

# International Cohorts Summit

Summary of Cohorts in Attendance

Lena Dolman and Teri Manolio



# Introductions & Acknowledgements

## Program Committee



**Rory Collins**  
University of Oxford  
UK



**Geoff Ginsburg**  
Duke University  
USA



**Teri Manolio**  
NHGRI  
USA



**Masayuki Yamamoto**  
Tohoku Medical Megabank  
Japan

## G2MC Staff Team



Justina Chung



Lena Dolman



Ashley Hobb



Teji Rakhra-  
Burris

## Sponsors

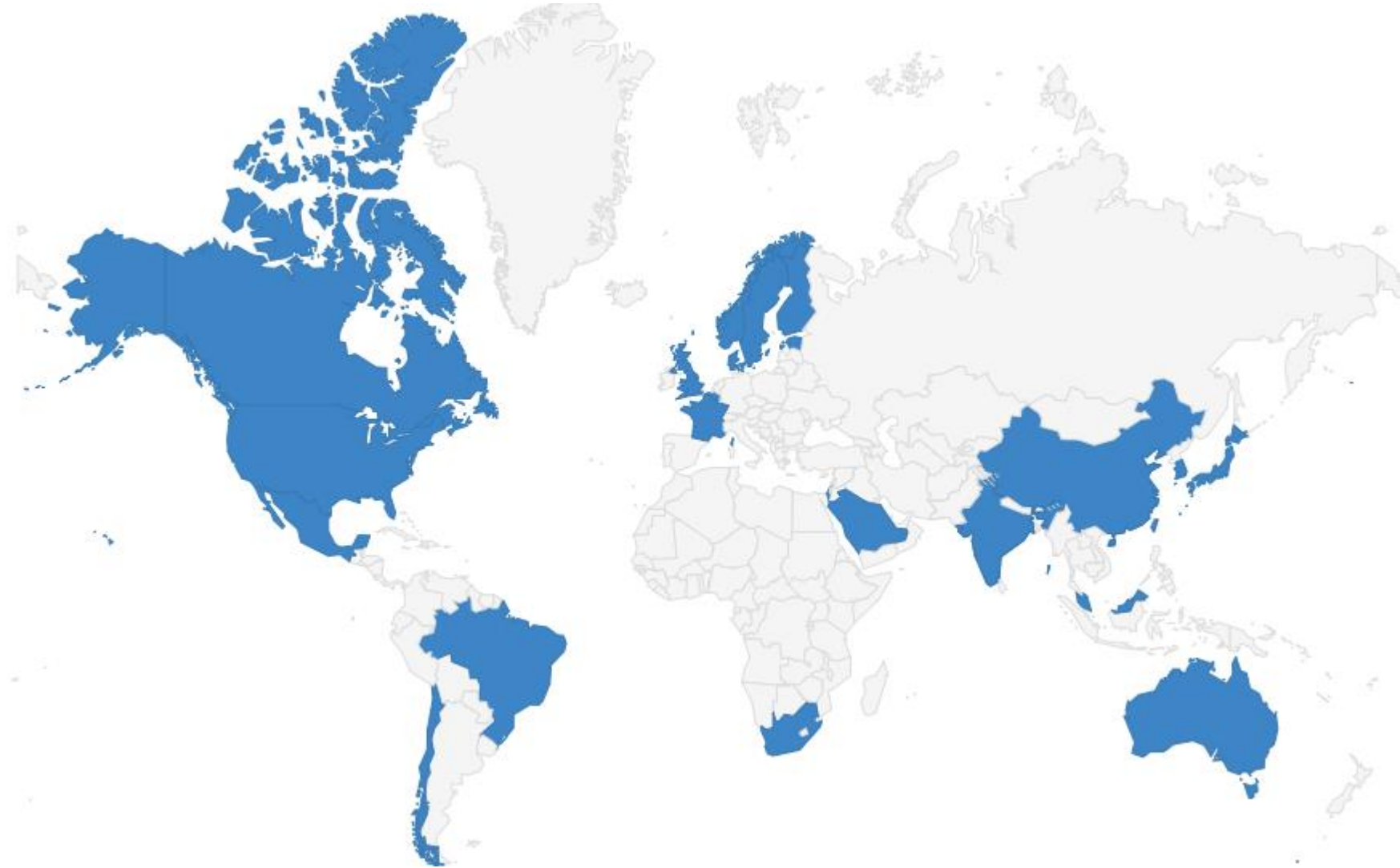




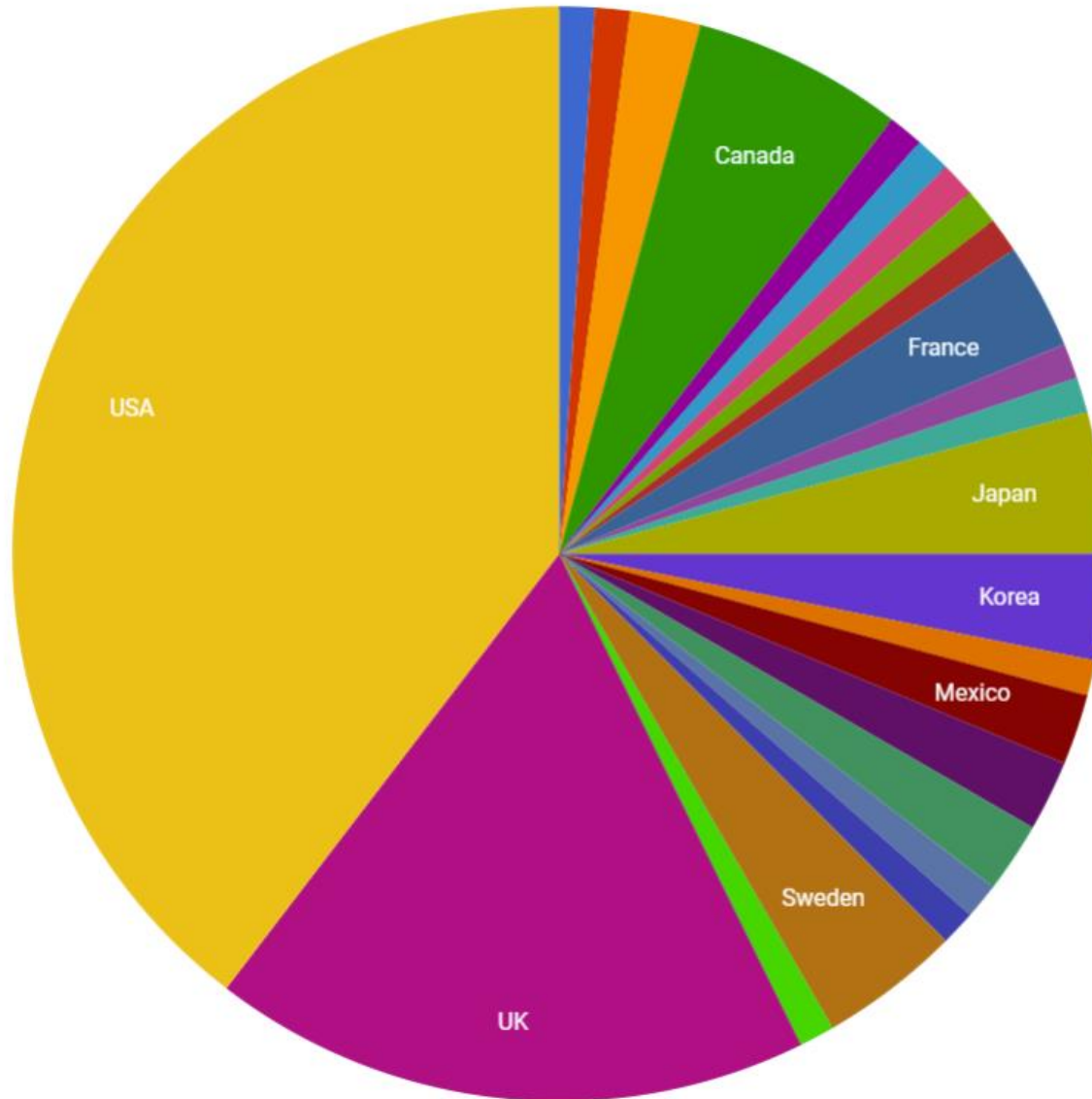
# Who is here?

