



# Cohort Overview

**Mia M. Gaudet, Ph.D.**

Connect Chief Scientist

Trans Divisional Research Program  
Division of Cancer Epidemiology and Genetics  
National Cancer Institute



April 1, 2022

# Study Objectives

**Comprehensive research resource using new technologies and methods for the scientific community to study:**

## **Cancer Etiology**

- Precursor to Tumor Progression
- Cancer Risk Prediction
- Early Detection of Cancer
- Survivorship

## Connect Designed to Address Key Research Priorities



Emerging exposures



Novel biomarkers, genomics



Cutting-edge methodology

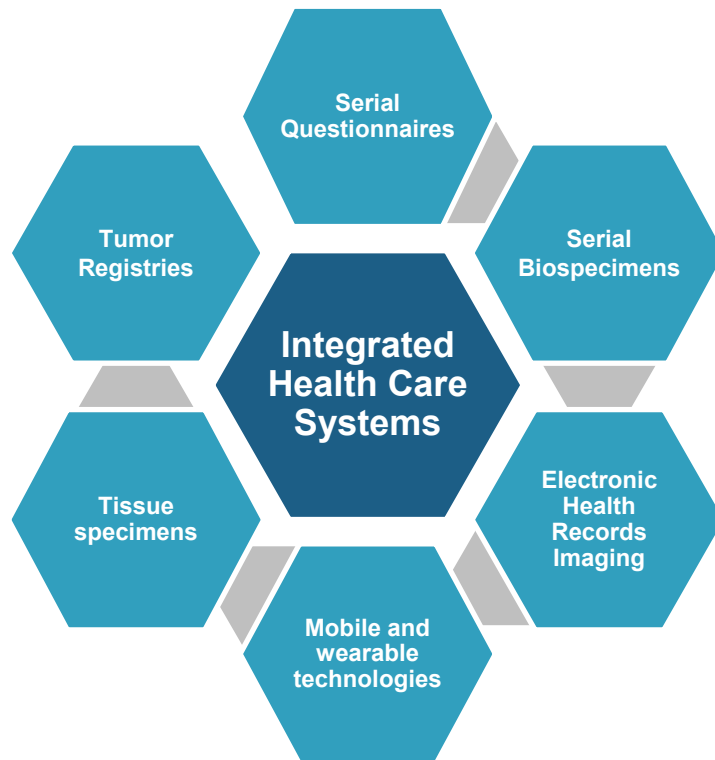


Diverse and special populations

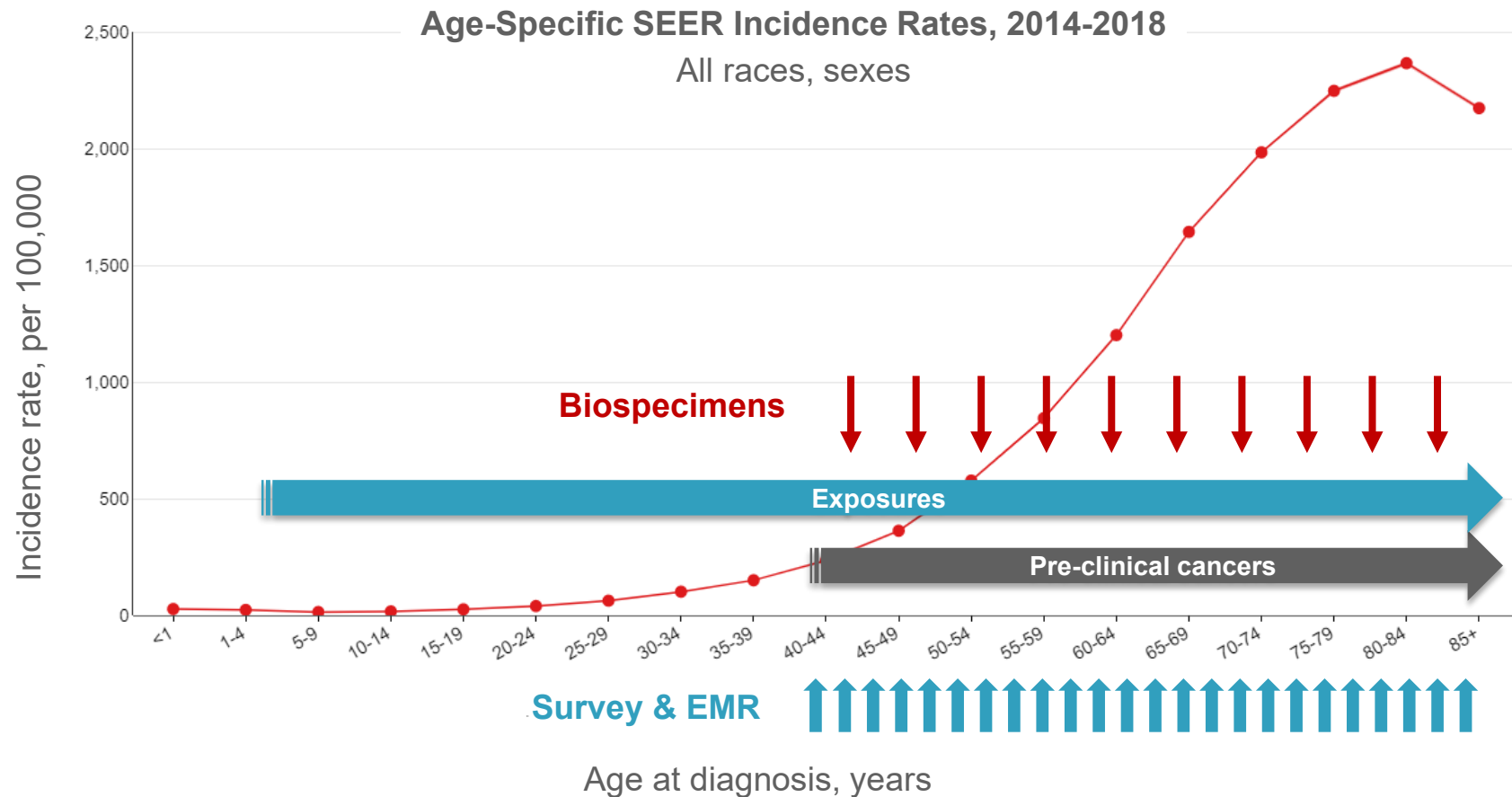
# Cohort Study Design Overview

# Key Features of Connect as a Trans-NCI Resource

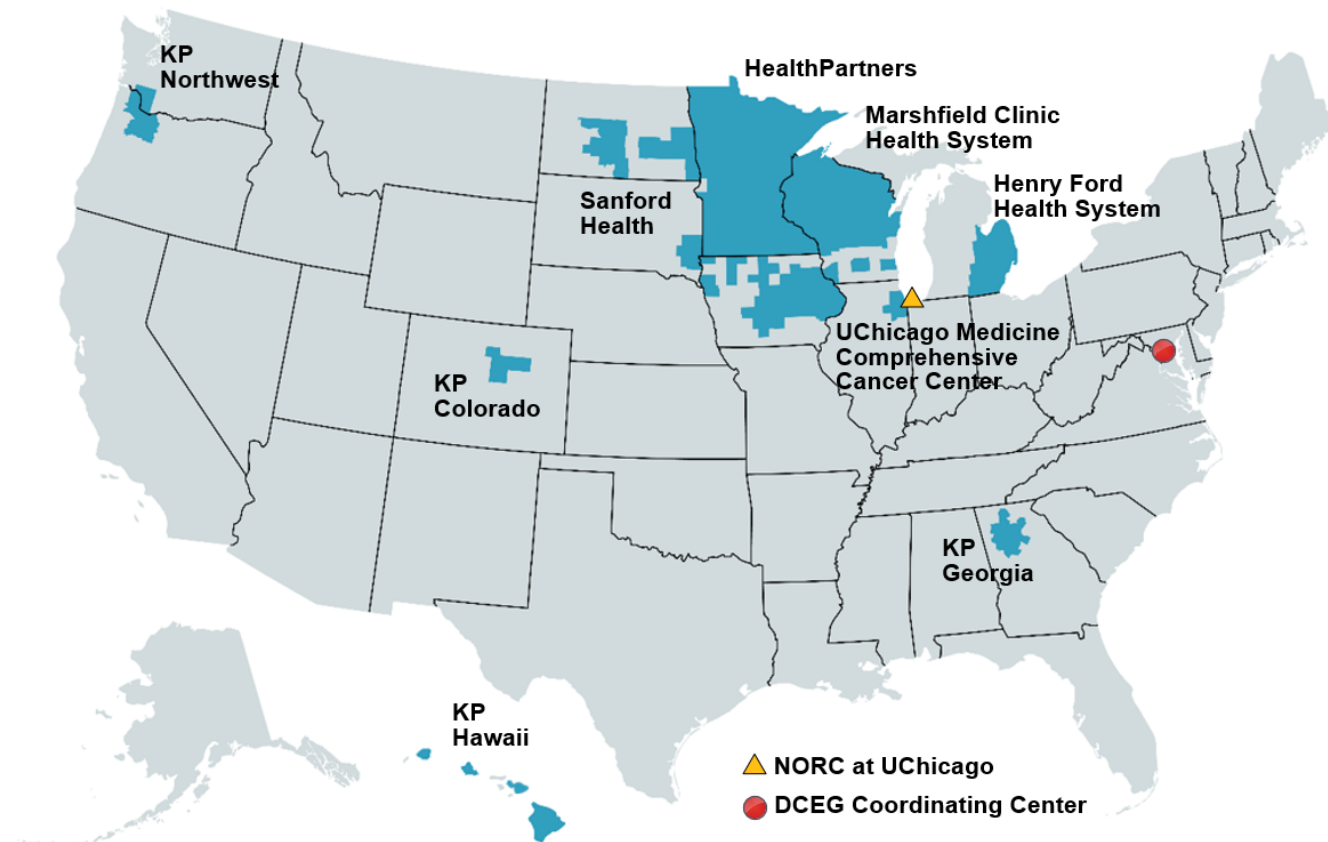
- ▶ 200,000 adults across the US
  - ✓ Aged 40-65 years
  - ✓ No history of cancer
  - ✓ Patients or members of partner health care systems
- ▶ Long-term follow-up
- ▶ Serial assessments
- ▶ Comprehensive cancer outcomes
- ▶ Flexible infrastructure for enhancement studies



