Determinants of knowledge gain in evidence-based medicine short courses: an international assessment

Regina Kunz, Karl Wegscheider, Lutz Fritsche, Holger J Schünemann, Virginia Moyer, Donald Miller, Nicole Boluyt, Yngve Falck-Ytter, Peter Griffiths, Heiner C Bucher, Antje Timmer, Jana Meyerrose, Klaus Witt, Martin Dawes, Trisha Greenhalgh, Gordon H Guyatt

ABSTRACT

Background: Health care professionals worldwide attend courses and workshops to learn evidence-based medicine (EBM), but evidence regarding the impact of these educational interventions is conflicting and of low methodologic quality and lacks generalizability. Furthermore, little is known about determinants of success. We sought to measure the effect of EBM short courses and workshops on knowledge and to identify course and learner characteristics associated with knowledge acquisition.

Methods: Health care professionals with varying expertise in EBM participated in an international, multicentre before–after study. The intervention consisted of short courses and workshops on EBM offered in diverse settings, formats and intensities. The primary outcome measure was the score on the Berlin Questionnaire, a validated instrument measuring EBM knowledge that the participants completed before and after the course.

Results: A total of 15 centres participated in the study and 420 learners from North America and Europe completed the study. The baseline score across courses was 7.49 points (range 3.97–10.42 points) out of a possible 15 points. The average increase in score was 1.40 points (95% confidence interval 0.48–2.31 points), which corresponded with an effect size of 0.44 standard deviation units. Greater improvement in scores was associated (in order of greatest to least magnitude) with active participation required of the learners, a separate statistics session, fewer topics, less teaching time, fewer learners per tutor, larger overall course size and smaller group size. Clinicians and learners involved in medical publishing improved their score more than other types of learners; administrators and public health professionals improved their score less. Learners who perceived themselves to have an advanced knowledge of EBM and had prior experience as an EBM tutor also showed greater improvement than those who did not.

Interpretation: EBM course organizers who wish to optimize knowledge gain should require learners to actively participate in the course and should consider focusing on a small number of topics, giving particular attention to statistical concepts.

Regina Kunz is nephrologist, associate professor and vice director, Basel Institute for Clinical Epidemiology and Biostatistics, University Hospital Basel, Basel, Switzerland. **Karl Wegscheider** is statistician, biometrician, professor and director, Institute of Medical Biometry and Epidemiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. **Lutz Fritsche** is vice medical director, Charité, Berlin, Germany. **Holger J Schünemann** is internist, professor and chair, Department of Clinical Epidemiology and Biostatistics, CLARITY research group, McMaster University, Hamilton, Ontario, Canada. **Virginia Moyer** is pediatrician and professor, Department of Pediatrics, Baylor College of Medicine, Texas Children's Hospital, Houston, Texas, US. **Donald Miller** is professor and chair, Pharmacy Practice Department, College of Pharmacy, Nursing and Allied Sciences, North Dakota State University, Fargo, North Dakota, US. **Nicole Boluyt** is pediatrician, Department of Pediatric Clinical Epidemiology, Academic Medical Center, Amsterdam, The Netherlands. **Yngve Falck-Ytter** is gastroenterologist, assistant professor of medicine and director of hepatology, Case Western Reserve University and Louis Stokes VA Medical Center, Cleveland, Ohio, US. **Peter Griffiths** is professor and director, National Nursing Research Unit, King's College London, London, UK. **Heiner C Bucher** is internist, professor and director, Basel Institute for Clinical Epidemiology and Biostatistics, University Hospital Basel, Basel. **Antje Timmer** is internist and senior researcher, Helmholtz Center, German Research Center for Environmental Health, Institute for Epidemiology, Neuherberg, Germany. **Jana Meyerrose** is general practitioner trainee, Practice Center Rudolf Virchow, Berlin. **Klaus Witt** is general practitioner and associate professor, Department of General Practice, University of Copenhagen, Copenhagen, Denmark. **Martin Dawes** is general practitioner, professor and chair, Department of Family Medicine, McGill University, Montréal, Quebec, Canada. **Trisha Greenhalgh** is general practitioner and professor, Department of Primary Health Care, University College London, London. **Gordon H Guyatt** is internist and professor, Department of Clinical Epidemiology and Biostatistics and CLARITY research group, McMaster University, Hamilton.

