The ethics of data and sample-based genomic research projects

Yann Joly, LLB, LLM, DCL, ADE, FCAHS

McGill

Outline

- Introduction
- Consent
- Confidentiality
- Re-contact of research participants
- Communication of research results
- Data sharing
- Benefit sharing
- Standardization (GA4GH)

Introduction

- I started working on the ethical, legal and social issues of genetic research in 1999 with Prof. Bartha M. Knoppers at the Université de Montréal. Director of the Centre of Genomics and Policy (CGP)
- Became interested in large-scale genomic research early in my career. I contributed to the development of governance policies and data access of genomics and health data for one of the first truly international genetic projects: The International Cancer Genome Consortium (ICGC) from 2009 to 2020
- I am an active member of GA4GH, I led the REWS (2019-24) and am now on the Strategic Leadership Committee of the organization. I am also Co-Chair of the ethics committee of the International Association for Forensic Genetics (IAFG)

Global Alliance for Genomics & Health

ICGC

Consent

Challenge: Can consent to a populational biobank, or database project, be truly informed?

- Community engagement
- Broad consent
- Dynamic consent
- The Future: GA4GH DUO

DISCOVER

Researchers can use DUO terms to discover relevant datasets

REQUEST

Researchers submit their data access request, encoded using DUO terms

MATCH

Researcher access requests are evaluated against the DUO-annotated datasets

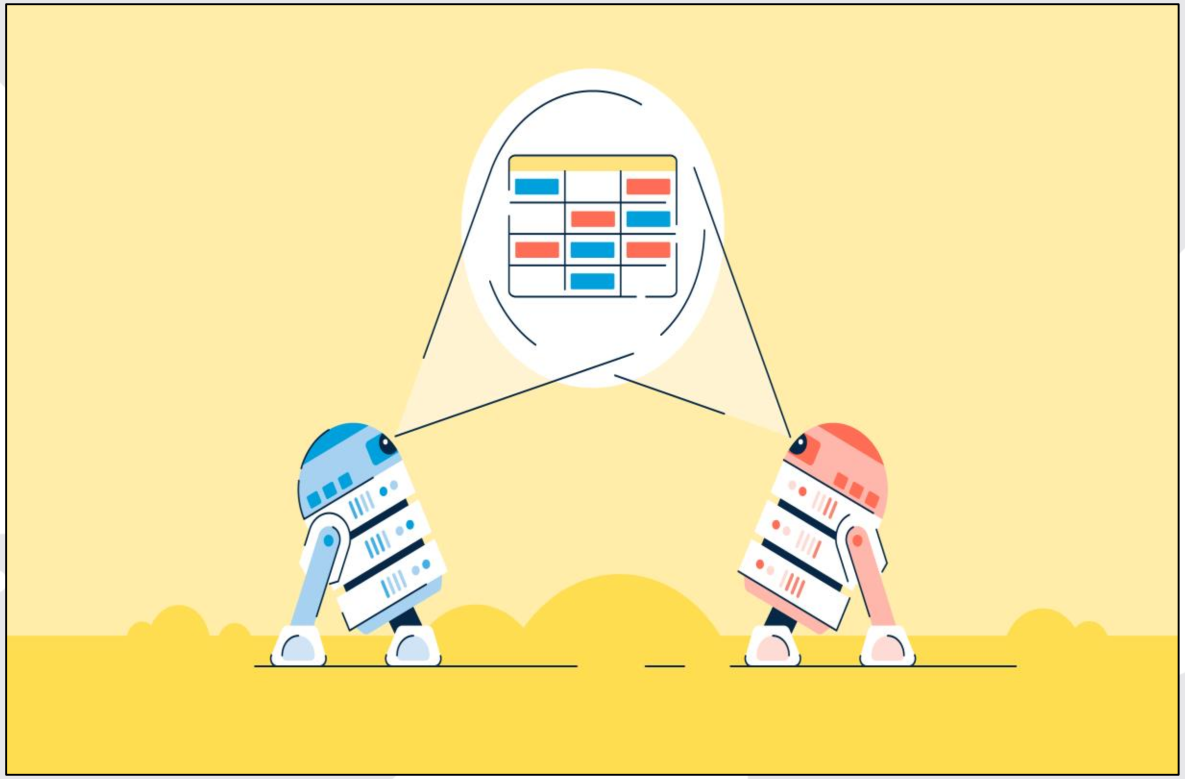
ACCESS

If approved, researchers receive access to datasets

Confidentiality

- **Challenge:** Genetic data is uniquely identifying, sensitive and a growing number of these data are available online
- **Challenge:** Large amount of genetic and other personal data is now available on the internet facilitating data matching aimed at re-identifying individuals

- De-identification (coding, double - coding, anonymization)
- IT security (firewall, encryption)
- Administrative control (controlled access)
- Secured analysis on protected server: data visiting
- **Current trend:** Data protectionism



