**PRACTICAL OncoArray data files**

**This document provides a description of the summary statistics files**

**For any query please contact the PRACTICAL team (PRACTICAL@icr.ac.uk)**

Before analysing the data, please read the **guidelines for PRACTICAL OncoArray QC** document with a description of the QC steps that have been used for the OncoArray data.

**Description of the OncoArray summary statistics files**

Files named:

results\_onco\_sample\_euro\_bycountry\_nooverlap\_imp\_chr\*\_varname\_se.fixed.txt

Included in the data release is the summary results of the **imputed data** for the overall analysis of samples excluding any overlap with other previous GWAS. This is the most appropriate set for merging results with other GWASes. e.g. iCOGS, CRUK 1&2, CAPS 1&2, BPC3, Pegasus GWASes.

For the overall analysis, effects were adjusted for 7 European principal components. A fixed effects model was used. The data was analysed stratified by country. Those countries with less than 50 samples (cases or controls) were excluded from analysis. The country variable is based on the country of the Study (not the country of the individual). Similar criteria were used for the rest of the analyses.

It is recommended that the LRT scores are used for rare variants.

The fields included in the summary statistics file for the **imputed data** are:

|  |  |
| --- | --- |
| **Variable** | **Description** |
| varname | Unique identifier for variant (Composed of chr\_position\_build37\_a0\_a1) |
| SNP | Name the variant was imputed as – either a 1000 Genomes identifier or the Oncoarray SNP Name. |
| Position or position\_b37 | Position\_Build37 |
| Effect | Effect allele (which refers to allele used to calculate the OR)  |
| Baseline | Alternative allele |
| EAFcontrols | Effect allele frequency in controls |
| EAFcases | Effect allele frequency in cases |
| r2 | Imputation r2 |
| NumCalled | No. of samples, 72,729 (27,904 controls / 44,825 cases) |
| OR | Beta (refers to effect allele)  |
| SE | Standard Error |
| chi2 | Chi2 test  |
| LRT | Likelihood ratio test |
| ScoreStat | Score test |
| se\_lrt | Corrected standard error based on LRT (se\_lrt=sqrt(beta^2/LRT) ) for meta-analysis purposes (if available) |