ORIGINAL ARTICLE

Competence in evidence-based medicine of senior medical students following a clinically integrated training programme

NM Lai 賴乃鳴
CL Teng 鄧昌鍊

Objective To assess the impact of a structured, clinically integrated evidence-based undergraduate medicine training programme using a validated tool.

Design Before and after study with no control group.

Setting A medical school in Malaysia with an affiliated district clinical training hospital.

Participants Seventy-two medical students in their final 6 months of training (senior clerkship) encountered between March and August 2006.

Intervention Our educational intervention included two plenary lectures at the beginning of the clerkship, small-group bedside question-generating sessions, and a journal club in the paediatric posting.

Main outcome measures Our primary outcome was evidence-based medicine knowledge, measured using the adapted Fresno test (score range, 0-212) administered before and after the intervention. We evaluated the performance of the whole cohort, as well as the scores of different subgroups that received separate small-group interventions in their paediatric posting. We also measured the correlation between the students' evidence-based medicine test scores and overall academic performances in the senior clerkship.

Results Fifty-five paired scripts were analysed. Evidence-based medicine knowledge improved significantly post-intervention (means: pre-test, 84 [standard deviation, 24]; post-test, 122 [22]; P<0.001). Post-test scores were significantly correlated with overall senior clerkship performance (r=0.329, P=0.014). Lower post-test scores were observed in subgroups that received their small-group training earlier as opposed to later in the clerkship.

Conclusions Clinically integrated undergraduate evidence-based medicine training produced an educationally important improvement in evidence-based medicine knowledge. Student performance in the adapted Fresno test to some extent reflected their overall academic performance in the senior clerkship. Loss of evidence-based medicine knowledge, which might have occurred soon after small-group training, is a concern that warrants future assessment.

Key words
Education, medical, undergraduate; Evidence-based medicine

Hong Kong Med J 2009;15:332-8

Department of Paediatrics, School of Medicine and Health Sciences, Monash University Malaysia, JKR 1235, Bukit Azah, 80100, Johor Bahru, Johor Darul Takzim, Malaysia
NM Lai, MRCP, MRCPCH
Department of Family Medicine, International Medical University, Clinical School Seremban, Jalan Rasah, 70100, Seremban, Negeri Sembilan, Darul Khusus, Malaysia
CL Teng, FRACGP, Mmed

Correspondence to: Dr NM Lai
E-mail: lainm123@yahoo.co.uk

Introduction

Nearly two decades after its introduction, evidence-based medicine (EBM) has become a core training domain in postgraduate and undergraduate medical curricula.[1] However, to date, EBM training especially at undergraduate level has been driven mainly by enthusiasm and belief, rather than clear evidence of benefits.[2] One possible reason is the difficulty in establishing the role of EBM among myriad of other factors that change clinical decision-making behaviour and lead to improved patient care.[3] The other reason might be that in its early days, there was a lack of a validated tool to assess a comprehensive range of knowledge and skills in EBM, as most reports focused either on specific domains like critical appraisal, or relied on subjective measures like confidence and self-perception.[4]

Although strong empirical evidence is still lacking, it is widely agreed that EBM training should be integrated into actual clinical practice, rather than solely as classroom

teaching.[5] A study conducted in a community hospital suggests that teaching residents EBM on-the-ground leads to more clinical decisions in line with high-level evidence, lending support to a clinically integrated approach in teaching the subject.[6] On the other hand, limited evidence shows that undergraduates appear to enjoy a greater learning gain in specific EBM-related skills compared to practising clinicians. This provides support for initiating EBM training at the undergraduate level.[2]

Recently, EBM assessment tools that are reported to be valid and reliable have emerged.[7,8] The Fresno test, developed by Ramos et al in 2003,[8] purports to assess a comprehensive range of knowledge in EBM. The test comprises seven open-ended, four short-answer, and one multiple-choice questions. There are published reports describing its administration to medical residents[9] and occupational therapists.[10]

