



Cohort Overview

Mia M. Gaudet, Ph.D.

Connect Chief Scientist

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



September 5, 2023

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

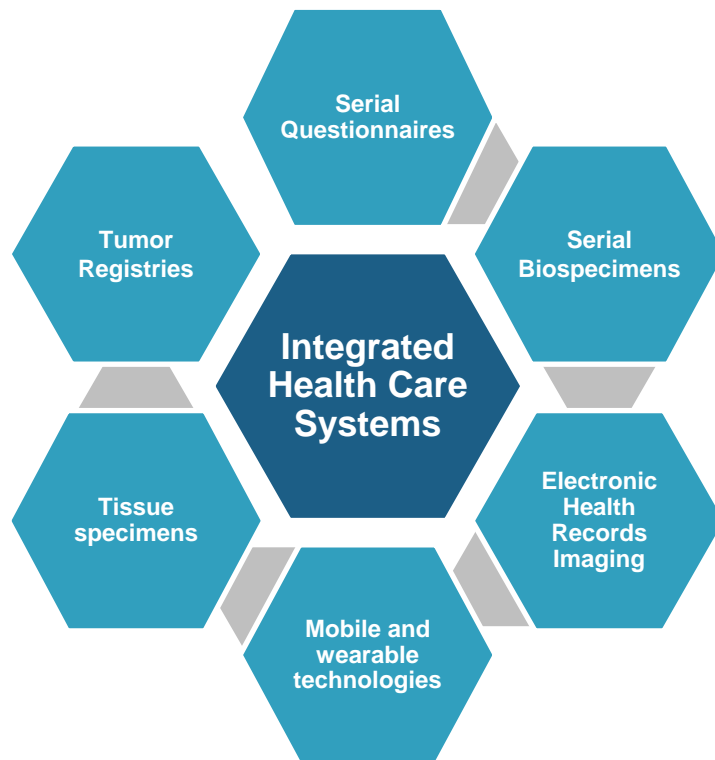


Diverse and special populations

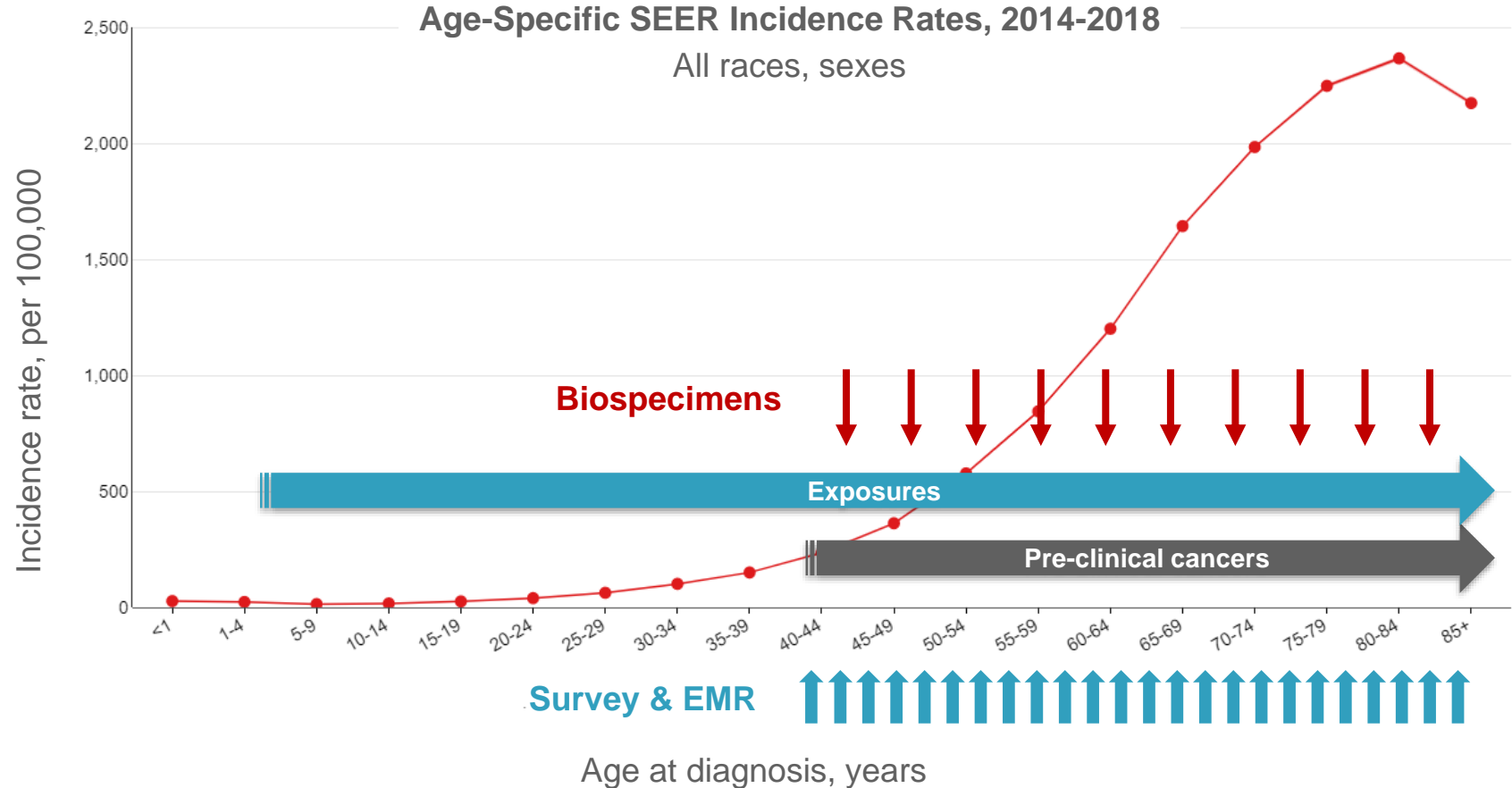
Cohort Study Design Overview & Recruitment Update

Key Study Design Features of Connect

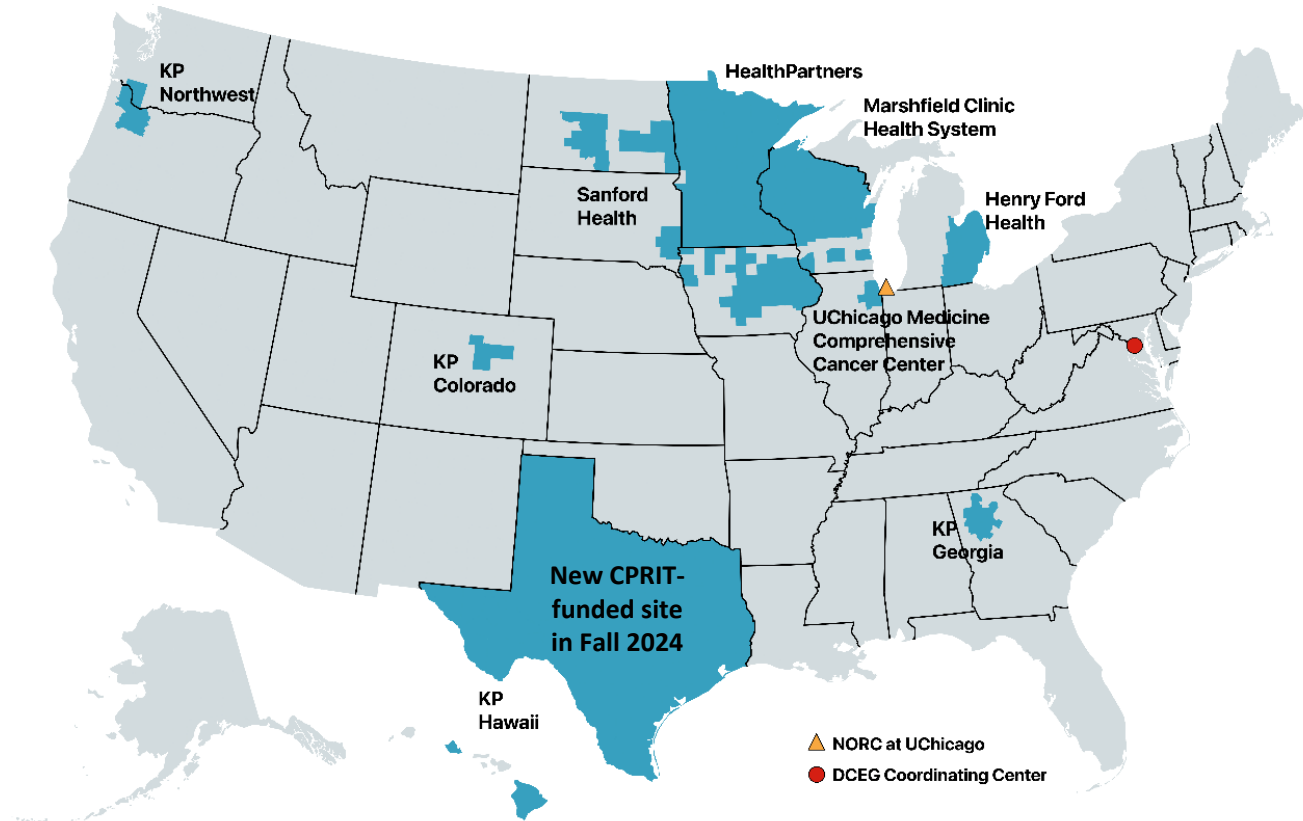
- ▶ 200,000 adults across the US
 - ✓ Aged 40-65 years
 - ✓ No history of cancer
 - ✓ Patients or members of partner health care systems
- ▶ Long-term follow-up
- ▶ Serial assessments
- ▶ Comprehensive cancer outcomes
- ▶ Flexible infrastructure for enhancement studies



