assays within an adjunctive stool-based screening pathway.

Lesion location, morphology, and the serrated pathway

Advanced adenoma is a useful translational endpoint, but it is not biologically uniform. Conventional adenomas may be tubular, tubulovillous, or villous, may occur in the proximal or distal colon, and may be single or multiple. Advanced serrated lesions arise through a partly distinct pathway and are frequently considered within the broader category of advanced precancerous lesions even when a study reserves the term advanced adenoma for conventional histology. Combining all of these lesions can increase the number of endpoint events, but it can also conceal clinically impor-

Table 3. Key microbial taxa across the colorectal adenoma-carcinoma sequence

Taxon	Direction of change	Disease stage	Sample type	Potential mechanism	Biomarker relevance	Evidence strength	Key references
Fusobacterium nucleatum	Often enriched	More consistent in CRC than early adenoma; reported in some adenoma studies	Fecal and mucosa-associated	FadA-mediated E-cadherin/β-catenin signaling, inflammation, immune modulation.	Candidate qPCR marker; more useful in panels than alone.	Moderate for CRC; limited for advanced adenoma.	[17, 18, 37]
Enterotoxigenic *Bacteroides fragilis*	Strain-specific enrichment reported	Adenoma and CRC, depending on cohort and assay	Fecal and mucosal	BFT-mediated E-cadherin cleavage, Wnt/β-catenin signaling, inflammatory signaling.	Virulence-marker assays may improve specificity.	Mechanistically strong; clinical biomarker evidence variable.	[19, 20]
pks+ *Escherichia coli*	Strain-specific enrichment reported	Colorectal neoplasia and CRC; stage-specific value remains uncertain	Mainly mucosal/tissue; fecal detection possible	Colibactin-associated DNA damage and mutational signatures.	Functional marker candidate rather than genus-level marker.	Mechanistically strong; screening evidence limited.	[21, 22]
Peptostreptococcus, *Parvimonas*, *Porphyromonas*	Often enriched	More consistent in CRC; may increase along progression	Fecal and mucosa-associated	Oral-associated anaerobic expansion, inflammation, proteolytic metabolism.	Useful components of multi-taxa signatures.	Moderate for CRC; limited for advanced adenoma.	[23, 24]
Bacteroides	Variable; strain-dependent	Adenoma and CRC	Fecal and mucosal	Species- and strain-specific effects; ETBF differs from commensal strains.	Requires species/strain or toxin-level resolution.	Variable.	[19, 20]
Faecalibacterium and *Roseburia*	Often depleted	Adenoma and CRC in some cohorts	Fecal	Reduced butyrate production and anti-inflammatory metabolic capacity.	Potential negative markers in panels.	Moderate but variable.	[23, 28]
Bifidobacterium and *Lactobacillus*	Variable	Adenoma and CRC; affected by diet and probiotic exposure	Fecal	Barrier support, immune modulation, fermentation; effects are context dependent.	Not reliable as single diagnostic markers.	Limited and inconsistent.	[9, 38]

Note: BFT, *Bacteroides fragilis* toxin; CRC, colorectal cancer; ETBF, enterotoxigenic *Bacteroides fragilis*; FIT, fecal immunochemical test; qPCR, quantitative polymerase chain reaction; rRNA, ribosomal ribonucleic acid; SCFA, short-chain fatty acid.

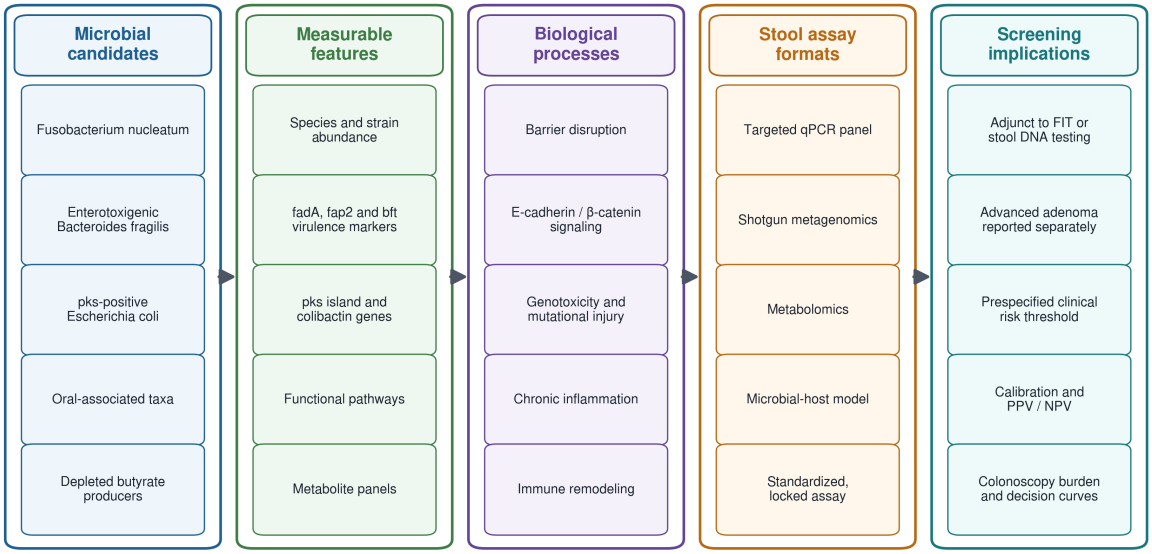


Figure 1. Mechanistic basis and assay implications of microbiota-assisted colorectal neoplasia screening. Candidate microbial taxa and strain-level features can be translated into functional or host-linked signals and then into a fixed stool assay. Clinical translation requires advanced adenoma to be reported separately and the assay to be evaluated as an adjunct to existing screening, with a prespecified threshold, calibration, PPV, NPV, colonoscopy burden, and decision-curve analysis. BFT, Bacteroides fragilis toxin; FadA, Fusobacterium adhesin A; Fap2, Fusobacterium autotransporter protein 2; pks, polyketide synthase; DNA, deoxyribonucleic acid; FIT, fecal immunochemical test; NPV, negative predictive value; PPV, positive predictive value; qPCR, quantitative polymerase chain reaction.

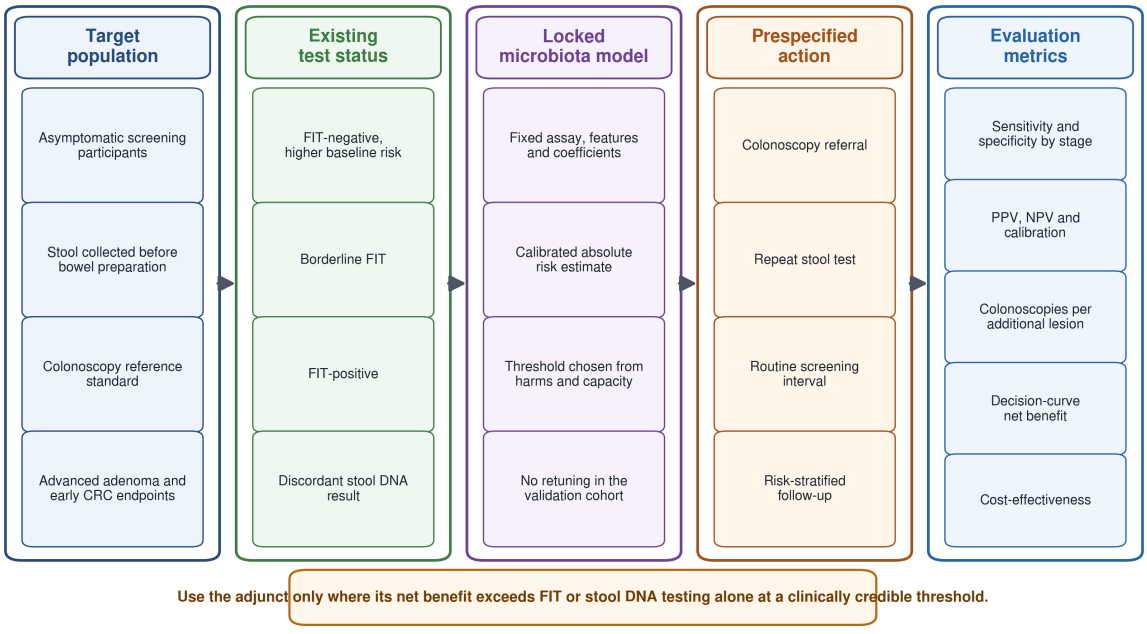


Figure 2. Proposed clinical pathway for microbiota-assisted stool-based screening and colonoscopy referral. The score is applied within an existing screening pathway to a defined population whose stool is collected before bowel preparation. The assay, features, coefficients, and action threshold are fixed before external evaluation. Referral is justified only when the adjunct produces greater net benefit than FIT or stool DNA testing alone at a clinically credible threshold. CRC, colorectal cancer; DNA, deoxyribonucleic acid; FIT, fecal immunochemical test; NPV, negative predictive value; PPV, positive predictive value.

tant differences in microbial exposure, shedding, bleeding, and detectability. Screening studies should therefore state whether the primary endpoint is conventional advanced adenoma, advanced serrated lesion, or a composite, and should provide the component results even when the composite is used for statistical power.

Anatomical location deserves similar attention. Proximal and distal lesions differ in molecular pathways, luminal conditions, and contact with mucosa-associated communities. Organized microbial biofilms have been described particularly in proximal colorectal cancers [36], supporting the possibility that spatial organization rather than stool abundance alone carries biological information. A stool sample, however, mixes material from the whole intestinal tract and cannot reliably localize a signal. A classifier that performs well for distal lesions may therefore fail to identify proximal or flat lesions, and a model trained on tissue-associated organisms may not transfer directly to fecal measurements. Location-stratified sensitivity should be reported whenever event counts permit, with uncertainty intervals rather than unsupported claims of equivalence.

Lesion morphology also affects the proposed complementarity with FIT. FIT detects human hemoglobin and is strongest when a lesion bleeds into the lumen; a microbial assay is attractive precisely because microbial adhesion, virulence functions, or metabolic disturbance could be present without continuous bleeding. This hypothesis should be tested rather than assumed. The relevant analysis is a paired one that identifies advanced lesions missed by FIT but detected by the locked microbial rule and compares that gain with the false-positive referrals added among participants without advanced neoplasia. Analyses limited to the AUC of each test cannot reveal whether the assays make different errors in the participants who matter for an adjunctive strategy.

