Big Data meets Big Science and Big Ethics: Emerging Trends, Regulatory Challenges and Opportunities for Engagement

THE UNIVERSITY of EDINBURGH

Edward S. Dove
University of Edinburgh, School of Law

mason institute for medicine, life sciences & the law

Ethics Review Equivalency Task Team Coordinator,
Global Alliance for Genomics and Health

Big Science, Big Data…Big Ethics?

- Rise of the bioeconomy brings new practices of knowledge co-production, and new understandings about the nature of knowledge, to the fore
- Fosters new changes in Big Science, namely collective innovation and large-scale global science that work alongside key actors (e.g. bioethicists) in a distributed, data- and computation-intensive consortia system to solve large—and well-funded—scientific problems

Big Science, Big Data…Big Ethics?

- In Big Ethics, a hitherto dispersed group of bioethicists and related scholars now collaborate not just to share ideas, but to build knowledge together as one well-funded integrated unit with Big Science.
- There is now a symbiotic relationship between science and ethics as the funding systems for consortia science also fund consortia ethics under the integrated research model. Big Science and Big Ethics become intertwined, and indeed, rarely can one now exist without the other.

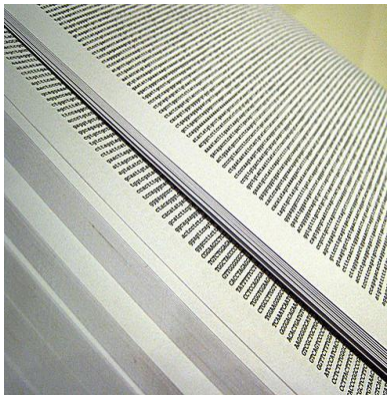
Big Science, Big Data…Big Ethics?

- Let us explore a particularly important area at the intersection of these three "Bigs"—namely, the sharing of genomic data within consortia to drive precision medicine, and the role of human rights.

The Challenge

Global Alliance for Genomics & Health

Unparalleled generation of human genetic data



How do we unlock its potential?

In a way that allows data to be shared on a global level

Thus empowering new knowledge, new diagnostics and new therapeutics for patients and populations

The Challenge

Global Alliance for Genomics & Health

Data from **millions of samples** may be needed to achieve results and progress - showing patterns that would otherwise remain obscure.

That will take new methods and organizational models.

Right now:

- Data is typically in silos: by type, by disease, by country, by institution
- Analysis methods are non-standardized, few at scale
- Approaches to regulation, consent and data sharing limit interoperability

If we don't act: risk an overwhelming mass of fragmented data, such as electronic medical records in many countries

Global Alliance
for Genomics & Health

Collaborate. Innovate. Accelerate.

What is the Global Alliance for Genomics and Health?

Mission of the Global Alliance

Global Alliance for Genomics & Health

To accelerate progress in human health by helping to establish a common framework of harmonized approaches to enable effective and responsible sharing of genomic and clinical data, and by catalyzing data sharing projects that drive and demonstrate the value of data sharing

Role

Global Alliance for Genomics & Health

Convene stakeholders

Catalyze sharing of data

Create harmonized approaches

Act as a clearinghouse

Foster innovation

Commit to responsible data sharing

Organizational Members

Global Alliance for Genomics & Health

Global Alliance members include:

1. Universities and research institutes (33%)
2. Academic medical centers and health systems (12%)
3. Disease advocacy organizations and patient groups (5%)
4. Consortia and professional societies (6%)
5. Funders and agencies (6%)
6. Life science and information technology companies (38%)

Global Alliance
for Genomics & Health

Collaborate. Innovate. Accelerate.

How does the Global Alliance operate?

Working Groups

Global Alliance for Genomics & Health

The **Clinical Working Group** aims to enable compatible, readily accessible, and scalable approaches for sharing clinical data and linking genomic data. Clinical Working Group strives to address both research and clinical use scenarios and be physician-oriented, researcher-focused, and patient-centered.

