Global Alliance for Genomics & Health
Collaborate. Innovate. Accelerate.

POL 008 / v1.0: October 27 2021

Global Alliance for Genomics and Health: Data Access Committee Guiding Principles and Procedural Standards Policy

Preamble

This document is the Global Alliance for Genomics and Health's (GA4GH's) Data Access Committee Guiding Principles and Procedural Standards Policy, produced by the GA4GH's Data Access Review Standards (DACReS) working group. It builds on the mission of the GA4GH and provides greater direction for the interpretation of the GA4GH *Framework for Responsible Sharing of Genomic and Health-Related Data* (the "Framework"). It seeks to complement the ongoing work in the GA4GH's Regulatory and Ethics Work Stream (REWS) and Data Use & Researcher Identities (DURI) Work Stream on facilitating responsible sharing of genomic and associated data through the development of harmonized standards and a data use ontology.

Both inappropriately restrictive and overly permissive policies governing access to genomic and associated data challenge underlying principles of research ethics. The former can prevent persons from accessing data that meaningfully advances research. The latter can place the privacy and confidentiality interests of participants/data subjects at increased risk and erode public trust in research. Data access committees (DACs) represent one institutional safeguard charged with applying rules meant to ensure an ethically permissible balance between data protection and accessibility. There are, however, no procedural standards that apply across DACs. The absence of such standards can invite inconsistencies in DAC reviews, compromising their quality and effectiveness. Conversely, standardizing DAC processes can foster trust and mutual recognition, paving the way to greater coordination, collaboration, and delegation between DACs and other oversight bodies that improves the efficiency of data access without sacrificing protection.

For the purpose of this Policy, a DAC is defined as a body comprised of one or more members who are responsible for overseeing access to genomic and associated data for research subsequent to the primary purpose for which the data were collected, where requested by internal and external applicants to one or more repositories. DACs may oversee multiple distinct datasets or repositories. They may govern access by researchers within a particular institution, as well as external requestors from other institutions, sectors, or countries. DACs may grant direct access to datasets sent digitally in secure form to data users or that are stored in secure computing environments (e.g. a data safe haven). DACS may also grant indirect access to data (i.e. model-to-data where data users submit algorithms to run on secure data sets that remain hidden).

I. Context

1. Purpose of this Policy. Supplementing the *Framework*, the purpose of this Policy is to outline guiding principles and provide procedural standards for DACs. Where applicable, the procedural standards are categorized as either "essential" or "desired." Operationalizing these guiding principles and procedural standards should engender greater trust in the DAC review process across institutions, repositories, and jurisdictions and thereby promote more efficient, more secure, and more consistent procedures for access to data.

2. Resources. The GA4GH has developed tools and resources to support its Policies. They are designed in accordance with the Foundational Principles of the *Framework*. The latest tools and resources developed by the GA4GH are available on its website.

3. Recognizing differences. Differences in the content or application of ethics principles, policies, or norms arise locally among institutions, repositories, and jurisdictions but are not insurmountable barriers to achieving harmonized guiding principles and procedural standards for DACs. This Policy aims to provide baseline or "essential" standards that all DACs should endeavor to apply, as well as additional "desired" standards that DACs may adopt as contextually relevant.

II. Guiding Principles and Procedural Standards

The following Guiding Principles and Procedural Standards are intended to complement and build upon existing laws, regulations, guidelines, and practices. They are not intended to supersede any human rights instruments or national laws and regulations. Across institutions, repositories, and jurisdictions, DACs and managing organizations of DACs should work together to implement these Guiding Principles and Procedural Safeguards. One or more of the Procedural Standards in this Policy may fall under the responsibility of the DAC and/or the host data repository and/or their managing organization, as per the applicable governance arrangements, terms of reference, and management structure.

Guiding Principles

1. Guiding Principles

The Guiding Principles of this Policy address the responsible oversight of access to genomic and associated data for research subsequent to the primary purpose for which the data were collected. They also facilitate compliance with the obligations and norms set by international and national laws and policies to:

i. Make data accessible to advance research and scientific knowledge, as well as to improve health outcomes, through responsible oversight practices;

ii. Promote health, wellbeing, and the fair distribution of expected benefits, as well as protect against risks;

iii. Respect the reasonable expectations of data producers and research participants/data subjects, and the communities to which participants/data subjects belong; and

iv. Maintain procedural fairness for applicants seeking access to data.

