PROSPR

Population-based Research to Optimize the Screening PRocess

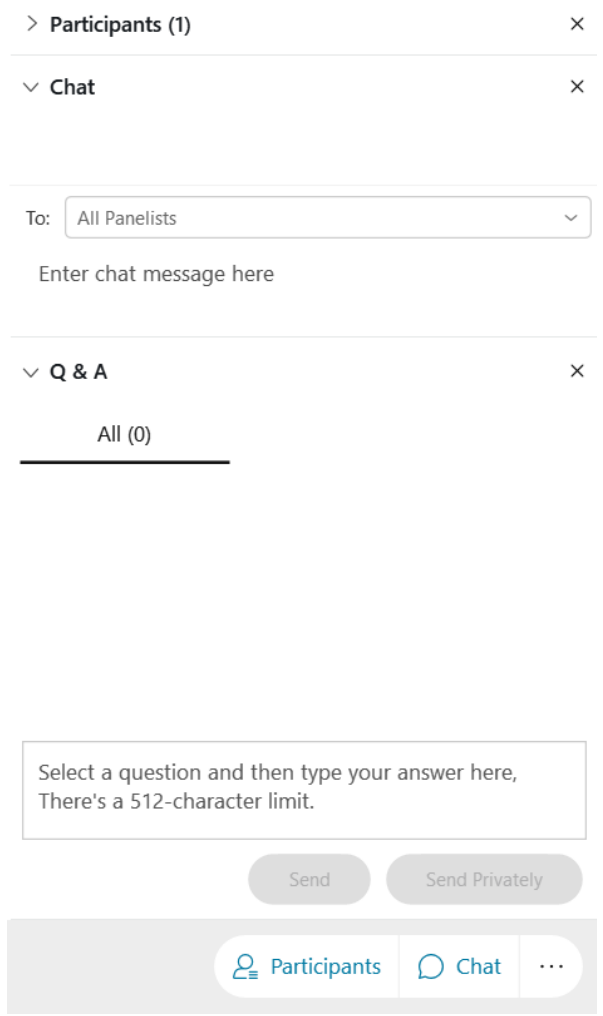
Population-based Research to Optimize the Screening Process (PROSPR) Initiative: Research Activities and Data Sharing

Pamela Marcus, PhD, MS, ELS

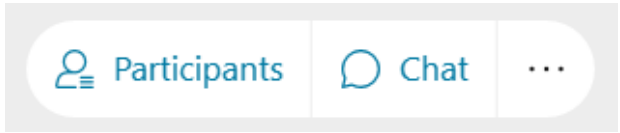
Paul Doria-Rose, PhD

September 20, 2022

Using WebEx and webinar logistics



- All lines will be in listen-only mode
- Submit questions at any time using the Q&A or Chat Panel and select *All Panelists*
- You may need to activate the appropriate box using the floating navigation panel found at the bottom of your screen



- This webinar is being recorded

Speakers:

Pamela Marcus, PhD, MS, ELS
Senior Scientist and Program Director
Healthcare Assessment Research Branch
marcusp@mail.nih.gov

Paul Doria-Rose, PhD
Branch Chief
Healthcare Assessment Research Branch
doriarop@mail.nih.gov

The PROSPR Initiative: Research Activities and Data Sharing

Pamela Marcus and V. Paul Doria-Rose

September 20, 2022

NIH NATIONAL CANCER INSTITUTE

Webinar Overview

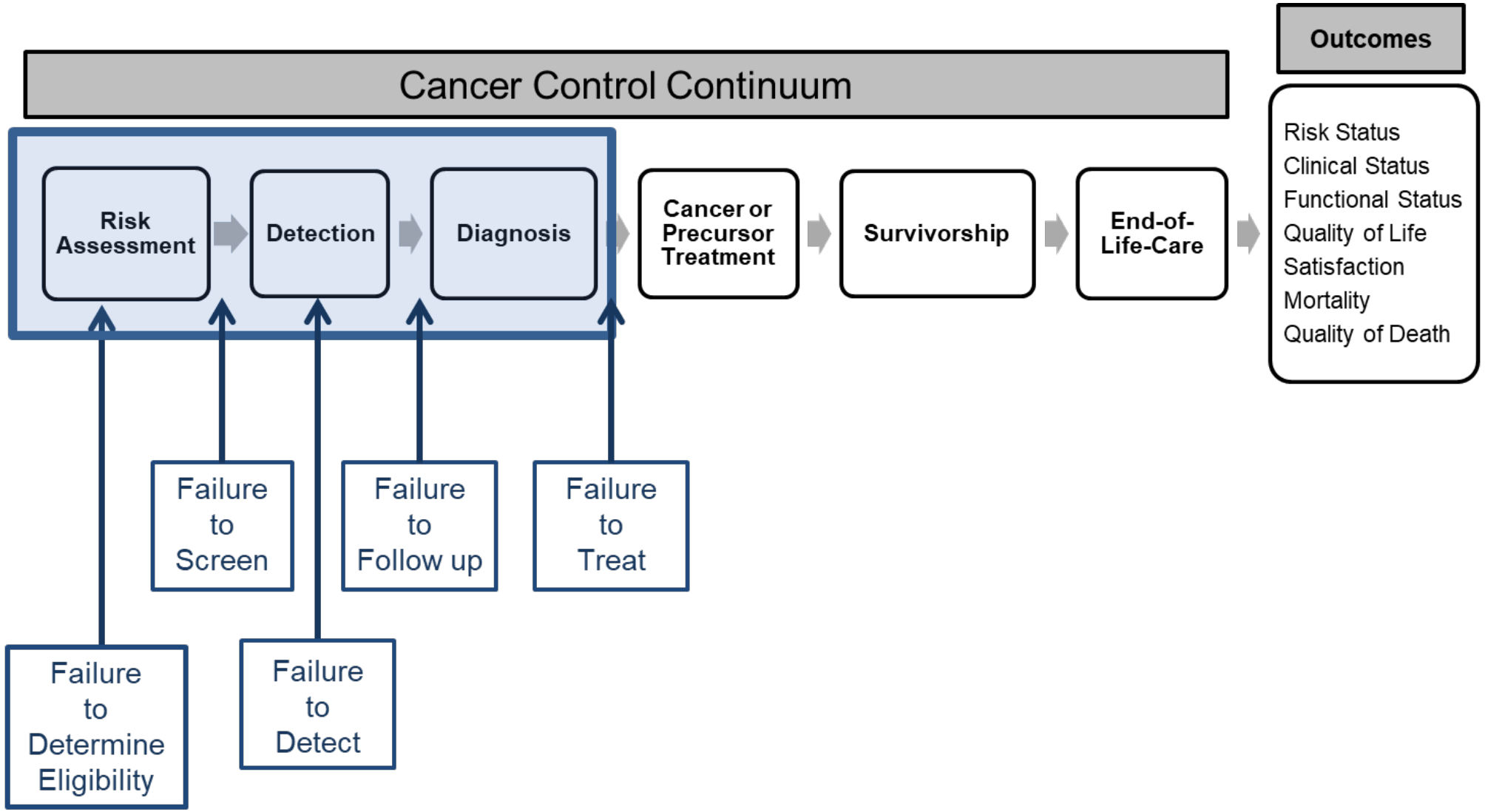
- PROSPR structure and goals
- Research examples
- Data sharing opportunities

PROSPR's Structure and Goals

The goal of PROSPR is to better understand how to improve the cancer screening process in community healthcare settings in the United States.

PROSPR is conducting multilevel observational research to evaluate factors that affect the quality and outcomes of the screening processes for cervical, colorectal, and lung cancer.

Rationale for PROSPR: Breakdowns Can Occur at Multiple Points in the Cancer Screening Process



PROSPR 2 (2018-2024)

- 3 PROSPR Research Centers (PRCs): cervical, colorectal, and lung
- At least 3 health systems per PRC
- PRC aims
  - Collect multilevel screening process data from electronic records systems
  - Conduct PRC-specific research
  - Participate in trans-PROSPR research
- PROSPR Coordinating Center (PCC)

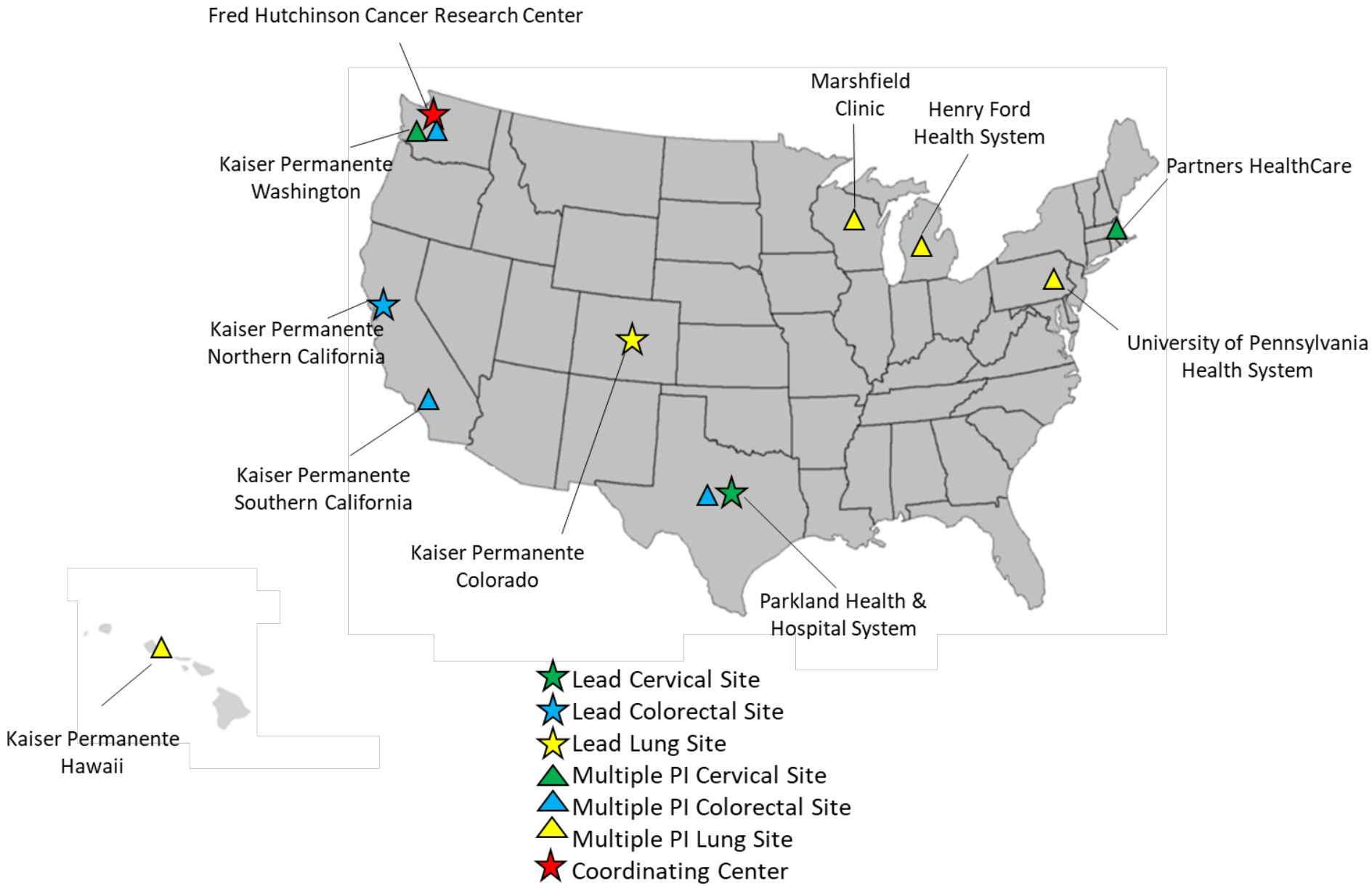
Emphases for PROSPR 2

- Underserved patient populations and social determinants of health
- Trans-PROSPR research to develop common conceptualizations and measures for assessing
  - Systems-level factors
  - Screening quality

