CME in the *Deutsches Ärzteblatt* and the development of multiple choice questions for medical educational purposes

Catrin Marx[1] and **Christopher Baethge**[1,2]

1 Deutsches Ärzteblatt, Cologne, Germany
2 Department of Psychiatry and Psychotherapy, University of Cologne Medical School, Germany

Correspondence to:

Catrin Marx
Deutsches Ärzteblatt
Ottostraße 12
D-50859 Köln
7011-135
Catrin.Marx@aerzteblatt.de

Abstract:

Continuing medical education (CME) in Germany, as in other countries, is an established instrument for the delivery of ongoing training in medicine. Since September 2004, *Deutsches Ärzteblatt* has published 155 CME papers, and, as of the 12th of August 2016, readers have logged on more than three million times to fill out a final evaluation form. The multiple-choice questions included in the CME modules in *Deutsches Ärzteblatt* have been analysed in a number of studies on medical education and compared with other German-language CME units from other providers. This article will present the mechanisms used at *Deutsches Ärzteblatt* to improve the quality of our CME units and reveal the aspects of medical education pedagogy that fed into this process.

CME

Since 2004, the modernisation of healthcare law (GKV-Modernisierungsgesetz) requires doctors working in Germany to acquire 250 continuing medical education (CME) points every five years. CME needs to be accessible and effective because if a physician fails to acquire 250 points, it can cause financial losses and, in the worst case, the withdrawal of their medical license.

CME points can be earned in various ways: by attending lectures, visiting conferences, or completing CME modules in certified journal articles. To offer CME modules in print, a journal must be certified by the relevant regional medical association. In addition, to be considered for inclusion in the CME programme, scientific papers have to pass an independent review process and their authors must declare their conflicts of interest.

Deutsches Ärzteblatt is the largest medical journal in Germany and is the official publication of the German Federal Medical Association (Bundesärztekammer) and the National Association of Statutory Health Insurance Physicians (Kassenärztliche Bundesvereinigung). Every doctor in Germany receives *Deutsches Ärzteblatt*. Including a special edition for doctors who are retired or no longer in clinical practice and another for practising clinical psychologists, the circulation is about 470,000. CME articles appear in the issues regardless of edition. As proof of participation in the module, the reader has to answer 10 multiple-choice (MC) questions relating to the content of the article. *Deutsches Ärzteblatt* readers can go online to find further information compiled by the authors together with a case example that illustrates the topic in a practical way. Readers who answer 70% of questions correctly are awarded three points. After answering the questions, participants are asked to score the article for quality and for relevance to their own clinical work. Unlike in medical school finals and medical specialisation exams, the MC-questions are intended merely to demonstrate whether an article has been read; they are not intended to function as an examination in the strict sense.

As *Deutsches Ärzteblatt* is the only free of charge journal that reaches every doctor working in Germany, CME modules must be targeted at an interdisciplinary audience and emphasise topics that are relevant to all specialties. Each
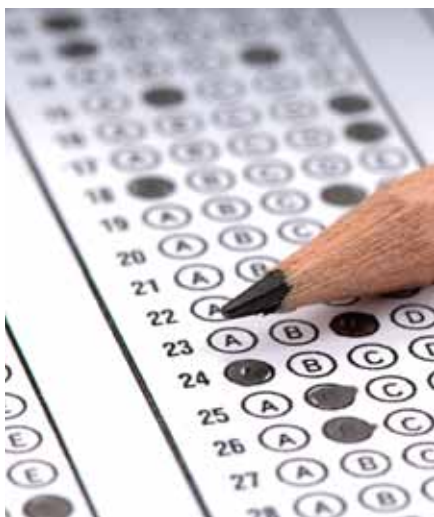
module can be accessed for three months. At the end of the three months, all successful participants receive a certificate. If they have any problems relating to the content of the module, participants can contact the editorial office. Thus, problems with the answers to the MC-questions usually become clear within a week of the release of the module. By this time, the editorial office has usually logged over 1,500 participants. These early data, combined with the proportion of correct answers, make it clear where the wording is ambiguous or there are errors in the questions.

If there is an error in the actual content of a question, an application must be made to the certificating body – in our case the Academy for Continuing Medical Education of the North Rhine Medical Association – to withdraw the question, and an erratum is published after the end of the 3-month access period. Furthermore, if too many participants fail the test the North Rhine Medical Association has the authority to remove approval for CME. Therefore, it is important to have high-quality, correct MC-questions.