- ~100 individuals
- ~25 countries
- ~60 cohort studies

Australia	Japan
Belgium	Korea
Brazil	Malaysia
Canada	Mexico
Chile	Norway
China	Saudi Arabia
Denmark	Singapore
Estonia	South Africa
Finland	Sweden
France	Taiwan
India	UK
Israel	USA



# Who is here?



- Australia
- Belgium
- Brazil
- Canada
- Chile
- China
- Denmark
- Estonia
- Finland
- France
- India
- Israel
- Japan
- Korea
- Malaysia
- Mexico
- Norway
- Saudi Arabia
- Singapore
- South Africa
- Sweden
- Taiwan
- UK
- USA

# Who is here?



23andMe  
45 and Up Study  
All of Us Research Program  
Apolipoprotein-related MOrtality RISk (AMORIS)  
AstraZeneca Integrated Genomics Initiative  
Barshi Cohort  
Biobank Japan  
BioVU Vanderbilt  
Brazilian Longitudinal Study of Adult Health (ELSA-Brasil)  
Cancer Prevention Study-II (CPS-II) & CPS-II Nutrition subcohort  
Children's Hospital of Philadelphia (CHOP) Biorepository  
China Kadoorie Biobank  
China PEACE Million Persons Project  
Clalit Israeli Genome Project  
Cohort of Norway (CONOR)  
Constances  
Danish National Biobank  
East London Genes and Health  
Environmental influences on Child Health Outcomes (ECHO)  
Estonian Genome Project  
European Prospective Investigation into Cancer & Nutrition (EPIC)  
Finnish Maternity Cohort Serum Bank  
Generations Study  
Genomics England  
Golestan Cohort Study  
Japan Public Health Center-based Prospective Study (JPHC)  
Japan Public Health Center-based Prospective Study for the Next generation (JPHC-NEXT)  
Kaiser Permanente Research Program on Genes, Environment, and Health (RPGEH)  
Korea Biobank Project

Korean Cancer Prevention Study-II (KCPS-II) Biobank  
Korean Genome and Epidemiology Study (KoGES)  
LifeGene (and EpiHealth)  
LIFEPATH  
Malaysian Cohort  
Maule Cohort (MAUCO Study)  
Mexico City Prospective Study  
Million Veteran Program  
Million Women Study  
Multiethnic Cohort Study (MEC)  
MyCode Community Health Initiative  
Newfoundland and Labrador Genome Project  
Northern Sweden Health and Disease Study  
Norwegian Family Based Life Course Study  
Norwegian Mother and Child Cohort Study (MoBa)  
Nurses' Health Study (NHS)  
Nurses' Health Study II (NHSII)  
OHS (Ontario Health Study) and the Canadian Partnership for Tomorrow Project (CPTP)  
PERSIAN Cohort Study  
Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO)  
Saudi Human Genome Program  
Saudi National Biobank  
Singapore National Precision Medicine Program  
Taiwan Biobank  
Tohoku Medical Megabank Project  
UK Biobank  
UKCTOCS (UK Collaborative Trial of Ovarian Cancer Screening)  
Longitudinal Women's Cohort – (UKLWC)  
Women's Health Initiative (WHI)

For more details,  
refer to meeting  
booklet at:  
[http://bit.ly/  
cohorts-summit](http://bit.ly/cohorts-summit)

