Building a Biobank for Clinical Research

Roberto Caricchio, MD
Professor of Medicine
Chief, Section of Rheumatology

10 Ideas Changing the World

March 2009 Issue

1. Why Your Job is Your Most Valuable Asset
2. Repurposing the Suburbs
3. Survival-Store Shopping
4. **Biobanks; Saving Your Parts**
5. Need Land? Rent-A-Country
6. The New Calvinism
7. Ecological Intelligence
8. Amortality; Forever Young
9. Africa; Open for Business
10. Reinventing the Highway

What are biobanks?

A collection of biological material and the associated data and information stored in an organized system, for a population or a large subset of a population.

Organization for the Economic Cooperation and Development

Organized collections and storage of human biological samples and associated data of great significance for research and personalized medicine.

BBMRI-European Commission (EC)

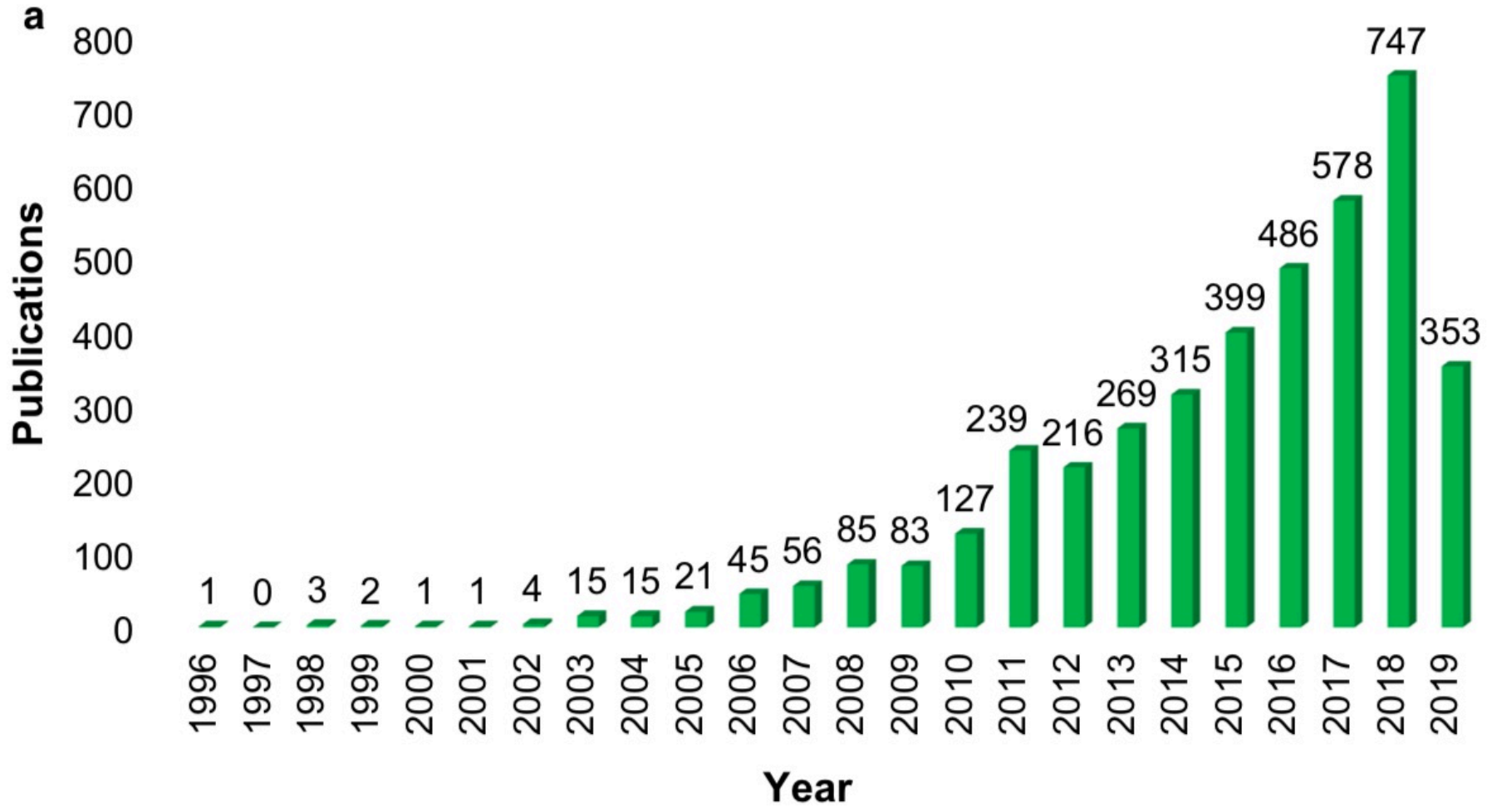
The rise of Biobanks

Human studies support the experimentally based notion of oxidative DNA damage as an important mutagenic and apparently carcinogenic factor. However, the proof of a causal relationship in humans is still lacking…..

…..This could possibly be supported by demonstration of the rate of oxidative DNA damage as an independent risk factor for cancer in a prospective study of **biobank** material using a nested case control design……

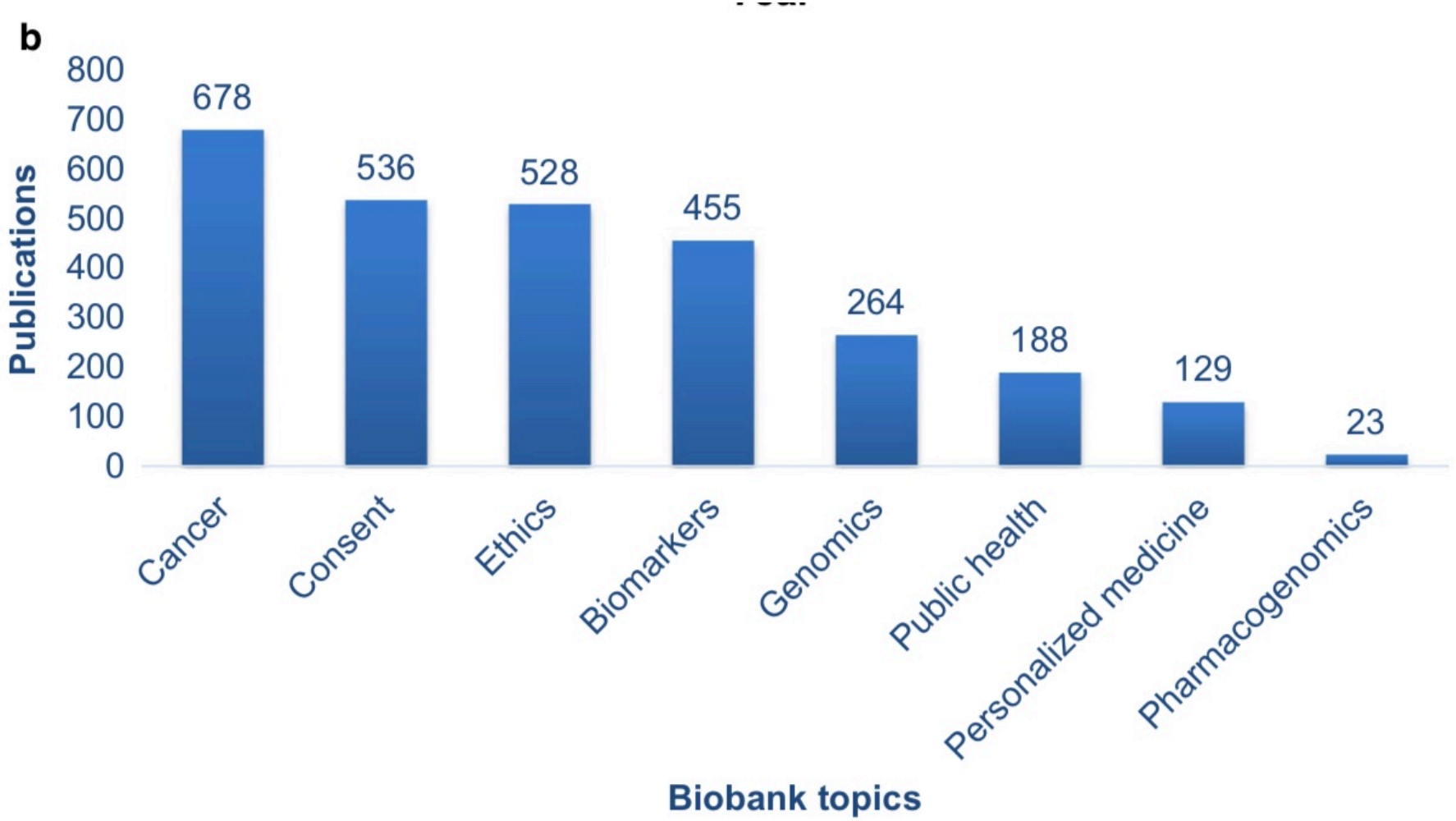
Loft S, Poulsen HE. Cancer risk and oxidative DNA damage in man.
J Mol Med (Berlin, Germany). 1996;74:297–312.

Biobank in the scientific literature



Coppola *et al. J Transl Med (2019) 17:172*

Biobank in the scientific literature



Coppola *et al. J Transl Med (2019) 17:172*

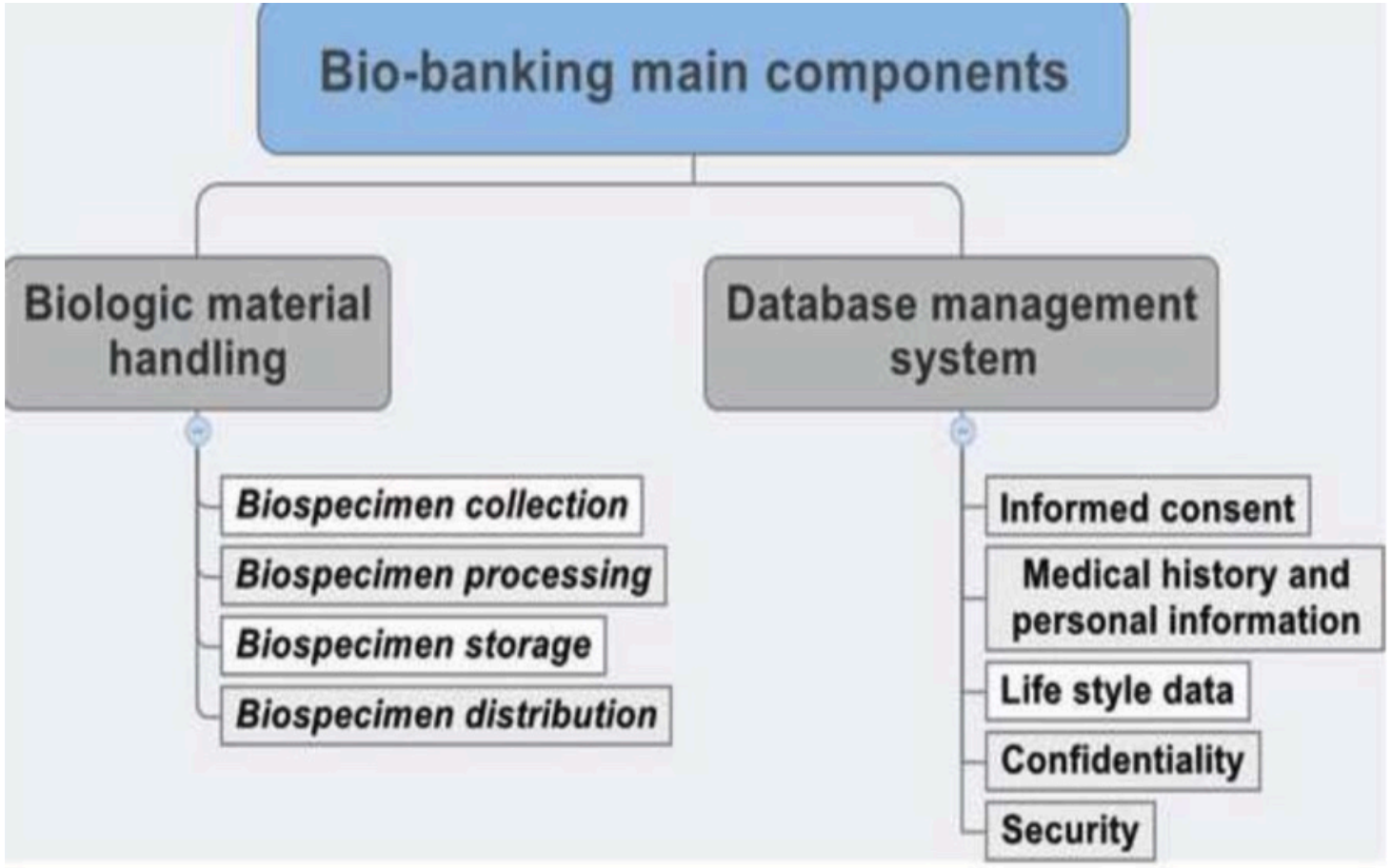
Biobank in the scientific literature

- Understanding of genomic information and genetic mechanisms in diseases,
- Developments in IT sector and bioinformatics.

