Q&A

(5 points; indicate all correct answers; a correct set of answers given to a question adds up your score with 1 point – there is no penalty; questions are quite similar to exam multiple choice questions)

1. Under Hardy-Weinberg Equilbrium (HWE), the likelihood of an individual in a population carrying two different alleles of a human DNA marker, each of which has a frequency of 0.2, will be
   a. 0.4
   b. 0.16
   c. 0.08

2. The power of a genome-wide association analysis is affected by
   a. the design (family, case-control, cohort)
   b. genetic effect size
   c. linkage between the marker locus and the disease susceptibility locus

3. Also population stratification may affect the power of a genome-wide association study. In the GWAS context population stratification best refers to
   a. the existence of different ethnic groups in the population
   b. the existence of subgroups of individuals that are on average more related to each other than to other members of the wider population
   c. differences in allelic distribution between different population subgroups

4. The common disease-common variant model suggests that the genetic basis of of complex genetic phenotypes is explained by
   a. low frequency alleles of high penetrance
   b. low frequency allele of variable penetrance
   c. common alleles of moderate to low effects

5. Which of the following statements best describe the relationship between association and causation?
   a. Association does not imply causation
   b. Correlation and causation are synonymous
   c. Association is unrelated to causation

(10 points)

6. Open question: Explain where the application of Data Interoperability in healthcare would be most successful and why (minimum ½ page)
   A good answer must include: a brief overview, choice of the domain (business, company, technical), explanation of benefits with respect to the chosen domain, surmountable obstacles justifying the choice, technical considerations on how to develop DI in that domain and hints on possible other paths.

Literature style

(15 points)

Papers are organized around the following 5 big themes, adding complexity to the basic GWAS settings seen in class. Each group selects one paper. Further instructions: see class of October 26.

Meta-analysis

Research
Open Access
07 Jul 2021
Nature Communications
Volume: 12, P: 1-8

Cross-ancestry GWAS meta-analysis identifies six breast cancer loci in African and European ancestry women

GWAS have enhanced our understanding for the genetic basis of breast cancer, but the majority of them were performed for European ancestry populations. Here, the authors use a cross-ancestry approach and report seven new variants associated with breast cancer risk among women of African ancestry.

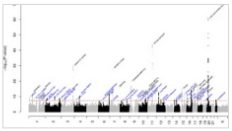
Babatunde Adedokun, Zhaohui Du ... Dezheng Huo

Research
Open Access
14 Jun 2021
Nature Communications
Volume: 12, P: 1-12

A large multiethnic GWAS meta-analysis of cataract identifies new risk loci and sex-specific effects

The genetic basis of cataract is not well understood. Here, the authors perform a genome-wide association multiethnic meta-analysis of cataract, finding 37 new loci and replicating known and new loci. They additionally perform sex-specific analyses, identifying new associations specific to women.

Hélène Choquet, Ronald B. Melles ... Eric Jorgenson



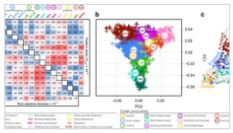
Population structure

Research
Open Access
11 Sept 2020
Nature Communications
Volume: 11, P: 1-11

Dutch population structure across space, time and GWAS design

Genetic variation in modern humans can reveal information about a population's history and migration patterns. Here, the authors describe the ancestry and geospatial genetic structure of the Netherlands, and demonstrate the utility of haplotype-based covariates in genome-wide association studies.

Ross P. Byrne, Wouter van Rheenen ... Russell L. McLaughlin

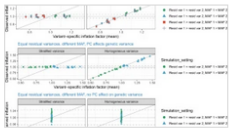


Research
Open Access
09 Jun 2021
Nature Communications
Volume: 12, P: 1-14

Variant-specific inflation factors for assessing population stratification at the phenotypic variance level

Pooling participant-level genetic data into a single analysis can result in variance stratification, reducing statistical performance. Here, the authors develop variant-specific inflation factors to assess variance stratification and apply this to pooled individual-level data from whole genome sequencing.

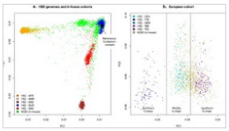
Tamar Sofer, Xiuwen Zheng ... Kenneth M. Rice



Research
Open Access
24 Sept 2021
Scientific Reports
Volume: 11, P: 1-14

Controlling for human population stratification in rare variant association studies

Matthieu Bouaziz, Jimmy Mullaert ... Aurélie Cobat



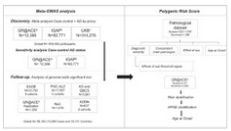
Polygenic risk scoring

Research
Open Access
07 Jun 2021
Nature Communications
Volume: 12, P: 1-16

Common variants in Alzheimer's disease and risk stratification by polygenic risk scores

Known genetic loci account for only a fraction of the genetic contribution to Alzheimer's disease. Here, the authors have performed a large genome-wide meta-analysis comprising 409,435 individuals to discover 6 new loci and demonstrate the efficacy of an Alzheimer's disease polygenic risk score.

Itziar de Rojas, Sonia Moreno-Grau ... Agustín Ruiz



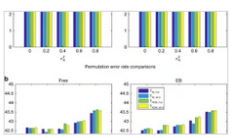
Reducing computational complexity and "alternative" phenotypes

Research
Open Access
14 Aug 2018
Nature Communications
Volume: 9, P: 1-13

Fast and powerful genome wide association of dense genetic data with high dimensional imaging phenotypes

Genome-wide association studies (GWAS) of neuroimaging data pose a significant computational burden because of the need to correct for multiple testing in both the genetic and the imaging data. Here, Ganjgahi et al. develop WLS-REML which significantly reduces computation running times in brain imaging GWAS.

Habib Ganjgahi, Anderson M. Winkler ... Thomas E. Nichols



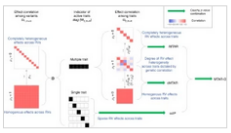
Multiplicity in the trait

Research
Open Access
05 Jun 2020
Nature Communications
Volume: 11, P: 1-11

Multi-trait analysis of rare-variant association summary statistics using MTAR

Methods to integrate association evidence across multiple traits often focus on individual common variants GWAS. Here the authors present multi-trait analysis of rare-variant associations (MTAR), a framework for joint analysis of association summary statistics between multiple rare variants and different traits.

Lan Luo, Judong Shen ... Zheng-Zheng Tang



Research
Open Access
04 Feb 2021
Scientific Reports
Volume: 11, P: 1-11

An evaluation of approaches for rare variant association analyses of binary traits in related samples

Ming-Huei Chen, Achilleas Pitsillides & Qiong Yang