PROSPR 1 (2011-2017)

- Similar in basic structure to PROSPR 2
  - Research centers
  - Coordinating center
- Centers focused on breast, cervical, and colorectal cancer screening

PROSPR 2 PRCs and PCC



PROSPR 2 Includes Data from Large, Diverse Cohorts

	Cervical (%) n=862,524	Colorectal (%) n=6,384,599	Lung (%) n=2,130,716
Age (years)			
<40	50	--	22
40-49	17	45	25
50-69	27	45	41
70+	6	11	12
Female	100	52	54
Race/Ethnicity			
Hispanic	25	27	7
White	51	49	65
Black	13	9	14
Asian/PI	8	14	7
Other/Unknown	3	1	5

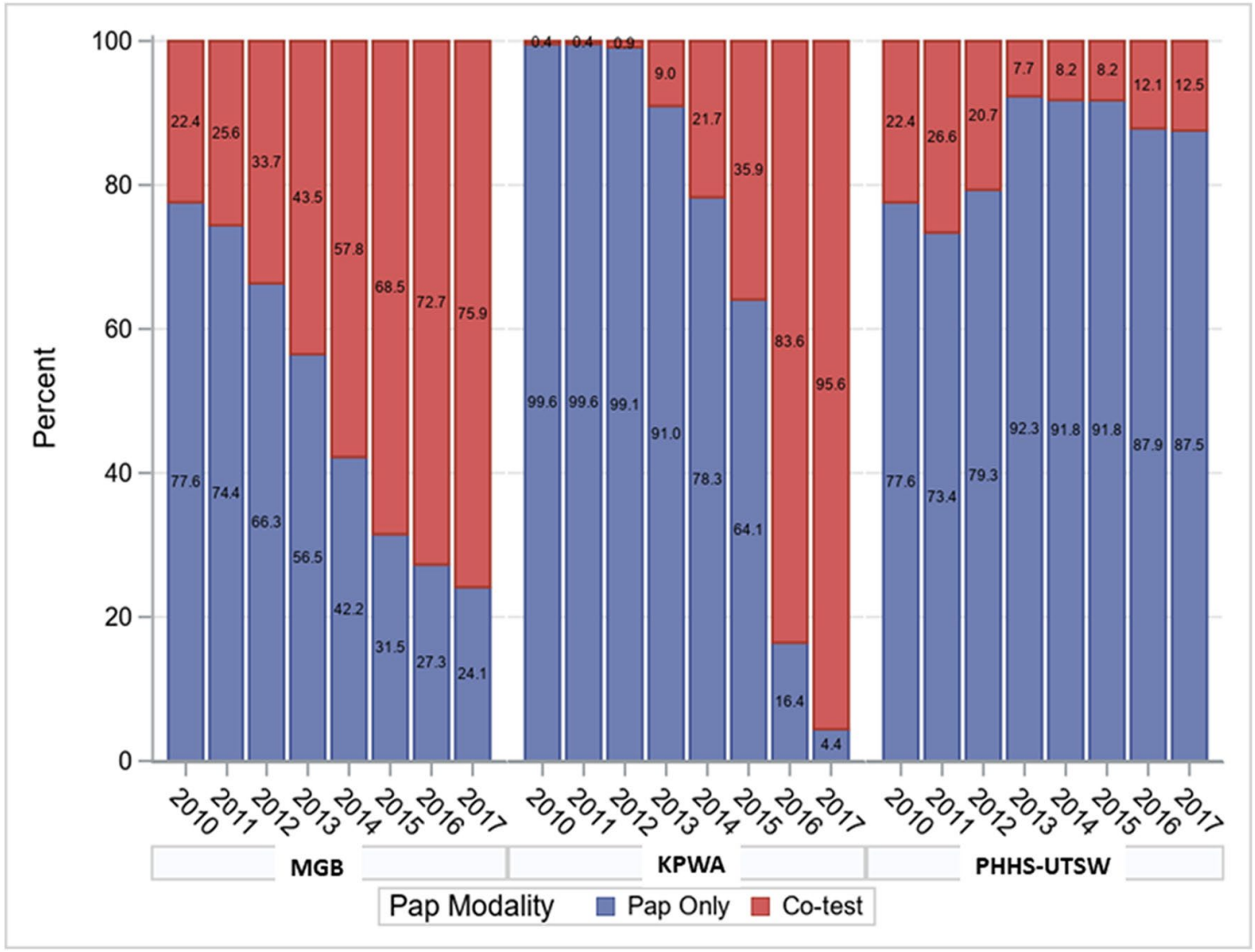
Beaber and Kamineni et al. CEBP 2022;31:1521-31.