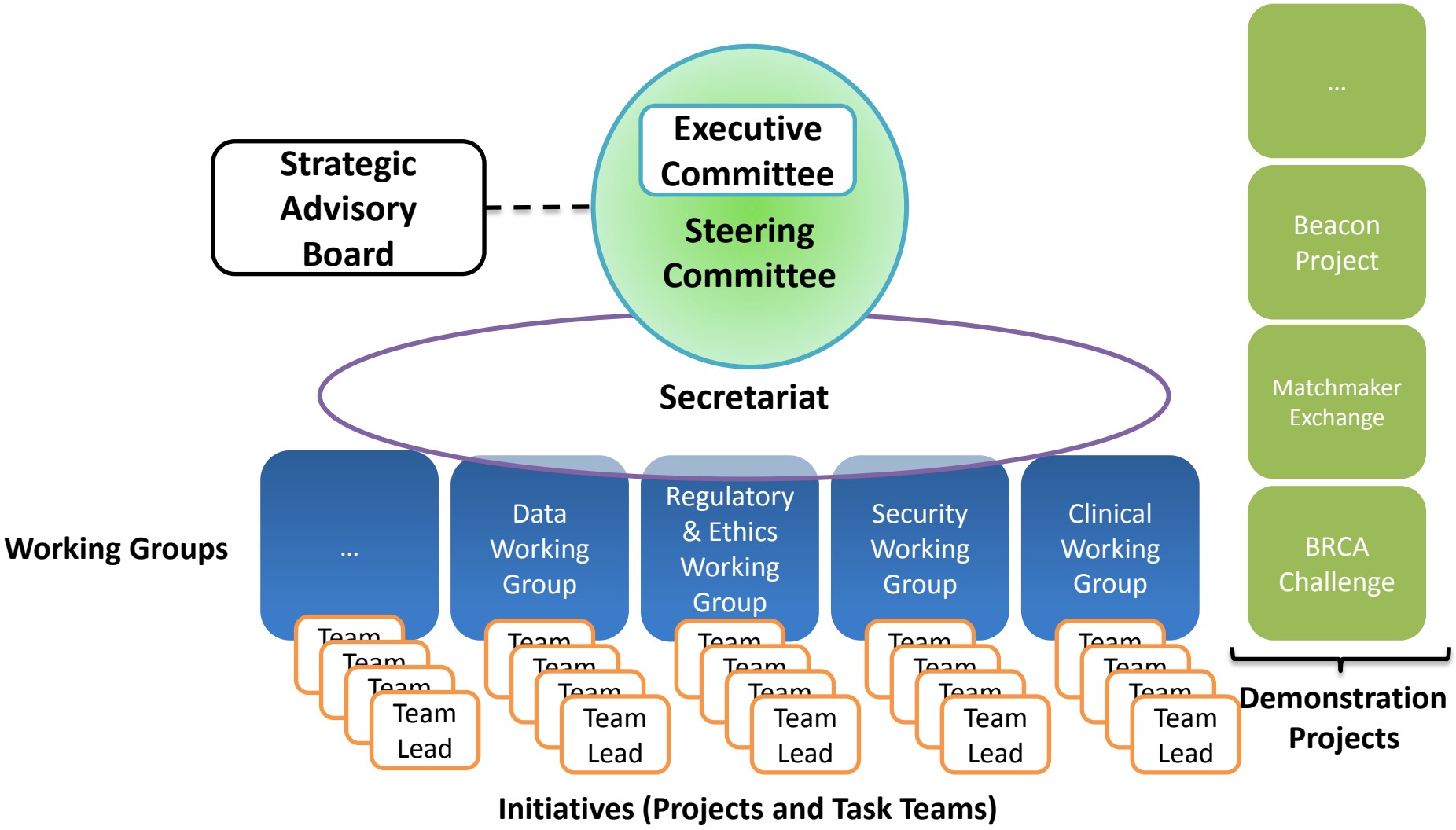
The **Data Working Group** concentrates on data representation, storage, and analysis of genomic data, including working with academic and industry leaders to develop approaches that facilitate interoperability.

The **Regulatory and Ethics Working Group** focuses on ethics and the legal and social implications of the Global Alliance, including harmonizing policies and standards, and developing forward-looking consent, privacy procedures, and best-practices in data governance and transparency.