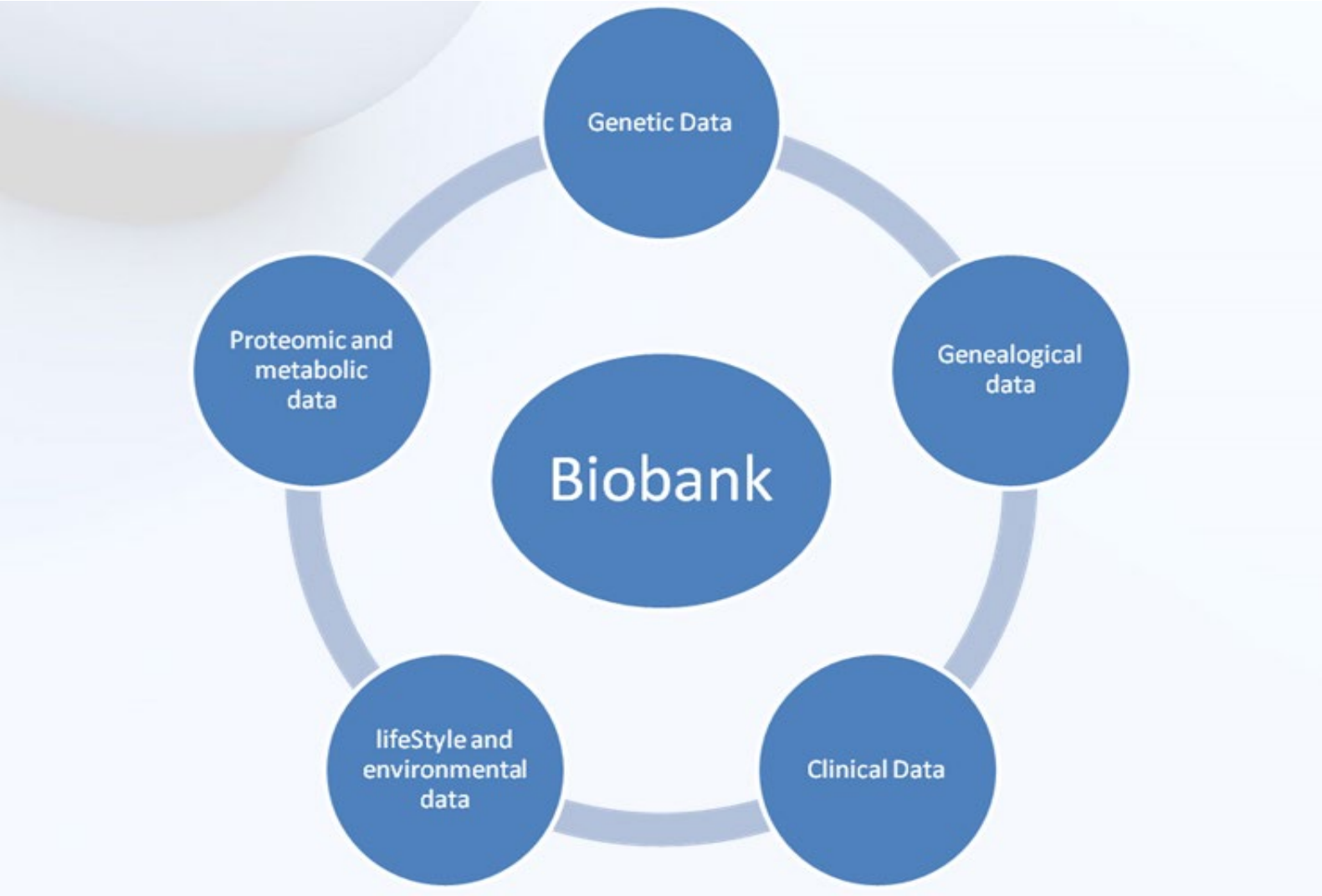
What are biobanks?

- Biological Human sample (Biospecimen)
- Attached or connected information
- Consent and patient data safety and protection

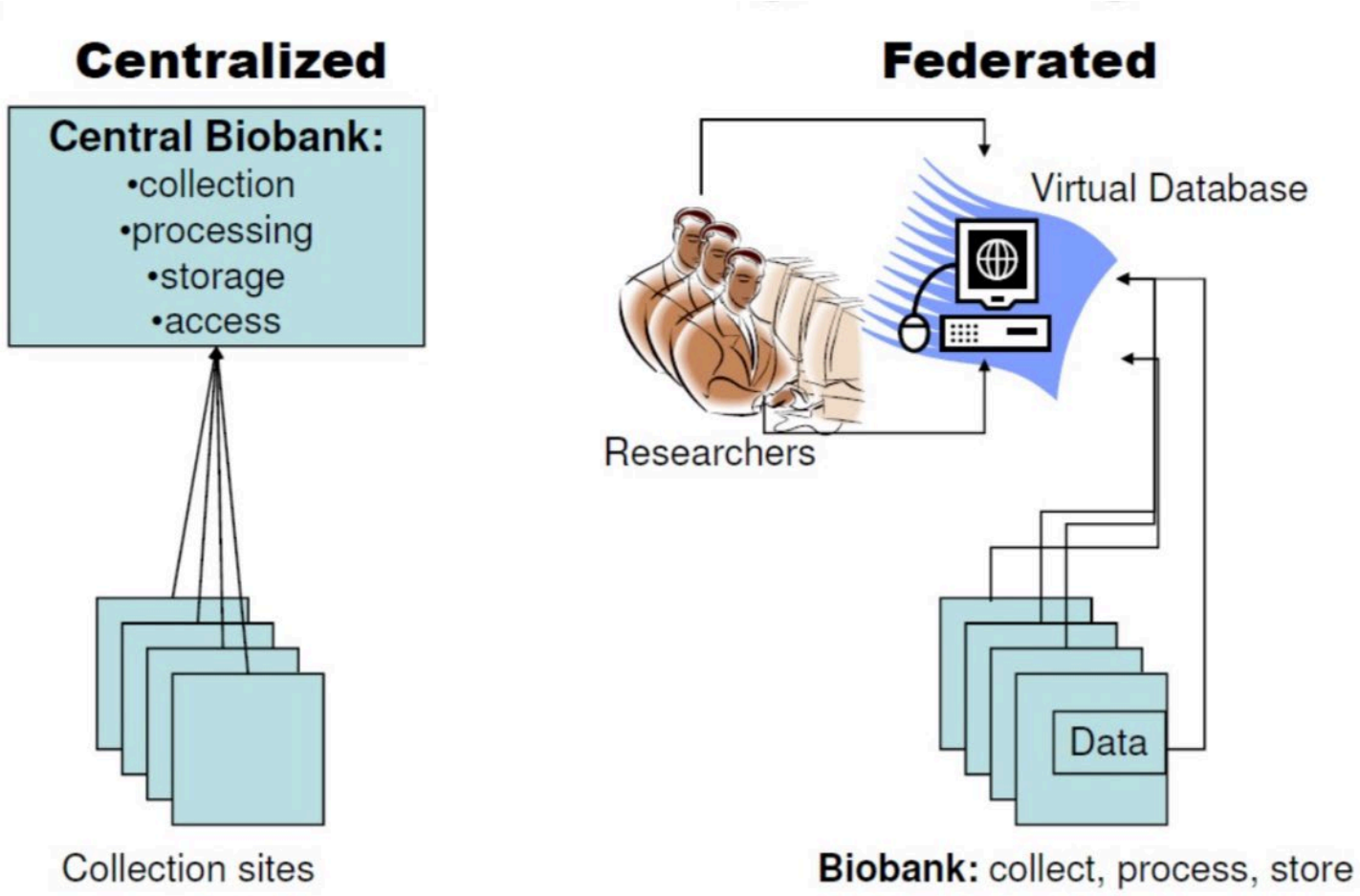
What are biobanks?



