ICSN2023
Cancer screening: in the present, the future
June 21-23, 2023 | Turin (Italy)

Centro di Riferimento per l'Epidemiologia e la Prevenzione Oncologica in Piemonte

AZIENDA OSPEDALIERO - UNIVERSITARIA
Città della Salute e della Scienza di Torino

An efficient strategy for evaluating new non-invasive screening tests for colorectal cancer. The guiding principles

Carlo Senore[1], Robert Bresalier[2], Graeme P Young[3]

[1] Epidemiology and screening unit – CPO
University hospital Città della Salute e della Scienza, Turin, Italy

[2] MD Anderson Cancer Center, USA

[3] Flinders University, Australia

Thank you to

James Allison
Robert Benamouzig
Sally Benton
Patrick M.Bossuyt
Beatriz Carvalho
Luis Caro
Han Mo Chiu
Veerle Coupe
Evelien Dekker
Wilhelm DeKlaver
Clasine M DeKlerk
Sunil Dowlani
Callum G Fraser
William Grady
Lydia Guittet
Samir Gupta
Stephen Halloran
Ulrike Haug
Geir Hoff
Steven Itzkowitz
Tim Kortlever
Anastasios Koulaouzidids
Uri Ladabaum
Beatrice Lauby-Secretan
Marcis Leja
Theodore R. Levin
Bernard Levin
Finlay Macrae
Gerrit Meijer
Joshua Melson
Colm O'Morain
Susan Parry
Linda Rabeneck
David Ransohof
Roque Saenz
Hiroshi Saito
Silvia Sanduleanu
Robert Schoen
Kevin Selby
Harminder Singh
Robert JC Steele
Joseph Sung
Erin Symonds
Sidney Winawers

Background

New screening tests for colorectal cancer (CRC) are rapidly emerging.

Conducting trials with mortality-reduction as the endpoint supporting their adoption, is challenging and may not be feasible.

The epidemiology of CRC is changing over time, also a result of the introduction of screening, which may alter the composition of the target populations of screening interventions

A «one size fit all» approach may not be relevant

Scope of 2016 recommendations

To develop practical advice on how best to compare “new” with proven screening tests, the ideal context, the informative endpoints and the appropriate study design.

Recommendations for a Step-Wise Comparative Approach to the Evaluation of New Screening Tests for Colorectal Cancer

Graeme P. Young, MD, FRACP, FTSE, AGAF[1]; Carlo Senore, MD, MSc[2]; Jack S. Mandel, PhD, MPH[3]; James E. Allison, MD, FACP, AGAF[4]; Wendy S. Atkin, MPH, PhD[5]; Robert Benamouzig, MD, PhD[6]; Patrick M. M. Bossuyt, PhD[7]; Mahinda De Silva, MB, BS, FRACP[8]; Lydia Guittet, MD, PhD[9]; Stephen P. Halloran, MBE, FRCPath[10]; Ulrike Haug, PhD[11]; Geir Hoff, MB, ChB, PhD[12]; Steven H. Itzkowitz, MD, FACP, FACG, AGAF[13]; Marcis Leja, MD, MBA, PhD, AGAF[14]; Bernard Levin, MB, BCh, FACP[15]; Gerrit A. Meijer, MD, PhD[16]; Colm A. O'Morain, MD[17]; Susan Parry, MbCHB, FRACP18[18]; Linda Rabeneck, MD, MPH, FRCPC[19]; Paul Rozen, MD[20†]; Hiroshi Saito, MD, PhD[21]; Robert E. Schoen, MD, MPH[22]; Helen E. Seaman, BSc, PhD[23]; Robert J. C. Steele, MD, FRCS[24]; Joseph J. Y. Sung, MD, PhD[25]; and Sidney J. Winawer, MD[26]

Cancer 2016; 122(6):826-39

What has transpired?

ICSN International Cancer Screening Network
ICSN2023
Cancer screening: in the present, the future
June 21-23, 2023 | Turin (Italy)

- New developments in biomarker technologies.
- Widespread implementation of organized population screening that makes test evaluation difficult in intended-to-use populations.
- Differing goals of screening programs around the world.
- The evidence required by regulatory authorities differs from that of health-care providers.
- Omissions and updating:

  Algorithm complexity and associated challenges were not included.

