

Dear Researcher,

As per UK Biobank's [Access Procedures](#) and [Data Management and Sharing Plan](#), you are required to return to UK Biobank the results of your research project within six months of publication or twelve months of the end-date of the research project, whichever comes first.

The reasons for requesting these results data are two-fold:

- To incorporate important derived data-fields into the database so they can be used by others;
- To make available the underlying methodology used to generate the main findings of the research (i.e. the analytical code), so that others can confirm, refute or expand on your published analyses.

What 'Results Data' do I need to return?

We would be grateful if you could return the following, where applicable

1. All key derived data-fields, plus a description of the methods used to generate them (if not available in sufficient detail in the publication);
2. The underlying syntax/code (e.g., SAS, R, STATA) used to generate the main results of the paper;
3. The RUN ID of the dataset used to generate the main results. (The RUN ID can be found in the data release notification e-mail. Please note that if multiple datasets were used to generate the main results, all RUN IDs should be provided);
4. Approximate dates when you accessed the bulk data (e.g., genomics, imaging) and record linkage data (i.e., HES) that generated the main results;
5. A copy of the final manuscript;
6. A short and simple paragraph that summarises your published findings, so that we can keep participants informed about research achievements.

In the event that you are unable to make certain analytical tools available – for example a confidential proprietary algorithm or third party software – then we would ask that you a) consider carefully whether you are able to do so by providing a fair, reasonable and consistent licence of the analytical tools, and/or b) provide appropriate assistance to researchers who may contact you, by for example providing details of the third party software provider.

What 'Results data' do I need to return for GWAS analyses?

In order to make GWAS results as consistent as possible for future meta-analytical purposes, we suggest the following derived genetic fields are returned, where applicable:

- **SNP**: SNP rs number
- **CHR**: chromosome number
- **POS**: base pair position of the SNP on the chromosome
- **A1**: effect allele (i.e. A, C, G, T)
- **A2**: non-effect allele (i.e. A, C, G, T)

- **A1/A2:** to denote the sequence reference allele
- **EAF:** effect allele frequency
- **Beta/OR** (effect estimate (log odds ratio))
- **se** (standard error of beta)
- **P** (nominal P-value)
- **N** (sample size details, e.g, N_cases, N_controls or N_analyzed)
- **INFO** (imputation quality score)

Please can you also include details of what software model was used to conduct the analyses, what covariate adjustment have been made (if any), and whether P-values and standard errors are corrected for genomic control (or not).

How do I return results?

Please inform the access team (access@ukbiobank.ac.uk) within 6 months of publication, or 12 months of the end-date of the research project, whichever comes first, that you wish to return Results Data. They will provide further details on how to upload the data.

How will my results data be used by UK Biobank?

Derived data-fields that may be of use to other researchers will be incorporated into the resource and made available via the Data Showcase (<http://biobank.ctsu.ox.ac.uk/crystal/>). This process will be undertaken in consultation with you (e.g., to help provide explanatory documentation) and full acknowledgement of the provenance of the data will be provided.

Should you have any questions, or require further clarification, please contact the Access Team at access@ukbiobank.ac.uk.

Many thanks for your co-operation – it's much appreciated.