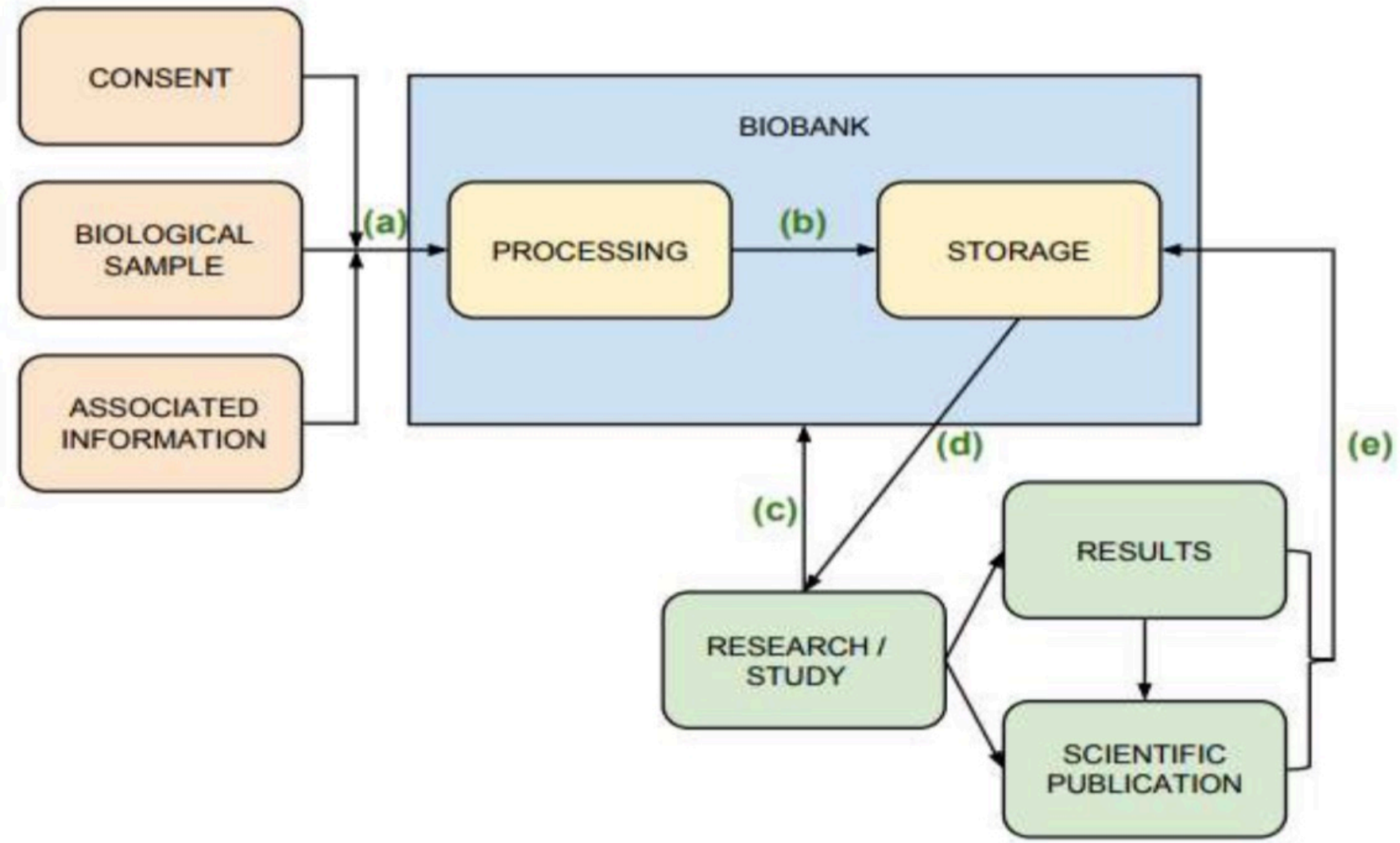
The ideal Biobank



Centralized vs Federated biobanking



The process of biobanking



Biospecimen

- Blood
- Plasma
- Serum
- RBC
- White cells
- DNA
- RNA
- proteins
- Cell lines
- urine
- CSF
- SF
- Amniotic fluid
- Buffy coat
- BMSC
- BM Tissue
- Tissue

Storage in biobanks

Samples are stored in a way appropriate for the sample material and the intended research purpose.

- Blood, plasma, serum, and DNA are stored in -80°C freezers.

- Tissues and cell lines are preserved in liquid nitrogen freezers at -196^0 C.

Types of Biobanks

Human biobank classification is based on:

- **Tissue type** (tumor tissue, cells, blood, DNA or RNA);
- **Purpose/intended use** (research, forensics, transplantation, source for therapeutics, e.g., umbilical blood, stem cell biobanks for individual or community use, or diagnostics);
- **Ownership** (academic and research institutions, hospitals, biotechnology and pharmaceutical companies or government run)

Types of Biobanks

- **Population-based biobanks**

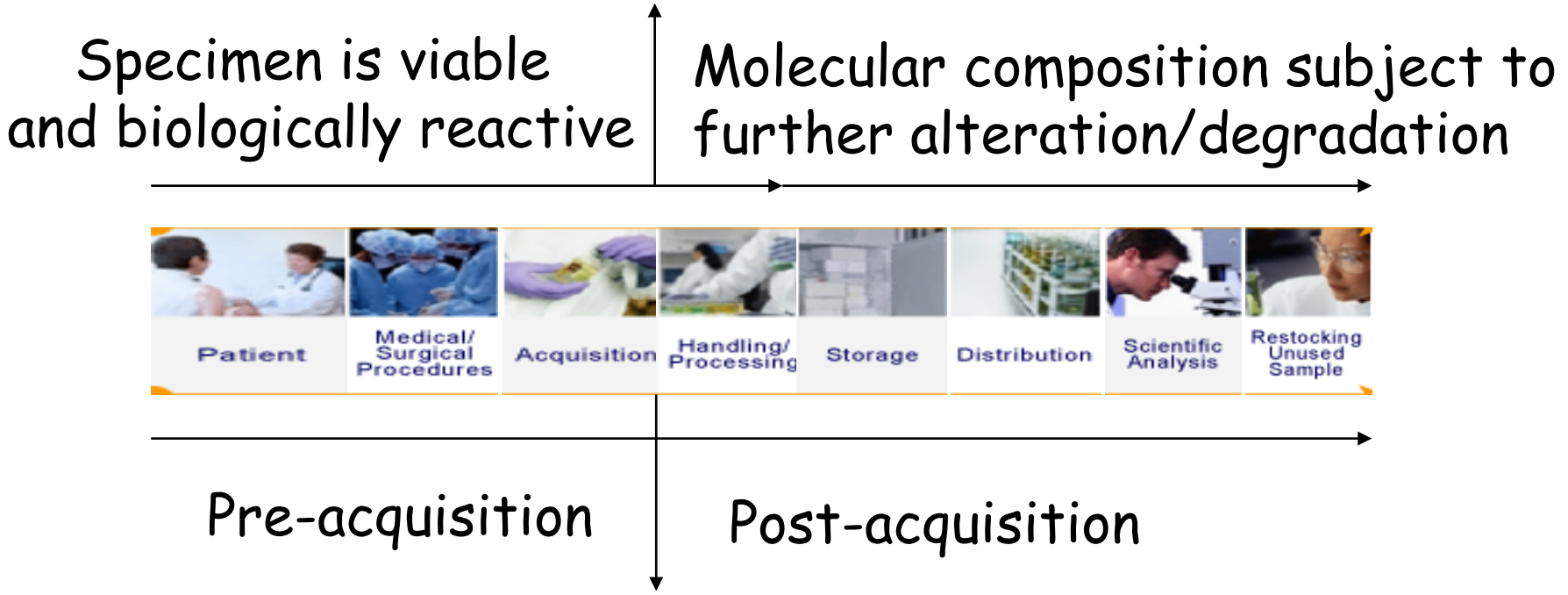
focused on the study of development of common, complex diseases over time

- **Disease oriented biobanks**

-biobanks of tissue samples and clinical data also referred to as disease oriented or clinical biobanks.

European Biobanking and BioMolecular resources Research Infrastructure (**BBMRI**)

Lifecycle of the Biospecimen



Biobanks requirements

- Consenting processes
- Collection services
- Storage infrastructure
- Specimen Access and Distribution
- Accounting
- Regulatory compliance

Biospecimen Availability and Quality: Major Challenge for Research

The lack of high quality, appropriately collected, stored and annotated specimen is the limiting factor for translational research or basic research

High quality specimen enables:

- Drug discovery
- Translational Research and basic research
- Clinical development
- Assay variability
- Specimen variability
- Diagnostic development
- Personalized medicine
- Targeted therapies

Why is it so difficult to obtain high quality specimens and data?

- Collection, processing, storage procedures differ
- Degree and type of data annotations varies
- Scope and type of patient consent differs
- Access policies are lacking or unknown to potential users
- Materials transfer agreement conditions differ
- Supporting IT structures differ in capacity and functionality
- Funding sources varies

Pre-analytical Variables

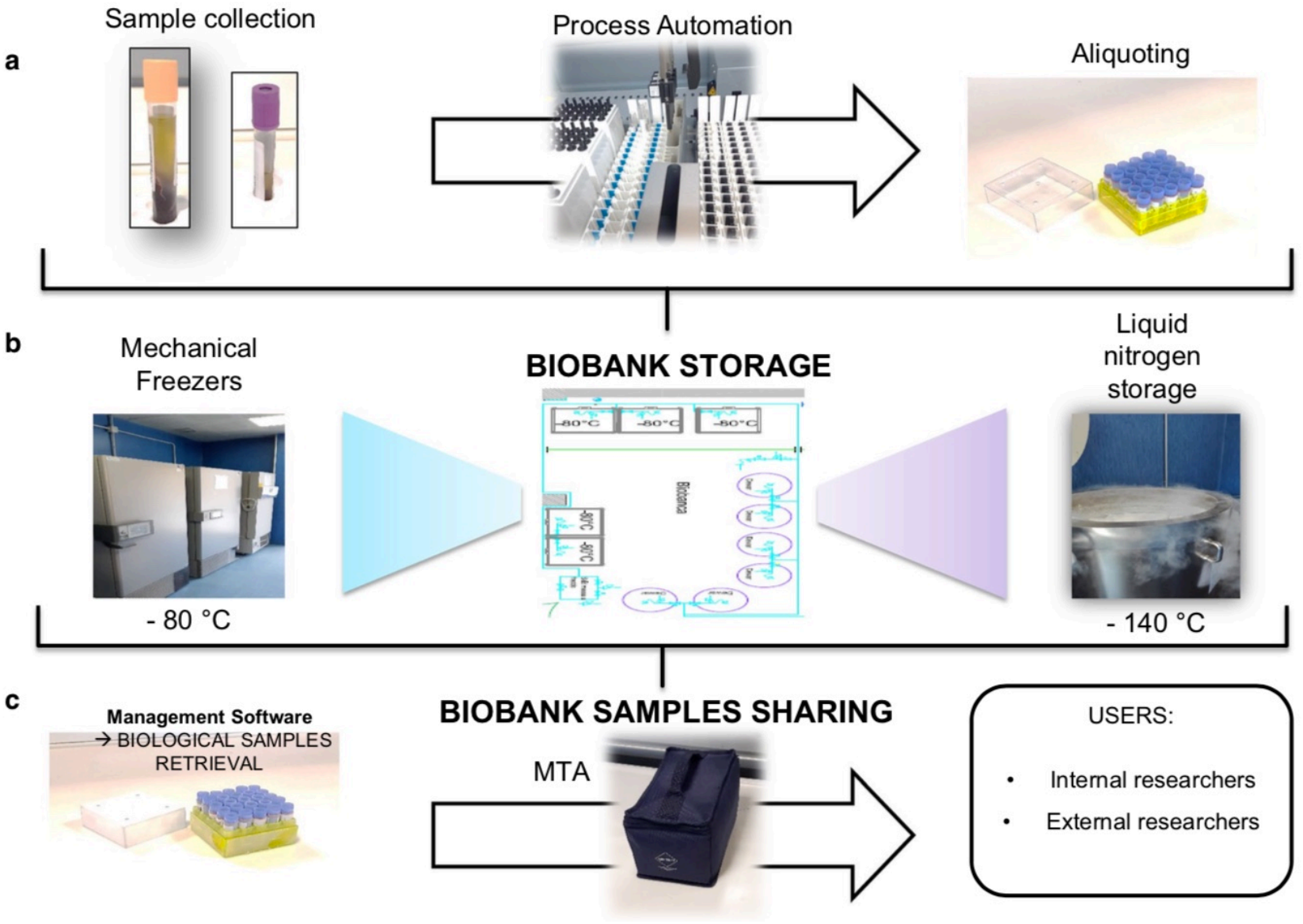
Pre-acquisition variables:

- Antibiotics
- Other drugs
- Type of Anesthesia
- Duration of Anesthesia
- Arterial Clamp time
- Blood pressure variations
- Intra-op blood loss
- Intra-op blood administration
- Intra-op fluid administration
- Pre-existing medical conditions
- Patient gender
- Ethnicity

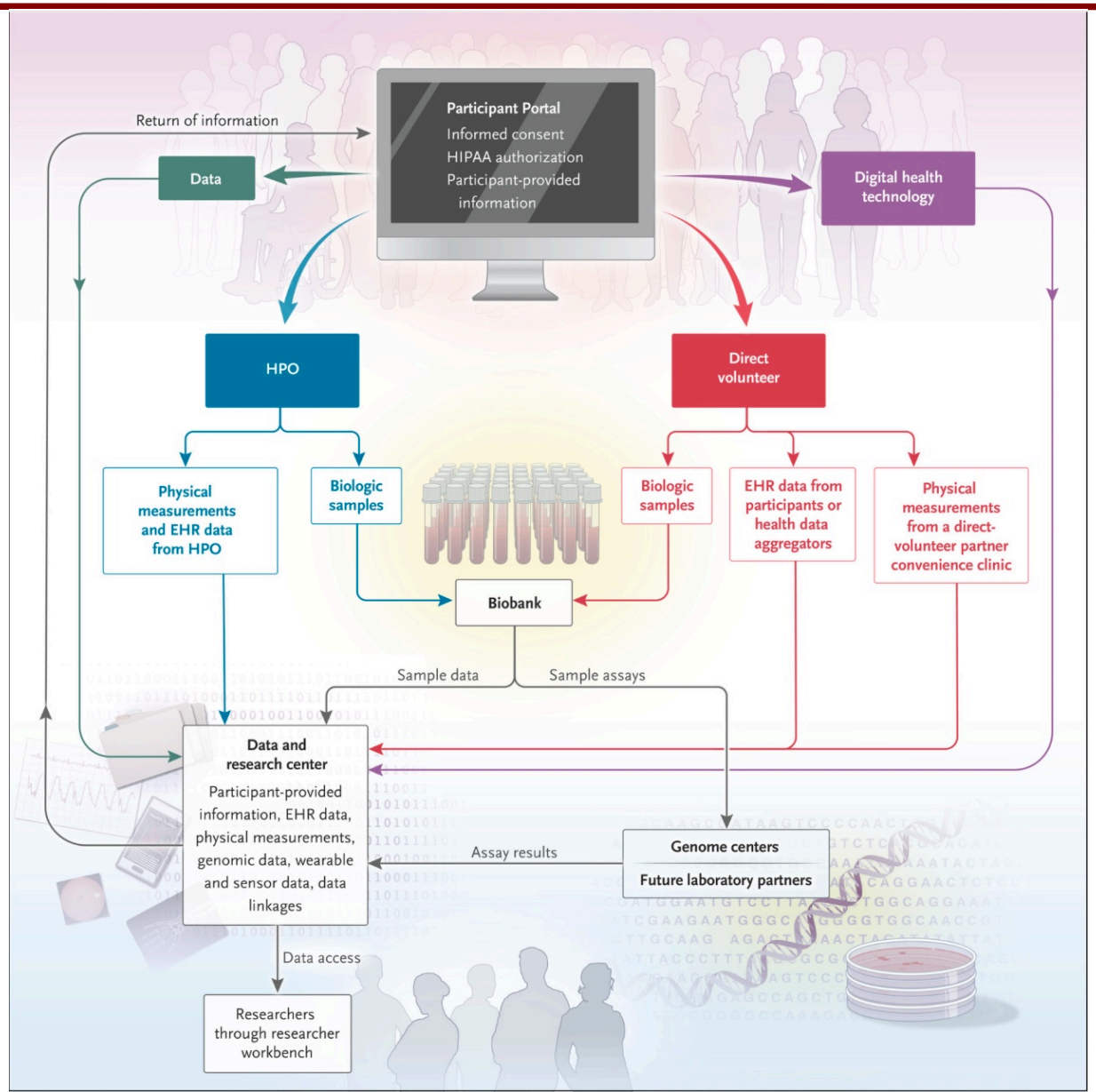
Post-acquisition variables:

- Time at room temperature
- Temperature of room
- Type of fixative
- Time of fixative
- Rate of freezing
- Processing method
- Size of aliquots
- Type of collection container
- Biomolecule extraction method
- Storage temperature
- Storage duration
- Storage in vacuum

The Ideal Work-flow



All of US


Return of information
Participant Portal
Informed consent
HIPAA authorization
Participant-provided information
Data
Digital health technology
HPO
Direct volunteer
Physical measurements and EHR data from HPO
Biologic samples
Biologic samples
EHR data from participants or health data aggregators
Physical measurements from a direct-volunteer partner convenience clinic
Biobank
Sample data
Sample assays
Data and research center
Participant-provided information, EHR data, physical measurements, genomic data, wearable and sensor data, data linkages
Assay results
Genome centers
Future laboratory partners
Data access
Researchers through researcher workbench

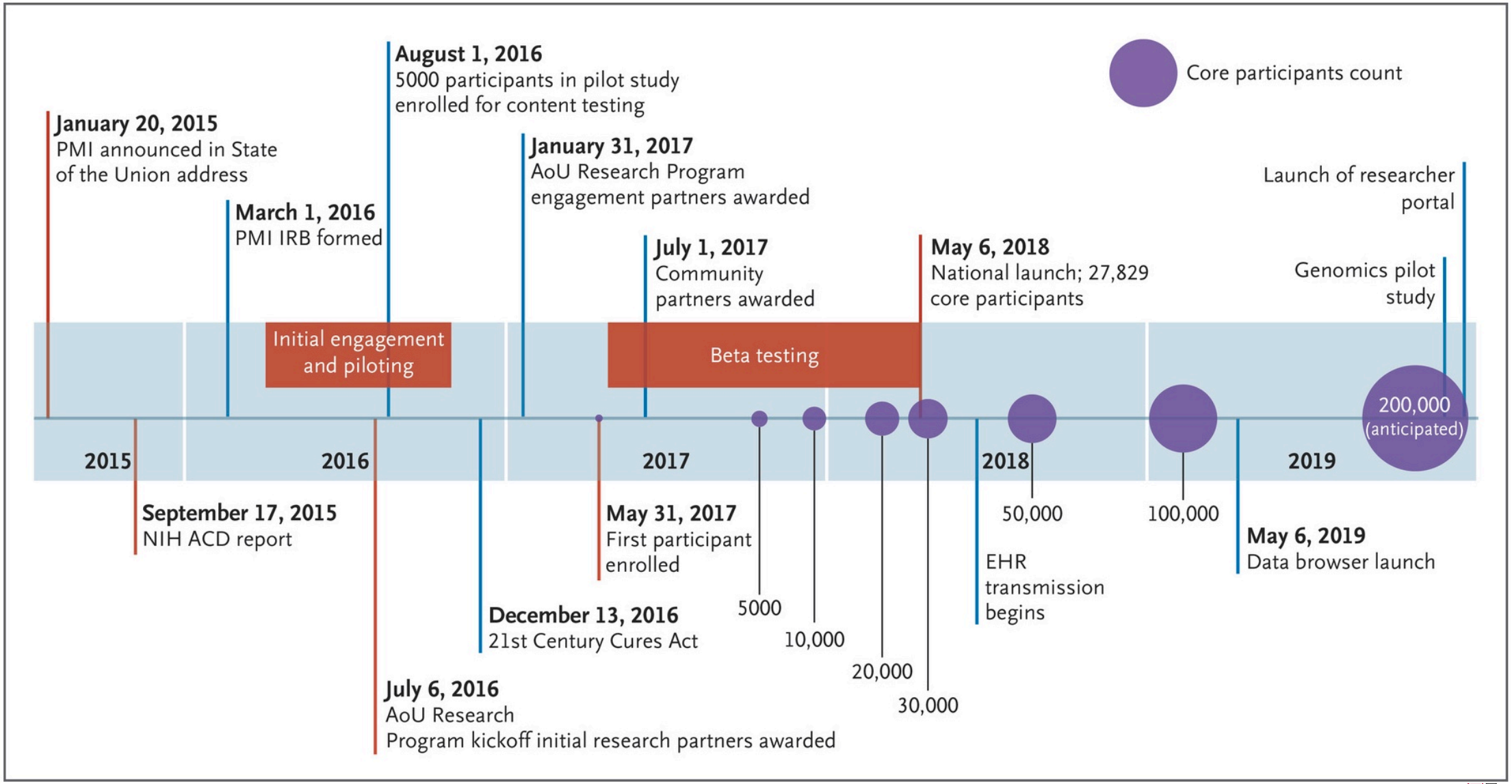
All of US

Table 1. Data Available to Researchers from the All of Us Cohort.*

Data Source	Details
Current sources	
Health surveys	Initial surveys include information on sociodemographic characteristics, overall health, lifestyle, and substance use, with subsequent modules covering personal and family medical history and access to health care.
Physical measurements	Per-protocol measurements include blood pressure, heart rate, weight, height, body-mass index, and hip and waist circumferences.
Biospecimens†	Blood and urine samples are tested for DNA, RNA, cell-free DNA, serum, and plasma. If blood specimens cannot be obtained, saliva specimens are obtained.
Electronic health records	Initial capture of structured data includes billing codes, medication history, laboratory results, vital signs, and encounter records from health care provider organizations. Records will be expanded to include narrative documents. Pilot studies are testing data collection through Sync for Science and other health data aggregators.
Digital health information	Data can be captured from compatible participant-owned devices such as Fitbit. Pilot studies of other devices and linkage to health apps are being explored.
Future sources	
Health surveys	Additional modules, including surveys regarding social behavioral determinants of health, are under development.
Bioassays	Pilot studies for genotyping and whole-genome sequencing are expected to begin by early 2020. Additional pilot studies of bioassays are planned.
Health care claims data	Systems for the use of claims data, including billing codes and medication data, are under development.
Geospatial and environmental data	These data include geospatial linkage to measures such as weather, air quality, pollutant levels, and census data. Assays and sensor-based measurements of exposure are under consideration.
Other sources	Voluntary contributions of data from social networks (e.g., Twitter feeds) and additional biospecimen collections are under consideration.

