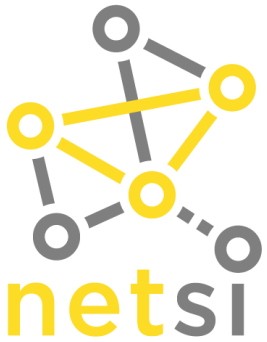


Northeastern University

Network Science Institute

Quantifying Biobank Impact

Rodrigo Dorantes-Gilardi, John Gaziano and Albert-László Barabási

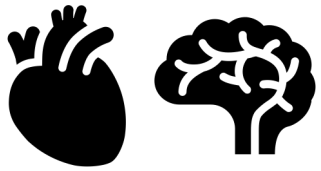
Email: r.dorantesgilardi [at] northeastern [dot] edu

Twitter: @ruydg

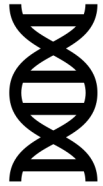
Biobanks are biorepositories

For a given population

1)SAMPLES



Bio samples

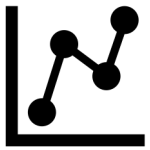


Genomic sequence

Medical history and lifestyle records

2)DATA

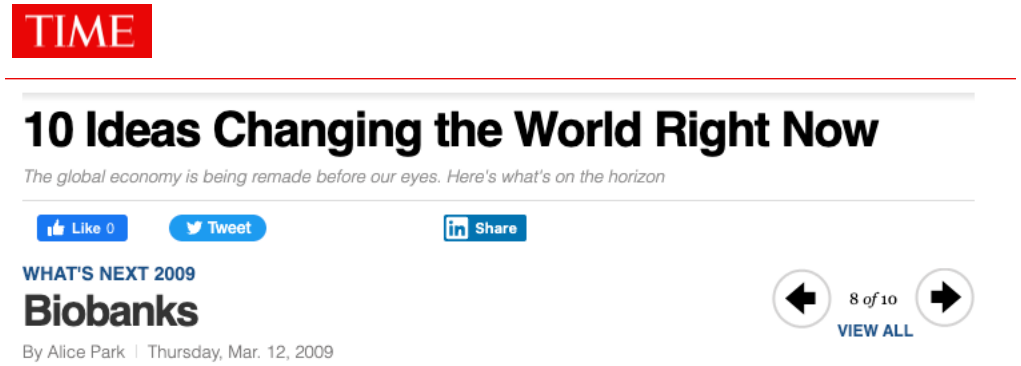
To study health and disease



3)Research

Biobanks essential for biomedical research

- Important diseases studied
- Diverse populations sampled
- General public or specific phenotype


TIME

10 Ideas Changing the World Right Now

The global economy is being remade before our eyes. Here's what's on the horizon

Like 0 Tweet Share

WHAT'S NEXT 2009

Biobanks

By Alice Park | Thursday, Mar. 12, 2009

8 of 10

VIEW ALL

A biobank helped uncover the risk factors of cardiovascular diseases

Framingham Heart Study

Est. in Massachusetts in 1948

Other examples include

Parkinson's Progression Markers Initiative

biobank uk

Enabling scientific discoveries that improve human health

THE CANCER GENOME ATLAS

MILLION VETERAN PROGRAM

Biobank numbers are hard to quantify

No central biobank database exists

How Many Health Research Biobanks Are There?

Sheila O'Donoghue, Simon Dee, Jennifer A. Byrne, and Peter Hamilton Watson

Published Online: 28 Sep 2021 | https://doi.org/10.1089/bio.2021.0063

Numbers and **Impact** hard to track

Sometimes biobank recognition goes lacking

Cambon-Thomsen et al. Proposed a biobank impact factor

The role of a bioresource research impact factor as an incentive to share human bioresources

