Session 2
MCD Test Validation

Opportunities and Challenges for the Development and Adoption of Multicancer Detection Tests: A Workshop

National Cancer Policy Forum

October 28-29, 2024

Session Objectives

PERFORMANCE

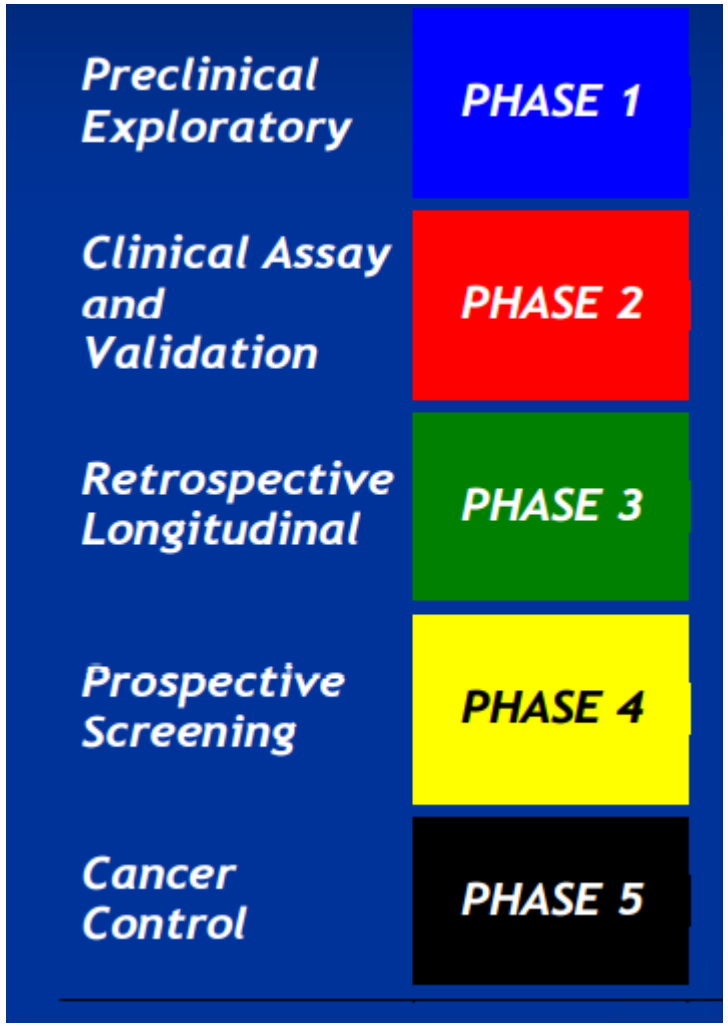
Sensitivity
Specificity
PPV
NPV

OUTCOMES

Cancer incidence
Late-stage incidence
Cancer mortality
Overdiagnosis

Focusing on the translation from performance to outcomes and how we can validate that MCD screening tests demonstrate adequate performance (clinical validity) and impact on clinical outcomes (clinical utility) to support their adoption for population screening

EDRN blueprint for biomarker-based tests



NIH NATIONAL CANCER INSTITUTE
Early Detection Research Network

NCI Cancer Screening Research Network
A program of the National Cancer Institute of the National Institutes of Health

The gold standard for the evaluation of a new screening modality is a randomized clinical trial. The CSRN is expected to conduct a variety of randomized control trials and other studies related to cancer screening.

Session outline

Ruth Etzioni

- Fred Hutch Cancer Center — *Diagnostic performance: the many faces of sensitivity*

Hormuzd Katki

- National Cancer Institute — *Clinical Trials for Evaluating the Mortality Benefits of MCD Testing*

- Hilary Robins
  - IARC — *Alternative Endpoints for MCD Clinical Trials*

Jane Lange

- Knight Cancer Institute/OHSU — *Addressing the Gap between Performance and Outcomes in the Absence of MCD Trials*

