ESMO
GOOD SCIENCE
BETTER MEDICINE
BEST PRACTICE

ANNALS OF ONCOLOGY
driving innovation in oncology

ORIGINAL ARTICLE

Clinical validation of a targeted methylation-based multi-cancer early detection test using an independent validation set

E. A. Klein[1*], D. Richards[2], A. Cohn[3], M. Tummala[4], R. Lapham[5], D. Cosgrove[6], G. Chung[7], J. Clement[8], J. Gao[9], N. Hunkapiller[9], A. Jamshidi[9], K. N. Kurtzman[9], M. V. Seiden[10], C. Swanton[11,12] & M. C. Liu[13]

[1]Glickman Urological and Kidney Institute, Cleveland Clinic, Cleveland; [2]The US Oncology Network, Tyler; [3]The US Oncology Network, Denver; [4]Mercy Clinic Cancer Center, Springfield; [5]Spartanburg Regional Healthcare System, Spartanburg; [6]Compass Oncology, Vancouver; [7]The Christ Hospital Health Network, Cincinnati; [8]Hartford HealthCare Cancer Institute, Hartford; [9]GRAIL, Inc., Menlo Park; [10]US Oncology Research, The Woodlands, USA; [11]The Francis Crick Institute, London; [12]University College London Cancer Institute, London, UK; [13]Mayo Clinic, Rochester, USA

Background: A multi-cancer early detection (MCED) test used to complement existing screening could increase the number of cancers detected through population screening, potentially improving clinical outcomes. The Circulating Cell-free Genome Atlas study (CCGA; NCT02889978) was a prospective, case-controlled, observational study and demonstrated that a blood-based MCED test utilizing cell-free DNA (cfDNA) sequencing in combination with machine learning could detect cancer signals across multiple cancer types and predict cancer signal origin (CSO) with high accuracy. The objective of this third and final CCGA substudy was to validate an MCED test version further refined for use as a screening tool.
Patients and methods: This pre-specified substudy included 4077 participants in an independent validation set (cancer: $n = 2823$; non-cancer: $n = 1254$, non-cancer status confirmed at year-one follow-up). Specificity, sensitivity, and CSO prediction accuracy were measured.
Results: Specificity for cancer signal detection was 99.5% [95% confidence interval (CI): 99.0% to 99.8%]. Overall sensitivity for cancer signal detection was 51.5% (49.6% to 53.3%); sensitivity increased with stage [stage I: 16.8% (14.5% to 19.5%), stage II: 40.4% (36.8% to 44.1%), stage III: 77.0% (73.4% to 80.3%), stage IV: 90.1% (87.5% to 92.2%)]. Stage I-III sensitivity was 67.6% (64.4% to 70.6%) in 12 pre-specified cancers that account for approximately two-thirds of annual USA cancer deaths and was 40.7% (38.7% to 42.9%) in all cancers. Cancer signals were detected across >50 cancer types. Overall accuracy of CSO prediction in true positives was 88.7% (87.0% to 90.2%).
Conclusion: In this pre-specified, large-scale, clinical validation substudy, the MCED test demonstrated high specificity and accuracy of CSO prediction and detected cancer signals across a wide diversity of cancers. These results support the feasibility of this blood-based MCED test as a complement to existing single-cancer screening tests.
Clinical trial number: NCT02889978.
Key words: cancer, multi-cancer early detection, liquid biopsy, methylation, cell-free nucleic acids, machine learning

INTRODUCTION

Cancer will soon be the leading cause of mortality globally, and while development of more effective therapies is needed,[1] many may only prolong survival by a few months, if at all.[2] In addition to potentially reducing mortality, improving population-scale early detection also reduces disease- and treatment-related morbidity,[3-10] increases the likelihood of treatment success,[6,9] improves quality of life,[11] and reduces treatment cost and complexity.[6,11]

Currently, only five cancer screening tests are available in the USA (breast, colorectal, cervical, lung, and prostate), collectively accounting for only 42% of annual cancer incidence in people aged 50-79 years.[12-16] While these screening tests reduce cancer-specific mortality,[5,17] they are associated with high false-positive rates, overdiagnosis and overtreatment,[18,19] disparities in adherence,[20-25] and low positive predictive value (PPV; proportion of true positives among those with a positive test result)[26-28] as well as a high cumulative false-positive rate when used sequentially.[29]

It has been estimated that cancer detection before stage IV could reduce cancer-related deaths by ≥15% within 5 years.[7] A new approach called multi-cancer early detection (MCED) has the potential to achieve this goal by detecting signals for multiple cancers from cell-free DNA (cfDNA) or

other circulating analytes in the blood shed by tumors. These assays use genomic sequencing or other approaches, sometimes in combination with machine learning to detect signals from cfDNA methylation,[30-33] mutations,[34,35] and/or fragmentation,[36] or other circulating analytes such as proteins.[34,35] Results from early studies of these tests have shown promise in detecting cancers at earlier stages,[30-35] although none are yet available outside of clinical trials.

The Circulating Cell-free Genome Atlas (CCGA; NCT02889978) study was designed to develop and validate an MCED test to detect cancer signals across multiple cancer types and predict cancer signal origin (CSO) via a single blood draw. Modeled data from this test have shown that its use in the general population could shift cancer detection from stage IV to earlier stages (stages I-III), potentially reducing cancer mortality.[37] CCGA was divided into three substudies; in the first, a comprehensive comparison of genomic sequencing approaches identified that whole-genome bisulfite sequencing (WGBS; detecting genome-wide DNA methylation status) outperformed other methods. In the second substudy,[30] the selected WGBS assay was refined into a targeted methylation assay, and machine learning classifiers for cancer detection and CSO prediction were developed.[30] The third and final CCGA substudy, reported herein, is a large clinical validation study of this MCED test.

PATIENTS AND METHODS

Study design

CCGA (NCT02889978) is a prospective, multi-center, case-control, observational study with longitudinal follow-up that enrolled 15 254 participants (8584 with cancer; 6670 without cancer) from 142 sites in North America between August 2016 and February 2019.[30] All participants were required to provide written informed consent. The study was approved by the Institutional Review Board or an independent ethics committee at each participating trial site and conducted in accordance with the International Conference on Harmonization for Good Clinical Practice guidelines and the Declaration of Helsinki.

CCGA was divided into three pre-specified substudies[30] (Figure 1): (i) discovery,[38,39] (ii) training and validation with the selected and updated assay and classifiers,[30] and (iii) clinical validation within an independent validation set with a further refined assay and classifiers optimized for screening (reported herein and detailed in Supplementary Material, available at https://doi.org/10.1016/j.annonc.2021.05.806). A total of 5309 qualified participants were included in the third substudy.

Participants

Adults (≥20 years of age) were enrolled as previously described.[30] Participants eligible for the cancer arm included individuals diagnosed with cancer and/or who were scheduled to undergo biopsy and/or surgical resection for known or highly suspected malignancy

(Supplementary Table S1, available at https://doi.org/10.1016/j.annonc.2021.05.806). Individuals who had received chemotherapy or radiotherapy, undergone definitive local therapy, or received more extensive surgery than that required to establish the diagnosis, before study blood draw, were ineligible. Non-cancer participants were enrolled from participating sites to control for confounding factors.

Study objectives and corresponding measures of test performance

The primary objectives of this substudy were to evaluate test performance for cancer signal detection, CSO prediction, and both combined (see Supplementary Material, available at https://doi.org/10.1016/j.annonc.2021.05.806 for details on combined measure). Test performance was defined as (i) cancer signal detection — measured by sensitivity (proportion of participants with a positive test result among all cancer participants in an analysis set) and specificity (proportion of participants with a negative test result among non-cancer participants); and (ii) CSO prediction — measured by overall accuracy of CSO prediction (proportion of participants with a correct predicted CSO label among true positive participants, excluding those who had an unknown origin) and displayed in a confusion matrix (comparing target CSO label to predicted CSO label).