# Serial Assessments to Observe Longitudinal Changes



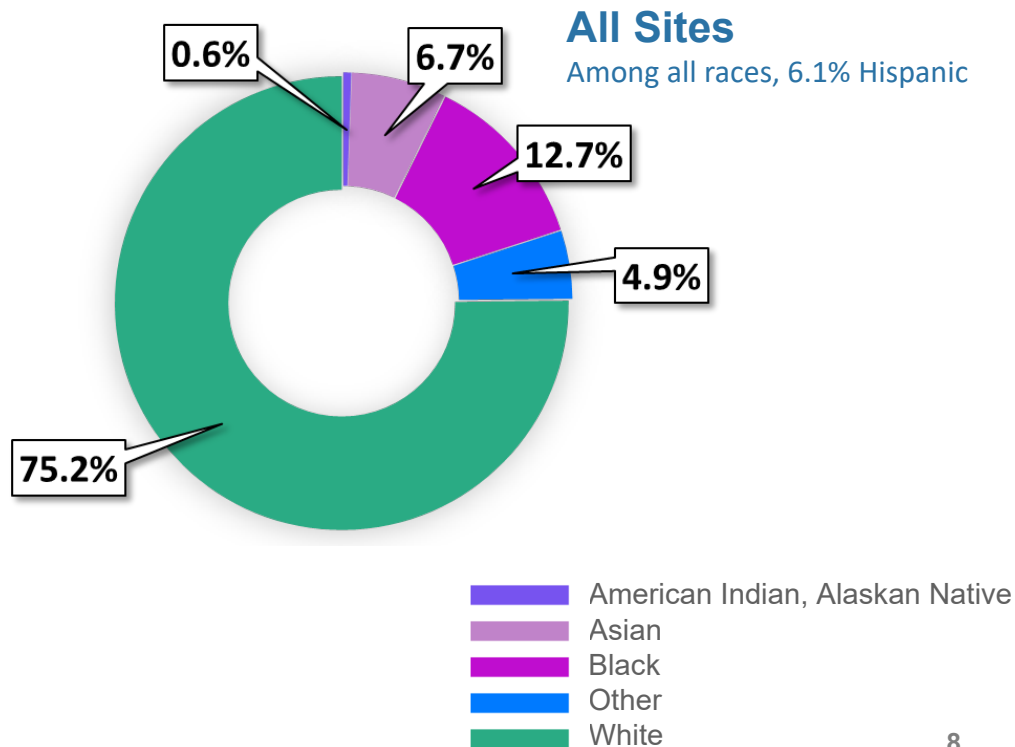
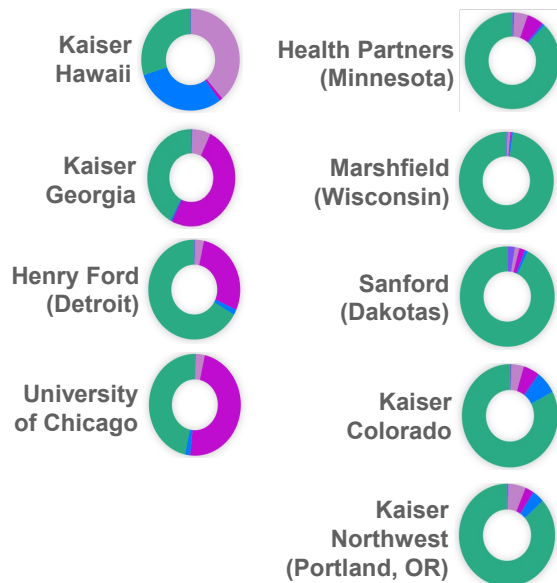
# Recruitment at Partner Health Care Systems



# Diversity of Catchment Population: Race and Ethnicity

(Based on 2018 estimates)

## Individual Recruitment Sites





# Participant Engagement

## Recruitment Materials & Strategy

- Value proposition to anchor communications
- Prevention-focused messaging
- Materials designed to engender trust, enthusiasm, and inclusivity
- Monetary and non-monetary incentives
- Active and passive recruitment

**Connect today.  
Prevent cancer  
tomorrow.**

**connect**  
for cancer prevention study

# Retention Strategies

- Candid messaging about participant burden for the long-term
- Responsive communication using multiple channels
- Incentives for retention, such as return of results



# Ongoing Engagement Strategies

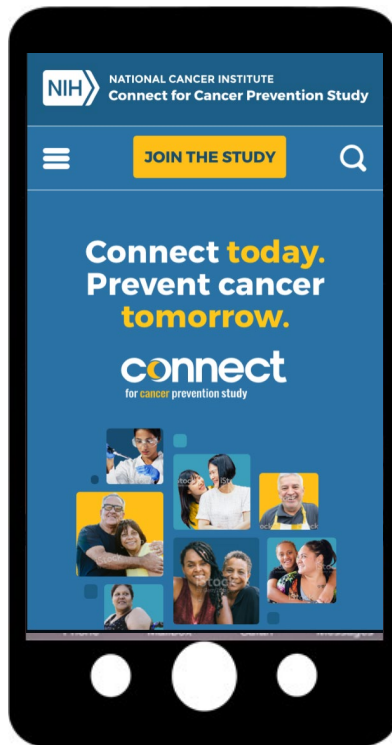
- Digital engagement
- Regular contact every 6 months beyond those for study activities
- “Study anniversary” mailing
- Monitor real-time metrics to pivot engagement strategy



# MyConnect Participant App

# MyConnect Participant App: Communication and Study Activity Hub

Reminders &  
Communication



Healthcare  
Provider App



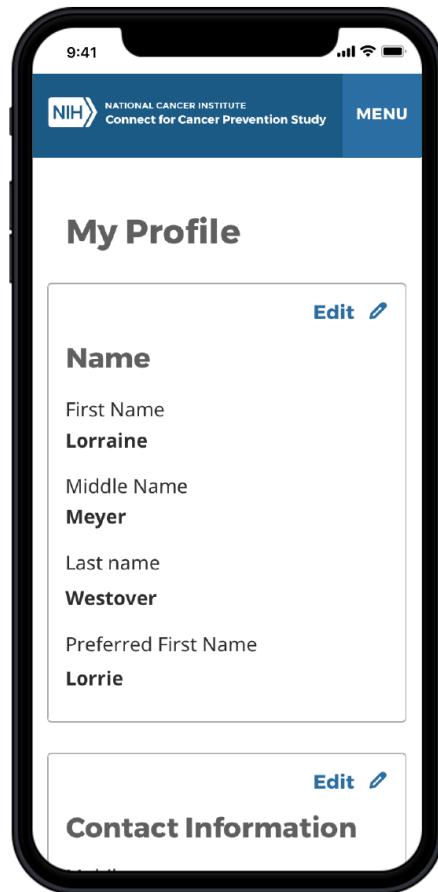
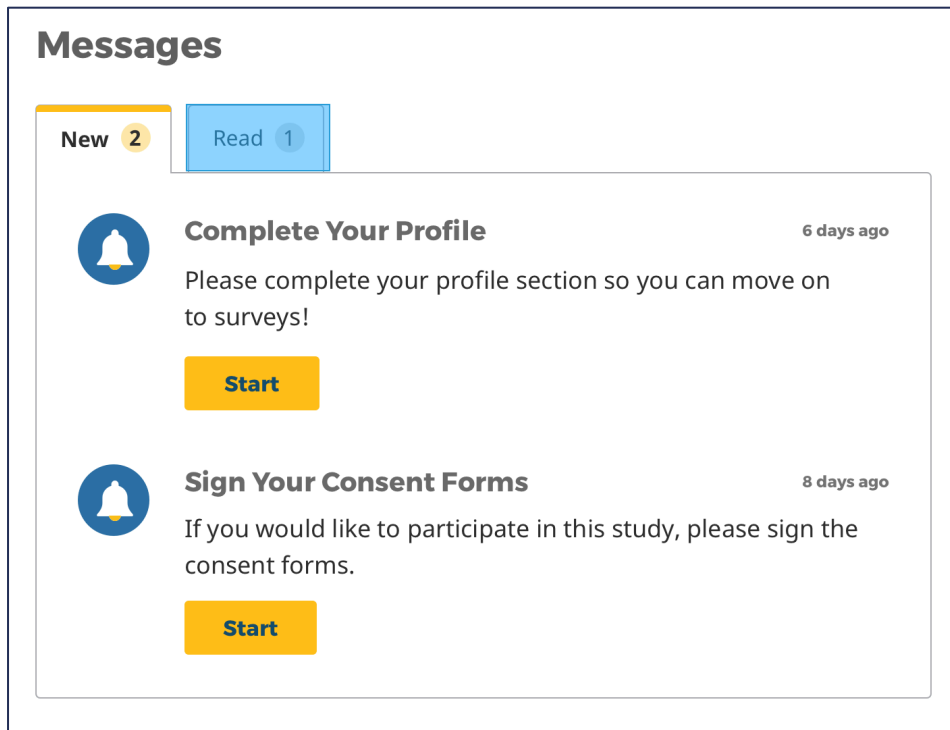
Health App