  The biomarker section was very basic and did not allow for marker panels.

CPO Centro di Riferimento per l'Epidemiologia e la Prevenzione Oncologica in Piemonte
AZIENDA OSPEDALIERO - UNIVERSITARIA Città della Salute e della Scienza di Torino

Methods

Review based on the Glaser and Delphi approaches to achieving consensus.

- 47 experts (gastroenterologists, epidemiologists, surgeons, public health physicians, clinical biochemists, tumor biologists) with knowledge, or experience in practice or research relevant to CRC screening
- 12 principles progressively redrafted based on feed-back of consensus rounds, webinars, semi-structured discussions

Consensus goal of 80% agreement (agree or strongly agree – on a 5 points Likert scale) was achieved after 4 rounds for all 12 principles

Methods

Industry representatives were asked to provide their views as to how regulatory bodies approve tests, but they were not involved in the definition of the content of the revised recommendations.

Explanatory text drafted for each principle based on extensive feedback and comments from the experts

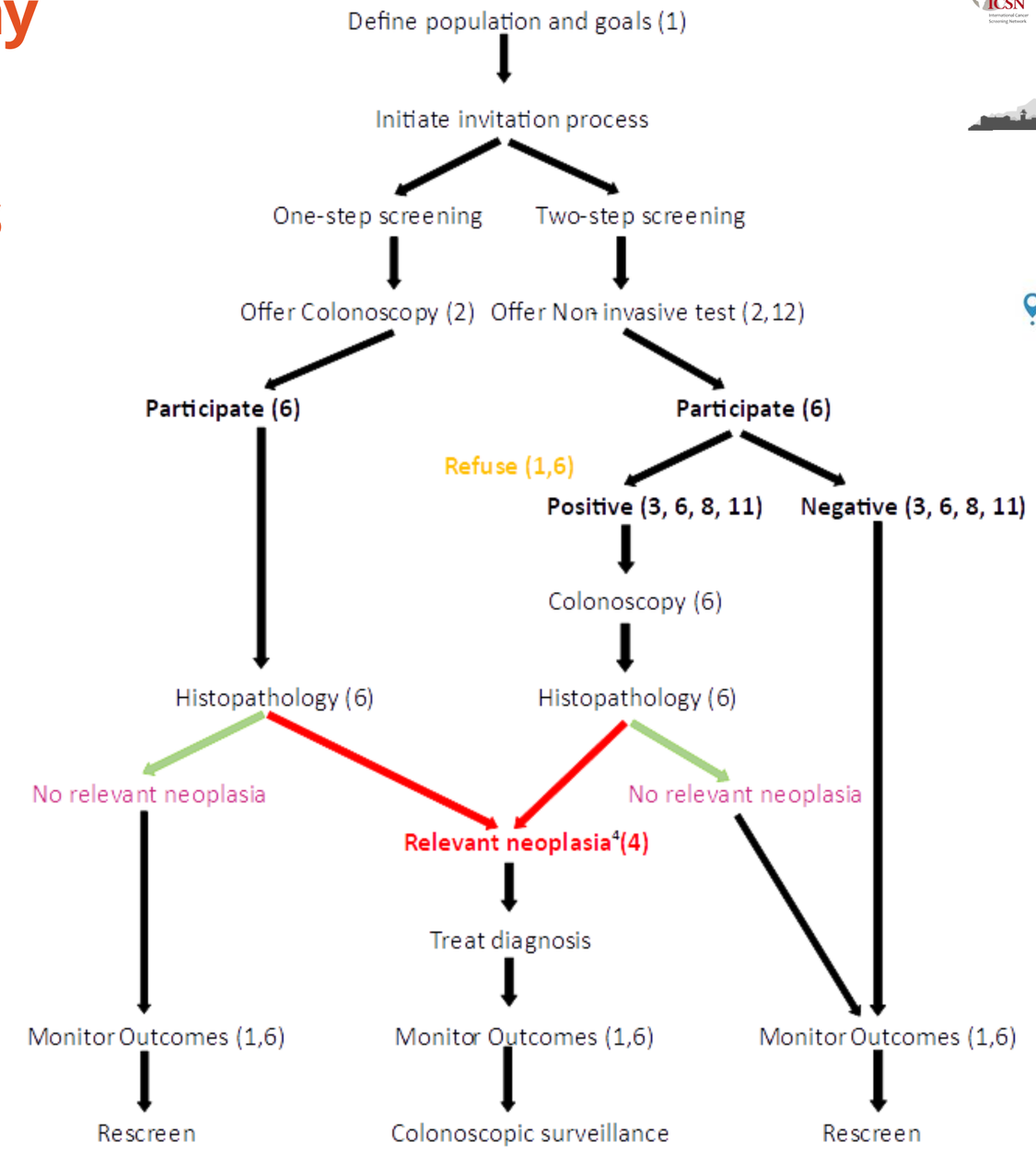
Topics Addressed in Each of the Principles Established by the Consensus Process

Principle Number	Topic
1	Desired outcomes of CRC screening
2	Screening is a multi-step process
3	A screening test identifies individuals with an increased likelihood of CRC and/or advanced precursor lesions
4	Nature of precursor lesions most important to detect
5	New biomarkers might detect lesions with a different natural history
6	Outcomes to be estimated in a screening population
7	Expectations of a new non-invasive test
8	An adjustable test positivity threshold accommodates different program goals
9	Predicting value by paired comparison to a proven non-invasive test
10	Evaluation proceeds through increasingly complex phases
11	Accuracy required for evaluation in a screening population
12	Analytic specifications, standards, and performance

Multistep screening pathway characteristic of organized screening programs one-and two-step strategies

Performing a screening test is just one event in a complex process that starts with an invitation to get tested and proceeds through diagnostic follow-up, and treatment for identified lesions, with further screening and surveillance as indicated

Demonstrating the value of a screening test must be rigorous and verified at all relevant steps of the screening pathway



ICSN2023
Cancer screening: in the present, the future
June 21-23, 2023 | Turin (Italy)

Guiding principles

Adoption of a new screening test requires evidence of effectiveness from a direct comparison with a proven test using intermediate outcomes, as long as the association of these outcomes to the expected health impact of screening was previously documented.

Cancer-specific mortality is not essential as an end-point, **provided that the mortality benefit of the comparator has been demonstrated** and that the biologic basis of detection is similar.

A rigorous and efficient four-phased approach is proposed

Phased evaluation

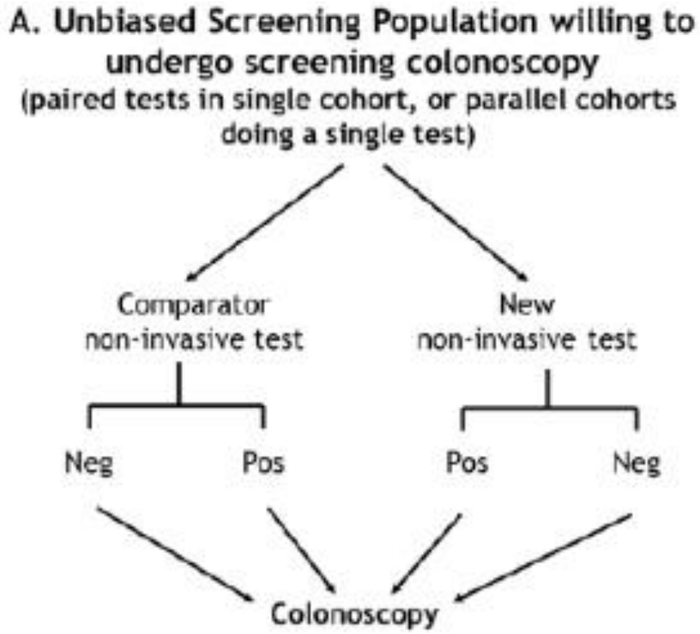
Phase	Goal(s)	Context	Approach and measures	Hurdle for progression
1	**Main:** Differentiates between CRC and non-neoplastic states?	Prescreening cohorts – limited	Distribution of test results in cohorts with and without CRC	• Test result must differ significantly in cancer cases.
2	**Main:** Detects early cancer and precursor lesions? **Others:** Initial positivity threshold? Accuracy relative to comparator? Causes of false-positives.	Prescreening cohorts - extensive	Distribution of test results in cohorts with CRC relevant precursor lesions, other colorectal diagnoses and no disease. Parallel or paired testing of new and comparator tests will be informative.	• Preliminary (although biased) estimates of accuracy are shown to be promising. • ROC analysis identifies a suitable positivity threshold.