Secondary objectives included (i) test performance by age group, (ii) test performance for cancer signal detection by method of cancer diagnosis [screening test (breast, cervical, colorectal, lung, and prostate) or clinical presentation], and (iii) test performance for cancer signal detection in a pre-specified group (Supplementary Table S2, available at https://doi.org/10.1016/j.annonc.2021.05.806) of 12 cancer classes (anus, bladder, colon/rectum, esophagus, head and neck, liver/bile duct, lung, lymphoma, ovary, pancreas, plasma cell neoplasm, and stomach) that account for approximately two-thirds of annual USA cancer deaths.[40] Previous studies identified these cancers as having higher cancer signal detection, consistent with their ability to release higher amounts of cfDNA into circulation.[30,39,41,42]

Exploratory objectives included an extrapolation of the PPV and negative predictive value (NPV; defined as the proportion of non-cancer participants among those with a negative test result) adjusted by the Surveillance, Epidemiology, and End Results Program (SEER) incidence rates for those aged 50-79 years (Supplementary Table S3, available at https://doi.org/10.1016/j.annonc.2021.05.806) and an evaluation of the test positive rate by American Joint Committee on Cancer (AJCC) cancer types.[43]

Other pre-specified secondary and exploratory objectives, measures, and their results are reported in the Supplementary Material, available at https://doi.org/10.1016/j.annonc.2021.05.806. A *post hoc* analysis of test performance for cancer signal detection was carried out for the following three categories: hematologic malignancies, solid tumors with common screening options, and solid tumors without common screening option.
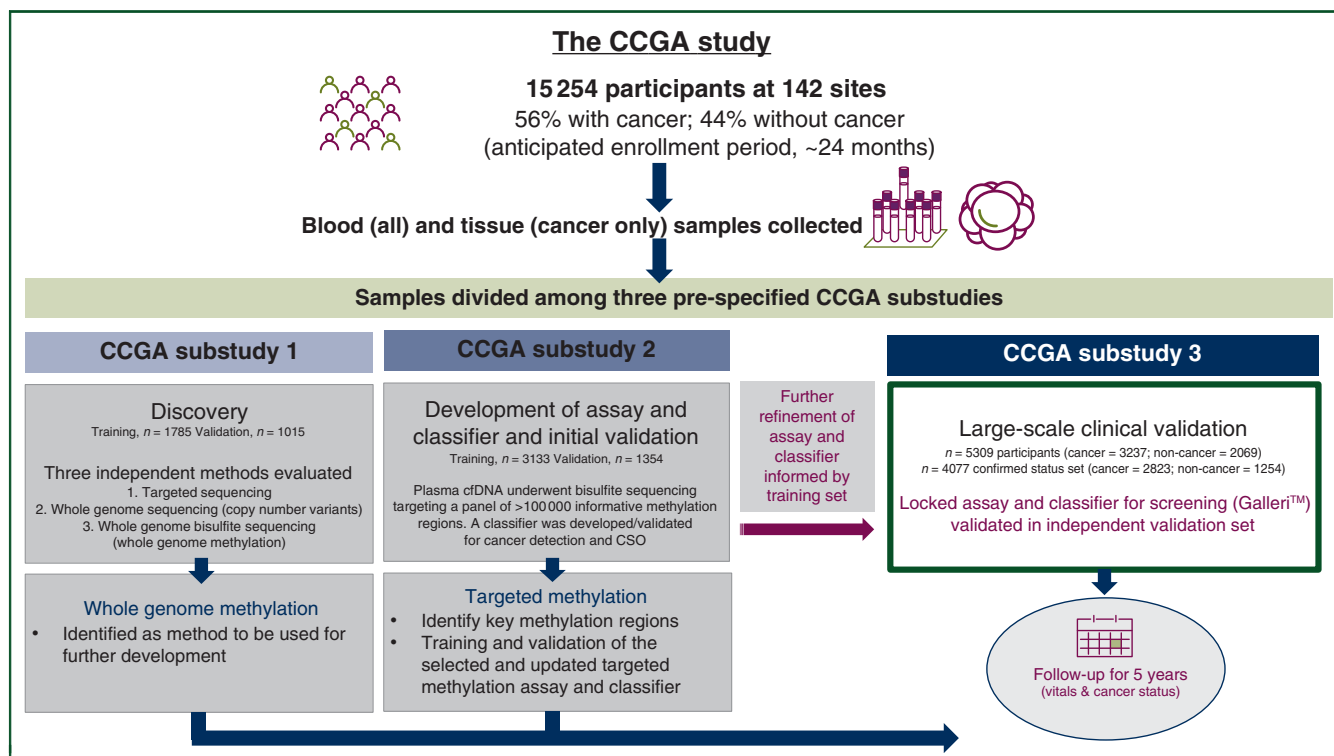
ARTICLE IN PRESS

E. A. Klein et al.   Annals of Oncology

Figure 1. Study design.
The CCGA study design is summarized. The study enrolled 15 254 participants with and without cancer to develop and validate a multi-cancer early detection test. The study was divided into three pre-specified substudies; CCGA participants not included in CCGA substudies ($n = \sim 2200$) were excluded for a variety of reasons, mainly due to incomplete or irregular clinical data at time of selection preventing selection into a substudy, availability of plasma samples, and miscellaneous other reasons. A two-stage classifier further refined for use as a screening tool relative to the one developed and validated in the prior CCGA2 substudy was trained on the data from the training set participants (see also Supplementary Methods, available at https://doi.org/10.1016/j.annonc.2021.05.806). Following the training procedure, all parameters, including thresholds, were fixed and the final two-stage classifier was applied to the independent samples from the validation set to assign cancer/non-cancer and signal origin labels to each sample.
CCGA, Circulating Cell-free Genome Atlas; cfDNA, cell-free DNA; CSO, cancer signal origin.

Sample collection, processing, and analysis

Plasma and tumor tissue sample collection, accessioning, storage, and processing were carried out as described previously and in Supplementary Material, available at https://doi.org/10.1016/j.annonc.2021.05.806.[30] To minimize bias, blood samples from cancer and non-cancer participants were randomized for processing across batches, operators, and reagent lots.

The targeted methylation assay was conducted as previously described,[30] with adjustments (Supplementary Material, available at https://doi.org/10.1016/j.annonc.2021.05.806). Briefly, plasma cfDNA (up to 75 ng) was subjected to customized bisulfite conversion reaction prepared as a dual indexed sequencing library, and enriched using standard hybridization capture conditions, for 150-bp paired-end sequencing on the Illumina NovaSeq.

Clinical data collection

Clinical, pathology, and radiology data were collected from participant questionnaires and abstracted from medical records, including reports of adverse events from the study blood draw.[30] Additionally, the World Health Organization (WHO) International Classification of Diseases for Oncology (ICD-O) morphologic and behavior codes[44] were assigned

to cancers by pathologists. Clinical stage was assigned by the treating physician or a certified cancer registry professional according to the AJCC Staging Manual (7th or 8th edition).[43,45] Cancers without an AJCC staging classification were analyzed without staging information. Additional details regarding handling of dropouts and missing data are provided in Supplementary Material, available at https://doi.org/10.1016/j.annonc.2021.05.806. Participant follow-up for clinical information was carried out annually (within ±2 months from anniversary of enrollment) from a search of medical records or direct contact with participants by clinical research staff.

Classification of cancer versus non-cancer and CSO

Briefly, custom software was built[30] to classify samples using source models that recognized methylation patterns per region as similar to those derived from a particular cancer class but potentially shared across multiple cancer classes, followed by a pair of machine learning modules — one to determine cancer/non-cancer status and the other to predict the CSO label (Supplementary Table S4, available at https://doi.org/10.1016/j.annonc.2021.05.806).

Three key classifier modifications were implemented to improve performance in a screening application (Supplementary Material, available at https://doi.org/10.

1016/j.annonc.2021.05.806): (i) the specificity threshold (i.e. false-positive rejection) was refined to account for cancer-like signals from prevalent non-malignant hematological conditions in non-cancer individuals; (ii) CSOs were refined to improve signal identification, resulting in a new CSO class — 'Neuroendocrine Cells of Lung or Other Organs' (Supplementary Figure S1, available at https://doi.org/10.1016/j.annonc.2021.05.806); and (iii) the CSO classifier was modified to remove an 'indeterminate' CSO category and return a CSO prediction for all test positive samples.