# Cohort Survey Responses

<http://bit.ly/cohorts-summit>



Global Genomic  
Medicine Collaborative

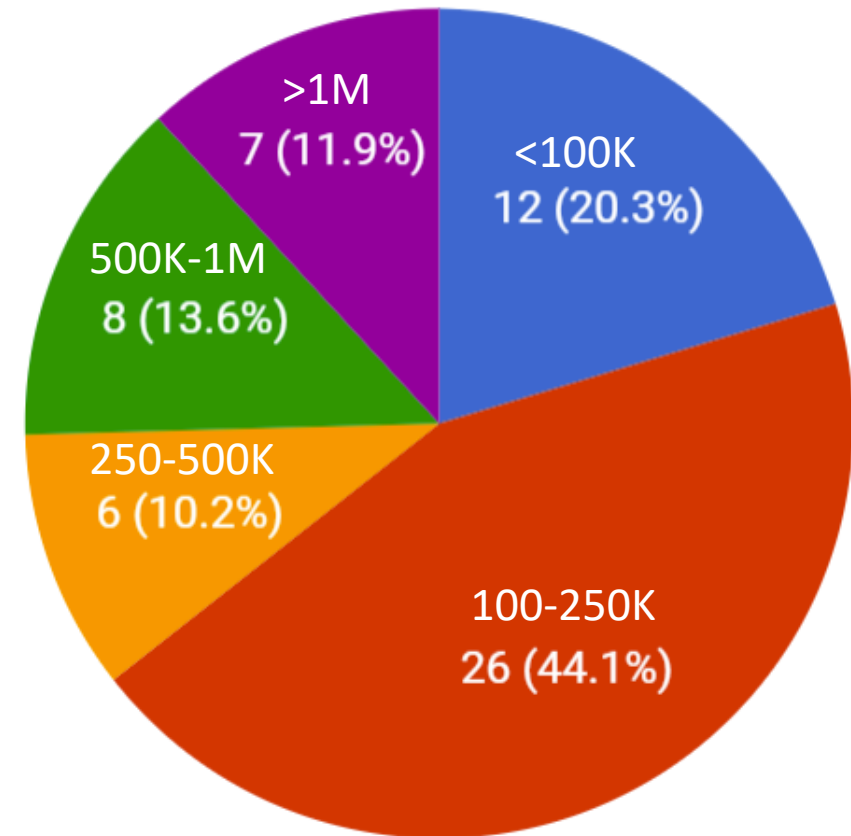


International Cohorts Summit

## Cohort Sizes

- Based on **current** sizes:
  - Collectively, >25.6 million people
- Based on **target** sizes:
  - Collectively, >84.6 million people
- Enrollment:
  - Started anywhere between 1960s and 2018
  - 6 studies commenced enrollment in past 5 years

## Cohorts of Various Sizes



# Cohort Survey Responses

<http://bit.ly/cohorts-summit>



Global Genomic  
Medicine Collaborative



International Cohorts Summit

## Samples, Analysis & Banking

- **What is being collected?**
  - 39 studies: blood
  - 23 studies: urine
  - 8 studies: tumor tissue
  - 30 studies: multiple types
- **Central biobank?**
  - 46 studies: Yes (now, or soon)
- **DNA available now/soon for all?**
  - 42 studies: Yes
  - 10 studies: For a subset
  - 3 studies: No
- **Genotyping?**
  - 45 studies: For some or all
  - 5 studies: No
  - (Some performed for purposes of ancestry testing)
- **Genomic sequencing?**
  - 28 studies: For some or all (WGS or WES)
  - 8 studies: No

# Cohort Survey Responses

<http://bit.ly/cohorts-summit>



Global Genomic  
Medicine Collaborative



International Cohorts Summit

## Patient Consent & Re-contact

- **Consent to share data beyond initial study investigators?**
  - 49 studies: Yes
    - 10: For research
    - 6: Type not specified
    - 1: Tiered
  - 2 studies: Varies by cohort
  - 6 studies: No
- **Study info/data returned to or accessible by participants?**
  - 28 studies: Yes
  - 25 studies: No



# Cohort Survey Responses

<http://bit.ly/cohorts-summit>



Global Genomic  
Medicine Collaborative



International Cohorts Summit

## Willing to Share Data?

- **Almost all (52 studies) willing to share - sometimes with limitations**
  - Redacted individual data?
    - # of studies: 21 yes with limitations/restrictions, 18 yes, 2 no
  - Summary data?
    - # of studies: 38 yes, 6 yes with limitations/restrictions
  - Metadata?
    - # of studies: 36 yes, 6 yes with limitations/restrictions
  - Case report forms, data collection materials?
    - # of studies: 33 yes, 6 yes with limitations/restrictions