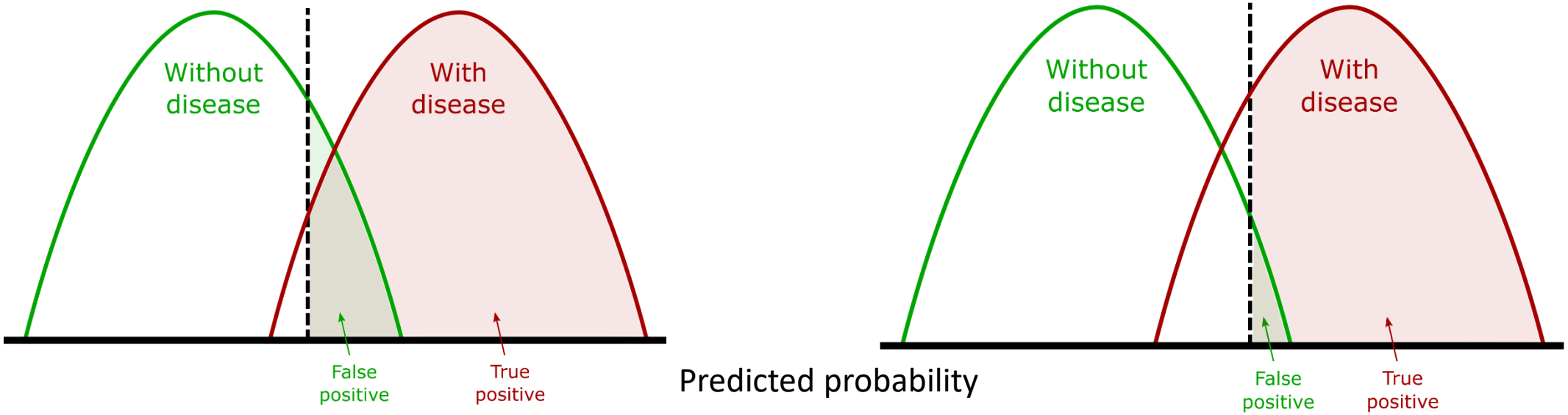
Diagnostic Performance
The Many Faces of Screening Test Sensitivity

Ruth Etzioni PhD
Rosalie and Harold Rea Brown Chair
Fred Hutch Cancer Center

Fred Hutch Cancer Center

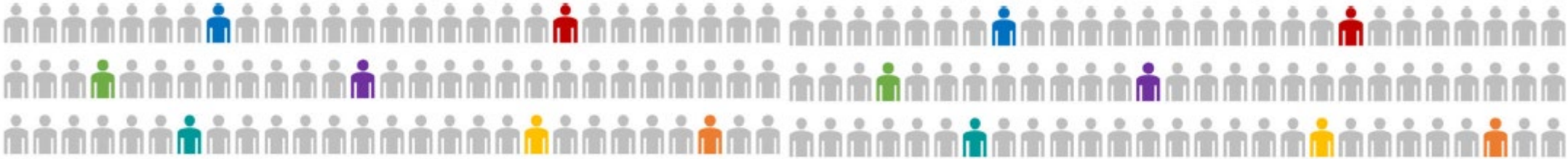
Algorithms for cancer screening

- Produce a score or **predicted probability** that a patient has cancer
- Define a **threshold** above which test is declared positive
  - A high threshold will limit false positives ➡ high specificity
  - But will reduce true positives as well ➡ low sensitivity



Most first-line tests are highly specific
High threshold for calling a positive

Because only a small minority of the population has cancer at any time



- Must limit false positive rate so that the number of false positives is not too high ➡ this would generate too many unnecessary biopsies
- Want to “believe the positive”
- PPV is the chance your positive test is a true positive

MCED screening tests tend to have high PPV

Inherited from their high specificity

MCED TEST DETECTED BROAD RANGE OF CANCER SIGNALS, INFORMING DIAGNOSTIC WORKUP, WITH 45% PPV

Diagnostic Workup

Table 1. Diagnostic Workup Procedures Per Participant with Diagnostic Resolution

	Median (Q1, Q3)		
	True Positive (n = 27[a])	False Positive (n = 36)	Total (n = 63)[a]
All Imaging Tests/Invasive Procedures	2.0 (1.5, 3.0)	1.5 (1.0, 2.2)	2.0 (1.0, 3.0)
All Imaging Tests	1.0 (1.0, 1.5)	1.0 (1.0, 2.0)	1.0 (1.0, 2.0)
Functional[b]	1.0 (0, 1.0)	1.0 (0, 1.0)	1.0 (0, 1.0)
Anatomic[c]	1.0 (0, 1.0)	1.0 (0, 1.0)	1.0 (0, 1.0)
All Invasive Procedures	1.0 (1.0, 1.0)	0 (0, 0.2)	0 (0, 1.0)
Minimally Invasive[d]	1.0 (0.5, 1.0)	0	0 (0, 1.0)
Surgical[e]	0	0	0
Clinical Lab Tests	3.0 (1.0, 5.5)	3.0 (1.0, 6.0)	3.0 (1.0, 6.0)
Days to Diagnostic Resolution	50.0 (27.0, 76.5)	49.0 (30.2, 153.8)	50.0 (28.0, 91.0)