At the International Medical University, Malaysia, a basic level of competence in EBM is expected of all medical graduates. To achieve this, students are progressively exposed to EBM in their undergraduate training from the pre-clinical phase. Such exposure consists of lectures, problem-based learning, research projects, and short EBM summaries. In the final 6 months of training—the senior clerkship, these activities culminate in a structured, clinically integrated training programme encompassing all major tenets of EBM. In the senior clerkship, students consolidate their clinical experience predominantly through independent bedside learning, akin to a shadow housemanship, supplemented by more focused and in-depth learning via the development of detailed structured case reports, which form their portfolios.[11] The EBM reports with critical appraisal of patient-relevant journal articles are incorporated into the portfolios, which serve as examination documents in each student's final oral assessment.

To assess the impact of our structured, clinically integrated EBM training programme, we evaluated a cohort of medical students in their senior clerkship. We measured their EBM knowledge before and after EBM training (about 5 months apart), using an adapted version of the Fresno test of competence in EBM.[8] We hypothesised that our training programme would produce a significant, educationally important gain in EBM knowledge, as measured by the adapted Fresno test. As a secondary objective, we assessed the correlation between student competence in EBM and overall performance in the senior clerkship. We postulated that competence in EBM would show a significant correlation with overall performance in the senior clerkship. As our EBM training was predominantly conducted in small groups, we also compared the performance of different subgroups that received the small-group training at different periods, so as to assess retention of the EBM knowledge gained.

透過一項結合臨床循證訓練計劃，評估醫科學生對循證醫學知識的程度

目的 透過已被確認的工具，評估一項結構性結合臨床循證的醫科學生訓練計劃。

設計 不設控制群的教育干預前後結果研究。

安排 一所馬來西亞附屬地區臨床訓練醫院的醫科大學。

參與者 2006年3月至8月期間，72位正接受最後6個月訓練的醫科學生（高階實習）。

干預 教育干預包括開始實習時的兩次特別演講、臨床研究討論小組，以及從病例找出相關文獻討論。

主要結果測量 研究的主要目的是透過改良版弗雷斯諾測試（Fresno test）〔分數範圍，0-212〕，在教育干預進行前後測試學生對循證醫學知識的程度。除對整隊表現作出評估，也把學生分成小組，在兒科部門參與訓練。研究也評估循證醫學知識水平與於高階實習時整體學術表現的關係。

結果 共分析55個被配對的訓練前後測試答案。循證醫學的知識水平於教育干預後有顯著改善（干預前中位數：84〔標準差，24〕；干預後中位數：122〔標準差，22〕；P<0.001）。干預後的得分與高階實習時整體學術表現有顯著關係（r=0.329，P=0.014）。相對實習時較遲接受小組訓練的，較早受訓練的學生的教育干預後得分則較低。

結論 結合臨床循證的醫科學生訓練計劃有助增進循證醫學的知識。學生於弗雷斯諾測試的表現某程度也反映高階實習的整體學術表現。不過，學生對循證醫學的知識也許在小組訓練後便很快忘記，須作進一步評估。

Methods

Study design

A before and after study design was used. Our subjects were a cohort of final-year medical students (n=72) posted to a peripheral district hospital (affiliated to our clinical school), during their final 6 months of training (senior clerkship), from March to August 2006. All students were invited to participate in the study. They received a study information leaflet and briefing from the first author before the pre-test. Written consent was obtained from students who agreed to participate. The signing of consent was overseen by an independent administrator, without any investigator being present. Whether or not they participated in the study, all students received the educational intervention, which was a standard programme in the university curriculum.

