Correlation of self-assessment with attendance in an evidence-based medicine course

Beatriz U. Ramirez
Faculty of Medical Sciences, Universidad de Santiago de Chile (Usach), Santiago, Chile

Submitted 6 May 2015; accepted in final form 2 August 2015

Ramirez BU. Correlation of self-assessment with attendance in an evidence-based medicine course. Adv Physiol Educ 39: 378–382, 2015; doi:10.1152/advan.00072.2015.—In previous studies, correlations between attendance and grades in lectures have given variable results and, when statistically significant, the correlation has been weak. In some studies, a sex effect has been reported. Lectures are a teacher-centered learning activity. Therefore, it appeared interesting to evaluate if a stronger correlation between attendance and grades would occur in a face-to-face “evidence-based medicine” course with few lectures and more time dedicated to active learning methods. Small-group work and peer learning were used to foster deep learning and to engage students in their own learning process. Most of the time, students worked in small groups solving contextualized problems and critically analyzing the quality of published medical literature. Peer learning was also developed in collaborative evaluations, and constant feedback was provided. Therefore, it was hypothesized that high attenders would develop a higher self-perception of learning and obtain higher marks than low attenders. Student self-perceptions of their capacity to apply evidence-based medicine were measured by the application of an online self-assessment survey, and objective learning was measured as the grades obtained in a final accumulative individual test. It was found that male students obtained higher grades and were more confident in their achievements than their female peers, despite male and female student attendance being similar. In addition, attendance was correlated with the perceived capacity to apply evidence-based medicine only in male students and was not correlated with academic outcome.

active learning; attendance; large classes; medicine students; evidence-based medicine

When only lecture attendance is evaluated, the influence of other factors that may be relevant for student success, such as collaborative peer learning or different sensory modalities used for learning (2, 7, 18, 21, 23), are not taken into account. Currently, in most undergraduate curricula for medical and other health-related careers, students’ active participation in their learning process means much more than just attending lectures. Despite lectures still being widely used, other more interactive learning activities that provide opportunities for peer learning, as students working in small groups searching for the solution of contextualized problems or cases in a clinical scenario, are increasingly being used.

Having an interactive model of the teaching-learning process in mind, a course of “evidence-based medicine” for undergraduate medical students was recently implemented at the Universidad de Santiago in Chile. The course was focused on the development of critical analysis capacity, oriented to the evaluation of the level of confidence and credibility of medicine papers. Because this course had a strong focus on the development of higher order skills and collaborative peer learning had more dedicated time than lectures, it was interesting to know if attendance in the programmed learning activities was related with outcome and/or with student perceptions of their learning. This was evaluated in male and female students.

METHODS

A group of 63 fourth-year undergraduate medical students attended a new, specially designed, evidence-based medicine course in the second semester 2014 at the School of Medicine of the Universidad de Santiago. The course was offered in a “face-to-face” modality [3-h sessions every week for 12 wk (learning sessions) plus two sessions for individual evaluation: first and final tests] and had a webpage with supporting material in the institutional Moodle platform. Handouts, PowerPoint slides used in lectures, clinical cases, contextualized problems, links to evidence-based medicine webpages, and other learning materials were available to download and print from the course webpage. Standard forms for the critical evaluation of papers dealing with treatment, diagnosis, or prognosis or systematic reviews (4) were uploaded to the course website.

Every session, students worked in small groups (3–6 students) solving problems and clinical cases or analyzing published papers extracted from the PubMed database. The groups were formed freely at the first session. In many but not all sessions, a short lecture (20–30 min) about the relevant concepts of the current course topic was first provided. During the sessions, open discussions of the results and the corresponding conclusions among students of all groups were guided by a tutor. Synthesis and analysis were encouraged to challenge students’ thinking. Free access to the internet and printed material was granted during the learning sessions and collaborative evaluations.

Four short written examinations with open-ended questions that were answered in peer collaboration (with the same grade for each student-group member) were applied during the semester, in the last 20 min of the corresponding session. In addition, two written indi-

Address for reprint requests and other correspondence: B. U. Ramirez, Escuela de Medicina, Facultad de Ciencias Medicas, Universidad de Santiago de Chile, Alameda 3363, Estacion Central, Santiago 9170019, Chile (e-mail: beatriz.ramirez@usach.cl).
individual tests consisting of open-ended questions were applied, in the middle and at the end (final test) of the semester. The final test was accumulative and intended to measure the student’s ability to apply evidence-based medicine. Examples of questions applied in the final test are shown in Table 1. Students were allowed to bring any written material they wanted to use during the individual tests, but multimedia tools with access to the internet and cell phones were prohibited. After all but the final test, feedback was provided through open discussions of the wrong and right answers (sometimes there were different ways to get a right answer) at the beginning of the next session. Student final grades (for promotion) were calculated as the weighted mean of the individual tests (35% each) and collaborative tests (30% for the mean of the four collaborative tests). The course was approved with a final grade of 57%, according to university standards.

