Reviewer’s report

Title: Teaching of Evidence-Based Medicine to medical students in Mexico: A randomized controlled trial.

Version: 1 Date: 24 August 2012

Reviewer: Nai Ming Lai

Reviewer’s report:

Researchers in evidence based medicine (EBM) education today face a double-challenge: the expectations on them to demonstrate the effectiveness of EBM training that they champion in accordance with the standards that they preach, and a common difficulty for medical education researchers in demonstrating “educationally important” learning gains, i.e. learning gains that are likely to impact on the student’s future practice. The authors of this paper addressed one of these challenges by designing a randomised controlled trial comparing a block EBM curriculum versus a controlled, non-EBM curriculum. I believe the main value of this study is that it presents a higher-level evidence in confirming what previous studies have consistently shown that EBM training improves knowledge and confidence at the completion of the training. I am less convinced, however, that the degree of learning gain demonstrated in the study was educationally important as stated by the authors, as the significant loss of knowledge in the students who received EBM training six months to one year earlier left one to wonder whether there would be any material amount of knowledge retained that could influence clinical decision-making in the future. In fact, the over-optimistic and sometimes misleading interpretation of their findings, which is found consistently in the abstract and discussion, is to me a major criticism of the paper. I would also suggest that the authors tighten up their discussion to focus on what is directly substantiated by the findings and not elaborate too much on general issues in teaching and application of EBM.

Following are specific comments pertaining to different parts of the paper:

• Major Compulsory Revisions (which the author must respond to before a decision on publication can be reached)

Abstract,

Background: “The objective was to assess the learning of EBM by undergraduate medical students, and retention one year later.”

Results

“Confidence in critical appraisal skills was higher in the intervention groups and was maintained one year later.”

Conclusions

“The educational intervention has a positive effect that may persist one year later.”
In my opinion, there is misleading use of wording in the sentences quoted above. I think words such as “retention”, “maintained” and “persist” should not be used, as retention was only implied, and not directly assessed in the study, because the study evaluated EBM knowledge and confidence among different groups of students, and not the same group over time. I suggest that the authors simply describe the results without making any assumption, for example, “confidence in critical appraisal skills was higher in the intervention group and the group of students who received EBM instruction in the previous year”.

Methods, study design

“The outcomes were measured at the end of the first semester of the academic year, after the EBM course ended, comparing the fifth year half-group that took the course with the fifth year half without the course, the fourth year students without the course, and the sixth year students that had the course one year previously.”

If there were no changes in the curriculum in the authors’ institution, I figure that there would be two distinct groups of sixth-year students in relation to EBM exposure: one who received EBM in the first half of their fifth year and one in the second half. It seems more appropriate to separate these two groups to more usefully evaluate the amount of knowledge and confidence in students who were six months and one year post-intervention.

It was also stated that students in year four were exposed to some elements of EBM (statistics, medical informatics, research methodology and epidemiology) and the authors distinguished this from their EBM course but made no further description of the difference between the delivery of these subjects in year four versus year five. This is important in light of the authors’ claims in their discussion that EBM intervention in year five appeared more effective than teaching of similar subjects such as clinical epidemiology and research methodology in year four, which should lead to an exploration on possible reasons for the difference. A description of the methods of delivery in the subjects of epidemiology and research methodology in year four should be included here.

“The assignment of students to groups was done by the medical school using Microsoft Excel for block randomization (two blocks with 48 students each in this case, for the 5th year class)”

I am unclear how block randomization was actually achieved in this study. It appears to me that the methods of randomization as described by the authors, using “two blocks each of 48 participants” was an unlikely way of randomization. It seems more appropriate in this case to randomize using blocks of two (instead of blocks of 48). Otherwise, using two blocks of 48 would be akin to cluster randomization, with different implications on analysis and reporting of results. The authors may wish to clarify this and describe their methods of randomization more accurately.
Methods, Statistics, data analysis and ethical aspects.

“The smallest meaningful difference (SMD) for the EBM Test instrument was estimated to be a difference of 10 questions between the intervention group and the control group.”

There should be more description here on how the “smallest meaningful difference” was determined. In medical education research, it is a challenge to define a meaningful effect size, mainly because there is a lack of literature on how great a learning gain would translate into an impact in students' future practice. As such, any threshold would seem arbitrary unless it is supported by a process which involves careful consideration from a panel of teachers who were closely involved in student teaching and had reasonable expectation on student performance. Was this the case? If not, this should be acknowledged as a limitation of the study.

Results, EBM summative test-Knowledge score

“The scores were similar between the non-EBM groups, the M4 group had a mean score of 30.6 +/- 5.6, and the M5 non-EBM group had 32.6 +/-6.6 (p=0.18). The M5 EBM group had a test score of 58.5 +/-7.9, higher than the M4 and M5 non-EBM groups.”

Converting to percentages, there was a 26% difference between the intervention group and the control group, which convinced me that the EBM training produced a substantial learning gain immediately post-intervention. However, year-six students only gained less than 9% compared to the pre-intervention group, and that was measured midway through their year-six training. It was worrying whether there would be any material learning gain at the end of their six-year course. I would like to see the authors putting forward their thoughts on this in discussion.