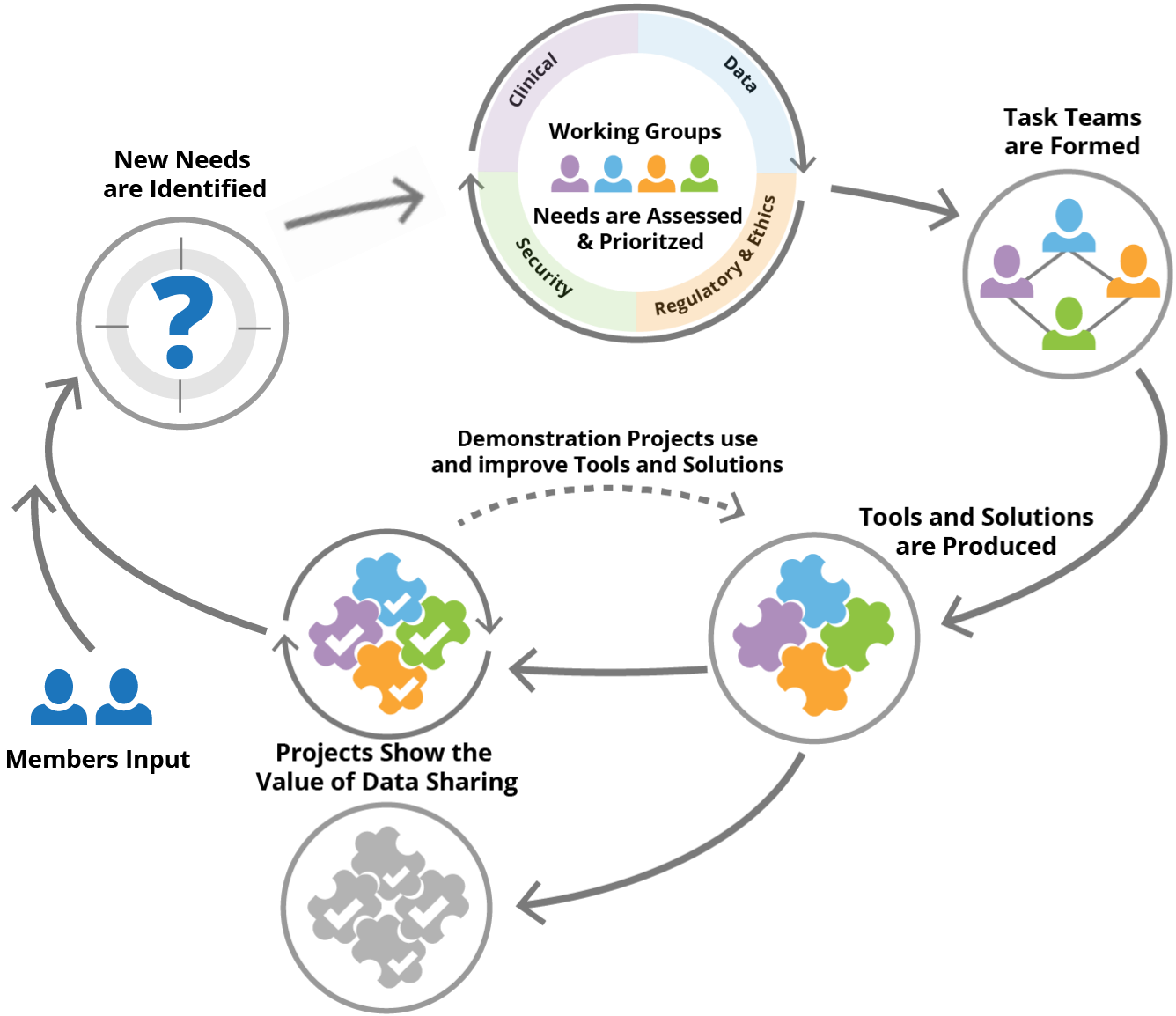
The **Security Working Group** leads the thinking on the technology aspects of data security, user access control, and audit functions, working to develop or adopt standards for data security, privacy protection, and user/owner access control.

Leadership model

Global Alliance for Genomics & Health



How We Work


New Needs
are Identified
Clinical
Data
Working Groups
Needs are Assessed
& Prioritzed
Security
Regulatory & Ethics
Task Teams
are Formed
Demonstration Projects use
and improve Tools and Solutions
Tools and Solutions
are Produced
Members Input
Projects Show the
Value of Data Sharing

Current Initiatives

Global Alliance for Genomics & Health

Clinical

Clinical Cancer Genome - Cancer Data Sharing
- White Paper + Survey/paper

Clinical Cancer Genome - Actionable Cancer Genome Initiative (P)
- Opinion paper + Standard

eHealth - Pedigree Consent
- Standard

eHealth - Family History
- Tool + Standard

eHealth - Federated Queries
- Standard

eHealth - Catalogue of Activities
- Tool

eHealth - Data Sharing
- White paper

Phenotype Ontologies - Rare Diseases
- Use cases

Phenotype Ontologies - Cancer / Complex Diseases

Data

Benchmarking
- Standard (variant calling benchmark toolkits)

Containers and Workflows
- Standard (interoperable GA4GH APIs with open source)

File Formats
- Standard (file formats)

Genotype2Phenotype Association
- Standard

Metadata
- Standard (metadata schema)

Reference Implementation
- Standard (GA4GH API)

Reference Variation
- Standard (graph-based reference)
- Implementation

RNA and Gene Expression
- Standard (RNA sequence reads)

Variant Annotation
- Standard (data representation)

Regulatory and Ethics

Accountability (Policy)

