DATA USE CERTIFICATION AGREEMENT

This Data Use Certification (DUC) Agreement outlines the terms of use for requested controlled-access datasets maintained in NIH-designated data repositories under the NIH Genomic Data Sharing Policy (e.g., the NIH database of Genotypes and Phenotypes (dbGaP), a NIH controlled-access data repository). The Addendum to this Agreement outlines additional terms and information which are specific to each requested dataset such as:

- Data Use Limitation(s)
- Sponsoring NIH Institute or Center
- Responsible Data Access Committee
- Study Description
- Suggested Acknowledgement Statement

INTRODUCTION AND STATEMENT OF POLICY

The National Institutes of Health (NIH) Genomic Data Sharing (GDS) Policy expects investigators generating large-scale human genomic data as well as relevant associated data to submit these data to a NIH-designated data repository. Respect for, and protection of the interests of, research participants are a key tenet of the GDS Policy and fundamental to NIH's stewardship of human genomic data. As such, access to controlled-access human genomic data will be provided only to research investigators who, along with their institutions, agree to meet the expectations and terms of access detailed below and to use the data according to participant informed consent, actualized as applicable Data Use Limitations established by the Submitting Institution through the Institutional Certification.

Definitions of the underlined terminology in this document are found in section 14.

The parties to this Agreement include: the Principal Investigator (PI) requesting access to controlled-access genomic and associated data (an "Approved User"), the PI's home institution (the "Requester") as represented by the Institutional Signing Official, and the NIH. The effective date of this Agreement shall be the data access request (DAR) Approval Date, as specified in the notification of Data Access Committee (DAC) approval.

TERMS OF ACCESS

1. Research Use

The Requester agrees that if access is approved, (1) the PI named in the DAR and (2) those named in the "Senior/Key Person Profile" section of the DAR, including the Information Technology Director and any trainee, employee, or contractor[1] working on the proposed research project under the direct

[1] If contractor services are to be utilized, the PI requesting the data must provide a brief description of the services that the contractor will perform for the PI (e.g., data cleaning services) in the research use statement of the DAR. The PI is expected to include in any contract agreement requirements that any of the contractor's employees who have access to the data adhere to the NIH GDS Policy, this Data Use Certification Agreement, and the NIH Security Best Practices for Users of Controlled-Access Data. Note that any scientific collaborators, including contractors, who are not at the Requester must submit their own DAR. These requirements apply whether the contractor carries out the work at the PI's facility or at the contractor's facility.

oversight of these individuals, shall become Approved Users of the requested dataset(s). The Requester and Approved Users acknowledge responsibility for ensuring the review and agreement to the terms within this Agreement and the appropriate research use of controlled-access data obtained through the attached DAR and any Data Derivatives of controlled-access datasets by research staff associated with any approved project, subject to applicable laws and regulations. Research use will occur solely in connection with the approved research project described in the DAR, which includes a 1-2 paragraph description of the proposed research (i.e., a Research Use Statement). The Requester further certifies that the Research Use Statement's description of the proposed research is truthful and accurate.

If the DAR process expects a Cloud Use Statement for investigators interested in using Cloud Computing, investigators must provide a Cloud Use Statement about the Cloud Service Provider (CSP) and/or third-party IT system and agree to secure the data according to the NIH Security Best Practices for Users of Controlled-Access Data. The Cloud Use Statement should at least state the name of the CSP and/or third-party IT system, the security standard, and how the CSP and/or third-party IT system will be used to carry out the work described in the Research Use Statement. If applicable, the investigator should describe the role of any Collaborators in using the CSP and/or third-party IT system. If the Approved User(s) plans to collaborate with investigators outside the Requester, the investigators at each external site must submit an independent DAR using the same project title and Research Use Statement, and if the DAR process expects when using the cloud, a Cloud Use Statement. New uses of these data outside those described in the DAR will require submission of a new DAR; modifications to the research project will require submission of an amendment to this application (e.g., adding or deleting Requester, Collaborators from the Requester, adding datasets to an approved project). Access to the requested dataset(s) is granted for a period of **one (1) year**, with the option to renew access or close-out a project at the end of that year.

Submitting Investigator(s), or their Collaborators, who provided the data or samples used to generate controlled-access datasets subject to the NIH GDS Policy and who have Institutional Review Board (IRB) approval, as applicable, and who meet any other study specific terms of access, are exempt from the limitation on the scope of the research use as defined in the DAR.