Abbreviations: CT, computerized tomography; MRI, magnetic resonance imaging; PET, positron emission tomography; Q1, first quartile; Q3, third quartile.
[a] participants with 'signal detected' MCED test result (true positives) were excluded from the diagnostic workup analysis because diagnostic testing was initiated before MCED test results were returned.
[b] Functional imaging includes PET-CT, PET-MRI, bone scan.
[c] Anatomic imaging includes CT, MRI, ultrasound, mammography, plain film X-ray (including skeletal survey).
[d] Minimally invasive procedures include esophagogastroduodenoscopy, colonoscopy, endoscopic ultrasound, endoscopic retrograde cholangiopancreatography, bronchoscopy, cystoscopy, hysteroscopy, fine needle aspiration of the thyroid gland, liver biopsy, thoracentesis, pulmonary arterio-venous malformation embolization.
[e] 4 surgical procedures were performed, 3 in true positive, 1 in false positive set.

- Most participants with diagnostic resolution (57/63, 90.5%) had at least 1 imaging test
  - Median number of imaging tests per participant was the same in true and false positive groups (**Table 1**)
- Most invasive procedures were minimally invasive (28/32 procedures, 87.5%)
  - 26/30 (86.6%) participants had only minimally invasive procedures
- There were 4 reports of study-related adverse events: 3 of anxiety and 1 bruise at venipuncture site, all were of mild severity

Test Performance

Table 2. MCED Test Performance

	≥50 y With Additional Risk	≥50 y Without Additional Risk	Total
Cancer Signal Detection, No.	n=3695	n=2934	N=6629
Detected, No. (%)	56 (1.5)	36 (1.2)	92 (1.4)
True Positive	20 (0.5)	9 (0.3)	29 (0.4)
False Positive	15 (0.4)	21 (0.7)	36 (0.5)
No Current Diagnostic Resolution	21 (0.6)	6 (0.2)	27 (0.4)
Not Detected	3639 (98.5)	2898 (98.8)	6537 (98.6)
PPV for Cancer Signal Detection, No.	n=35	n=30	n=65
% (95% CI)	57.1 (40.9-72.0)	30.0 (16.7-47.9)	44.6 (33.2-56.7)
CSO Prediction Accuracy, No.	n=19[a]	n=8[a]	n=27[a]
First CSO,[b] % (95% CI)	84.2 (62.4-94.5)	87.5 (52.9-99.4)	85.2 (67.5-94.1)
First or Second CSO,[c] % (95% CI)	100 (83.2-100)	87.5 (52.9-99.4)	96.3 (81.7-99.8)

Abbreviations: CI, confidence interval; CSO, cancer signal origin; PPV, positive predictive value.
[a] Excludes 1 participant with unknown cancer type and 1 with indeterminate CSO from the true positive set.
[b] Proportion of correctly predicted first CSO among true positive participants with determinate CSO.
[c] Proportion of correctly predicted first or second CSO among true positive participants with determinate CSO.

- Of the 6629 analyzable participants, the MCED test detected cancer signal in 92 (1.4%); 1.5% of participants with additional risk and 1.2% without (**Table 2**)
- The PPV of the MCED test for participants with cancer signal detected who achieved diagnostic resolution was 44.6% (**Table 2**)
  - PPV was 57.1% in the "additional risk" vs 30.0% in the "without additional risk" cohort

Table 3. Cancer Stage at Diagnosis Following a Positive MCED Result (n=28[a] True Positives)