The classifier was trained on 17 339 samples [12 185 from the first and second CCGA substudies, 4891 from a separate clinical study (STRIVE; NCT03085888), and 263 obtained commercially] from 6383 distinct individuals; 1014 samples were analyzed using WGBS,[30] and 16 325 samples were analyzed using a targeted methylation assay. The locked classifier was trained to target 99.4% specificity. Samples from participants in the third CCGA study were entirely reserved for independent validation.

Statistical analyses

For demographics and baseline characteristics, descriptive statistics are reported. For categorical variables, the number and percentage of participants in each category were calculated; for continuous variables, the total number of participants (n), mean, standard deviation (SD) or standard error (SE), median, first quartile (Q1), third quartile (Q3), minimum, and maximum were calculated. The 95% confidence intervals (CIs) for these test performance measures (sensitivity, specificity, and overall accuracy of CSO prediction) were calculated using the Wilson (score) method, unless otherwise specified. No formal statistical tests were conducted. All analyses were carried out using R software, version 3.6[46] or higher.

RESULTS

Participant disposition

A total of 5309 participants (enrolled as cancer, $n = 3237$; enrolled as non-cancer, $n = 2069$; missing enrollment status, $n = 3$) were included in this third CCGA validation substudy. Of these, 4077 (cancer, $n = 2823$; non-cancer, $n = 1254$) were included in the Confirmed Status analysis set (Figure 2). The most common reasons for exclusion were incomplete year-one follow-up for non-cancer participants ($n = 324$), presence of non-malignant conditions at enrollment ($n = 283$), and unconfirmed cancer or treatment status at blood draw ($n = 171$). All exclusion categories were pre-specified before unblinding. The assay failure rate was low [0.8% (45/5309)]. A total of 0.4% (20/5309) of participants reported an adverse event related to the blood draw; 17 of 20 were mild in severity, and 3 of 20 were moderate. No serious adverse events related to the blood draw were reported.
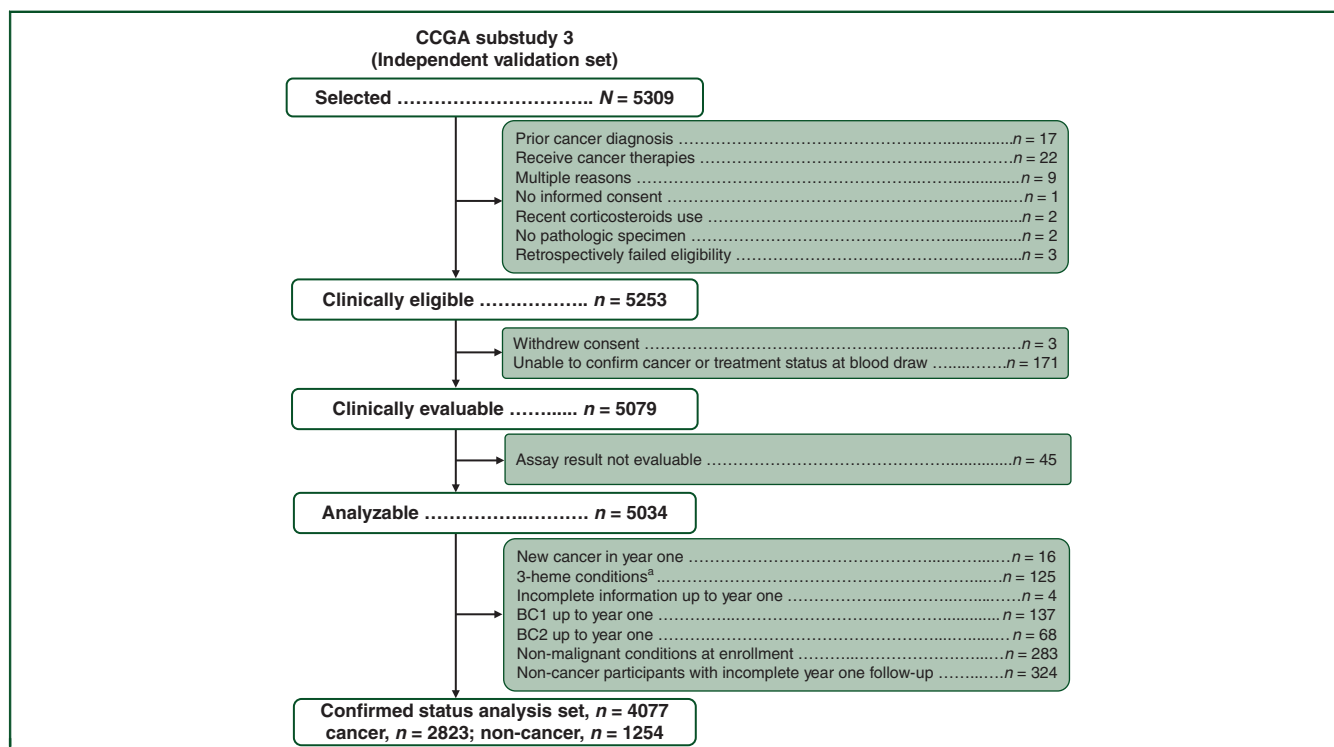

CCGA substudy 3
(Independent validation set)

Selected ………………………….. $N = 5309$

> Prior cancer diagnosis ……………………………………………...............$n = 17$
> Receive cancer therapies ………………………………………......…….....$n = 22$
> Multiple reasons …………………………………………………….................$n = 9$
> No informed consent …………………………………………………….........$n = 1$
> Recent corticosteroids use …………………………………………….........$n = 2$
> No pathologic specimen …………………………………………….............$n = 2$
> Retrospectively failed eligibility ………………………………….............$n = 3$

Clinically eligible ……………….. $n = 5253$

> Withdrew consent ………………………………………………….............$n = 3$
> Unable to confirm cancer or treatment status at blood draw …….…......$n = 171$

Clinically evaluable ………..…. $n = 5079$

> Assay result not evaluable ………………………………………….........$n = 45$

Analyzable ……………...……… $n = 5034$

> New cancer in year one ………………………………......…....…..$n = 16$
> 3-heme conditions[a] ………………………………………....……..$n = 125$
> Incomplete information up to year one …………………...…………….$n = 4$
> BC1 up to year one …………………………………………….…........$n = 137$
> BC2 up to year one ……………………………………………......….....$n = 68$
> Non-malignant conditions at enrollment ………....…………….............$n = 283$
> Non-cancer participants with incomplete year one follow-up …….....….$n = 324$

Confirmed status analysis set, $n = 4077$
cancer, $n = 2823$; non-cancer, $n = 1254$

Figure 2. Participant disposition.
The clinically eligible and evaluable dataset included participants meeting the study inclusion/exclusion and evaluability criteria; of these, participants with evaluable assay results from at least one plasma sample were included in the analyzable set. The Confirmed Status analysis set included participants who were analyzable and were confirmed as either a cancer participant or a non-cancer participant with non-cancer status confirmed at year-one follow-up. All data presented in this report are for the Confirmed Status analysis set, unless otherwise specified.
[a]3-Heme, selected malignant hematologic conditions representing proliferative or dysplastic disorders or neoplasms with BC3; BC1, Behavior Code 1 (borderline); BC2, Behavior Code 2 (in situ); BC3, Behavior Code 3 (malignant); CCGA, Circulating Cell-free Genome Atlas.

ARTICLE IN PRESS

E. A. Klein et al.   Annals of Oncology

Participant demographics and baseline characteristics were comparable between groups (Table 1), with expected differences in age group distribution between the cancer and non-cancer groups (i.e. there were more cancers than non-cancers in the older age groups). Mean (SD) age was 60.6 (12.4) years, 55.4% (2258/4077) were female (with a higher percentage in the non-cancer versus cancer group), and 81.2% (3312/4077) were classified as non-Hispanic white. In the cancer group, 54.9% (1552/2823) had stage I/II cancer.