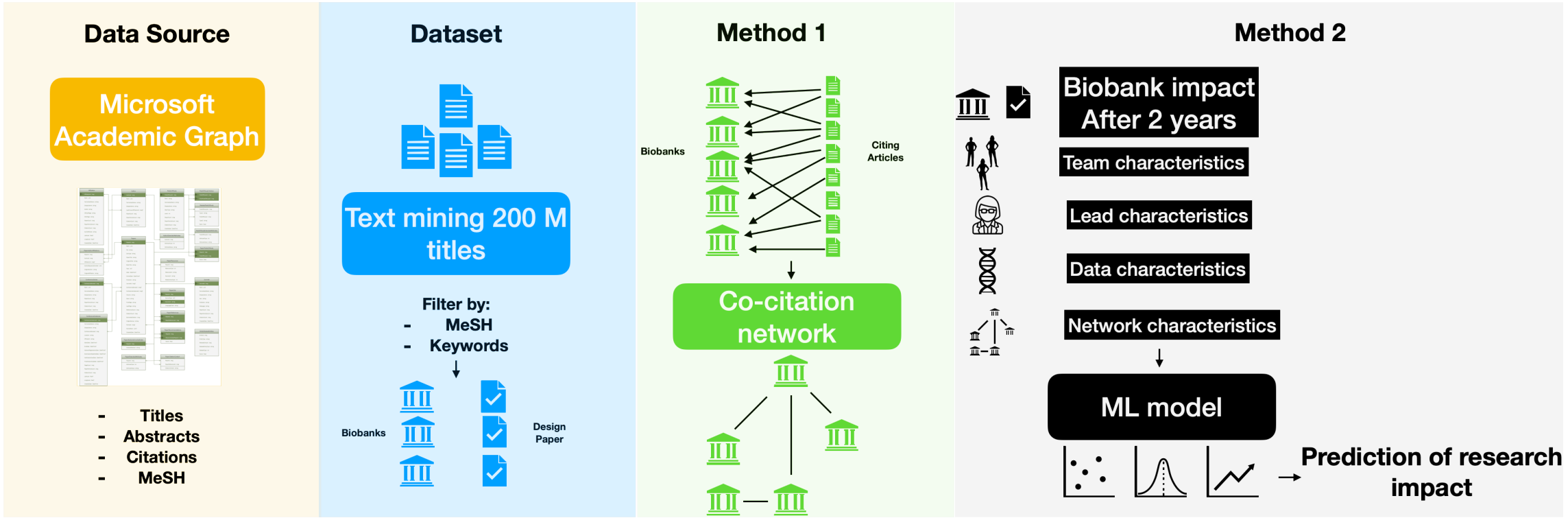
the BRIF workshop group

Nature Genetics **43**, 503–504 (2011) | Cite this article

Some unanswered questions to tackle

- Numbers: How many biobanks are there of each type?
- Incentives: How are researchers giving recognition to biobanks?
- Success: What are the characteristics of high-impact biobanks?