Mobile/ Wearable  
Technologies



# Messages and Data Input



# Educational Module before Long Form Consent

 NATIONAL CANCER INSTITUTE  
Connect for Cancer Prevention Study

Sign Out

Surveys Samples Agreements About the Study What to Expect Privacy My Profile Messages

1

About

2

Activities

3

Privacy

4

Leaving

5

Results

6

Benefits

7

Consent

8

Thank You

## What is the Connect for Cancer Prevention Study?

This research study will explore causes of cancer and learn how to better prevent cancer in adults. Since it takes time to understand what causes cancer, Connect will go on for many years. The longer you participate, the more we may learn.

Researchers will study things like habits, behaviors, and the environment you and others live in that can help us learn how cancer may form. Researchers hope to learn new ways to stop cancer from forming in the first place

Connect is studying cancer prevention. Researchers will not look for treatments for cancer, give medical care, or share medical advice.

Next



NATIONAL CANCER INSTITUTE  
Connect for Cancer Prevention Study

MENU

7 of 8 Consent

## Informed Consent

In order to join Connect, we need to you to give your informed consent by reviewing the full consent form and electronic health records release form, and signing your name.

### Informed consent form

This is a more detailed explanation of what it means to take part in Connect.



[Download an unsigned copy of the consent form](#)



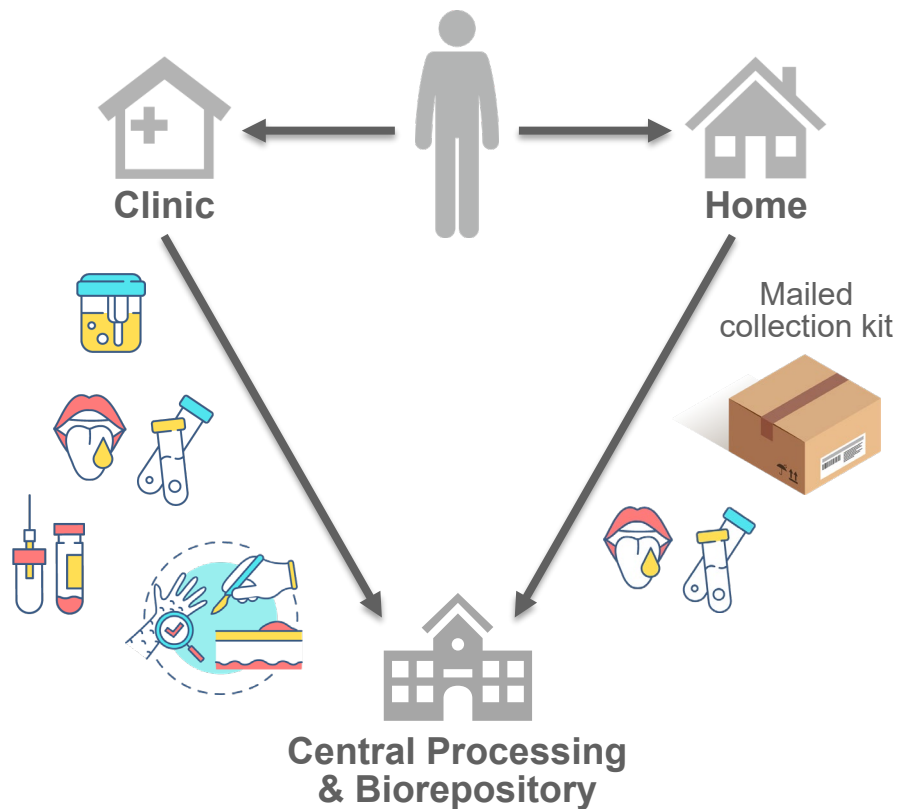
# Overview of Study Activities and Data Collection

# Study Activities

Timing	Type of Study Activity	
	Participant actively completing...	Occurring in the background...
Recruitment	Surveys	EHR data transfer
	Specimen donation	Benign breast biopsies, colon adenomas, other tissues
Follow-up	Surveys	EHR data transfer
	Specimen donation	Clinical archived/ discard specimens
	Physical measurements	Linkages
	Wearable devices, downloaded apps	Pathology and radiology imaging

# Biospecimen Collections

# Biospecimen Collections at Recruitment



	No. Tubes	Fraction	Tube Type
<b>Buccal</b>	1		Saliva-mouthwash kit
<b>Blood</b>	2	Serum	10ml Red top
	1	Plasma/buffy coat	10ml Lavender top (EDTA)
	1	Plasma/buffy coat	10ml Green top (Heparin)
	1	Whole blood	6ml ACD
<b>Urine</b>	1		10ml Urine kit
<b>Tissue</b>			Precursor tissue blocks, if available

## Biospecimen Collections at Follow-up

- Timed on average two to three years
  - More frequent for participants at higher risk of cancer
- Series of same biospecimens and new types of biospecimens (e.g., liquid biopsy, stool)
- Located at healthcare system, external clinic, home
- Paired with questionnaires and other relevant data collection

# Surveys

## Background and Overall Health

- Background Information
- Medical History
- Family History of Cancer
- Education and Occupation

## Medications, Reproductive Health, Exercise, Sleep

- Medications
- Pregnancy History
- Physical Activity
- Sleep


## Smoking, Alcohol, Sun Exposure

- Tobacco
- Marijuana
- Alcohol
- Sun Exposure

## Where You Live and Work

- Residential History
- Commuting

# Questionnaire Module Topics and Display in MyConnect at Recruitment

 NATIONAL CANCER INSTITUTE  
Connect for Cancer Prevention Study

Sign Out

Surveys | Samples | Agreements | About the Study | What to Expect | Privacy | My Profile | Messages

1 Introduction

2 **Background Information**

3 Medical History

4 General Health

5 Family History

6 Education and Occupation

How often do you have children in the household from birth to 5 years of age in your household?

☒ Every day, at least 2 hours each day

☐ Every day, less than 2 hours each day

☐ Most days of the week

☐ One Day per week

☐ A few days per year

☐ Never

Clear

# Questionnaire Topics to Accompany Biospecimen Collection at Recruitment

## Blood

- Covid-19 exposure and vaccination
- Menstrual cycle dates (last, next)
- Recent exposures to food, drink, pharmaceuticals

## Urine

- Menstrual cycle dates (last, next)
- Recent exposures to food, drink, pharmaceuticals

## Mouthwash

- Oral health and care
- Recent exposures to food, drink, pharmaceuticals
- Collection parameters, if home collection

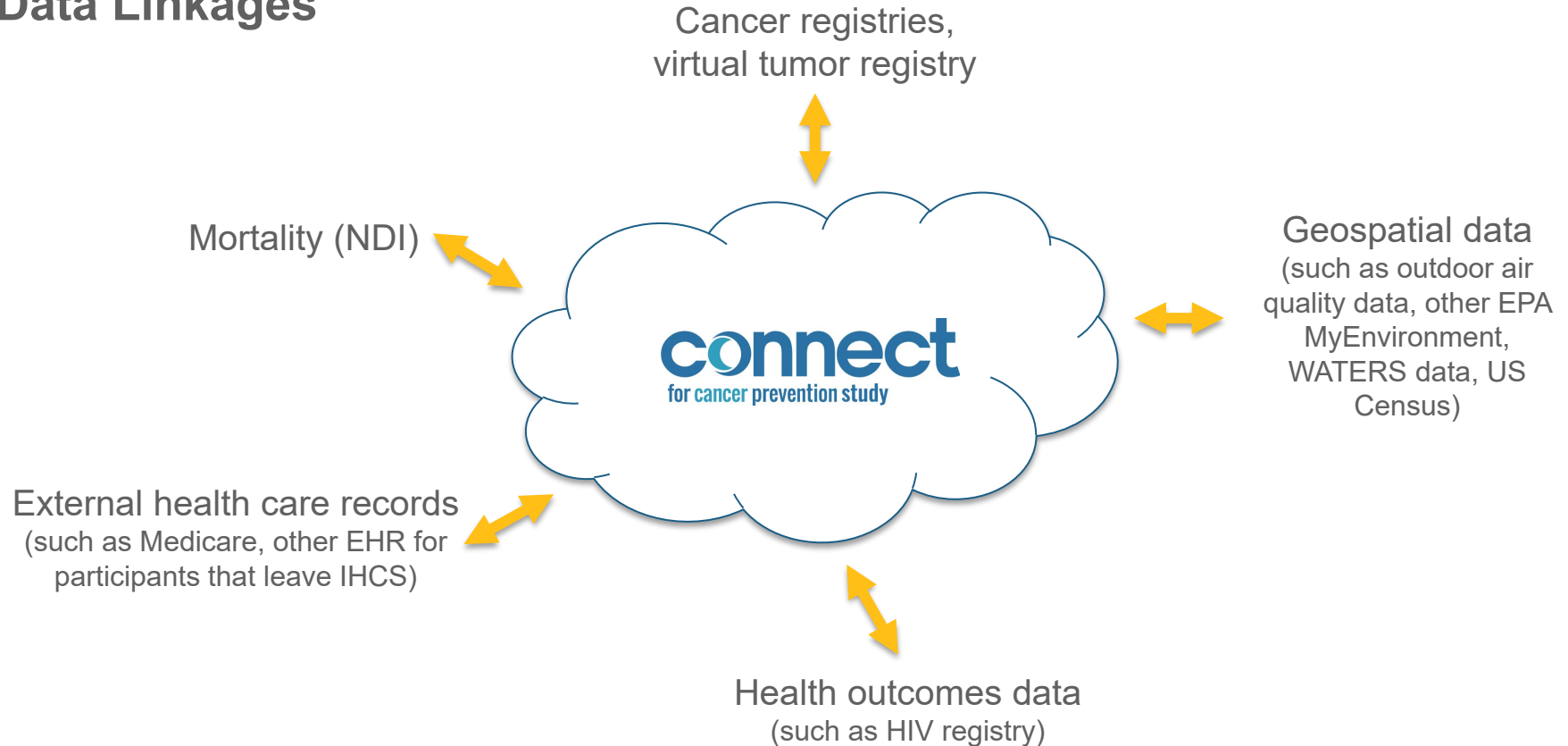


## Planned Follow-up Surveys

6 Months	Within 24 Months	12 Months	2 Years
Diet food frequency questionnaire (DHQ)	Repeated 24-hour recalls for diet (ASA24) & physical activity (ACT24)	Screening Quality of life Psychosocial	Update on baseline exposures, additional questions of covered and other topics
Survey 3-6 months after cancer diagnosis			

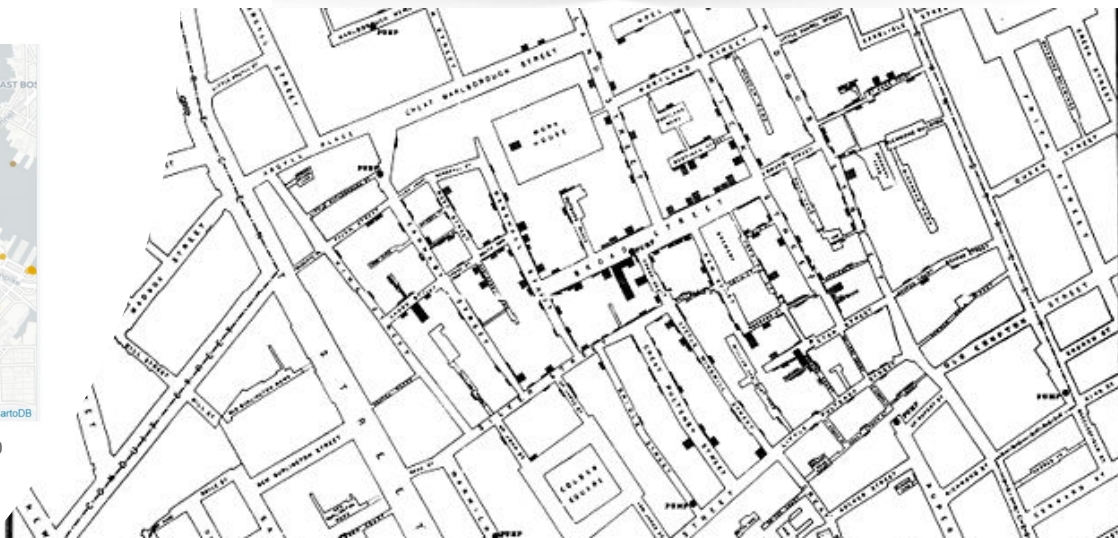
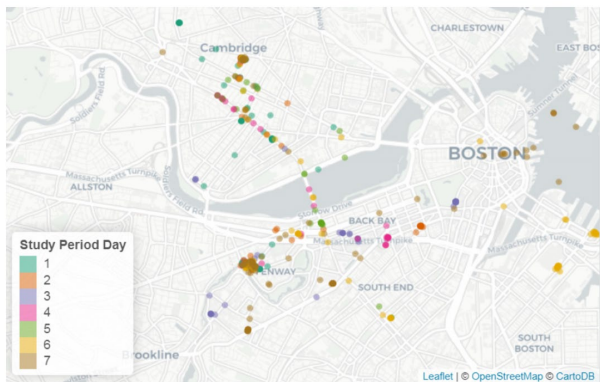
# Data Linkages

## Data Linkages



# Geographic Information

- Geocode 20-year address histories (collected on baseline survey)
- Enable linkages to datasets for numerous external risk factor and exposure information
- Useful for planning ancillary field studies
- Add daily movement and travel to exposure information via MyConnect



# Wearable Devices and Apps

- Connect will have research-grade and consumer-oriented devices available
- Key domains where tech can improve assessment of exposures and risk factors:
  - sleep
  - light at night
  - air pollution
  - physical activity



# Estimated Outcomes

## Estimated total number of incident invasive cancers

Cancer site	Total Number of Expected Incident Cancers	
	10 years of follow up	15 years of follow up
Prostate (males only)	1,951	3,595
Breast (females only)	1,566	2,553
Lung & Bronchus	1,151	2,165
Colorectal	877	1,566
Melanoma of the Skin	410	681
Non-Hodgkin Lymphoma	369	649
Uterine Corpus	345	579
Urinary Bladder	327	622
Thyroid	232	356
Pancreas	221	410
Leukemia	203	365
Ovary	146	246

## Estimated Total Number of Precursor Lesions

Lesions	Total Number of Expected Precursor Lesions	
	Retrospective at Baseline	Prospective at 10 Years of Follow Up
Colorectal Adenomas	5,778	18,311
Benign Breast Biopsies (females only)	2,339	4,942
Dysplastic Nevus	2,311	3,779
Cervical Biopsies (females only)	2,137	2,822
High-Grade Prostatic Intraepithelial Neoplasia (males only)	238	384
Lung Nodules	165	349



# Data System Infrastructure & Digital Tools

Designed to maximize F.A.I.R. principles

# F.A.I.R. Data Ecosystem Makes F.A.I.R. Data

## Principles



### Findable

- Unique data identifiers and metadata that is searchable using web-based tools



### Accessible

- Retrievable data by identifiers based on a shared governance model



### Interoperable

- Data architecture that allows for integration of data from a range of sources and applications