Course attendance was recorded by means of sign-up sheets passed around during each session. Regular attendance was encouraged, but it was not mandatory (no penalty for absences).

An online self-assessment with 16 closed questions that covered the main course objectives (Table 2) was applied the day after the final test, before the corresponding grades were known by the students. Questions were formulated as “Can you verb...subject?” The closed answers were as follows: “No,” “I have heard/read how to do it,” “I have seen others doing it,” “I needed help to do it,” and “I have done it alone and can repeat it.” The answers to the options of “No”, “I have heard/read how to do it,” and “I have seen others doing it” were collected together into a category named “unsatisfactory.” Each student self-assessment then received points, calculated as the sum of the points obtained in all questions, according to the category of the answer, with a maximum of 48 points, as follows: proficient (I did it alone and can reproduce it) (3 points), needs help (I needed help to do it) (2 points), and unsatisfactory (1 point).

Only those students for whom the final test, the attendance record, and a self-assessment survey were available included in the study (n = 53, 84.13%).

Descriptive statistics are presented as means ± SD. Unpaired t-tests were used to identify differences in variable scores between male and female students, and Bonferroni’s multiple-comparison test after ANOVA was used when more than two samples were compared. Pearson correlations were calculated between attendance and points scored in the self-assessment, between attendance and grades, and between points scored in the self-assessment and grades. Two-tailed P value of <0.05 were considered statistically significant.

RESULTS

There were 30 male students and 23 female students that fulfilled the requirements to participate in the study. The mean attendance in course activities, expressed as percent of total learning sessions, was similar between male and female students (Table 3).

Student academic outcome was evaluated here as the grade obtained in the final test, because it measured the acquired ability to apply evidence-based medicine. As shown in Table 3, male students obtained significantly higher grades than female students in the final test.

In addition to an objective measurement, the student’s own perception of learning was also measured through a self-

<table>
<thead>
<tr>
<th>Questions</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example 1. Which of the following terms are characteristics of this study? Randomized/not randomized Blind/Double blind/Other Controlled/Without control group Multicenter/Not multicenter Prospective/Retrospective/Transversal</td>
<td>The effectiveness of tamoxifen in breast cancer was assessed in a trial in which 20 hospitals participated. Altogether, ten thousand women immediately after breast cancer surgery were included. In each participating hospital, a list with numbers that had been generated by a random program was put up. If a woman was included in the study, she was allocated to the next available number on the list, which was then crossed off. If the number was even, the patient was treated with 2 × 30 mg tamoxifen daily for 3 yr. If the number was odd, the patient was observed for the same time period but not treated. A drug representative tells you about the results of a study showing that a new lipid-lowering drug produced a 50% reduction in the risk of dying from myocardial infarction. The study was done with a large number of patients (healthy company employees) using a randomized, double-blind prospective protocol. Of the 4,000 employees treated with the drug, 4 employees died (0.1%). In the control untreated group, 8 employees died (0.2%). At a gastroenterology department, the prevalence of colon carcinoma was 30%. This department evaluated a new, noninvasive diagnostic test for colon carcinoma. The study was done with 1,000 consecutive patients. Of these, 630 patients were truly negative for cancer. The number of false positive and false negative patients was identical.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>New Test</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gold Standard</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Questions applied in the online survey for the self-assessment of learning