Future cohorts should consequently record lesion number, size, histology, dysplasia, morphology, and segment of the colon using prespecified definitions. Conventional advanced adenoma and advanced serrated lesion should be tabulated separately, followed by a clearly labeled composite if required. Paired fecal and mucosal sampling in a subset can help distinguish a stable luminal screening signal from a lesion-localized mechanism, but the clinically intended specimen must determine the primary performance estimate. This phenotyping discipline would make apparently discordant microbiome studies more interpretable and would clarify which precursor lesions a proposed assay can realistically prevent from progressing.

Mechanistic links between gut dysbiosis and colorectal carcinogenesis

As summarized in **Figure 1**, the mechanistic evidence is most useful when it is translated into measurable assay components: taxa, virulence genes, functional pathways, metabolites, and host-response signals. The figure does not imply that dysbiosis alone can diagnose advanced adenoma; it identifies which biological features could be tested as stool-based adjunctive biomarkers. Major microbiota-based screening strategies, plausible clinical use cases, and validation considerations are summarized in **Table 4**.

Barrier disruption

The intestinal barrier includes mucus, epithelial tight and adherens junctions, antimicrobial peptides, immunoglobulin A (IgA), and immune surveillance. Dysbiosis can alter mucus degradation, epithelial permeability, and microbial contact with the mucosa. Barrier disruption may increase exposure to microbial products and metabolites, thereby amplifying inflammatory and proliferative signaling.

Mucus biology is particularly relevant to the distinction between fecal and mucosal signals. Fecal profiling may detect organisms shed from the lumen, but mucosa-associated bacteria are positioned to interact with epithelial glycans, antimicrobial peptides, and immune cells. Biofilm-like organization may increase bacterial density near the epithelial surface and create microenvironments that concentrate metabolites or virulence factors. These spatial features are difficult to infer from bulk stool sequencing. Future studies that combine stool profiling with paired mucosal sampling, spatial imaging, metatranscriptomics, and host epithelial readouts may clarify which fecal markers reflect true epithelial contact.

Table 4. Microbiota-based screening strategies: plausible use cases, key evidence, and minimum validation needs

Strategy	Plausible clinical use case	Current evidence base	Key failure point	Minimum evidence required before implementation
Single microbial marker	Simple qPCR add-on when a strong organism-specific signal is expected.	Fusobacterium nucleatum and selected taxa show repeated CRC associations.	Species abundance is not specific enough for advanced adenoma screening and may miss strain-level biology.	Locked assay, predefined threshold, stage-specific performance, and comparison with FIT or stool DNA testing.
Multi-taxa microbial panel	Risk refinement after non-invasive stool testing or in colonoscopy referral triage.	Cross-cohort CRC classifiers and adenoma studies support feasibility.	Performance often falls when tested outside CRC-enriched discovery cohorts.	External validation in asymptomatic screening participants with advanced adenoma reported separately.
Microbiota plus FIT	Most realistic near-term adjunct, especially for FIT-negative higher-risk or borderline FIT scenarios.	Some studies show microbial markers can add information to FIT for advanced adenoma.	Small gains may increase false positives and colonoscopy demand.	FIT-stratified performance, decision-curve analysis, and colonoscopy burden per advanced neoplasia detected.
Microbiota plus stool DNA testing	Additional biological layer when stool DNA is already used.	Conceptually attractive because host DNA, hemoglobin, and microbial ecology capture different signals.	Increment beyond modern stool DNA testing may be small and costly.	Head-to-head testing against established and next-generation stool DNA assays with calibration and cost analysis.
Microbiome-metabolome panel	Functional readout for early lesion biology and assay prioritization.	Metagenomic-metabolomic studies show stage-related functional shifts.	Metabolites can reflect diet, host metabolism, medications, and sample handling.	Standardized preanalytics, reproducible metabolite panel, and clinical thresholds tied to advanced adenoma or early CRC.
Integrated risk model	Decision support combining microbiota, FIT, stool DNA, age, family history, and other risk factors.	Machine-learning models can combine complementary features.	Overfitting and poor transportability are common in high-dimensional microbiome data.	Locked model, external validation, calibration, transparent reporting, and implementation-impact analysis.

Note: AUC, area under the receiver operating characteristic curve; CRC, colorectal cancer; DNA, deoxyribonucleic acid; FIT, fecal immunochemical test; NPV, negative predictive value; PPV, positive predictive value; qPCR, quantitative polymerase chain reaction.

Chronic inflammation

Chronic inflammation links microbial dysbiosis with epithelial injury, compensatory proliferation, and immune-cell recruitment [9]. CRC-associated microbial communities can activate NF-κB, STAT3, IL-6, IL-17, and chemokine pathways, although the strength of evidence differs by organism and model [19]. This mechanism supports assay designs that pair microbial markers with host inflammatory or hemoglobin-based signals, especially when triaging FIT-negative or borderline stool-test results.

Virulence factors and genotoxicity

Several CRC-associated microbes matter because they carry strain-specific mechanisms rather than only broad taxonomic associations. F. nucleatum expresses FadA, an adhesin that can bind E-cadherin and activate β-catenin signaling in experimental models [17]. ETBF secretes BFT, which can disrupt E-cadherin and influence β-catenin, NF-κB, MAPK, and cytokine signaling [19]. pks+ E. coli can produce colibactin, a genotoxin linked to DNA damage and mutational processes [21, 39]. These observations favor targeted virulence-gene assays or shotgun metagenomic features over genus-level calls when the goal is a clinically reproducible biomarker.

Among these mechanisms, colibactin-related DNA damage provides one of the strongest links between a bacterial product and cancer-relevant genomic injury, although the timing of exposure and its screening implications remain unresolved.

Metabolic dysregulation

The gut microbiota contributes to short-chain fatty acids, bile acids, amino acid metabolites, polyamines, sulfur metabolites, and other small molecules. Butyrate-producing taxa such as Faecalibacterium and Roseburia are often depleted in colorectal neoplasia, while secondary bile acids, tryptophan metabolites, and sulfur metabolism may connect diet, microbial function, epithelial stress, and immunity. For assay design, metabolite readouts can help distinguish passive taxonomic association from active pathway disturbance, making microbiome-metabolome panels particularly relevant to early lesion biology.

Metabolite-based interpretation is attractive because it can connect taxonomic observations to functions. A fecal sample enriched for an organism does not necessarily show that a pathway is active; metagenomics, metatranscriptomics, metabolomics, and targeted biochemical assays can provide more direct functional evidence. For example, detection of a bile acid-transforming gene set, altered secondary bile acid concentrations, and a compatible dietary pattern would provide stronger biological coherence than taxonomy alone. In screening models, metabolites may also be more stable or more directly linked to phenotype than rare taxa, but they are vulnerable to preanalytical variation and short-term diet. For this reason, metabolomic biomarkers require the same level of validation as microbial markers.

Immune microenvironment remodeling

Microbiota can shape local and systemic immunity, including T-cell polarization, macrophage recruitment, myeloid-derived suppressor cells, natural killer (NK) cell activity, antigen presentation, and immune-checkpoint expression. *F. nucleatum* and ETBF have been associated with immune and inflammatory remodeling in experimental and observational studies [19, 37, 40]. Because immune remodeling is bidirectional, biomarker panels should treat immune-linked features as context-dependent signals that may need host or lesion-context markers to preserve specificity.

Microbiota-based biomarkers for early screening

Single-taxon biomarkers

Single-taxon markers such as fecal *F. nucleatum* are attractive because they can be measured by quantitative polymerase chain reaction (qPCR) and interpreted readily. *F. nucleatum* is associated with CRC in many cohorts and has plausible mechanisms involving adhesion, beta-catenin signaling, inflammation, and immune modulation [17, 37]. ETBF and pks+ *E. coli* are also attractive because their virulence genes can be targeted directly. Single markers, however, rarely capture the heterogeneity of colorectal neoplasia, and their specificity can be reduced by oral disease, inflammatory con-

ditions, antibiotic exposure, geography, diet, and analytical variation.

Multi-taxa microbial signatures

Multi-taxa signatures use combinations of enriched and depleted taxa, often generated by random forest, logistic regression, penalized regression, or other machine-learning approaches. Cross-cohort metagenomic studies have identified reproducible CRC-associated signatures and have shown that classifiers can generalize better when trained across diverse populations [23, 24]. Multi-taxa models can capture ecological patterns more effectively than single organisms, including enrichment of oral-associated anaerobes and depletion of commensal SCFA producers. Yet many reported models are based on case-control designs, small sample sizes, incomplete covariate adjustment, or internal validation only. Overfitting is a major risk, especially when high-dimensional microbial features greatly outnumber outcome events.

Model interpretation must follow the intended decision. A case-control area under the curve averages performance across every possible cut-off and is largely insensitive to the disease prevalence created by the investigators. Screening programs instead need sensitivity at one or more clinically acceptable specificities, calibration of predicted absolute risk, positive predictive value (PPV) and negative predictive value (NPV) in the target population, and the number of referrals produced per lesion detected. Results should be shown separately in FIT-negative and FIT-positive strata because a microbiota score used to rescue FIT-negative lesions has a different pretest risk and a different tolerance for false-positive colonoscopy than a score used to prioritize FIT-positive referrals.

Microbiota combined with FIT

Combining microbiota with FIT remains the most plausible near-term strategy because FIT is already embedded in screening and microbial markers may capture biology unrelated to bleeding. In **Table 2**, Zhang et al. showed an advanced adenoma-specific gain when bacteria were added to FIT, while Wong et al. showed improved detection of advanced colorectal neoplasia with fecal *F. nucleatum* plus FIT. **Figure 2** illustrates how such a score would be most useful in FIT-negative higher-risk, borderline FIT, FIT-positive prioritization, or stool DNA-discordant scenarios rather than as a universal replacement test.

Together, these FIT-adjunctive findings show why microbiota-assisted screening should be evaluated by incremental sensitivity, specificity, and downstream colonoscopy burden rather than by case-control AUC alone.

Microbiota combined with stool DNA testing

Stool DNA testing detects host-derived molecular alterations and hemoglobin. Adding microbial markers could theoretically capture tumor-microbe interactions or premalignant ecological shifts. However, incremental benefit over established stool DNA panels must be demonstrated, particularly for advanced adenoma, specificity, cost, and colonoscopy burden. Any proposed combined assay should report calibration, decision-curve analysis, false-positive colonoscopy implications, and performance across age, sex, race or ethnicity, geography, diet, and tumor location.