### Reusable

- Goal of F, A, and I to allow reuse of data, code, and tools

## Implementation in Connect

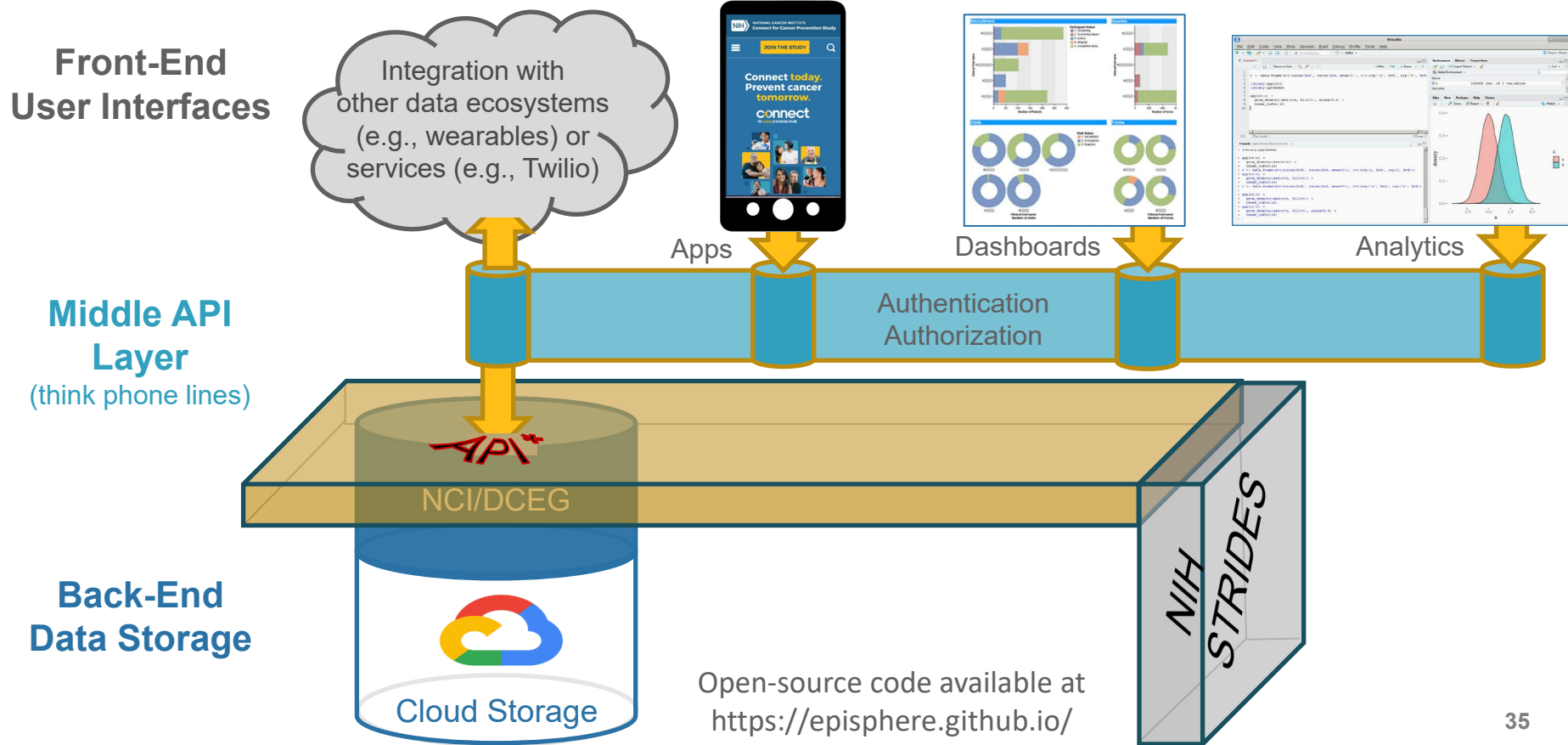
- Online, searchable data dictionary
- Map variables to common thesaurus to aid pooling (planned)

- Authentication and authorization parameters to enforce governance
- Online governance and instructions for data access (when data available)

- Cloud-based data available through governed APIs

- Data, metadata, code, tools and infrastructure available through public, versioned, open-source code on [GitHub](https://github.com)

# Connect Data Ecosystem provides an interoperable infrastructure





# Example of Front-End User Interface: MyConnect Participant App


## Study Activity Communication

### Messages

**New 2** **Read 1**

**Complete Your Profile** 6 days ago  
Please complete your profile section so you can move on to surveys!  
**Start**

**Sign Your Consent Forms** 8 days ago  
If you would like to participate in this study, please sign the consent forms.  
**Start**

 **NATIONAL CANCER INSTITUTE**  
Connect for Cancer Prevention Study

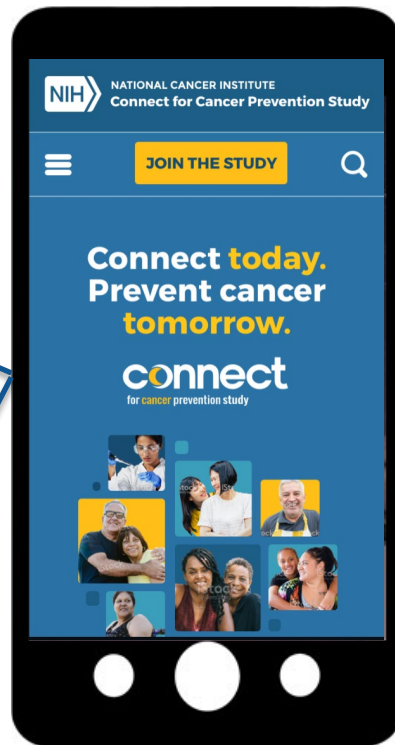
**Sign Out**

**Surveys** Samples Agreements About the Study What to Expect Privacy My Profile Messages


**1** Introduction **2** Background Information 3 Medical History 4 General Health 5 Family History 6 Education and Occupation

How often do you have children in the household from birth to 5 years of age in your household?  
☒ Every day, at least 2 hours each day  
☐ Every day, less than 2 hours each day  
☐ Most days of the week  
☐ One Day per week  
☐ A few days per year  
☐ Never  
**Clear**

## Survey Completion



## Informed Consent

 **NATIONAL CANCER INSTITUTE**  
Connect for Cancer Prevention Study

**MENU**


**7** of 8 **Consent**

**Informed Consent**

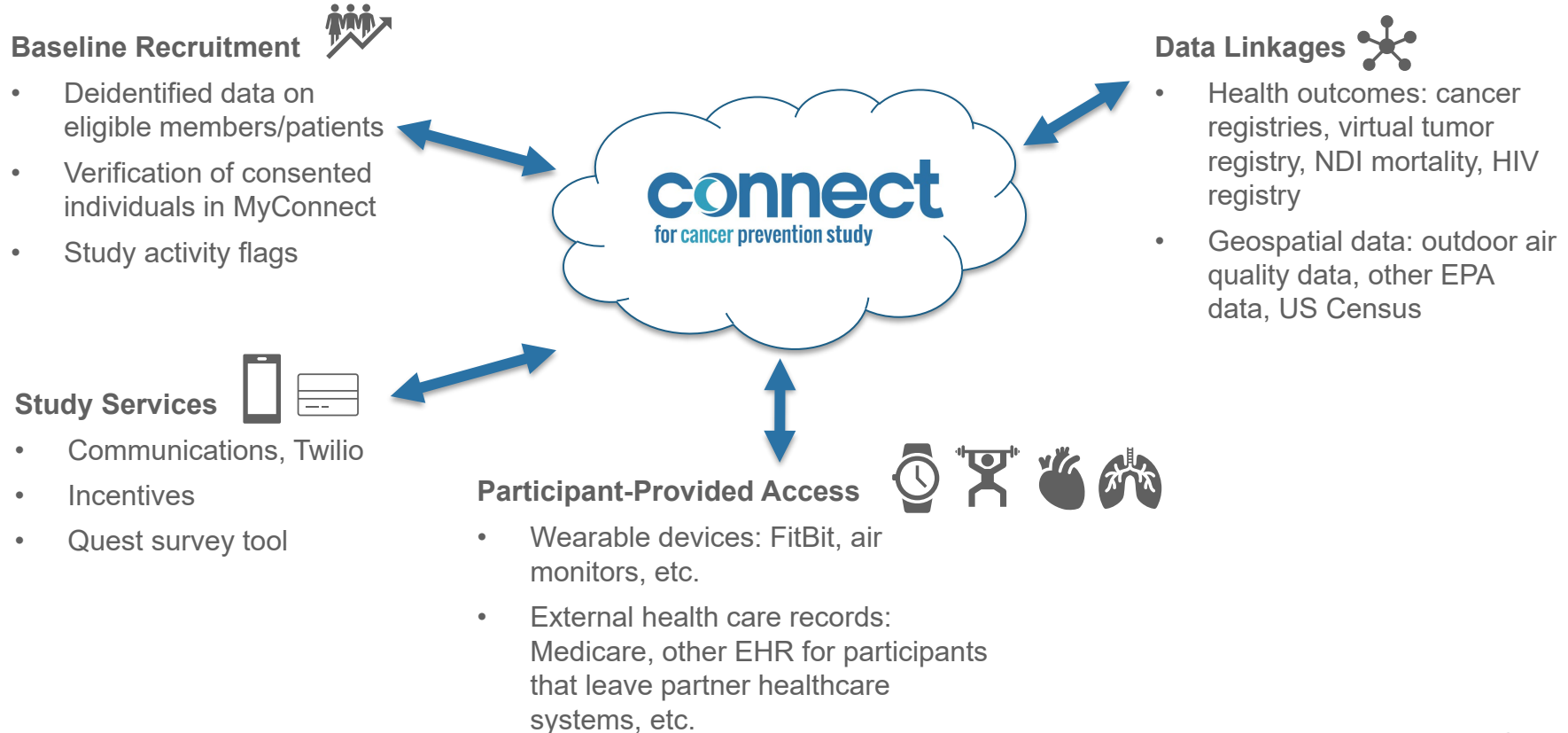
In order to join Connect, we need to you to give your informed consent by reviewing the full consent form and electronic health records release form, and signing your name.