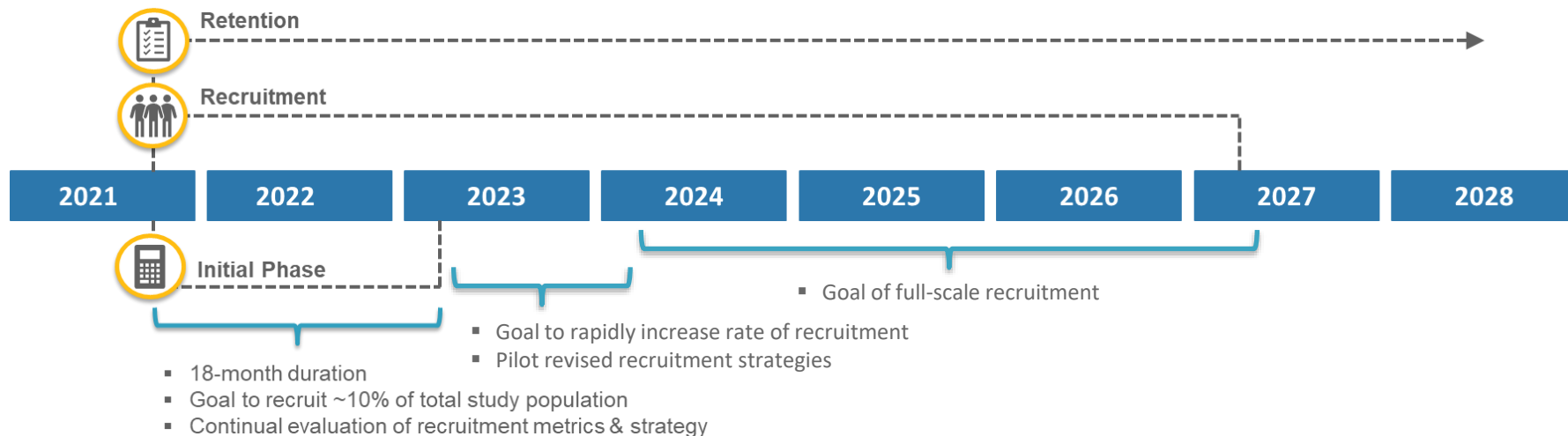
Serial Assessments to Observe Longitudinal Changes



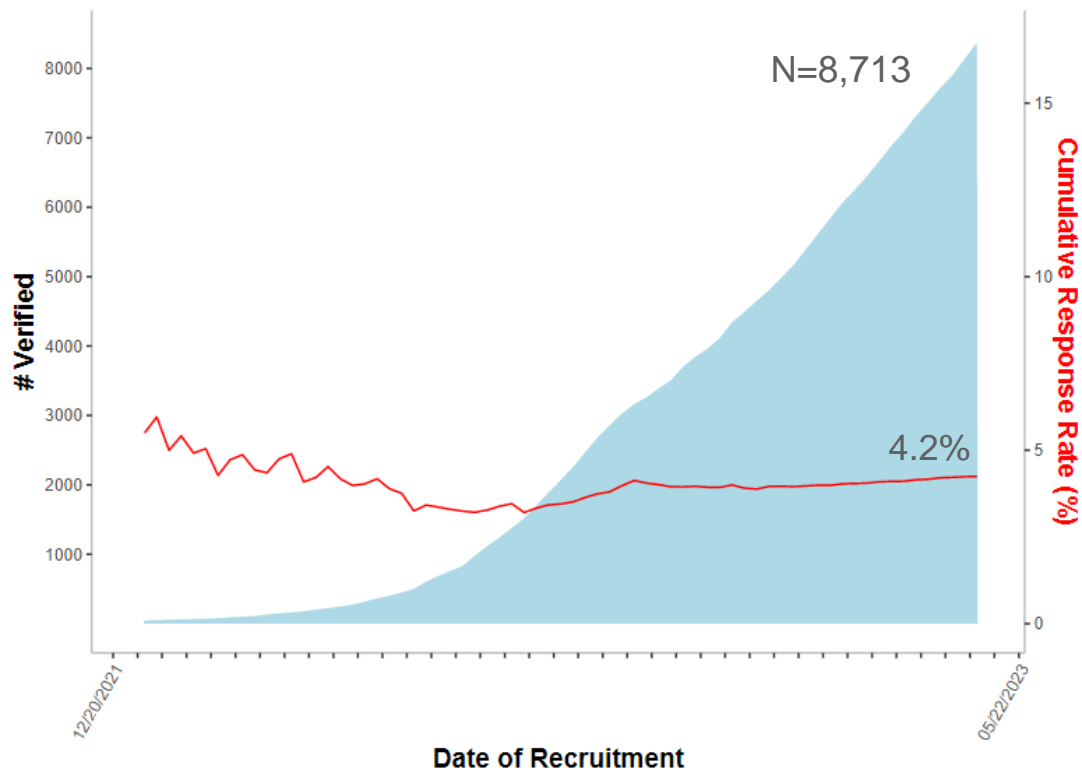
Expansion of recruitment to CPRIT-funded site expected to increase diversity of catchment population