Methodology



Analyze bibliographic data

Classify biobanks

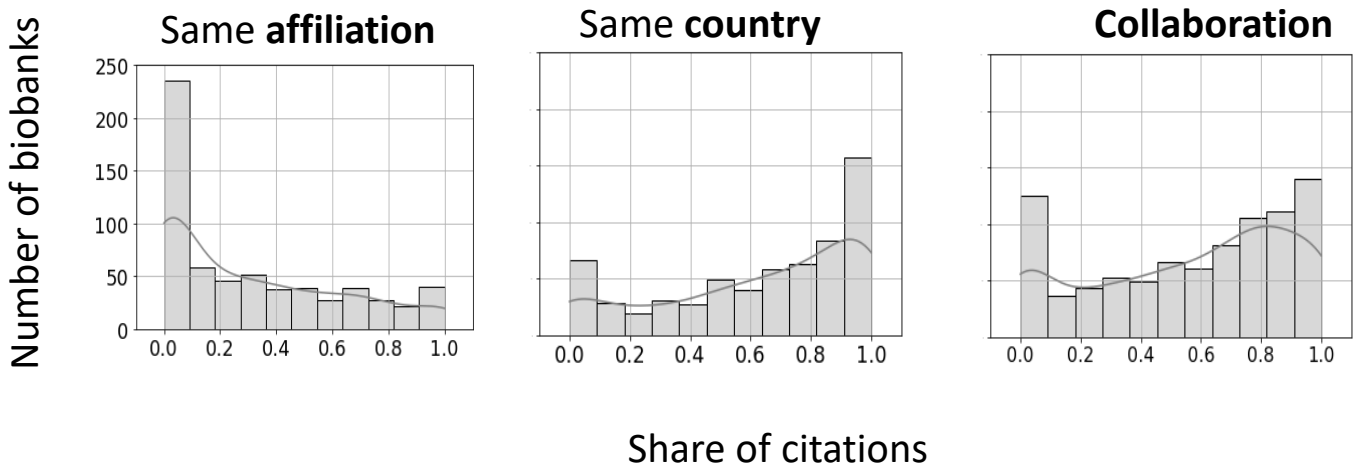
Obtain important characteristics

The impact of biobanks is local

56K+ Citations to 1k biobanks.

On average:

- 31% same affiliation
- 63% national
- 56% collaborations

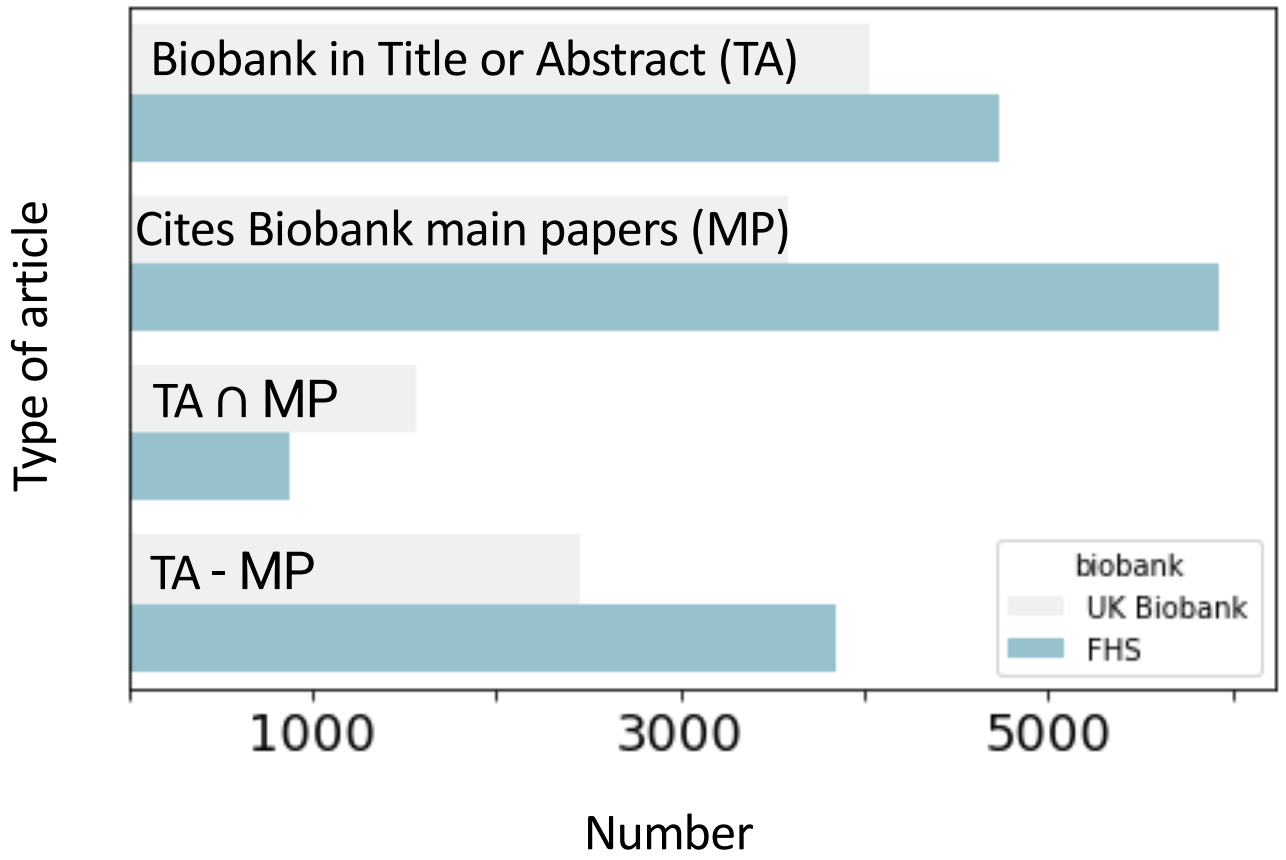


Citations are not a vehicle to biobank recognition

Observation:
Less than **40% of articles** with the biobank in the title or abstract cite the biobank

Potential consequence:
- Co-authorship as a possible incentive to produce biobanks.