Educational intervention

The students received a structured, clinically

integrated EBM training programme first developed in May 2003, in accordance with the university curricular requirements in the senior clerkship, and used principles delineated by Sackett et al.[12]

There were two major parts in the programme. Part I consisted of two introductory plenary lectures in the first week of clerkship. These provided an overview of EBM, and covered searching for evidence and critical appraisal using two case scenarios, one on therapy and one on prognosis. Part II entailed intensive, small-group training, during which each subgroup of 13 to 15 students rotated to the 4-week paediatric posting. This comprised six bedside clinical sessions with exercises on formulating clinical questions (averaging 2 hours per session), and three journal club sessions (averaging 2 hours per session) where students presented the clinical articles that they acquired following the clinical queries generated from the bedside discussion. Each journal club presentation included a clinical question, search strategies, a critical appraisal of the article(s) retrieved, followed by discussion on how the findings apply to patients. We used a modified critical appraisal worksheet developed by the Centre for Evidence Based Medicine, Toronto.[13] Where applicable, the sessions also included role-playing of physician-patient communication. On average, four students presented in each session. The journal club presentations were adapted by the students as their EBM portfolio reports. Each student was required to develop at least five EBM reports in different disciplines in the senior clerkship.

Although the small-group training was conducted at different periods for different student groups, we expected all students to acquire and maintain their EBM knowledge throughout the clerkship, whatever the timing of their small-group training. This was because they were required to apply EBM principles to all disciplines and produce corresponding reports in their portfolios. In the introductory plenary lectures, all students were directed on the use of EBM learning resources, including critical appraisal tools with comprehensive explanatory notes, made available in the virtual drive ("I" Drive) of the clinical school, and that the drive could be accessed anytime by students and staff. We expected all students to have a similar level of competence in EBM at the end of the clerkship, regardless of when they received their small-group training.

The training programme was jointly developed by both authors who were tasked with developing the EBM curriculum in the clinical phase student training. All training sessions in the senior clerkship, including the introductory lectures and small-group sessions, were facilitated by the first author, who was then the resident coordinator and assigned teacher of EBM in the senior clerkship programme. Resident supervisors from other disciplines also received prior instructions on EBM.

Outcome measurement

The adapted Fresno test of competence in evidence-based medicine

The primary outcome, the change in EBM knowledge was measured using an adapted version of the Fresno test of competence in EBM (score, 0-212).[8] We made the following adaptations:

1. Both clinical scenarios on question 1 were changed to reflect common clinical conditions in Malaysia, with corresponding changes made in question 3.
2. The order of questions from 5 to 7 was reversed to reflect the format in which critical appraisal was taught in our training programme, namely assessing validity, clinical importance, and finally, applicability.
3. The order of questions was re-arranged so all open-ended (questions 1-7), short-answer (questions 8-11), and multiple-choice questions (question 12) were grouped together.
4. For questions 10 and 11, while all the statistical concepts tested remained unchanged, clinical scenarios were simplified to the level appropriate for undergraduates.
5. For question 12, while the concept tested remained unchanged (ie understanding of 95% confidence interval [CI]), the format of question was changed from short-answer to multiple-choice, with two acceptable answers out of the seven choices offered.
6. The wording of some questions was simplified to improve readability for our undergraduates.
7. Grading rubrics were revised according to the above changes.

The adaptation was undertaken by the first author with editorial input from the second author. After the adaptation, a pilot test was conducted on 12 final-year medical students in December 2005, which resulted in further changes to the wording of some questions.