**Informed consent form**

This is a more detailed explanation of what it means to take part in Connect.

 [Download an unsigned copy of the consent form](#)

# Benefit of Middle Layer of APIs: Pulling from multiple resources



# Quest provides open-source code to render surveys into MyConnect and capture data into Connect's backend data

## Quest

[Quest.js](#) | [Wiki](#) | [Issues](#) | [Project Page](#) | [Gitter](#)

### Questionnaire Options

Choose File No file chosen

Enter URL:

File Name:  .txt

Upload Questionnaire:

### Styling

☐ No Style ☒ With Style 1

### Logic

☐ Not Active ☒ Active

### Previous Results

```
{
  "firstName": "Daniel",
  "age": "51"
}
```

added json...

### Markup (Demo)

Change Font Size

```
{
  "name": "D_726699695"
} [INTROM1] Welcome, {${firstName}}! This survey is split into sections. Each section has questions that ask you about a wide range of topics. Our goal is to collect information about your medical history, family, work, and health behaviors. You can answer all of the questions in each survey section at one time, or answer some questions, pause, and return to answer the rest later. If you pause, your answers will be saved and you can pick up where you left off. You can also skip any questions that you do not want to answer. For some questions, you may see a word or phrase that appears as a button. Clicking the button will show more information that might help you answer the question. Here is an example of how additional information will be displayed. Let's get started. [INTROBAC] First, we are interested in learning some general information about you, your medical history, and your family history. This information will help us better understand your
```

### Rendering

To start, please tell us a bit about yourself.

Based on the information you provided when you enrolled in this study, you are 51 years old today. Is that correct?

☐ NO

☒ YES

## NATIONAL CANCER INSTITUTE

### Connect for Cancer Prevention Study

[Surveys](#) | [Samples](#) | [Agreements](#) | [About the Study](#) | [What to Expect](#) | [Privacy](#) | [My Profile](#) | [Messages](#)

1 Introduction 2 **Background Information** 3 Medical History 4 General Health 5 Family History 6 Education and Occupation

How often do you have children in the household from birth to 5 years of age in your household?

☒ Every day, at least 2 hours each day

☐ Every day, less than 2 hours each day


☐ Most days of the week

☐ One Day per week

☐ A few days per year

☐ Never

NCI/DCEG



# Governance & Organizational Structure

# Connect Governance

## Connect Leadership and Coordinating Center

### **Governing Board**

Stephen Chanock (Chair)

### **Executive Leadership**

#### **Steering Committee**

Chief Scientist (Chair), DCEG Connect PIs,  
Coordinating Center Leadership, DCEG Associate  
Directors

#### **Executive Committee**

Chief Scientist (co-Chair), Rotating Site PI (co-  
Chair), Steering Committee Members, Site PIs,  
Participant Advisory Board Lead,  
Support Service Contractor Leads

## External Advisory and Review Groups

### **International Scientific Advisory Board**

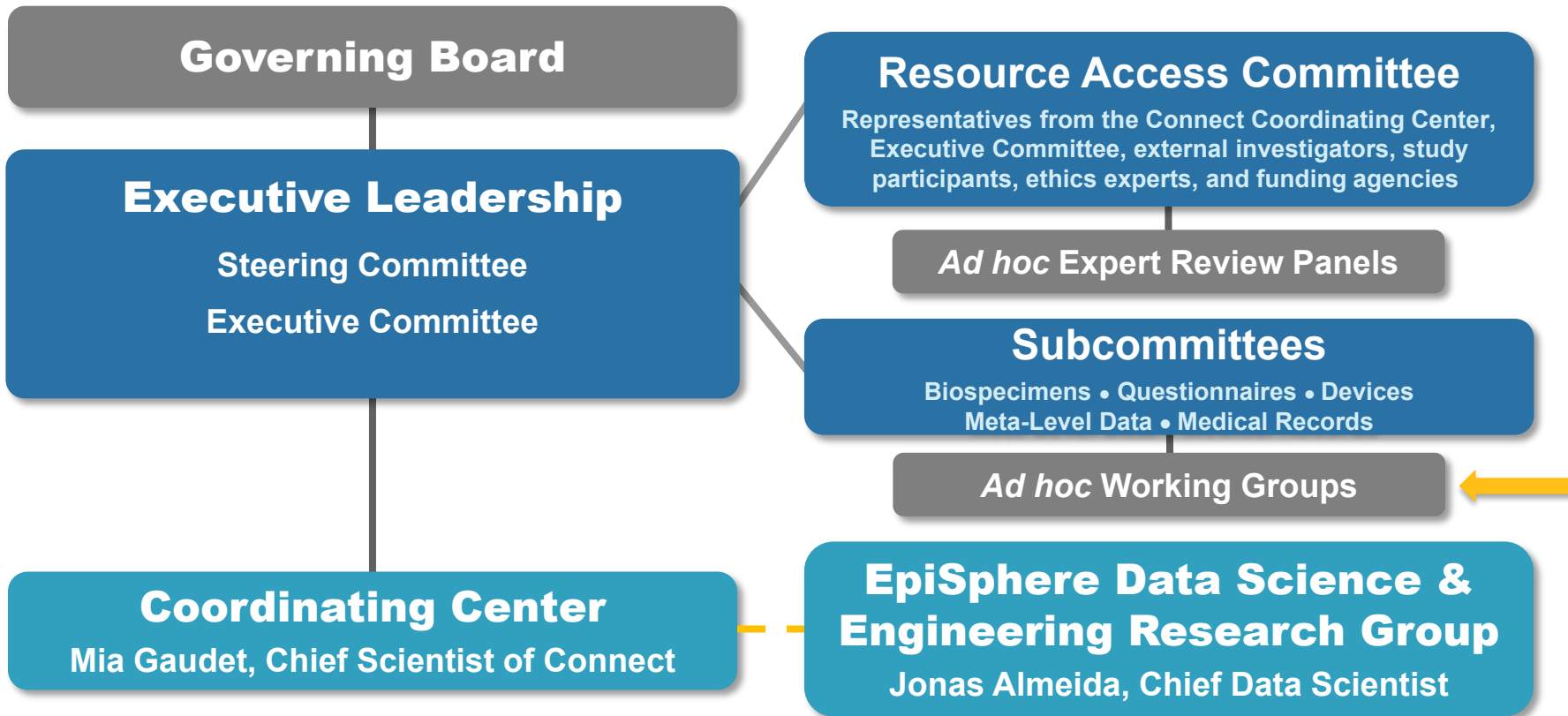
Naomi Allen, Eric Boerwinkle, Julie Buring, Carmen  
Guerra, Chris Haiman, David Hunter, Roger Milne,  
Alpa Patel, Karriem Watson

### **National Cancer Institute (NCI) Board of Scientific Counselors (BSC) – Clinical Sciences and Epidemiology**

### **Participant Advisory Board (PAB)**

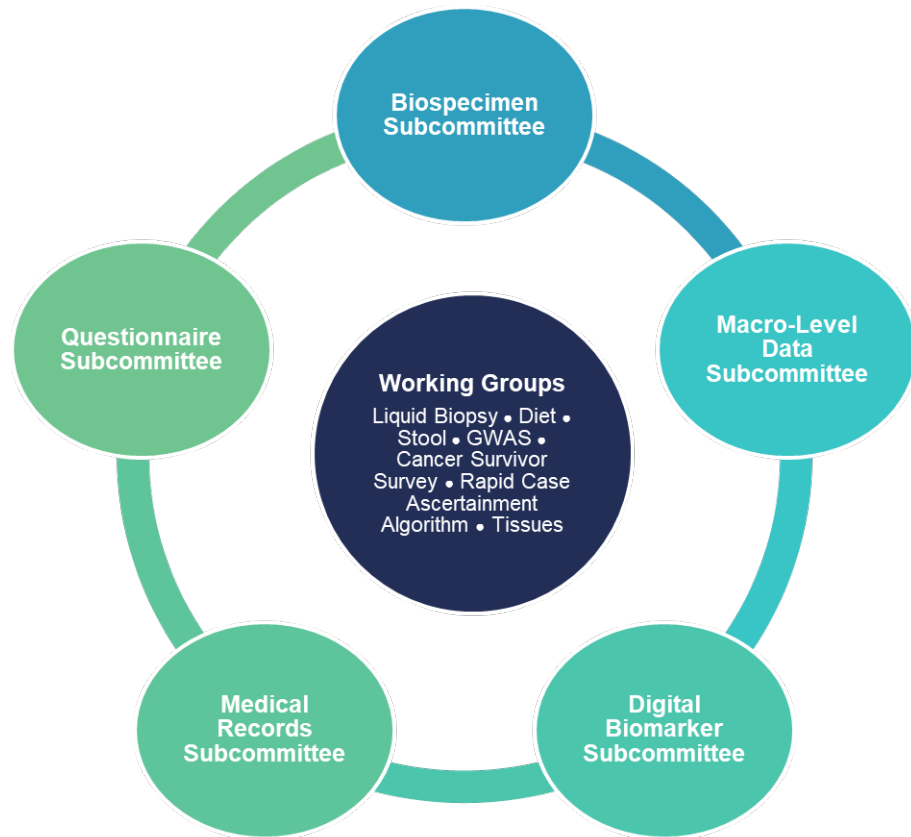


# Connect Leadership and Organization



# Connect Resource Development Structure

- Subcommittees (standing)
  - Guide development of Connect infrastructure and enhancements
  - Led by DCEG with NIH investigators
  - Report to Steering Committee
- Working Groups (*ad hoc*)
  - Led by nominator with investigators from NIH, recruitment sites, and scientific community
  - Develop scientific and technical details of proposals and protocols
  - Report to relevant Subcommittees



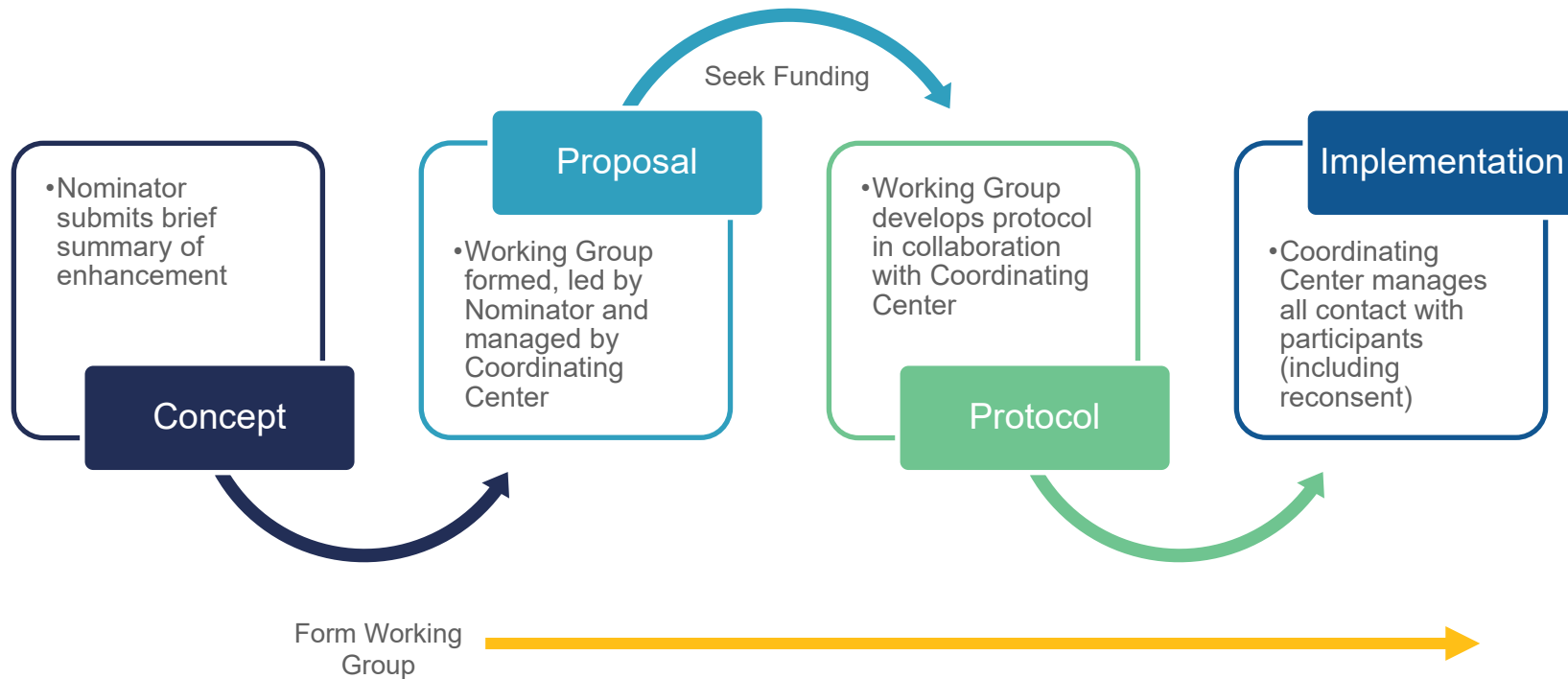
# Resource Enhancements

New questionnaire modules, biospecimens, other sources of data

## Connect Resource Building Principles



# Connect Resource Enhancement Approval Process

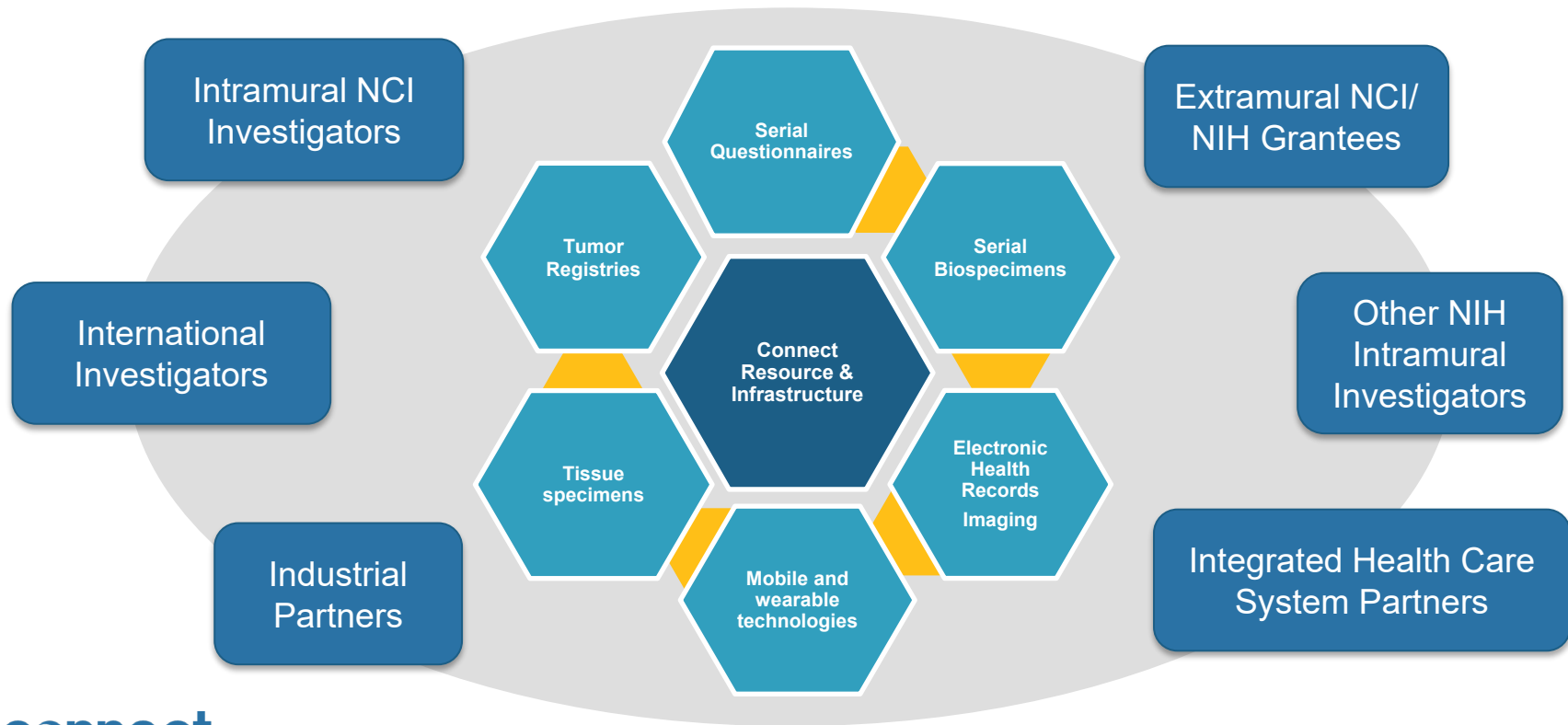


# Resource Access Principles

## Connect Resource Access Principles

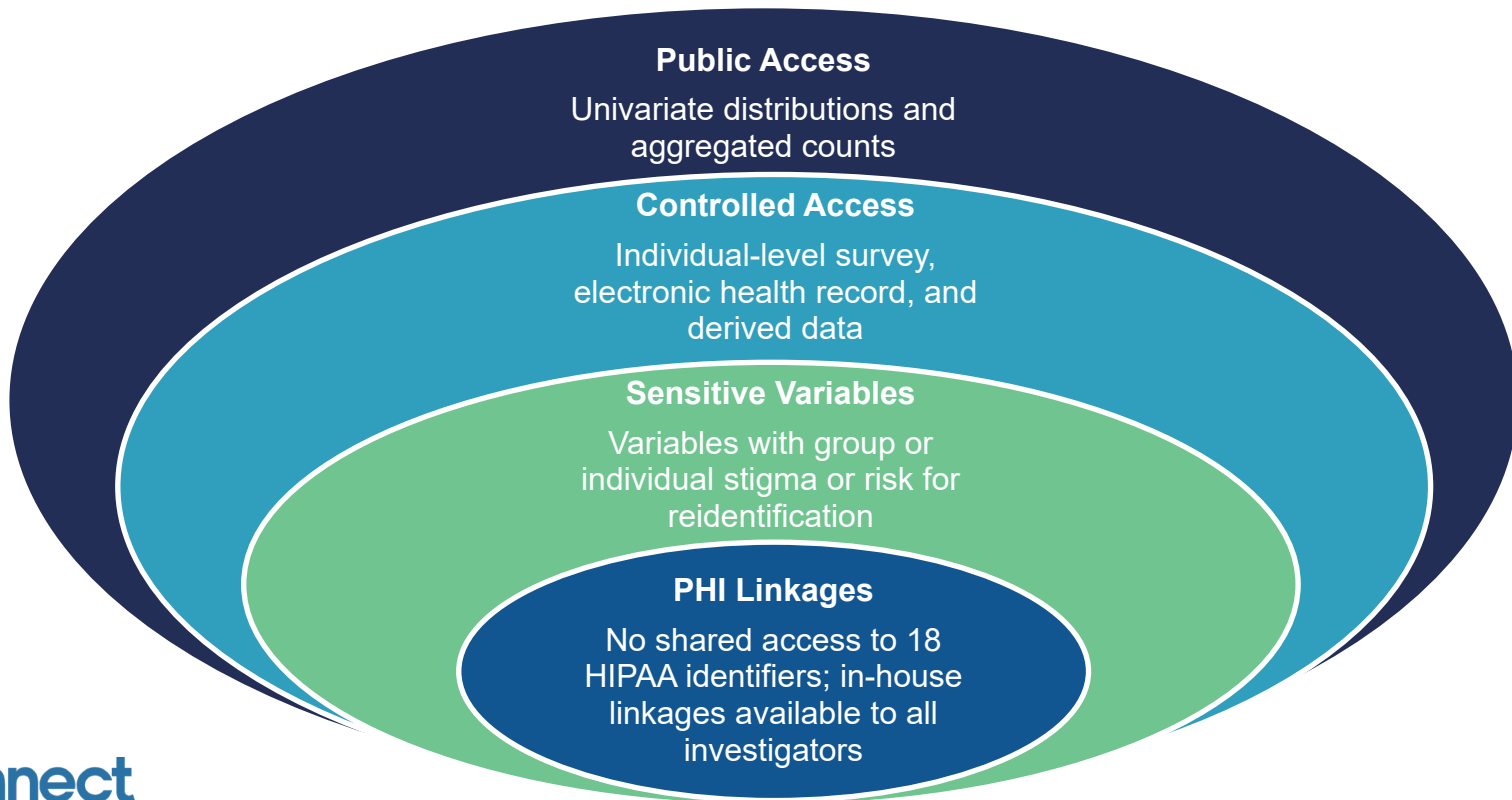


# Resource Access for All\*



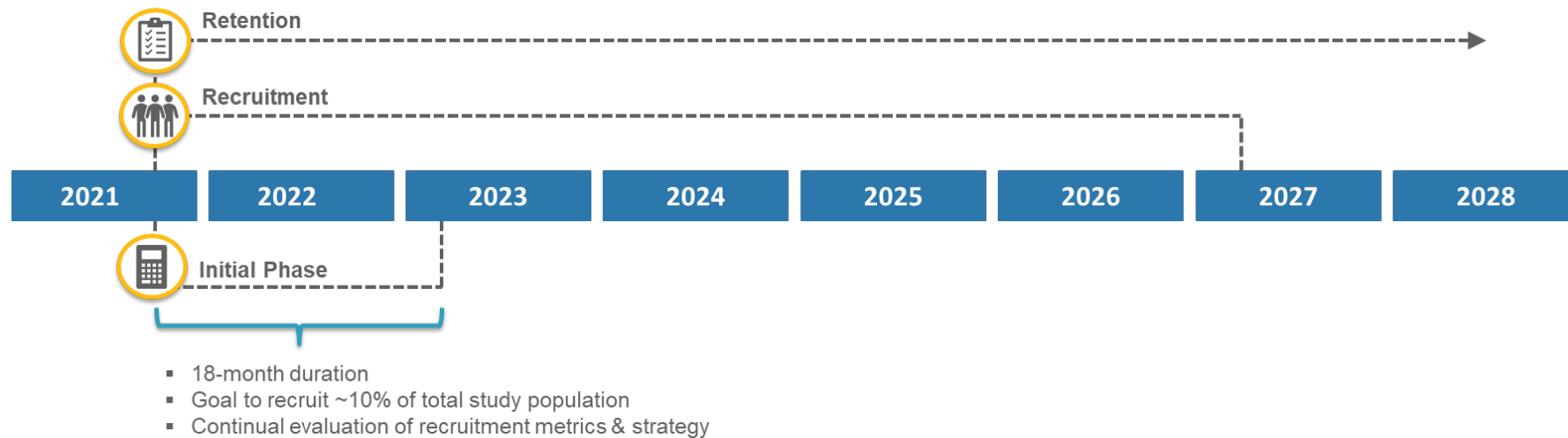


# Levels of Data Access

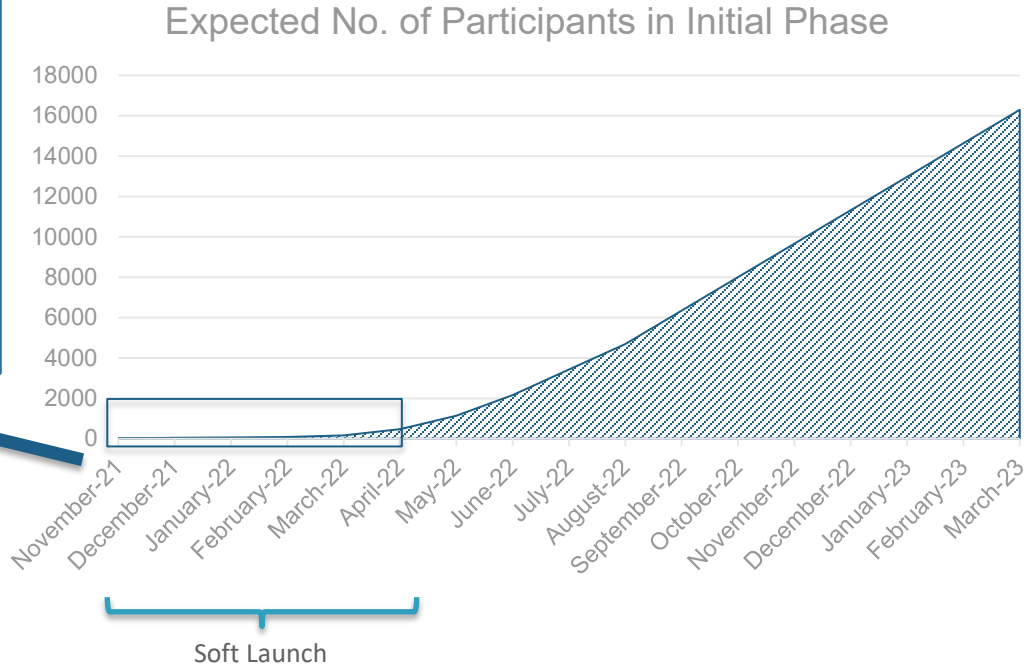
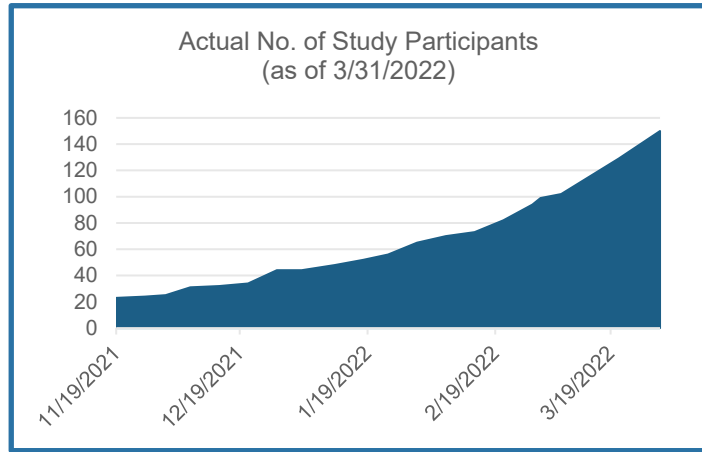


# Recruitment Timeline

# Recruitment Began!



## Initial Phase of Recruitment: slow but steady to reach 16,000



connect  
for cancer prevention study