Re-contact of research participants

- **Benefit:** Re-contact optimizes the value of a repository by 1) allowing researcher(s) to collect additional information and materials; 2) allowing the longitudinal follow-up of patients; 3) recruitment for targeted clinical trials for precision medicine
- **Challenge:** What is the benefit for participants?
- **Challenge:** Increased privacy risk
- **Challenge:** Lost to follow-up
- BS Return of significant clinical findings
- Draft a short protocol on re-contact (who does?, who authorizes?, how?, familial implications?)
- Adequate re-contact mechanism provided for in the informed consent document

Communication of research results

Challenge: We are entering the era of genomic medicine and precision therapies. Actionable results and incidental findings need to be communicated to 'research' participants. The clear distinction between research and medical care is becoming increasingly blurred

Challenge: Biobanks and databases created over 10-15 years ago often did not plan for the necessity of such re-contact

Challenge: Returning results can be a very expensive process. Especially when it comes at an unanticipated expense

- Not possible for anonymized studies
- Ask REB/IRB for permission to re-contact
- Online disclosure of results (with human operator or AI tool) Ex. Genetics Adviser https://www.geneticsadviser.com/

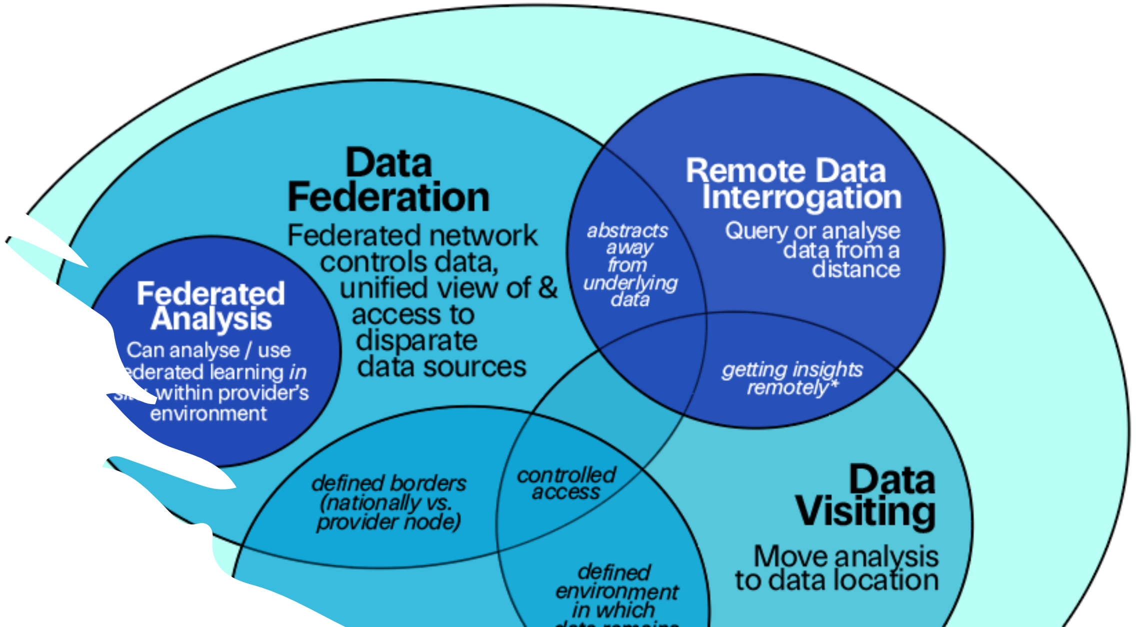
Data Sharing

Challenge: respecting data privacy

Challenge: lack of incentive for data producers

Challenge: data protectionism

- Transparency in consent documents
- Data access committee
- Acknowledgements and publication policy
- Follow GA4GH DACRES standards and read their monthly health data forum blog
- Awareness of cyber security policy applicable in your country/institution



Benefit sharing

- **Definition of benefit :** ''A good that contributes to the well-being of an individual and/or a given community'' HUGO, 1997
- **Challenge:** Benefit sharing is now expected for genetic research with vulnerable population groups/developing countries
- **Challenge:** Benefits are not explicitly provided for by funding agencies

❖ Benefit can be monetary or non-monetary

❖ Possibility to obtain funding for a specific benefit well-aligned with the objectives of your study (ex. training of local specialists in genetic counselling)

❖ Possibility to make benefit contingent to commercial success of project (but not necessarily recommended)

Standardization (GA4GH)

- International Society for Biological and Environmental Repository (ISBER)
- Global Alliance for Genomics and Health (GA4GH)
- World Health Organization (WHO)
- UNESCO International Bioethics Committee (IBC)
- Organization for Economic Co-operation and Development (OECD)

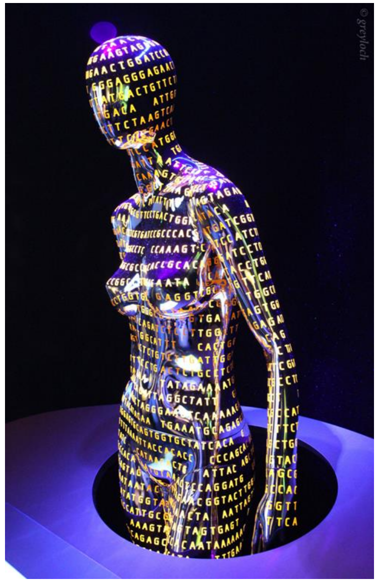
Standardization (GA4GH)