From the time the term evidence-based medicine (EBM) was coined,[1] proponents of this approach to health care delivery have offered courses and workshops to teach and disseminate it. Some observers have expressed skepticism regarding the effectiveness of teaching EBM,[2] and multiple systematic reviews have not clarified which teaching approaches are optimal.[3–7] Nevertheless, the demand for courses and learning opportunities for EBM continues to increase, and EBM concepts have been disseminated worldwide.

Traditionally, EBM courses have involved problem-based, self-directed learning in which learners work in small groups supported by tutor-facilitators. Although the ingenuity of medical educators has led to the development of variations on the teaching approach and new educational formats targeted to particular settings, learner characteristics and local resources,[8–10] short courses and workshops remain a pillar of EBM dissemination. Empirical evidence to guide the design and delivery of short courses and workshops continues, however, to be limited. To close this gap, we undertook an international study to determine whether teaching EBM can produce changes in knowledge across a heterogeneous sample of course formats, course content and target groups and to explore which course features and learner characteristics are particularly suited to the acquisition of EBM knowledge.

Methods

Recruitment of workshops and learners. Through course announcements on the Internet, in an EBM mailing list (EVIDENCE-BASED-HEALTH@JIS-CMAIL.AC.UK) and in publications,[11] as well as through personal contacts, we identified short courses and workshops designed to teach EBM and invited the coordinators to participate in the study. We did not restrict participation by country, course format, setting or teaching style. We targeted courses for health care professionals but did not exclude courses for students.

Design. We performed a multicentre before–after study addressing the impact of short courses and workshops on EBM knowledge.[12] The primary outcome measure was the score on the Berlin Questionnaire, a validated instrument measuring EBM knowledge that the participants completed before and after the course.[11] To balance any potential difference in difficulty between the 2 alternative versions of the Berlin Questionnaire (designated sets A and B), we randomly assigned the learners to complete either set A or set B immediately before the course and then the other set immediately after the course. We advised course organizers to describe the study design to the learners and to inform them that participation was voluntary and that individual results would not be disclosed to learners.

Course characteristics. Coordinators of eligible courses provided information regarding the format and content of their course, including its duration, total teaching time, the number of tutors, the mix of small- and large-group activities, the number of learners in the entire course, the number of learners in each small group, and the target groups. Coordinators also indicated the learners' expected level of engagement: compulsory engagement (e.g., case preparations or presentations), optional engagement (e.g., discussions) or no expectation of engagement (e.g., lectures). We asked about topics covered (therapy, diagnosis, prognosis, harm, meta-analysis, guidelines and study design), whether additional sessions devoted to searching the literature, statistics and teaching techniques were offered, and about the cost.

Learner characteristics. We asked learners to provide personal information regarding their age and sex, education, current position, occupation, prior exposure to EBM (e.g., had attended a session lasting less than 1 day, participated in a workshop, read a book, or taught as a tutor) and their assessment of their knowledge of EBM. We also asked them whether their participation in the course was mandatory or voluntary.

Assessment tool: the Berlin Questionnaire. The Berlin Questionnaire[11] consists of 2 separate sets (A and B) of 15 multiple-choice questions built around typical clinical scenarios and focusing on therapy and diagnostics (4 questions each) and meta-analysis, prognosis and harm (1 question each), with 4 questions that cover a variety of categories. Each question has 5 response options. The questionnaire addresses the following domains of EBM: relating a clinical problem to a clinical question, identifying study design and interpreting the evidence.[13] The Berlin Questionnaire has documented content validity and internal consistency. The instrument is able to discriminate different levels of knowledge and is responsive to change.[11]

Analysis. We recorded the number of correct answers for each questionnaire administration and calculated the difference in respondents' scores before and after the course. To identify factors associated with degree of change in questionnaire score (the dependent variable),

we constructed a 2-level linear mixed model with course as the upper (cluster) level and learner as the lower (individual) level. Potential confounders and predictors were entered into or deleted from the model step by step as described below. All models included questionnaire sequence (A or B first) and baseline score.

