Statistical interaction in epidemiology and interaction among epidemiologists and statisticians

Dirk Taeger[1]
Jürgen Wellmann[2]

1 Institut für Prävention und Arbeitsmedizin der Deutschen Gesetzlichen Unfallversicherung, Institut der Ruhr-Universität Bochum (IPA), Bochum, Deutschland

2 Institut für Epidemiologie und Sozialmedizin, Westfälische Wilhelms-Universität Münster, Deutschland

Editorial

In the fall of 2011 some 40 German epidemiologists and statisticians convened in Bochum for a workshop on such diverse topics as "interactions in epidemiology" and "reproducible research" (http://www.ipa.ruhr-uni-bochum.de/specials/workshop2011.php). The scope of the workshop was broadened by a session for open topics in the area of statistics and epidemiology. This special issue of MIBE presents elaborated versions of a small selection of contributions to a multifaceted workshop.

The workshop itself was an example of one of its main topics – interaction – as it was a joint collaboration of five working groups, each of which belongs to one, two, or three out of a set of four parental academic societies, namely the German Region of the International Biometric Society (Internationale Biometrische Gesellschaft, Deutsche Region, IBS-DR), the German Society for Epidemiology (Deutsche Gesellschaft für Epidemiologie, DGEpi), the "Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V." (gmds) and the German Association for Social Medicine and Prevention (Deutsche Gesellschaft für Sozialmedizin und Prävention, DGSMP). The workshop was hosted by the Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr-Universität Bochum (IPA). This institute is also characterized by interaction of scientists of diverse fields.

The workshop was characterized by a productive atmosphere with lots of interesting presentations. Three of them have been prepared for this special edition of MIBE. André Scherag [1] discussed gene-environment interactions in the context of genome wide association studies (GWAS). He addressed the phenomenon known as "missing heritability", which means that a GWAS indentifies some single nucleotide polymorphisms (SNPs) to be associated with a complex trait, but the explained variation of the outcome due to these SNPs remains to be unexpectedly low, compared to the heritability estimated by studies like twin studies, family studies or adoption studies. One possible explanation for this discrepancy might be some interaction between genes and the environment. Scherag cites an example of variants that have been shown in GWAS to have an effect on the body mass index, but this effect is attenuated in physically active individuals.

Heiner Claessen and colleagues [2] focused on the impact of diabetes on the mortality in patients with stroke. For this analysis data of one big statutory health insurance company in Germany were analyzed. The advantage of using health service data is their availability. Setting up cohort studies consumes more time and money. A disadvantage of such routine data is that they are not collected for scientific research and may lack important information. Claessen and his co-authors have shown that men have a time dependent mortality risk of diabetes following their first stroke.

In the open topics session Anne Lotz [3] contributed an interesting talk on the statistical analysis of measurements that are subject to detection limits. In some areas epidemiologists are quite often confronted with outcomes that cannot reliably be measured if they fall below a certain threshold. One may be tempted to tackle this problem by ad-hoc methods, but Lotz and her coworkers demonstrate that modern imputation methods or Tobit regression outperform these simplistic approaches. Tobit regression is a parametric technique for censored outcomes. It has been introduced in econometrics, but epidemiologists know its principles from survival analysis. In the context of measurements with detection limits the data

are left censored. This can also apply to survival data, where interval censored data might also occur, and right censored survival times are the rule rather than the exception.

Dirk Taeger and Jürgen Wellmann for the local organizers and the working groups

References

1. Scherag A. Missing heritability of complex traits and G-E interactions. GMS Med Inform Biom Epidemiol. 2013;9(2):Doc06. DOI: 10.3205/mibe000134

2. Claessen H, Icks A, Morbach S, Glaeske G, Hoffmann F. Time dependent impact of diabetes on mortality in patients with stroke: Survival up to 5 years in a health insurance population cohort in Germany. GMS Med Inform Biom Epidemiol. 2013;9(2):Doc07. DOI: 10.3205/mibe000135

3. Lotz A, Kendzia B, Gawrych K, Lehnert M, Brüning T, Pesch B. Statistical methods for the analysis of left-censored variables. GMS Med Inform Biom Epidemiol. 2013;9(2):Doc05. DOI: 10.3205/mibe000133

Corresponding author:
PD Dr. Jürgen Wellmann
Institut für Epidemiologie und Sozialmedizin, Westfälische Wilhelms-Universität Münster, Albert-Schweitzer-Campus 1, Gebäude D3, 48149 Münster, Deutschland, Phone: +49 (0)251/83-55648
wellmann@uni-muenster.de