The multiple-choice questions and answers

The test questions are designed to be answerable solely on the basis of reading the article, so that even doctors not working in the specialist field of the article in question should be able to participate in the CME simply by reading the article. The overall objective of the CME programme is to bring doctors up to date about new treatment strategies in other specialties than their own. In 2008 almost 95% of all MC-questions on the CME articles were answered correctly by the test takers.[2,3] Nevertheless, when Kühne-Eversmann *et al.* analysed the quality of MC-Questions in German-language medical journals, including *Deutsches Ärzteblatt*, she

found formal errors in the questions of all journals.[4]

In her study Lisa Kühne-Eversmann and co-authors analysed the quality of MC-questions in the journals *Der Internist*, *Deutsche Medizinische Wochenschrift*, and *Deutsches Ärzteblatt*.[4] Although the *Deutsches Ärzteblatt* preformed the best of these three journals, the study found faults in 60% of the questions. The main sources of error were described thus: "The stem was worded in such a general way that it failed to indicate what the topic was in any detail. As far as possible, the reader should be able to answer the stem without looking at the answer options." This means that questions such as 'Which of these statements is true?' or 'Which statement about XXXX is correct?' are poor questions because of the 'unfocused stem'. The main idea of the question should be placed in the stem, not in the choices. The test-taker should always know what he is asked for after having read the stem.[5]

Another common error was questions that require negative answers ('Which of the following statements is untrue…?'). Tamir (2013) states that this is an error in form because: "negatively phrased items require about twice as much working memory as equivalent positively phrased forms of the same item. Negative words appearing both in the stem and in one or more options might require four times as much working memory as a positively phrased forms of the same item".[6]

Another point of criticism related to the existence of 'cues' in the questions or answer options. Cues are unintentional hints that indicate the correct answer. Cue words include absolute terms such as 'always', 'never', 'only', 'exclusively', or 'alone'. An example of the use of such cues might be: 'In XY, the first-line medication is [A] always glucocorticoids, [B] immunosuppressants only, [C] always to use cytostatics.' Specific determiners such as 'only' and 'always' are so extreme that they cannot be the right answer.

Furthermore, an empirical study carried out by Rotthoff *et al.*[7] examined 200 MC-questions from 20 CME units in the journals *Der Internist*, *Deutsches Ärzteblatt*, *Medizinische Klinik*, and *Klinikarzt*. In all these journals, the authors found that there were too few questions that were geared to practical clinical use and actual clinical decisions.

Taken together, these findings prompted the

editorial staff at the *Deutsches Ärzteblatt* to develop some 'Rules for the construction of multiple-choice tests in *Deutsches Ärzteblatt*'. This guide was sent to every CME author. The following rules were developed partly based on the analysis of the above-mentioned two studies, [4,8] and partly on the research results of Krebs and Haladnya in the field of medical education.[5,9]

Style and format criteria
- The stem should be no longer than three sentences. The answers should be short.
- Negative words should be avoided in the stem and in the answers.
- The main idea of the question is placed in the stem not in the answers. It should be written in clear and understandable language.
- The items should not have options that give clues to the right answer (e.g. 'always', 'never', 'only', 'exclusively', or 'alone').
- Only one answer is correct.
- Do not use the choice "all of the above".
- The answers should be homogeneous in terms of content and grammatical form.
- The lengths of the choices should be the same.
- Avoid verbal associations between stem and the right answers.
  For example:
  Which *test* is to exclude Cushing syndrome?
  a. Dexamethasone suppression *test*
  b. Aldosterone renin quotient
  c. Cortisol determination
- The stem or an item of one question should not include the information for solving another question of the test.

Since then, Deutsches Ärzteblatt has also increased the emphasis on relevance to practice during the editorial development of the questions in order to ensure that the CME unit is fit for purpose in terms of the achievement of learning goals. Each set of questions must now contain at least two questions that relate to specific clinical situations, ask about clinical decisions, or are geared towards use in real clinical life.

To further increase the effectiveness of CME modules, three learning goals are now formulated at the beginning of each article. The contents of these learning goals are reflected in the MC-questions. Teaching of the most important contents is supported by special panels on each

page, giving visual prominence to short statements summarising the practice-relevant management guidelines, diagnostic classifications, and important advice about treatment.

To minimise the errors identified by Kühne-Eversmann et al., all MC-questions are checked by the editorial staff against the points of criticism and revised if necessary. Then, all the questions are given a test run before publication by a team from a variety of medical specialities. In addition to the quality criteria, the team also assesses whether the questions correspond properly to the defined learning goals and whether it is possible to answer them on the basis of reading the article alone. After this test run, if required, the questions undergo final modification.

Results after the changes

In a study in 2013, Drossard et al. investigated the proportion of formally incorrect questions in print CME publications dating from 2012[8] and compared them to data collected by Kühne-Eversmann et al. in 2006.