Example: Framingham Heart Study (FHS) and UK Biobank

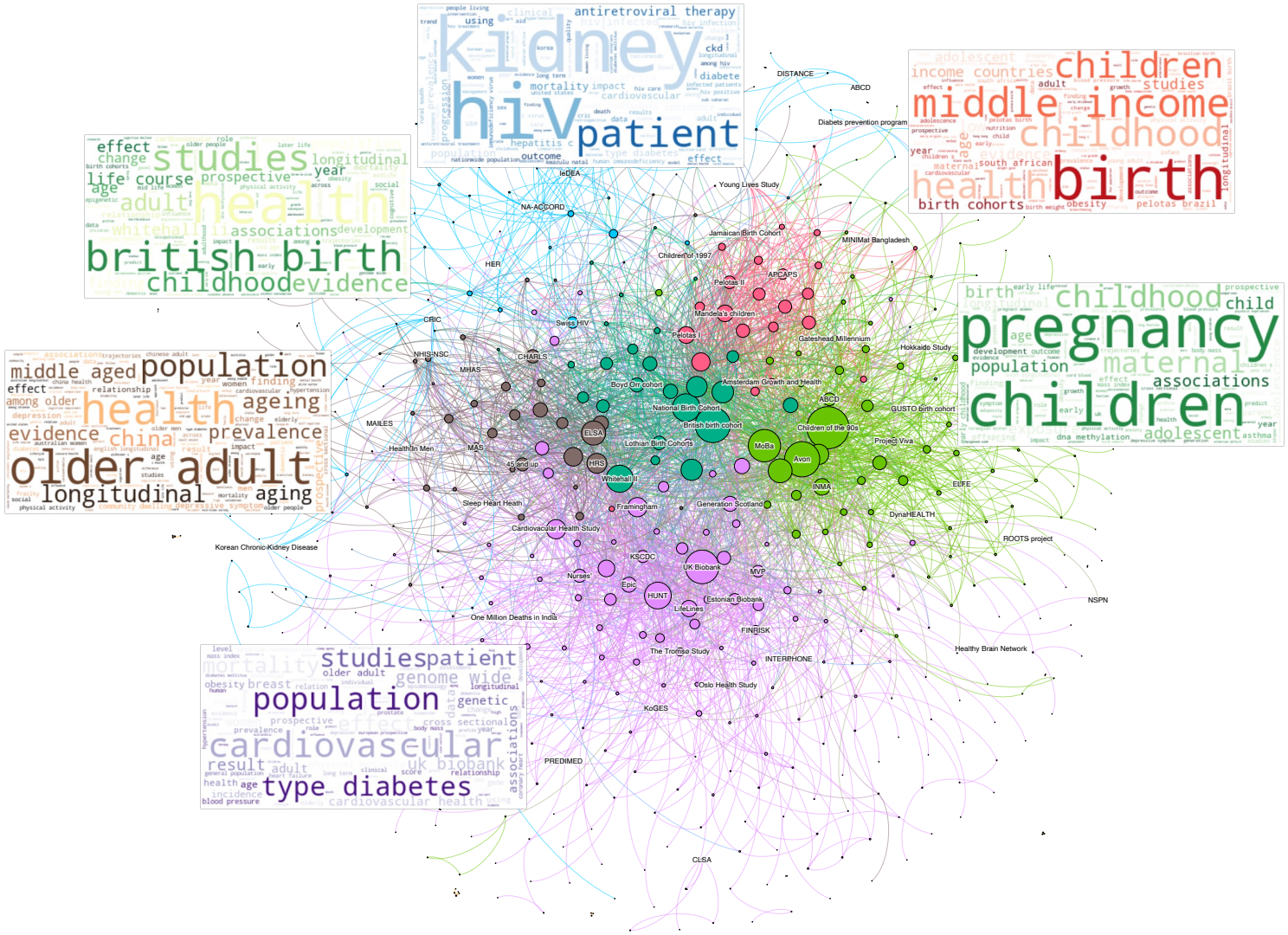


Biobank co-citation network

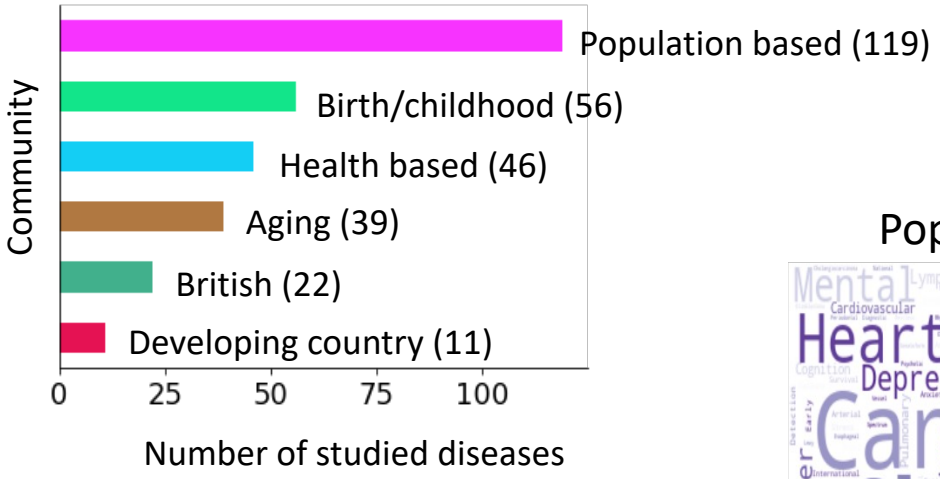
Sample of **1,209** biobanks with their official design article.

Biobank co-citation network
- Nodes: **Biobanks**
- Edges: **Co-citations**

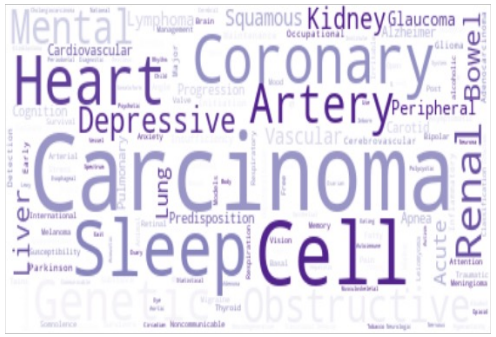
Communities based on titles from citing papers (LDA).
- Health based
- Population based
- Aging
- British biobanks
- Developing countries
- Birth/childhood



