

## GWAS Central Submission Guidance Notes

All data submitted to GWAS Central will remain the property of the data generators and/or submitters, and all records will be presented to database with links and acknowledgements leading back to the original data source. Any users who might wish to obtain non-aggregated data will be instructed to make suitable requests to the relevant submitter and their data access authorities.

Submissions can be submitted with embargo dates or conditions attached. We will still immediately process such datasets to ensure the submission is complete and useable, but we will not release the submitted data to the public until instructed to do so.

When submitting genetic association data and/or allele/genotype frequency data to GWAS Central, we require that the utilised Markers are all present in a major public marker/variation database (e.g., dbSNP). If this is not the case, we can assist you in depositing the Markers into a suitable database.

To submit genetic association and/or allele/genotype frequency data, please gather together the information listed below (fields marked \* and **in bold** are required). An [Excel Data Submission Form](#) and a completed [submission example](#) are provided to help you with this task. Each submission will equate to one Study in GWAS Central, but each Study (i.e., each submission) can include one or more Experiments.

Completed Excel submission forms and accompanying data files should be e-mailed to: [submissions@gwascentral.org](mailto:submissions@gwascentral.org)

## Study

The Study should be thought of as a databased version of a journal article. Therefore, send in one submission for each set of information you believe tells a single, useful story. All fields are free text.

<b>*Study Title</b>	Brief title for the study
Abstract	Brief summary of the Study and its findings
Background	Describe the background to the project
Objectives	Main objective of the study
Key Results	Important results from the study
Conclusions	Important conclusions
Study Design	Specify the Study design
Acknowledgements	Acknowledgements sentences
<b>*Author Names</b>	List of Study authors (1 = corresponding author (CA))
CA Department	Corresponding author's department
<b>*CA Institution</b>	Corresponding author's institution
CA Address	Corresponding author's address
<b>*CA E-mail</b>	Corresponding author's e-mail address
CA WWW	Web link to the corresponding author's institutional webpage
Project WWW	List of web links related to the project responsible for the Study
Pubmed IDs	List of PubMed IDs for any papers describing the Study

## Phenotypes

Submissions may include any number of Phenotypes per Study, and these are broken down into the name of the phenotype trait (e.g., 'blood pressure', or 'Alzheimer's Disease') and the method used to test for that phenotype trait (e.g., 'manual use of blood pressure cuff', or 'clinical evaluation of AD using DSM3 criteria'). The observation or result obtained by applying the method is not listed here, but in the Sample Panel and Assayed Panel section. All fields are free text.

<b>*Phenotype Name</b>	A name for the Phenotype trait
<b>*Method Name</b>	A name for the method used to test the Phenotype
Method Details	Description of details of the method ( <i>include such things as the assayed biological sample, the measured attribute, the circumstance of the test, the time instant and period of operation, the units of measurement, etc</i> )
Description	Additional comments to delineate the Phenotype

## Panels (1): Sample Panels

Submissions may include any number of Sample Panels per Study (but at least one is required), and the details for each of these should be separated one above the other towards the top of the 'Panels' sheet in the Excel submission document. Fields are free text unless otherwise indicated.

<b>*Sample Panel Name</b>	Unique name for the Sample Panel
<b>*Description</b>	General description of the Sample Panel
<b>*Phenotypes Available</b>	List of Phenotypes for which measurements are available for the Sample Panel ( <i>list of values from the list of submitted Phenotype Names</i> )
<b>*No. of Individuals</b>	Total number of people in the Sample Panel ( <i>integer</i> )
No. of Males	Total number of males in the Sample Panel ( <i>integer</i> )
No. of Females	Total number of females in the Sample Panel ( <i>integer</i> )
Mode of Recruitment	Describe how the individuals were recruited
Sampling Period	When were the individuals recruited
Sampling Age Range	How old are the individuals
Population Info	What 'population' label best fits this Sample Panel
Geographic Region Info	Where in the world do the individuals live
Ethnicity Info	What ethnicity are the Sample Panel members
Environment Info	Describe environment exposures of importance

## Panels (2): Assayed Panels

Submissions may include any number of Assayed Panels per Study (but at least one case and one control panel are required), and the details for each of these should be separated one above the other on the 'Panels' sheet in the Excel submission document. The complete list of Assayed panel details should be placed below the Sample Panel details. Fields are free text unless otherwise indicated.

<b>*Assayed Panel Name</b>	Unique name for the Assayed Panel
Description	General description of the Assayed Panel
<b>*No. of Individuals</b>	Total number of people in the Assayed Panel <i>(integer)</i>
No. of Males	Total number of males in the Assayed Panel <i>(integer)</i>
No. of Females	Total number of females in the Assayed Panel <i>(integer)</i>
<b>*Sample Panel Names</b>	List of Sample Panels from which the Assayed Panel is assembled <i>(one or more Names from the submitted Sample Panel Names)</i>
<b>*Sample Panel No.s</b>	List of number of individuals taken from each of the Sample Panels listed in last data item <i>(one or more integers)</i>
<b>*Phenotype Criteria</b>	Phenotype criteria that were used as the basis for choosing which individuals to include in the Assayed Panel
Environment Criteria	Environment criteria that were used as the basis for choosing which individuals to include in the Assayed Panel

## General information for individual Experiments

At present we can only accept data from case-control type association studies, though other approaches (such as TDT) will be supported in the near future. Submissions may include any number of Experiments per Study, and the general information for each of these should be separated one above the other on a single sheet in the Excel submission document. An Experiment should be described for each each new dataset and an experiment must relate to the genetic basis of **one** phenotype. The Experiment Title can be reused only if the Statistical Method varies between Experiments (see the [submission example](#)). Fields are free text unless otherwise indicated.

<b>*Experiment Title</b>	Brief title for the Experiment
<b>*Phenotype Tested</b>	Phenotype tested for genetic association ( <i>one of the list of submitted Phenotype Names</i> )
Objective	Summarise the Experiments Objective
Outcome	Summarise the Experiments Outcome
<b>*Genotyping Platform</b>	List of technologies used to genotype the experiment participants
<b>*Individual Data Availability</b>	Is individual level data available upon request? ( <i>yes/no</i> )
<b>*Access To Source Data</b>	How to access the source data, with URL if possible
<b>*Statistics Method</b>	Name of the statistical method(s) employed
Statistics Details	Details of the statistical method(s) employed
Comments	Further information regarding the Experiment

## Association Data

For each experiment aggregated association results can be supplied in external data files. As a minimum, the data file should contain two columns: the SNP identifier (e.g. dbSNP rs number) and aggregated p-value (i.e., no individual level results) and be supplied as either Excel spreadsheets or plain text (e.g. TSV or CSV) formats (i.e. not PDFs or embedded Word tables). The data file(s) must either be downloadable online, available as online journal supplementary material or submitted to GWAS Central via e-mail along with the Excel template. Fields are free text unless otherwise indicated.

<b>*Data File Name</b>	Name of the file containing association data
<b>*Experiment Title</b>	Experiment that this data file relates to ( <i>from the submitted Experiment Titles</i> )
<b>*Statistical Method</b>	Statistical Method pertaining to the Experiment the data file relates to ( <i>from the submitted Experiment Statistical Method</i> )
<b>*Data Name</b>	Name to describe the data from an association experiment
URL	Where the file can be obtained if available online