To estimate the variance in knowledge gain between courses and between learners by restricted maximum likelihood (REML) estimates and the intraclass correlation (the variance ratio) that measures the cluster effect that results from the differences between courses, we began by constructing a model that included only sequence and baseline score. We calculated not only the main outcome (score change) but also the corresponding effect size (i.e., the difference in means divided by the square root of the pooled variance of the baseline score).[14] This allows comparison with the results of other studies and reflects a recent trend in evidence-based education.[15] In accordance with Cohen's interpretation, we labelled an effect size of 0.2 as small, an effect size of 0.5 as moderate and an effect size of 0.8 as large.[14]

We then entered 8 additional potential predictors (degree of active participation [required/encouraged], total number of learners, group size, ratio of tutors to learners, English or German language of questionnaire, number of topics, separate statistics session and total teaching time) at the course level and 11 potential predictors at the learner level (age, sex, degree in epidemiology or public health, current occupation, current position, education, previous exposure to EBM [had attended a session lasting less than 1 day, participated in a workshop, read a book, or taught as a tutor] and self-rated EBM knowledge) to the model and reduced the model step by step by deleting non-significant predictors, starting with the predictors with the highest p value. We stopped the stepwise backward elimination when only variables with p values below 0.05 remained in the model. We report the estimated parameter contrasts or regression coefficients of the final model with confidence limits and p values. All analyses were done using SPSS 13.5 (SPSS Inc., Chicago, Ill.).
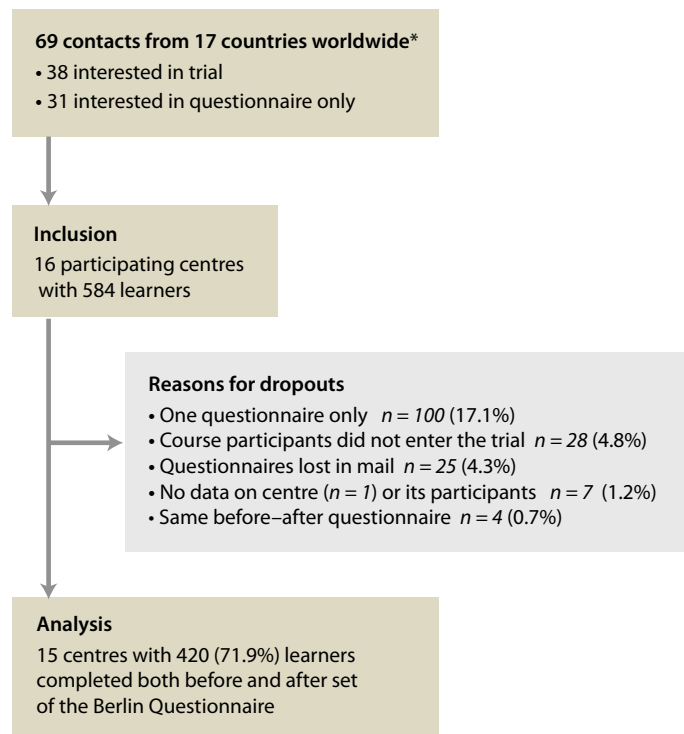
Ethical approval was granted by the King's College research ethics committee (PG) and by the McMaster University research ethics committee (HJS). At all other participating centres, waivers of review were obtained because the study involved educational research or there was no requirement for ethical approval.

Results

Selection of courses and participants. We established contacts with 69 course organizers from 17 countries; 16 of

these course organizers (United States, n = 2; Canada, n = 1; New Zealand, n = 1; and Europe, n = 12) agreed to participate. In total there were 584 learners in the participating courses. Figure 1 summarizes the recruitment of courses and learners. We excluded 1 course that did not provide centre and participant information. One hundred (17.1%) participants provided only 1 questionnaire; 34 of these single questionnaires came from a course that had a tight schedule and a high proportion of participants who had to catch air transport after the course. Another 25 questionnaires were lost in the mail. Our final sample included 15 courses with 420 learners (71.9%).