<table>
<thead>
<tr>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Identify the design of a biomedical or clinical study</td>
</tr>
<tr>
<td>2. Identify limitations in the quality in different types of clinical studies</td>
</tr>
<tr>
<td>3. Construct a relevant clinical question to search medical literature databases</td>
</tr>
<tr>
<td>4. Improve searching strategies for finding answers to clinical questions</td>
</tr>
<tr>
<td>5. Apply validity criteria to treatment studies</td>
</tr>
<tr>
<td>6. Calculate absolut risk reduction, relative risk reduction, and the number needed to treat</td>
</tr>
<tr>
<td>7. Critically assess an article dealing with questions of diagnosis</td>
</tr>
<tr>
<td>8. Calculate sensitivity and specificity using data from an article dealing with questions of diagnosis</td>
</tr>
<tr>
<td>9. Calculate the likelihood ratio from data in a paper dealing with questions of diagnosis</td>
</tr>
<tr>
<td>10. Apply a likelihood ratio to evaluate the probability that your patient has the diagnosed pathology</td>
</tr>
<tr>
<td>11. Identify the major differences between cohort and case-control studies</td>
</tr>
<tr>
<td>12. Understand results presented as odds ratios with confidence intervals</td>
</tr>
<tr>
<td>13. Critically assess the quality of a systematic review</td>
</tr>
<tr>
<td>14. Understand a meta-analysis and evaluate data heterogeneity</td>
</tr>
<tr>
<td>15. Understand the hierarchy of studies for answering clinical questions</td>
</tr>
<tr>
<td>16. Understand a forest plot</td>
</tr>
</tbody>
</table>

Questions were multiple choice type. All questions started with “Can you” followed by the phrase indicated in the corresponding numbers. The screen displayed one question and its five possible answers (“No,” “I have heard/read how to do it,” “I have seen others doing it,” “I needed help to do it,” and “I did it alone and can reproduce it”) at each time.

assessment applied online, after the course was finished but before final test grades were known by the students. The 16 questions of the survey, which are shown in Table 2, covered all the most relevant course topics. According to the outcomes in the final test, male students turned out to be more confident in their capacities to apply evidence-based medicine than female students and obtained significantly higher points in the survey (Table 3). As shown in Fig. 1, this result was due to a different distribution of the survey points between male and female students. Only female students had the lowest points, whereas the highest points were obtained by 13 male students but only by 1 female student. In the middle range (35–44 points), male and female students obtained similar results.

However, no correlation was found between final test grades and self-assessment in either the whole course (n = 53 data pairs, R = 0.183, P = 0.191) or in male students (n = 30, R = 0.089, P = 0.640) or in female students (n = 23, R = 0.041, P = 0.852).

When attendance in course activities was associated with the points obtained in the corresponding self-assessment survey, a statistically significant positive correlation was found in the whole course (n = 53, R = 0.294, P = 0.033), due to the highly significant correlation between both variables in male students (Fig. 2). However, no correlation was found between attendance and self-assessment in female students (Fig. 2).

In addition, no correlation was found between final test grades, in combined male and female data, and attendance (Fig. 3). Accordingly, no correlation was found between attendance and final test grades in male students (n = 30, R = 0.111, P = 0.559) or in female students (n = 23, R = 0.174, P = 0.426).

Since no report was obtained from students concerning the amount of time they spent learning in this course, other than the lesson-reserved hours, visits payed by each student to the webpage were the only hints to estimate students’ interest in their self-learning process. However, to avoid memorization, students were allowed to bring any written material to the tests.

Therefore, it was expected that most, if not all, students downloaded from the course website the lecture slides, the math form, and the forms used to evaluate the quality of the analyzed papers. As expected, the math form, which was used in many learning sessions to work out data from papers (for example, absolute risk, relative risk reduction, and other) was accessed by all except one student and she could have photocopied the material. In addition, the different standard forms used for the evaluation of the quality of medical papers, besides being uploaded in the webpage, were handed out in the learning sessions (1 form/group), and were photocopied at will.

Therefore, the website visits to these materials were useless as indicators of the time used by students in self-directed learning, because they may have studied using copies of the corresponding materials without having visited the website. However, looking around in the final test, it could be seen that the students had colored slide prints, suggesting that they printed (and not photocopied) the slides used in lectures directly from the course webpage. Therefore, a webpage review allowed an estimation of how many students accessed the uploaded slides provided in the course. It was found that some students visited several lectures many times and others visited some of the lectures only once. Since one visit is enough to print the slides, only one visit per lecture was accounted for each student. As shown in Fig. 4, lectures in subjects 1–3 received a large amount of visits (81–95% of total possible visits), but subject
received only 60% of the possible visits. All lectures were given by the same teacher, and subject 4, the last subject in the semester, occurred when all courses were approaching their exams.

**DISCUSSION**

In other studies, a statistically significant but weak correlation between lecture attendance and marks has been described, and when the sex effect was considered, a correlation between both variables was found only in the sex with a higher outcome (5, 9). A similar result was obtained here when the correlation between attendance and self-assessment was measured. Despite attendance being similar in both sexes, a statistically significant correlation between these variables was found only in male students, but not in female students, and male students had significantly higher values than female students in the self-perception survey.