Biobanks Focus on different conditions



Population based

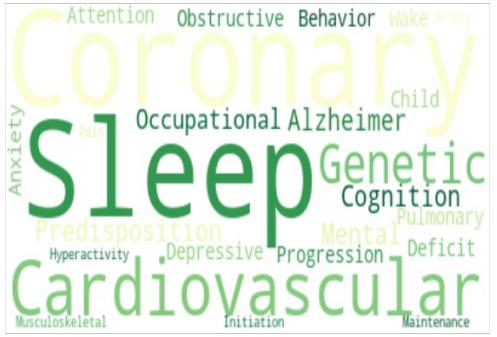


Birth/childhood

Health based

Aging

British

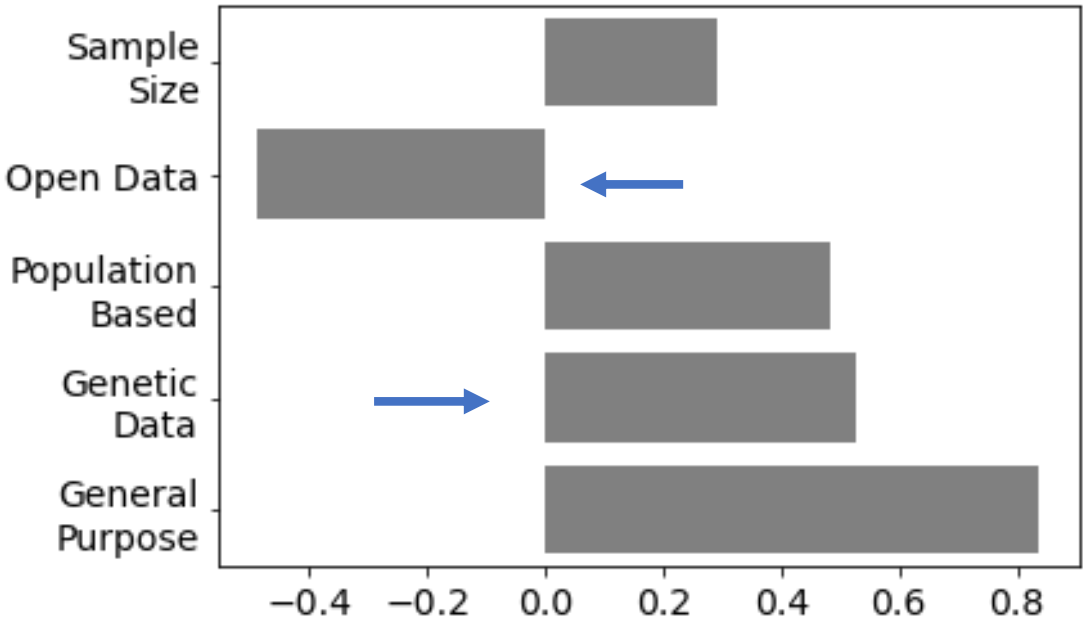


Developing country

Characteristics of high impact biobanks

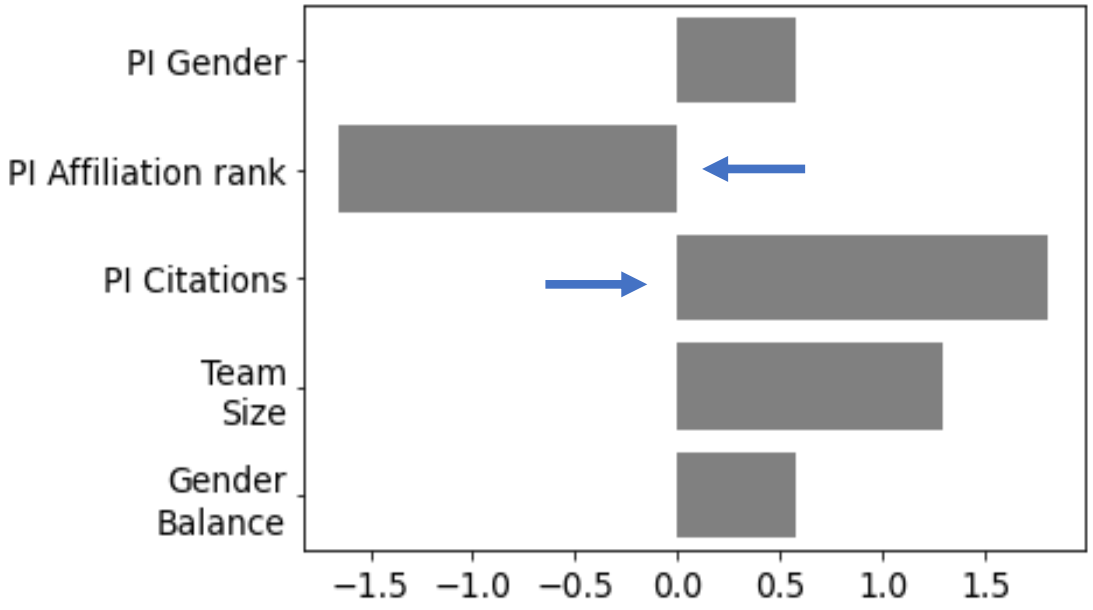
$$\log\frac{P(a_i=1)}{1-P(a_i=1)} \propto \beta_1 \times n_i + \beta_2 \times d_i + \beta_3 \times k_i + \beta_4 \times g_i + \beta_5 \times p_i.$$