Discussion

Overall, I think while the section is very comprehensive, it may flow better if the authors focus on discussing the findings of this study and not elaborate too much on general issues in teaching and application of EBM. The authors may wish to consider removing paragraphs 1,3 and the last paragraph all together and trimming the other paragraphs to omit excessive commentaries on EBM in general.

Paragraph 2, “The educational effect of the course appears to decrease with time, but lasts at least a year after the educational intervention compared to the control groups.”

This statement is misleading for two reasons. First, from my understanding, not all year-six students received EBM training a year previously, as some received the training six months previously (assuming the curriculum remained unchanged). Next, “lasting at least a year” was an over-statement as the performance of year-six students did not suggest that they had kept a substantial
portion of the knowledge learnt in year five. I think the second half of the
sentence should be replaced by something more neutral, such as “…and year-six
students appeared to still retain some EBM knowledge learnt in year five”.

Also, what did the findings mean in relation to the expected outcomes of EBM
training in the authors’ institutions?

Paragraph 9, “The effect size immediately after the course in the critical appraisal
skills score was higher than 1.0, which reveals a large effect that is maintained
one year later (Table 6).”

Paragraph 13, “These increases are educationally significant, with effect sizes
ranging from 0.88 to 4.6, which reveal an important effect of the educational
intervention.”

I think these statements should be omitted as they are difficult to interpret and
contain misleading statement on knowledge retention and educational
significance, as commented earlier.

Paragraph 15, “This apparent decrease in knowledge needs to be moderated by
the fact that our one year post-course measure was done in a different group of
students, not the randomized 5th year class, so it may not represent a true
measure of knowledge decay. On the other hand, the knowledge and attitude
scores were significantly higher than the pre-EBM course scores in M4 and M5
students, which suggest that the course had a long-term effect in this setting.”

In the statements, the authors are not consistent in their arguments. They chose
to attribute the positive learning gain to their EBM course and lower scores in
year-six students to possible differences in student characteristics. The truth is,
one would not know whether the differences in scores among the groups
assessed were due to the effects of EBM training or to the groups themselves, as
there were no follow-up assessments performed in this study. Rather than
making postulation either way, I would suggest that the authors acknowledge the
uncertainties and include them as a limitation of the study.

• Minor Essential Revisions (such as missing labels on figures, or the wrong use
of a term, which the author can be trusted to correct)

Abstract, RESULTS: “M5 EBM=48, M5 non-EBM=47, M4=87, and M6=107.”
Expand each term in full when used for the first time.

Methods, outcomes and instrumentation

“The knowledge portion of the questionnaire includes six multiple-choice
questions”
Should it be “multiple true-false”?

“incorrect responses are negatively scored (-1) to try to persuade the students
from guessing”
Should it be “try to prevent” or “dissuade” instead of “persuade”?

Results
In general, the report contains excessive technical details and the authors should omit this in favour of a more descriptive presentation. For instance, statistical tests employed need not be mentioned in the results.

Table 6: comparison between cohorts
It would be good to also include “M6 vs M5 EBM” and “M6 student who received EBM training in the first semester versus second semester”.

• Discretionary Revisions (which are recommendations for improvement but which the author can choose to ignore)

Abstract, methods: “EBM attitudes, knowledge and self-reported skills were measured using Taylor’s questionnaire and a summative objective test.”
The authors may wish to add the following description after the term “summative objective test”: “, which comprised of a 100-item multiple-choice test”.

Methods, intervention
“The course faculty were six professors with training in EBM teaching”

What was the teachers’ background? Were they involved in clinical service, and were they involved in the training of year four and year six students as well?

I think how the EBM training programme was clinically-linked should be described in greater detail in the methods, including specific methods of delivery for each domain. A convenient way of including this is by adding an extra column in Table 1 on how teaching of each subject was delivered, i.e. whether via lecture, small-group tutorial, bedside session or journal club etc.

Results, Instructional sensitivity and reliability.

I don’t think these needs to be described in full in the results. These technical details can be appended, and only a brief mention is sufficient at the beginning of the results, with a link to the appendix.

Discussion
Paragraph 16, “As with any implementation of a new course in a medical school, there was an intense interest from the course director and instructors to develop and implement as effective an educational intervention as possible, so there could be a tendency for confirmation bias. This bias would be expected in this type of educational intervention experimental study, where it is impossible to blind either the instructors or the students to the educational intervention.”

One other possible reason that could lead to bias would be the Hawthorn effect, in which students in the intervention group were aware that they were being evaluated on the effectiveness of their current training programme, unlike M6
students who were not "watched" as closely when the EBM training was administered in the previous year. The authors may wish to incorporate this in their discussion.

Paragraph 17, “…it provides evidence that the course can impact knowledge, attitudes, critical appraisal confidence, and self-reported skills and behaviors about EBM and its related concepts, in a statistically and educationally significant manner.”

As in my previous comments, I think describing the degree of learning gain from year five to year six as “educationally significant” is an overstatement. A more cautionary note on possible degree of learning gain that would remain at the end of the students training and in their clinical practice would be more appropriate here.

**Level of interest:** An article of importance in its field

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests