

Blinded validation demonstrates early detection and localization with a three-level methylation-based multi-cancer blood test

Johanna B. Withers, PhD¹; Elie Massaad, MD, MSc¹; A. Gregory DiRienzo, PhD¹; Dorna KashefHaghighi, PhD¹; Jocelyn Charlton, PhD¹; Anthony P. Shuber, MS¹; Hutan Ashrafian, MD, PhD, MBA¹
¹Harbinger Health, Cambridge, MA

INTRODUCTION

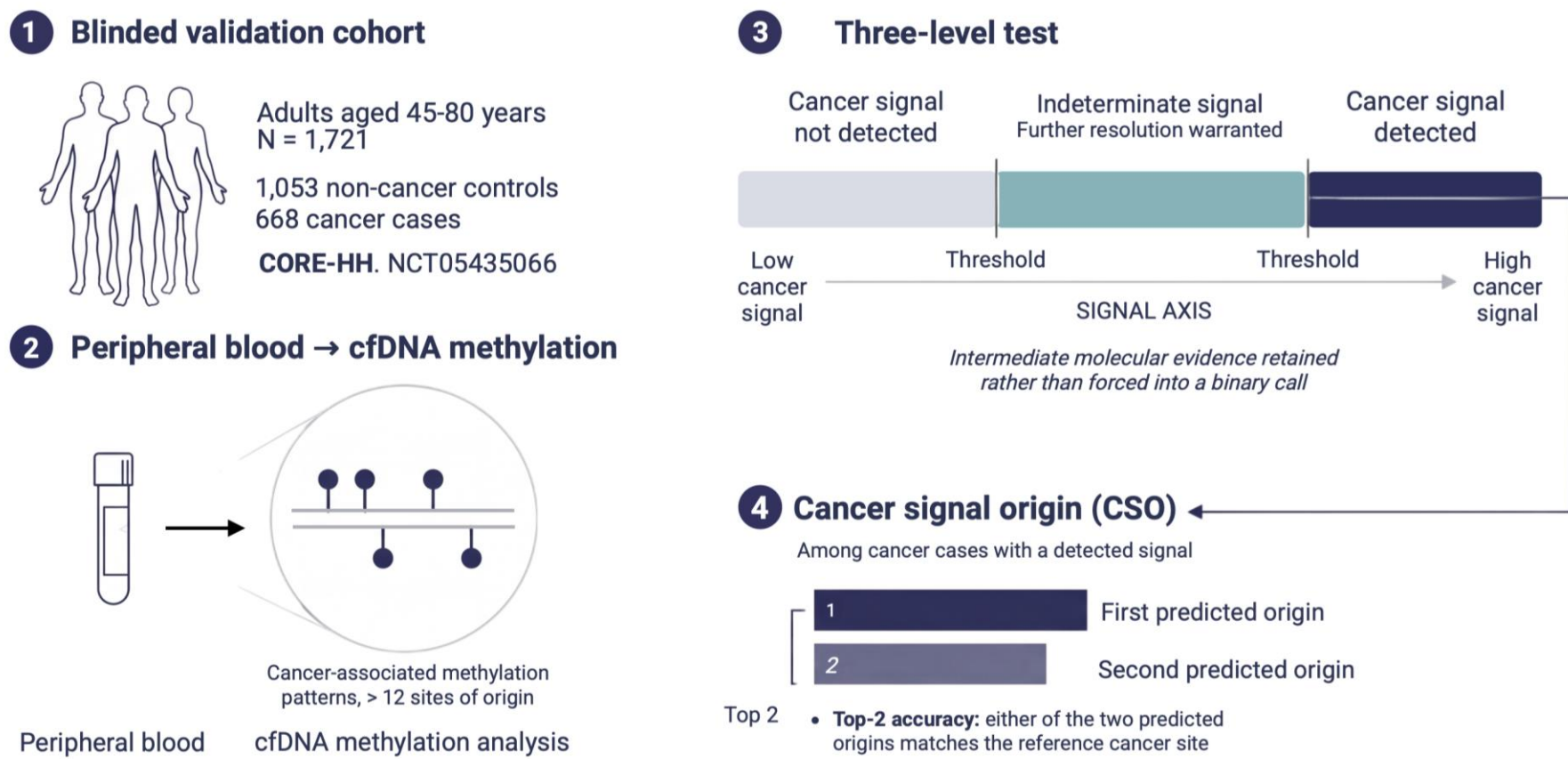
- In patients with clinical suspicion of cancer based on nonspecific symptoms, laboratory abnormalities, or indeterminate imaging, sequential organ-directed evaluation may be required to confirm malignancy and identify its site of origin, particularly for pancreatobiliary, hepatobiliary, and upper gastrointestinal cancers.
- Cell-free DNA methylation can support this evaluation by distinguishing cancer-associated signals and informing cancer signal origin through tissue-specific methylation patterns.
- Cancer-associated cfDNA exists along a continuum of molecular signal, particularly in earlier-stage or lower-burden disease where tumor-derived DNA may be limited. A three-level classifier (Not Detected, Indeterminate, Detected) preserves intermediate evidence near the classification boundary rather than forcing uncertain signal into a binary result, potentially reducing false-negative assignment while maintaining a low false-positive proportion in the Cancer Signal Detected category.
- We performed blinded validation of a three-level cfDNA methylation-based multi-cancer test for cancer detection and localization across >12 sites, with stage-specific evaluation including GI malignancies.