Cancer signal detection

Specificity was 99.5% (95% CI: 99.0% to 99.8%; 1248/1254), indicating a low false-positive rate of 0.5%. Overall sensitivity across cancer classes and stages was 51.5% (49.6%-53.3%; 1453/2823) (Figure 3A). Sensitivity of cancer signal detection by cancer class is shown in Figure 3B. Sensitivity in the pre-specified group of 12 cancer classes was 76.3% (74.0%-78.5%) across all stages and was 67.6% (64.4%-70.6%) for stages I-III. Across cancer classes, sensitivity was 51.9% (50.0%-53.8%) for stages I-IV and was 40.7% (38.7%-42.9%) for stages I-III. As expected, sensitivity of cancer signal detection increased with increasing stage [stage I, 16.8% (14.5%-19.5%); stage II, 40.4% (36.8%-44.1%); stage III, 77.0% (73.4%-80.3%); stage IV, 90.1% (87.5%-92.2%)] (Table 2). Sensitivity of cancer signal detection by both individual cancer class and clinical stage is reported in Figure 3C (12 pre-specified cancers) and in Supplementary Table S5, available at https://doi.org/10.1016/j.annonc.2021.05.806.

Test positive rate by AJCC cancer type

Cancer types defined by the more granular AJCC criteria[43] were assigned to each cancer participant; because of the small sample sizes for certain cancer types, the test positive rate for each cancer type (not sensitivity) is reported. Participants with multiple primaries or unknown primary were excluded because there was insufficient information to assign an AJCC cancer type. The test positive rate across all AJCC cancer types and clinical stages was 51.0% (1420/2786; 49.1%-52.8%). Test positive rates by AJCC cancer type are shown in Supplementary Table S6, available at https://doi.org/10.1016/j.annonc.2021.05.806. Overall, cancer signals were detected across >50 AJCC cancer types.

CSO prediction

The overall accuracy of CSO prediction was 88.7% (87.0%-90.2%) in true positives (excluding participants with an unknown origin) based on the top CSO label prediction. Figure 4 demonstrates accuracy of the predicted CSO across individual CSOs. Sensitivity for combined cancer signal detection and CSO prediction is presented in Supplementary Table S7, available at https://doi.org/10.1016/j.annonc.2021.05.806.

Performance in subgroups

Specificity, sensitivity, and accuracy of CSO prediction showed similar results across age groups (<50 years, ≥50 years, ≥65 years; Supplementary Table S8, available at https://doi.org/10.1016/j.annonc.2021.05.806). When evaluating test performance in groups by method of cancer diagnosis, overall sensitivity in cancers identified by clinical presentation [63.9% (61.8%-66.0%)] was higher than that in cancers identified by screening tests [18.0% (15.5%-20.8%)], likely due to a preponderance of early-stage prostate and breast cancers in the screen-detected cancer classes. In a *post hoc* analysis that categorized cancers into three groups to better characterize findings in solid versus hematological cancers, overall sensitivity for solid tumors without common screening options was nearly twice that for solid tumors with common screening options, including breast, colorectal, cervical, and prostate [65.6% (876/1336; 63.0%-68.1%) and 33.7% (396/1175; 31.1%-36.5%)]. Overall sensitivity for hematologic malignancies was 55.1% (156/283; 49.3%-60.8%). Lung cancer is included in the category without common screening options because no broadly adopted guideline-recommended screening for the average risk population currently exists, and <13% of the 55-80-year-old population meet current U.S. Preventive Services Task Force high-risk criteria for lung cancer screening.[47]

Extrapolated positive and negative predictive values

To further understand performance in a potential screening population, PPV and NPV were extrapolated (adjusted to SEER cancer incidence and stage distribution in the 50-79 years age group). In this analysis, PPV for cancer signal detection was 44.4% (28.6%-79.9%), and NPV was 99.4% (99.4%-99.5%). To understand the value of the test in the event of accurate cancer signal detection but incorrect CSO prediction, residual PPV of incorrect predicted CSO (conditional probability given a positive test result but an incorrect top CSO prediction when one or two CSO predictions were generated) was found to be 8.0% (4.2%-29.7%); when implemented clinically, this could potentially be high enough to warrant a workup.

DISCUSSION

The ability to detect cancer at earlier stages has the potential to reduce cancer mortality.[37] The introduction of MCED tests together with current screening protocols on a population scale may increase the absolute number of cancers detected through screening and shift detection to earlier stages when outcomes are better and mortality is lower. In prior work that modeled potential stage shifts and mortality reductions, an earlier version of this test[30] was predicted to provide an absolute reduction of 104 deaths per 100 000 (26% of all cancer deaths).[37]

CCGA is a large-scale study (N = 15 254) that included systematic comparison of genomic methods to identify the best analytic approach and large-scale validation of a refined assay and classifiers to support the development of

Table 1. Participant demographics and baseline characteristics

	Cancer (n = 2823)	Non-cancer (n = 1254)	Total (N = 4077)
Age, years			
Mean (SD)	62.6 (11.76)	56.2 (12.63)	60.6 (12.39)
Median (Q1, Q3)	64.0 (56.0, 71.0)	57.0 (47.0, 65.0)	62.0 (52.0, 69.0)
Min, max	21, 85	21, 85	21, 85
Age group, n (%)			
<50 years	385 (13.6)	404 (32.2)	789 (19.4)
≥50 years	2438 (86.4)	850 (67.8)	3288 (80.6)
<65 years	1492 (52.9)	916 (73.0)	2408 (59.1)
≥65 years	1331 (47.1)	338 (27.0)	1669 (40.9)
Sex, n (%)			
Female	1394 (49.4)	864 (68.9)	2258 (55.4)
Male	1429 (50.6)	390 (31.1)	1819 (44.6)
Race-ethnicity, n (%)			
White, non-Hispanic	2316 (82.0)	996 (79.4)	3312 (81.2)
Black, non-Hispanic	193 (6.8)	85 (6.8)	278 (6.8)
Hispanic	192 (6.8)	103 (8.2)	295 (7.2)
Asian, Native Hawaiian or Pacific Islander	49 (1.7)	26 (2.1)	75 (1.8)
American Indian or Alaska native	8 (0.3)	7 (0.6)	15 (0.4)
Other	65 (2.3)	37 (3.0)	102 (2.5)
Smoking status, n (%)			
Never smoker	1322 (46.8)	719 (57.3)	2041 (50.1)
Former smoker	1080 (38.3)	402 (32.1)	1482 (36.4)
Current smoker	378 (13.4)	126 (10.0)	504 (12.4)
Missing	43 (1.5)	7 (0.6)	50 (1.2)
Family history of cancer, n (%)			
Yes	2254 (79.8)	1014 (80.9)	3268 (80.2)
No	493 (17.5)	210 (16.7)	703 (17.2)
Not sure	76 (2.7)	30 (2.4)	106 (2.6)
Immediate[a] family members' history of cancer, n (%)			
Yes	1792 (63.5)	736 (58.7)	2528 (62.0)
No	949 (33.6)	485 (38.7)	1434 (35.2)
Not sure	76 (2.7)	30 (2.4)	106 (2.6)
Missing	6 (0.2)	3 (0.2)	9 (0.2)
Extended[b] family members' history of cancer, n (%)			
Yes	1258 (44.6)	739 (58.9)	1997 (49.0)
No	1470 (52.1)	481 (38.4)	1951 (47.9)
Not sure	76 (2.7)	30 (2.4)	106 (2.6)
Missing	19 (0.7)	4 (0.3)	23 (0.6)
Body mass index (kg/m^2)			
n	2817	1249	4066
Mean (SD)	29.3 (6.83)	30.1 (7.24)	29.6 (6.96)
Median (Q1, Q3)	28.3 (24.4, 33.0)	28.7 (24.7, 34.3)	28.4 (24.5, 33.4)
Min, Max	12.8, 69.3	14.4, 79.2	12.8, 79.2
Body mass index category, n (%)			
Underweight	53 (1.9)	7 (0.6)	60 (1.5)
Normal	732 (25.9)	323 (25.8)	1055 (25.9)
Overweight	908 (32.2)	376 (30.0)	1284 (31.5)
Obese	1124 (39.8)	543 (43.3)	1667 (40.9)
Missing	6 (0.2)	5 (0.4)	11 (0.3)
Clinical cancer stage, n (%)			
I	849 (30.1)		849 (20.8)
II	703 (24.9)		703 (17.2)
III	566 (20.0)		566 (13.9)
IV	618 (21.9)		618 (15.2)
Not expected to be staged	67 (2.4)		67 (1.6)
Missing	20 (0.7)		20 (0.5)
Method of cancer diagnosis, n (%)			
Identified by clinical presentation	2036 (72.1)		2042 (50.1)
Identified by screening test[c]	787 (27.9)		787 (19.3)
Site region, n (%)			
Canada	41 (1.5)	10 (0.8)	51 (1.3)
Midwest	654 (23.2)	284 (22.6)	938 (23.0)
Northeast	254 (9.0)	175 (14.0)	429 (10.5)
South	1396 (49.5)	570 (45.5)	1966 (48.2)
West	478 (16.9)	215 (17.1)	693 (17.0)

Age was truncated at 85 years, per HIPPA policy.

CRF, case report form; HIPPA, Health Insurance Portability and Accountability Act; Q1, Q3, first quartile, third quartile; SD, standard deviation.

[a] Immediate family members included parents, siblings, and children.

[b] Extended family members included grandparents, aunts, uncles, nieces, and nephews per the Subject Questionnaire eCRF versions A-D. The extended family members included grandparents, aunts, uncles, nieces, nephews, and cousins of blood per the Subject Questionnaire eCRF version E.

[c] Cancers reported to be identified by a screening test included anus, breast, cervix, colon/rectum, lung, prostate, and uterus (anus and uterus cancers incidentally identified).
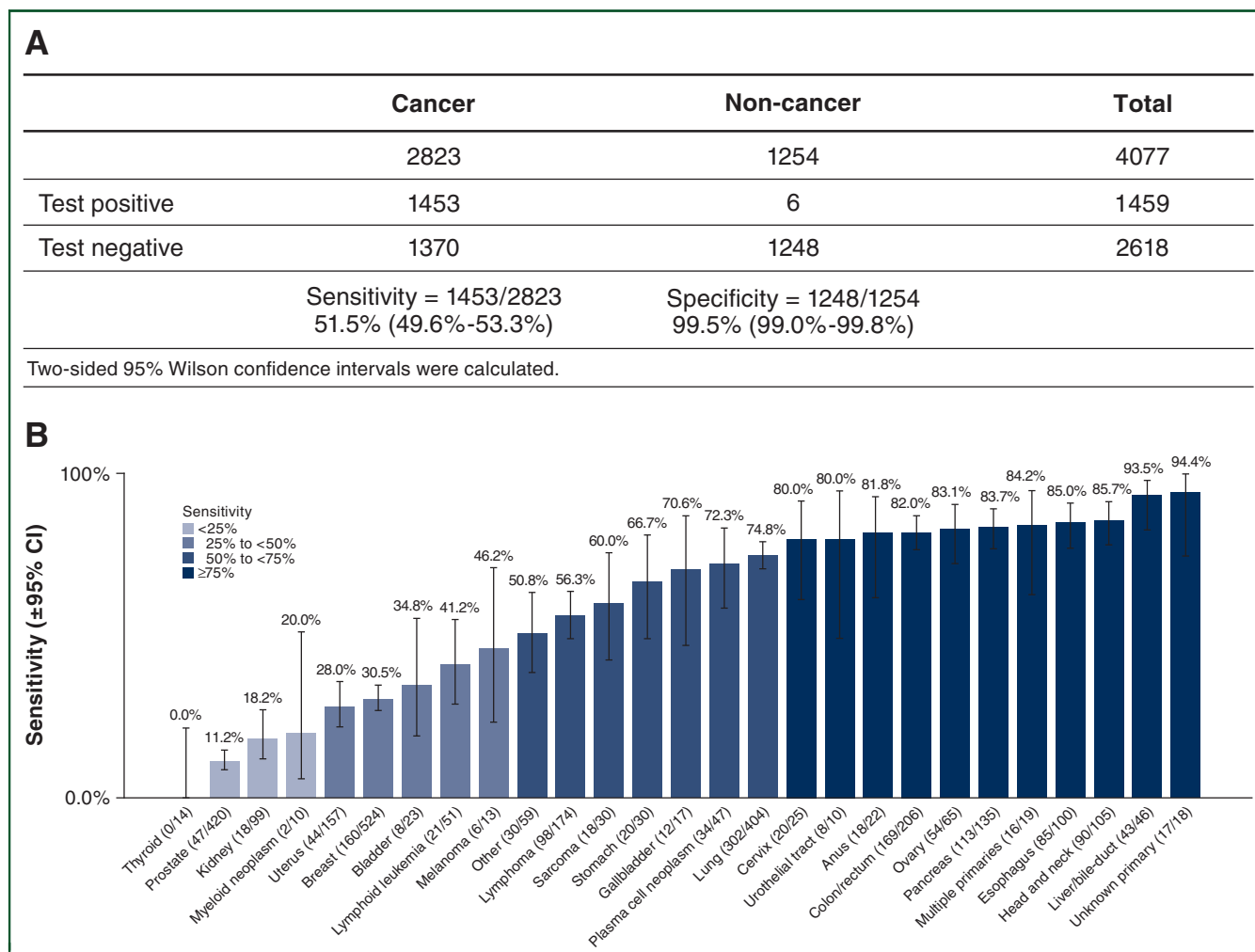
A

	Cancer	Non-cancer	Total
	2823	1254	4077
Test positive	1453	6	1459
Test negative	1370	1248	2618

Sensitivity = 1453/2823 51.5% (49.6%-53.3%)	Specificity = 1248/1254 99.5% (99.0%-99.8%)

Two-sided 95% Wilson confidence intervals were calculated.

B

Figure 3. MCED test performance for cancer signal detection (A) overall sensitivity and specificity, (B) sensitivity by cancer class, and (C) sensitivity by stage in 12 pre-specified cancers.
(A) The 2 × 2 contingency table summarizes overall sensitivity and specificity. (B) Sensitivity (y-axis) by cancer class based on individual cancer classes (x-axis), including other, unknown primary, and multiple primaries. Cancer classes are ordered based on increasing sensitivity; bars indicate 95% CI. (C) Sensitivity by stage is depicted in each box for each of the 12 pre-specified cancer classes; bars indicate 95% CI.
CI, confidence interval; MCED, multi-cancer early detection.

an MCED test (Galleri™). The low false-positive rate of 0.5% suggests that the test may limit additional harms due to unnecessary diagnostic workups when implemented clinically; this compares favorably to the false-positive rates associated with current recommended single-cancer screening tests (9.0%-14.5%), which generally optimize sensitivity over specificity.[26-28,48] By contrast, this MCED test was designed to maintain a high specificity while detecting common signals across many cancer types, allowing for an overall increase in the population cancer detection rate. The extrapolated PPV reported here based on SEER cancer incidence and clinical stage distribution was 44.4% in the screening-eligible 50-79-year age group, which is higher than that of currently recommended screening tests, as PPV is driven by specificity and population incidence.[26-28,48,49] Studies in intended-use populations that will provide more accurate PPV estimates are ongoing (STRIVE, NCT03085888; SUMMIT, NCT03934866; PATHFINDER, NCT04241796). It should be noted that the CCGA study includes 5-year longitudinal follow-up, and these data will provide additional insight into test performance.

A true 'multi-cancer' test should be able to detect cancer signals in as many cancer types as possible to maximize the population cancer detection rate (the fraction of cancers detected from the total expected cancers in the population). There are limitations to using sensitivity to measure performance of an MCED test in that the absolute number of cancers detected increases with each additional cancer class, even if the average sensitivity over all cancer classes decreases. In other words, overall sensitivity of 51.5% would represent more absolute cancer cases detected than the 76.3% sensitivity in the restricted pre-specified set of 12 cancer classes. These observations reinforce the limitation of the sensitivity metric, which may not reflect the total clinical utility of an MCED test. Thus, PPV may be a more clinically relevant metric.

Cancer incidence increases with age, so it is important to ensure consistent performance across age groups. In this
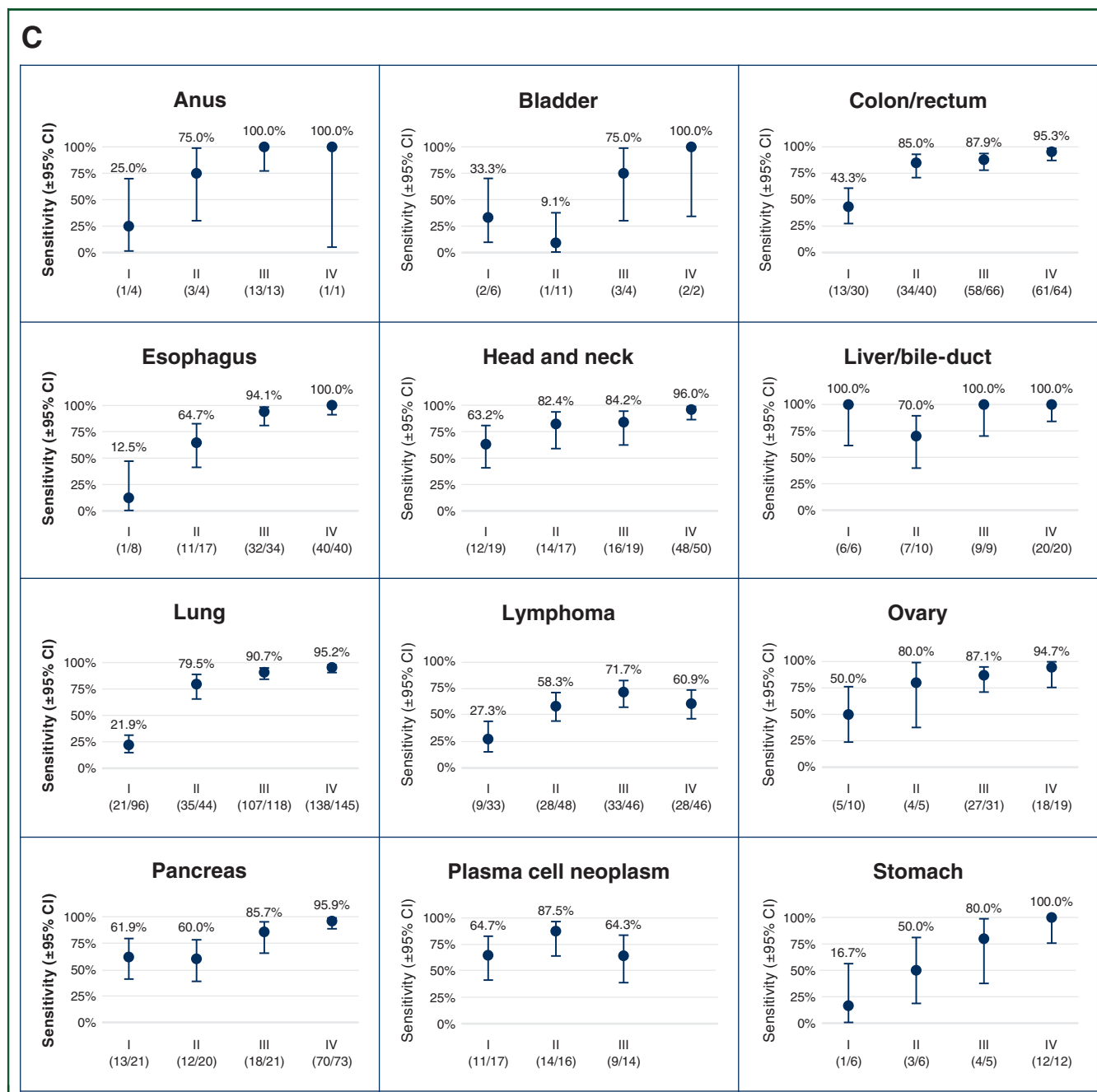
Figure 3. Continued.

substudy, test performance, including specificity, was similar across age groups, such that no particular age group would be expected to be at higher risk of additional harms associated with false positives.

This MCED test is meant to complement, and not replace, existing screening tests. Sensitivity of this test was higher for cancers identified by clinical presentation than by existing single-cancer screening tests; this observed lower sensitivity in screen-detected cancers was driven by a preponderance of early-stage breast and prostate cancers, both of which have a recommended screening test with proven survival benefit. The intended complementarity is also supported by the *post hoc* analysis showing a higher

sensitivity in solid tumors without common screening methods relative to solid tumors with common screening methods, including breast and prostate. For context, a reported 71% of cancer deaths in those aged 50-79 years are attributed to cancers without recommended screening options.[50] Lastly, this test may also provide a complementary screening option for individuals ineligible for or non-compliant with current screening tests, as well as for underserved communities with poor access to screening facilities.

Cancer classes that tend to be aggressive, such as pancreatic and esophageal cancers, were included in the pre-specified group of 12 cancers that contribute to a large

Table 2. Sensitivity of cancer signal detection by clinical stage

Clinical stage	Total N	Test positive	Sensitivity % (95% CI)[a]
All	2823	1453	51.5 (49.6% to 53.3%)
I	849	143	16.8 (14.5% to 19.5%)
II	703	284	40.4 (36.8% to 44.1%)
III	566	436	77.0 (73.4% to 80.3%)
IV	618	557	90.1 (87.5% to 92.2%)
I-II	1552	427	27.5 (25.3% to 29.8%)
I-III	2118	863	40.7 (38.7% to 42.9%)
I-IV	2736	1420	51.9 (50.0% to 53.8%)
III-IV	1184	993	83.9 (81.7% to 85.9%)
Not expected to be staged	67	23	34.3 (24.1% to 46.3%)
Missing	20	10	50.0 (29.9% to 70.1%)

CI, confidence interval.
[a] Two-sided 95% Wilson CIs were calculated.

proportion of cancer deaths and are more likely than others to shed more cfDNA into circulation.[42] Indeed, sensitivity in the pre-specified group of 12 cancer classes, most of which currently lack screening tests, was higher than that observed in all cancers. These 12 cancers account for ~two-thirds of US cancer deaths, underscoring the potential for this test to provide population-scale benefits.[40]

Conversely, some indolent cancers, like early-stage prostate cancer, shed less and are thus less detectable by this approach. Together, this suggests that this MCED test has the potential to minimize overdiagnosis.