Course characteristics. Many centres used a modified "McMaster course" format:[16] there was a strong emphasis on small-group sessions (a median of 67% of total teaching time in the courses in this study), a high ratio of tutors to learners (a median of 3.5 learners per tutor) and unambiguous expectations of active participation (56% of all learners were required to participate in their course) (Table 1). The courses had between 7 and 98 participants and lasted between 1.3 and 6 days; the

69 contacts from 17 countries worldwide*
- 38 interested in trial
- 31 interested in questionnaire only

↓

Inclusion
16 participating centres
with 584 learners

↓

Reasons for dropouts
- One questionnaire only $n = 100$ (17.1%)
- Course participants did not enter the trial $n = 28$ (4.8%)
- Questionnaires lost in mail $n = 25$ (4.3%)
- No data on centre ($n = 1$) or its participants $n = 7$ (1.2%)
- Same before–after questionnaire $n = 4$ (0.7%)

↓

Analysis
15 centres with 420 (71.9%) learners
completed both before and after set
of the Berlin Questionnaire

*Reasons given by course organizers for not partipating in the trial: the timing was not right, the course situation was not conducive to participation in the study (e.g., difficult group in the previous course), the target group of the course differed from that of the study (e.g., medical students) or the course was too different (e.g., too short). The latter 2 reasons were not formal exclusion criteria for the study.

Figure 1: Flowchart of course and learner recruitment.

Table 1: Course characteristics	
Characteristic	Mean, SD*
Duration of course, days	3.9 (1.3)
Total number of teaching hours	23.8 (9.4)
Number of learners per course	54 (22.4)
Number of learners per tutor, median (range)	3.5 (0.78–71)
Number of learners per group	7.4 (1.9)
Number of topics (out of 7)	4.2 (1.3)
Separate session, topic†	
Searching, no. (%)	14 (93.3)
Statistics, no. (%)	11 (73.3)
Teaching techniques, no. (%)	4 (26.7)
Active participation‡	
Required, no. (%)	236 (56)
Encouraged, no. (%)	184 (44)
Small-group activities (% of total teaching time), median (range)	66.7 (17–82)
Fee‡	
Paid, no. (%)	270 (64.3)
Sponsored, no. (%)	51 (12.1)
Not reported, no. (%)	99 (23.6)
Attendance‡	
Mandatory, no. (%)	123 (29.3)
Voluntary, no. (%)	249 (59.3)
Not reported, no. (%)	48 (11.4)

SD = standard deviation.
* Unless indicated otherwise.
† The denominator is the number of courses (n = 15).
‡ The denominator is the number of learners (n = 420).

median teaching time was 23.8 hours. They covered an average of 4.2 (range 3–6) of 7 possible topics: therapy (n = 14; 93%); meta-analysis (n = 13; 87%); diagnosis (n = 10; 67%); study design, harm and prognosis (each n = 6; 40%); and guidelines (n = 4; 27%).

Learner characteristics. Two-thirds of the learners had completed medical school (Table 2). Few of the 420 learners had previous exposure to EBM; 42.1% (177) rated their knowledge on EBM as little and 22.4% (94) as non-existent.

Gain in knowledge. The mean baseline score across the 15 courses was 7.49 points (range 3.97–10.42 points). Set B (mean score 7.29 points, standard deviation [SD] 3.34) proved to be slightly more difficult than set A (mean score 7.71 points, SD 3.20; difference 0.42 points, 95% confidence interval [CI] 0.85–0.02). The mean post-test score was 8.89 points (range 4.77–10.91 points).

The average increase in score, adjusted for baseline score and questionnaire sequence, was 1.40 points (95% CI 0.48–2.31). This gain corresponds to an effect size of 0.44 of the baseline SD of the pooled sample of participants. There was significant variation in knowledge gain among

courses (SD 1.50, $p = 0.031$) and highly significant variation among learners within courses (SD 2.14, $p < 0.001$).

A total of 100 learners filled in only 1 questionnaire. Of these, 81 participants from 13 courses answered the pre-course questionnaire only. Their baseline score was 7.04 points (95% CI 6.28–7.80). Another 19 participants from 9 courses completed only the post-course questionnaire. They achieved a post-course score of 8.68 points (95% CI 7.34–10.02). The scores of learners who completed only the pre- or the post-course questionnaire did not differ significantly from the corresponding scores of those who completed both ($p > 0.05$ in both cases).