* Additional information is available at https://www.researchallofus.org/data.
† Types of biospecimens are listed in Table S3 in the Supplementary Appendix.

All of US


August 1, 2016
5000 participants in pilot study enrolled for content testing
Core participants count
January 20, 2015
PMI announced in State of the Union address
January 31, 2017
AoU Research Program engagement partners awarded
Launch of researcher portal
March 1, 2016
PMI IRB formed
July 1, 2017
Community partners awarded
May 6, 2018
National launch; 27,829 core participants
Genomics pilot study
Initial engagement and piloting
Beta testing
200,000 (anticipated)
2015
2016
2017
2018
2019
September 17, 2015
NIH ACD report
May 31, 2017
First participant enrolled
50,000
100,000
May 6, 2019
Data browser launch
EHR transmission begins
December 13, 2016
21st Century Cures Act
5000
10,000
20,000
30,000
July 6, 2016
AoU Research Program kickoff initial research partners awarded

All of US

Table 2. Scientific Goals of the All of Us Program and Expected Timelines.*

Goal	Years				
	End of 2018 (N = 94,000)	End of 2019 (N = >200,000)	2020–2022 (N = <650,000)	2023–2027 (N = >1 million)	After 2027 (N = >1 million)
Return data to participants	+	+	+++	+++	++++
Establish demonstration projects†		+	+++	+	+
Discover genetic and environmental correlates with disease			++	+++	++++
Improve predictions of therapeutic safety and efficacy			++	+++	+++
Discover disease biomarkers			++	+++	+++
Connect mobile health, digital health, and sensor data with clinical outcomes			++	+++	+++
Develop new disease classifications			+	+++	++++
Support clinical trials			+	+++	+++
Enable machine-learning applications			++	+++	++++
Improve understanding of health disparities			++	+++	+++
Develop and test new therapeutic agents					++

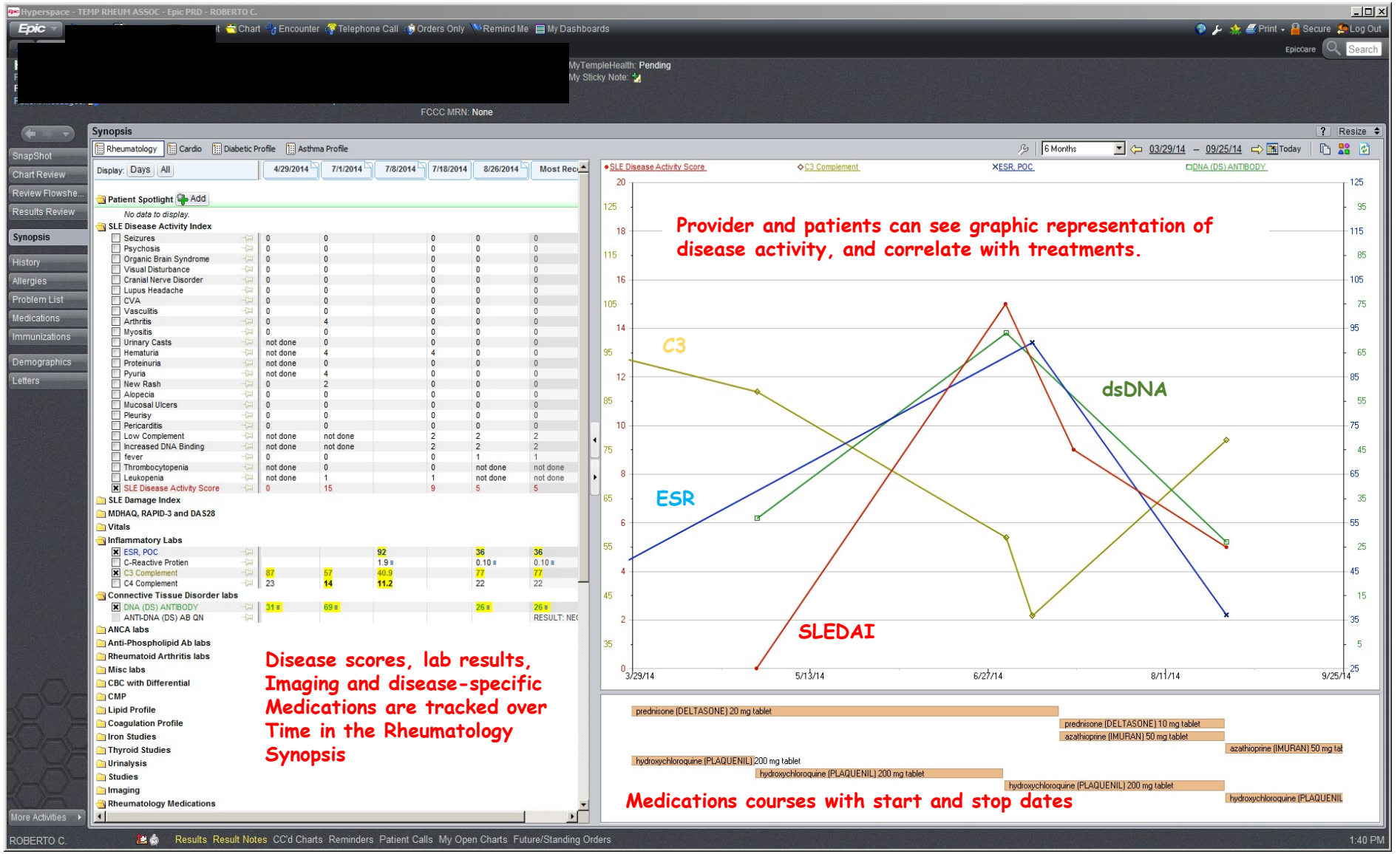
* The expected number of participants in the cohort is shown for each time period. The number of plus signs in each cell indicates the anticipated relative degree to which each goal may be accomplished during the estimated timeline for focused research.

† Demonstration projects are scientific studies implemented by the All of Us program to show the quality, usefulness, validity, and diversity of the All of Us research data set and platform. In these projects, the population and data are further characterized, and the data are evaluated with a view to determining whether known associations can be replicated.