As expected, accuracy of CSO prediction was slightly lower in this substudy compared with the second substudy (88.7% versus 93.3%),[30] in part because indeterminate CSO predictions were removed in the refined test. However, accuracy was still high, and the few incorrect CSO calls were often the result of a biological phenomenon that complicated CSO assignment [e.g. mismatched CSO predictions between human papillomavirus (HPV)-driven cancers, like cervical and anal]. Providing CSO predictions is intended to help health care providers define diagnostic workups after a positive test result. One exception is the 'Neuroendocrine Cells of Lung or Other Organs' CSO class, which may require a whole-body computed tomography (CT) or positron emission tomography (PET)-CT scan to localize the primary tumor. Importantly, the estimated likelihood of having cancer even with an unresolved positive signal (i.e. workup of the CSO did not confirm a cancer diagnosis) was still sufficiently high (8.0%, a PPV that is at least twice as high as
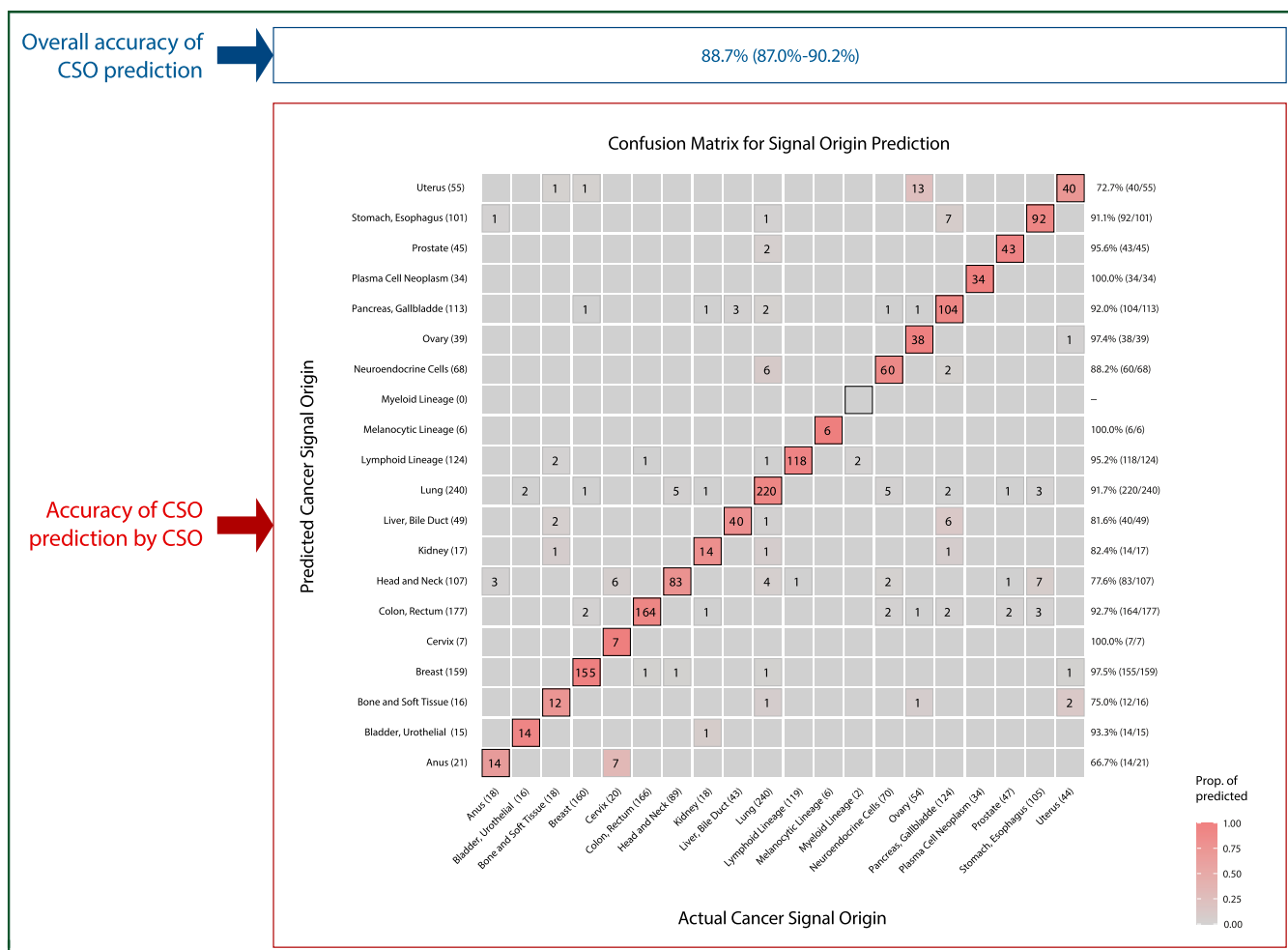

Figure 4. Accuracy of CSO prediction (confusion matrix).
The top panel indicates overall accuracy of CSO prediction. The bottom panel depicts a confusion matrix showing accuracy (top horizontal axis) and precision of CSO prediction by CSO (right vertical axis) among true positive participants with a known cancer signal origin. The proportion of each call is indicated by the strength of the colored signal within each individual box. Correct CSO calls are indicated on the diagonal.
CSO, cancer signal origin; Prop., proportion.

that reported for current screening modalities for breast and colorectal cancers)[13,26] to potentially warrant further diagnostic evaluation.

The assay and classifiers used here underwent multiple refinements for use as a multi-cancer screening tool. In previous work with an earlier version of this test,[30] a hematopoietic CSO was associated with a larger number of false positives. To mitigate this, the cancer detection classifier was adjusted depending on whether the CSO prediction was consistent with a solid tumor or a hematological cancer. This allowed for increased sensitivity for solid tumors while maintaining specificity and was supported by the *post hoc* analysis reported here separating solid cancers (with or without common screening options) and hematological cancers. Additionally, neuroendocrine tumors of all organs were associated with a common methylation pattern; as such, 'Neuroendocrine Cells of Lung or Other Organs' was created as a separate, distinct CSO label. Finally, improvements in classification removed the need for an indeterminate CSO category, meaning that all positive test results had an associated CSO prediction, which could help physicians direct diagnostic workups.

Limitations of this study include that the blood samples collected from participants with cancer after biopsies had been carried out could increase the possibility that the tumor cfDNA fraction may increase relative to before the biopsy. Another limitation is that CCGA is a case-control study, and as such, is not reflective of performance in a screening population; a larger clinical development program that includes other studies evaluating test performance and/or clinical utility in target-use populations (STRIVE, NCT03085888; SUMMIT, NCT03934866; PATHFINDER, NCT04241796) is underway.[51] Importantly, an ongoing interventional return-of-results study (PATHFINDER, NCT04241796) will also assess clinical implementation (e.g. time to diagnostic resolution) as well as safety.

Conclusion

Taken together, these results demonstrate that this targeted methylation-based MCED test has high specificity that is generalizable across study populations, detects cancer signals across a broad range of cancer types with diverse biologic features (including those that currently lack screening tests), and provides accurate CSO prediction that may inform patient management. These results support that this blood-based MCED test may complement existing single-cancer screening tests and result in reduced cancer mortality.

DISCLOSURE

EAK is a consultant for GRAIL, Inc.; DC is a consultant for Pfizer and Merck; AJ and KNK are full-time employees of GRAIL, Inc. and have stock in Illumina and GRAIL, Inc.; JG and NH are full-time employees of GRAIL Inc. and own stock in GRAIL, Inc. MVS has stock in McKesson Corporation, is a clinical adviser for GRAIL, Inc. and a director of Next Oncology and Nemucore Medical Innovations; CS has stock in GRAIL, Inc., Epic Biosciences and Apogen Biotech, grants from Pfizer and AstraZeneca, received honoraria or consultant fees from Roche Ventana, Celgene, Pfizer, Novartis, Genentech, and BMS, and is a co-founder of Achilles Therapeutics; MCL participated as an advisory board member for GRAIL, Inc. All other authors have declared no conflicts of interest.

REFERENCES

1. Dagenais GR, Leong DP, Rangarajan S, et al. Variations in common diseases, hospital admissions, and deaths in middle-aged adults in 21 countries from five continents (PURE): a prospective cohort study. *Lancet*. 2020;395:785-794.
2. Raza A. *The First Cell: And the Human Costs of Pursuing Cancer to the Last*. New York, NY: Basic Books; 2019.
3. Ofman JJ, Raza A. Taking early cancer detection to the next level. *New York, NY: Scientific American*; 2020.
4. Ofman JJ, Raza A. The cancer detection rate—a public health approach to early detection. *Cancer Lett*. 2020;46:48-50.
5. Ahlquist DA. Universal cancer screening: revolutionary, rational, and realizable. *Npj Precis Oncol*. 2018;2:23.
6. Etzioni R, Urban N, Ramsey S, et al. The case for early detection. *Nat Rev Cancer*. 2003;3:243-252.

7. Clarke CA, Hubbell E, Kurian AW, Colditz GA, Hartman A-R, Gomez SL. Projected reductions in absolute cancer-related deaths from diagnosing cancers before metastasis, 2006—2015. *Cancer Epidemiol Biomarkers Prev*. 2020;29:895-902.

