Correction to: Regression to the mean: what it is and how to deal with it

Adrian G Barnett,1* Jolieke C Van Der Pols2 and Annette J Dobson2

1Institute of Health and Biomedical Innovation, Queensland University of Technology, Kelvin Grove, QLD, Australia and 2School of Public Health, University of Queensland, Herston, QLD, Australia
*Corresponding author. Institute of Health and Biomedical Innovation (IHBI) & School of Public Health and Social Work, Queensland University of Technology, 60 Musk Avenue, Kelvin Grove, QLD 4059, Australia. E-mail: a.barnett@qut.edu.au

Back in 2005 we wrote about the ubiquitous problem of regression to the mean.1 We wrote our paper in reaction to seeing multiple cases in the literature, so thanks to regression to the mean we might expect fewer problems today. However, regression to the mean will always exist. It becomes a problem when researchers misinterpret the movements due to regression to the mean as a real change, and this is still happening.2

There was an error in our formula for calculating the expected regression to the mean (RTM) effect. The equation in the left-hand column of page 217 should have been:

\[ C(z) = \varphi(z)/(1-\Phi(z)) \]

The correlation in the example below the equation should have been 0.36 (1 minus the original figure of 0.64), giving an expected regression to the mean of 17.3 mg/dl instead of 9.6 mg/dl.

Our formula did not include the direction of the change, which depends on whether participants are chosen according to whether their baseline measurement is above or below the cut-off, \( c \). This can be rectified by changing the formula to:

\[ RTM \text{ effect} = \begin{cases} \sigma(1-\rho)C(z), & \text{selected to be } > c \\ \sigma(1-\rho)C(z), & \text{selected to be } \leq c \end{cases} \]

Acknowledgements

Thanks to Jamie Daw at Harvard University for highlighting the errors.

References