- Formed in 2013, the Global Alliance for Genomics and Health (GA4GH) unites an international community dedicated to advancing human health through genomic data.
- Objective: To build technical standards and policy frameworks and tools to expand responsible, voluntary, and secure use of genomic and other related health data.
- Scope: 500+ organizations collaborate to set international standards, guidelines and tools for genomic data use. 3,000,000+ genomes stewarded by our global partners with a commitment to responsible and effective access.
- 8 Work Streams create products that encourage responsible international genomic data use. These teams from around the globe build standards, frameworks, and tools that overcome technical and regulatory hurdles.
- 28 Driver Projects shape GA4GH products and apply them to real genomic data.

Standardization (GA4GH)

Data Diversity

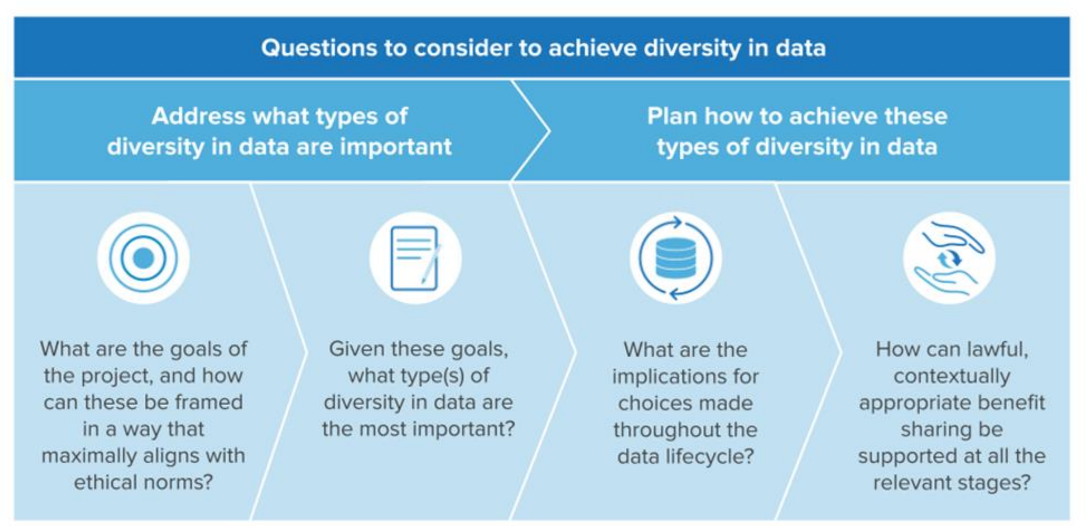
- The need to better understand the relevance of the human genome to human health and wellbeing has led to pleas for more ‘diverse’ data, and for changed practices, from a range of players in the global genomics field including funders, professional societies, genomic initiatives and publishers
- The benefits of increased diversity in datasets are twofold. First, increased power of scientific advances e.g. locus discovery, fine mapping, and causal inference. Second, health equity in the context of genomic medicine (the global and unbiased applicability of genomic knowledge to patients’ care)
- Yet, how can we properly assess the specific type of diversity that is needed by genomic researchers for a specific project?



GA4GH (Standardization)

Data Diversity

- Stage: Policy guidance adopted by GA4GH on Jan 2024, published in Nature Genetics
- Lead: Anna Middleton
- Methods: Literature review, expert workshops
- Proposed outcome : Policy guidance and scientific manuscript on diversity in datasets for genomic studies



Questions to consider to achieve diversity in data

Address what types of diversity in data are important

Plan how to achieve these types of diversity in data

What are the goals of the project, and how can these be framed in a way that maximally aligns with ethical norms?

Given these goals, what type(s) of diversity in data are the most important?

What are the implications for choices made throughout the data lifecycle?

How can lawful, contextually appropriate benefit sharing be supported at all the relevant stages?

GA4GH (Standardization)

Data Diversity

- There is no single definition of “diversity” in datasets
- To identify what types of differences are relevant, and why: start with the goals of the project, ensuring these are in pursuit of ethical norms
- Equity will often be the most important concept
- Diversity is not necessarily about representativeness
- Inclusion — i.e. making it as easy as possible for differently situated people to participate should they wish to do so — remains a constant mandate for every project

For considerations on the Data diversity policy framework in East Asia : Sato M., Muto K., Momozawa Y., Joly Y. *(Not so) Lost in Translation: Considering the GA4GH Diversity in Datasets Policy in the Japanese context ABR (2024). https://doi.org/10.1007/s41649-024-00305-5*

Thank you

Questions?

VICTOR PHILLIP DAHDALEH
INSTITUTE OF GENOMIC MEDICINE