Cooperative Agreement Funding Mechanism

- Substantial involvement from NCI program staff (Project Scientists)
- RFA: NCI specifies basic structure and goals of the consortium
- Awardees determine focus/details of organ-specific data collected and research
- Trans-PROSPR research prioritized and conducted by PRCs, PCC, and NCI and coordinated by the PCC

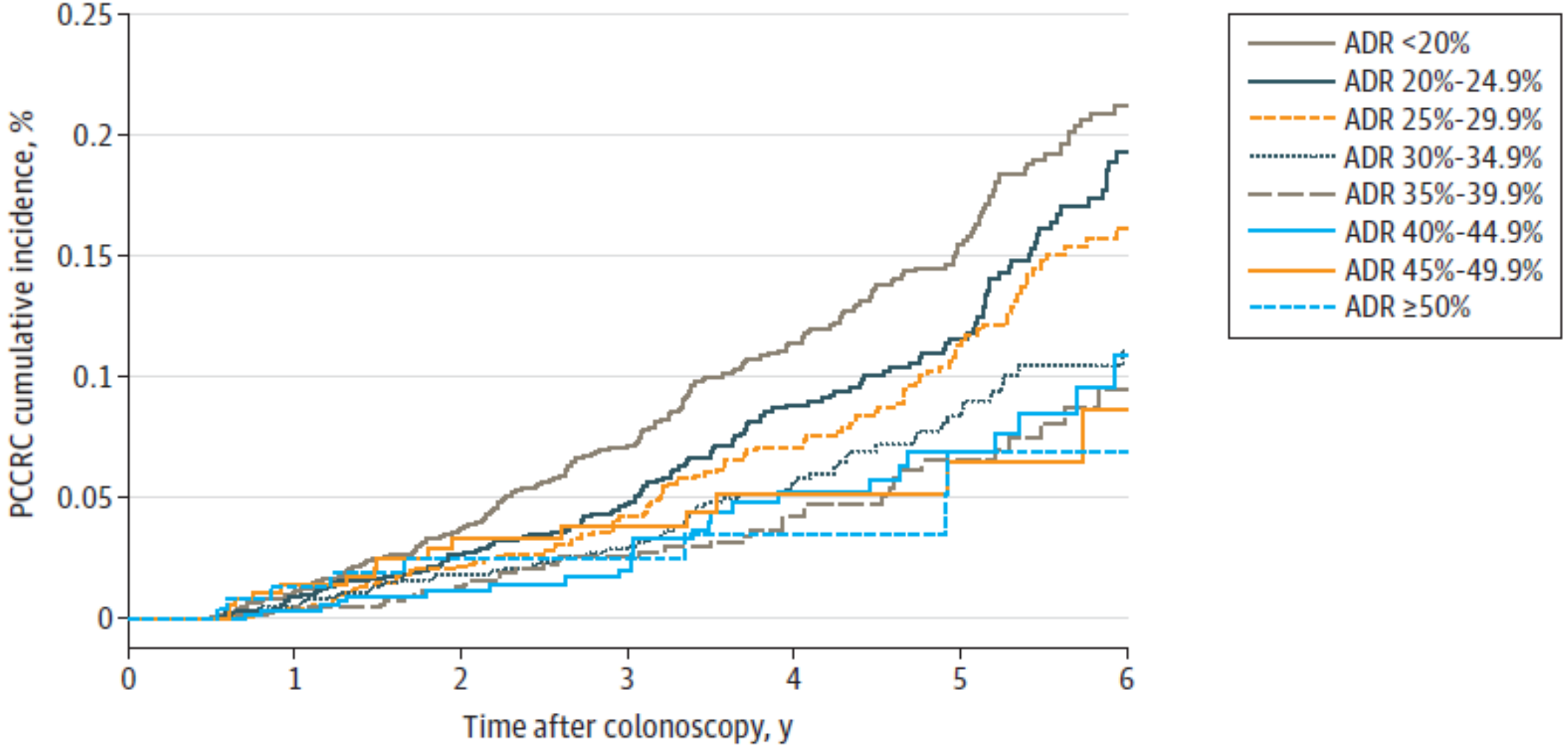
PROSPR Research Examples

Co-testing use from 2010 to 2017 across 3 PROSPR cervical health systems



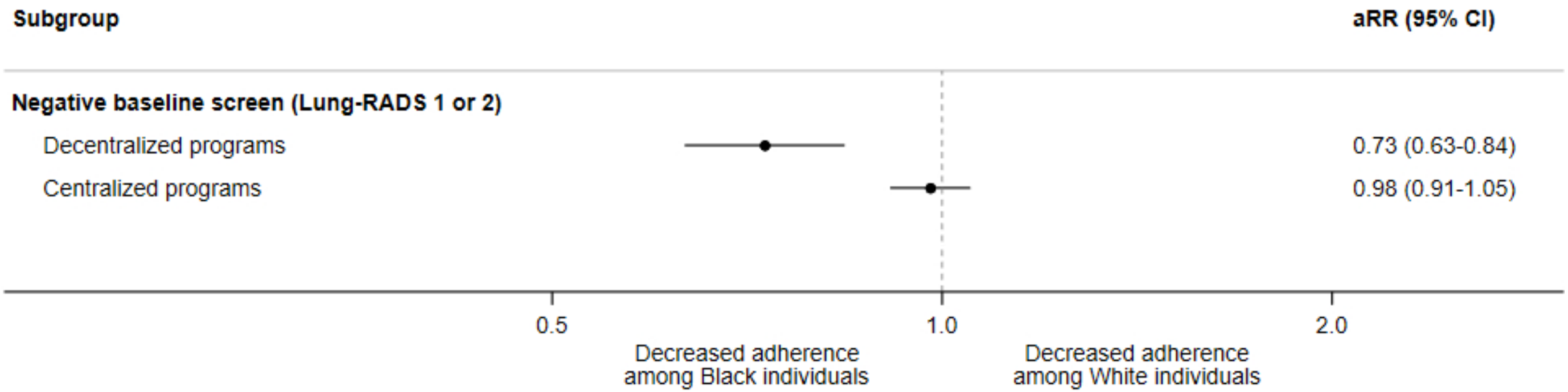
Haas et al. Prev Med 2022;154:106871.