The Temple Lupus Cohort

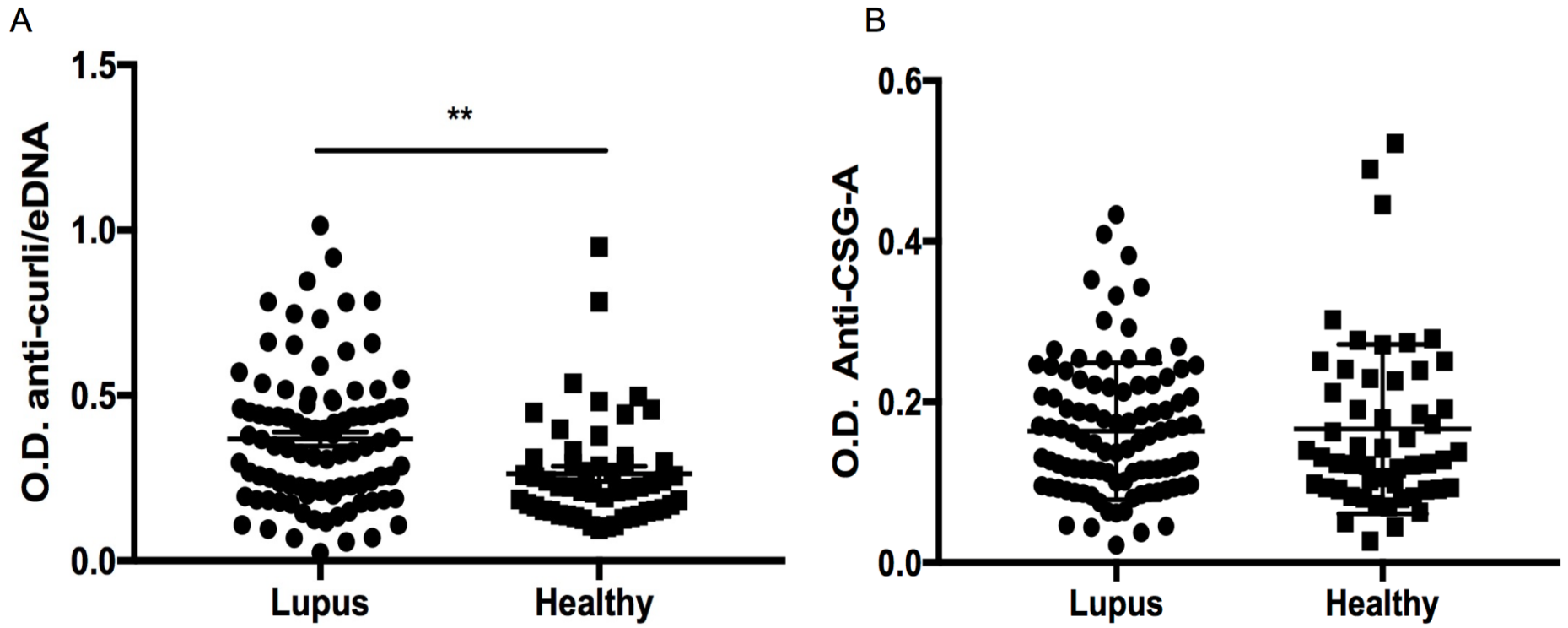
- Born from the need of a systematic database of Lupus patients to improve Clinical Trials and Quality of Care.
- Implemented along with the electronic medical record (EMR) which allows to collect systematic and consistent information on lupus patients at each visit.
- Integration of SLE disease activity (SLEDAI) and SLE damage index (SDI) measure outcomes with electronic charts during clinic encounters since November 2010
- Systematic collection of blood, urine samples and renal tissue to foster translational research and clinical trials

Temple Lupus Clinic and Electronic Medical Record

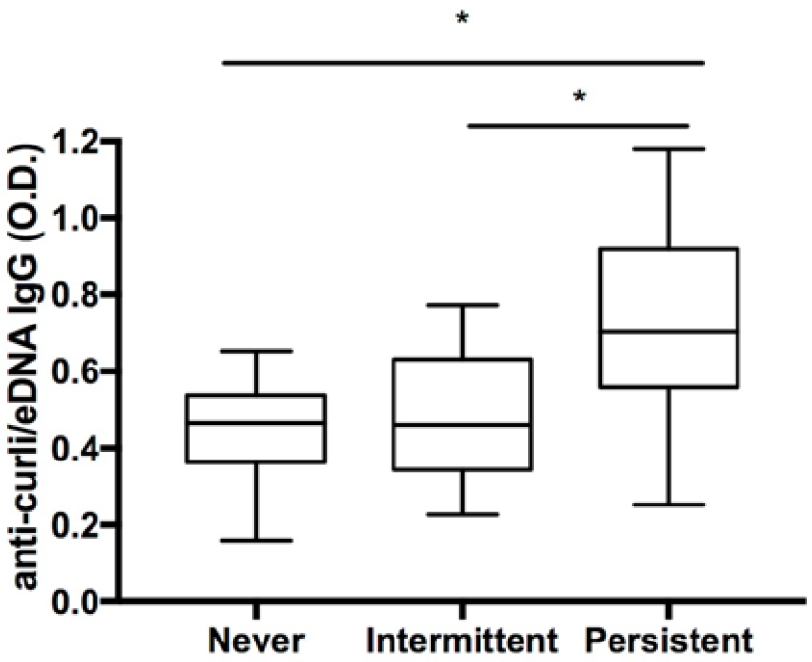


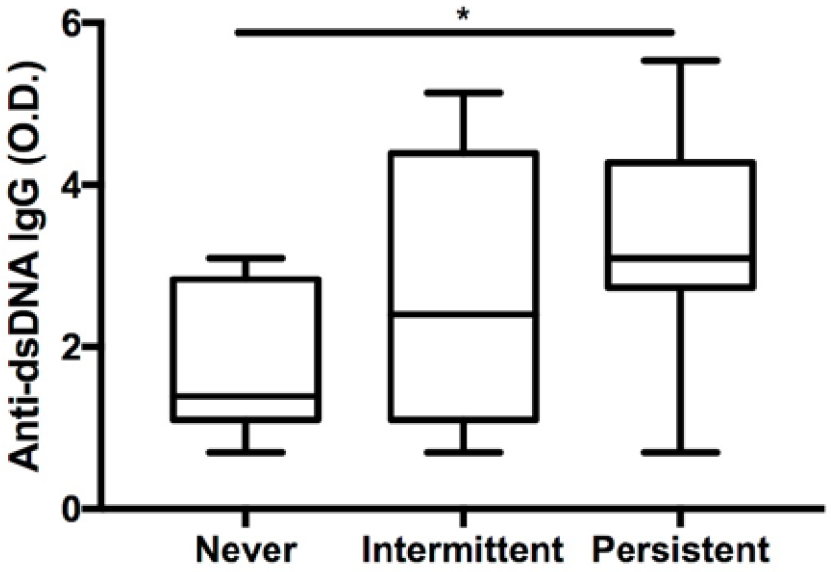
TEMPLE HEALTH
Lupus Program

SLE patient and healthy individuals have anti-Curli antibodies



Frequency of bacteriuria correlate with anti-curli/eDNA antibodies





Collaborators

New York University

Penn State

LKSoM

Janssen

LuCIN

NIH

Conclusion

- Biobanks are complex systems of systematically programmed storage of human material and associated data.

- In the past 20 years the science of biobanks has became an integral part of personalized medicine

- A great number of biobanks have been established all over the world to support the dramatic development in diseases prevention, prediction, diagnosis and treatment.

Acknowledgements

TEMPLE UNIVERSITY
Lewis Katz School of Medicine

- **Caricchio Laboratory**
  - Ryan Pachuki.
  - Xinyan Zhang
  - Lynne Kohler
- **Rheumatology Fellows**
- **Lupus Patients**

TEMPLE HEALTH
Lupus Program

NIAID

NIAMS

LUPUS RESEARCH ALLIANCE