2. Requester and Approved User Responsibilities

The Requester agrees, through the submission of the DAR, that the Approved Users have reviewed and understand the principles for responsible use and data management of controlled-access data as defined in the GDS Policy and the NIH Security Best Practices for Users of Controlled-Access Data. The Requester and Approved Users acknowledge that the NIH (including NIH DACs) may reject DARs, request revisions to DARs, and terminate ongoing research described in the Research Use Statement if NIH assesses the project has significant potential to cause harm to research participants, their families, groups and populations of which they are a part, or the national security of the United States, or for any reason at NIH's discretion. The Requester and Approved Users further acknowledge that they are responsible for ensuring that all uses of the data are consistent with national, Tribal, and state laws and regulations, as appropriate, as well as relevant institutional policies and procedures for managing controlled-access data. The Requester and Approved Users agree that in using the data, they are not aware of significant potential for the research to cause harm to participants, their families, groups and populations of which they are a part (e.g., from stigma associated with the research results), or the national security of the United States. The Requester and Approved Users agree that in using the data, if they become aware of significant potential for the research to cause harm to participants, their

families, groups and populations of which they are a part, or the national security of the United States that they will notify NIH within 24 hours. The Requester certifies that the PI is in good standing (i.e., no known sanctions) with the institution, relevant funding agencies, and regulatory agencies and is eligible to conduct independent research (i.e., is not a postdoctoral fellow, student, or trainee). The Requester and any Approved Users may use the dataset(s) only in accordance with the parameters described on the study page and in the Addendum to this Agreement for the appropriate research use, as well as any limitations on such use of the dataset(s) as described in the DAR, and as required by law.

Through the submission of this DAR, the Requester and Approved Users acknowledge receiving and reviewing a copy of the Addendum which includes Data Use Limitation(s) for requested controlled-access data. The Requester and Approved Users agree to comply with the terms listed in the Addendum.

Through submission of the DAR, the PI and Requester agree to submit a Project Renewal or Project Close-out prior to the expiration date of the one (1) year data access period. The PI also agrees to submit an annual Progress Update prior to the one (1) year anniversary of the project, as described under *Research Use Reporting* (Term 11) below.

By approving and submitting the attached DAR, the Institutional Signing Official provides assurance that relevant institutional policies and applicable local, state, Tribal, and federal laws and regulations, as applicable, have been followed, including IRB approval, if required. Approved Users may be required to have IRB approval if they have access to personal identifying information for research participants in the original study at their institution, or through their Collaborators. The Institutional Signing Official also assures, through the approval of the DAR, that other institutional departments with relevant authorities (e.g., those overseeing human subjects research, information technology, technology transfer) have reviewed the relevant sections of the NIH GDS Policy and the associated procedures and are in agreement with the principles defined.

The Requester acknowledges that controlled-access datasets subject to the NIH GDS Policy may be updated to exclude or include additional information. Unless otherwise indicated, all statements herein are applicable to the access and use of all versions of these datasets.

3. Public Posting of Approved Users' Research Use Statement

The PI agrees that information about themselves and the approved research use will be posted publicly on the dbGaP website. The information includes the PI's name and Requester, project name, Research Use Statement, and a Non-Technical Summary of the Research Use Statement. In addition, and if applicable, this information may include the Cloud Use Statement and name of the CSP and/or third-party IT system. Citations of publications resulting from the use of controlled-access data obtained through this DAR may also be posted on the dbGaP website.

4. Non-Identification

Approved Users agree to make no attempt to identify or contact, either directly or indirectly, individual participants or their families. Approved Users agree not to use the requested datasets, either alone or in concert with any other information, to identify or contact individual participants from whom data and/or samples were collected. Approved Users also agree not to generate information (e.g., facial images or comparable representations) that could allow the identities of research participants to be

readily ascertained. These provisions do not apply to research investigators operating with specific IRB approval, pursuant to 45 CFR 46, to contact individuals within datasets or to obtain and use identifying information under an IRB-approved research protocol. All investigators including any Approved User conducting "human subjects research" within the scope of 45 CFR 46 must comply with the requirements contained therein.