Incidence of post-colonoscopy colorectal cancer according to endoscopist adenoma detection rate



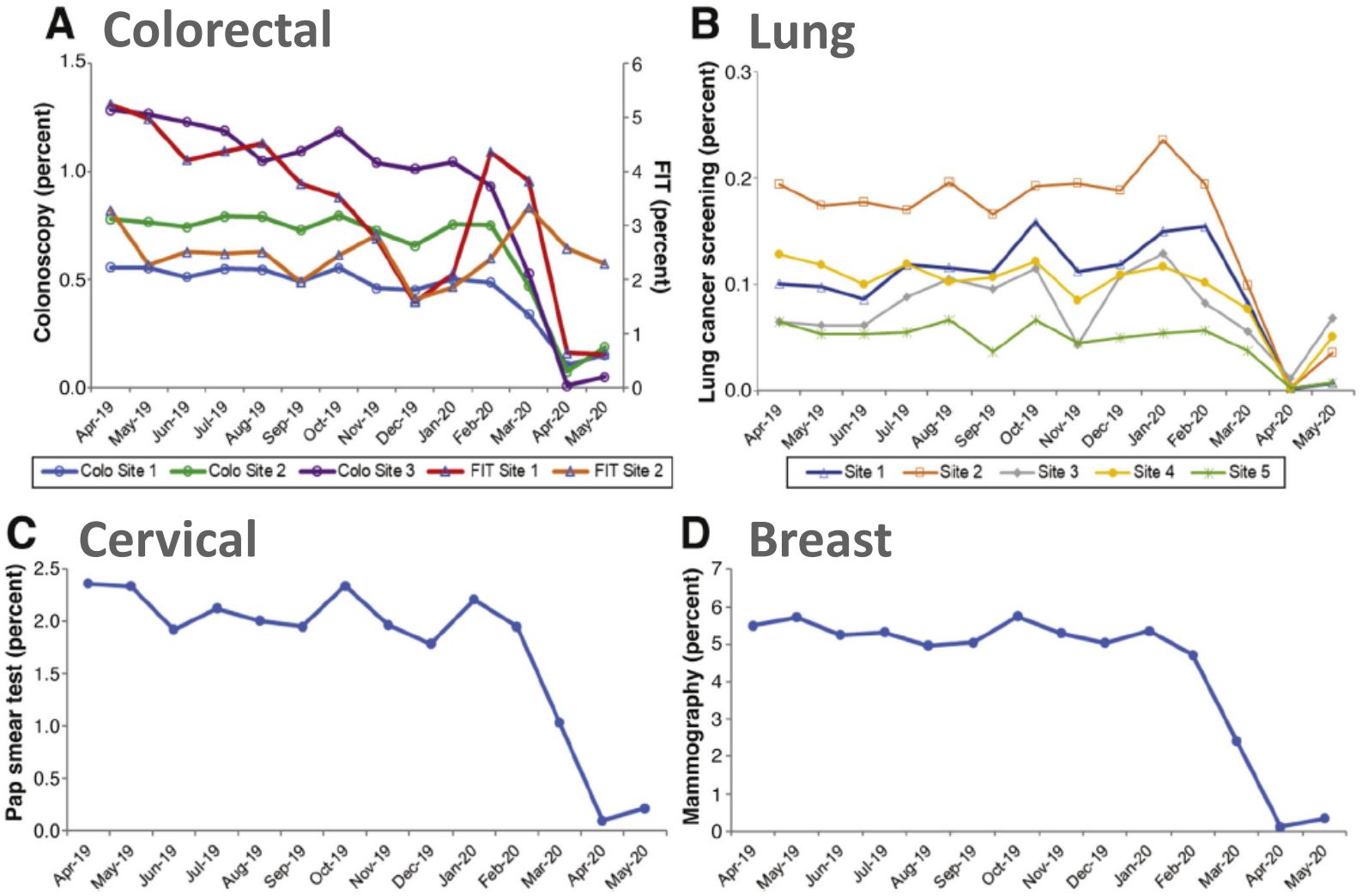
Schottinger et al. JAMA 2022;327:2114-22.

Association of race with adherence to lung cancer screening, by program type



Kim et al. Ann ATS, 2022 Feb 15.

Decline in Cancer Screening During the Early Months of the COVID-19 Pandemic



Corley et al. Gastroenterol 2021;160:999-1002.

Healthcare Organizational Survey: Domains of Interest

A. Organizational structure
(e.g., health plans, payment models, centralized IT/QI/population health)

B. Clinical data infrastructure
(e.g., EHR & ancillary systems, social determinants of health data, EHR defaults and hard stops, clinical decision support)

C. Cancer screening policies
(e.g., organizational formal policies, ordering provider policies, provider-facing and patient-facing reminders)

D. Incentive programs related to cancer screening
(e.g., financial or non-financial based programs, outcomes measured by programs, program oversight)

→ Collected via a written survey completed by key stakeholders at each healthcare organization

E. Quality improvement

F. Culture/climate related to preventive care delivery

→ Did not pursue due to feasibility (would require widespread sampling)

Social Determinants of Health Dataset

- Being used for PROSPR research including
  - Residential segregation and screening
  - SDoH and screening and abnormal follow-up
- Available for download from the HDRP website*
- Dataset includes area-level variables from the American Community Survey (for every US census tract)
  - Race/ethnicity
  - Education
  - Poverty
  - Urban/rural
  - Socioeconomic status
  - Racial segregation indices

*https://healthcaredelivery.cancer.gov/social-determinants; must agree to certain standard data use conditions

Population-based Research to Optimize the Screening Process (PROSPR) Initiative: Data Sharing

Pamela Marcus, PhD, MS, ELS

NIH NATIONAL CANCER INSTITUTE

September 20, 2022

Today's talk

- Rationale for sharing PROSPR data
- PROSPR DataShare
- Types of data requests
- Closing thoughts

Rationale for sharing PROSPR data

Overarching rationale for sharing any data

- From the NIH Scientific Data Sharing website*:

…Sharing data accelerates research discovery, enhances research rigor, and promotes data reuse for future research studies…

…Ultimately, the sharing of scientific data expedites the translation of research results into knowledge, products, and procedures to improve human health…

*some words omitted in the interest of brevity. To see the entire statement, visit https://sharing.nih.gov/data-management-and-sharing-policy/about-data-management-and-sharing-policy/data-management-and-sharing-policy-overview

Rationale for sharing PROSPR data

- PROSPR grantees are governed by two data sharing requirements
  - 2003 NIH Data Sharing Policy
    - In force at the time of awards
    - Required of all awards with more than $500,000 in direct costs (annual)
  - Additional stipulations set forth in the PROSPR RFA

2003 NIH Data Sharing Policy

Expectations Under NIH's 2003 Data Sharing Policy

NIH expects that data be made as widely and freely available as possible while safeguarding the privacy of participants and protecting confidential and proprietary data. Sharing is particularly important for unique data that cannot be readily replicated.

