

# MEDICAL GENOME REFERENCE BANK DATA ACCESS APPLICATION FORM

Medical Genome Reference Bank Data Access Application

Prepared by the MGRB Data Access Committee

Version 1.5 – 1 November 2021



Health



Garvan Institute  
of Medical Research



SaxInstitute



45 and Up  
Study

## SECTION 1 – BACKGROUND

### 1.1 The Medical Genome Reference Bank

The Medical Genome Reference Bank (MGRB) is comprised of whole genome sequencing (WGS) data and phenotypic information from 4010 healthy Australians over 70 years of age. MGRB participants, consented through contributing studies, 45 and Up (Sax Institute, Sydney), and the ASPIrin in Reducing Events in the Elderly (ASPREE) clinical trial (Monash University, Melbourne), are free from cardiovascular disease, degenerative neurological disorders and of a history of cancer at the time of consent into the study. The dataset can function as a powerful filter to distinguish between causal genetic and population-based genetic variation and be a resource to maximise the efficiency of genomic discovery in both the research and clinical setting.

WGS was performed on the Illumina HiSeq X-Ten platform at the Garvan Institute of Medical Research (Sydney, Australia) under clinically accredited conditions (ISO 15189). Data was aligned and variant files were generated using best practice (BWA, GATK) pipelines to make results comparable with other cohorts. Joint-called variants were loaded into the [Sydney Genomics Collaborative](#) data portal built on the Garvan Institute's Vectis platform. Key features of Vectis (a lever in Latin) include the support of both summary statistics (Summary and Explore tabs) as well as a scale-out Variant Store that supports the filtering of patients based on genomic features (genes and chromosome coordinates) as well as clinical information. Access to the Clinical Filtering features of Vectis are restricted to users whose Data Access Applications have been approved. Researchers are invited to complete this **Data Access Application** to gain access to comprehensive genotypic and clinical information, to support high-level integrative analysis.

Information on the comprehensive clinical data that is available through ASPREE and 45 and UP (pending a successful application) can be found in the following documents:

#### ASPREE

[ASPREE Protocol AUS Version# 9 Nov 2014](#)

#### 45 and Up

Download the [Baseline Questionnaire for Women](#) (PDF 537KB)

Download the [Baseline Questionnaire for Men](#) (PDF 537KB)

## 1.2 Data Access Policy

To maintain participant privacy and confidentiality, whilst maximising MGRB utility, we have deployed a tiered data management system that determines the depth of data that is made available to researchers (as summarised in the schematic below). This consists of 3 access tiers; Open access, Controlled access and Restricted access. Completion of the MGRB DAA (Section 2), and execution of the Data Transfer Agreement (Sections 3, 4 & 5) is required before access to Controlled and Restricted access data will be granted. The MGRB Data Access Committee (DAC) will review completed applications within 6 weeks. Upon approval, Controlled access can then be immediately made available to the applicant. Should Restricted access data be requested, the application will be immediately forwarded to the governing body of the participating study, and will be considered in parallel with the MGRB DAC review. The study governing body may insist on an independent application process, the time-line for which dependent solely on the governing body – more information can be found here for [ASPREE](#) and [45 and UP](#). A copy of the full MGRB Data Access Policy can be found [here](#).

## 1.3 Data Access Policy Schematic

Tier	Open Access	Controlled Access	Restricted Access
Access	Institutional email address required for MGRB data-portal access ( <b>not</b> required for Beacon) <a href="http://www.sgc.garvan.org.au/mgrb">www.sgc.garvan.org.au/mgrb</a>	Data Access Application (DAA) must be approved by the MGRB Data Access Committee (DAC)	DAA must be approved by the MGRB DAC <b>and</b> referred to the applicable cohort governing body for <b>further</b> approval
Clinical Data	Basic demographic data are provided - genomic queries can be filtered according to these fields	Basic demographic data and minimal clinical information (where available) are provided per individual record	Comprehensive clinical data that is potentially specific to a participating cohort is provided per individual record
Genomic Data	Beacon and preprocessed population-scale variant analysis	Individual record data provided – either processed (VCF/ gVCF format), or unprocessed (FASTQ or BAM format) (dependent on justification criteria being met)	

## 1.4 Data Access

All approved users will be granted immediate access to the MGRB on the [Sydney Genomics Collaborative](#) Vectis Platform. Applicants wishing to access data files can access these from the [European Genome-phenome Archive \(EGA\)](#) as well as [NCI-Australia](#). The EGA resource is intended for approved MGRB users who wish to download the data as lossless CRAM files (approximately 150 Terabytes). Access to the MGRB at NCI-Australia is intended for users of this Australian Tier 1 supercomputing facility but does not support the downloading of data.