5. Certificate of Confidentiality

Certificates of Confidentiality (Certificate) protect the privacy of research participants by prohibiting disclosure of protected information for non-research purposes to anyone not connected with the research except in specific situations. Data that are stored in and shared through the NIH data repositories are protected by a Certificate. Therefore, Approved User(s), whether or not funded by the NIH, who are approved to access a copy of information protected by a Certificate, are also subject to the requirements of the Certificate of Confidentiality and subsection 301(d) of the Public Health Service Act (codified at 42 U.S.C. 241(d))

Under Section 301(d) of the Public Health Service Act and the *NIH Policy for Issuing Certificates of Confidentiality*, recipients of a Certificate of Confidentiality shall not:

- Disclose or provide, in any Federal, State, or local civil, criminal, administrative, legislative, or other proceeding, the name of such individual or any such information, document, or biospecimen that contains identifiable, sensitive information about the individual and that was created or compiled for purposes of the research, unless such disclosure or use is made with the consent of the individual whom the information, document, or biospecimen pertains; or
- Disclose or provide to any other person not connected with the research the name of such an individual or any information, document, or biospecimen that contains identifiable, sensitive information about such an individual and that was created or compiled for purposes of the research.

Disclosure is permitted only when:

- Required by Federal, State, or local laws (e.g., as required by the Federal Food, Drug, and Cosmetic Act, or state laws requiring the reporting of communicable diseases to State and local health departments), excluding instances of disclosure in any Federal, State, or local civil, criminal, administrative, legislative, or other proceeding;
- Necessary for the medical treatment of the individual to whom the information, document, or biospecimen pertains and made with the consent of such individual;
- Made with the consent of the individual to whom the information, document, or biospecimen pertains; or
- Made for the purposes of other scientific research that is in compliance with applicable Federal regulations governing the protection of human subjects in research.

For more information see: Certificates of Confidentiality (CoC) | Grants & Funding

6. Non-Transferability

The Requester and Approved Users agree not to distribute controlled-access data and any Data Derivates to any entity or individual not identified in the approved request without appropriate written

approvals from the NIH. If the Approved Users are provided access to controlled-access datasets subject to the NIH GDS Policy for inter-institutional collaborative research described in the Research Use Statement of the DAR, and all members of the collaboration are also Approved Users through their home institution(s), data obtained through the attached DAR may be securely transmitted within the collaborative group. Each Approved User will secure the data according to the NIH Security Best Practices for Users of Controlled-Access Data, the terms of this Agreement, and the Requester's IT security requirements and policies.

Requester and Approved Users agree that controlled-access datasets obtained through the attached DAR and any Data Derivatives of controlled-access datasets, in whole or in part, may not be sold to any individual at any point in time for any purpose.

The PI agrees that if they change institutions during the approved access period, they will complete the Project Close-out process (See Term 13 for more details) before moving to their new institution. A new DAR, in which the new Requester agrees to the Data Use Certification Agreement and the Genomic Data User Code of Conduct, must be approved by the relevant NIH DAC(s) before controlled-access data may be re-accessed.

7. Data Security and Unauthorized Data Release

The Requester and Approved Users acknowledge NIH's expectation that they have reviewed and agree to manage the requested controlled-access data and any Data Derivatives according to NIH's expectations set forth in the current NIH Security Best Practices for Users of Controlled-Access Data and the Requester's IT security requirements and policies.

The Requester and PI agree to notify the NIH Incident Response Team, NIH DAC(s) on the project request, and NIH Office of Extramural Research Data Sharing Policy Implementation (OER/DSPI) Team of any unauthorized data sharing, breaches of data security, or inadvertent data releases that may compromise data confidentiality within 24 hours of when the incident is identified. For the NIH Incident Response Team notifications can be made by phone (301) 496-HELP (4357); Toll Free Number: (866) 319-4357or TTY: (301) 496-8294 and can also be sent by email to NIHInfoSec@nih.gov or via the Report an Incident Link: https://irtportal.ocio.nih.gov/. For OER/DSPI Team, notifications can be sent to DMI_OER@mail.nih.gov.

As permitted by law, notifications should include any known information regarding the incident and a general description of the activities or process in place to define and remediate the situation fully. Within 3 business days of the DAC notification, the Requester agrees to submit to the DAC(s) and the OER/DSPI Team a detailed written report including the date and nature of the event, actions taken or to be taken to remediate the issue(s), and plans or processes developed to prevent further problems, including specific information on timelines anticipated for action. The Requester agrees to provide any additional documentation requested by the NIH DAC(s) or the OER/DSPI Team on the incident, including verifying that the remediation plans have been implemented. Repeated violations or unresponsiveness to NIH requests may result in further compliance measures affecting the Requester.