Under NIH's 2003 Data Sharing Policy, investigators are expected to:

- **Include a data sharing plan** in research proposals seeking $500,000 or more in direct costs describing how final research data will be shared. Alternatively, the investigator is expected to explain why data sharing is not possible.
- **Release and share the data**, as described in the approved application, no later than the acceptance for publication of the main findings from the final dataset.
- **Report any progress** made on data sharing progress in the annual submission of their Research Progress Performance Report (RPPR).

Given the breadth and variety of science that NIH supports, the NIH Data Sharing Policy does not require specific ways of documenting, formatting, presenting, or transporting data.

https://sharing.nih.gov/data-management-and-sharing-policy/about-data-management-and-sharing-policy/data-management-and-sharing-policy-overview

PROSPR RFA data sharing conditions

- PROSPR investigators are expected to make data resources available to extramural investigators for research purposes
- Research will generally be conducted in collaboration with PROSPR investigators
- PROSPR investigators, supported by the PROSPR Coordinating Center, will participate in the development of a process for data sharing

https://grants.nih.gov/grants/guide/rfa-files/rfa-ca-16-016.html

The challenge

- To the best of our knowledge, the specifics for how to share large scale PROSPR-like data collected using a federal grant mechanism had not been worked out
  - Community-based clinical practice data
    - Primary source: Electronic health records
    - Primary reason for collection: not research
    - Federal grants, not contracts
    - Observational data

The solution

- The PROSPR Team worked together to build a web-based data sharing system
- We called it PROSPR DataShare

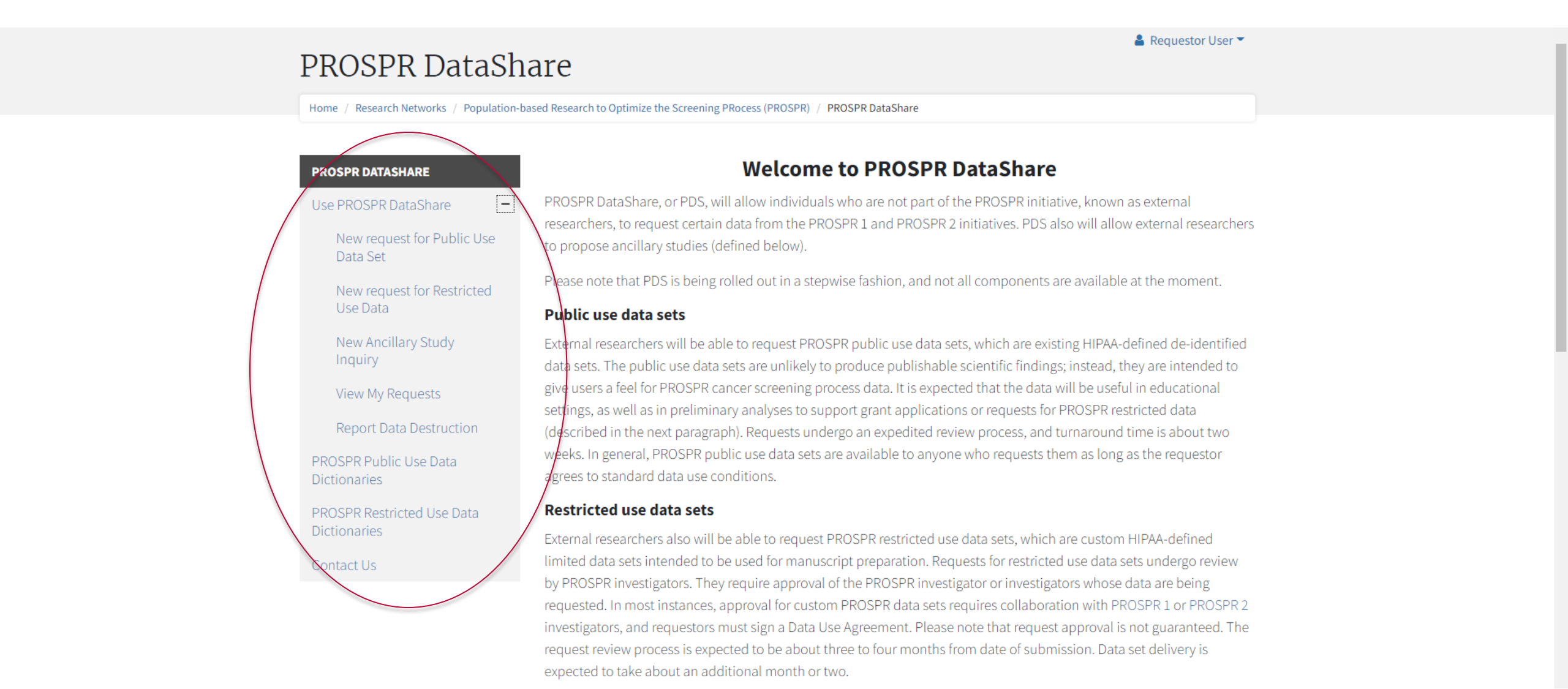
PROSPR DataShare (PDS)