OBJECTIVES

- To validate the performance of a three-level cfDNA methylation-based multi-cancer test for cancer signal detection and cancer signal origin (CSO) localization, including stage-specific performance across GI and other malignancies.
- To assess whether an Indeterminate result category improves separation of Cancer Signal Detected and Cancer Signal Not Detected results by separately classifying samples that do not meet the thresholds for either definitive category.

METHODS

- A held-out, blinded validation cohort from the CORE-HH case-control study (NCT05435066) included adults aged 45–80 years without a reported cancer diagnosis and participants with treatment-naïve cancer. Peripheral blood samples were analyzed using a cfDNA methylation-based multi-cancer test that evaluates cancer-associated methylation patterns across >12 potential sites of origin and assigns one of three prespecified outputs: Cancer Signal Not Detected, Indeterminate, or Cancer Signal Detected. For Cancer Signal Detected results, the model also reports the predicted cancer signal origin. Test performance was evaluated by cancer type and stage, including early-stage (I–II) and late-stage (III–IV) disease.¹



RESULTS

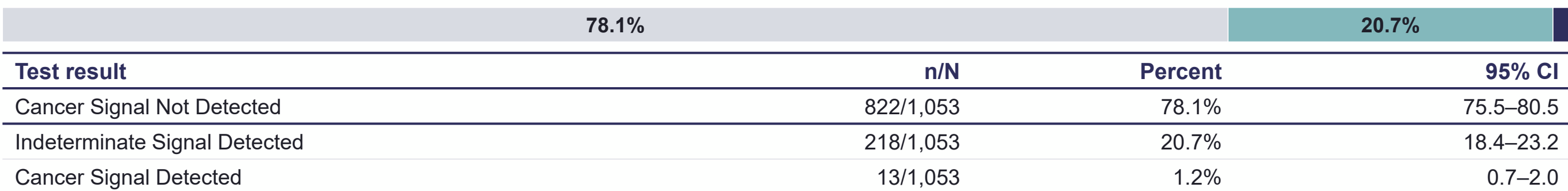
Validation cohort: 1,721 participants (1,053 non-cancer controls and 668 cancer cases, including 214 gastrointestinal cancer cases)

Table 1. Distribution of Cancer Cases by Gastrointestinal Cancer Site and Stage in the Validation Cohort

Cancer site group	All stages	Stage I–II	Stage III–IV	Unstaged
Colorectal cancer	39	19	18	2
Esophageal, GEJ, and gastric cancers	74	28	36	10
Pancreatic and gallbladder cancers	75	29	41	5
Liver and biliary tract cancers	26	7	16	3

Unstaged cases are excluded from the stage-specific analyses in Figure 2.

Figure 1. Distribution of Test Results Among Non-Cancer Controls



"Specificity" at the Cancer Signal Detected threshold was 98.8% (1,040/1,053; 95% CI, 98.0–99.3). For this calculation, Indeterminate and Cancer Signal Not Detected results were counted as nonpositive.

