ATELIER « INFÉRENCE CAUSALE EN RECHERCHE SUR LA SANTÉ »
9 – 13 MAI 2011

**WORKSHOP ON "CAUSAL INFERENCE IN HEALTH RESEARCH"**
**MAY 9 – 13, 2011**

# The random discontinuity design in epidemiology

Sara Geneletti [*]
S.Geneletti@lse.ac.uk

Gianluca Baio [†]
gianluca@stats.ucl.ac.uk

It is important to determine whether a particular drug is being prescribed at the point in disease progression where they will be most beneficial to the patients and result in lower overall costs incurred in the treatment. While clincal trials go some way towards understanding ideal dosages and prescription guidelines, those are not their principal goals. Setting up prescription guidelines falls to expert clinicians in the field who base these on available data from experimental and observational studies and experience.

The regression discontinuity design (RDD) described below provides a way of evaluating whether a particular prescription rule is successful or whether it should be amended. RDD was first introduced in the educational psychology literature in the 1960s Thistlethwaite and Campbell (1960) and recently become popular in the economics Van der Klaauw (2008), Imbens and Lemieux(2008).

RDD is a natural experiment that exploits the fact that many treatments are assigned according to pre-defined rules. An example is the prescription of statins, a class of cholesterol-lowering drugs. Individuals are prescribed statins if their risk of developing CVD in the subsequent 10 years, as calculated by an appropriate risk calculator, exceeds 20%. The idea of the RDD is that individuals on either side of this threshold are likely to be very similar and can therefore be regarded as forming two groups of randomly treated and untreated individuals. Consider for example two men of 50, both smokers with systolic blood pressure of 200 and total cholesterol of 4 mmol/L, one with HDL cholesterol of 2 mmol/L and the other of 1.9 mmol/L. These two men do not belong to different populations. However using the British National Formulary cardio vascular risk calculator, the former has risk of developing CVD in the next 10 years of 19.4%, meaning he is not prescribed statins while the latter's risk is 20% and he is prescribed statins. The small difference in HDL cholesterol can be seen as a random fluctuation which has resulted in one man receiving treatment and the other going without.

Generally, if we can plausibly assume that individuals within a certain distance of the threshold belong to the same population with respect to the characteristics that inform the assignment rule and determine the outcome, then the threshold can be seen as a quasi-random intervention which assigns the treatment to those that are just above the threshold and assigns no treatment to those that fall just below the threshold. We can then exploit this random assignment to estimate the (causal) effect of the treatment for individuals in the region around the

[*]Department of Statistics, London School of Economics, Houghton Street, London, WC2A 2AE, United Kingdom.
[†]Department of Statistical Science, University College London, 1 - 19 Torrington Place, London, WC1E 6BT, United Kingdom.

threshold. The estimated effect can tell us whether the drug is being prescribed at a suitable threshold. For example if no effect is found then this indicates that perhaps the threshold should be moved. If the effect is very large, the threshold should be lowered to allow more individuals to benefit from it. In this talk we describe the RDD and its potential applications in epidemiology by using a subsample of the THIN database which contains anonymised patient records for a sample of UK general practitioners.

International Biometric Conference, Floripa, Brazil, 5-10 December 2010