Microbiota combined with metabolomics

Metabolomics may provide functional readouts of microbial activity and host-microbe interaction. Multi-omics studies suggest that microbial and metabolomic alterations can appear in early tumorigenesis and may improve discrimination between controls, adenoma, and CRC [28, 41]. Metabolites such as SCFAs, bile acids, amino acid derivatives, polyamines, and sulfur compounds may reflect pathway-level changes more directly than taxonomic abundance. Nevertheless, metabolomic profiles are highly sensitive to diet, fasting status, sample handling, storage, platform, and normalization. Without standardized collection and external validation, metabolomic signals may simply reproduce diet, handling, or platform effects rather than lesion biology.

Artificial intelligence and machine-learning risk models

Artificial intelligence (AI) and machine-learning models can combine microbial taxa, functional genes, metabolomics, FIT concentration, stool DNA results, age, family history, medica-

tion exposure, and other clinical risk factors. The breadth of input is appealing, but it also creates opportunities for leakage: preprocessing, feature selection, imputation, batch correction, and cut-off selection can all transfer outcome information into validation folds. A credible report keeps the complete analytical pipeline inside resampling, separates model development from evaluation, publishes the fixed feature set and coefficients, evaluates calibration as well as discrimination, and compares the resulting decision with the existing screening pathway rather than with healthy controls alone.

Representative microbiome-derived biomarker studies are summarized in Supplementary Table 2. These studies illustrate the diversity of sample types, analytical platforms, target stages, and validation approaches, while reinforcing the need to report advanced adenoma separately from broader colorectal neoplasia endpoints.

Translational challenges and future directions

Evidence gaps in advanced adenoma detection

Advanced adenoma is the prevention endpoint on which microbiota-assisted screening will largely stand or fall. Many datasets either merge advanced adenoma with smaller adenomas or analyze it together with CRC, making sensitivity, specificity, calibration, and added clinical value difficult to interpret. A more useful study would define advanced adenoma consistently, enroll asymptomatic screening participants before colonoscopy, and report this endpoint separately from non-advanced adenoma and established cancer. **Table 2** and Supplementary Table 1 summarize representative evidence and highlight where advanced adenoma-specific results are still missing.

Prevalence makes PPV and NPV indispensable for screening interpretation. Sensitivity and specificity are conditional on disease status, whereas PPV and NPV answer the questions faced after a result is known: how often a positive result represents advanced neoplasia and how much reassurance a negative result provides. A case-control dataset that deliberately contains many cancers or adenomas cannot yield a PPV that is transportable to an average-risk program unless results are reweighted to the intended prevalence. Reporting only sensitivity, specificity, and area under the curve therefore leaves the colonoscopy consequence unresolved.

A simple illustration shows the size of the effect. If sensitivity is 70% and specificity is 90%, PPV is 26.9% when prevalence is 5% but only 12.5% when prevalence is 2%; the corresponding NPVs are 98.3% and 99.3%. At the lower prevalence, seven of every eight positive results would be false positives even though the operating characteristics are unchanged. Conversely, a high NPV in a low-prevalence cohort can coexist with clinically important missed lesions and should not be mistaken for high sensitivity. The prevalence used in this calculation must match the deployment stratum: an all-comers screening population, FIT-negative higher-risk participants, and surveillance patients will not have the same post-test probabilities.

Future reports should therefore provide the observed advanced-adenoma prevalence, PPV and NPV with 95% confidence intervals, and expected outcomes per 1,000 people screened. The latter should include true-positive and false-positive referrals, missed advanced adenomas, and total colonoscopies. For a single strategy, approximately 1/PPV colonoscopies are generated per true-positive finding; for an adjunct, the more informative measure is the paired increase in colonoscopies divided by the paired increase in advanced neoplasia detected relative to FIT or stool DNA testing alone. This prevalence-aware presentation directly links diagnostic performance to the capacity and harms of a screening program.

Preanalytical heterogeneity

Preanalytical variation is not a technical footnote in this field; it can change the signal being measured. Stool collection kits, preservatives, time at ambient temperature, freezing, transport, and aliquoting can shift microbial profiles [42-46]. DNA extraction and processing pipelines add further variability [42-46], and STORMS provides a reporting framework for capturing these details [46]. Mucosal sampling introduces additional variables, including biopsy site, lesion proximity, bowel preparation, and endoscopic technique. These details should be

recorded before signatures are compared across cohorts.

Analytical platform differences

Platform choice should follow the intended use. 16S rRNA sequencing is accessible for broad community profiling but has limited species- and strain-level resolution and cannot reliably capture virulence determinants such as FadA, BFT, or pks-associated genes. Shotgun metagenomics offers stronger taxonomic and functional resolution, at higher cost and analytical complexity. For near-term clinical testing, a locked qPCR panel may be more practical, provided that assay calibration, quality control, and interlaboratory reproducibility are demonstrated [42-44, 47, 48].

Population heterogeneity and confounding

Population heterogeneity is a major reason that microbial classifiers fail to travel. Diet, geography, ethnicity, antibiotics, probiotics, proton pump inhibitors, metformin, bowel habits, stool consistency, comorbidity, and recent colonoscopy preparation can all influence gut microbiota. A signature trained in one region or sequencing workflow may therefore perform poorly elsewhere unless metadata, preprocessing, and validation are handled transparently.

External validation should mirror the intended screening pathway. Testing a model in another case-control dataset is less informative than evaluating it among asymptomatic people who are already undergoing FIT, stool DNA testing, or colonoscopy-based screening.

Case-control studies remain useful for discovery, but they are a weak basis for implementation. Known CRC cases may have larger tumors, symptoms, altered diet, bleeding, inflammation, or treatment-related exposures. These differences can inflate apparent diagnostic performance. Screening utility should be tested before colonoscopy or within established non-invasive pathways, where the clinical decision is still unresolved.

Inter-cohort reproducibility and model transportability

Reproducibility is a central barrier because a stool signature can encode geography, diet, medication exposure, sequencing chemistry, DNA extraction, or case definition as readily as lesion biology. A classifier that performs well in one discovery cohort may lose both discrimination and calibration when the specimen workflow or population changes. This matters especially in screening, where advanced-neoplasia prevalence is lower than in most case-control datasets and even a small decrease in specificity can generate a large number of additional colonoscopies.

Piccinno et al. provide the largest recent test of cross-cohort reproducibility [54]. Their pooled dataset comprised 3,741 stool metagenomes from 18 cohorts, including 1,471 CRC samples, 702 adenoma samples, and 1,568 controls. The CRC-versus-control classifier achieved a mean leave-one-dataset-out (LODO) area under the curve of 0.85 (range, 0.71-0.97), and the analysis identified reproducible species-, pathway-, and strain-level features. The result is important for CRC biology and transportability. It is not, however, an advanced-adenoma accuracy estimate: adenomas were analyzed as a pooled category, advanced and non-advanced adenomas were not separated, and no advanced-adenoma-specific sensitivity, specificity, PPV, NPV, or area under the curve was reported. At the global community level, the adenoma group did not differ significantly from controls. **Table 2** now makes this limitation explicit.

Internal k-fold or leave-one-out validation within a single dataset remains useful for model development but does not demonstrate clinical transportability. Stronger evidence requires a fixed assay, feature set, coefficients, preprocessing pipeline, and clinical cut-off before an independent cohort is opened. Leave-one-cohort-out analysis can identify fragile sites during development, but the decisive evaluation is prospective application in participants who resemble the intended screening population. Transportability should be judged by preserved calibration and net benefit as well as by a similar area under the curve.

Machine-learning overfitting and validation

High-dimensional microbiome data can produce apparently strong internal performance when the number of candidate features is large relative to the number of advanced-adenoma

events. The most common errors are selecting taxa before cross-validation, estimating normalization parameters from the full dataset, optimizing hyperparameters and the clinical cut-off on the test data, and allowing samples from the same participant or batch to appear in both training and validation sets. Each error makes the validation set less independent than its label implies.

The unit of model development should be the participant, and every outcome-informed step should be nested inside the training data. Sample-size planning should be based on the number of advanced adenomas and early cancers rather than the total cohort size. Performance should be accompanied by uncertainty intervals, calibration plots, calibration intercept and slope, and a description of missing-data and invalid-assay handling. If a model is updated after external testing, that cohort becomes part of development and a new untouched evaluation cohort is required.

A locked model does not mean that laboratory quality control is frozen indefinitely. Reagent lots, extraction platforms, reference databases, and taxonomic pipelines can change. Any permitted analytical update should be governed by a prespecified bridging study, acceptance criteria, and version control. TRIPOD and STARD provide a reporting foundation [49, 50], but microbiome studies also need the specimen and analytical detail required by STORMS [46].

Setting clinical thresholds with decision-curve analysis

A locked model and a prespecified threshold solve different problems. Locking fixes the assay, preprocessing, features, coefficients, and output calculation. The threshold converts that output into an action such as colonoscopy, repeat stool testing, or routine follow-up. It should not be selected simply because it maximizes the Youden index or gives the most favorable sensitivity in the evaluation cohort. Those choices implicitly give false-positive and false-negative classifications equal weight and ignore colonoscopy capacity, lesion prevalence, and the intended position of the test in the pathway.

Threshold setting should begin with the clinical use case. For a microbiota score applied after a negative FIT in a participant with elevated baseline risk, missing an advanced adenoma may justify a relatively low action threshold, but the acceptable threshold cannot be so low that false-positive referrals overwhelm endoscopy services. In FIT-positive participants, the microbiota score should not be used to deny a guideline-indicated colonoscopy without direct safety evidence; a threshold may instead prioritize referral timing. A model used after a discordant stool DNA result requires yet another balance because the pretest probability and existing false-positive burden differ.

The threshold probability, denoted p_t, is the predicted probability at which the expected benefit of colonoscopy is judged to equal its expected harm or burden. Decision-curve analysis translates that preference into net benefit: net benefit = (true positives/N) - (false positives/N) × p_t/(1-p_t) [55]. Plotting net benefit across a clinically credible range of p_t compares the microbiota-assisted strategy with FIT or stool DNA testing alone and with refer-all and refer-none policies. The useful range should be specified before analysis through input from clinicians, patients, screening-program managers, and local capacity planning; presenting a wide post hoc range can make almost any model appear useful somewhere.

Advanced adenoma and CRC should not be combined uncritically in this analysis. A program may assign different consequences to missing an advanced adenoma, an early cancer, and a late cancer, and those weights should be explicit if a composite endpoint is used. An alternative is to present separate decision curves and consequences per 1,000 screened for advanced adenoma, early CRC, and all advanced neoplasia. The report should also show the cut-off, sensitivity, specificity, PPV, NPV, referral proportion, and additional colonoscopies per additional advanced neoplasia detected at each proposed operating point.

The chosen cut-off must be fixed in the development protocol and applied without retuning in external validation. If prevalence differs across sites, a prespecified intercept recalibration may be reasonable, but changing the feature set, coefficients, or cut-off after outcomes are known is model redevelopment, not validation. A two-threshold policy can be considered

when the clinical pathway genuinely has three actions - for example, routine interval, repeat testing, and immediate colonoscopy - but both cut-offs and the intermediate action must be defined in advance. This approach makes the term "prespecified threshold" operational and ties it to the clinical risk of advanced adenoma rather than to a favorable retrospective statistic.

Repeated-round performance and temporal stability

Screening is repeated, whereas most microbiome classifiers are evaluated from one stool sample per participant. Within-person microbial profiles can change with antibiotics, acute gastrointestinal illness, bowel preparation, major dietary change, travel, medication, and stool consistency. A marker that separates cases from controls at one time point may therefore have limited repeatability near the clinical cut-off. Analytical reproducibility from the same aliquot and biological stability across independently collected samples should be reported separately.

Repeated-round evaluation should quantify the proportion of participants who change risk category without an intervening lesion diagnosis, the within-person variability of the score, invalid-test rates, and the effect of a defined repeat-sampling rule. Protocols should state whether testing is deferred after antibiotics or acute illness, how long the deferral lasts, and whether an indeterminate result triggers recollection or colonoscopy. Without such rules, apparent sensitivity gains may be accompanied by repeated false-positive referrals in the same individuals.

Program-level performance also differs from single-round accuracy. Investigators should estimate cumulative detection, false-positive colonoscopies, adherence, and interval cancers or advanced neoplasia across more than one screening round. A model that adds a small number of false positives each year may create a substantial cumulative burden even if its single-round specificity appears acceptable. Conversely, a stable adjunct that detects lesions missed by intermittent bleeding may contribute value over repeated FIT rounds that is not visible in a one-time comparison.

Temporal monitoring is also required after deployment. Positive and negative controls, lot-bridging samples, score-distribution surveillance, and predefined drift alarms can distinguish a population shift from an analytical failure. Any software, database, or wet-laboratory change that alters the score distribution should trigger a documented bridging assessment before the original clinical threshold is reused.

Cost-effectiveness and implementation barriers

Implementation costs extend beyond the price of sequencing or a quantitative polymerase chain reaction (qPCR) reaction. They include the collection device, ambient-temperature stability, accessioning, DNA extraction, controls, failed-sample repeats, bioinformatics, result reporting, staff time, and the colonoscopies generated by false-positive results. Turnaround time must also fit the screening workflow; an accurate result returned after the referral decision has already been made has no incremental utility.

Platform choice should follow the intended product. Shotgun metagenomics is valuable for discovery and strain-level analysis but may be difficult to standardize and reimburse at population scale. A targeted qPCR or other fixed molecular panel is more plausible for routine use if it preserves the relevant signal and passes interlaboratory reproducibility, limit-of-detection, interference, and stability studies. A broad sequencing assay should not be retained merely because it performed best during development if a smaller locked panel offers similar net benefit with lower cost and faster reporting.

Economic evaluation should compare complete strategies rather than test prices in isolation. The comparator may be annual FIT, stool DNA testing at its recommended interval, colonoscopy, or a risk-adapted pathway. The model should include adherence, repeat rounds, surveillance after polypectomy, stage shift, complications, treatment costs, quality-adjusted survival, and colonoscopy capacity. The principal uncertainty is likely to be advanced-adenoma performance because small changes in sensitivity and specificity can alter both future cancer incidence and immediate referral volume.

Before adoption, a screening program should be able to state who is tested, when the sample

Table 5. Methodological and translational barriers in microbiota-based CRC screening

Barrier	Examples	Impact on biomarker development	Suggested solutions	Relevant references
Preanalytical variation	Collection kit, preservative, transport time, freezing, bowel preparation.	Can shift microbial profiles and reduce reproducibility.	Standardized collection, storage metadata, harmonized DNA extraction.	[42-46]
Analytical platform differences	16S rRNA, qPCR, shotgun metagenomics, metabolomics.	Different taxonomic/functional resolution and batch effects.	Use locked assays, reference materials, and cross-platform validation.	[42-44, 47, 48]
Population heterogeneity	Diet, geography, ethnicity, medications, antibiotics, probiotics.	Limits transferability of signatures across cohorts.	Multicenter cohorts with detailed metadata and subgroup reporting.	[42-44, 46, 51]
Endpoint heterogeneity	Adenoma mixed with advanced adenoma or CRC.	Obscures prevention-relevant performance.	Report non-advanced adenoma, advanced adenoma, early CRC, and advanced CRC separately.	[14, 16, 18, 27]
Machine-learning overfitting	High-dimensional features, small cohorts, data leakage.	Inflates apparent diagnostic performance.	Prespecified models, external validation, calibration, transparent reporting.	[48-50]
Implementation and cost	Unclear added clinical value beyond FIT/stool DNA, cost of sequencing.	Limits clinical adoption despite promising biology.	Incremental utility studies and cost-effectiveness analyses.	[6, 52, 53]
Prevalence and threshold mismatch	Case-control enrichment, post hoc cut-offs, and referral thresholds disconnected from colonoscopy capacity.	Inflates apparent PPV and may create excessive false-positive referrals.	Report target prevalence, PPV, NPV, and consequences per 1,000; prespecify the action threshold and evaluate decision-curve net benefit.	[49, 50, 55]
Temporal and batch instability	Within-person variation, antibiotics, acute illness, reagent lots, and pipeline drift.	A locked score may change between rounds for reasons unrelated to a lesion.	Use repeated-sample studies, bridging controls, invalid-test rules, and prospective drift monitoring.	[42-48]

Note: CRC, colorectal cancer; DNA, deoxyribonucleic acid; FIT, fecal immunochemical test; NPV, negative predictive value; PPV, positive predictive value; qPCR, quantitative polymerase chain reaction; rRNA, ribosomal ribonucleic acid; STORMS, Strengthening The Organization and Reporting of Microbiome Studies.

is collected, which result changes management, how an invalid result is handled, and what additional endoscopy demand is expected. Major methodological and implementation barriers are summarized in **Table 5**, and **Table 6** compares the resulting strategies with current stool-based options.

A concise comparison of FIT, stool DNA testing, and microbiota-assisted strategies is provided in **Table 6**, with representative quantitative data from FIT-adjunctive microbiota studies where advanced adenoma-specific results were available.

Regulatory readiness, equity, and governance

Regulatory and laboratory readiness begin with an explicit intended-use statement. A test developed for average-risk adults completing routine stool screening cannot be assumed to apply to symptomatic patients, hereditary cancer syndromes, inflammatory bowel disease, post-polypectomy surveillance, or triage after an abnormal FIT. The specimen type, collection device, transport window, extraction method, analytical platform, software version, reportable range, invalid-result rule, and clinical action must form one fixed system. Changing any of these elements after validation may alter the score distribution and should trigger a proportionate bridging or revalidation study rather than an undocumented software update.

Analytical validation should be separated from clinical validation. Precision within and between runs, operator and laboratory effects, lot variation, carryover, contamination, interfering substances, limit of detection, reportable range, sample stability, and failure rates should be established with appropriate controls [42-47]. A sequencing workflow also requires rules for read depth, host-read removal, taxonomic or functional database versions, batch correction, and features below the quantification limit. A targeted panel requires comparable documentation for primer or probe specificity and amplification efficiency. These studies do not demonstrate clinical benefit, but without them an apparent change in clinical risk may reflect an unstable measurement process.

Equity is a performance requirement rather than a post-implementation aspiration. Diet, geography, medication use, health conditions, access to sample collection, and the background distribution of microbial features can differ across populations. A single threshold can therefore produce different referral rates or calibration errors even when the rank-order AUC is similar. Recruitment should represent the intended screening population, and reporting should compare calibration, sensitivity, specificity, PPV, invalid-test frequency, and follow-up completion across prespecified groups where sample size and ethical governance allow. If subgroup estimates are imprecise, that limitation should be stated and addressed through planned post-market evidence rather than interpreted as proof of equal performance.

The software component also needs lifecycle governance. The locked model should have a traceable version, documented feature definitions, deterministic preprocessing, quality checks, and an audit trail linking each reported score to the assay and database versions used. Score-distribution drift, assay failure, referral volume, and advanced-neoplasia yield should be monitored after deployment. Recalibration may be appropriate when baseline risk changes, but changing coefficients, features, or thresholds creates a modified test and requires new evidence. Access control, privacy protection, retention rules, and procedures for correcting an erroneous report are part of the clinical product even though they are rarely visible in a retrospective accuracy paper.

Clinical scenarios for implementation

FIT-negative but clinically higher-risk individuals

A microbiota adjunct could be most useful among people with a negative FIT but higher baseline risk because of age, family history, prior neoplasia, metabolic risk, or other clinical factors. As shown in **Figure 2**, the score would need a locked threshold and calibrated risk output before it could justify colonoscopy despite a negative hemoglobin-based test. The main danger is false-positive expansion: if specificity is poor, the assay would consume colonoscopy capacity without preventing many additional cancers.