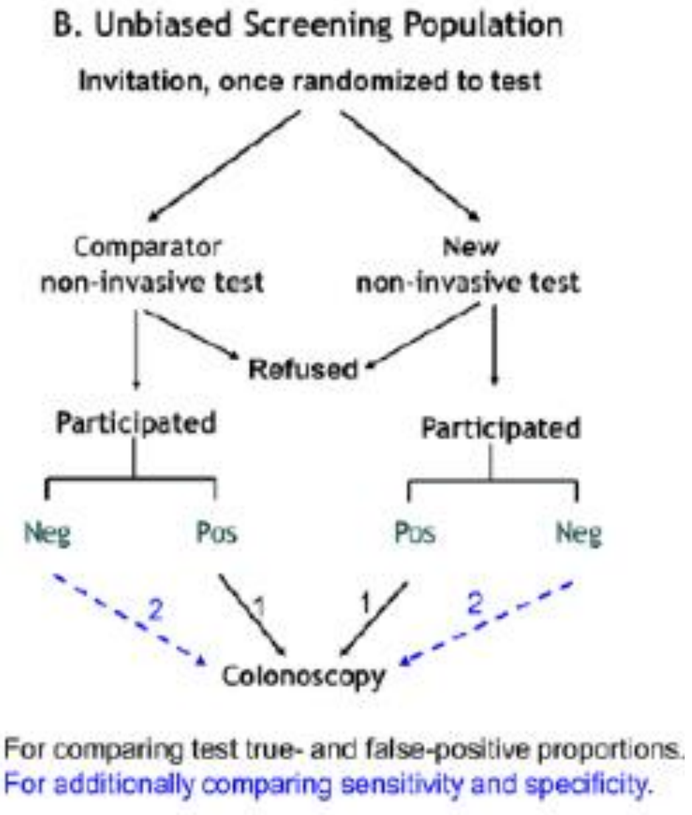
Results from Phases 1 and 2 studies justify moving to a subsequent (necessary) step in the evaluation, to be performed in a screening population, but decisions to approve, recommend and use a new test should not be made on the basis of results in Phases 1 and 2.

Phased evaluation

3	**Main:** Test accuracy in a typical screening evaluation? Test acceptance? **Others:** Test failure rate? Other variables for modelling effectiveness and cost-effectiveness.	Screening populations – single round	Apply test prospectively to a typical unbiased intended-use population. Choose study design appropriate to program goal and jurisdictional context: e.g., colonoscope all for estimating test accuracy, parallel testing for comparing non-invasive tests and intention-to-screen outcomes.	• A significant improvement in some aspect of screening. • Non-inferior in accuracy to a comparator test, OR • Accuracy likely delivers benefit. • Feasible colonoscopy workload. • Modeled effectiveness and cost-effectiveness are satisfactory.
4	**Main:** Missed lesions or adverse events? **Others:** Participation rates over time and re-test intervals?	Screening population – multiple rounds	Apply the test prospectively to an intended-use screening population over multiple rounds, with careful monitoring of population program outcomes.	

Study design frameworks applicable to Phase 3 studies





1 For comparing test true- and false-positive proportions.
2 For additionally comparing sensitivity and specificity.

A: design appropriate to determine test accuracy where all cases undergo colonoscopy, but intention-to-screen outcomes cannot be ascertained (comparison of a comparator with the new test can be paired in a single cohort or parallel in separate cohorts). B: design appropriate for estimating intention-to-screen outcomes and where the accuracy of the new test can be compared to that of a non-invasive comparator either when colonoscoping only test-positive individuals (compare true- and false-positive fractions) or all participants (sensitivity and specificity).

Only test-positive cases undergo colonoscopy

Test result	Diagnostic verification	Result of diagnostic verification	Related accuracy characteristic	Program consequence
Positive	**YES**	True, hence true-positive (TP_1, TP_2)	**Relative sensitivity**; TP_1/TP_2	**Detection of neoplasia.**
			Positive predictive value; $TP_1 / (TP_1 + FP_1)$ $TP_2 / (TP_2 + FP_2)$	Efficiency of detection.
		False, hence false-positive (FP_1, FP_2)	**Relative false-positive rate**; FP_1/FP_2	**Burden associated with detection.**
Negative	**NO**			

All cases undergo colonoscopy

Test result	Diagnostic verification	Result of diagnostic verification	Related accuracy characteristic	Program consequence
Positive	**YES**	True, hence true-positive (TP)	**Sensitivity** TP / (TP + FN).	**Detection of neoplasia.**
			Positive predictive value TP / (TP + FP)	Efficiency of detection.
		False, hence false-positive (FP)		Burden associated with detection.
Negative	YES	True, hence true-negative (TN)	**Specificity** TN / (TN + FP)	**Accuracy of detection.**
			Negative predictive value TN / (TN + FN)	Exclusion of neoplasia
		False, hence false-negative (FN)		Miss rate (missed lesions).

Comparator test

Guiding principles 2016

gFOBT is the **minimum standard** (29-47% sensitivity for CRC; 87%-98% specificity)

FIT also **acceptable** (superior to gFOBT)

FIT New standard 2023:

- improved sensitivity for CRC and better capacity to detect (advanced) adenomas
- lower interval cancer rate
- repeated testing improves detection
- preliminary reports from population based studies are suggestive for the effectiveness of FIT in reducing CRC incidence and mortality

(vanRossum et al. 2008; Hol et al. 2009; Wieten et al. 2019; Crotta et al 2012; Kapidzic et al 2014; Zorzi et al. 2018; Ventura et al. 2014; Zorzi et al. 2015; Levin et al. 2018)

Comparator test

USPSTF suggests (single test) *(Bibbins-Domingo et al., 2016)*

- acceptable **sensitivity for CRC** (all stages) to be at least **70%**
- acceptable **specificity** to be at least **90%**

Desirable standards have not been set for advanced adenomas

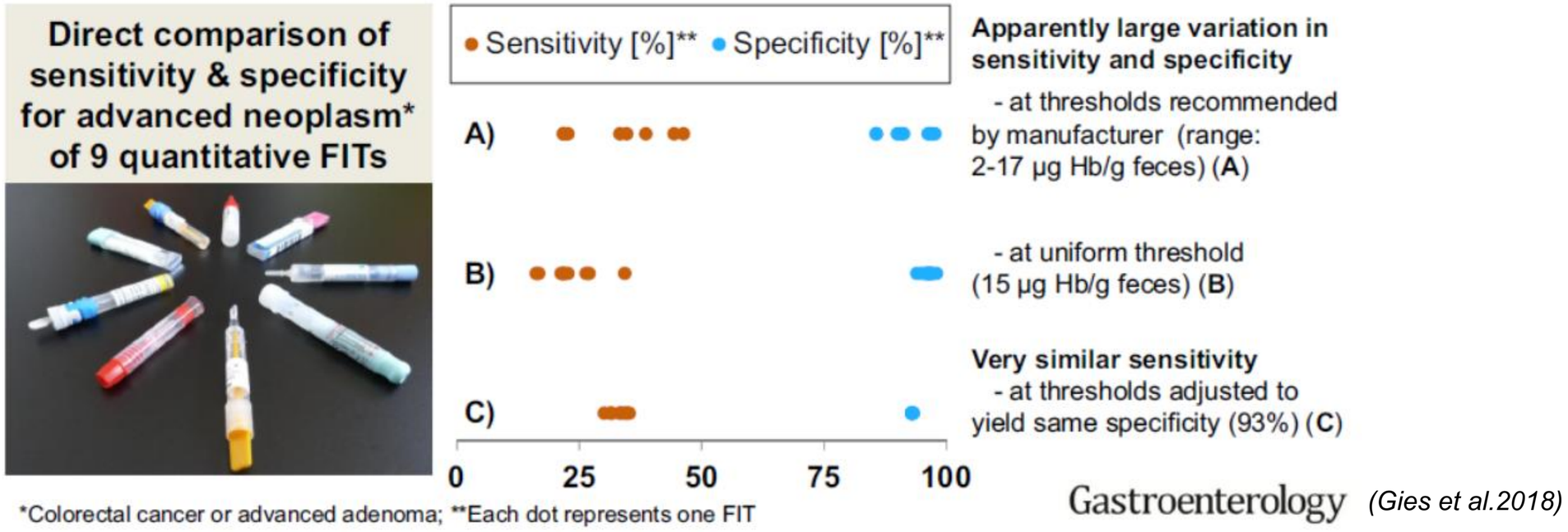
FIT (cut-off 20 µg Hb/gr.)

- sensitivity for CRC (all stages): 75% (61-86%)
- specificity: 95% (92-98%)

(Imperiale et al.,2019)

FIT as a comparator

Different quantitative methods are showing different performance characteristics



Gastroenterology *(Gies et al.2018)*

Comparing FIT methods performance at identical positivity rates

(Grobbee et al. 2016; Passamonti et al. 2018)

Positivity threshold should ideally be low, in studies with colonoscopy follow-up only for subjects testing positive at either test, to allow simulating comparisons for different cut-off levels

Adjustable test positivity threshold

New test: Cut-off needs to be flexible

performance should be measured **simulating different cut-off levels**, focusing on the **positivity range of established population based programs**

Established standards will depend upon the nature of a healthcare system, what it wants to achieve and what is feasible (3).

Limitations in endoscopy resources could lead providers to set the positivity threshold to match colonoscopy capacity, even though test sensitivity might be compromised (8)

Target population should be informed of estimated test accuracy and its expected benefit, using well designed communication strategies to empower individuals to make their own decision (3,4)

Characteristics of the new test

Based on the discussion of these principles, a new non-invasive test should achieve at least some of the following, when compared to an established test that has been demonstrated to have a positive impact on CRC-specific mortality:

- Be flexible
- Improve sensitivity for relevant neoplasia, while maintaining acceptable specificity.
- Improve precursor lesion detection and hence reduce CRC incidence.
- Improve participation rates over initial and subsequent rounds.

Conclusions

We re-examined the **principles underlying evaluation of new non-invasive tests** in view of technological developments and identification of new biomarkers.

The consensus process reaffirmed the view that **comparative evaluation of a new with a proven non-invasive comparator test**, with well-establish benefits for incidence and mortality reduction, is a powerful approach for new test evaluation.

This framework allows for a dynamic process that has a broad application.

This process is not bound to a specific test.

Conclusions

By recognizing that screening is a complex multistep process, this revision of the recommendations shifts the adequacy of the evidence in support of a new test from Phase 2 to Phase 3 and 4 and it provides more detail on the expected outcomes assessed in these phases.

Phased evaluation in a stepwise manner is an efficient way first to establish the potential value of a new test and then to gather the evidence that will lead to its acceptance by professionals, healthcare providers, and regulatory bodies.

An efficient strategy for evaluating new non-invasive screening tests for colorectal cancer: The guiding principles

Authors: Bresalier* R.S., Senore*, C., Young*, G.P., Allison, J., Benamouzig, R, Benton, S., Bossuyt, P.M., Caro, L., Carvalho, B., Chiu, H.M., Coupe, V.M.H., de Klaver, W., de Klerk, C.M., Dekker, E., Dolwani, S., Fraser, C.G., Grady, W.M., Guittet, L., Gupta, S., Halloran, S.P., Haug, U., Hoff, G., Itzkowitz, S.H., Kortlever, T.L., Koulauzidis, A., Ladabaum, U., Lauby-Secretan, B., Leja, M., Levin, B., Levin, T.R., Macrae, F., Meijer, G.A., Melson, J., O'Morain, C., Parry, S., Rabeneck, L., Ransohoff, D.F., Saenz, R., Saito, H., Sanduleanu, S., Schoen, R.E., Selby, K., Singh, H., Steele, R.J.C., Sung, J.J.Y., Symonds, E., Winawer, S.J. (Members of the WEO CRC Screening New Test Evaluation Expert Working Group)

Thank you for your attention

carlo.senore@cpo.it