Factors associated with a change in knowledge and skills. We found 7 course variables that were independently associated with change in knowledge (Table 3). Greater improvements in knowledge were associated with (in order of decreasing magnitude) a requirement for active engagement in the learning process, a separate statistics session, fewer topics covered in a given time, less teaching time, smaller numbers of learners per tutor, more participants in the entire course and smaller groups. The 7 course variables together explain one-third of the total variation in knowledge gain (intraclass correlation coefficient 33%).

Predicting learning success on an individual level proved more difficult. Prior knowledge was the strongest learner determinant for change in knowledge: participants with a lower baseline score improved their score more than did those with higher baseline scores (Table 4). Of the other 11 learners' variables, 3 independently influenced change: current occupation (we observed a gradient in scores across learners with jobs in the pharmaceutical industry or medical publishing, learners involved in patient care, and administrators and epidemiologists or public health professionals), self-rating (we observed a gradient of change from learners describing themselves as having advanced knowledge, who showed the greatest improvement, to those describing themselves as having average knowledge, who showed less improvement, to those describing themselves as having little or no knowledge, who showed the least improvement) and previous experience as a tutor (those who had previous experience improved more than those who did not). Once baseline knowledge entered the model, additional learner characteristics explained 7.2% of the residual variance.

Interpretation

In this study of 15 EBM courses with diverse settings, formats and participants, we demonstrated an increase

in EBM knowledge among most participants and identified key course features that influence knowledge acquisition.

What is the significance of the mean effect size of 0.44 (corresponding to an increase of 1.40 correct answers)? This change is greater than that achieved by most single-centre courses. A recent systematic review examining 24 randomized, non-randomized and before–after studies of postgraduate EBM interventions[7] found a median effect size of 0.36 with an interquartile range of 0.31–0.42. We observed a larger effect even though the simultaneous consideration of courses (clusters) and individual learners in our model inflates the variance and therefore provides a more conservative estimate of the effect size than would a consideration of a single course. Furthermore, the outcome assessment tool we used, the Berlin Questionnaire, includes questions addressing studies of harm and prognosis that were not covered in all course programs. Thus, an outcome measure targeted specifically to course content may have shown even larger effects. However, some courses had a substantially greater impact than others. Course differences contributed one-third to the total variability in knowledge gain between learners.

Another way of thinking about the change of 1.40 is in terms of the differences found in the scores of different categories of clinicians. We previously demonstrated that clinicians without formal EBM training have a mean score of 4.2, those enrolling in EBM courses (presumably because they have a particular interest in the subject and have done some prior reading) have a mean score of 6.3 and experts have a mean score of 11.9.[11] Thus, a mean change of 1.40 would take an individual without prior knowledge of EBM approximately one-fifth of the way to achieving the score that one might expect from an expert.

This study has a number of strengths. Because of the diversity in course format, content, participants and geographic location, this study provides a picture of the impact of a broad range of EBM courses. We used the same well-validated instrument, the Berlin Questionnaire, for outcome assessment in all courses, which allowed a valid comparison of the learning effect across courses. The study design, with learners randomly assigned to complete one or the other version of the Berlin Questionnaire first, minimized possible bias. Our calculation of effect size allowed comparison with the impact of courses previously reported in the literature.[7] Most importantly, we collected detailed information about each course, which allowed us to construct a sophisticated statistical model

that explained the variability in knowledge gain across courses. The association of knowledge gain with some variables was very strong: required active engagement of learners was associated with a difference of 2.8 items answered correctly. This is a very large effect in a questionnaire with 15 items (Table 3). The ability to explain the variability in outcomes to this extent is unusual and

