How well calibrated is EuroSCORE II?

Graeme L. Hickeya and Ben Bridgewatera,b,*

a University of Manchester, Manchester Academic Health Science Centre, The Northwest Institute for BioHealth Informatics, Manchester, UK
b Department of Cardiothoracic Surgery, University of Manchester, Manchester Academic Health Science Centre, University Hospital of South Manchester, Manchester, UK

* Corresponding author. Department of Cardiothoracic Surgery, University of Manchester, Manchester Academic Health Science Centre, University Hospital of South Manchester, Manchester, UK. Tel.: +44-161-3067875; e-mail: Ben.Bridgewater@uhsm.nhs.uk (B. Bridgewater).

Received 16 April 2012; accepted 7 May 2012

Keywords: EuroSCORE II • Model calibration • Model validation • Statistics

It has come to our attention that there appears to be an error in the recent article by Nashef et al. [1]. In the section ‘Calibration and discrimination of the new model’, it is stated that a validation dataset of 5553 records was used (of which 232 died). However, in Table 7, it states that the Hosmer–Lemeshow test was applied to 15 952 records (of which 605 died). A reader might conclude that the authors used the model training dataset; this was 16 828 records and cannot, therefore, be the case.

The $P$-value associated with the Hosmer–Lemeshow test was 0.0505, making it minutely greater than the clinically accepted de facto value of $P = 0.05$. It is conceivable that had the authors applied the test to the subset of data held back from the model validation, they would have rejected the null hypothesis of the model being calibrated. If the null hypothesis were to be rejected, this would not necessarily suggest that the model was poorly calibrated, and we would advise the consideration of a more comprehensive approach to the assessment of the model calibration [2]. We do, however, feel that an erratum should be published to fully inform practitioners of the actual model validity, especially if it is to be used for health-care decision making or governance.

Funding

This research was partly funded by Heart Research UK.

Conflict of interest: none declared.

REFERENCES