Procedural Standards

2. Purpose

i. The primary purpose of a DAC is to oversee access to genomic and associated data for research subsequent to the primary purpose for which the data were collected, where requested by internal and external applicants to one or more repositories, and to assure responsible data use. Such data may identify individual research participants/data subjects. DACs often make decisions on access requests and oversee the management and administration of data access. DAC roles include ensuring that data users who are granted access to the data comply with the terms of a Data Access Agreement (DAA) [1] and to use the data only in approved ways.

ii. Secondary purposes of DACs include verifying an applicant's identity and affiliation; adhering to the terms of participant consent; determining if the requested data uses are relevant and appropriate to the applicant's proposed research project; ensuring responsible downstream data uses; and complying with applicable legal, regulatory, and data protection provisions. [2]

iii. As distinct from a research ethics committee or institutional review board, which assesses the full range of ethical issues associated with a research project involving humans before it begins, the purpose of a DAC is typically to assess if the proposed use of the data is within the bounds of the data's permitted uses and to vet applicant qualifications.

iv. It is possible that DACs also oversee access to biospecimens.

v. DACs are normally established by the host data repository and/or their managing organization to evaluate the merit and feasibility of secondary/ancillary research projects that are proposed to make use of the data.

vi. Each DAC should clearly state its purpose(s) in its own Terms of Reference (see below), as suitable for the context in which the DAC operates.

3. Transparency

i. As outlined in the GA4GH Ethics Review Recognition Policy, which applies to research ethics committees, this Policy also recognizes that transparency ought to serve as a standard that drives committee work in the genomic and health-related data context. For DACs, this includes transparency in both the application process to request data access and in the criteria/access conditions that underpin the DAC's decisions, as well as in the ethics approvals for the data access it oversees.

ii. The DAC's host data repository and/or their managing organization should provide transparent mechanisms for an applicant to apply for data access. This includes providing clear, publicly accessible information about the terms of access (i.e. how applicants who may wish to use the data should seek access), the criteria by which requests for data access are assessed, as well as contact details for applicants who may wish to find out more about the data available for access.

iii. The DAC's host data repository and/or their managing organization should provide a clear, publicly accessible statement about ethics approvals, if applicable, for the data access arrangements in place.

iv. Information about approved access requests, including short summaries in plain language of the research purposes (which may be the responsibility of the data users as part of the application process) and the principal data user(s), should be made publicly available on the DAC or the host data repository's website.

v. If feasible, the permitted purposes for subsequent research projects, including associated limits and conditions should be described for all resources hosted within a repository (data or biospecimens) using a common ontology e.g. GA4GH Data Use Ontology (DUO). This will allow for semantic tagging of datasets with usage restrictions, which can help interpret the conditions in a transparent, consistent, and structured manner.

4. Terms of Reference

A DAC should establish and communicate Terms of Reference in codified form to define its purpose, scope and structure, and its roles and responsibilities. The Terms of Reference ought to provide a documented basis for making decisions and for confirming or developing a common understanding of the scope among the DAC members. Such terms should reflect local circumstances (e.g. different models of access) and key aspects of membership, whether a single member or small/large committee structure.

Items that may be addressed in Terms of Reference include:

i. The role and purpose(s) of the DAC;

ii. Degree of independence of the DAC from the host data repository and/or their managing organization to review data access applications;

iii. The applicable duration of the Terms of Reference;

iv. DAC Membership, including:

   a. How many members comprise the DAC
   b. Expertise necessary to fulfill the purposes of the DAC (where applicable: scientific, ethical, legal, patient/community representatives)
   c. Ad hoc members and other invited experts where needed (including opportunities for community consultation)
   d. Member competency and experience (i.e. familiarity with the database)
   e. Appropriate induction or and/or training of DAC members (including on a periodic basis) in e.g. data privacy, data security, technical and organizational safeguards, access management
   f. Process for appointing new DAC members
   g. Length of service (e.g. contractual, renewable, rotating)
   h. Terms of membership (i.e. professional, voluntary, ability to resign)

v. Roles and responsibilities of members within the DAC;

vi. Declaration of both perceived and real conflict(s) of interest before each access review process;

vii. The nature and structure of DAC meetings and correspondence, including frequency of meetings and technology used for communication;

viii. Processes for amending, modifying, or varying the Terms of Reference.

5. Standard Operating Procedures (SOPs)

A standard operating procedure (SOP) is a set of instructions that guides routine tasks of an organization or group. For DACs, SOPs can assist members to perform their tasks in efficient, consistent, and robust ways. SOPs can also guide internal audits or other performance evaluations. The SOPs of a DAC should reflect local circumstances (e.g. different models of access).