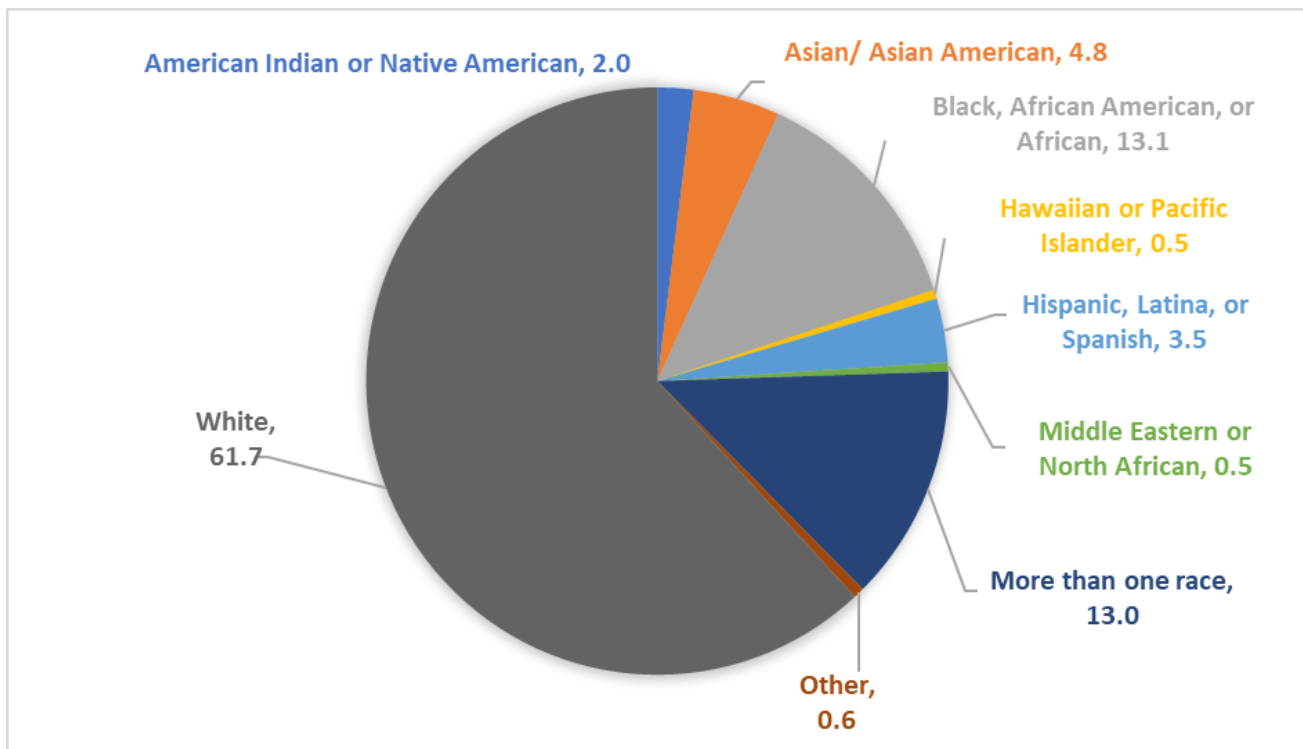
Currently, scaling up recruitment



Current Number of Verified Participants and Response Proportion



Distribution of Race & Ethnicity (%) Among Connect Verified Participants



Distributions among Connect participants in comparison with other populations

Demographic/ Risk Factor	Connect Participants	Compared with US Representative Data ¹	Compared with Cohorts of Similar Birth Cohorts ²
Males	30%		
Gender Minorities	<1%		
Sexual Minorities	7%		
Black	14%		
Hispanic	3%		
Asian/ Pacific Islander	5%		
College graduate or higher	62%		
Income, \$100k+	46%		
Obesity	36%		
Alcohol, never consumed	5%		
Tobacco user, current	8%		

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

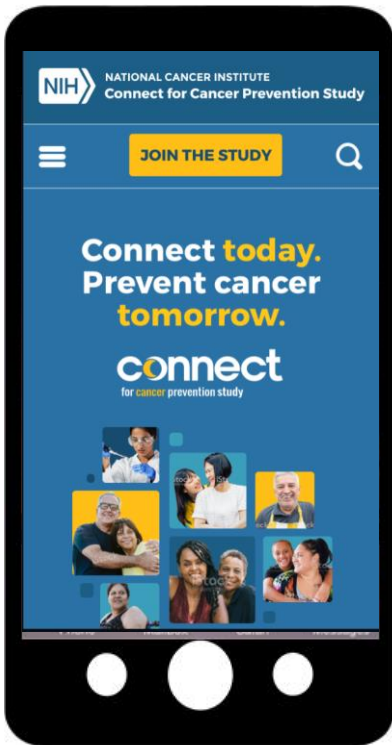
MyConnect Participant App: Communication and Study Activity Hub

Existing Functionality

Reminders & Communication



Study Activities



Planned Functionality


Return of Results



Mobile/ Wearable Technologies



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.

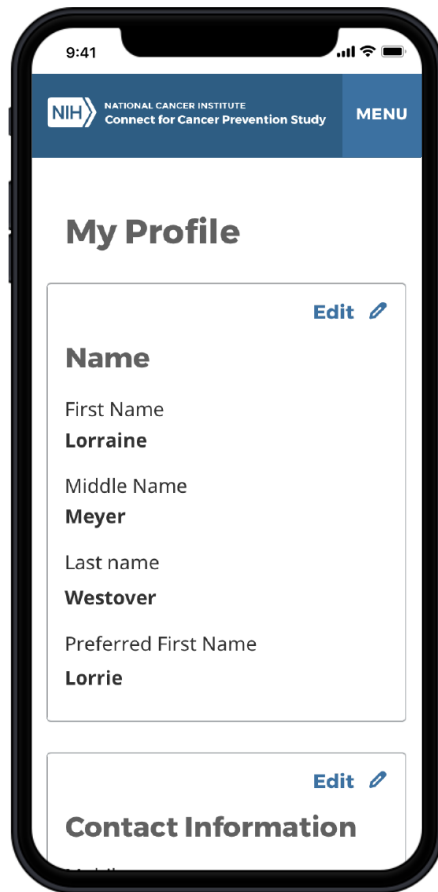
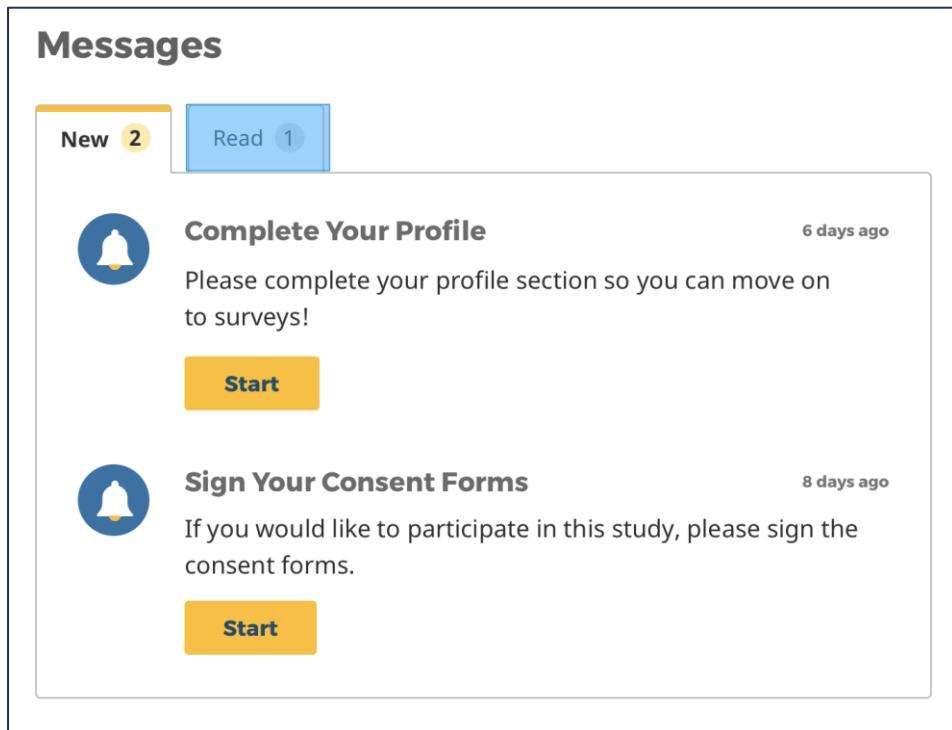
Informed consent form

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



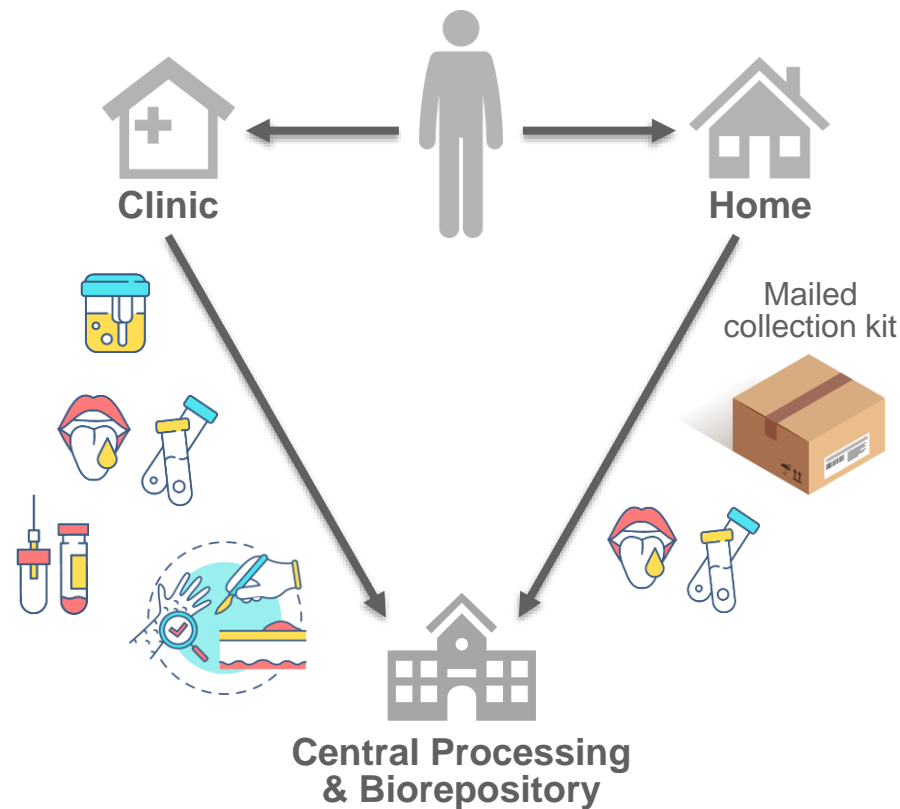
[Download an unsigned copy of the consent form](#)

Messages and Data Input



Biospecimen Collections

Biospecimen Collections at Recruitment



	No. Tubes	Fraction	Tube Type
Buccal	1		Saliva-mouthwash kit
Blood	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
Urine	1		10ml Urine kit
Tissue			Precursor tissue blocks, if available

Biospecimen Collections at Follow-up

- Timed on average every 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

Baseline 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

First Survey

This survey is split into four sections that ask about a wide range of topics, including information about your medical history, family, work, and health behaviors. You can answer all of the questions at one time, or pause and return to complete the survey later. If you pause, your answers will be saved so you can pick up where you left off. You can skip any questions that you do not want to answer.



Background and Overall Health

Questions about you, your medical history, and your family history.

Start

Estimated Time: 20 to 30 minutes



Where You Live and Work

Questions about places where you have lived and worked, and your commute to school or work.

Start

Estimated Time: 20 to 30 minutes



Medications, Reproductive Health, Exercise, and Sleep

Questions about your current and past use of medications, your exercise and sleep habits, and your reproductive health.

Start

Estimated Time: 20 to 30 minutes



Smoking, Alcohol, and Sun Exposure

Questions about your use of tobacco, nicotine, marijuana, and alcohol, as well as your sun exposure.

Start

Estimated Time: 20 to 30 minutes

Surveys Accompanying Biospecimen Collection

Blood & Urine

- Recent exposures to food, drink, pharmaceuticals
- Reproductive Health
 - Last menstrual period date



Menstrual Cycle

- Next menstrual period date

Mouthwash

- Oral health and care
- Recent exposures to toothpaste, mouthwash, gum, tobacco
- Permanent teeth lost

COVID-19

- Exposure
- Complications
- Long-COVID
- Vaccination

Surveys available in MyConnect

Baseline	Background & Overall Health	Demographics, medical history, family history of cancer, education & occupation
	Medications, Reproductive Health, Exercise, Sleep	OTC/rx, menstrual cycle, pregnancy, fertility, hormonal meds & contraception, physical activity, shiftwork
	Smoking, Alcohol, Sun Exposure	Tobacco/nicotine, marijuana, alcohol, tanning
	Where You Live & Work	Residential, seasonal, & childhood addresses, water source, tap water intake, work & school commute
	SSN	Full or partial SSN
Biospecimens	Blood, Urine, Mouthwash (Research Appt)	Symptoms/behaviors on day of donation, meds/antibiotics, recent reproductive health, oral hygiene & health
	Blood/Urine (Clinical Collection)	Symptoms/behaviors on day of donation, meds, recent reproductive health
	COVID-19	# Infections, symptoms, hospitalization, vaccination
	Menstrual Cycle Survey	Last period, triggered by research appt/clinical collection response

Surveys In Development

Participant Experience	MyConnect, study communications, activities completed, overall experience, motivation for joining
Cancer Screening History	Organ Inventory (born with/current), history of cancer screening tests
Cancer Diagnosis Surveys (17)	Dx, symptoms, Patient-Provider interaction, medical history repeat assessment
Menstrual & Intimate Care Products	Vaginoplasty, powder, douching, vaginal cleansing products, menstrual products
Fecal Collection	Donation, bowel movements, meds, supplements, probiotics
Social Determinants of Health (SDOH)	Discrimination, police interaction, medical mistrust, social support, and financial, food, and housing insecurity
Hair Products	Dyes, relaxers, straighteners, perms, oils
Menstrual Experience Survey	Menstrual problems, endometriosis dx and treatment
Mothballs & Scented Products	Household exposure to p-DCB & Napthalene

Detailed Status Available [Here](#)

Planned Survey Cadence

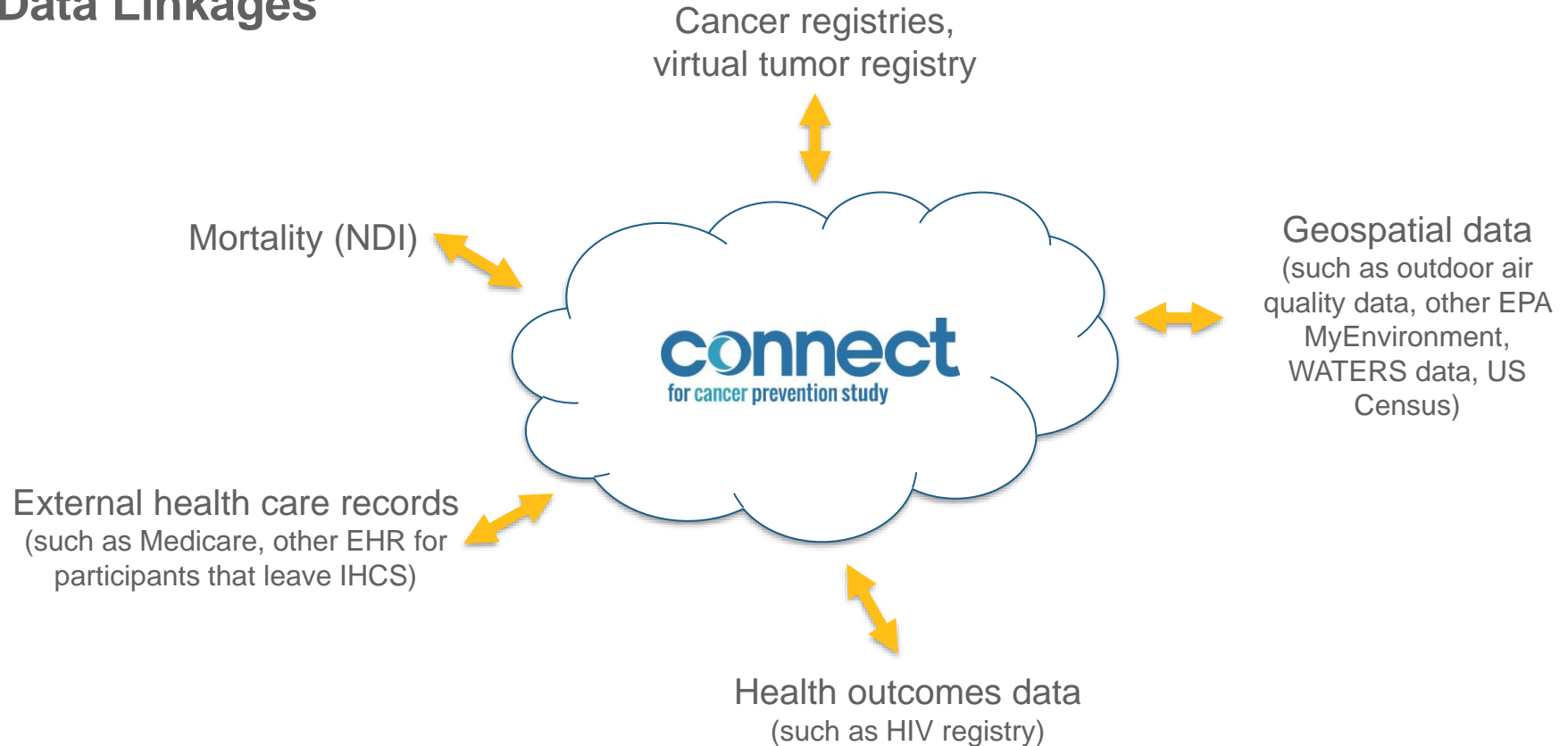
6 Months	Within 24 Months	12 Months	2 Years
<ul style="list-style-type: none">• Diet food frequency questionnaire (DHQ)	<ul style="list-style-type: none">• Repeated 24-hour recalls for diet (ASA24) and physical activity (ACT24)	<ul style="list-style-type: none">• Screening History• Quality of life (PROMIS)• Psychosocial• Social determinants of health	<ul style="list-style-type: none">• Baseline exposures repeat assessment• Consumer products (Menstrual & intimate care, Hair, mothballs & scented)• Menstrual experience (Endometriosis)