Figure 2. Test Results by Gastrointestinal Cancer Site and Stage

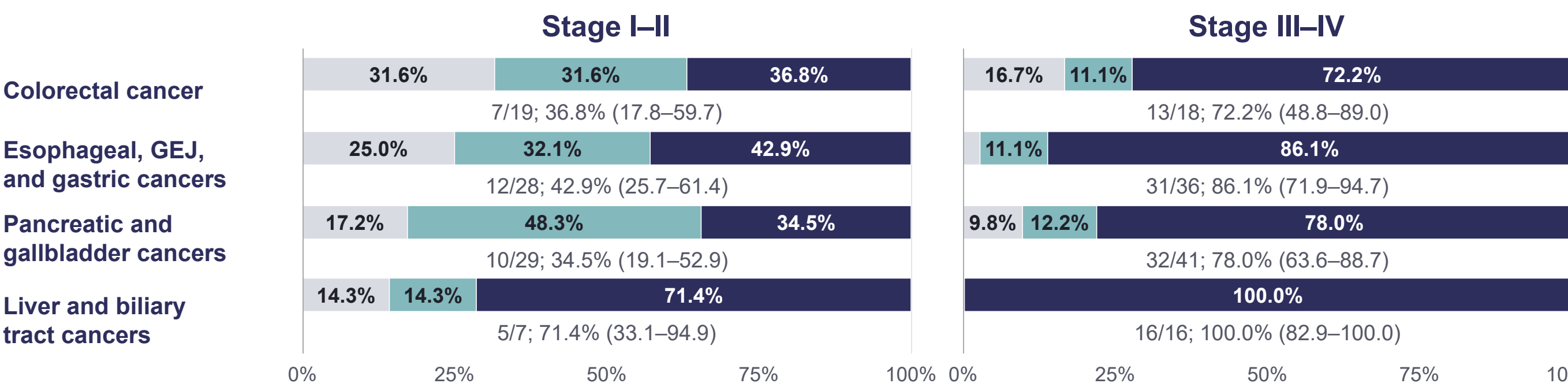
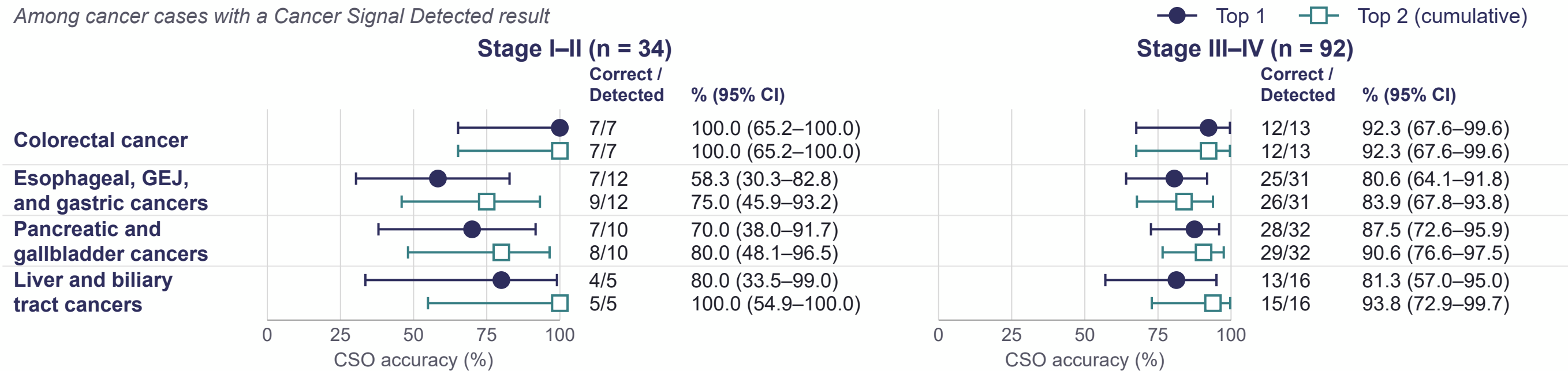


Figure 3. Cancer signal origin (CSO) accuracy by gastrointestinal cancer site and stage



DISCUSSION

- The proportion of cases with a Cancer Signal Detected result increased with stage across GI malignancies, consistent with the expected relationship between tumor burden and circulating tumor-derived signal. Importantly, clinically meaningful detection was observed in stage I–II disease, including cancers for which early diagnosis remains challenging.
- The three-level test preserved intermediate molecular evidence that would be lost with a binary threshold. This was particularly apparent in early-stage pancreatic/gallbladder cancers, for which 48.3% of cases were classified as Indeterminate, with 34.5% classified as Detected. These findings suggest that the Indeterminate category identifies a biologically distinct region of the molecular signal continuum rather than simply representing an analytical failure or false-negative result.
- Among Cancer Signal Detected results, cancer signal origin accuracy (Top 2) for early-stage cancers ranged from 75% (Esophageal, GEJ and gastric) to 100%, and for late-stage cancers from 84% (Esophageal, GEJ and gastric) to 94%, evidencing the potential for providing clinically useful direction for subsequent diagnostic evaluation.
- The low observed false-positive proportion (1.2%) in the Cancer Signal Detected category supports the potential for high PPV in clinical settings, depending on cancer incidence and sensitivity, while the separate Indeterminate category captures intermediate classifier results. The clinical value of this architecture will ultimately depend on how Indeterminate results modify the post-test cancer distribution and are incorporated into subsequent evaluation, including longitudinal or repeat-testing strategies.

CONCLUSIONS

- In blinded validation, the cfDNA methylation-based multi-cancer test demonstrated high specificity (98.8%) with stage- and cancer type-dependent detection and localization capability, including clinically relevant performance in upper GI, pancreatobiliary, hepatobiliary, and colorectal cancers.
- The three-level classification framework preserves the molecular signal that may otherwise be assigned as false-negative result in a binary model, while maintaining a Cancer Signal Detected category with the potential for high PPV. These findings support further prospective evaluation, including whether longitudinal testing can clarify the clinical pathway for Intermediate Signal Detected results and inform subsequent diagnostic evaluation.
- Limitations include the case–control study design and small counts within individual GI sites (wide confidence intervals).

REFERENCES

A. G. DiRienzo, E. Massaad, and H. Ashrafian, "Classification-Specific Predictive Performance: A Unified Estimation and Inference Framework for Multi-Category Tests," *Statistics in Medicine* 45, no. 3–5 (2026): e70431, <https://doi.org/10.1002/sim.70431>.

DISCLOSURES

This study was sponsored by Harbinger Health, Cambridge, MA

ACKNOWLEDGEMENTS

We gratefully acknowledge all participants for their contributions, without whom this research would not have been possible. We also thank all cross-functional teams for their dedicated efforts that made this work and the resulting data possible.