Essential items address in SOPs include:

i. Method of DAC operation and communication (email, virtual meetings, face-to-face);

ii. Meeting quorum, if the DAC adopts a committee structure;

iii. Frequency and approximative meeting length;

iv. Estimated time from application review to decision;

v. Terms of consensus for DAC decisions (e.g. unanimity, majority);

vi. Range of acceptable DAC decisions (e.g. approval, deferral, refusal, incomplete);

vii. Process for submitting and reviewing amendments to an approved project;

viii. Progress reporting and monitoring of projects that have received DAC approval;

ix. Method of correspondence, including the DAC decision to the applicant;

x. Template/generic application forms for access requests;

xi. Template agreements/conditions (e.g. signed by project PIs).

Desired items addressed in SOPs include:

i. Written guidance outlining the responsibilities of PIs and project members (e.g. non-disclosure);

ii. Guidance for writing plain language descriptions of projects submitted for access requests;

iii. Process for submitting an appeal or reapplication if the original data access request is denied;

iv. Method of tracking and auditing decisions (e.g. compiling statistics on DAC applications and decisions);

v. Structure and update of user fees;

vi. Delineation of elements of the data access request that could qualify for expedited DAC review using algorithms, software, and other automated systems;

vii. Where DACs employ computational tools that include, but are not limited to those above, SOPs should clearly address what aspects of the access adjudication process that machine assistance can be applied to supplement human decision making, and make source code as well as software transparent and publicly accessible;

viii. Communication with other review bodies, including research ethics committees or institutional review boards.

6. Criteria for Assessing Access Applications

Criteria should be established for assessing data access applications in a standardized and transparent manner. The criteria should be written in accessible language, publicly available, and reviewed on a semi-regular basis.

Essential assessment criteria include:

i. Evidence of ethics approval or ethics approval exemption/waiver for the project if required by underlying consent or regulatory framework;

ii. Data user identity and affiliation (e.g. via institutional signing official, granting agency) are verified and confirmed to demonstrate that requestors are *bona fide* researchers with sufficient experience to manage the uses of data consistent with ethical practice, relevant laws, and data sharing agreements;

iii. Justification for data access;

iv. Application/project scope aligns with the data use request and access request fully adheres to applicable data use restrictions (i.e. the application addresses topics that fall within the acknowledged remit of the project, as understood by participants);

v. Assurances that practices, methods, or mechanisms are in place to protect participants'/data subjects' rights and interests in the project, including assurances that privacy, confidentiality, and data uses comply with consent permissions for data use or the requirements of the research funding agency and with applicable laws (e.g. ensuring there are appropriate safeguards against the risk that the application will produce information that may identify individual project participants/data subjects, and by assessing what types of linkage with external datasets are permitted);

vi. Assurances that the application/project does not violate or potentially violate participant consent(s) or consent of their guardians where applicable, along with reasonable expectations of their data's use (i.e. ensuring that the data will not be used for something a reasonable participant would find objectionable, or uses that the DAC would have reason to believe participants within the dataset would find objectionable);

vii. Assurances that any of the ethical permissions granted to the project(s) for which data are requested cover the applicant's proposed research uses of the requested data;

viii. Assurances that the application/project for which the data are requested poses minimal risk of re-identifying individuals within the dataset, as well as avoids risk of stigmatization and risk of dignitary harms to both individuals and the communities implicated in the dataset;

ix. Any criteria specific to access to biospecimens if relevant to the DAC (e.g. a review of projects for finite samples given that any use may preclude subsequent future uses).

Desired assessment criteria include:

i. The application includes a summary of the research project in plain language that addresses, among others: (1) what is the research question and why is it important? (2) how will the participants'/data subjects' data be used to investigate the research question? (3) what is the method, in plain language? and (4) what are the potential benefits or implications of the proposed research (including both short-term outcomes and longer-term impact)?

7. Progress Reporting

It is important for the DAC to promote safe and secure access to data while building trust with members of the research community. Monitoring the progress and evaluating the outcomes of DACs are central to building trustworthy systems and institutions for participants/data subjects and research collaborators alike across the data lifecycle. Taken together, progress reporting facilitates and increases opportunities for researchers to share, collaborate, and advance scientific aims.

Essential elements of progress reporting include:

i. The number of contributed projects and/or collection size available to the research community by project and/or data type, demonstrating applicability of data resources to the research community;

ii. The number of access requests, both approved and denied, and the average length of time from user submission to DAC decision, demonstrating transparency in DAC review performance;

iii. The number of users approved for access, their institutional affiliation, and/or their geographical distribution, demonstrating efficacy of the DAC in fulfilling access requests;

iv. Confirmation from the DAC that all conditions of review have been met and the project is complete in terms of data access requirements, demonstrating DAC accountability to data contributors.