### Please indicate where you would like to access the MGRB data

**EGA**

<https://ega-archive.org/studies/EGAS00001003511>

Associated phenotype data are also available:

```
pyega3 -c 1000 -cf ~/.ega2 -d fetch EGAF00002440978
```

**NCI-Australia**

Once downloaded, minimum protection measures are required to protect all MGRB data. Data can be held in unencrypted files on an institutional computer system, with the equivalent of Unix user group read/write access for one or more appropriate groups but not Unix world read/write access behind a secure firewall. Laptops holding this data should have password protected logins and screenlocks (set to lock after 5 min of inactivity). If held on USB keys or other portable hard drives, the data must be encrypted.

## SECTION 2 - DATA ACCESS APPLICATION

### 2.1 Project Overview

**Level of data required**

**Controlled**

**Restricted**

**Provide a brief overview of the proposed project; with specific emphasis on what level of information is being requested and how MGRB data will be used (500 words maximum)**

## 2.2 Genomic Data

Select the MGRB data required for the proposed project (tick both if required)

45 and Up

ASPREE

Provide detail as to how MGRB data will be stored and what security measures are in place to ensure conformity with the conditions stipulated in Section 1.4)

### 2.3 Clinical Information

**Clinical data (Controlled Access) for 45 and UP and ASPREE are listed below and contained within the Phenotype file on the European Genome-phenome Archive**

(Should additional clinical information be required from either study, please state what data are required and justify the release of this data – release of this information is at the discretion of the external governing board for the study in question and maybe subject to an independent proposal)

#### **45 and UP**

**Sex**

**Year of Birth**

**Height**

**Weight**

Additional clinical information required from 45 and Up and justification thereof;

#### **ASPREE**

**Sex**

**Year of Birth**

**Height**

**Weight**

**Systolic Blood Pressure**

**Evidence of Macular Degeneration**

**Abdominal Circumference**

**Resting Glucose**

Additional clinical information required from ASPREE and justification thereof;	
---	--



## SECTION 3 – DATA TRANSFER AGREEMENT TERMS AND CONDITIONS

The MGRB Research Governance Committee (RGC) has been established by the three organisations that generated and collected the clinical and genomic data that comprises the MGRB: the Garvan Institute of Medical Research, Monash University and the Sax Institute. The RGC has delegated responsibility for assessing applications to access data to the MGRB Data Access Committee (DAC). Garvan has been delegated by the DAC to sign the Data Transfer Agreement on behalf of the MGRB Partners. Garvan cannot amend the terms and conditions of this agreement.

The **User Institution** (named in Section 5.1) agrees with the Garvan Institute of Medical Research of 384 Victoria St, Darlinghurst, New South Wales, 2010 Australia ABN 62 330 391 937 (**Garvan**), that access to and use of the MGRB dataset specified in Sections 2.2 and 2.3 will be governed by the following terms and conditions.

**The User Institution agrees to be bound by these terms and conditions.**

### 3.1 Definitions

**Authorised Personnel** means the individuals at the User Institution to whom the MGRB Data Access Committee grants access to the Data including the User listed in Section 5.2, the individuals listed in Section 5.3 and any other individuals for whom the User Institution subsequently requests access to the Data.

**Data** means the managed access datasets to which the User Institution has requested access as set out in Sections 2.2 and 2.3.

**External Collaborator** means a collaborator of the User, working for an institution other than the User Institution.

**MGRB Partners** means the Garvan Institute of Medical Research, the Sax Institute and Monash University.

**MGRB Data** means the open and controlled access genomic and clinical datasets generated by the MGRB Partners.

**Project** means the project for which the User Institution has requested access to these Data as set out in Section 2.1.

**Publication** means without limitation, articles published in print journals, electronic journals, reviews, books, posters and other written and oral presentations of research.

