The first 10-year population-based coronary disease genome-wide association study (GWAS) in more than 100,000 participants to personalize cardiovascular prevention in Spain.

A project of the CORDELIA Study
(Collaborative cOhorts Reassembled Data to study mEchanisms and Longterm Incidence of cArdiovascular diseases)

Isaac Subirana (PhD), Anna Camps (MSc) & Jaume Marrugat (PhD, PI) on behalf of CORDELIA investigators
contact: isubirana@imim.es

PURPOSE

Southern Europe is lacking sufficiently large cohorts and DNA biobanks to organize genome-wide association studies (GWAS) with cardiovascular disease (CVD), which come mainly from case-control studies.

1) To identify the genetic characteristics associated with the 10-year incidence of CVD in the Spanish population using a GWAS on ~102,000 participants form 24 Spanish cohorts already created, followed and duly combined.
2) To test new 10-year genetic risk score (GRS) for CVD adapted to the characteristics of the Spanish population, and to validate previously proposed GRS.

METHODS

Study design and population

- A collaborative and multicenter prospective cohort of 31 population-based pooled cohorts recruited in Spain in the last 30 years, with more than 167,000 participants, of whom > 102,000 have DNA samples still available.
- Spanish natives or residents 35 and 84 years old, 50% women, free of acute myocardial infarction, stroke or peripheral arterial disease at the time of recruitment.

Statistical analyses

- Data management: data collection, data merging, quality control reports, ...
- Genetic analyses: Genome Wide Association (GWAS) of observed and imputed SNPs, multiple testing, population stratification adjustment, Hardy-Weinberg Disequilibrium, Genetic and Polygenetic Risk Scores (GRS, PGR), ...
- Predictive models: assessment of prediction capacity and accuracy of multivariate linear, logistic regression and Cox models.
- Elaborate reports of intermediate and final results.

Work plan

- Data management
  - Data from 31 Spanish cohorts
  - 84 variables joined from ~170,000 participants
- Genetic analyses
  - 33.152 DNA non-extracted
  - 68.245 ADN extracted
- Predictive models
  - DNA Extracted & QC from 101,397 participants
- Complete data from ~170,000 participants

CONCLUSIONS

Student will:
- Clean and prepare complex data from big multicentric cohort study.
- Perform Genome-Wide-Association Study on more than 100,000 individuals.
- Evaluate prediction capacity of genetic scores using multivariate regression models and survival techniques.

CORDELIA CVD GWAS is:
- Feasible at 3 years
- Cost-effective
- Promising