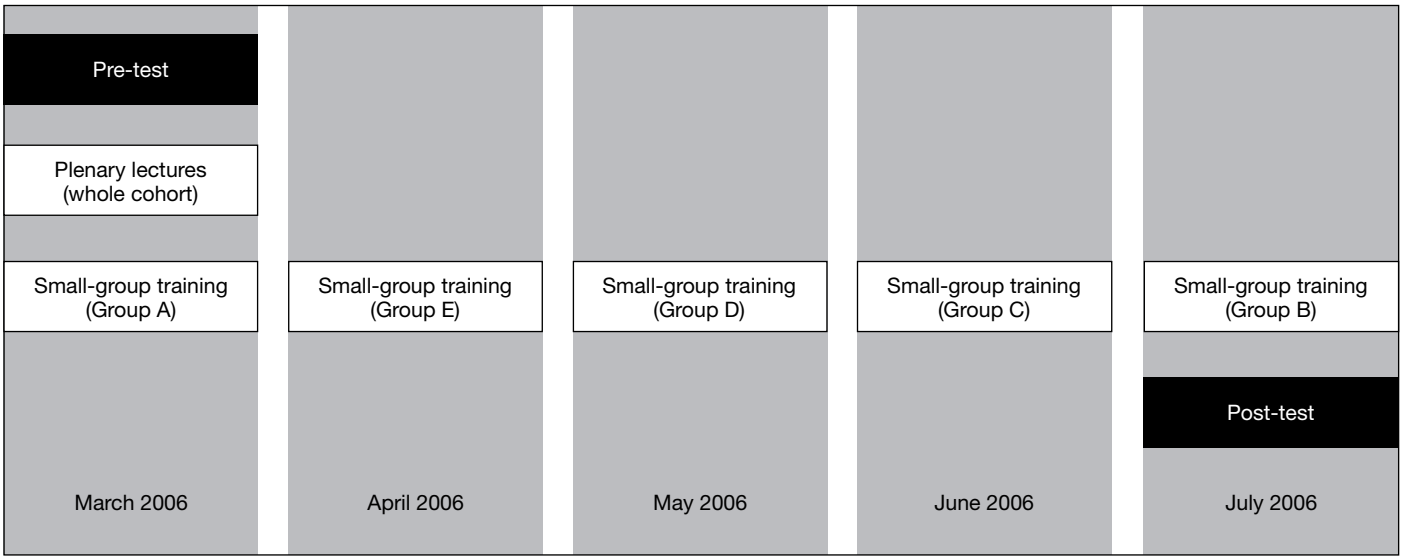
We also determined a 'pass mark' appropriate for undergraduate students, first independently using the principles of the Angoff method of standard setting,[14] and then through discussions leading to a consensus, aided by the results of the pilot test. The final 'pass mark' was set at 106 (50%).

The pre-test was conducted in the first week of March 2006, and the post-test in the last week of July 2006. The conduct of our study is graphically illustrated in Figure 1.

Scoring

Both authors independently scored all test scripts. We analysed inter-rater difference in the final scores and

FIG 1. Illustration of the educational interventions and assessment within the senior clerkship period



obtained a mean difference and standard deviation (SD). If the inter-rater difference was more than two SDs apart, we discussed the scripts concerned question by question, leading to a consensus score. We averaged the final scores of all other scripts. We took the following measures to ensure blinding of the raters to the status of the completed scripts:

1. Prior to the study, an independent administrator pre-coded all empty scripts with a 3-digit serial number (000 to 150). The scripts were hand-reshuffled so the serial numbers were no longer in order. After reshuffling, the first 75 scripts were assigned as "pre" scripts, and next 75 "post" scripts. Their serial numbers and corresponding "pre" or "post" status were only released to the investigators after scoring was completed.
2. Identical test scripts were used for both pre- and post-tests.
3. No investigator was present during the tests.
4. Scoring was commenced only after completion of the post-test.
5. Prior to scoring, completed scripts from both tests were combined and hand-reshuffled again by the independent administrator.

Pairing of the scripts

To enable pairing while preserving script anonymity for the investigators, students recorded their identification numbers beside their script serial numbers on a recording sheet, during both rounds of the test. An independent administrator performed the pairing by matching the pre- and post-serial numbers using the student identification numbers. Student's grades in three major assessments in the senior clerkship were also collated using the student identification numbers, which were then deleted and released to the authors for analysis.

Overall performance in the senior clerkship

The students' overall performance in the senior clerkship was determined from three major assessments: community and family case study (0-100%), in-course continuous assessment (6-scaled global rating), and portfolio-based oral assessment (4-scaled global rating). All global ratings were converted to percentages, and the average percentage of these three assessments was taken as the overall performance.

Sample size estimation

We referred to the validation study to estimate our sample size.[8] We assumed that senior medical students had slightly lower levels of EBM knowledge than practising health care staff who were new to EBM. We therefore estimated a pre-test mean score of our sample as 40%, 5% lower than that of the novice group in the validation study. We considered an improvement of at least 15% to be educationally important. To demonstrate this with a power of 90%, alpha of 0.05, and assuming an SD of the difference of 30%, the required sample size (number of paired readings) was 44, following the approach of Dupont and Plummer[15] using the PS Power and Sample Size Calculation software (Vanderbilt University Medical Centre, Nashville [TN], US).

