Meta-analysis of individual participant data from observational studies

Thursday 21st March 2013, 1pm – 4 pm

Simon G. Thompson, Professor of Biostatistics
Department of Public Health and Primary Care, University of Cambridge, UK
Email: sgt27@medschl.cam.ac.uk

Meta-analyses of multiple studies, for which individual participant data (IPD) are available, are becoming more common. The aim of this session is to update participants on statistical methods that can be used for such analyses, and the pitfalls to be avoided. The focus will be on observational studies rather than randomised trials. Available software will be discussed. The session will be organised as four 30-minute presentations, each allowing 10 minutes for discussion and questions. Specifically, the presentations will cover the following topics:

1. The basics of meta-analysis

To ensure participants have the relevant background, the session starts with a résumé of conventional statistical methods used in meta-analysis:

- The assumptions behind fixed-effect and random-effects analyses
- The interpretation of overall estimates, confidence and prediction intervals
- Quantifying and investigating heterogeneity
- The advantages of IPD over summary data
- One-step and two-step meta-analysis methods

2. Meta-analysis of observational epidemiological studies

Using examples from the Emerging Risk Factors Collaboration, a consortium of over 100 epidemiological studies that has shared individual data on over 1 million participants, the following issues will be addressed:

- Summarising data in each study in a consistent way
- Meta-analysis of log hazard ratios
- Adjusting for covariates and analysing interactions
- Investigation of heterogeneity: separating within- and between-study information
- Handling confounders that are completely missing in some studies

3. Meta-analysis of risk prediction metrics

Whether risk factors aid the prediction of clinical events is often assessed by metrics such as Harrell's C-index, Royston-Sauerbrei's D-statistic, and Pencina's net reclassification index (NRI). Meta-analysis of these metrics across studies presents some problems in practice:

- Deriving C and D in individual studies
- Choice of weighting in a meta-analysis
- The change in C and D, and the NRI, on adding a new risk factor
- Heterogeneity between studies
- Examples from the Emerging Risk Factors Collaboration

4. Meta-analysis of Mendelian randomisation studies

Mendelian randomisation is the application of instrumental variable techniques using genetic variants to estimate the causal effect of a risk factor on an outcome from observational data. Large studies are typically needed to estimate causal effects precisely, and meta-analysis is often required. Topics to be addressed include:

- The principles of Mendelian randomisation, and instrumental variable analysis
- Two-stage analyses and weak instrument bias
- A one-stage analysis that minimizes weak instrument bias
- Dealing with partially missing risk factor data
- Example of C-reactive protein and coronary heart disease

References

Burgess S, Thompson SG; CRP CHD Genetics Collaboration. Bayesian methods for meta-analysis of causal relationships estimated using genetic instrumental variables. *Statistics in Medicine* 2010; **29**: 1298-1311.

Burgess S, Thompson SG; CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. *International Journal of Epidemiology* 2011; **40**: 755-764.

Burgess S, Thompson SG; CRP CHD Genetics Collaboration. Methods for meta-analysis of individual participant data from Mendelian randomisation studies with binary outcomes. *Statistical Methods in Medical Research*, online June 2012.

Emerging Risk Factors Collaboration. Major lipids, apolipoproteins and risk of vascular disease. *Journal of the American Medical Association* 2009; **302**: 1993-2000.

Emerging Risk Factors Collaboration. C-reactive protein, fibrinogen, and cardiovascular disease prediction. *New England Journal of Medicine* 2012; **367**: 1310-1320.

Jackson D, White IR; Fibrinogen Studies Collaboration. Systematically missing confounders in individual participant data meta-analysis of observational cohort studies. *Statistics in Medicine* 2009; **28**: 1218-1237.

Jackson D, White IR, Thompson SG. Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. *Statistics in Medicine* 2010; **29**: 1282-1297.

Thompson SG, Kaptoge S, White IR, Wood AM, Perry PL, Danesh J; Emerging Risk Factors Collaboration. Statistical methods for the time-to-event analysis of individual participant data from multiple epidemiological studies. *International Journal of Epidemiology* 2010; **39**: 1345-1359.

Wood AM, White IR, Thompson SG; Fibrinogen Studies Collaboration. Correcting for multivariate measurement error by regression calibration in meta-analyses of epidemiological studies. *Statistics in Medicine* 2009; **28**: 1067-1092.