Student perceptions of their capacity to apply evidence-based medicine after the course was clearly different between sexes, with male students being more confident in their capacity than female students. In addition, this perception was coincident with male and female student academic outcomes. Therefore, there was a match between student perceptions and outcomes in relation to sex.

One of the abilities needed to apply evidence-based medicine is the capacity to solve some basic math calculations, which is required to verify if the author conclusions are based on the right data (4). All math formulae required for these simple calculations were available in the webpage. However, the fact that some students may have been more skilled than others in the use of math logic may have affected the final test grades. In this connection, a national study compared skills in male and female young students and concluded that male student achieved higher marks than their female peers in science and math (1). This may be due to social pressure during growth, which associates sex with certain skills, or to other factors, because, despite there being individual female students with good marks, female students had a lower mean mark than male students. This could also have been a characteristic of the cohort that was studied here. Results in future courses may enlighten this point.

The large variability in outcomes obtained in the final test by students that had attended a similar percent of the total course learning activities may be due to several factors that affect

Fig. 2. Attendance in course learning activities was associated with self-assessment points in male students but not in female students. No correlation was found in female students ($R = 0.205, P = 0.35$). In contrast, there was a statistically significant correlation between attendance and self-assessment in male students ($R = 0.43, P = 0.018$).

Fig. 3. No correlation was found between attendance and final test grades in combined male and female data. Pearson correlation $R = 0.099$ ($n = 53$ data pairs, $P = 0.480$).

Fig. 4. No correlation was found between attendance and self-assessment in male and female students. Pearson correlation $R = 0.205$ ($n = 53$ data pairs, $P = 0.35$). In contrast, there was a statistically significant correlation between attendance and self-assessment in male students ($R = 0.43, P = 0.018$).

Fig. 4. Visits to the slides used in lectures. Most students visited the slides in subjects 1–3, but subject 4, which was the last subject, received significantly fewer visits (60% of possible visits). $P < 0.0001$ by ANOVA for subjects 1–4. $P < 0.001$ for subject 1 vs. 4 and $P < 0.05$ for subjects 2 and 3 vs. subject 4 by Bonferroni’s multiple-comparison tests. There were no differences among subjects 1, 2, and 3. Data are means ± SE.
learning. Among them are the different velocities at which students learn (3) and the different amounts of time dedicated to self-directed learning (6, 19). The latter was not directly evaluated here. However, competition for the time dedicated to learning in the several courses that are taken at the same time in medical curriculum is huge, and this evidence-based medicine course probably did not have a high priority compared with internal medicine or other clinical courses. In addition, some students may have had a greater than real sense of accomplishment and confidence in their capacities to apply evidence-based medicine, based on the results obtained in the collaborative tests, because it has been shown that the marks in collaborative tests are higher than those obtained when the same test is done individually (10). Both factors may have affected self-directed learning time mainly in the final course stage, when all courses were close to their final exam date.

This course was focused on the development of abilities for applying concepts, not on memorization of facts, and the learning strategies recommended to achieve this goal are active learning and peer instruction with immediate feedback (7, 16). Therefore, the programmed learning activities dedicated most of the time to hands-on application of concepts, with the students working in small groups to foster a collaborative learning environment (14). Even during the evaluation of learning was peer learning facilitated, by the application of collaborative tests. In addition, feedback was constantly provided through the tutor-directed discussion of questions and right and wrong answers among students in the different groups, to promote meaningful learning. This teaching methodology provided a student-centered learning environment and offered a wide range of opportunities to engage students in an active learning process (2, 14, 15, 20).

According to expectations, a positive correlation between self-perception of learning and attendance occurred in this course, but only in male students. In contrast, there was no correlation between academic outcome and attendance, despite there being a match between student perceptions of their capacity to apply evidence-based medicine and outcome between sexes. Therefore, it is possible that a much larger number of students be needed to detect, in this course, an association between attendance and outcome or between attendance and self-assessment of learning in female students.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: B.U.R. conception and design of research; B.U.R. performed experiments; B.U.R. analyzed data; B.U.R. interpreted results of experiments; B.U.R. prepared figures; B.U.R. drafted manuscript; B.U.R. edited and revised manuscript; B.U.R. approved final version of manuscript.

REFERENCES