Data characteristics



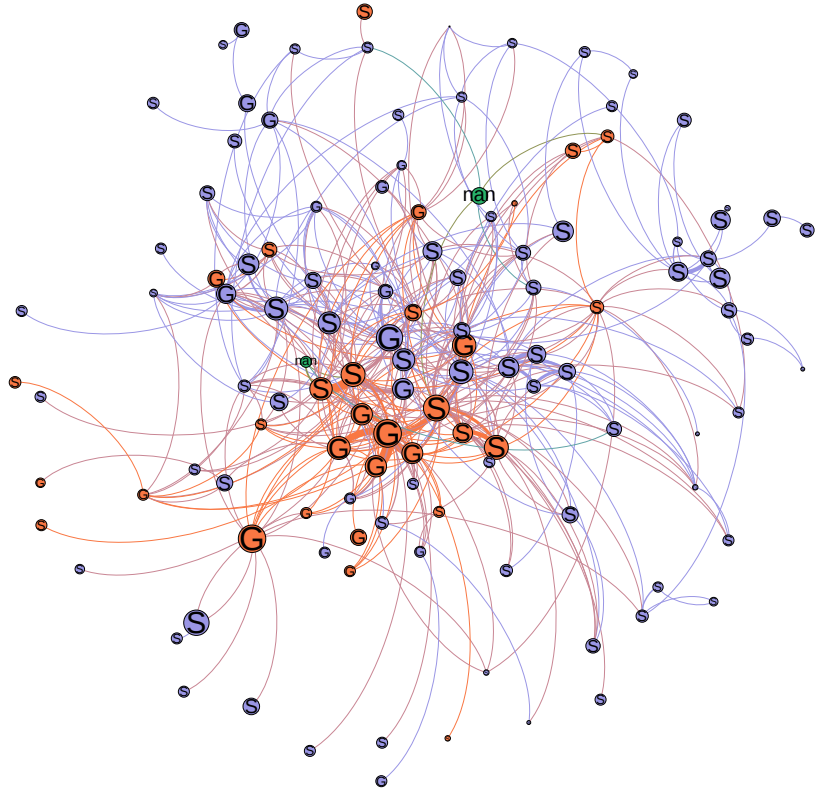
Model coefficient

Team characteristics



Model coefficient

Why?