Statistical analyses

The following tests were used: paired *t* tests to analyse the pre- and post-test scores, Chi squared tests for the proportion that attained the pass mark, Kruskal-

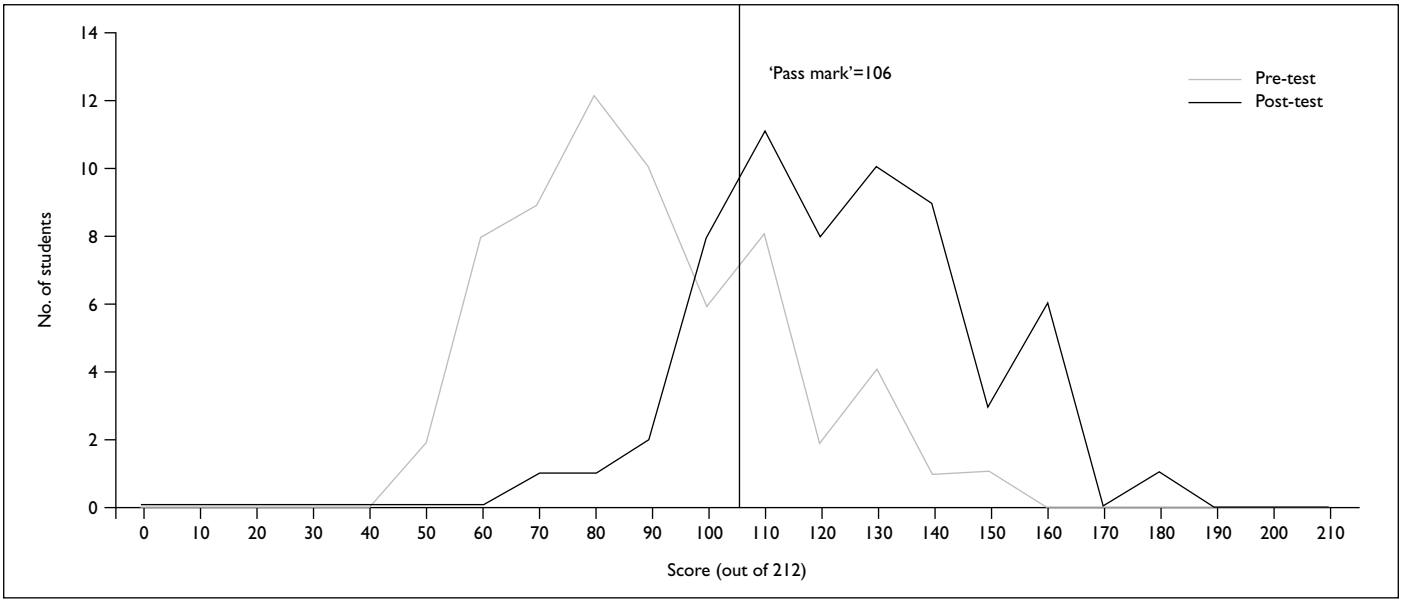


FIG 2. Score distribution of the adapted Fresno test of competence in evidence-based medicine pre- and post-intervention

TABLE. Performance in each question in the adapted Fresno test of competence in evidence-based medicine