Table 2: Learner characteristics (n = 420)	
Characteristic	No. (%)*
Age, yr, mean (SD)	34.8 (10.3)
20–25	57 (13.6)
26–30	80 (19.1)
31–35	75 (17.8)
36–40	95 (22.6)
> 40	110 (26.2)
Missing	3 (0.7)
Female	215 (51.2)
Baseline score (across courses), mean (range)	7.49 (3.97–10.42)
Sequence of questionnaires in baseline assessment	
Set A first	214 (50.5)
Set B first	206 (49.5)
Education	
Attended medical school	278 (66.2)
Completed other health professional training	42 (10.0)
Current student	77 (18.3)
Not reported	23 (5.5)
Previous experience in EBM (multiple selections possible)	
Taken a session (< 1 d)	79 (18.8)
Taken a workshop	62 (14.8)
Read a book	84 (19)
Taught as a tutor	8 (1.9)
Self-rated knowledge of EBM	
None	94 (22.4)
Little	177 (42.1)
Average	118 (28.1)
Advanced	22 (5.2)
Not reported	9 (2.1)
Occupation	
Direct patient care provider	251 (60.0)
Academic (public health or epidemiology)	25 (5.9)
Administrator	8 (1.9)
Medical publishing or pharmaceutical industry personnel	15 (3.6)
Student	77 (18.3)
Not reported	44 (10.5)
Current position	
Student	77 (18.3)
Other health professional	30 (3.3)
Intern or resident	84 (20)
Fellow	56 (13.3)
Consultant	77 (18.3)
Health professional in private practice	27 (6.4)
Other	69 (16.4)
SD = standard deviation	

Table 3: Course characteristics associated with a change in score

Characteristic	Change in score, points (95% CI)	p value
Total number of teaching hours*	−0.25 (−0.32 to −0.19)	< 0.001
Number of learners per tutor*	−0.07 (−0.09 to −0.04)	< 0.001
Number of learners per course*	0.03 (0.01 to 0.05)	0.001
Number of learners per group*	−0.25 (−0.44 to −0.06)	0.010
Program		
Number of topics*	−1.09 (−1.52 to −0.66)	< 0.001
Separate statistics session†		
Yes	1.9 (1.09 to 2.68)	< 0.001
No	Reference	
Learner engagement†		
Active participation required	2.8 (1.46 to 4.15)	< 0.001
Active participation encouraged	Reference	

CI = confidence interval.
* Change in score per unit.
† Change in score compared with the reference category.
Note: Regression analysis was conducted on change in score. A positive value indicates an increase in knowledge compared with the reference category and a negative value indicates a decrease in knowledge compared with the reference category. For example, assuming that all other parameters stayed the same, learners in courses that required active participation improved their score by 2.8 more points than learners in courses that only encouraged active participation. As another example, assuming that all other parameters stayed the same, learners scored 1.09 point less for every additional topic that was covered in a course.

provides planners of subsequent courses and workshops, one of whose goals is to improve knowledge, with guidance in structuring their program.

Many of the course items that proved to be a significant determinant of knowledge gain in our analysis are consistent with theories of problem-based learning.[17] An exploratory matching of the characteristics of problem-based learning with similar terms from a synthesis of 302 meta-analyses on educational and psychological interventions[18] found small to moderate effect sizes for the following characteristics: "individualized learning" (effect size 0.23), "cooperative learning" (effect size 0.54), "small group process" (effect size 0.31), "using problems" (effect size 0.2) and "instructions for problem-solving provided" (effect size 0.54).[19]

One limitation of our study is the before–after design, which produces weaker inferences than a randomized trial. Thus, one might question the extent to which courses were actually responsible for the learners' improved questionnaire performance after they finished the course (e.g., pre-course practice with the alternate form of the questionnaire could have been responsible for the change instead). If another factor were responsible for the change, however, it is extremely implausible that course characteristics would have been so powerful in explaining variability in knowledge gain across courses.

Another potential limitation is the risk of selection bias. Only a minority of course organizers expressed interest in participating in our study, and some felt that the

questionnaire was inappropriate for their course and withdrew after they reviewed the questionnaire. Non-participating courses may have been less successful than participating courses; thus, selection bias provides another potential explanation for the large size of effect we observed.

Furthermore, 100 learners filled in only 1 part of the questionnaire (pre- or post-course). Scores from individuals who completed only one part of the questionnaire were, however, similar to the scores of those who completed both pre- and post- versions. Although the group that completed only the pre-course questionnaire scored somewhat lower overall, the confidence interval around this score included the entire confidence interval around the mean score of the 420 participants who completed both questionnaire administrations.

The generalizability of our results is limited in that we studied only formal courses involving direct contact between educators and learners. Thus, our results do not address either courses embedded in the clinical setting (known as integrated EBM courses)[6,9] or e-learning EBM courses, which represent an emerging format.[20–22] Nevertheless, the course formats we studied are likely to remain both prevalent and popular. Finally, we did not assess long-term outcomes, the use of the new knowledge in clinical practice, or attitude or behaviour changes among learners. Recent evidence from a systematic review and 3 randomized trials shows that the translation from knowledge gain to change in behaviour is complex.[6,23–25]

Our results also provide information concerning outcome measurement for EBM courses and workshops. A number of reviews have highlighted the shortage of validated instruments for assessing the outcomes of EBM teaching.[8,26,27] A recent systematic review of assessment tools noted the attractive psychometric properties of the Berlin Questionnaire.[13] Our results provide further evidence that the Berlin Questionnaire is responsive across different settings, participants and course formats and can differentiate between individuals and courses.

Although incorporating some of the factors associated with higher scores (smaller group size, higher ratio of tutors to learners and larger overall course size) would require more resources, 1 factor (less teaching time) would require fewer resources, and others, including the 3 most powerful predictors (required active participation of the learners, a separate statistics session and fewer topics) would require no increase in resources. In terms of the

Table 4: Learner characteristics associated with a change in score		
Characteristic	Change in score, points (95% CI)	*p* value
Baseline score*	−0.66 (−0.74 to −0.58)	< 0.001
Sequence of Berlin Questionnaire†		
Set A first	−0.43 (−0.85 to −0.02)	0.04
Set B first	Reference	
Occupation†		0.001
Direct patient care provider	Reference	
Academic (epidemiology, public health)	−0.34 (−1.21 to 0.53)	
Administrator	−2.88 (−4.4 to −1.39)	
Medical publishing or pharmaceutical industry personnel	1.04 (−0.07 to 2.14)	
Student or not reported	−0.28 (−0.89 to 0.33)	
Previous experience as a tutor†		0.008
Yes	2.2 (0.58 to 3.84)	
No	Reference	
Self-rated knowledge†		0.020
None	0.08 (−0.57 to 0.73)	
Little	Reference	
Average	0.56 (0.04 to 1.09)	
Advanced	1.09 (−0.04 to 2.14)	
Not reported	−1.61 (−3.14 to −0.08)	

CI = confidence interval.
* Change in score per unit.
† Change in score compared with the reference category.
Note: Regression analysis was conducted on change in knowledge score. A positive value indicates an increase in knowledge compared with the reference category and a negative value indicates a decrease in knowledge compared with the reference category. For example, assuming that all other parameters stayed the same, learners with previous experience as a tutor improved their knowledge by 2.2 more points than learners who did not have this experience.

importance of the overall effect, there may be other benefits (enthusiasm for evidence-based practice, stimulation of subsequent learning and ultimate improvement in practice) not measured here that would further justify the investment in resources.

In conclusion, our results provide guidance for EBM course organizers interested in maximizing participant gains in knowledge. If it is feasible, they should use a small-group structure with a high tutor-to-learner ratio. Course organizers may welcome the news that long courses are not necessary; indeed, shorter courses may be more effective. Organizers should focus on a small number of topics, have special sessions devoted to statistical issues and, most importantly, insist on active participation from the learners.