# Cohort Survey Responses

<http://bit.ly/cohorts-summit>



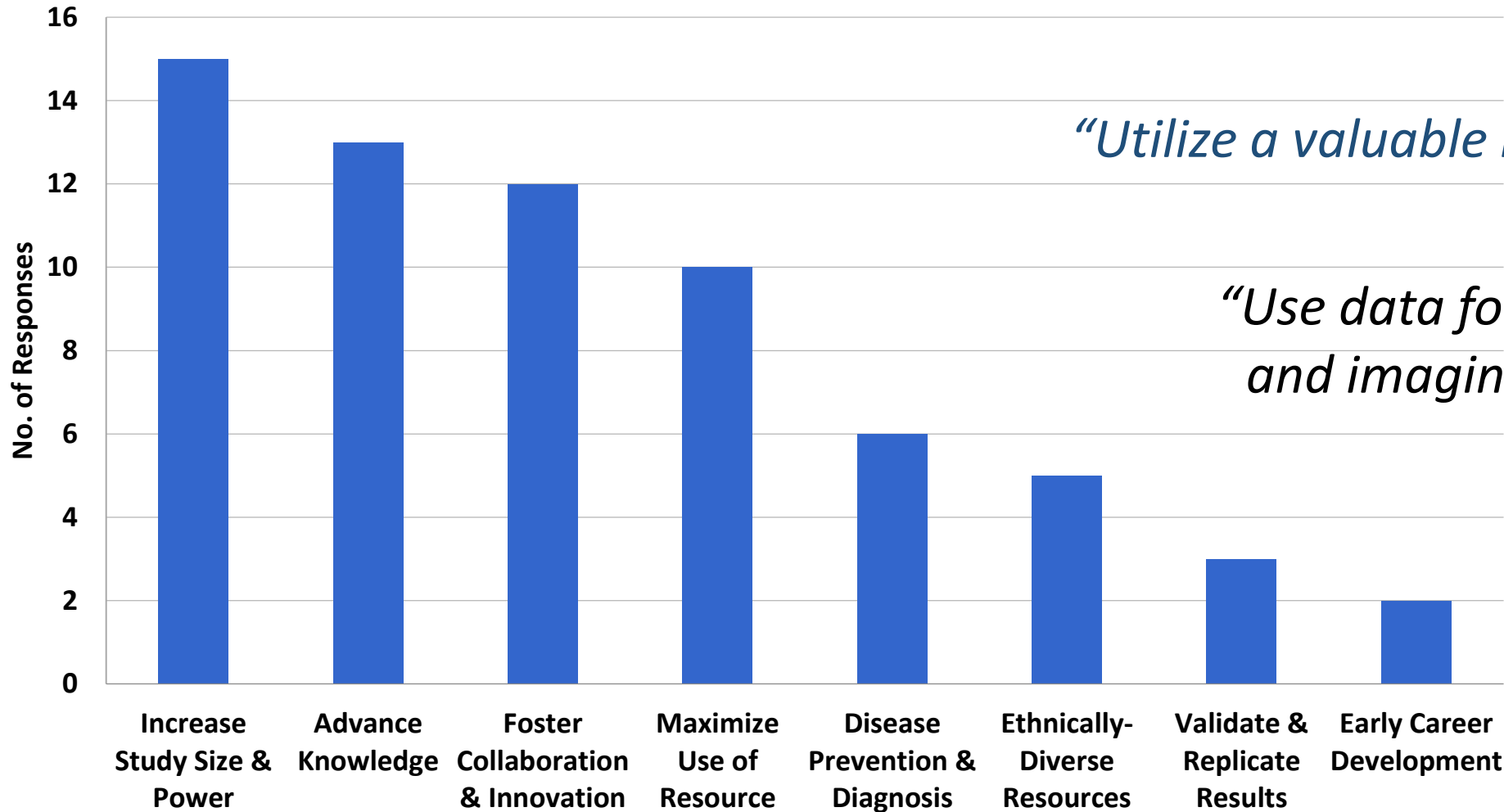
Global Genomic  
Medicine Collaborative



International Cohorts Summit

## What are the **VALUES** of sharing data?

*“The values are infinite”*



*“Utilize a valuable resource to its fullest”*

*“Use data for an increasingly wide and imaginative set of purposes”*

*“Compare different populations to each other”*

# Cohort Survey Responses

<http://bit.ly/cohorts-summit>

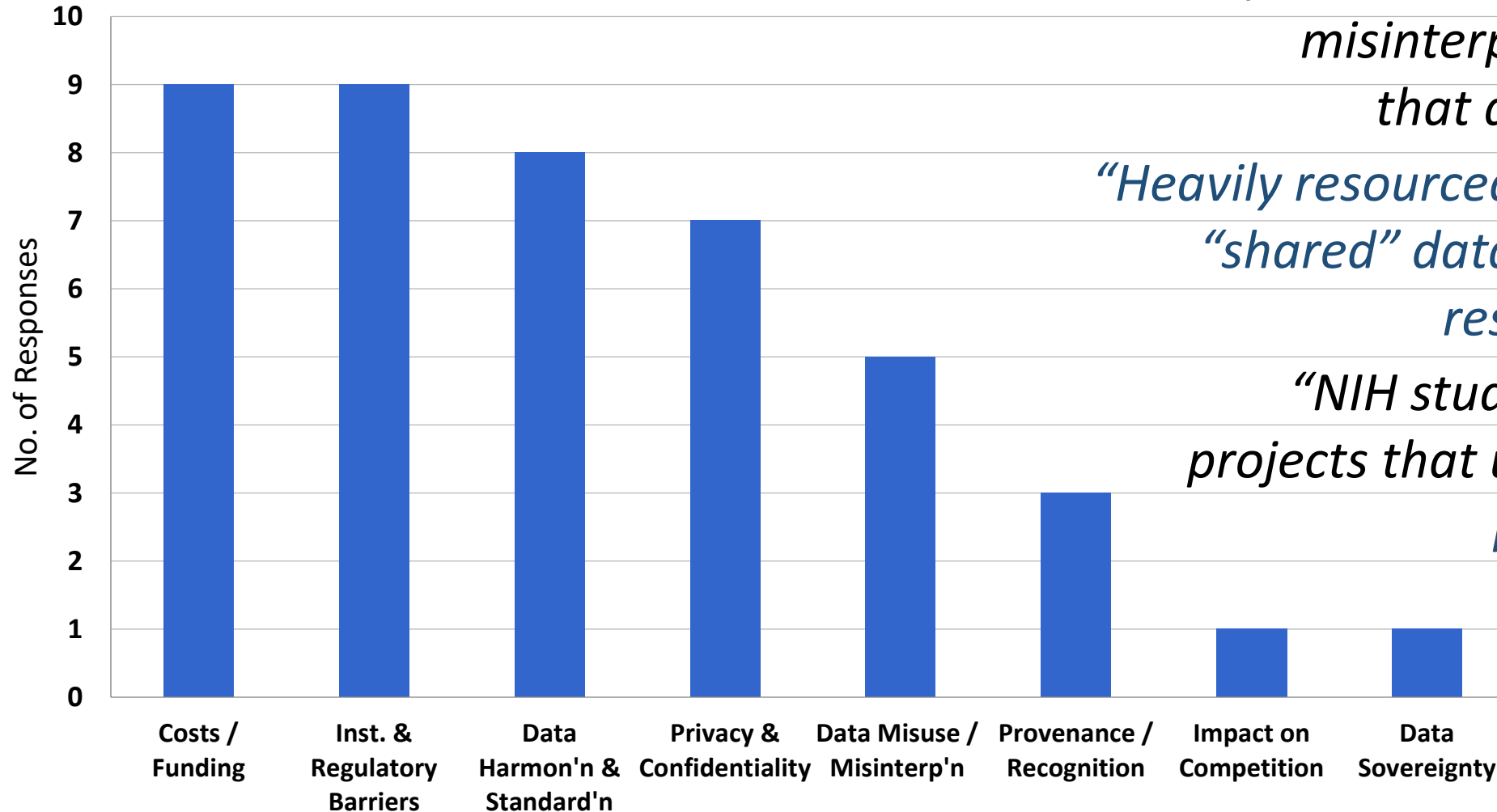


Global Genomic  
Medicine Collaborative



International Cohorts Summit

## What are the **CHALLENGES** of sharing data?



*“Complex data made public is open to misinterpretation, or analyses that are not well-planned”*

*“Heavily resourced centers concentrate “shared” data, have lion’s share of research opportunities”*