Question No.	Area of knowledge tested (format)	Total score allocated	Mean (SD*) pre-test score	Mean (SD) post-test score	P value
1	Formulation of clinical questions (open-ended)	24	17.6 (3.2)	22.7 (1.6)	<0.001
2	Sources of evidence (open-ended)	24	9.7 (4.4)	12.0 (5.1)	0.001
3	Study design to answer clinical question in No.1 (open-ended)	24	9.9 (5.8)	15.5 (4.1)	<0.001
4	Search techniques (open-ended)	24	10.6 (4.2)	14.0 (4.0)	<0.001
5	Critical appraisal: internal validity (open-ended)	24	7.0 (5.7)	17.0 (5.5)	<0.001
6	Critical appraisal: clinical importance (open-ended)	24	5.5 (4.6)	7.7 (6.0)	0.014
7	Critical appraisal: applicability (open-ended)	24	7.0 (5.6)	13.4 (5.5)	<0.001
8	Study design: study on diagnosis (short-answer)	4	0.8 (1.6)	0.6 (1.4)	0.315
9	Study design: study on prognosis (short-answer)	4	2.5 (2.0)	3.0 (1.8)	0.182
10	Statistical expression on the value of diagnosis/screening tools (short-answer)	20	9.3 (6.9)	10.1 (6.4)	0.462
11	Statistical expression on the value of therapeutic tools (short-answer)	12	3.7 (2.7)	4.5 (3.4)	0.136
12	Understanding of 95% confidence interval (multiple-choice)	4	0.4 (0.9)	0.8 (1.4)	0.129

* SD denotes standard deviation

Wallis tests to compare scores of subgroups that received the small-group intervention at different periods, Wilcoxon's signed rank tests to compare the pre- and post-test scores within each subgroup, and Pearson's product-moment correlation for inter-rater agreement and correlation between the test scores and students' overall clerkship performance. The Statistical Package for the Social Sciences (Windows version 11.0; SPSS Inc, Chicago [IL], US) was used to conduct these analyses.

Research and Ethics Committees approval

This study was approved by the Research and Ethics Committees of the International Medical University, Malaysia.

Results

Sixty-three (88%) and 60 (83%) of the students sat for the pre- and post-tests, respectively. Only 55 pairs of scripts were analysed as identifying information for pairing was incomplete for the remainder. Overall, inter-rater correlation was 0.90 (95% CI, 0.87-0.93). Excluding question 12 for which inter-rater correlation was perfect as expected, correlation for individual questions ranged from 0.52 to 0.98. The scores from the two raters differed by an average of 13 points (SD, 13). Overall internal reliability of the test (Cronbach's alpha from the average of all possible split-half

