Wednesday, 27th August 2014 – 9:00-11:00

Invited session

S1 STRATOS (Strengthening Thinking about Analyses of Observational Studies) initiative: first results & future steps

Organizers: Willi Sauerbrei and Harbajan Chadha-Boreham

S1.1
Setting the stage with initial data analyses

M Huebner[1], S Le Cessie[2], W Vach[3], M Blettner[4], D Bodicoat[5]
[1]Michigan State University, East Lansing, United States, [2]Leiden University Medical Center, Leiden, The Netherlands, [3]Universitaet Freiburg, Freiburg, Germany, [4]Universitaet Mainz, Mainz, Germany, [5]University of Leicester, Leicester, United Kingdom

The importance of initial data analyses in observational studies has been recognized but is often neglected. Careful data preparation and description is crucial before embarking on complex analyses to avoid spurious results.
Obtaining high quality data starts far before data collection and includes a careful database design with variable definitions, within and between variable plausibility checks and date checks. Data should be cleaned systematically and carefully, especially when integrating multiple data sources. Changes to the data such as corrections, transformations, definitions of categories, or treatment of missing data need to be integrated in the programming code for reproducibility, rather than changes to the raw dataset. Data cleaning can take as much as 80% effort of the analysis, and the process may need to be automated for large data sets. The inclusion and exclusion criteria in the process of selecting the subset of data to be analyzed in the study should be described with an overview of missing measurements and follow-up data. Numerical and graphical descriptions include table summaries, illustrating correlations and confounding factors, or examining distributions of variables and homogeneity of groups, keeping in mind the objectives of the study.
It is especially important that the complete initial data analysis process is transparent and that researchers document all steps for reproducibility. STRATOS topic group 3 aims to provide guidance on this process based on an overview of existing literature with examples and feedback from experienced statisticians.

S1.2
Evaluation of incremental value of a marker: a historic perspective on the Net Reclassification Improvement

EW Steyerberg[1], P Macaskill[2], AV Vickers[3]
[1]Erasmus MC, Rotterdam, The Netherlands, [2]University of Sydney, Sydney, Australia, [3]Memorial Sloan Kettering Cancer Center, New York, United States

The net reclassification improvement (NRI) is an increasingly popular measure for evaluating improvements in risk predictions. In a recent review, 67 publications were considered from high-impact general clinical journals that considered the NRI. Incomplete reporting of NRI methods, incorrect calculation, and common misinterpretations were found. To aid improved applications of the NRI, the article elaborated on aspects of the computation and interpretation in various settings [1].
An accompanying Editorial emphasized conceptual problems (such as weighting reclassification inappropriately), and statistical problems (such as artificial inflation of NRI values and Type I error). It doubted that anything would be gained by reporting the NRI, either overall or in its components. The Editorial suggested that investigators should move away from statistical abstractions, such as the NRI, and illustrate the consequences of using a marker or model in straightforward clinical terms [2].
In this presentation, we will discuss the use of NRI and related performance measures from a historical perspective and suggest directions for improvement, particularly with respect to the use of decision analytic measures.

References:

1. Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. Ann Intern Med. 2014 Jan 21;160(2):122-31.
2. Vickers AJ, Pepe M. Does the net reclassification improvement help us evaluate models and markers? Ann Intern Med. 2014 Jan 21;160(2):136-7.

S1.3
Review of methods used in recent observational epidemiological studies to select variables and their functional forms [STRATOS Task Group 2]

M Abrahamowicz[1], RP Kyle[1]
[1]McGill University, Montreal, Canada

The over-arching STRATOS goal is to improve statistical methodology used in real-life observational studies, focusing on selected 'generic' issues. Task Group 2 deals with selection of independent variables, and functional forms for continuous variables, in multivariable explanatory models. To convince end-users to adapt more sophisticated statistical methods, and demonstrate the weaknesses of 'conventional', currently applied methods, we reviewed methods in 50 papers published in 2013 in high-ranking epidemiology and clinical journals, which focused on some continuous variable(s).
Whereas several studies selected independent variables *a priori,* on substantive grounds, many used arbitrarily selected data-dependent criteria or procedures, and failed to account for, or even mention, their impact on the estimation and inference. As expected, most studies imposed *a priori* linearity of the effects of continuous variable(s), and did not test or evaluate this assumption. Review of additional papers revealed that linearity was often imposed even for associations consistently demonstrated to be non-linear, in previous flexible analyses (e.g. BMI versus mortality; age at diagnosis versus recurrence or death in various cancers). Studies that did consider possibly non-linear relationships employed different flexible methods (polynomials, fractional polynomials, and various spline-based approaches), and used different criteria to assess (non-)linearity.
In conclusion, current applied research will benefit from evidence-based guidance, and a systematic comparison of methods, for selection of variables and their functional forms.

S1.4
Causal questions and principled answers: a guide through the landscape for practicing statisticians

E Goetghebeur[1], E Moodie[2], I Waernbaum[3], S Le Cessie[4]
[1]Ghent University, Gent, Belgium, [2]Mc Gill University, Montreal, Canada, [3]Umea University, Umea, Sweden, [4]Leiden University Medical Center, Leiden, The Netherlands

Causal inference came a long way over the past decade. New methodological approaches casting assumptions, models and results in terms of potential outcomes find their way into the clinical literature. This opens great potential for deepened understanding, but also for misunderstandings if subtle assumptions or interpretation of results are misunderstood. When choosing the causal analysis method or synthesizing evidence from different approaches one should be clear about the specific questions they aim to answer and about what can (not) be achieved from available

observations. The choices lead to a focus on direct or indirect, conditional or marginal effects of a particular type of exposure for specific (sub)populations over a time horizon. We examine what is kept fixed and let loose for the practical question `what if exposure had been different'.

The choices come with different meaning and distinct technical challenges. We consider how practical questions and answers differ under the no unmeasured confounders assumption relying on outcome and/or propensity score models and/or matching; or the instrumental variables assumption in a marginal or (double) conditional set-up including principle strata. We point to tutorials on the separate methods, line up overlap and differences in a principled fashion and by example. We zoom in on effectiveness research aiming to learn about drug effects (such as statins) from electronic health records and on the evaluation of quality of care in terms of hospital outcomes.

Presented by Els Goetghebeur for STRATOS TG7 including Saskia Le Cessie, Erica Moodie, Ingeborg Waernbaum et al.