**Recipient** means the User Institution listed in Section 5.1.

**Research Participant** means an individual whose data form part of the Data.

**Research Purpose** means an academic research not for commercial application that is seeking to advance the understanding of genetics and genomics, including the treatment of disorders, and work on statistical methods that may be applied to such research.

**User** means the Principal Investigator for the Project listed in Section 5.2.

- 3.2 The User Institution agrees to only use the Data for the purpose of the Project and only for Research Purposes. The User Institution provides a warranty that the Project for which MGRB data are being sought, has been approved by a human research ethics committee.
- 3.3 The User Institution agrees to preserve, at all times, the confidentiality of the Data. In particular, it undertakes not to use, or attempt to use the Data to compromise or otherwise infringe the confidentiality of information on Research Participants. Without prejudice to the generality of the foregoing, the User Institution agrees to use at least the measures set out in Section 1.4 to protect the Data.
- 3.4 The User Institution agrees to protect the confidentiality of Research Participants in any research paper or publication that is prepared using the Data by taking all reasonable care to limit the possibility of identification.
- 3.5 The User Institution agrees not to link or combine the Data to other information or archived data available in a way that could re-identify the Research Participants, even if access to that data has been formally granted to the User Institution or is freely available without restriction.
- 3.6 The User Institution agrees only to transfer or disclose the Data, in whole or part, or any material derived from the Data, to the User and Authorised Personnel (as defined in Section 5). Should the User Institution wish to share the Data with an External Collaborator, the External Collaborator must complete a separate application for access to the Data.
- 3.7 The User Institution agrees that the MGRB Partners, and all other parties involved in the creation, funding or protection of the Data: a) make no warranty or representation, express or implied as to the accuracy, quality or comprehensiveness of the Data; b) bear no responsibility for the unavailability of, or break in access to, the Data for whatever reason; and c) bear no responsibility for the further analysis or interpretation of these Data. Except to the extent prohibited by law, the User institution assumes all risk and liability for all damages, losses, expenses (including reasonable legal expenses), claims, demands, suits or other liability (Loss) arising from or in relation to the User's use, storage or disposal of the Data. The MGRB Partners will not be liable to the User Institution for any Loss incurred by the User or User Institution, or made against the User or User Institution by any other party, due to or arising from the use, storage or disposal of the Data by the User or User Institution, except to the extent permitted by law when caused by the gross negligence or willful misconduct of the MGRB Partners.

- 3.8 The User Institution agrees to follow the [Fort Lauderdale Guidelines](#) and the [Toronto Statement](#). This includes but is not limited to recognising the contribution of the MGRB Partners and including a proper acknowledgement in all reports or publications resulting from the use of these Data.
- 3.9 The User Institution agrees to follow the *Publication Policy* described in Section 4.
- 3.10 None of the MGRB Partners will have any right, title or interest in any intellectual property (IP) that is developed, created or invented in the course of the Project unless one or more of the MGRB Partners has contributed to its development creation or invention. If that is the case the IP will be jointly owned by the contributing parties as tenants in common in shares proportionate to their respective intellectual contribution to the development creation or invention of that IP. The joint owners will consult and decide on measures to be taken to protect and exploit that IP.
- 3.11 The User Institution agrees not to make intellectual property claims on the Data and not to use intellectual property protection in ways that would prevent or block access to, or use of, any element of the Data, or conclusions drawn directly from the Data or any of it.
- 3.12 The User Institution must notify the MGRB Data Access Committee of any patent application filed in respect of the Project, at the time of filing and provide the MGRB Data Access Committee (or its nominee) with sufficient information to determine whether there has been a breach of clause 3.11. If there has been a breach of clause 3.11 then the User Institution must amend its patent application to eliminate the breach or withdraw it. The information provided to MGRB Data Access Committee pursuant to this clause will be treated as confidential to the User Institution.
- 3.13 The User Institution agrees to destroy/discard the Data held, once it is no longer required for use in the Project, unless obliged to retain the Data for archival purposes in conformity with audit or legal requirements. The User will not be required to delete or destroy Data that resides on electronic back-up tapes or other electronic back-up files that have been created solely by the User Institution's automatic or routine archiving and back-up procedures but agrees to preserve the confidentiality of such Data.
- 3.14 The User Institution will notify the MGRB Data Access Committee within 30 days of any changes or departures of Authorised Personnel.
- 3.15 The User Institution will notify the MGRB Data Access Committee as soon as it becomes aware of a breach of the terms or conditions of this agreement.
- 3.16 The MGRB Data Access Committee may terminate this agreement by written notice to the User Institution. If this agreement terminates for any reason, the User Institution will be required to destroy any Data held, including copies. This clause does not prevent the User Institution from retaining the Data for archival purposes in

conformity with audit or legal requirements. The User will not be required to delete or destroy Data that resides on electronic back-up tapes or other electronic back-up files that have been created solely by the User Institution's automatic or routine archiving and back-up procedures but agrees to preserve the confidentiality of such Data.