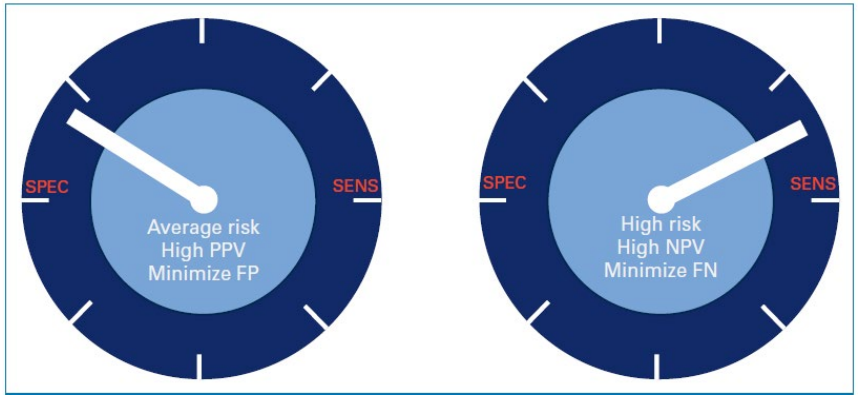
Cancer Type Diagnosed	Clinical AJCC Stage[b] of New Cancers					Extent of Recurrent Cancers		First Predicted Cancer Signal Origin
	I	II	III	IV	Other	Local	Distant	
Colon or rectum				1	1 Unknown[c]			Upper GI Tract (stage IV pt) Colon/Rectum (unk pt)
Head and Neck		1		1				Head and Neck
Liver, bile duct	1		1					Liver, bile-duct
Lung			1					Lung
Lymphoid leukemia					2 NA[d]			Lymphoid Neoplasm
Lymphoma	2	3	1	2				Lymphoid Neoplasm
Ovary, peritoneum or fallopian tube			1					Uterus (ovary second CSO)
Pancreas		1						Pancreas/Gallbladder
Plasma cell neoplasm					1 NA[d]			Plasma Cell Neoplasm
Prostate				1				Indeterminate
Small intestine	1							Colon/Rectum (upper GI second CSO)
Waldenstrom macroglobulinemia					1 NA[d]			Lymphoid Neoplasm
Breast							4	3 cases Breast 1 case Breast (first CSO), lymphoid (second)
Prostate						1		Lymphoid (first CSO), prostate (second)
Total	4	5	4	5	5	1	4	

Abbreviations: AJCC, the American Joint Committee on Cancer; CSO, cancer signal origin; GI, gastrointestinal.
[a] Information not available for cancer type/stage/recurrence for one true positive participant at time of analysis.
[b] AJCC version 8.
[c] Unknown stage at time of analysis.
[d] No AJCC stage expected.

Turning the Knobs on Screening Liquid Biopsies for High-Risk Populations: Potential for Dialing Down Invasive Procedures

Sana Raoof, MD, PhD[1] and Razelle Kurzrock, MD[2]

JCO 2024



In the case of published MCEDs, the threshold has been chosen to produce high specificity—in excess of 99%. High test specificity has always been important for producing screening tests with high PPV and limiting unnecessary medical workup in average-risk populations, but often comes at the cost of lower sensitivity

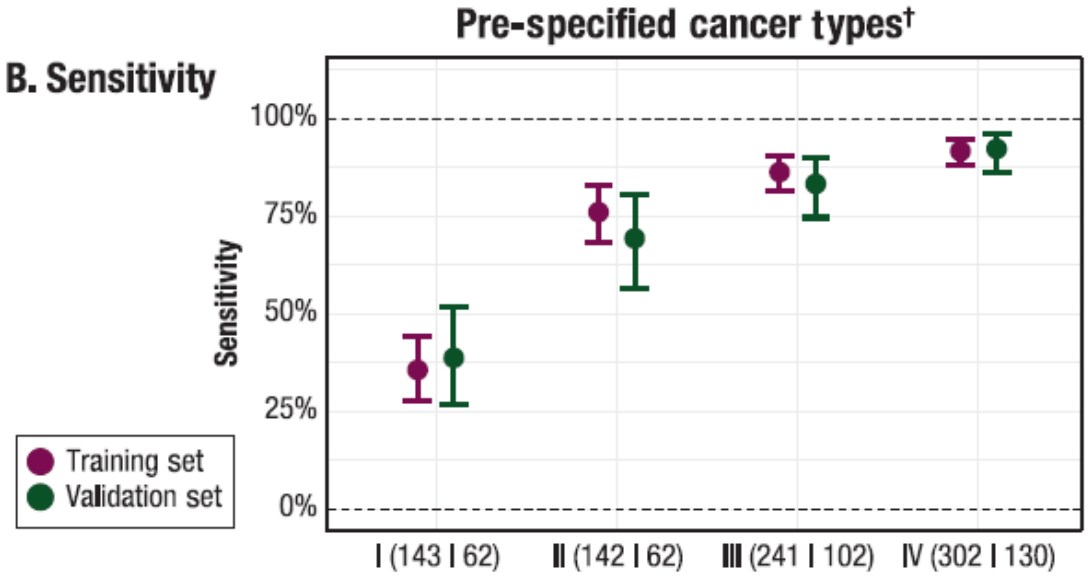
Sensitivity: likelihood that a test conducted in someone who has cancer yields a positive result