NIH, or another entity designated by NIH may, as permitted by law, also investigate any data security incident. Approved Users and their associates agree to support such investigations and provide any information, within the limits of applicable local, state, Tribal, and federal laws and regulations. In

addition, Requester and Approved Users agree to work with the NIH to assure that plans and procedures that are developed to address identified problems are mutually acceptable and consistent with applicable law.

8. Terms of Access Violations

The Requester and Approved Users acknowledge that the NIH may terminate the DAR, including this Agreement and immediately revoke or suspend access to all controlled-access datasets subject to the NIH GDS Policy at any time if the Requester is found to be no longer in agreement with the principles outlined in the NIH GDS Policy, the terms described in this Agreement, the Genomic Data User Code of Conduct or the policies, principles and procedures of NIH. The Requester and Approved Users agree to notify the OER/DSPI Team, and the NIH DAC(s) indicated in the project request to this Agreement of any violations of the NIH GDS Policy, this Agreement, or the Genomic Data User Code of Conduct, hereinafter referred to as data management incidents (DMIs), within 24 hours of when the incident is identified. For OER/DSPI Team, notifications can be sent to DMI_OER@mail.nih.gov. Repeated violations or unresponsiveness to NIH requests may result in further compliance measures affecting the Requester.

As permitted by law, notifications should include any known information regarding the incident and a general description of the activities, corrective actions, or process in place to define and remediate the situation fully. Within 3 business days of the notification(s), the Requester agrees to submit to the NIH DAC(s)and the OER/DSPI Team a detailed written report including the date and nature of the event, actions taken or to be taken to remediate the issue(s), and plans, preventive actions or processes developed to prevent future incidents, including specific information on timelines anticipated for action. The Requester agrees to provide documentation verifying that the remediation plans have been implemented. The Requestor agrees to incorporate any changes to corrective or preventive actions or to make any additional corrective and preventive actions requested by NIH. Repeated violations or unresponsiveness to NIH requests may result in further compliance measures affecting the Requester.

NIH, or another entity designated by NIH may, as permitted by law, also investigate any DMI. Approved Users and their associates agree to support such investigations and provide information, within the limits of applicable local, state, Tribal, and federal laws, and regulations. In addition, Requester and Approved Users agree to work with the NIH to assure that plans and procedures that are developed to address identified problems are mutually acceptable and consistent with applicable law.

9. Intellectual Property

By requesting access to dataset(s), the Requester and Approved Users acknowledge the intent of the NIH that anyone authorized for research access through the DAR follow the intellectual property (IP) principles as summarized below:

- Achieving maximum public benefit is the ultimate goal of data distribution through the NIH controlled-access data repositories. The NIH encourages broad use of NIH controlled-access data that is consistent with a responsible approach to management of intellectual property derived from downstream discoveries and expects that the Requestor and Approved User(s) adhere to licensing practices consistent with the NIH Research Tools Policy.

The NIH considers these data as pre-competitive and urges Approved Users to avoid making IP claims derived directly from the dataset(s). It is expected that these NIH-provided data, and conclusions

derived therefrom, will remain freely available, without requirement for licensing. However, the NIH also recognizes the importance of intellectual property in promoting the development of new therapies and products; as such, there is no restriction on development of commercial products resulting from the knowledge gained from the research project. Ownership of all intellectual property generated by activities under the research project will be governed by applicable patent law.

10. Dissemination of Research Findings and Acknowledgement of Controlled-Access Data Subject to the NIH GDS Policy

It is NIH's intent to promote the dissemination of research findings from use of controlled-access data subject to the NIH GDS Policy as widely as possible through scientific publication or other appropriate public dissemination mechanisms. Approved Users are strongly encouraged to publish their results in peer-reviewed journals and to present research findings at scientific meetings.

Approved Users agree to acknowledge the Submitting Investigator(s) who submitted data from the original study to an NIH-designated data repository, the primary funding organization that supported the Submitting Investigator(s), and the NIH-designated data repository, in all oral and written presentations, disclosures, and publications resulting from any analyses of controlled-access data obtained through the attached DAR. Approved Users further agree that the acknowledgment shall include the dbGaP accession number to the specific version of the dataset(s) analyzed. A sample acknowledgment statement is provided for each dataset in the Addendum to this Agreement.