3-6 Months Post-Cancer Diagnosis Rapid Case Ascertainment

- Cancer Diagnosis Surveys (Dx, symptoms, patient-provider interaction, medical history)
 - Quality of life (PROMIS: repeat assessment & patient specific)

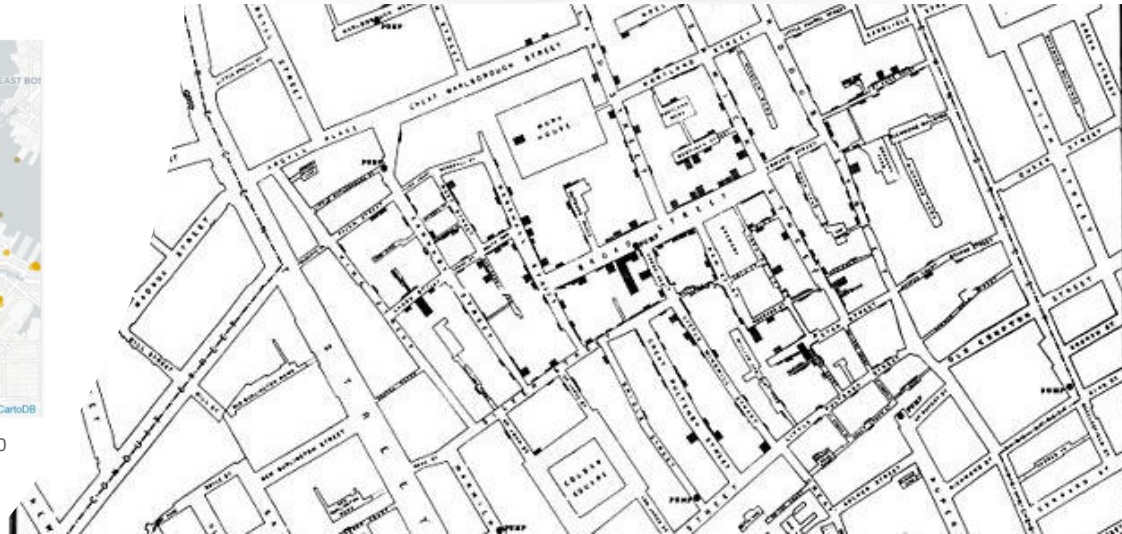
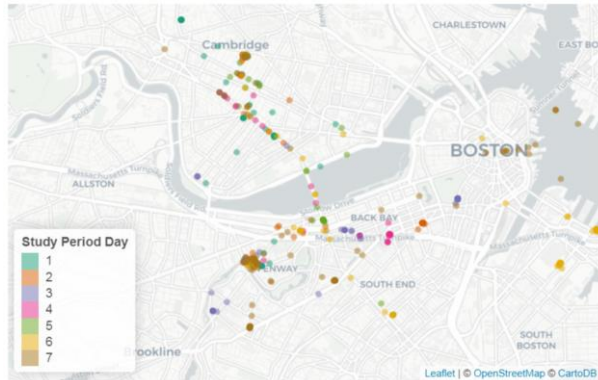
Data Linkages

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

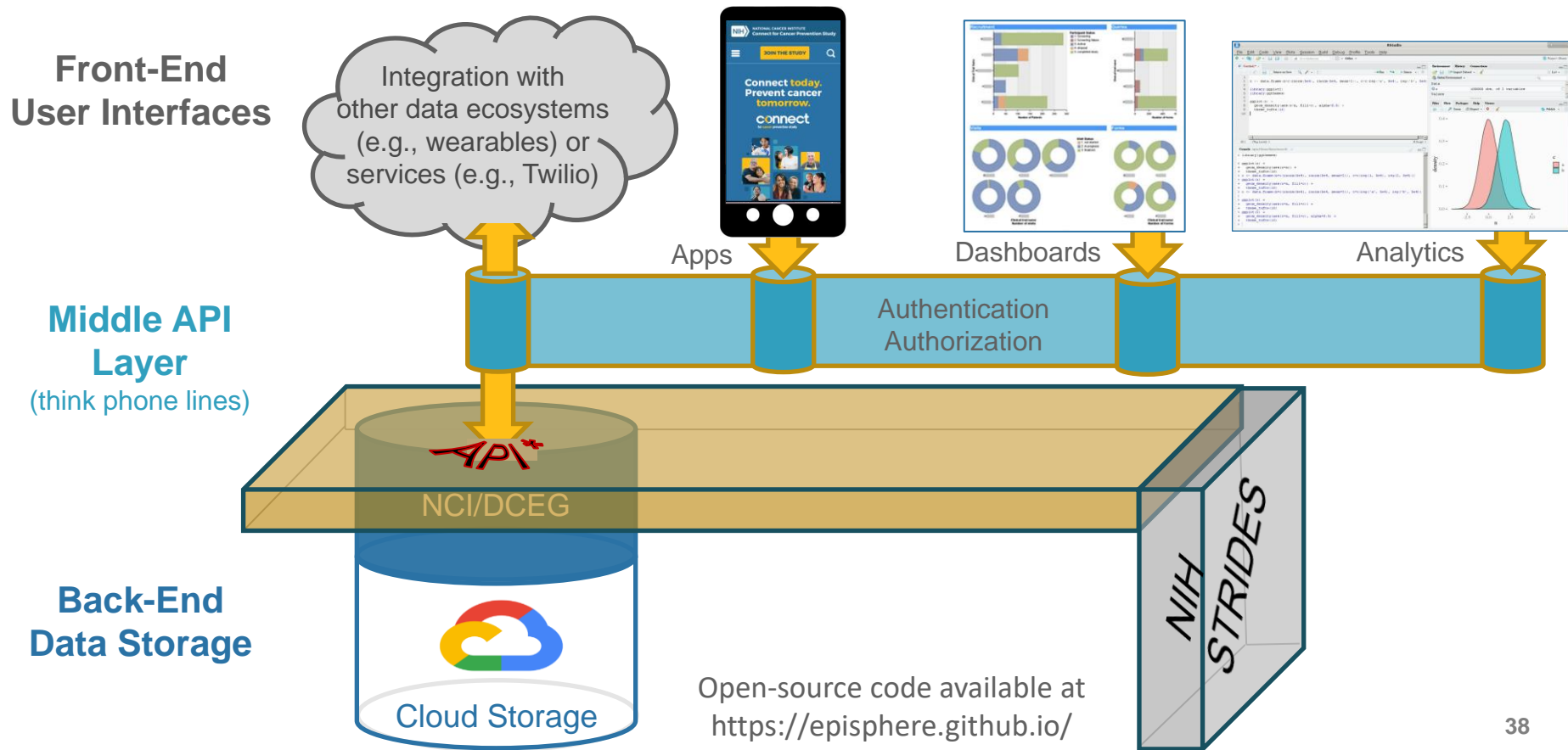
- Online, searchable data dictionary
- Map variables to common thesaurus to aid pooling (planned)

- 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](#)

Survey Completion

NIH 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

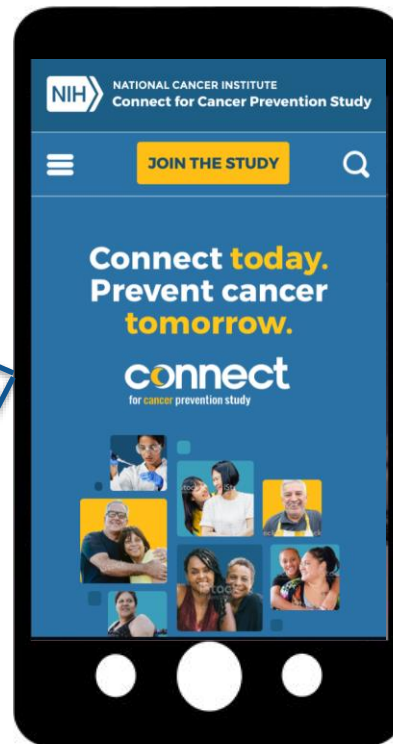
☐ Most days of the week

☐ One Day per week

☐ A few days per year

☐ Never

[Clear](#)



Informed Consent

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

7 of 8 **Consent**

Informed Consent

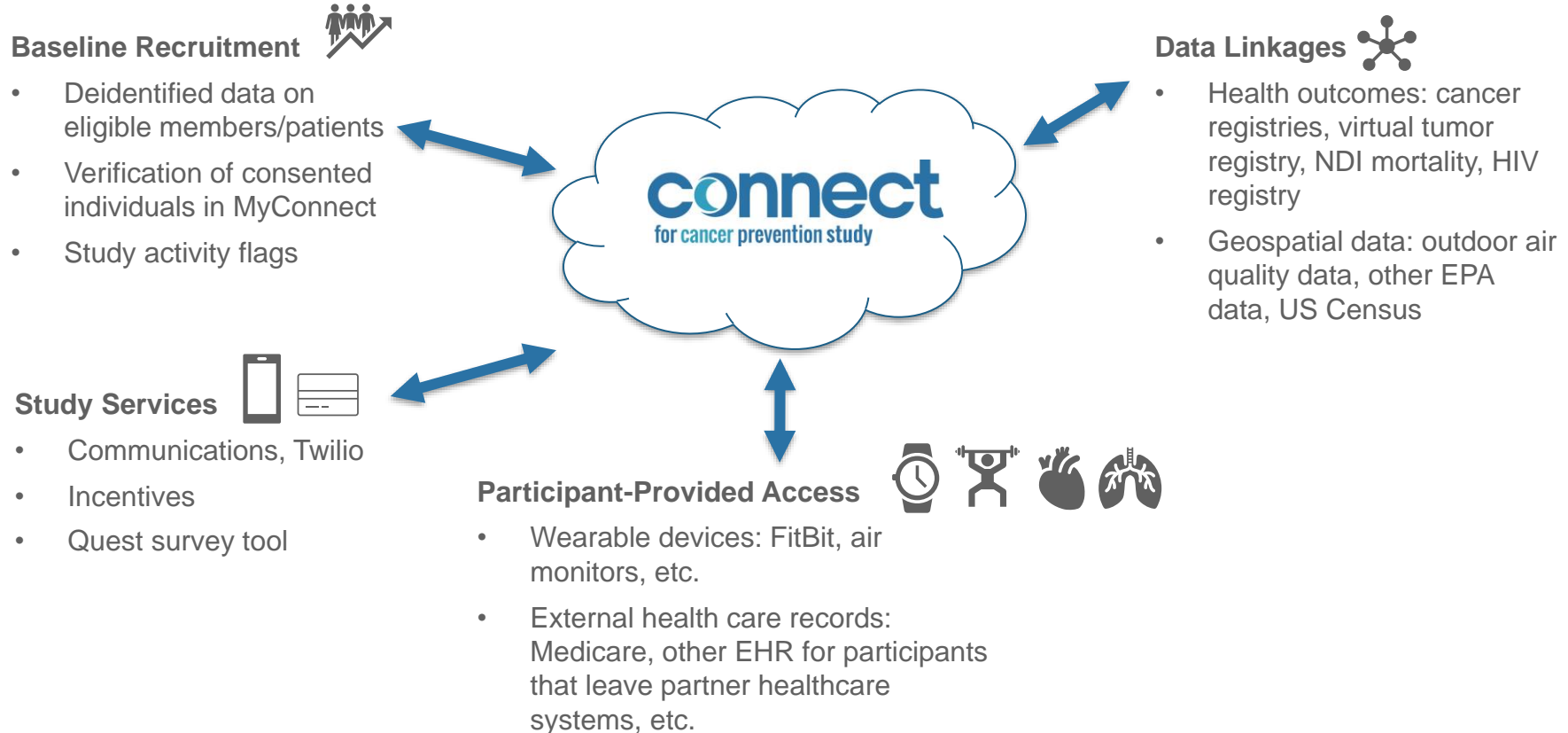
In order to join Connect, we need to you to give your informed consent by reviewing the full consent form and electronic health records release form, and signing your name.

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, {${u: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

[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

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

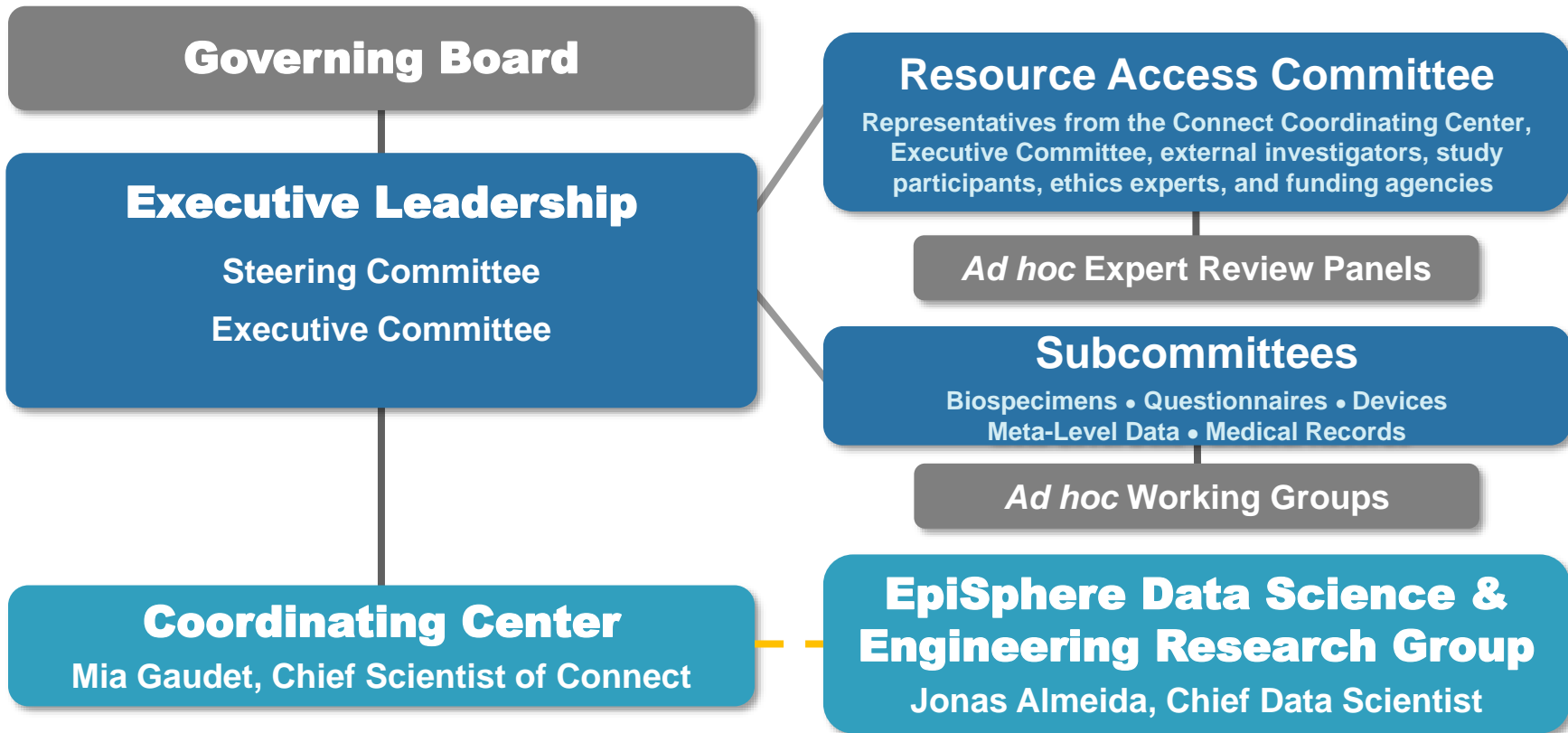
International Scientific Advisory Board

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

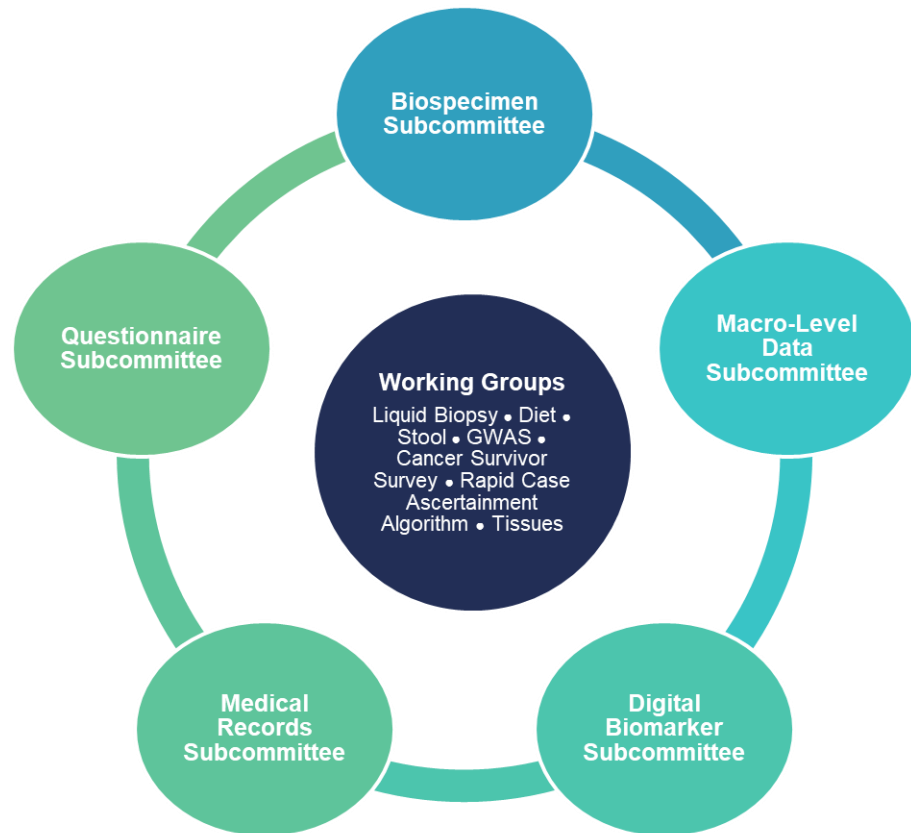
Participant Advisory Board (PAB)

Connect Leadership and Organization



Connect Resource Development Structure

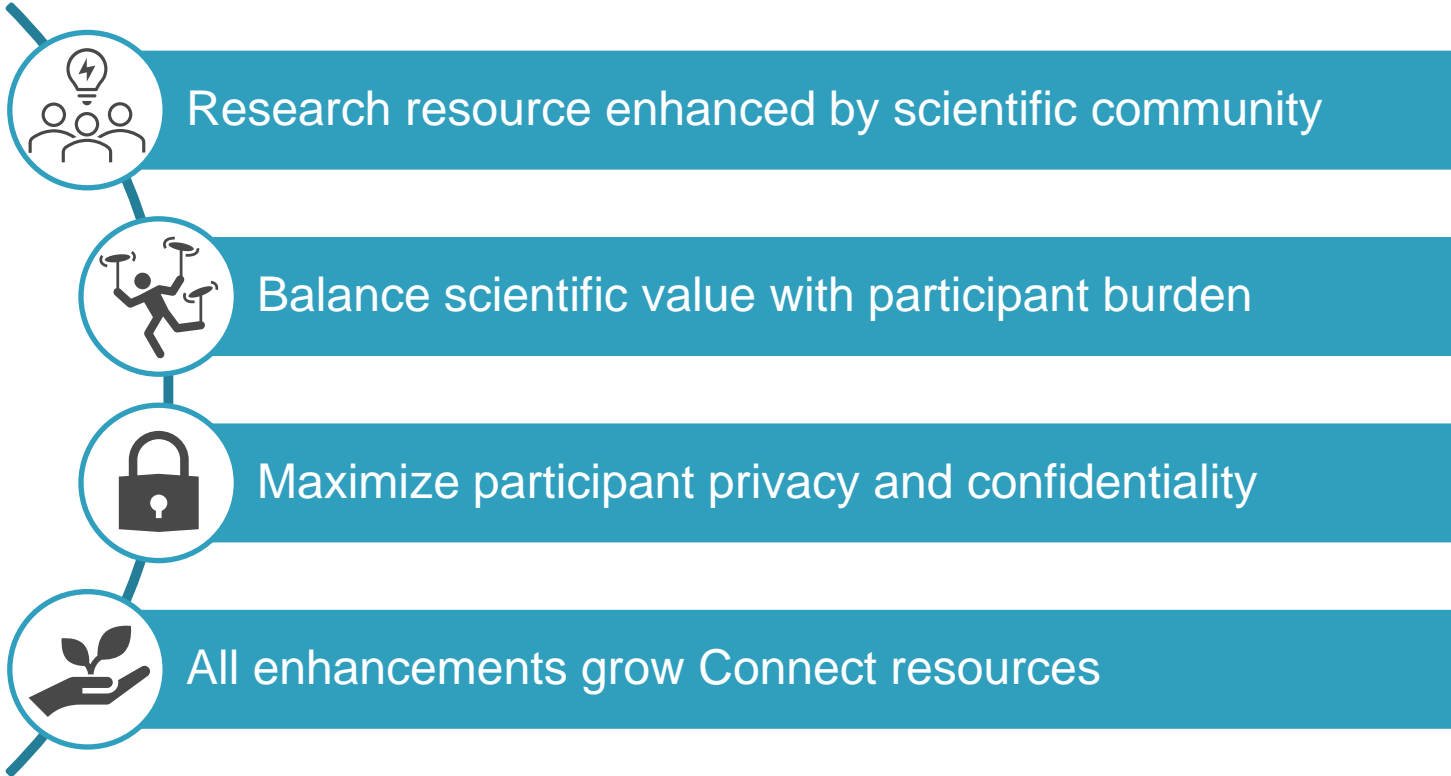
- Subcommittees (standing)
 - Guide development of Connect infrastructure and enhancements
 - Led by DCEG with NIH investigators
 - Report to Steering Committee
- Working Groups (*ad hoc*)
 - Led by nominator with investigators from NIH, recruitment sites, and scientific community
 - Develop scientific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



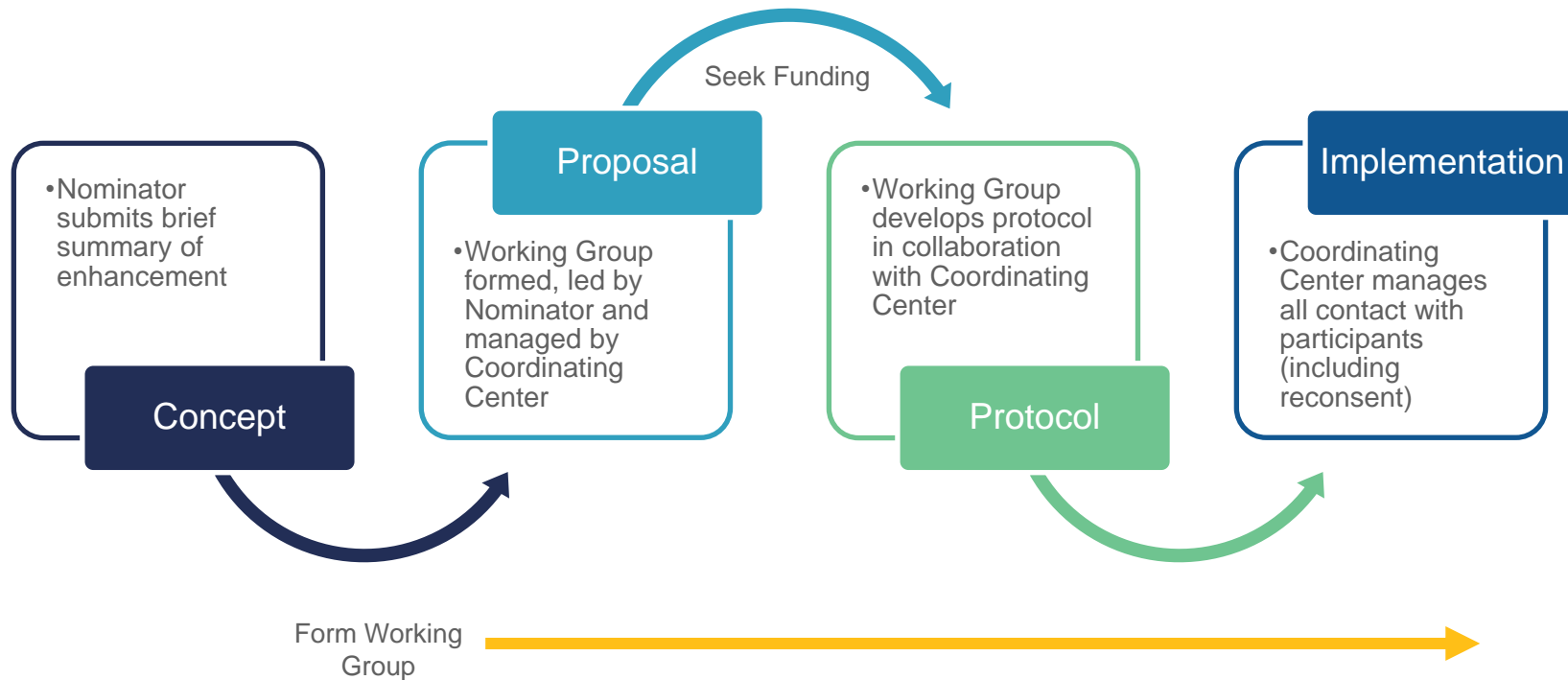
Resource Enhancements

New questionnaire modules, biospecimens, other sources of data

Connect Resource Building Principles



Connect Resource Enhancement Approval Process

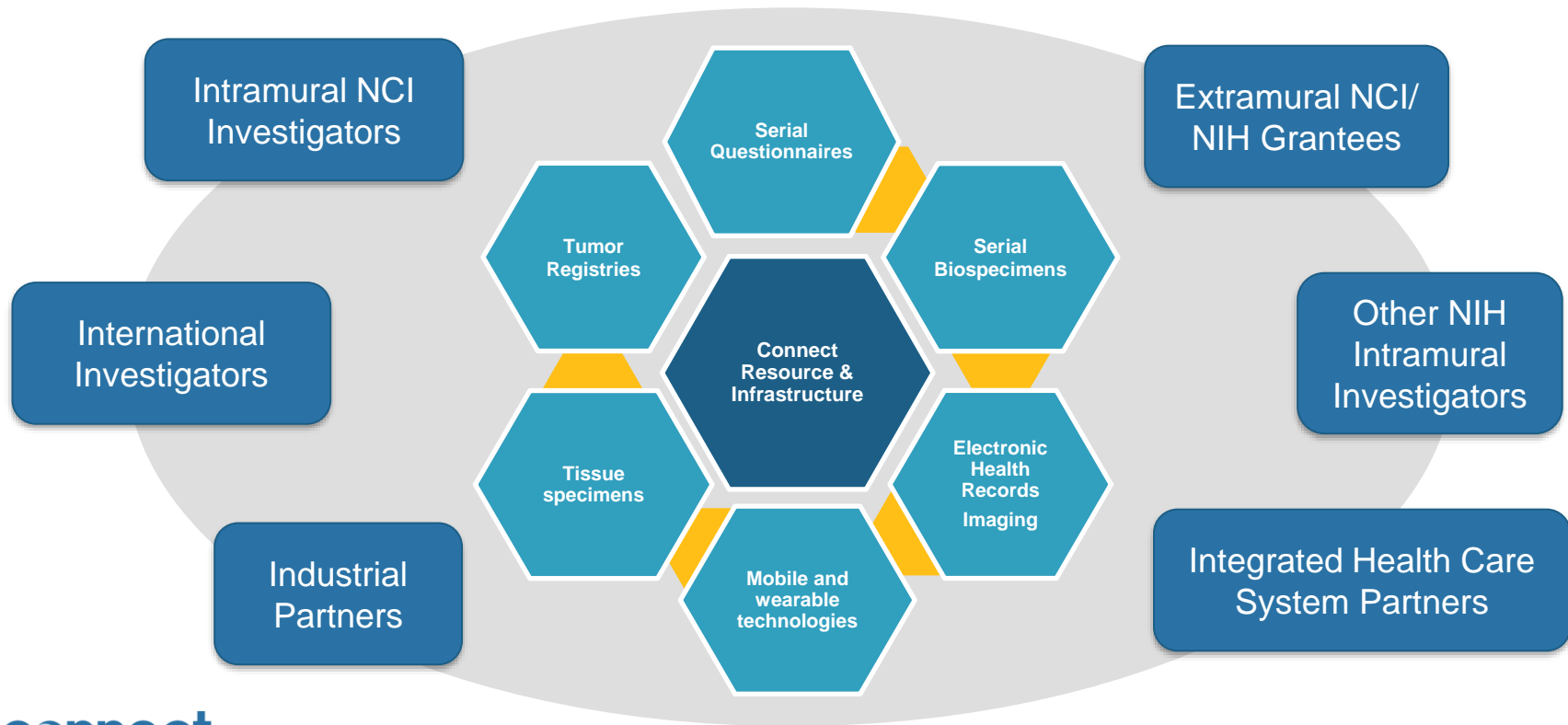


Resource Access Principles

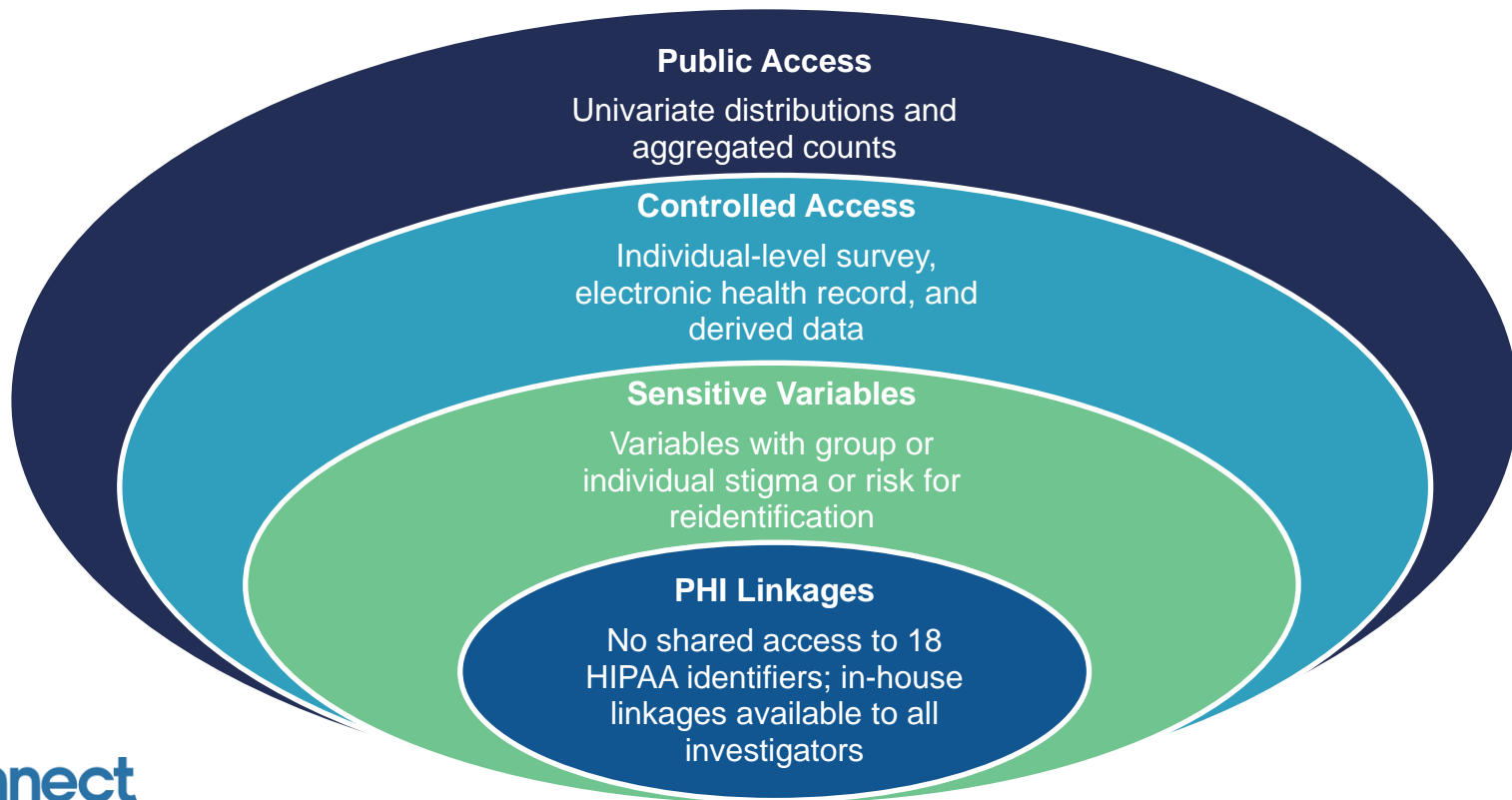
Connect Resource Access Principles



Resource Access for All*



Levels of Data Access



connect
for cancer prevention study