Desired elements of progress reporting include:

i. Rationale for denying data access requests, demonstrating DAC fairness in the review or decision-making process for data access;

ii. Annual reports, based on self-reports from data users, about the following metrics, demonstrating overall impact of the DAC and its management of data resources:

- Project progress to date as well as any data-related issues arising and how these have been addressed;
- Whether new data have been contributed, returned, or deposited to the repository relating directly to the data for which access was granted;
- Whether new scientific collaborations have been developed relating to the data for which access was granted;
- Knowledge products developed by data users relating to the data for which access was granted (e.g. publications, associated grants, intellectual property);

iii. An end of project report based on self-reports from data users that details how primary/metadata will be/have been destroyed or anonymised as per the terms of the data access agreement, if required, demonstrating DAC accountability to data contributors.

8. Data Management Incidents

A data management incident is any event that affects the integrity or security of data and/or the privacy of participants/data subjects whose data are in the affected dataset. They include situations such as policy violations, unauthorized data access or sharing, breaches of data security, or inadvertent data releases that may compromise data confidentiality. A DAC or the host data repository (and/or their managing organization) should establish a process or set of procedures:

i. For reporting to the host data repository and/or their managing organization all serious data management incidents as soon as they are discovered by the DAC and/or reported by data users or others to the DAC;

ii. To allow data users to report data management incidents (e.g. data breaches, data loss) to the DAC. Users should be made aware of such procedures upon data release as well as any penalties for non-compliance;

iii. To acknowledge receipt of incident reports from data users in a timely fashion, assess the extent of the incident, minimize further data damage or loss, ascribe responsibility, and apply remedies or penalties as needed.

9. DACs as Living and Learning Organizations

DACs iteratively adapt their operations and procedures as data availability and infrastructures evolve in line with advances in the field. These real-time adaptations reinforce the role of DACs as living and learning organizations that should ideally seek, where possible and appropriate, to:

i. Understand and contribute to the scientific literature on improving the quality and effectiveness of data access processes and governance;

ii. Undertake participant and public outreach that engages individuals and communities more fully in learning and teaching about how their data are used, incorporates and respects their views about acceptable uses and access decisions, and integrates diverse voices into governance approaches and processes;

iii. Act as a guiding body for data producers and data users, collating lessons learned and educating producers and users about safe and responsible uses of data within the dataset's use restrictions.

III. Implementation Mechanisms and Amendments

1. Organizations and bodies supporting this Policy should take all reasonable and appropriate measures, whether of a regulatory, contractual, administrative, or other means, to give effect to this Policy and promote its implementation, monitoring, and enforcement.

2. Procedures and policies should be transparent and accessible. Attention should be paid to the interrelation of this Policy with other GA4GH Policies (e.g. Consent Policy, Data Privacy and Security Policy, Ethics Review Recognition Policy), Frameworks (e.g. GA4GH Framework for Participants, Patients and Publics (PPP) Engagement in Genomics Research and Health Implementation), and the Data Use Ontology (DUO) for semantic tagging of datasets with usage restrictions.

3. Any entity or individual supporting this Policy may propose one or more amendments by communicating with the GA4GH's Regulatory and Ethics Work Stream (REWS). The REWS shall publicly circulate any amendments for comments for possible inclusion in this Policy.

4. The REWS, in collaboration with organizational members and other GA4GH Foundational and Technical Work Streams, will track and report on the adoption of this Policy and its application. The REWS will also routinely review the Policy's provisions; be aware of

advances in basic research and technology; monitor ethical and legal developments, strive to ensure that this Policy is fit for purpose; and encourage the creation of fora or other forms of communication to advance the practice of responsible data access review.

Acknowledgements

This policy was developed by the Data Access Committee Review Standards (DACReS) work group of the GA4GH, and is the result of the collaborative work, comments, and input of many individual and organizational contributors.

Policy Revision History

Policy Number/Version	Date Effective	Summary of Revisions
POL 008 / v. 1.0	27 October 2021	Original document

[1] This may also be known by the terms “Data Use Certificate” or “Data Transfer Agreement”.

[2] Adapted from Expert Advisory Group on Data Access, “Governance of Data Access: Annexes” (June 2015), available at <https://wellcome.org/sites/default/files/governance-of-data-access-annexes-eagda-jun15.pdf> 8.