PROSPR DataShare

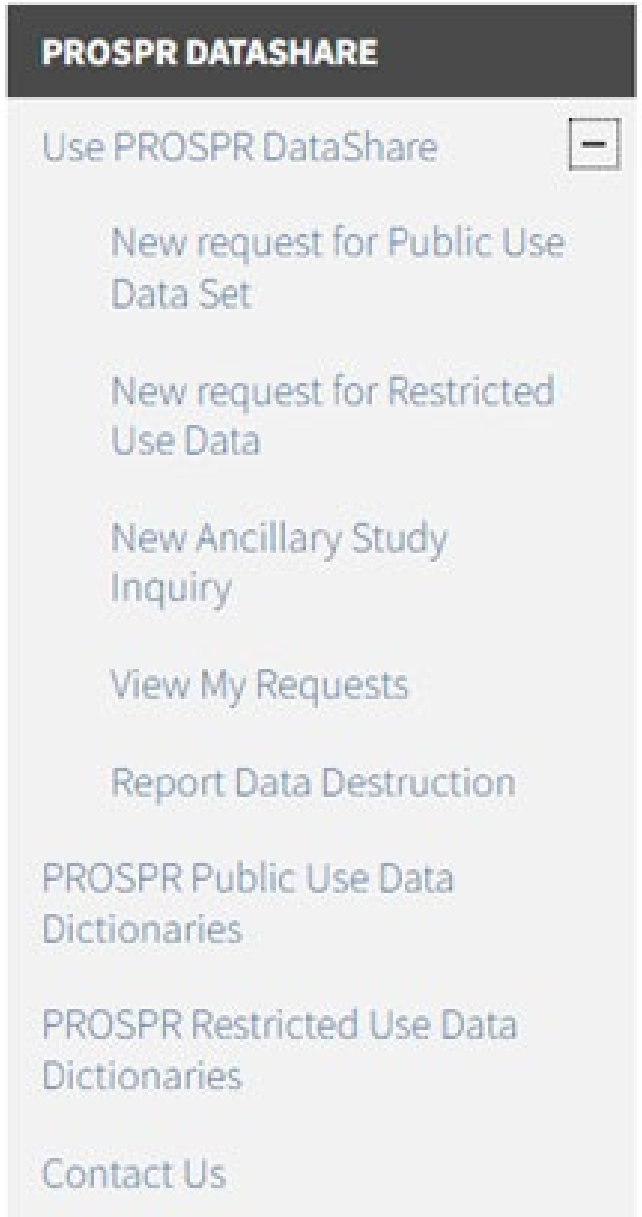
- What is PROSPR DataShare?
  - Web-based system through which non-PROSPR researchers can request PROSPR data and explore ways to collaborate with PROSPR researchers
  - Uses the PROSPR Data Repository to fulfill most data requests

PROSPR DataShare capabilities

- Data dictionaries
- Information about application processes
- Request entry
- Request review
- Data downloads (for certain approved requests)
- Notification/tracking
- Contact us



Full page available at https://healthcaredelivery.cancer.gov/prospr/datashare/


PROSPR DATASHARE
Use PROSPR DataShare
New request for Public Use Data Set
New request for Restricted Use Data
New Ancillary Study Inquiry
View My Requests
Report Data Destruction
PROSPR Public Use Data Dictionaries
PROSPR Restricted Use Data Dictionaries
Contact Us

Creating PROSPR DataShare

- All PROSPR entities worked together on numerous aspects of PROSPR DataShare development
  - Policies
  - Practices
  - Request forms
  - Documentation

- Each also had a well-described primary role

Creating PROSPR DataShare: primary roles

- PROSPR Research Centers (PRCs): contribute data and documentation to the PROSPR data repository

- PROSPR Coordinating Center (PCC): led development of the process by which external researchers could make data requests/propose collaborative projects

- NCI: data storage and system development (via IMS, a DCCPS contractor)

Other role: wrangling

- All team members wrangled with challenges inherent to data sharing, including but not limited to:
  - HIPAA-protected data
  - Facility and system level data
  - Providing enough documentation so that others can use data correctly
  - Balancing one's own research interests
  - Stewardship of data once it is shared

Types of requests

Three types of PROSPR requests

- Request for a public use data set (PUDS)
- Request for a custom-made restricted use data set (RUDS)
- Propose an ancillary study

Public use data sets (PUDS)

- Existing data sets
- Individual-level cohorts, HIPAA-defined de-identified data elements only
- Contents:
  - Cancer screening process information (exam, results, diagnostic evaluation, diagnosis)
  - Some covariates (e.g.: birth year, race, comorbidities, BMI)
  - 1 or 2 years of data

Public use data sets (PUDS)