Ageing and Dementia (P)
- Policy
- Tools (authorization/consent)

BRCA Ethico-Legal and Advocacy (P)
- Policy (data protection / liability)
- Tool (tiered access model)
- Tool (patient advocacy network)

Data Protection Regulation

Data Sharing Lexicon
- Tool (Catalogue)

Ethics Review Equivalency
- Resource (mutual recognition model)

Individual Access
- Tool (international survey)
- Resource (review of laws)

Machine-Readable Consent
- Standard (data use types in consent)

Paediatric
- Policy (WGS and Newborn Screening)

Privacy Breach Notification (P)
- Resource (review of laws)

Registered Access (Tool) (P)

Security

Cloud Security
- Policy (guidance for cloud security providers)
- Implementation

Incident Response
- Policy

Security Infrastructure
- Policy

Software Security
- Policy
- Implementation

Global Alliance
for Genomics & Health

Collaborate. Innovate. Accelerate.

Regulatory and Ethics Working Group

Mandate and Scope

Global Alliance for Genomics & Health

The Regulatory and Ethics Working Group (REWG) is tasked with addressing the ethical, legal and social implications of enabling responsible genomic and clinical data sharing. Specifically, the REWG:

- Prepares the overall policy framework for responsible data sharing;
- Harmonizes policies, procedures, standards and codes of conduct for the storage, analysis and sharing of genomic and clinical data;
- Develops forward-looking consent and privacy/security procedures that responsibly engage participants and researchers; and
- Develops best practices in governance and transparency of data repositories.

Global Alliance
for Genomics & Health

Collaborate. Innovate. Accelerate.

Building the *Framework for Responsible Sharing of Genomic and Health-Related Data*

Obstacles to data sharing in genomics

- Lack of policy harmonization
- Lack of structural support
- Legal and ethical hurdles
- Cultural barriers

Hum Genet (2014) 133:895–903
DOI 10.1007/s00439-014-1432-6

ORIGINAL INVESTIGATION

A human rights approach to an international code of conduct for genomic and clinical data sharing

Bartha M. Knoppers · Jennifer R. Harris ·
Isabelle Budin-Ljøsne · Edward S. Dove

Why human rights for a Framework?

- Universalizing force
- Political and legal dimensions that reach beyond the moral appeals of bioethics
- International legal force
- Belong to groups as well as individuals (reciprocity)
- Impose positive duties on governments and private actors


THE UNIVERSAL DECLARATION
OF Human Rights

Universal Declaration of Human Rights (1948)

"The Right to Science"

27(1) "Everyone has the right freely to participate in the cultural life of the community, to enjoy the arts and to share in scientific advancement and its benefits."

"The Right to Recognition for Scientific Production"

27(2) "Everyone has the right to the protection of the moral and material interests resulting from any scientific, literary or artistic production of which he is the author."

Legal force of the right to science

- Dual rights rendered legally binding by the International Covenant on Economic, Social and Cultural Rights (ICESCR) (1966) – Article 15.
- 164 States have ratified the ICESCR.
- States are bound to implement the treaty in their national laws.

The Right to Science → Actionable

Renewed Interest

The right to science "has acquired an increased importance in today's globalized world."
(UNESCO, *Venice Statement*, 2009).

"The scope, normative content and obligations of the State under … 'the right to science', remain underdeveloped while scientific innovations are changing human existence in ways … inconceivable a few decades ago."
(UN Special Rapporteur, A/HRC/20/26, 2012).

Content of the Right to Science

1. access by everyone without discrimination to the benefits of science and its applications;
2. opportunities for all to contribute to the scientific enterprise
3. the freedom indispensable for scientific research;
4. participation of individuals and communities in decision-making;
5. conservation, development and diffusion of science and technology.

(UNESCO, *Venice Statement*, 2009)

Right to Science = Access to Data

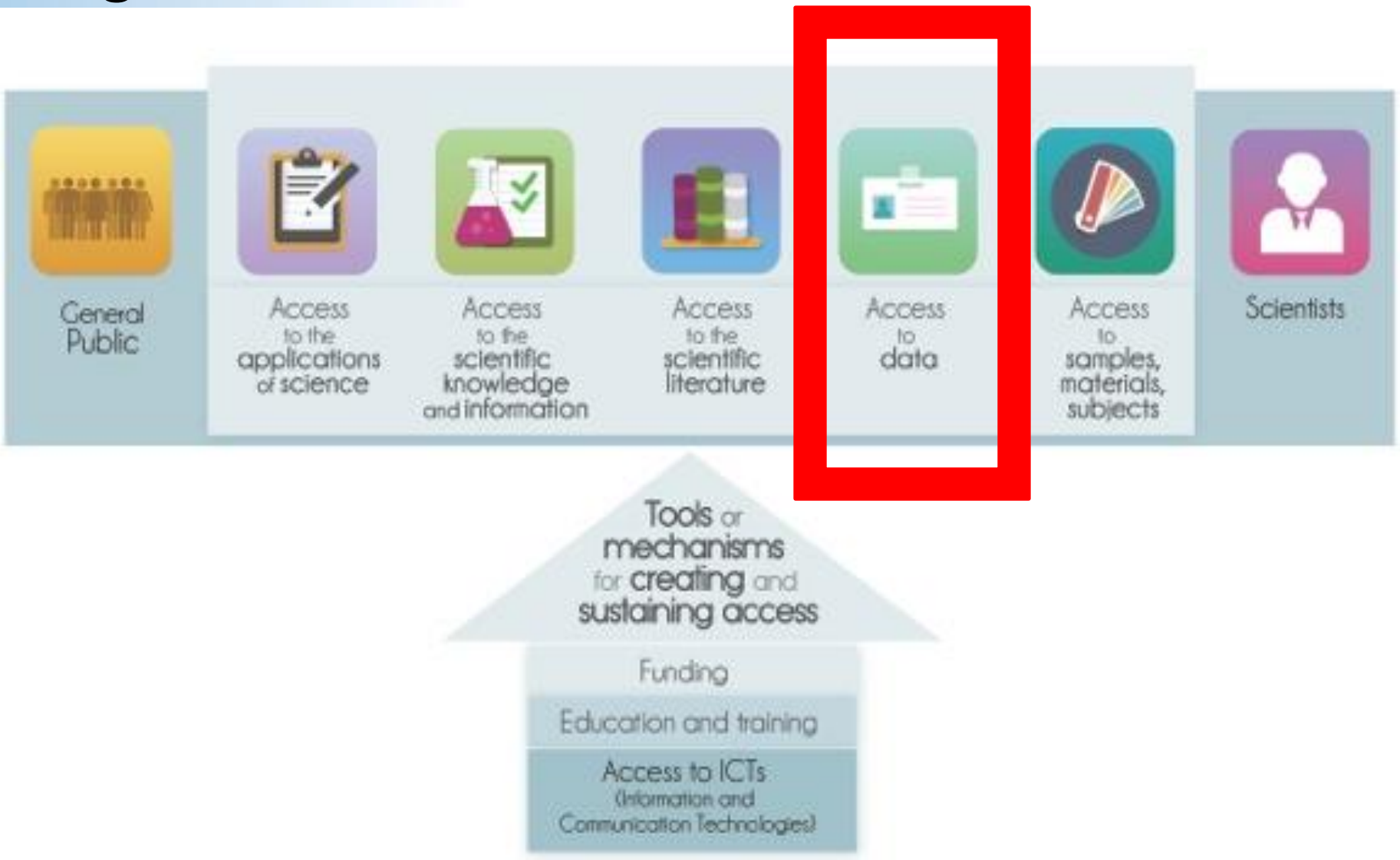


Image based on American Association for the Advancement of Science (AAAS, 2010)

Global Alliance for Genomics & Health

GA4GH *Framework for Responsible Sharing of Genomic and Health-Related Data*

- Current frameworks are founded on the principle of protection from harm. In contrast,
- The GA4GH Framework aims to **activate** the human right to science and the human right to recognition for scientific production by promoting responsible data sharing.

(http://genomicsandhealth.org/framework)

Development of the *Framework*

Global Alliance for Genomics & Health

- **October 2013 – February 2014** : Initial discussions and early drafts (Drafts 1-3) between the Global Alliance for Genomics and Health (GA4GH) Regulatory and Ethics Working Group (REWG) and international organizations
- **March 4-5, 2014** : GA4GH and REWG Meetings in London (Draft 4)
- **April 2, 2014** : P3G/BioSHaRE/INSERM Meeting in Paris (Draft 5)
- **May 1, 2014** : Draft 6 released to international organizations and on Global Alliance website for public comment until July 1
- **July 24, 2014** : Draft 7 released for public comment; endorsed by the Transitional Steering Committee of the GA4GH
- **September 10, 2014** : *Framework* finalized
- **October 10, 2014**: *Framework* translated into 5 languages: French, Spanish, Arabic, Chinese, Japanese
- **October 18, 2014**: *Framework* published (open access) in *The HUGO Journal*

