

Prequential model selection methods in the assessment of genetic effects on long-term development of coronary artery disease.

Carlo Berzuini

Medical Research Council, Biostatistics Unit, Cambridge, UK

An important target for future studies of genetic association will be to assess the impact and role of genetic variants in the development of disease events. An ideal setting for this kind of investigation are studies based on the prospective follow-up of large cohorts of extensively genotyped patients.

In a motivating study on the genetics of coronary artery disease, 2000 patients, ascertained upon a first infarction before age 45, have been prospectively followed-up for a total of 20000 person-years. Recorded events include re-infarction, death and revascularisation interventions. Whole genome genotyping has been performed in each patient.

Prequential (predictive sequential) methods of model selection, implemented by using sequential Monte Carlo (particle filtering) algorithms, have been used to analyze the complex set of longitudinal data. This has allowed us to assess the contribution of genetic variants upon prediction of prospectively observed, pathway-specific, disease events. We shall discuss the assumptions under which these results may be interpreted causally, leading to important hypotheses about the role of specific variants in specific molecular pathways involved in the disease process.