11. Research Use Reporting

To assure adherence to NIH GDS Policy, the PI agrees to provide annual Progress Updates as part of the annual Project Renewal or Project Close-out processes, prior to the expiration of the one (1) year data access period. The PI who is seeking renewal or close-out of a project agree to complete the appropriate online forms and provide specific information such as how the data have been used, including publications or presentations that resulted from the use of the requested dataset(s), a summary of any plans for future research use (if the PI is seeking renewal), any violations of the terms of access described within this Agreement and the implemented remediation, and information on any downstream intellectual property generated from the data. The PI also may include general comments regarding suggestions for improving the data access process in general. Information provided in the Progress Updates helps NIH evaluate program activities and may be considered by the NIH GDS governance committees as part of NIH's effort to provide ongoing stewardship of data sharing activities subject to the NIH GDS Policy.

12. Non-Endorsement, Indemnification

The Requester and Approved Users acknowledge that although all reasonable efforts have been taken to ensure the accuracy and reliability of controlled-access data obtained through the attached DAR, the NIH and Submitting Investigator(s) do not and cannot warrant the results that may be obtained by using any data included therein. NIH and all contributors to these datasets disclaim all warranties as to performance or fitness of the data for any particular purpose.

No indemnification for any loss, claim, damage, or liability is intended or provided by any party under this agreement. Each party shall be liable for any loss, claim, damage, or liability that said party incurs as a result of its activities under this agreement, except that NIH, as an agency of the United States, may be

liable only to the extent provided under the Federal Tort Claims Act, 28 USC 2671 et seq.

13. Termination and Data Destruction

A Project Close-out must be completed when an approved project is completed. Upon Project Close-out, the Requester and Approved Users agree to destroy all copies, versions, and Data Derivatives of the data retrieved from NIH-designated data repositories, on both local servers and hardware, and if Cloud Computing was used, delete the data and cloud images from Cloud Computing provider storage, virtual and physical machines, and databases in accord with the NIH Security Best Practices for Users of Controlled-Access Data . However, the Requester may retain only encrypted copies of the minimum data necessary at their institution to comply with institutional scientific data retention policy, law, and scientific transparency expectations for disseminated research results, and/or journal policies. A Requester who retains data for any of these purposes continues to be a steward of the data and is responsible for the management of the retained data in accordance with the NIH Security Best Practices for Users of Controlled-Access Data , and any institutional policies. Any retained data may only be used by the PI and Requester to support the findings (e.g., validation) resulting from the research described in the DAR that was submitted by the Requester and approved by NIH. The data may not be used to answer any additional research questions, even if they are within the scope of the approved DAR, unless the Requester submits a new DAR and is approved by NIH to conduct the additional research. If a Requester retains data for any of these purposes, the relevant portions of Terms 4, 5, 6, 7, 8, and 13 remain in effect after termination of this Data Use Certification Agreement. These terms remain in effect until the data is destroyed. In instances where NIH provides written notification that Data Derivatives should be transferred to a NIH controlled-access data repository, the transfer must be completed prior to Project Close-out.

NIH may terminate this agreement at any time for any reason at its discretion with written notice to the Requestor.

14. Definitions

Approved User: A user approved by the relevant Data Access Committee(s) to access one or more datasets for a specified period of time and only for the purposes outlined in the Principal Investigator (PI)'s approved Research Use Statement. The Information Technology (IT) Director indicated on the Data Access Request, as well as any staff members and trainees under the direct supervision of the PI are also Approved Users and must abide by the terms laid out in the Data Use Certification Agreement.

Collaborator: An individual who is not under the direct supervision of the PI (e.g., not a member of the PI's laboratory) who assists with the PI's research project involving controlled-access data subject to the NIH GDS Policy. Internal Collaborators are employees of the Requester and work at the same location/campus as the PI. External Collaborators are not employees of the Requester and/or do not work at the same location as the PI, and consequently must be independently approved to access controlled-access data subject to the NIH GDS Policy.

Cloud Computing: The National Institute for Standards and Technology defines cloud computing as a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. For more information see NIST Special Publication 800-145.