Borderline or discordant stool-test results

Borderline FIT concentrations, repeated weak positives, or discordance between FIT, stool

Table 6. Comparison of current stool-based screening strategies and microbiota-assisted approaches

Strategy	Advanced adenoma/advanced precursor detection	Specificity/false-positive burden	Colonoscopy referral implication	Implementation considerations
FIT	Generally limited single-round sensitivity for advanced adenoma; better suited to detecting bleeding cancers than non-bleeding precursor lesions.	Generally high specificity compared with multitarget stool DNA testing.	Lower referral burden, but repeated testing is needed to maintain program-level effectiveness.	Low cost, scalable, widely implemented; performance depends on threshold, adherence, and repeated rounds.
Multitarget stool DNA testing	Improves detection of CRC and some advanced precancerous lesions compared with FIT, but advanced adenoma detection remains imperfect.	Lower specificity than FIT in major screening studies.	More positive results and potentially more colonoscopy referrals than FIT.	Higher cost; useful where stool DNA testing is available, but incremental benefit must be balanced against specificity and resource use.
Next-generation stool DNA testing	Recent studies suggest improved detection of advanced neoplasia compared with earlier stool DNA approaches, but implementation impact depends on specificity and referral burden.	Specificity remains central because even modest reductions can increase colonoscopy demand at population scale.	Requires evaluation of colonoscopies generated per additional advanced neoplasia detected.	Requires head-to-head comparison with FIT, established stool DNA testing, and cost-effectiveness analysis.
Microbiota + FIT	Potential incremental sensitivity for advanced adenoma or ACN when microbial markers capture non-bleeding biology. Wong et al. 2017: FIT sensitivity 15.5%, AUC 0.55 (95% CI 0.48-0.63); FIT + fecal *F. nucleatum* sensitivity 38.6%, specificity 89.0%, AUC 0.63 (95% CI 0.56-0.70). Zhang et al. 2021: FIT adjusted AUC 0.527; FIT + 13 microbial markers adjusted AUC 0.641; FIT + markers + APCS adjusted AUC 0.706.	Must preserve acceptable specificity to avoid excessive false-positive colonoscopy referrals. Zhang et al. 2021 reported FIT specificity 98.4%; at 90% specificity, the combined FIT + marker model reached 28.4% adjusted sensitivity.	Most useful if additional advanced neoplasia detected offsets added colonoscopy burden; quantitative reporting should include colonoscopies generated per additional advanced neoplasia detected.	Requires locked microbial assays, FIT-stratified reporting, external validation, calibrated thresholds, and prospective screening evaluation.
Microbiota + stool DNA or multi-omics	Conceptually attractive because host DNA, hemoglobin, microbial ecology, virulence genes, and metabolites may capture complementary biology.	Incremental specificity and net benefit over stool DNA alone remain uncertain.	Should be adopted only if decision-curve analysis and cost-effectiveness support clinical utility.	Not ready for routine screening; requires standardized preanalytics, external validation, and implementation-impact analysis.

Note: ACN, advanced colorectal neoplasia; APCS, Asia-Pacific Colorectal Screening score; AUC, area under the receiver operating characteristic curve; CRC, colorectal cancer; DNA, deoxyribonucleic acid; FIT, fecal immunochemical test; NPV, negative predictive value; PPV, positive predictive value.

DNA testing, and clinical risk create uncertainty that is common in real screening programs. A microbial score could act as a tie-breaker only if thresholds are prespecified and the effect on referral decisions is measured. In this scenario, calibration matters more than discovery-model AUC because clinicians need to know how the result changes post-test probability and downstream colonoscopy use.

Post-polypectomy surveillance or risk refinement

Microbial profiles could eventually help refine surveillance intervals after polypectomy or identify people whose risk remains elevated despite removal of visible lesions. This use case is biologically plausible because mucosal ecology, residual field effects, and recurrent adenoma risk may be linked. It is also risky: surveillance populations differ from average-risk screening cohorts, bowel preparation can perturb microbiota, and overuse could intensify colonoscopy demand without proven outcome benefit.

Design of a prospective implementation study

The most informative next study is a prospective, paired evaluation nested within an established screening program. Consecutive asymptomatic participants should provide stool before bowel preparation, undergo the same FIT and microbiota workflow, and receive colonoscopy as the reference standard irrespective of the investigational result. This design avoids verification bias and permits a within-person comparison of FIT alone with FIT plus the microbial score. Symptomatic referrals and surveillance patients may be studied, but they should be enrolled and reported as separate cohorts rather than mixed with average-risk participants.

The primary endpoint should be advanced adenoma, defined before recruitment by size, villous features, and high-grade dysplasia. Advanced serrated lesions should be prespecified and reported separately because their biology, location, bleeding pattern, and fecal signal may differ. Early CRC, all CRC, non-advanced adenoma, lesion location, multiplicity, and interval outcomes are important secondary endpoints. Central or quality-controlled pathology review would reduce endpoint misclassification.

Development and evaluation data must be separated. If feature discovery is still required, it should occur in an earlier cohort, and the final assay, coefficients, normalization, missing-value rules, and clinical threshold should be frozen before the prospective samples are analyzed. Laboratory personnel should be blinded to colonoscopy and pathology, and endoscopists should be blinded to the investigational score unless the study is explicitly designed to measure score-guided management. A public protocol and analysis plan should identify the primary comparison and permissible recalibration.

Sample size should be driven by the expected number of advanced-adenoma events and the desired precision of sensitivity, PPV, and the incremental detection difference, not simply by total enrollment. Because the proposed strategy is paired, estimates of additional true-positive and false-positive referrals should retain that pairing. Confidence intervals are needed for stage-specific sensitivity, specificity, PPV, NPV, calibration, and the difference in referral proportions. Prespecified subgroup analyses should address age, sex, race or ethnicity where legally and ethically appropriate, geography, medication exposure, tumor location, and FIT stratum, while recognizing that most subgroup analyses will be underpowered unless planned deliberately.

The implementation analysis should present consequences per 1,000 participants at the locked cut-off and across the prespecified threshold-probability range. At minimum, these consequences are additional advanced adenomas and CRCs detected, missed lesions, false-positive colonoscopies, total colonoscopies, invalid assays, and additional colonoscopies per additional advanced neoplasia detected. Decision curves should compare FIT alone, the adjunctive strategy, refer-all, and refer-none policies [55]. A budget-impact analysis can then translate referral proportions into local endoscopy demand.

Finally, the study should capture the operational outcomes that determine whether a technically valid assay can function in practice: kit return, time to laboratory receipt, sample stability, assay failure, turnaround time, uptake of recommended colonoscopy, adverse events, and participant acceptability. Reporting a favor-

Table 7. Minimum evidence standards for microbiota-assisted colorectal neoplasia screening

1. Enroll asymptomatic screening participants before colonoscopy.
2. Collect stool before bowel preparation whenever possible.
3. Report non-advanced adenoma, advanced adenoma, early CRC, and advanced CRC separately.
4. Compare microbial models against FIT and/or stool DNA testing.
5. Report FIT-stratified performance.
6. Use locked models and external validation.
7. Provide calibration, thresholds, and decision-curve analysis.
8. Estimate additional colonoscopies required per additional advanced neoplasia detected compared with FIT or stool DNA testing alone.
9. Include cost-effectiveness assessment.
10. Report preanalytical and analytical metadata according to STORMS.
11. Report target prevalence, PPV, NPV, and outcomes per 1,000 participants with confidence intervals.
12. Prespecify the clinical action threshold and evaluate net benefit across a clinically credible threshold range.
13. Evaluate repeat-round stability, invalid-test rates, and analytical drift.

Note: CRC, colorectal cancer; DNA, deoxyribonucleic acid; FIT, fecal immunochemical test; NPV, negative predictive value; PPV, positive predictive value; STORMS, Strengthening The Organization and Reporting of Microbiome Studies.

able area under the curve without these data would repeat the current evidence gap. Reporting the complete pathway would show whether the microbial signal improves prevention rather than only classification.

Clinical reporting and follow-up safeguards

A clinical report should communicate the decision, not the development statistic. The recipient needs to know whether the result is valid, which screening pathway and time point it applies to, the estimated or categorized risk under the validated model, and the recommended next action. AUC, feature importance, or a list of detected organisms is not an adequate patient-level interpretation. If an absolute risk is reported, the reference population and time horizon should be clear, and the report should avoid implying that the detected organisms caused a lesion in that individual.

Invalid and discordant results require prespecified handling. The protocol should distinguish an inadequate specimen from analytical failure, define whether and when recollection is allowed, and prevent repeated testing from becoming an unmeasured source of selection. When FIT and the microbial score disagree, the action should follow the validated decision rule rather than an ad hoc judgment based on whichever result appears more concerning. Programs should measure whether participants actually complete the recommended colonoscopy because an accurate triage rule cannot improve outcomes if referral is delayed or declined.

A positive microbial result followed by a high-quality negative colonoscopy is particularly sensitive. Current evidence does not justify repeated colonoscopy, broad antimicrobial treatment, or microbiome-directed therapy solely because the score remains high. Unless a prospective protocol has validated another action, subsequent care should return to the applicable screening or surveillance guideline. This safeguard limits cascades from false-positive results and keeps the investigational biomarker within the decision for which its benefit and harms were evaluated.

Minimum evidence standards for future studies

The next generation of studies should be designed around decisions made in screening programs. The field does not need another internally validated CRC-versus-control classifier; it needs a locked adjunctive model applied before colonoscopy in a clearly defined screening population. **Table 7** summarizes the minimum evidence needed for clinicians, payers, and guideline panels to interpret the result.

These standards are deliberately stricter than ordinary biomarker discovery. Advanced adenoma must be reported separately, the comparator must reflect current practice, and the output must be an absolute risk or decision

rule whose calibration can be checked. PPV and NPV should be calculated at the observed prevalence and, where relevant, standardized to the prevalence of the intended deployment population. Threshold selection and decision-curve analysis should be completed without access to the external validation outcomes.

A study can satisfy these reporting requirements and still yield a negative result. Such results are valuable: they define the limits of transportability, prevent an unhelpful assay from consuming colonoscopy capacity, and show whether the failure lies in the biological signal, the laboratory platform, or the clinical decision rule. Full reporting of negative and indeterminate findings is therefore part of the evidence standard, not an optional supplement.

Minimum analytical evidence should accompany the clinical estimates. Investigators should disclose the collection kit, stabilization conditions, transport time, extraction protocol, controls, sequencing or amplification method, bioinformatic workflow, database versions, normalization, missing-data rules, batch structure, quality thresholds, and failed-sample denominator. Replicate precision and interlaboratory reproducibility are especially important when the intended test is transferred from a discovery laboratory to a clinical laboratory. The final validation report should use the same specimen-handling and computational pipeline that would be used in practice.