Framework partners

Global Alliance for Genomics & Health

- Global Alliance for Genomics and Health (Regulatory and Ethics Working Group, other Working Groups and Task Teams);
- Biobank Standardisation and Harmonisation for Research Excellence project (BioSHaRE);
- Centre for Law and Genetics (University of Tasmania);
- Centre of Genomics and Policy (McGill University);
- ELSI 2.0;
- H3Africa;
- Health Research Authority (UK);
- HeLEX (University of Oxford);
- Human Variome Project (HVP);
- INSERM;
- International Cancer Genome Consortium (ICGC);
- International Rare Disease Research Consortium (IRDiRC);
- International Society for Biological and Environmental Repositories (ISBER);
- Personal Genome Project (USA);
- PErsonalised Risk Stratification for Prevention and Early deteCTIon of breast cancer project (PERSPECTIVE);
- PHG Foundation (UK); and
- Public Population Project in Genomics and Society-International Policy interoperability and data Access Clearinghouse (P3G-IPAC).

Framework: Goals

Global Alliance for Genomics & Health

Purpose. The purpose of this Framework is to provide a principled and practical framework for the responsible sharing of genomic and health-related data. Its primary goals are to:

i. Protect and promote the welfare, rights, and interests of individuals from around the world in genomic and health-related data sharing, particularly those who contribute their data for biomedical research;

ii. Complement laws and regulations on privacy and personal data protection, as well as policies and codes of conduct for the ethical governance of research;

iii. Foster responsible data sharing and oversight of research data systems;

iv. Establish a framework for greater international data sharing, collaboration and good governance;

v. Serve as a dynamic instrument that can respond to future developments in the science, technology, and practices of genomic and health-related data sharing;

vi. Serve as a tool for the evaluation of responsible research by research ethics committees and data access committees; and

vii. Provide overarching principles to be respected in developing legally-binding tools such as data access agreements.

Framework: Foundational Principles

Global Alliance for Genomics & Health

Foundational Principles for Responsible Sharing of Genomic and Health-Related Data

- Respect Individuals, Families and Communities
- Advance Research and Scientific Knowledge
- Promote Health, Wellbeing and the Fair Distribution of Benefits
- Foster Trust, Integrity and Reciprocity

Global Alliance for Genomics & Health

Core Elements for Responsible Data Sharing

Transparency

Accountability

Engagement

Data Quality and Security

Privacy, Data Protection and Confidentiality

Risk-Benefit Analysis

Recognition and Attribution

Sustainability

Education and Training

Accessibility and Dissemination

Framework for Responsible Sharing of Genomic and Health-Related Data

Global Alliance for Genomics & Health

The *Framework* is currently available in 10 languages. Thank you to all the volunteers!

- Arabic — إطار لتبادل مسؤول للمعلومات الجينومية والمتصلة بالصحة
- Chinese — 基因组学与健康相关数据负责任的共享框架
- French — Cadre pour un partage responsable des données génomiques et des données de santé
- Greek — Πλαίσιο για την Υπεύθυνη Κοινοχρησία Γονιδιωματικών και άλλων Ιατρικών Δεδομένων
- Japanese — **ゲノム及び健康関連データの責任ある共有に関する枠組み**
- Portuguese — Framework para Compartilhamento Responsável de Dados Genômicos e Relacionados à Saúde
- Spanish — Marco de actuación para el uso compartido responsable de datos genómicos y relativos a la salud
- German — Rahmenkonzept für die den verantwortungsvollen Datenaustausch genomischer und gesundheitsbezogener Daten
- Hindi — जीनोमिकी और स्वास्थ्य संबंधी डेटा को उत्तरदायित्वपूर्ण रूप से साझा करने के लिए रूपरेखा

Next steps after *Framework*?

Global Alliance for Genomics & Health

This Framework will be elaborated by subsequent Policies (Appendix 2) on particular issues such as ethical governance, consent, privacy and security. The Framework and its subsequent Policies should be used in projects around the world (whether Global Alliance "inspired" or not) such that they become the tools that approval entities, recognized by different jurisdictions, will turn or refer to for guidance. Recognizing diversity of legal and ethical approaches and being responsive to emerging issues, both this Framework and its Policies are intended to provide leadership in this domain for wider discussion.

Policies: Consent

Global Alliance for Genomics & Health

- **Purpose:** To guide international data sharing in a way that respects autonomous decision making while promoting the common good.
- Gives principled and practical guidance on consent issues:
  - Transparency
  - Privacy Safeguards
  - Withdrawal
  - Consent to access policies and terms
  - Sharing of Legacy Data
    - Notification and opt-out, or re-consent, for international data sharing if possible and practical. Otherwise seek authorization from competent authority.