*“NIH study sections don’t fund projects that utilize “shared” data”*

*Impact of EU General Data Protection Regulation?*



# Cohort Survey Responses

<http://bit.ly/cohorts-summit>



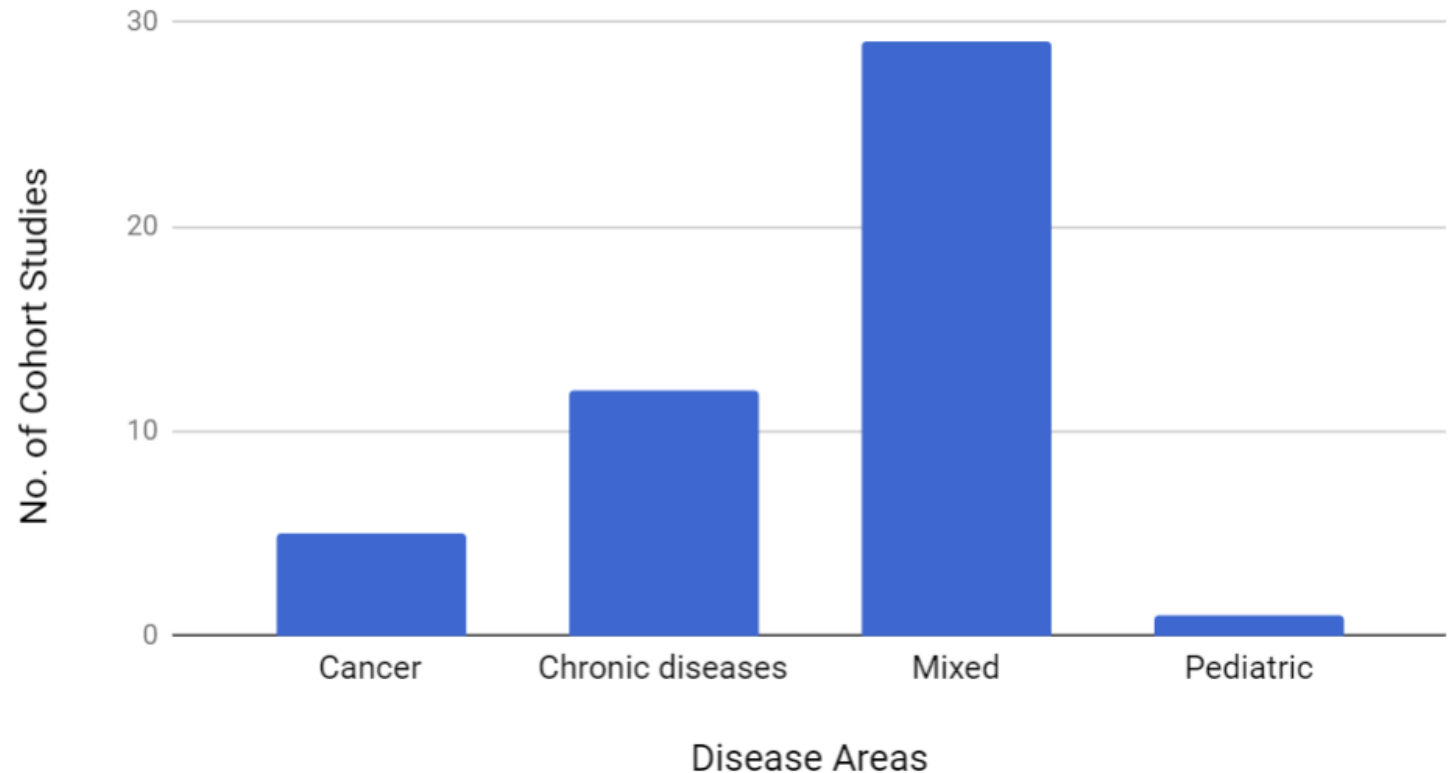
Global Genomic  
Medicine Collaborative



International Cohorts Summit

## Disease Areas of Focus

- Selected for disease?
  - 45 studies: **no**
  - 3 studies: **yes**
  - 3 studies: **yes & no**
- Areas of focus:
  - 29: **multiple diseases**
  - 12: **chronic only**
  - 5: **cancer only**
  - 1: **pediatric only**



# Cohort Survey Responses

<http://bit.ly/cohorts-summit>



Global Genomic  
Medicine Collaborative



International Cohorts Summit

## Period of Follow-up

- **Follow-up**
  - 6 studies: up to 5 years
  - 9 studies: 5-10 years
  - 9 studies: 10-20 years
  - 3 studies: 20-40 years
  - 4 studies: indefinite