Minimum transparency also includes a dated protocol, a prespecified statistical analysis plan, a complete model specification or an auditable executable implementation, and reporting consistent with STORMS, TRIPOD, and STARD [46, 49, 50]. Participant flow should identify exclusions, missing reference standards, invalid investigational tests, and incomplete colonoscopy. Performance should not be calculated only after excluding failed assays, because the failure rate is part of real-world effectiveness. When proprietary restrictions prevent public release of coefficients, an independent evaluator should still receive a genuinely locked implementation before outcomes are unblinded.

Evidence for adoption should progress in stages: analytical validity, prospective clinical validity, comparative clinical utility, implementation feasibility, and economic value. Success at one stage should not substitute for the next. A reproducible microbial association is not yet an assay; an accurate assay is not yet a useful decision rule; and a useful decision rule is not yet a sustainable screening program. This staged standard allows promising signals to advance while preventing premature claims based on retrospective case-control discrimination.

Conclusion

Fecal microbial differences in established CRC are reproducible across populations, but that achievement does not yet solve the prevention problem. Advanced adenoma produces a weaker and less consistently reported signal, and the largest pooled 2025 analysis did not provide advanced-adenoma-specific accuracy. The most directly relevant evidence comes from screening-based studies that combine microbial markers with FIT, where modest gains have been observed but external, locked validation remains limited.

Translation now depends on a change in emphasis: from retrospective discrimination to a prevalence-aware decision. A clinically credible study must report advanced adenoma separately, preserve PPV and specificity at the intended prevalence, calibrate absolute risk, fix its threshold before evaluation, and demonstrate net benefit relative to FIT or stool DNA testing. It must also quantify the added colonoscopy burden, repeat-round stability, operational failure, and cost. Until those conditions are met, microbiota-derived markers should be regarded as candidate adjuncts within established stool-based pathways, not as replacement screening test.

Disclosure of conflict of interest

None.

Abbreviations

ACN, advanced colorectal neoplasia; AI, artificial intelligence; APCS, Asia-Pacific Colorectal Screening score; AUC, area under the receiver operating characteristic curve; BFT, *Bacteroides fragilis* toxin; CRC, colorectal cancer; DCA, decision-curve analysis; DNA, deoxyribonucleic

acid; ETBF, enterotoxigenic *Bacteroides fragilis*; FIT, fecal immunochemical test; IgA, immunoglobulin A; LODO, leave-one-dataset-out; NK, natural killer; NPV, negative predictive value; PPV, positive predictive value; qPCR, quantitative polymerase chain reaction; rRNA, ribosomal ribonucleic acid; SCFA, short-chain fatty acid; STARD, Standards for Reporting Diagnostic Accuracy Studies; STORMS, Strengthening The Organization and Reporting of Microbiome Studies; TRIPOD, Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis.

Address correspondence to: Junying Liu, Department of Gastroenterology, Zhoukou Central Hospital Affiliated to Henan Medical University, No. 26, East Section of Renmin Road, Zhoukou 466000, Henan, China. Tel: +86-17839477428; E-mail: Liujunying-781025@163.com

References

[1] Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I and Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024; 74: 229-263.

[2] Fearon ER. Molecular genetics of colorectal cancer. Annu Rev Pathol 2011; 6: 479-507.

[3] Biller LH and Schrag D. Diagnosis and treatment of metastatic colorectal cancer: a review. JAMA 2021; 325: 669-685.

[4] US Preventive Services Task Force, Davidson KW, Barry MJ, Mangione CM, Cabana M, Caughey AB, Davis EM, Donahue KE, Doubeni CA, Krist AH, Kubik M, Li L, Ogedegbe G, Owens DK, Pbert L, Silverstein M, Stevermer J, Tseng CW and Wong JB. Screening for colorectal cancer: us preventive services task force recommendation statement. JAMA 2021; 325: 1965-1977.

[5] Wolf AMD, Fontham ETH, Church TR, Flowers CR, Guerra CE, LaMonte SJ, Etzioni R, McKenna MT, Oeffinger KC, Shih YT, Walter LC, Andrews KS, Brawley OW, Brooks D, Fedewa SA, Manassaram-Baptiste D, Siegel RL, Wender RC and Smith RA. Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society. CA Cancer J Clin 2018; 68: 250-281.

[6] Imperiale TF, Ransohoff DF, Itzkowitz SH, Levin TR, Lavin P, Lidgard GP, Ahlquist DA and Berger BM. Multitarget stool DNA testing for colorectal-cancer screening. N Engl J Med 2014; 370: 1287-1297.

[7] Helsingen LM, Vandvik PO, Jodal HC, Agoritsas T, Lytvyn L, Anderson JC, Auer R, Murphy SB, Almadi MA, Corley DA, Quinlan C, Fuchs JM, McKinnon A, Qaseem A, Heen AF, Siemieniuk RAC, Kalager M, Usher-Smith JA, Lansdorp-Vogelaar I, Bretthauer M and Guyatt G. Colorectal cancer screening with faecal immunochemical testing, sigmoidoscopy or colonoscopy: a clinical practice guideline. BMJ 2019; 367: l5515.

[8] Vacante M, Ciuni R, Basile F and Biondi A. Gut microbiota and colorectal cancer development: a closer look to the adenoma-carcinoma sequence. Biomedicines 2020; 8: 489.

[9] Dougherty MW and Jobin C. Intestinal bacteria and colorectal cancer: etiology and treatment. Gut Microbes 2023; 15: 2185028.

[10] Dekker E, Tanis PJ, Vleugels JLA, Kasi PM and Wallace MB. Colorectal cancer. Lancet 2019; 394: 1467-1480.

[11] Yu L, Zhao G, Wang L, Zhou X, Sun J, Li X, Zhu Y, He Y, Kofonikolas K, Bogaert D, Dunlop M, Zhu Y, Theodoratou E and Li X. A systematic review of microbial markers for risk prediction of colorectal neoplasia. Br J Cancer 2022; 126: 1318-1328.

[12] Zwezerijnen-Jiwa FH, Sivov H, Paizs P, Zafeiropoulou K and Kinross J. A systematic review of microbiome-derived biomarkers for early colorectal cancer detection. Neoplasia 2023; 36: 100868.

[13] Herlo LF, Salcudean A, Sirli R, Iurciuc S, Herlo A, Nelson-Twakor A, Alexandrescu L and Dumache R. Gut microbiota signatures in colorectal cancer as a potential diagnostic biomarker in the future: a systematic review. Int J Mol Sci 2024; 25: 7937.

[14] Zackular JP, Rogers MA, Ruffin MT 4th and Schloss PD. The human gut microbiome as a screening tool for colorectal cancer. Cancer Prev Res (Phila) 2014; 7: 1112-1121.

[15] Zeller G, Tap J, Voigt AY, Sunagawa S, Kultima JR, Costea PI, Amiot A, Böhm J, Brunetti F, Habermann N, Hercog R, Koch M, Luciani A, Mende DR, Schneider MA, Schrotz-King P, Tournigand C, Tran Van Nhieu J, Yamada T, Zimmermann J, Benes V, Kloor M, Ulrich CM, von Knebel Doeberitz M, Sobhani I and Bork P. Potential of fecal microbiota for early-stage detection of colorectal cancer. Mol Syst Biol 2014; 10: 766.

[16] Feng Q, Liang S, Jia H, Stadlmayr A, Tang L, Lan Z, Zhang D, Xia H, Xu X, Jie Z, Su L, Li X, Li X, Li J, Xiao L, Huber-Schönauer U, Niederseer D, Xu X, Al-Aama JY, Yang H, Wang J, Kristiansen K, Arumugam M, Tilg H, Datz C and Wang J. Gut microbiome development along the colorectal adenoma-carcinoma sequence. Nat Commun 2015; 6: 6528.

[17] Rubinstein MR, Wang X, Liu W, Hao Y, Cai G and Han YW. *Fusobacterium nucleatum* promotes colorectal carcinogenesis by modulating E-cadherin/β-catenin signaling via its FadA adhesin. Cell Host Microbe 2013; 14: 195-206.

[18] Wong SH, Kwong TNY, Chow TC, Luk AKC, Dai RZW, Nakatsu G, Lam TYT, Zhang L, Wu JCY, Chan FKL, Ng SSM, Wong MCS, Ng SC, Wu WKK, Yu J and Sung JJY. Quantitation of faecal Fusobacterium improves faecal immunochemical test in detecting advanced colorectal neoplasia. Gut 2017; 66: 1441-1448.

[19] Wu S, Lim KC, Huang J, Saidi RF and Sears CL. *Bacteroides fragilis* enterotoxin cleaves the zonula adherens protein, E-cadherin. Proc Natl Acad Sci U S A 1998; 95: 14979-14984.

[20] Wu S, Rhee KJ, Albesiano E, Rabizadeh S, Wu X, Yen HR, Huso DL, Brancati FL, Wick E, McAllister F, Housseau F, Pardoll DM and Sears CL. A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. Nat Med 2009; 15: 1016-1022.

[21] Pleguezuelos-Manzano C, Puschhof J, Rosendahl Huber A, van Hoeck A, Wood HM, Nomburg J, Gurjao C, Manders F, Dalmasso G, Stege PB, Paganelli FL, Geurts MH, Beumer J, Mizutani T, Miao Y, van der Linden R, van der Elst S; Genomics England Research Consortium, Garcia KC, Top J, Willems RJL, Giannakis M, Bonnet R, Quirke P, Meyerson M, Cuppen E, van Boxtel R and Clevers H. Mutational signature in colorectal cancer caused by genotoxic pks^{+} *E. coli*. Nature 2020; 580: 269-273.

[22] Nougayrède JP, Homburg S, Taieb F, Boury M, Brzuszkiewicz E, Gottschalk G, Buchrieser C, Hacker J, Dobrindt U and Oswald E. *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. Science 2006; 313: 848-851.

[23] Thomas AM, Manghi P, Asnicar F, Pasolli E, Armanini F, Zolfo M, Beghini F, Manara S, Karcher N, Pozzi C, Gandini S, Serrano D, Tarallo S, Francavilla A, Gallo G, Trompetto M, Ferrero G, Mizutani S, Shiroma H, Shiba S, Shibata T, Yachida S, Yamada T, Wirbel J, Schrotz-King P, Ulrich CM, Brenner H, Arumugam M, Bork P, Zeller G, Cordero F, Dias-Neto E, Setubal JC, Tett A, Pardini B, Rescigno M, Waldron L, Naccarati A and Segata N. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. Nat Med 2019; 25: 667-678.

[24] Wirbel J, Pyl PT, Kartal E, Zych K, Kashani A, Milanese A, Fleck JS, Voigt AY, Palleja A, Ponnudurai R, Sunagawa S, Coelho LP, Schrotz-King P, Vogtmann E, Habermann N, Niméus E, Thomas AM, Manghi P, Gandini S, Serrano D, Mizutani S, Shiroma H, Shiba S, Shibata T, Yachida S, Yamada T, Waldron L, Naccarati A, Segata N, Sinha R, Ulrich CM, Brenner H, Arumugam M, Bork P and Zeller G. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. Nat Med 2019; 25: 679-689.

[25] Yu J, Feng Q, Wong SH, Zhang D, Liang QY, Qin Y, Tang L, Zhao H, Stenvang J, Li Y, Wang X, Xu X, Chen N, Wu WK, Al-Aama J, Nielsen HJ, Kiilerich P, Jensen BA, Yau TO, Lan Z, Jia H, Li J, Xiao L, Lam TY, Ng SC, Cheng AS, Wong VW, Chan FK, Xu X, Yang H, Madsen L, Datz C, Tilg H, Wang J, Brünner N, Kristiansen K, Arumugam M, Sung JJ and Wang J. Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. Gut 2017; 66: 70-78.

[26] Dai Z, Coker OO, Nakatsu G, Wu WKK, Zhao L, Chen Z, Chan FKL, Kristiansen K, Sung JJY, Wong SH and Yu J. Multi-cohort analysis of colorectal cancer metagenome identified altered bacteria across populations and universal bacterial markers. Microbiome 2018; 6: 70.

[27] Baxter NT, Ruffin MT 4th, Rogers MA and Schloss PD. Microbiota-based model improves the sensitivity of fecal immunochemical test for detecting colonic lesions. Genome Med 2016; 8: 37.

[28] Yachida S, Mizutani S, Shiroma H, Shiba S, Nakajima T, Sakamoto T, Watanabe H, Masuda K, Nishimoto Y, Kubo M, Hosoda F, Rokutan H, Matsumoto M, Takamaru H, Yamada M, Matsuda T, Iwasaki M, Yamaji T, Yachida T, Soga T, Kurokawa K, Toyoda A, Ogura Y, Hayashi T, Hatakeyama M, Nakagama H, Saito Y, Fukuda S, Shibata T and Yamada T. Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. Nat Med 2019; 25: 968-976.

[29] Liang Q, Chiu J, Chen Y, Huang Y, Higashimori A, Fang J, Brim H, Ashktorab H, Ng SC, Ng SSM, Zheng S, Chan FKL, Sung JJY and Yu J. Fecal bacteria act as novel biomarkers for noninvasive diagnosis of colorectal cancer. Clin Cancer Res 2017; 23: 2061-2070.

[30] Liang JQ, Li T, Nakatsu G, Chen YX, Yau TO, Chu E, Wong S, Szeto CH, Ng SC, Chan FKL, Fang JY, Sung JJY and Yu J. A novel faecal *Lachnoclostridium* marker for the non-invasive diagnosis of colorectal adenoma and cancer. Gut 2020; 69: 1248-1257.

[31] Konishi Y, Okumura S, Matsumoto T, Itatani Y, Nishiyama T, Okazaki Y, Shibutani M, Ohtani N, Nagahara H, Obama K, Ohira M, Sakai Y, Nagayama S and Hara E. Development and evaluation of a colorectal cancer screening

method using machine learning-based gut microbiota analysis. Cancer Med 2022; 11: 3194-3206.
[32] Chen F, Dai X, Zhou CC, Li KX, Zhang YJ, Lou XY, Zhu YM, Sun YL, Peng BX and Cui W. Integrated analysis of the faecal metagenome and serum metabolome reveals the role of gut microbiome-associated metabolites in the detection of colorectal cancer and adenoma. Gut 2022; 71: 1315-1325.
[33] Coker OO, Liu C, Wu WKK, Wong SH, Jia W, Sung JJY and Yu J. Altered gut metabolites and microbiota interactions are implicated in colorectal carcinogenesis and can be non-invasive diagnostic biomarkers. Microbiome 2022; 10: 35.
[34] Zhang Y, Lu M, Lu B, Liu C, Ma Y, Liu L, Miao X, Qin J, Chen H and Dai M. Leveraging fecal microbial markers to improve the diagnostic accuracy of the fecal immunochemical test for advanced colorectal adenoma. Clin Transl Gastroenterol 2021; 12: e00389.
[35] Tjalsma H, Boleij A, Marchesi JR and Dutilh BE. A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects. Nat Rev Microbiol 2012; 10: 575-582.
[36] Dejea CM, Wick EC, Hechenbleikner EM, White JR, Mark Welch JL, Rossetti BJ, Peterson SN, Snesrud EC, Borisy GG, Lazarev M, Stein E, Vadivelu J, Roslani AC, Malik AA, Wanyiri JW, Goh KL, Thevambiga I, Fu K, Wan F, Llosa N, Housseau F, Romans K, Wu X, McAllister FM, Wu S, Vogelstein B, Kinzler KW, Pardoll DM and Sears CL. Microbiota organization is a distinct feature of proximal colorectal cancers. Proc Natl Acad Sci U S A 2014; 111: 18321-18326.
[37] Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, Michaud M, Clancy TE, Chung DC, Lochhead P, Hold GL, El-Omar EM, Brenner D, Fuchs CS, Meyerson M and Garrett WS. *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. Cell Host Microbe 2013; 14: 207-215.
[38] Mori G and Pasca MR. Gut microbial signatures in sporadic and hereditary colorectal cancer. Int J Mol Sci 2021; 22: 1312.
[39] Arthur JC, Perez-Chanona E, Mühlbauer M, Tomkovich S, Uronis JM, Fan TJ, Campbell BJ, Abujamel T, Dogan B, Rogers AB, Rhodes JM, Stintzi A, Simpson KW, Hansen JJ, Keku TO, Fodor AA and Jobin C. Intestinal inflammation targets cancer-inducing activity of the microbiota. Science 2012; 338: 120-123.
[40] Gur C, Ibrahim Y, Isaacson B, Yamin R, Abed J, Gamliel M, Enk J, Bar-On Y, Stanietsky-Kaynan N, Coppenhagen-Glazer S, Shussman N, Almogy G, Cuapio A, Hofer E, Mevorach D, Tabib A, Ortenberg R, Markel G, Miklić K, Jonjic S, Brennan CA, Garrett WS, Bachrach G and Mandelboim O. Binding of the Fap2 protein of *Fusobacterium nucleatum* to human inhibitory receptor TIGIT protects tumors from immune cell attack. Immunity 2015; 42: 344-355.
[41] Zhou L, Jiang Z, Zhang Z, Xing J, Wang D and Tang D. Progress of gut microbiome and its metabolomics in early screening of colorectal cancer. Clin Transl Oncol 2023; 25: 1949-1962.
[42] Costea PI, Zeller G, Sunagawa S, Pelletier E, Alberti A, Levenez F, Tramontano M, Driessen M, Hercog R, Jung FE, Kultima JR, Hayward MR, Coelho LP, Allen-Vercoe E, Bertrand L, Blaut M, Brown JRM, Carton T, Cools-Portier S, Daigneault M, Derrien M, Druesne A, de Vos WM, Finlay BB, Flint HJ, Guarner F, Hattori M, Heilig H, Luna RA, van Hylckama Vlieg J, Junick J, Klymiuk I, Langella P, Le Chatelier E, Mai V, Manichanh C, Martin JC, Mery C, Morita H, O'Toole PW, Orvain C, Patil KR, Penders J, Persson S, Pons N, Popova M, Salonen A, Saulnier D, Scott KP, Singh B, Slezak K, Veiga P, Versalovic J, Zhao L, Zoetendal EG, Ehrlich SD, Dore J and Bork P. Towards standards for human fecal sample processing in metagenomic studies. Nat Biotechnol 2017; 35: 1069-1076.
[43] Claesson MJ, Clooney AG and O'Toole PW. A clinician's guide to microbiome analysis. Nat Rev Gastroenterol Hepatol 2017; 14: 585-595.
[44] Knight R, Vrbanac A, Taylor BC, Aksenov A, Callewaert C, Debelius J, Gonzalez A, Kosciolek T, McCall LI, McDonald D, Melnik AV, Morton JT, Navas J, Quinn RA, Sanders JG, Swafford AD, Thompson LR, Tripathi A, Xu ZZ, Zaneveld JR, Zhu Q, Caporaso JG and Dorrestein PC. Best practices for analysing microbiomes. Nat Rev Microbiol 2018; 16: 410-422.
[45] Eisenhofer R, Minich JJ, Marotz C, Cooper A, Knight R and Weyrich LS. Contamination in low microbial biomass microbiome studies: issues and recommendations. Trends Microbiol 2019; 27: 105-117.
[46] Mirzayi C and Renson A; Genomic Standards Consortium; Massive Analysis and Quality Control Society, Zohra F, Elsafoury S, Geistlinger L, Kasselman LJ, Eckenrode K, van de Wijgert J, Loughman A, Marques FZ, MacIntyre DA, Arumugam M, Azhar R, Beghini F, Bergstrom K, Bhatt A, Bisanz JE, Braun J, Bravo HC, Buck GA, Bushman F, Casero D, Clarke G, Collado MC, Cotter PD, Cryan JF, Demmer RT, Devkota S, Elinav E, Escobar JS, Fettweis J, Finn RD, Fodor AA, Forslund S, Franke A, Furlanello C, Gilbert J, Grice E, Haibe-Kains B, Handley S, Herd P, Holmes S, Jacobs JP, Karstens L, Knight R, Knights D, Koren O, Kwon DS, Langille M, Lindsay B, McGovern D, McHardy AC, McWeeney S, Mueller NT, Nezi L,

Olm M, Palm N, Pasolli E, Raes J, Redinbo MR, Rühlemann M, Balfour Sartor R, Schloss PD, Schriml L, Segal E, Shardell M, Sharpton T, Smirnova E, Sokol H, Sonnenburg JL, Srinivasan S, Thingholm LB, Turnbaugh PJ, Upadhyay V, Walls RL, Wilmes P, Yamada T, Zeller G, Zhang M, Zhao N, Zhao L, Bao W, Culhane A, Devanarayan V, Dopazo J, Fan X, Fischer M, Jones W, Kusko R, Mason CE, Mercer TR, Sansone SA, Scherer A, Shi L, Thakkar S, Tong W, Wolfinger R, Hunter C, Segata N, Huttenhower C, Dowd JB, Jones HE and Waldron L. Reporting guidelines for human microbiome research: the STORMS checklist. Nat Med 2021; 27: 1885-1892.

[47] Quince C, Walker AW, Simpson JT, Loman NJ and Segata N. Shotgun metagenomics, from sampling to analysis. Nat Biotechnol 2017; 35: 833-844.

[48] Nearing JT, Douglas GM, Hayes MG, MacDonald J, Desai DK, Allward N, Jones CMA, Wright RJ, Dhanani AS, Comeau AM and Langille MGI. Microbiome differential abundance methods produce different results across 38 datasets. Nat Commun 2022; 13: 342.

[49] Collins GS, Reitsma JB, Altman DG and Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ 2015; 350: g7594.

[50] Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, de Vet HC, Kressel HY, Rifai N, Golub RM, Altman DG, Hooft L, Korevaar DA and Cohen JF; STARD Group. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015; 351: h5527.

[51] Vandeputte D, Kathagen G, D'hoe K, Vieira-Silva S, Valles-Colomer M, Sabino J, Wang J, Tito RY, De Commer L, Darzi Y, Vermeire S, Falony G and Raes J. Quantitative microbiome profiling links gut community variation to microbial load. Nature 2017; 551: 507-511.

[52] Imperiale TF, Porter K, Zella J, Gagrat ZD, Olson MC, Statz S, Garces J, Lavin PT, Aguilar H, Brinberg D, Berkelhammer C, Kisiel JB and Limburg PJ; BLUE-C Study Investigators. Next-generation multitarget stool DNA test for colorectal cancer screening. N Engl J Med 2024; 390: 984-993.

[53] Ladabaum U, Mannalithara A, Weng Y, Schoen RE, Dominitz JA, Desai M and Lieberman D. Comparative effectiveness and cost-effectiveness of colorectal cancer screening with blood-based biomarkers (Liquid Biopsy) vs fecal tests or colonoscopy. Gastroenterology 2024; 167: 378-391.

[54] Piccinno G, Thompson KN, Manghi P, Ghazi AR, Thomas AM, Blanco-Míguez A, Asnicar F, Mladenovic K, Pinto F, Armanini F, Punčochář M, Piperni E, Heidrich V, Fackelmann G, Ferrero G, Tarallo S, Nguyen LH, Yan Y, Keles NA, Tuna BG, Vymetalkova V, Trompetto M, Liska V, Hucl T, Vodicka P, Bencsiková B, Čarnogurská M, Popovici V, Marmorino F, Cremolini C, Pardini B, Cordero F, Song M, Chan AT, Derosa L, Zitvogel L, Huttenhower C, Naccarati A, Budinska E and Segata N. Pooled analysis of 3,741 stool metagenomes from 18 cohorts for cross-stage and strain-level reproducible microbial biomarkers of colorectal cancer. Nat Med 2025; 31: 2416-2429.

[55] Vickers AJ and Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006; 26: 565-574.

Supplementary Table 1. Advanced adenoma-related microbiome evidence

Study	Year	Country/ region	Population	Sample type	Microbiota method	Advanced adenoma definition	Advanced adenoma sample size	Comparator group	Marker/ model	Reported performance	Validation type	Main limitation	Reference number
Zackular et al.	2014	USA	Colonoscopy-based groups	Stool	16S rRNA sequencing	Not clearly reported	Not clearly reported	Controls and colorectal lesion groups	Microbial community model	Not clearly reported for advanced adenoma	Internal	Advanced adenoma-specific estimates not clearly separated	[14]
Baxter et al.	2016	USA	Colonoscopy participants	Stool	16S rRNA plus FIT	Not clearly reported	Not clearly reported	Participants with and without colonic lesions	Microbiota + FIT model	Improved FIT-based detection reported; Advanced adenoma-specific performance was not separately reported	Internal	External validation and advanced adenoma-specific reporting limited	[27]
Yachida et al.	2019	Japan	Stage-stratified colorectal neoplasia	Stool	Metagenomics and metabolomics	Not clearly reported	Not clearly reported.	Controls, adenoma, CRC stage groups	Microbial and metabolic phenotypes	Stage-specific changes reported; advanced adenoma-specific diagnostic performance not clearly reported	Internal	Screening implementation and advanced adenoma-specific diagnostic performance remain insufficiently established	[28]
Feng et al.	2015	China/ Europe	Adenoma-CRC sequence	Stool	Metagenomics	Not clearly reported	Not clearly reported	Controls, adenoma, CRC	Stage-associated microbiome profile	Microbial shifts along adenoma-carcinoma progression reported; advanced adenoma performance not clearly separated	Cross-cohort	Advanced adenoma-specific endpoints should be interpreted separately when reported	[16]
Liang et al.	2017	China/ Hong Kong	Colorectal cancer and adenoma evaluation	Stool	Fecal bacterial marker assay	Not clearly reported	Not clearly reported	Controls and colorectal neoplasia groups	Fecal bacteria biomarker panel	Not clearly reported for advanced adenoma	Not clearly reported	Advanced adenoma-specific sample size and performance were not separately reported	[29]
Wong et al.	2017	China/ Hong Kong	Advanced colorectal neoplasia	Stool	qPCR plus FIT	Advanced colorectal neoplasia; advanced adenoma component not clearly reported	Not clearly reported	Screening/ colonoscopy comparator groups	Fecal F. nucleatum + FIT	Improved FIT-based detection of advanced colorectal neoplasia reported; Specific numeric performance for isolated advanced adenoma was not separately reported	Validation cohort	Advanced adenoma-only interpretation remains limited because the reported endpoint was advanced colorectal neoplasia	[18]
Liang et al.	2020	China/ Hong Kong	Colorectal adenoma and cancer	Stool	Fecal marker assay	Not clearly reported	Not clearly reported	Controls, adenoma, CRC	Lachnoclostridium marker	Adenoma and CRC signal reported; advanced adenoma-specific performance not clearly reported	Not clearly reported	Advanced adenoma-specific subgroup performance was not separately reported	[30]
Konishi et al.	2022	Japan	Screening-oriented cohort	Stool	Machine-learning microbiota analysis	Not clearly reported	Not clearly reported	Screening cohort comparator groups	ML-based gut microbiota model	CRC screening method evaluated; advanced adenoma-specific performance not clearly reported	Study-specific	Generalizability and prevention-endpoint performance require further testing	[31]

Supplementary Table 2. Representative studies of microbiome-derived biomarkers for colorectal neoplasia

Author, year	Country	Study population	Sample type	Detection method	Target stage	Biomarker/ model	Main findings	Validation	Key limitation	Reference number
Zackular et al., 2014	USA	Colonoscopy-based groups	Stool	16S rRNA	CRC/adenoma	Microbial community model	Fecal microbiome features were associated with colorectal lesions	Internal	Limited advanced adenoma-specific reporting	[14]
Zeller et al., 2014	France/ Germany	CRC and controls	Stool	Metagenomics	Early CRC/ CRC	Metagenomic species signature	Fecal microbial signatures showed potential for early-stage CRC detection	Cross-cohort	Screening-population validation limited	[15]
Feng et al., 2015	China/ Europe	Adenoma-CRC sequence	Stool	Metagenomics	Adenoma/ CRC	Stage-associated microbiome profile	Microbial shifts were described along adenoma-carcinoma progression	Cross-cohort	Advanced adenoma endpoints require separation	[16]
Baxter et al., 2016	USA	Colonoscopy participants	Stool	16S rRNA plus FIT	Colonic lesions	Microbiota + FIT model	Microbiota improved FIT sensitivity in the study setting	Internal	External validation and advanced adenoma data limited	[27]
Yu et al., 2017	China	CRC and controls	Stool	Metagenomics	CRC	Fecal metagenomic marker panel	Species-level markers supported non-invasive CRC detection	Validation cohort	CRC-focused rather than prevention-endpoint focused	[25]
Wong et al., 2017	China	Advanced colorectal neoplasia	Stool	qPCR plus FIT	Advanced neoplasia	F. nucleatum + FIT	Quantitative F. nucleatum improved FIT-based detection in the reported cohort	Validation cohort	Needs broader screening validation	[18]
Thomas et al., 2019	Multinational	Public metagenomic cohorts	Stool	Shotgun metagenomics	CRC	Cross-cohort metagenomic signature	Reproducible CRC-associated microbial features were identified	Cross-cohort	Advanced adenoma performance not primary endpoint	[23]
Wirbel et al., 2019	Multinational	Metagenomic cohorts	Stool	Shotgun metagenomics	CRC	Global microbial signature	CRC-specific fecal metagenomic signatures were observed	Cross-cohort	Case-control bias remains relevant	[24]
Yachida et al., 2019	Japan	Stage-stratified colorectal neoplasia	Stool	Metagenomics/ metabolomics	Adenoma to CRC	Microbial and metabolic phenotypes	Stage-specific microbial and metabolomic changes were reported	Internal	Screening implementation requires external validation	[28]
Konishi et al., 2022	Japan	Screening-oriented cohort	Stool	Machine-learning microbiota analysis	CRC screening	ML-based gut microbiota model	A gut microbiota-based screening method was developed and evaluated	Study-specific	Generalizability requires further testing	[31]

Supplementary Table 3. Representative PubMed search strategy

Database	Search strategy example
PubMed	(“colorectal cancer” OR “colorectal neoplasia” OR “colorectal adenoma” OR “advanced adenoma” OR “advanced colorectal neoplasia” OR “adenoma-carcinoma sequence”) AND (“gut microbiota” OR “gut microbiome” OR “fecal microbiome” OR “faecal microbiome” OR “mucosa-associated microbiota” OR “microbial biomarker” OR “microbiota biomarker” OR “metagenomics” OR “metabolomics” OR “16S rRNA sequencing”) AND (“screening” OR “early detection” OR “fecal immunochemical test” OR “faecal immunochemical test” OR “FIT” OR “stool DNA” OR “multitarget stool DNA”)

Note: The search was adapted for Embase, Web of Science, Scopus, and the Cochrane Library using database-specific syntax.