Policies: Privacy and Security

Global Alliance for Genomics & Health

- How to manage privacy and security risks and related expectations?
- How to ensure data use is consistent with individuals' expectations, and respects the rights of individuals?
- **Consent:** Data should be used strictly in accordance with the Data Donor's consent for use and sharing, and/or the terms and conditions of authorization for use by competent bodies
- **Proportionate Safeguards:** Data privacy safeguards should be proportionate to the sensitivity, nature, and possible benefits, risks, and uses of the Data
- Security: **organizational, technical and physical measures** to manage risks to privacy and data integrity.

Policies: Accountability

Global Alliance for Genomics & Health

- Sets forth how Members of the Global Alliance can be governed and held accountable for the trust they engender in the responsible sharing of genomic and clinical data.
- Explores mechanisms to **promote** responsible data sharing, as well as to **prevent** and appropriately **sanction** data misuse.
- Possible Inclusion: Why are you NOT sharing, if the participant consented to sharing?

Policies: Ethics Review Mutual Recognition

Global Alliance for Genomics & Health

- Developing models that allow for mutual recognition of ethics review processes, particularly in data-intensive and international collaborative research. Irrespective of proposed models, would need to include:

- **Mutual recognition**. Requires harmonization/interoperability of:
    - Principles, forms, IT platform and operating procedures
    - Processes for ethical evaluation + oversight (per research type)
    - RECs - Latitude for some degree of local interpretation

International Charter of principles for sharing bio-specimens and data

EJHG Open

European Journal of Human Genetics (2015) 23, 721–728

www.nature.com/ejhg

POLICY

International Charter of principles for sharing bio-specimens and data

Deborah Mascalzoni[*,1,2], Edward S Dove[3], Yaffa Rubinstein[4], Hugh JS Dawkins[5,6,7,8], Anna Kole[9], Pauline McCormack[10], Simon Woods[10], Olaf Riess[11], Franz Schaefer[12,13], Hanns Lochmüller[10], Bartha M Knoppers[3] and Mats Hansson[1]

International Charter of principles for sharing bio-specimens and data

- **Five principles** for the custodianship of bio-specimen repositories and data constitute the common premise for the Charter:
  1. **Respect for privacy and autonomy**: custodianship implies protection of participants' privacy. Privacy protection measures should be in place and informed consent must provide provisions for future as yet unspecified research using data and bio-specimens.
  2. **Reciprocity**: custodianship also implies giving back. Feedback of general results should be channelled to institutions and patients.
  3. **Freedom of scientific enquiry**: custodianship should encourage openness of scientific enquiry, and should maximize data and bio-specimen use and sharing so as to exploit their full potential to promote health.
  4. **Attribution**: the intellectual investment of investigators involved in the creation of data registries and bio-repositories is often substantial, and could be acknowledged by mutual agreement.
  5. **Respect for intellectual property**: the sharing of data and biospecimens needs to protect proprietary information and address the requirements of institutions and third-party funders.

MTA/DTA Template

MTA/DTA TEMPLATE

Mutual agreement between provider and recipient

Provider of data……..

Recipient of data…….. (This must be an institution officially registered by national/regional authorities.)

MTA/DTA must be signed before any exchange takes place

1. Provider. Provider…….. hereby declares that:

The (name country) …….. from which the human biological material is collected has a legal and ethical framework providing a high level of quality, security and privacy protection concerning medical research involving human biological material; and that health data exchange is in accordance with local regulation (please note that some regulations, such as those from the EU, require compliance with their rules even if the research is conducted abroad).

Data/bio-specimens provided consist of the following: …….. (description including type of material and type of data: primary data and which type, genotypes, aggregate data, etc.)

The bio-specimens provided refer to …. (no. of individuals) and are composed of ….. (no. of tubes and quantity of material referred to the scope).

The material will be de-identified, stripped of all personally identifying information, without any direct means of identification. Bio-specimens will be double-coded (no direct identifier shall be on the tubes).

To ensure the confidentiality and security of the associated data, transfer and processing will be handled safely and associated data will not be transported together with bio-specimens.

To ensure traceability of the material to be de-identified, a code will be applied to the tubes. The mechanism to re-identify the data will remain with the provider.

To ensure exchange and transport security, biosafety (packaging, labelling description of transport means and insurance for bio-specimens) will be observed.

The project from which the data and bio-specimens were collected was approved by a local ethics committee or IRB.

Informed consent for storage and distribution was obtained (enclose a copy of the model in use), and that:

- the informed consent contained the clause necessary for allowing bio-specimens and data sharing abroad (in case of an international project).
- the informed consent allowed the research described in the project description section.