8. Hawkes N. Cancer survival data emphasise importance of early diagnosis. *Br Med J*. 2019;364:l408.

9. Cancer Research UK. Why is early diagnosis important? 2021. Available at: https://www.cancerresearchuk.org/about-cancer/cancer-symptoms/why-is-early-diagnosis-important. Accessed March 22, 2021.

10. National Cancer Institute. Research Areas: Detection and Diagnosis. 2020. Available at: https://www.cancer.gov/research/areas/diagnosis. Accessed March 22, 2021.

11. The ASCO Post. World Cancer Day 2019: Emphasis on Early Detection. 2019. Available at: https://ascopost.com/News/59711. Accessed March 2, 2021.

12. Curry SJ, Krist AH, Owens DK, et al. Screening for cervical cancer: US preventive services task force recommendation statement. *J Am Med Assoc*. 2018;320:674-686.

13. Bibbins-Domingo K, Grossman DC, Curry SJ, et al. Screening for colorectal cancer: US preventive services task force recommendation statement. *J Am Med Assoc*. 2016;315:2564-2575.

14. Grossman DC, Curry SJ, Owens DK, et al. Screening for prostate cancer: US preventive services task force recommendation statement. *J Am Med Assoc*. 2018;319:1901-1913.

15. Siu AL on behalf of the U.S. Preventive Services Task Force. Screening for breast cancer: U.S. preventive services task force recommendation statement. *Ann Intern Med*. 2016;164:279-296.

16. Moyer VA, LeFevre ML, Siu AL, et al. Screening for lung cancer: U.S. preventive services task force recommendation statement. *Ann Intern Med*. 2014;160:330-338.

17. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin*. 2018;68:7-30.

18. Black WC. Overdiagnosis: an underrecognized cause of confusion and harm in cancer screening. *J Natl Cancer Inst*. 2000;92:1280-1282.

19. Srivastava S, Koay EJ, Borowsky AD, et al. Cancer overdiagnosis: a biological challenge and clinical dilemma. *Nat Rev Cancer*. 2019;19:349-358.

20. Smith RA, Andrews KS, Brooks D, et al. Cancer screening in the United States, 2019: a review of current American Cancer Society guidelines and current issues in cancer screening. *CA Cancer J Clin*. 2019;69:184-210.

21. Hall IJ, Tangka FKL, Sabatino SA, Thompson TD, Graubard BI, Breen N. Patterns and trends in cancer screening in the United States. *Prev Chronic Dis*. 2018;15:170465.

22. Narayan A, Fischer A, Zhang Z, Woods R, Morris E, Harvey S. Nationwide cross-sectional adherence to mammography screening guidelines: national behavioral risk factor surveillance system survey results. *Breast Cancer Res Treat*. 2017;164:719-725.

23. Limmer K, LoBiondo-Wood G, Dains J. Predictors of cervical cancer screening adherence in the United States: a systematic review. *J Adv Pract Oncol*. 2014;5:31-41.

24. Daskalakis C, DiCarlo M, Hegarty S, Gudur A, Vernon SW, Myers RE. Predictors of overall and test-specific colorectal cancer screening adherence. *Prev Med*. 2020;133:106022.

25. Cossu G, Saba L, Minerba L, Mascalchi M. Colorectal cancer screening: the role of psychological, social and background factors in decision-making process. *Clin Pract Epidemiol Ment Health*. 2018;14:63-69.

26. Lehman CD, Arao RF, Sprague BL, et al. National performance benchmarks for modern screening digital mammography: update from the Breast Cancer Surveillance Consortium. *Radiology*. 2017;283:49-58.

27. Pinsky PF, Gierada DS, Black W, et al. Performance of Lung-RADS in the National Lung Screening Trial: a retrospective assessment. *Ann Intern Med*. 2015;162:485-491.

28. Kim JJ, Burger EA, Regan C, Sy S. Screening for cervical cancer in primary care: a decision analysis for the US Preventive Services Task Force. *J Am Med Assoc*. 2018;320:706-714.

29. Croswell JM, Kramer BS, Kreimer AR, et al. Cumulative incidence of false-positive results in repeated, multimodal cancer screening. *Ann Fam Med*. 2009;7:212-222.

30. Liu MC, Oxnard GR, Klein EA, Swanton C, Seiden MV. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. *Ann Oncol*. 2020;31:745-759.

31. Li B, Wang C, Xu J, et al. Abstract A06: Multiplatform analysis of early-stage cancer signatures in blood. *Clin Cancer Res*. 2020;26:A06.

32. Shen SY, Singhania R, Fehringer G, et al. Sensitive tumour detection and classification using plasma cell-free DNA methylomes. *Nature*. 2018;563:579-583.

33. Chen X, Gole J, Gore A, et al. Non-invasive early detection of cancer four years before conventional diagnosis using a blood test. *Nat Commun*. 2020;11:3475.

34. Cohen JD, Li L, Wang Y, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science*. 2018;359:926-930.

35. Lennon AM, Buchanan AH, Kinde I, et al. Feasibility of blood testing combined with PET-CT to screen for cancer and guide intervention. *Science*. 2020;369:eabb9601.

36. Cristiano S, Leal A, Phallen J, et al. Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature*. 2019;570:385-389.

37. Hubbell E, Clarke CA, Aravanis AM, Berg CD. Modeled reductions in late-stage cancer with a multi-cancer early detection test. *Cancer Epidemiol Biomarkers Prev*. 2021;30:460-468.

38. Klein EA, Hubbell E, Maddala T, et al. Development of a comprehensive cell-free DNA (cfDNA) assay for early detection of multiple tumor types: The Circulating Cell-free Genome Atlas (CCGA) study. *J Clin Oncol*. 2018;36:12021.

39. Liu MC, Klein E, Hubbell E, et al. Plasma cell-free DNA (cfDNA) assays for early multi-cancer detection: the Circulating Cell-free Genome Atlas (CCGA) study. *Ann Oncol*. 2018;29:500.

40. American Cancer Society. Cancer Facts & Figures. 2021. Available at: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2021/cancer-facts-and-figures-2021.pdf. Accessed March 10, 2021.

41. Liu MC, Jamshidi A, Venn O, et al. Genome-wide cell-free DNA (cfDNA) methylation signatures and effect on tissue of origin (TOO) performance. *J Clin Oncol*. 2019;37:3049.

42. Bettegowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med*. 2014;6:224ra24.

43. Amin MB, Greene FL, Edge SB, et al. The Eighth Edition AJCC Cancer Staging Manual: continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin*. 2017;67:93-99.

44. World Health Organization. International Classification of Diseases for Oncology, 3rd Edition (ICD-O-3). 2021. Available at: https://www.who.int/standards/classifications/other-classifications/international-classification-of-diseases-for-oncology. Accessed March 1, 2021.

45. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol*. 2010;17:1471-1474.

46. R. The R Project for Statistical Computing. Available at: https://www.r-project.org/. Accessed December 9, 2020.

47. Richards TB, Soman A, Thomas CC, et al. Screening for Lung Cancer—10 States, 2017. *MMWR Morb Mortal Wkly Rep*. 2020;69:201-206.

48. US Food and Drug Administration. FDA Summary of Safety and Effectiveness Data. PMA P130017. Silver Spring, MD: US Dept of Health and Human Services; 2014.

49. Wolf AMD, Wender RC, Etzioni RB, et al. American Cancer Society guideline for the early detection of prostate cancer: update 2010. *CA Cancer J Clin*. 2010;60:70-98.

50. American Cancer Society. Cancer Facts & Figures 2020. Atlanta: American Cancer Society; 2020. Accessed September 8, 2020.

51. GRAIL, Inc. GRAIL and UK Government to Make Galleri Multi-Cancer Early Detection Blood Test Available to Patients. 2020. Available at: https://grail.com/press-releases/grail-and-uk-government-to-make-galleri-multi-cancer-early-detection-blood-test-available-to-patients/. Accessed February 11, 2021.