Therefore, one data set was collected after the introduction of the new quality standard ('Rules for the construction of multiple-choice tests') at Deutsches Ärzteblatt, and one data set was collected prior to this introduction. This new study included Der Internist, Deutsche Medizinische Wochenschrift, and Deutsches Ärzteblatt. Of all of these, the Deutsches Ärzteblatt showed the best results, demonstrating a clear improvement in quality from 2006 to 2012 (39% correct questions in 2006 and 67% in 2012). Furthermore, in a study published in 2013, Wolfgang Öchsner and Anja Böckers[10] investigated the extent to which CME MC-questions fulfil formal quality criteria. Three journals were compared: one surgical journal, one medical (non-surgical, non-interventional) journal, and one interdisciplinary

journal. According to Öchsner and Böckers, the unnamed interdisciplinary journal "was in fact Deutsches Ärzteblatt, the journal that did better than the others in all items of the study," (personal communication). It is evident that making guidelines and improving the procedures to be followed during editorial development increases the quality of MC-questions.

Future perspectives

Doctors encounter MC-questions from medical school onwards; this form of testing is ubiquitous in training and in continuing education. Öchsner and Böckers[10] believe, "it is important that the multiple-choice questions presented to readers of medical journals are methodologically exemplary and of high formal quality, so that they can implicitly serve as templates ready for adoption by readers who are themselves authors with responsibility for multiple-choice questions."[10] Perhaps, in the long term, good MC-questions could become a building block in the improvement of medical school final exams, and thus also increase the learning – and the enjoyment of learning – to be had from continuing medical education.

Conflicts of Interest and Disclaimers

Catrin Marx and Prof. Baethge declare that they are full-time editors of Deutsches Ärzteblatt.

References

1. Bundesärztekammer. Struktur der Ärzteschaft 2015 [cited 2016 Aug 8]. Available from www.bundesaerztekammer.de/fileadmin/user_upload/downloads/pdf-Ordner/Statistik2015/Stat15AbbTab.pdf

2. Marx C. CME im Deutschen Ärzteblatt: 43 000 Teilnehmer – sehr gute Bewertung. Dtsch Arztebl 2005;102(37):A 2453–6.

3. Baethge C, Marx C, Griebenow R, Leuschner H. CME in a General Medical Journal: Three-Year Evaluation of CME in Deutsches Ärzteblatt. Teaching Learning Med; 2010;22(3):229-32.

4. Christ H, Baethge C, Lehmacher W, Loesche P, Griebenow R. Interdisciplinary CME: Is the need evident? Results of the evaluation of CME articles in the Journal of the German Medical Association. J Eur CME 2015; 4: 1-6. view/28331/pdf_13

5. Kühne-Eversmann L, Nussbaum C, Reincke M, Fischer MR. cme-Fortbildungsangebote in medizinischen Fachzeitschriften: Strukturqualität der MC-Fragen als Erfolgskontrolle. Am Beispiel der Fachzeitschriften Deutsches Ärzteblatt, deutsche Medizinische Wochenschrift und der Internist. Med Klinik 2007;102(12):993-1001.

6. Stagnaro-Green AS, Downing SM. Use of flawed multiple-choice items by the New England Journal of Medicine for continuing medical education. Med Teach 2006;28(6):566–8.

7. Rotthoff T, Fahron U, Baehring T, Scherbaum WA. Die Qualität von CME-Fragen in der ärztlichen Fortbildung – eine empirische Studie. Z Ärztl Fortbild Qual Gesundhwes 2008; 101(10): 667–73.

8. Krebs R. Anleitung zur Herstellung von MC-Fragen und MC-Prüfungen für die ärztliche Ausbildung. Universität Bern. Available from: www.iawf.unibe.ch.

9. Haladyna TM, Rodriguez MC. Developing and validating test items. London: Routledge; 2013.

10. Drossard S, Kühne-Eversmann L, Bauer D, Fischer M. Strukturqualität von MC-Fragen in der ärztlichen Fortbildung. Welchen Einfluss hat die Einführung von Qualitätsmaßnahmen? Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA). Graz 26-28 Sept 2013. Düsseldorf: German Medical Science Publishing House; 2013. DocP05_01 (P05_01).

11. Öchsner W, Böckers A. [Medical journals as implicit role models for good multiple choice questions: how far have CME tests so far fulfilled formal quality criteria?]. Z. Evid Fortbild Qual Gesundhwes 2013;107:468–74.

Author information

Catrin Marx is part of the editorial staff of Deutsches Ärzteblatt since 1994. Her work focuses on CME articles.

Professor Baethge is Chief Scientific Editor of Deutsches Ärzteblatt (German Medical Journal).