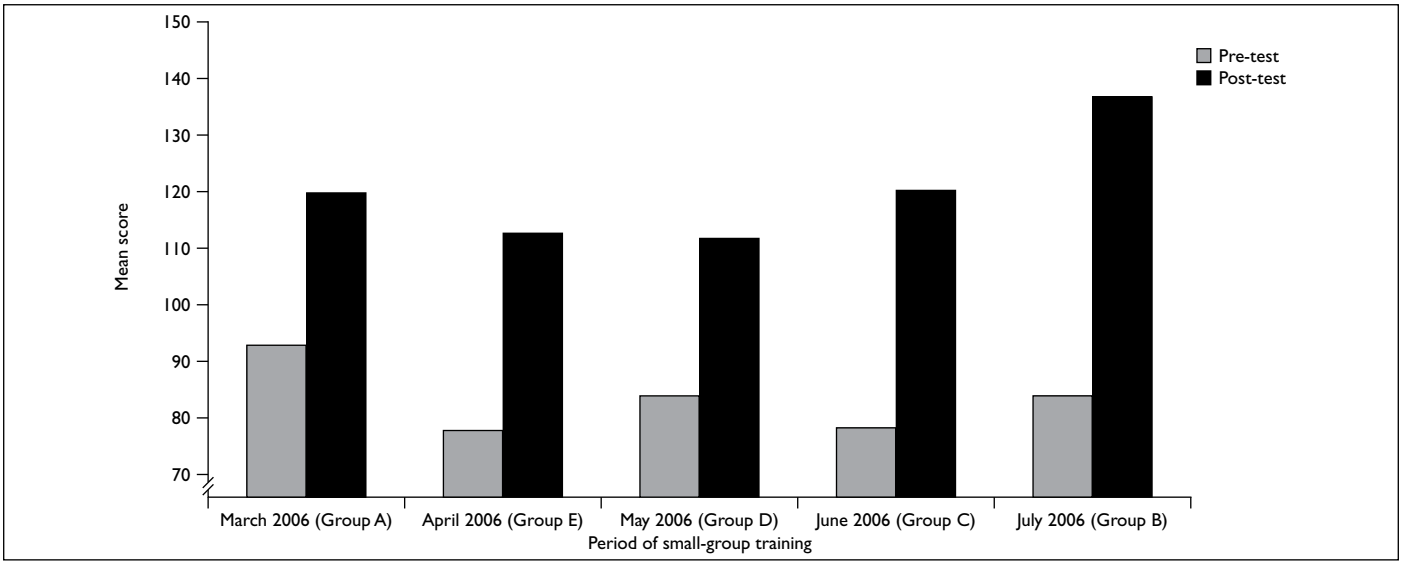


FIG 3. Pre- and post-test scores for each subgroup that received the small-group training in the paediatric posting at different periods in the senior clerkship training

correlations) was 0.55.

Overall, we found significantly improved test scores post-intervention (mean: pre-test, 84 [SD, 24]; post-test, 122 [SD, 22]; P<0.001). This represented an improvement of 18%, from 40% to 58%. The proportion that attained the 'pass mark' of 106 increased from 16% to 77% (P<0.001) [Fig 2]. As illustrated in the Table, most improvements were in questions 1 to 7.

For our secondary outcome, the students' overall score in the senior clerkship, as indicated by the average of the three major clerkship assessments, was 74% (SD, 3%). There was a significant correlation between EBM test scores and their senior clerkship assessment scores (r=0.329, P=0.014).

Figure 3 shows the scores of different subgroups that separately received the small-group training in their paediatric posting. As illustrated, among these subgroups pre-test scores were similar (P=0.563). However, despite significant post-test improvements from the pre-test for all the subgroups (P value, 0.001-0.017), the scores varied considerably; lower scores being observed in subgroups that received the small-group training earlier in the clerkship (P=0.03). The overall clerkship performance, as indicated by the average scores of three major assessments, was similar among the subgroups (P=0.992).

Discussion

This study shows that a period of structured, clinically integrated undergraduate training in EBM produces significant, educationally important gains in EBM knowledge, measured using an objective tool. This finding was consistent with previous studies on physicians and occupational therapists.[1,9,10] Senior medical students were chosen for our educational intervention based on our belief that their clinical maturity relative to junior students would enable them to better appreciate the clinical relevance of EBM, and hence they would be more receptive to our educational intervention. Notably, our students' post-test scores compared favourably with those of practising health care staff in the validation study, ie mid-way between the novice and the expert groups.[8]

The first seven questions of our test script covered core EBM concepts, including "asking clinical questions", "searching", and "critical appraisal". The next five questions covered relatively more in-depth concepts, like study design and statistical expressions with calculations. In this study, significant improvements were observed only in relation to the first seven questions. The greatest improvement was noted in question 5, which covered internal validity aspects of critical appraisal. These findings clearly reflect our EBM curricular design, with emphasis on core EBM concepts and critical appraisal in preference to more complex statistical concepts. Whether such selective emphasis is appropriate or a weakness in our curriculum merits ongoing evaluation. The moderate agreement between student performance in the adapted Fresno test and their overall performance in the senior clerkship implies that to some extent competence in EBM reflects the overall academic ability of a student, although there might be important differences between the ways competence in EBM is acquired as compared to clinical competence in general.

Comparison of different subgroups that received the small-group training separately yielded an interesting finding. Although all of them had improved test scores post-intervention, those that received training earlier seemed to score lower than those that received training later, ie nearer the

post-test. This could not be explained by differing academic abilities between subgroups, as their pre-test scores and overall academic performance in the senior clerkship were similar. Inconsistent teaching during the small-group training was also an unlikely contributing factor, as all the training sessions were facilitated by a single individual (the first author). To us, the finding suggests that contrary to our expectations, students were not able to retain their EBM knowledge for long after their small-group training. The loss of knowledge within such a short period of time is a concern about EBM knowledge acquired from this training programme. A possible reason may be a lack of consistent application of EBM principles across all disciplines in the senior clerkship, due to variable input from the supervisors, despite students being required to produce EBM reports in their portfolios. It was possible that the supervisors had different levels of confidence and motivation in practising EBM, despite receiving prior instructions. It appears that regular instructions and support of supervisors, as well as their feedback are necessary to maintain a culture to EBM throughout the clerkship. This may facilitate application and retainability of EBM knowledge in the long run.

