

SOME OFTEN FORGOTTEN ISSUES IN STATISTICS APPLIED TO EPIDEMIOLOGY

Ch. E. Minder^{†1}

¹*Department of Social and preventive Medicine, University of Berne, Berne, Switzerland*

[†] E-mail: *cminder@ispm.unibe.ch*

This paper is not about a new technique or a new application of a known technique. Its aim is rather to bring to attention some points often forgotten in the application of statistics to epidemiology. The points raised are selected based on experiences of the author in the field in the last 25 years.

- Often, it is forgotten that data may not be what they appear to be: affected by random and systematic errors, even with best available practices
- Despite available literature, there appears to be a general lack of understanding on how to proceed in order to obtain a useful statistical model for a given epidemiological setting
- Due to the widespread availability of statistical software of stupendous power and flexibility, ever more complex models are fitted to increasingly complex data sets. See also above. This is done far too often without proper checking of models
- Sometimes, the absence of model checking is due to the rarity and lack of prominence of papers on appropriate methods
- Intrinsic collinearities pose an obstacle difficult to surmount in some areas of study
- The all-important rule of the applied statistician: "Balance your efforts in the various necessary directions" needs to be recalled

These points will be illustrated by examples and hints from the literature .