REFERENCES

1. Guyatt GH. Evidence-based medicine [editorial]. *ACP J Club* 1991;114(2):A16.
2. Norman GR, Shannon SI. Effectiveness of instruction in critical appraisal (evidence-based medicine) skills: a critical appraisal. *CMAJ* 1998;158(2):177–181.
3. Parkes J, Hyde C, Deeks J, Milne R. Teaching critical appraisal skills in health care settings. *Cochrane Database Syst Rev* 2001;(3):CD001270.
4. Taylor R, Reeves B, Ewings P, Binns S, Keast J, Mears R. A systematic review of the effectiveness of critical appraisal skills training for clinicians. *Med Educ* 2000;34(2):120–125.
5. Ebbert JO, Montori VM, Schultz HJ. The journal club in postgraduate medical education: a systematic review. *Med Teach* 2001;23(5):455–461.
6. Coomarasamy A, Khan KS. What is the evidence that postgraduate teaching in evidence based medicine changes anything? A systematic review. *BMJ* 2004;329(7473):1017.
7. Flores-Mateo G, Argimon JM. Evidence based practice in postgraduate healthcare education: a systematic review. *BMC Health Serv Res* 2007;7:119.
8. Straus SE, Green ML, Bell DS, Badgett R, Davis D, Gerrity M, et al. Evaluating the teaching of evidence based medicine: conceptual framework. *BMJ* 2004;329(7473):1029–1032.
9. Khan KS, Coomarasamy A. A hierarchy of effective teaching and learning to acquire competence in evidenced-based medicine. *BMC Med Educ* 2006;6:59.
10. Guyatt GH, Meade MO, Jaeschke RZ, Cook DJ, Haynes RB. Practitioners of evidence based care. Not all clinicians need to appraise evidence from scratch but all need some skills. *BMJ* 2000;8;320(7240):954–955.
11. Fritsche L, Greenhalgh T, Falck-Ytter Y, Neumayer HH, Kunz R. Do short courses in evidence based medicine improve knowledge and skills? Validation of Berlin questionnaire and before and after study of courses in evidence based medicine. *BMJ.* 2002;7(325):1338–1341.
12. Anderson LW, Krathwohl DR, eds. *A taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of educational objectives.* New York: Longman; 2001.
13. Shaneyfelt T, Baum KD, Bell D, Feldstein D, Houston TK, Kaatz S, et al. Instruments for evaluating education in evidence-based practice: a systematic review. *JAMA* 2006;296(9):1116–1127.
14. Cohen J. *Statistical power analysis for the behavioral sciences.* 2nd ed. New Jersey: Lawrence Erlbaum; 1988.
15. Coe R. It's the effect size, stupid. What effect size is and why it is important. British Educational Research Association Annual Conference; Exeter. Sept 2002.
16. How To Teach Evidence-Based Clinical Practice Workshop. [Website of McMaster University]. Available: http://ebm.mcmaster.ca/ (accessed 28 Sept 2009).
17. Maudsley G. Do we all mean the same thing by "problem-based learning"? A review of the concepts and a formulation of the ground rules. *Acad Med* 1999;74(2):178–185.
18. Lipsey M, Wilson DB. The efficacy of psychological, educational, and behavioural treatment. Confirmation from meta-analysis. *Am Psychol* 1993;48(12):1181–1209.

19. Norman GR, Schmidt HG. Effectiveness of problem-based learning curricula: theory, practice and paper darts. *Med Educ* 2000;34(9):721–728.

20. Coppus SFPJ, Emparanza JI, Hadley J, Kulier R, Weinbrenner S, Arvanitis TN, et al. A clinically integrated curriculum in evidence-based medicine for just-in-time learning through on-the-job training: the EU-EBM project. *BMC Med Educ* 2007;27;7(1):46.

21. Kunz R, Avanitis T, Burls A, Capello-Lopez J, Coppus S, Decsi T, et al. EU EBM Unity Project: Implementing systematic reviews into patient care through an integrated European curriculum in evidence-based medicine [abstract P073]. 14th Cochrane Colloquium, Oct 22–26, 2006; Dublin.

22. Schilling K, Wiecha J, Polineni D, Khalil S. An interactive web-based curriculum on evidence-based medicine: design and effectiveness. *Fam Med* 2006;38(2):126–132.

23. Hugenholtz NIR, Schaafsma FG, Nieuwenhuijsen K, van Dijk FJ. Effect of an EBM course in combination with case method learning sessions: an RCT on professional performance, job satisfaction, and self-efficacy of occupational physicians. *Int Arch Occup Environ Health* 2008;82(1):107–115.

24. Shuval K, Berkovits E, Netzer D, Hekselman I, Linn S, Brezis M, et al. Evaluating the impact of an evidence-based medicine educational intervention on primary care doctors' attitudes, knowledge and clinical behaviour: a controlled trial and before and after study. *J Eval Clin Pract* 2007;13(4):581–598.

25. Kim S, Willett LR, Murphy DJ, O'Rourke K, Sharma R, Shea JA. Impact of an evidence-based medicine curriculum on resident use of electronic resources: a randomized controlled study. *J Gen Intern Med* 2008;23(11):1804–1808.

26. Green ML. Evaluating evidence-based practice performance. *Evid Based Med* 2006;11(4):99–101.

27. Hatala R, Guyatt G. Evaluating the teaching of evidence-based medicine. *JAMA* 2002;288(9):1110–1112.