Introduction to the statistical analysis of genome-wide association studies

DATE: 24 to 28 January 2022

VENUE: Online live via Zoom

AUDIENCE:

Geneticists facing the need to analyse large-scale human genotyping data in relation to their effect on common human traits and diseases

Scientists aiming to undertake genome-wide association studies and their meta-analyses

Researchers willing to understand better the statistical approaches and analytical procedures for the genetic association studies

COURSE LEADERS:

Inga Prokopenko, PhD, Professor of E-One, Health, University of Surrey, Guildford, UK

Andrew P. Morris, PhD, Professor of Statistical Genetics, University of Manchester, Manchester, UK

Reedik Mägi, PhD, Professor, Head of Bioinformatics workgroup, Estonian Genome Centre, University of Tartu, Tartu, Estonia

Krista Fischer, PhD, Professor of Statistics, Estonian Genome Centre, University of Tartu, Tartu, Estonia

Marika Kaakinen, PhD, Lecturer in Statistical Multi-Omics, University of Surrey, Guildford, UK

INFO: https://www.surrey.ac.uk/cpd-and-short-courses/introduction-statistical-analysis-genome-wide-association-studies

CONTACT: **gwascourse@surrey.ac.uk**

Course programme:

Day one:

Introduction to statistics for geneticists - Prof Inga Prokopenko and Dr Marika Kaakinen
Basics of probability theory, binomial and normal distribution, polygenic inheritance and complex traits, allele frequencies in population, Hardy-Weinberg equilibrium,

Introduction to Linux and R – Dr. Ayse Demirkan
Interface, command line and basic commands, functions, text editors, saving commands in scripts and running scripts, installing software tools for statistical analysis of genetic data, versions, data storage. Linux as environment for PLINK software tool. Basics of R usage to run graphical tools for genome-wide data and analysis results

Introduction to genome-wise association studies (GWAS) – Dr. Marika Kaakinen
Principles of linkage disequilibrium (LD) and SNP tagging for genome-wide genotyping array design, analysis and imputation; haplotypes, study design, sample size and statistical power, use UCSC browser and NHGRI GWAS catalog,

Day two:

Quality Control (QC) for GWAS – Prof Reedik Mägi
Sample and variant QC :on individuals (samples) for missingness, gender checks, duplicates and cryptic relatedness, population outliers, heterozygosity and inbreeding; and on SNPs for missingness, minor allele frequency and Hardy-Weinberg equilibrium.

Invited Speaker

Statistical models for genetic association studies – Prof Krista Fischer
linear and logistic regression, additive genetic model, test significance, type I error and multiple testing.

Day three:

Association analysis - Prof Inga Prokopenko
Analyses of data using PLINK software, including genetic models used for statistical analysis, covariates and adjustments, basic types of single-variant analyses, graphical representation of the output results

Population structure – Prof Andrew P. Morris
Identification of population outliers in GWAS and methods for detecting and accounting for structure within populations. Use of PLINK for principal components analysis and association analysis adjusting for structure.

Day four:

Imputation of GWAS - Prof Inga Prokopenko
GWAS reference panels, including HapMap and 1000 Genomes Projects, reference haplotypes, imputation with IMPUTE software, phasing and imputation steps, chromosome chunks, combining chinks for imputed data analysis, quality of imputation, imputed genotypes probability.

Invited Speaker

Meta-analysis of GWAS - Prof Andrew P. Morris
Combining association summary statistics across GWAS using fixed-and random-effects meta-analysis. GWAMA software to perform meta-analysis.

Day five:

Genetic risk scores, Mendelian Randomization - Prof Krista Fischer
Weighted and unweighted genetic risk scores. Dissecting causal relationships between exposures and complex traits using Mendelian Randomization. Instrumental variable approach.

Analysis of rare variants – Prof Andrew P. Morris
Rationale for rare variant analysis. Methods for assaying rare variation. Methods for the analysis of rare variants. GRANVIL software for testing association with rare variants.

Q&A session – all course leaders