

Quality control in the discipline of clinical anatomy of the faculty of medicine of the university of porto: application of modern item response theory to multiple choice exams

Severo, M.¹ and Tavares, MAF.²

¹Centro de Educação Médica, Serviço de Higiene e Epidemiologia,
Faculdade de Medicina da Universidade do Porto

²Instituto de Anatomia J. A. Pires de Lima, Faculdade de Medicina da Universidade do Porto

The assessment of the overall quality of an exam is essential to the implementation and development of quality in the teaching/learning process of any course. The aim of this paper is to evaluate the quality of the multiple-choice questions in the Clinical Anatomy examination of the undergraduate medical course. One hundred and seventy six students answered to the theoretical examination of Clinical Anatomy (June 2007). The exam had 100 multiple-choice questions with 5 options, where only one option was the correct answer. The exam was organized according to a process of thematic and equitable distribution of the areas covered in the program. Item Response Theory (IRT) was used to assess the quality of the exam. The IRT is a theory of measurement that assumes that the performance on a given item can be explained by the latent ability of the examinee. Two parameter models were used in this study, which allowed to estimate the parameters of difficulty and discrimination of every question allowing to evaluate its quality. 157 out of 176 (89.2%) achieved positive rating, with an average a rate of 12.6 out of 20 and a standard deviation of 2.2. The slope estimative ranged from -0.723 to 1.902 , indicating considerable variation in item discrimination (these values correspond to factor loadings of -0.586 -0.885). The location parameters for the 100 items reflect a sizeable range of underlying knowledge (-17.8 - 21.1), but the majority of the items response categories are endorsed by students who had lower than average levels of knowledge, implying that the item set as a whole is most useful in discriminating among individuals at the lower end of the knowledge continuum. Eleven items showed none or negative discrimination parameter; 8 items of the 11 had lower discrimination due to the lack of detail knowledge of the students, 2 because the items had a double response and 1 because the item was not well written. It was possible to identify the causes of the problems expressed in the IRT analysis of the examination. The theoretical test discriminates, effectively, students with less "ability", but shows some difficulties in distinguishing students with higher levels of knowledge. These results showed the importance of quality control on the multiple-choice exam, in order to obtain better and more effective assessment of knowledge.