2. Recipient. Recipient…….. hereby declares that:

- the data will be used solely for the scope of the project described below, and no attempt will be made to sell the data/material or share it with a third party. The data/samples will be used for the following purpose: (description of the biomedical research project of the Recipient (or of the joint project): ……..(aims should be clear, including the duration of the project……..
- authorization from the local ethics review committee or IRB ……(date and copy enclosed).
- when consented to, will ensure the return of results relevant to the health of individuals …….. and …….. ……..(description of the type of data that must be returned).
- will not harm the persons who provided bio-specimens by naming the provenience of the bio-specimens unless approved by the Provider.

3. Terms and conditions (for Recipient). Conditions of use include:

- no attempt to re-identify the participants
- adherence to use limitations stated in approved application
- no third-party data or sample sharing/selling without authorization from Provider
- primary data must not be patented
- that the use of the material has, for example, medical/public health objectives
- informing the resource of issues related to data integrity and/or the privacy of the participants as applicable
- compliance with original consents and applicable laws and institutional policies
- access granted for a limited time period (eg, 6 months or ……), after which Recipient must reapply.

Documents to be provided by the recipient institution:

1. Authorization from a local ethics committee or a regional or local competent authority for the project for which the data are provided (by Recipient or, in the case of a joint project, by both parties).
2. Documents by the national data protection authority that reference the applicable laws that allow research using sensitive and health data, including genetic data.

4. Receipt and handling of imported biological material. The Recipient must document and follow the procedures for the receipt and proper storage of the type of biological material handled.

Provider has to prepare and ship the biological material in accordance with postal regulations, such as IATA (International Air Transport Association) and ADR (European Agreement on International Carriage of Dangerous Goods).

5. Publication. Prior review of publications before submission may be required (eg, to ensure the privacy/confidentiality of data and that the results will not cause stigmatization). Recipients should acknowledge the biobank/data provider in any publication/presentation (or other clauses).

6. Is the material used to be returned or destroyed? Recipient will comply with the destruction/return of unused bio-specimens and of data related to the bio-specimens at the end of the project or of the duration stated above

Description of the requirement (destruction/return)……..

7. Intellectual property rights. (Requires a specific agreement case by case. See general Introduction)……..

8. Who controls the data/bio-specimens in the resource? Control of the bio-specimens remains with Provider, who can at any time demand the return or destruction of data and bio-specimens if a breach in the agreement occurs.

9. Obligation to report. Annual…….. (or other) and final reports to Provider are required. Reports should include …….. (specify required content).

10. Responsibilities of the biobank/consortia. Biobanks and research consortia have the right to terminate/alter this agreement with the researcher/institution, for the safety of the patients/participants or because of any infringements of the obligations stated in the present DTA/MTA. The Recipient understands and agrees that Provider does not bear any responsibility or accept any liability arising from the Recipient's use of the data or bio-specimens.

Time, place and signatures of the persons legally responsible for the institutions involved.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank the RD Connect Patient Ethics Council and RD Connect Patient Advisory Council (Jean Jacques Cassiman, Tracy Dudding, Muriel Gevrey, Emma Heslop, Joseph Irwin, Julian Isla, Sigurður Jóhannesson, Lydia Lemonnier, Chantal Loirat, Dorthe Lykke, Milan Macek, Caron Molster, Kay Parkinson, Odile Perrousseaux, Marita Pohlschmidt, Daniel Renault, Peter Reussner, Francoise Rouault, Balthasar Schaap, Inge Schwersenz, Chris Sotirelis, Volker Straub, Oliver Timmis, Johannes van Delden, Marieke van Meel, Elizabeth Vroom, Urban Wiesing) and all RD-connect partners for their valuable inputs. This work has been supported by the European Union Seventh Framework Programme (FP7/2007–2013) under grant agreements no. 305444 (RD-Connect), 305121 (Neuromics), and no. 305608 (EURenOmics) and RD Connect from the Australian National Health and Medical Research Council APP1055319 under the NHMRC-European Union Collaborative Research Grants scheme'. In addition, DM and MH have received funding from the Innovative Medicines Initiative project BTCure (grant agreement no. 115142-1), the BioBanking and Molecular Resource Infrastructure of Sweden project (financed by the Swedish Research Council), Euro-TEAM, BiobankCloud, BioSHaRE and Biobanking and Biomolecular Resources Research Infrastructure (BBMRI)LPC (313010). BMK and ESD wish to thank the Government of Canada through Genome Canada and the Canadian Institutes of Health Research, the Quebec Breast Cancer Foundation, and the Ministère de l'enseignement supérieur, de la recherche, de la science et de la technologie du Québec through Génome Québec. The funders had no influence on the design or the writing of the article.

Thank you!

edward.dove@ed.ac.uk

@EdwardSDove