Different versions of sensitivity

Clinical sensitivity	• Sensitivity to detect disease in known or clinically diagnosed cases
Preclinical sensitivity	• Sensitivity to detect disease in preclinical cases before time of clinical diagnosis: hard to measure in MCED
Empirical sensitivity	• A version of sensitivity typically reported in prospective screening studies or cohorts

*Here clinically diagnosed means diagnosed in the absence of the biomarker-based test

Clinical sensitivity in early MCED studies



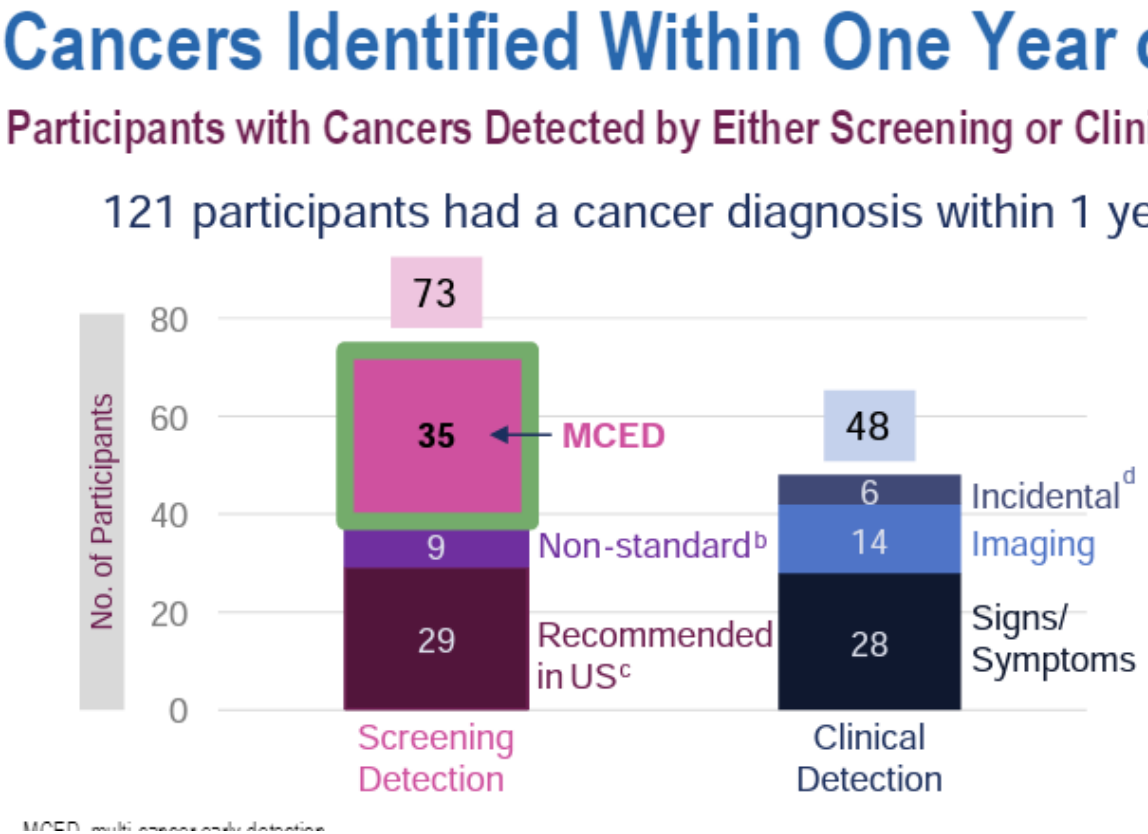
Sensitivity by stage

Overall 67.3% for 12 cancers

- Fraction of cases diagnosed for whom MCED test is positive
- These cases
  - Presented without the test
  - May be late in natural history
  - Stage distribution by convenience
  - Cancer mix does not reflect prevalence in population
- Expect clinical sensitivity to overestimate preclinical sensitivity

Liu et al Annals of Oncology 2020 for Grail test

Empirical sensitivity



- Out of 121 cancers diagnosed within a year 35 were detected by MCED screening
- Empirical sensitivity

29% = 35/121

Schrag et al ESMO 2022 and The Lancet 2023

Empirical sensitivity is commonly used

$$\frac{\#\ \text{screen detected}}{\#\ \text{screen detected} + \#\ \text{interval detected}}$$

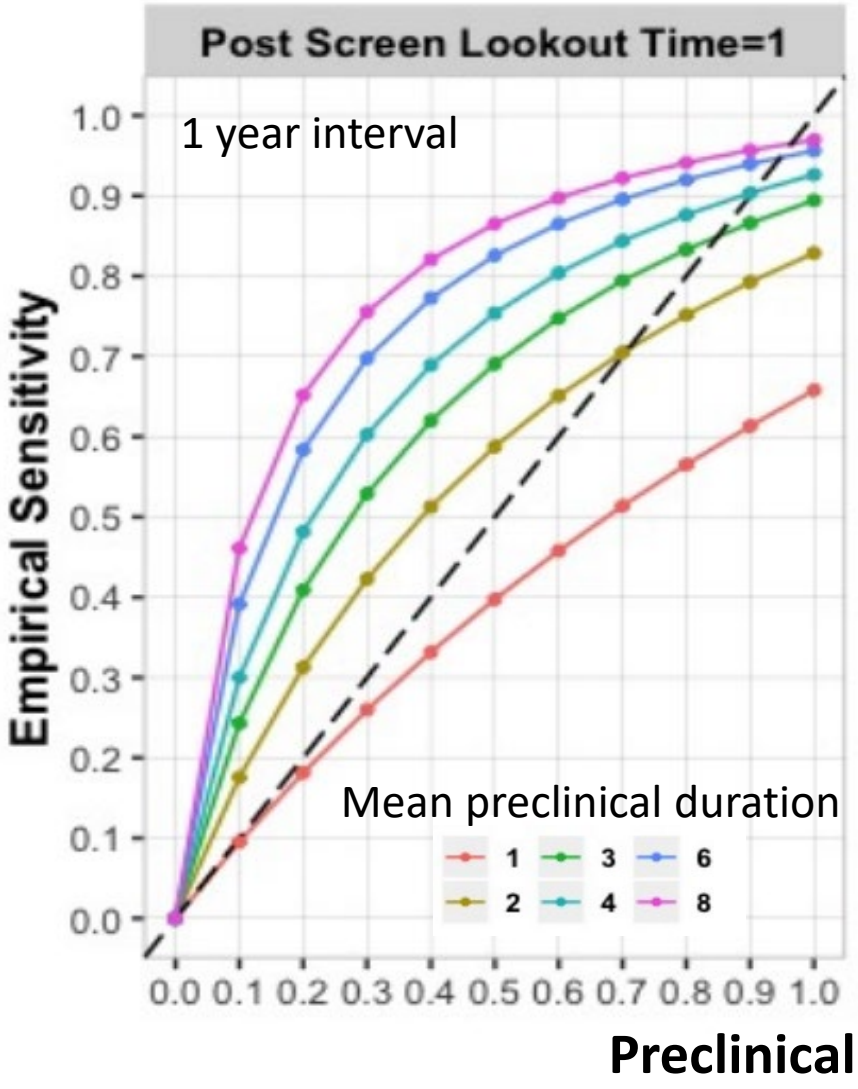
- Empirical sensitivity takes the cancers diagnosed within a year as proxy for cancers present at test
- May be a biased estimate of true preclinical sensitivity

BCSC

Based on BCSC data through 2013

	Number of Screening Exams
Sensitivity*	86.9%
True positives†	8,529
Cancers‡	9,812
Specificity§	88.9%
True negatives‖	1,486,553

Bias of empirical sensitivity depends on the cancer



Empirical sensitivity optimistic when

- Preclinical sensitivity is modest

  and

- Mean preclinical duration is long relative to the interval

Lange JA et al SMMR 2023

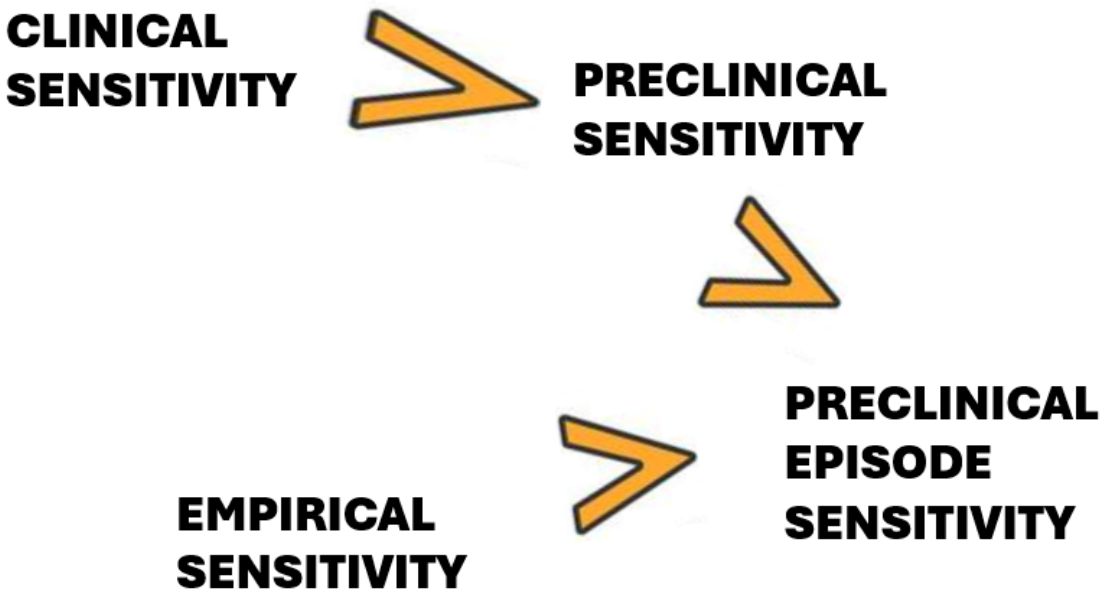
Preclinical test sensitivity versus preclinical episode sensitivity

Test sensitivity	Likelihood test returns a positive result if cancer is present
Episode sensitivity	Likelihood the testing episode returns a positive result i.e. detects cancer if it is present. Depends on gold standard testing

Empirical sensitivity is an episode estimate

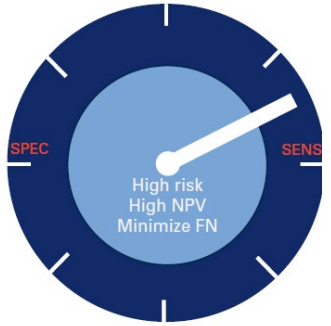
Summary so far

1. Biomarker-based tests including MCED
   - Performance determined by design
2. High PPV is due to the test being conservative
   - Inherited from high specificity
3. Preclinical sensitivity in first-line screening tests may be modest
   - Proxies and estimates may be optimistic for some cancers



Interpreting MCD tests when sensitivity is modest

1. The chance that cancer is there given a negative test is similar to the chance the cancer was there before the test
2. Can’t “believe the negative”
3. In some settings we might want to rule out cancer and believe the negative
4. In some of these settings the risk of having cancer is higher so can reduce specificity and increase sensitivity



- Individuals with symptoms
- In lieu of challenging or low-sensitivity biopsies
  - Ovarian cancer
  - Lung cancer

Take-home messages

- MCED screening tests and biomarker-based tests
  - Sensitivity is determined by test design and can be adjusted
- Many different versions of sensitivity not all discussed here
  - Let’s use different terms for different versions!
- Early-stage sensitivity is key for first-line screening
  - May not be degraded as much as overall sensitivity in prospective studies
- Other compelling use cases for MCED and other liquid-biopsy tests
  - Require thresholding tests differently so can believe the negative

Thank you!
Collaborators and support

- Yibai Zhao
- Jane Lange (OHSU)
- Roman Gulati
- Ziding Feng
- Yingye Zheng
- Sana Raoof (MSKCC)
- Sudhir Srivastava (NCI)
- Stuart Baker (NCI)

Rosalie and Harold Rea Brown chair at Fred Hutch

NCI EDRN DMCC

NCI R35 on Modeling and Analytics for Novel Cancer Diagnostics

NCI Cancer Screening Research Network

retzioni@fredhutch.org
https://research.fredhutch.org/etzioni/en.html

Fred Hutch Cancer Center

NCI Cancer Screening Research Network

NATIONAL CANCER INSTITUTE
Early Detection Research Network

NIH NATIONAL CANCER INSTITUTE