- 3.17 The User Institution accepts that it may be necessary for the MGRB Data Access Committee to alter the terms of this agreement from time to time. In the event that changes are required, the MGRB Data Access Committee or their appointed agent will contact the User Institution to inform it of the changes and the User Institution may elect to accept the changes or terminate the agreement by written notice to the other party to this agreement.
- 3.18 If requested, the User Institution will allow data security and management documentation to be inspected by an agent of the other party to verify that it is complying with the terms of this agreement.
- 3.19 The User Institution agrees to distribute a copy of these terms to the Authorised Personnel. The User Institution will procure that the Authorised Personnel comply with the terms of this agreement.
- 3.20 The parties acknowledge that this agreement does not address the law that governs disputes arising out of this agreement or the subject matter of this Agreement.

## **SECTION 4 - PUBLICATION POLICY**

The User Institution is required to include an appropriate acknowledgement (and authorship if deemed necessary) in any Publication that makes use of or reference to an MGRB dataset in accordance with the following terms. Manuscripts or presentations may be submitted to [MGRB@garvan.org.au](mailto:MGRB@garvan.org.au) for review.

### **4.1 Open Access and/or Controlled Access data**

Publications that make use of Open Access data will be required to acknowledge the MGRB Partners in the acknowledgements section of the Publication.

Should the User Institution wish to add any personnel from an MGRB Partner to the list of authors, the User Institution agrees to forward the Publication to the MGRB Data Access Committee for approval prior to submission. Manuscripts or presentations may be submitted to [MGRB@garvan.org.au](mailto:MGRB@garvan.org.au) for review. Authors are also encouraged to recognise the contribution of the appropriate cohort convenors via the acknowledgements section in their Publication.

An example of a proper attribution is:

"The results <published or shown> here are in whole or part based upon data generated by the MGRB Partners: <https://sgc.garvan.org.au/initiatives/mgrb>

### **4.2 Restricted Access (or Tier 3) data**

When a User Institution applies to use Restricted Access data the arrangements for publication review, acknowledgement and/or authorship will be determined by the entity in the MGRB Partners that owns the relevant study. The MGRB Data Access Committee must be notified of any Publication arising from use of or containing Restricted Access data at least one month prior to submission (or presentation).

## SECTION 5 - USER INFORMATION

### 5.1 USER INSTITUTION ("Recipient")

Institution Name:

Address:

Phone:

### 5.2 PRINCIPAL INVESTIGATOR ("User")

Name:

Address:

Phone:

Email:

### 5.3 AUTHORISED PERSONNEL

(personnel at the User Institution to be granted access to the Data).

<b>Name</b>	<b>Position/ Job Title</b>	<b>Department/ Division</b>	<b>Email/ Phone</b>	<b>NCI-Australia Username ( for applicants accessing the MGRB at NCI-Australia</b>

**Executed as an Agreement.**

Both parties confirm that they have read, understood and accept the terms and conditions outlined in this Application.

**SIGNED** on behalf of the **USER INSTITUTION**, by its duly authorised officer, in the presence of:

\_\_\_\_\_  
Signature of officer

\_\_\_\_\_  
Signature of witness

\_\_\_\_\_  
Name

\_\_\_\_\_  
Name

\_\_\_\_\_  
Date

**Electronically SIGNED and DATED on behalf of THE MGRB DATA ACCESS COMMITTEE which authorises access to the MGRB cohort by the Principal Investigator and named Authorised Personnel of this application. Authority to use this cohort is effective as of the date electronically stamped on each page of this document.**

**Please submit the completed application (as a PDF file) on Garvan’s Resource Entitlement Management System (REMS) instance: <https://rems.public.garvan.org.au/>**

**Instructions for the REMS submission are [here](#).**