Cloud Service Provider (CSP): A company or institution that offers some component of cloud computing to other businesses or individual, typically Infrastructure as a Service (IaaS), Software as a Service (SaaS) or Platform as a Service (PaaS), as defined by the National Institute of Standards and Technology. For more information see NIST Special Publication 800-145.

Data Access Request (DAR): A request submitted by a Principal Investigator to a NIH Data Access Committee for access to controlled-access data from a NIH-designated data repository. The DAR is signed by the PI requesting the data and their Institutional Signing Official.

Data Derivative: Data derived from controlled-access datasets obtained from NIH-designated data repositories. Examples of derived data include imputed datasets and single nucleotide polymorphisms, or any data explicitly designated as Data Derivatives by NIH.

Data Use Certification (DUC) Agreement: An agreement between the Approved User, the Requester, and NIH regarding the terms associated with access of controlled-access datasets subject to the NIH GDS Policy and the expectations for use of these datasets.

Genomic Data User Code of Conduct: Key principles and practices agreed to by all research investigators requesting access to controlled-access data subject to the NIH GDS Policy. The elements within the Genomic Data User Code of Conduct reflect the terms of access in the Data Use Certification Agreement.

Information Technology (IT) Director: An Approved User who is generally a senior IT official of the Requester with the necessary expertise and authority to affirm the IT capacities at the Requester. The IT Director is expected to have the authority and capacity to ensure that the NIH Security Best Practices for Users of Controlled-Access Data and the Requester's IT security requirements and policies are followed by all of the Requester's Approved Users.

Institutional Certification: Certification by the Submitting Institution that delineates, among other items, the appropriate research uses of the data and the uses that are specifically excluded by the relevant informed consent documents. Further information may be found here.

Institutional Signing Official: The label, "Signing Official," is used in conjunction with the NIH eRA Commons and refers to the individual that has institutional authority to legally bind the institution in grants administration matters. The individual fulfilling this role may have any number of titles in the institution but is typically located in its Office of Sponsored Research or equivalent.

Principal Investigator (PI): An investigator who is a permanent employee of their institution at a level equivalent to a tenure-track professor or senior scientist with responsibilities that most likely include laboratory administration and oversight. Additionally, the investigator has the authority to ensure those that they directly supervise adhere to the terms of access in this agreement.

Progress Update: Information included with the annual Data Access Request (DAR) renewal or Close-out summarizing the analysis of controlled-access datasets obtained through the DAR and any publications and presentations derived from the work.

Project Close-out: Termination of a research project that used controlled-access data from an NIH-

designated data repository (e.g., dbGaP) and confirmation of data destruction when the research is completed and/or discontinued. The project close-out process is completed in the dbGaP Authorized Access System.

Project Renewal: Renewal of a PI's access to controlled-access datasets for a previously approved project.

Requester: The home institution or organization of the Approved User that applies to dbGaP for access to controlled-access data subject to the NIH GDS Policy.

Submitting Institution: An organization who submitted a genomic dataset to an NIH-designated data repository (e.g., dbGaP).

Submitting Investigator: An investigator who submitted a genomic dataset to an NIH designated data repository (e.g., dbGaP).

Third-party IT system: A collection of computing and/or communications components and other resources that support one or more functional objectives of an organization.

Study specific DUC addendum

phs004480 : Combined Genomic and Phenotypic Classification of Inherited and Acquired Genetic Disease with Long-Read Sequencing

Public Posting of Genomic Summary Results - Allowed.

NIH Data Access Committee (DAC) : National Human Genome Research Institute

Important Contacts : nhgridac@mail.nih.gov; GDS@mail.nih.gov

In the event of a data management incident, within 24 hours, please contact emails above.

Acknowledgement Statement : This research was supported by the Snow Medical Research Foundation (Grant No. PF2019-040). The funders were not involved in the design of the study, collection, analysis, and interpretation of the data, the writing of this report, or the decision to submit the article for publication. The authors express their sincere appreciation to the Snow Medical Research Foundation for their generous support of this work. This dataset is described in greater detail in https://doi.org/10.1101/2025.05.08.25327265.

Name : Health/Medical/Biomedical

Consent Group # : 1

Abbreviation : HMB

Data Use Limitation : Use of this data is limited to health/medical/biomedical purposes, does not include the study of population origins or ancestry.