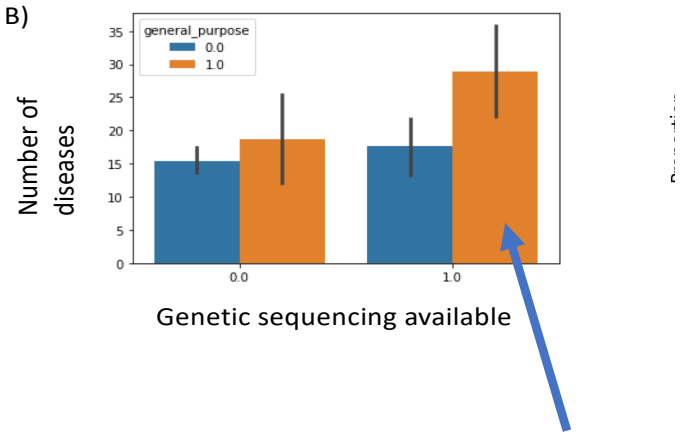


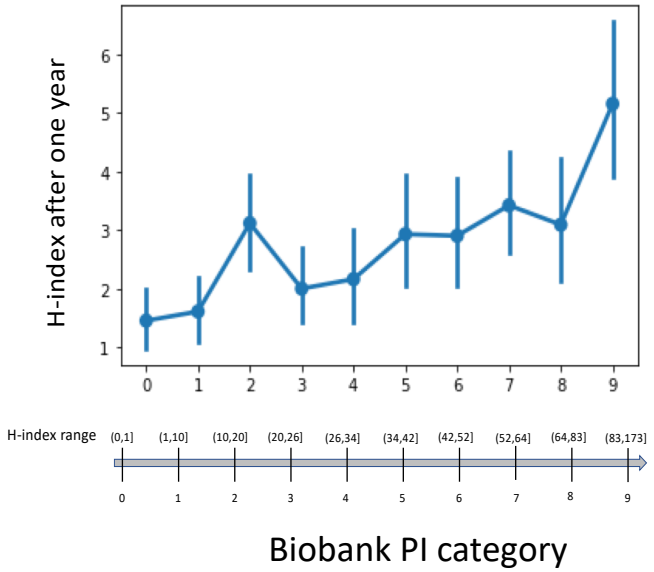
Genetic data
No genetic data

G – General purpose
S – Disease specific

Higher
biobank h-index

1. General purpose with genetic data are central nodes in co-citation network





2. They help study more diseases

3. Popular PIs push the biobank impact immediately

Conclusions

- 1) Incentive of Biobank creation is based on collaborations and not citations
- 2) Six biobank communities with different characteristics/diseases studied
- 3) Genetic data and electronic health records is important for biobank success
- 4) Popular leaders are important for biobank success

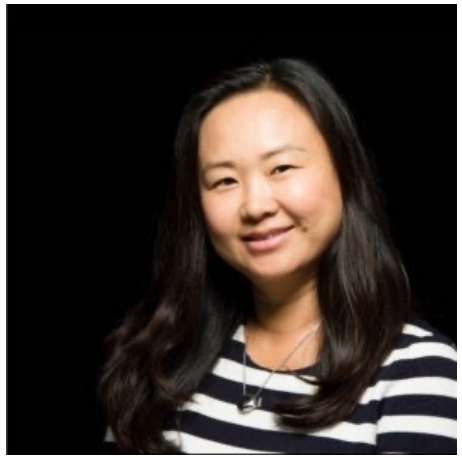
Thanks to the team

Albert-Laszlo Barabási

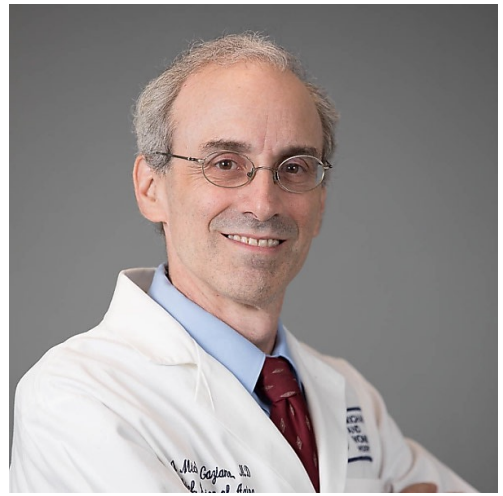
Kerry Ivey

Rachael Matty

Lauren Costa



Kelly Cho



John Gaziano