We acknowledge the following limitations in our study. First, as this was not a controlled trial, we cannot exclude the effect of senior clerkship training in general on the observed improvements. Second, in our setting EBM was a standard curricular requirement and a component in the summative assessment. Our findings may not be applicable to settings without a similar emphasis on EBM. Third, we could not reproduce the excellent internal reliability of the original Fresno test; our overall internal reliability (Cronbach's alpha) was 0.55 compared to 0.88 in the validation study.[8] It is worth noting that the original validation was performed on a group of practising health care staff and not on medical undergraduates. Our adaptations for the undergraduates, despite having gone through a round of piloting, could have been further improved.

Notwithstanding these limitations, we believe that overall, the degree of improvement demonstrated in this study lends sufficient weight to support our undergraduate EBM training programme. Our findings imply that there is a much greater challenge awaiting us. Namely, to ensure that such EBM knowledge as is acquired is retained when the students venture into clinical practice. Only then could we assess whether such a training programme has any long-term effect on improving clinical decision-making behaviour and patient care.

Acknowledgements

We would like to express our gratitude to Professor Ramesh Jutti, Academic Dean (Clinical Sciences), International Medical University for providing invaluable advice and input in the first abstract of this study, and to Mrs Hapiza Baharom and Ms Fei Chuen Liau from Administrative Department, International Medical University, for providing administrative support in the study. This research was funded internally by International Medical University, Malaysia.

References

1. Guyatt G, Cook D, Haynes B. Evidence based medicine has come a long way. BMJ 2004;329:990-1.
2. Dobbie AE, Schneider FD, Anderson AD, Littlefield J. What evidence supports teaching evidence-based medicine? Acad Med 2000;75:1184-5.
3. Reilly BM. The essence of EBM. BMJ 2004;329:991-2.
4. Norman GR, Shannon SI. Effectiveness of instruction in critical appraisal (evidence-based medicine) skills: a critical appraisal. CMAJ 1998;158:203-4.
5. Del Mar C, Glasziou P, Mayer D. Teaching evidence based medicine. BMJ 2004;329:989-90.
6. Straus SE, Ball C, Balcombe N, Sheldon J, McAlister FA. Teaching evidence-based medicine skills can change practice in a community hospital. J Gen Intern Med 2005;20:340-3.
7. 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;325:1338-41.
8. Ramos KD, Schafer S, Tracz SM. Validation of the Fresno test of competence in evidence based medicine. BMJ 2003;326:319-21.
9. Dinkevich E, Markinson A, Ahsan S, Lawrence B. Effect of a brief intervention on evidence-based medicine skills of pediatric residents. BMC Medical Education website: http://www.biomedcentral.com/1472-6920/6/1. Accessed 1 May 2006.
10. McCluskey A, Lovarini M. Providing education on evidence-based practice improved knowledge but did not change behaviour: a before and after study. BMC Med Educ 2005;5:40.
11. Lai NM, Ramesh JC. The product of outcome-based undergraduate medical education: competencies and readiness for internship. Singapore Med J 2006;47:1053-62.
12. Sackett DL, Straus SE, Richardson WE, Rosenberg WM, Haynes RB. Evidence-based medicine: how to practice and teach EBM. 2nd ed. Edinburgh: Churchill Livingston; 2000.
13. Teaching EBM: Teaching materials: Critical appraisal worksheets. Toronto: Centre for Evidence Based Medicine: University Health Network website: http://www.cebm.utoronto.ca/teach/materials/caworksheets.htm. Accessed 30 Dec 2008.
14. Ben-David M. AMEE Guide No. 18: Standard setting in student assessment. Medical Teacher 2000;22:120-30.
15. Dupont WD, Plummer WD Jr. Power and sample size calculations. A review and computer program. Control Clin Trials 1990;11:116-28.