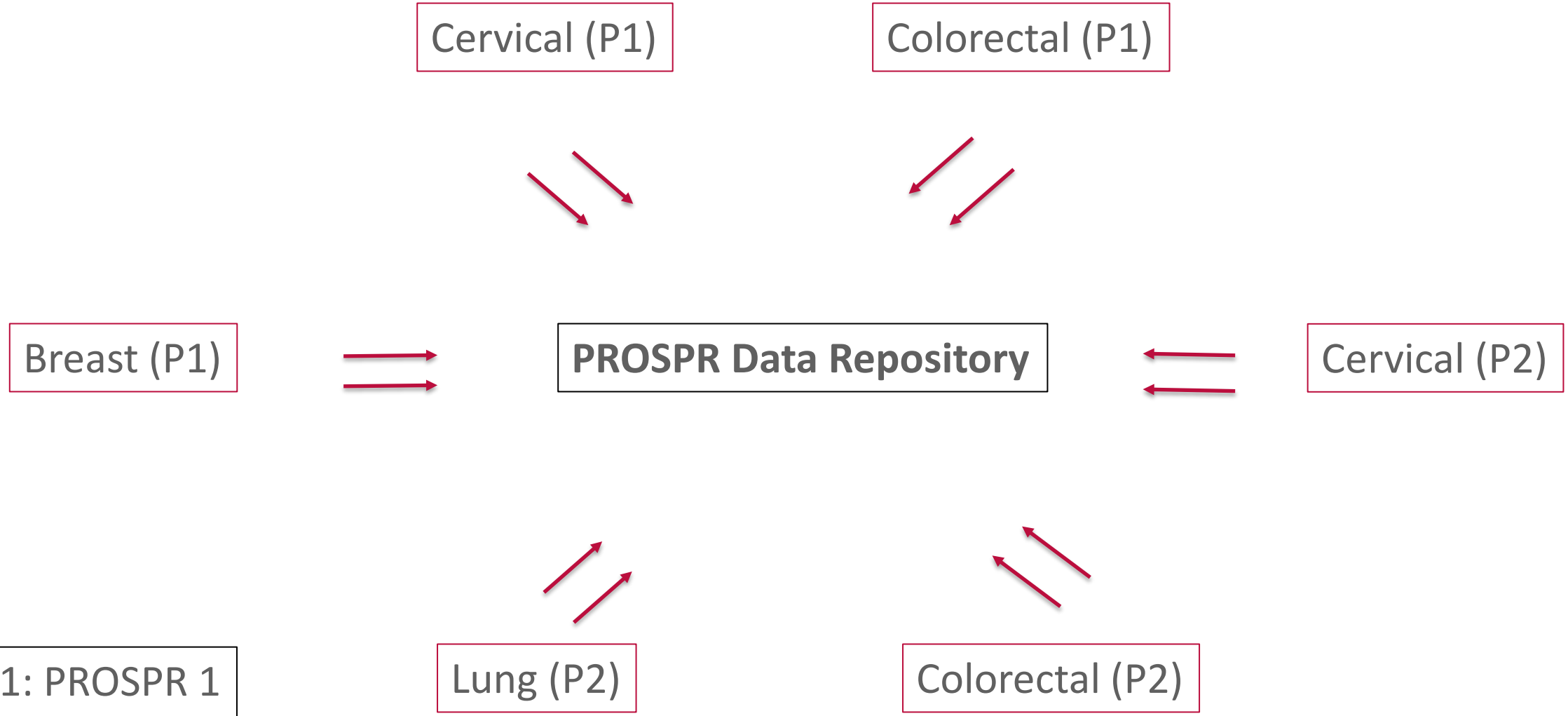
- Both PROSPR rounds and all organ sites
- Available to anyone who agrees to certain standard data sharing conditions

- Best use:
  - Give users a feel for PROSPR cancer screening process data
  - Educational settings; preliminary analyses to support grant applications; requests for PROSPR custom data sets
  - Unlikely to produce publishable scientific findings

Restricted use data sets (RUDS)

- Both PROSPR rounds and all organ sites
- Data sets made on a case-by-case basis, created from the PROPSR data repository
  - Classes of data elements provided on an upcoming slide
  - Best to consult data dictionaries to see the exact data elements
  - Availability of certain data elements is on a case-by-case basis

- Best use:
  - Scientific publication


Cervical (P1)
Colorectal (P1)
Breast (P1)
PROSPR Data Repository
Cervical (P2)
Lung (P2)
Colorectal (P2)
P1: PROSPR 1
P2: PROSPR 2

Multiple files per entity (examples)

- Participant
- Enrollment
- Screening episode
- Procedure
- Diagnosis
- Comorbidity
- Cancer registry
- Medical encounter
- Death
- Provider
- Facility
- Social determinants of health

Ancillary studies

- Use PROSPR infrastructure to collect or access data not currently in the PROSPR Data Repository

- External researchers are required to partner with at least one PROSPR investigator when proposing and conducting an ancillary study

Logistics/cost by type of request

	Public use data set	Restricted use data set	Ancillary study
Cost	Free	Almost always free	Cost
Review process	Minimal	Multi-step (inquiry and proposal)	Determined by relevant PRCs
Collaboration requirement	None	Usually	Always
IRB/DUA	No	Probably	Probably
Turnaround time	~2 weeks	4-6 months	Varies

Closing Thoughts

Status of PROSPR DataShare

- PROSPR DataShare is live
  - Data dictionaries are accessible and downloadable
  - Public use data sets can be requested
  - Inquiries regarding restricted use data set can be accepted
  - Ancillary study inquiries will be possible by the end of this week

- PROSPR DataShare hasn’t been available for very long

We’re only half-way through the story

- Can’t speak to successes or challenges just yet

- We hope that our experience with PROSPR DataShare will help extramural investigators as the 2023 NIH Data Management and Sharing Policy comes into force

2023 NIH Data Management & Sharing (DMS) Policy

- Applies to nearly all applications submitted for 1/25/23 due dates and later
- Regardless of level of funding

- *NIH expects that investigators and institutions:*

  Submit a DMS plan for review when applying for funding

  Comply with the approved DMS plan

https://sharing.nih.gov/data-management-and-sharing-policy/about-data-management-and-sharing-policy/data-management-and-sharing-policy-overview

What does this mean?

- Researchers need to plan for data sharing while preparing applications
  - Most researchers, not only those with large grants

- Less flexibility in how data are shared
  - 2003 plan – not specified
  - 2023 plan – appears that data need to be deposited in a repository

What do I expect*?

- Bring to light the difficulties in sharing community-based clinical practice data
  - EHR
  - HIPAA
  - Concerns of patients, providers, health care systems

- Provide impetus to develop the needed resources for sharing PROSPR-like data

*My opinion

Thank you to…

Thank you!

- IMS
  - Brad Ohm, Todd Gibson, and David Roney
- PROSPR grantees
  - PROSPR Coordinating Center
  - PROSPR 1 and 2 Research Sites and Centers
- NCI HDRP PROSPR Team
- NCI HDRP Leadership

NIH NATIONAL CANCER INSTITUTE

www.cancer.gov

www.cancer.gov/espanol