

BIAS FROM CATEGORISATION OF CONTINUOUS EXPOSURES MEASURED WITH ERROR

Darren C. Greenwood¹, Janet E. Cade², and Mark S. Gilthorpe¹

¹ *Biostatistics Unit, Centre for Epidemiology & Biostatistics, University of Leeds, UK*

² *Nutrition Epidemiology, Centre for Epidemiology & Biostatistics, University of Leeds, UK*

In epidemiology it is common practice for a continuous exposure to be categorised into discrete groups. This can lead to loss of information caused by placing individuals with different exposures into the same category. Secondly, if there is measurement error in the continuous exposure, then individuals may be attributed to the wrong category. We compare the relative roles of these two sources of error. One situation when epidemiologists often resort to categorisation for ease of presentation is when they expect an association may follow a nonlinear trend. We explore, through simulation, the potential bias associated with categorisation of a continuous exposure for linear and nonlinear dose-response curves. For the linear model, categorisation had only a small effect towards the null on average. For nonlinear trends, the bias was larger, and the mean bias could be in either direction. We illustrate situations where the average direction of the bias can be different for the extreme categories than for the intermediate ones. We conclude that categorisation does not aid the interpretation of data when there are nonlinear dose-response curves, but instead can introduce substantial bias, irrespective of whether the underlying continuous exposure is itself measured with error.