Teaching and Learning in Medicine: An International Journal
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/htlm20

Web-Based Learning Versus Standardized Patients For Teaching Clinical Diagnosis: A Randomized, Controlled, Crossover Trial
Michael K. Turner, Steven R. Simon, Kevin C. Facemyer, Lynn M. Newhall & Tracy L. Veach
Available online: 15 Jun 2010

To cite this article: Michael K. Turner, Steven R. Simon, Kevin C. Facemyer, Lynn M. Newhall & Tracy L. Veach (2006): Web-Based Learning Versus Standardized Patients For Teaching Clinical Diagnosis: A Randomized, Controlled, Crossover Trial, Teaching and Learning in Medicine: An International Journal, 18:3, 208-214
To link to this article: http://dx.doi.org/10.1207/s15328015tlm1803_4

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Web-Based Learning Versus Standardized Patients For Teaching Clinical Diagnosis: A Randomized, Controlled, Crossover Trial

Michael K. Turner
Department of Physical Medicine and Rehabilitation
Mayo Clinic
Rochester, Minnesota, USA

Steven R. Simon
Department of Ambulatory Care and Prevention
Harvard Medical School and Harvard Pilgrim Health Care
Boston, Massachusetts, USA

Kevin C. Facemyer
Department of Biochemistry and Molecular Biology
University of Nevada School of Medicine
Reno, Nevada, USA

Lynn M. Newhall
Department of Medical Education
University of Nevada School of Medicine
Reno, Nevada, USA

Tracy L. Veach
Psychiatry and Behavioral Sciences and Internal Medicine
University of Nevada School of Medicine
Reno, Nevada, USA

Background: Little evidence exists to guide the selection of methods for teaching clinical diagnosis.

Purpose: To compare the efficacy, student preference, and cost of a Web-based (WB) program versus a standardized patient (SP) encounter for teaching clinical diagnosis skills to 2nd-year medical students.

Methods: Randomized, controlled, crossover study comparing WB versus SP-based teaching for the clinical diagnosis of abdominal pain and headache. Outcome measures were performance on a 2-case SP examination (scored on the basis of a checklist completed by a faculty observer and an objective score on a postencounter subjective–objective assessment plan [SOAP] note), format preferences as assessed by end-of-course evaluations, and cost.

Results: Thirty students consented to participate. WB and SP training produced similar scores on both the Abdominal Pain checklist (66% vs. 62%; \(p = .17\)) and Headache checklist (56% vs. 63%; \(p = .07\)). WB training produced a higher score on the Abdominal Pain SOAP note (69% vs. 47%; \(p = .006\)), but not the Headache SOAP note (69% vs. 67%; \(p = .85\)). Students rated the SP format higher than the WB format on all 7 preference measures. Start-up costs were estimated at $2,190 for the SP format and $2,250 for the WB format. Ongoing costs per case per student were estimated to be $45 for the SP format and $30 for the WB format.

Conclusions: WB and SP learning outcomes were comparable, but students preferred the SP format. Start-up costs were comparable, but the ongoing costs of the WB format were less expensive, suggesting that WB teaching may be a viable strategy.
As methods of teaching clinical reasoning to medical students continue to evolve, there is an ongoing need to develop and refine methodological approaches. The use of both standardized patient (SP) and Web-based (WB) teaching methodologies has increased dramatically over the last decade.1,2 Both of these methods have been widely implemented by medical schools in the United States.3,4

The use of SPs rests on a foundation of empirical support. SPs have been shown to improve the clinical skills of physicians when used as a form of continuing medical education.5 SPs can effectively teach the normal physical examination to medical students and are a less expensive alternative to traditional faculty-led small-group teaching methods.6 Furthermore, SPs were shown to be more effective than a lecture—demonstration in teaching students how to approach a patient eligible for breast cancer screening.7

The literature regarding multimedia instruction in medical education is less clear. One study showed WB instruction to be more effective than paper-based approaches for teaching certain physical examination skills.8 Four other studies, however, found no difference in test scores between WB and paper formats.9–12 Similarly, some studies have reported greater student satisfaction with WB approaches,8,9,11 whereas others have found no difference when compared with traditional methods.13–15 Almost uniformly, however, WB learning has been lauded for advantages such as its remote access and interactivity,16 ability to deliver just-in-time clinical information,17 and ability to produce equivalent learning outcomes with less study time.11,12

The importance of teaching medical students to be reproducibly proficient in their clinical skills has been underscored by the addition of the clinical skills component to Step 2 of the United States Medical Licensing Examination, a prerequisite for national licensure in all 50 states. This test assesses two skill sets: doctor–patient communication and clinical reasoning. To assess clinical reasoning, students record each patient encounter in a subjective–objective assessment plan (SOAP) format, which is then scored by trained physician reviewers. Students’ scores reflect their ability to gather relevant data from the interview and physical examination, generate and investigate diagnostic hypotheses, reach appropriate diagnostic conclusions, advise appropriate treatment, and record their findings and impressions clearly.18

As medical schools address the question of how best to prepare students for Step 2 of the United States Medical Licensing Examination—and more broadly, how best to teach clinical reasoning—they encounter a set of difficult questions and a paucity of data on which to make rational decisions. Few controlled studies have been published, and to our knowledge, no controlled studies have directly compared the SP and WB formats. Which format demonstrates a greater ability to improve clinical reasoning? How robust is this influence? How do the costs of each method compare? Also, how do students respond to them? This article begins to address these questions.

Methods

Design, Setting, and Participants

This randomized, controlled, crossover trial was conducted during the 2002–2003 academic year at the University of Nevada School of Medicine, where the institutional review board approved the study protocol. Fifty-four 2nd-year medical students participated in the study via their enrollment in a required clinical problem-solving and physical diagnosis course. Thirty students consented to allow their data to be included in the analysis. None of the authors had any financial interests with the manufacturer of the software used to present the WB cases.

Interventions

The 54 students were randomized via computerized random-number generator into two cohorts: WB followed by SP instruction (Cohort 1) versus SP followed by WB instruction (Cohort 2). Students in both cohorts participated in two sequential teaching sessions. The first teaching session involved abdominal pain. Four weeks later, a headache case was presented, during which, each cohort was taught with the opposite modality (Figure 1).

WB cases. Students reported to the computer laboratory and logged on to a server that had cases with either abdominal pain or headache pain as a presenting complaint. These cases were obtained from a commonly available commercial software package, DxE Clinician© (DxR Development Group, Carbondale, IL). To prevent students not randomized to this condition from viewing the Web-format cases, an administrator made the Web-format cases available only to the selected students, and students were allowed to log in and access them only for the duration of the 45-min teaching session.

SP cases. Students participated in an observed SP case, which required a focused history and physical examination of a patient presenting with abdominal pain or headache pain. At the end of the case, a faculty observer gave 5 min of verbal feedback.

Every effort was made to make the two formats as comparable as possible. The content of the standard WB cases was reviewed and modified by an expert faculty panel to be consistent with the SP cases. SP actors were carefully trained to a high standard of reliability
and consistency with the WB cases. A common checklist of required history and physical examination components was created for each case. Each session was 45 min long and required students to perform a focused history and physical examination (a virtual one in the case of WB) and to write a postencounter note using a modified SOAP format (Figure 2). The SOAP note required students to list pertinent findings, generate a list of possible diagnoses, argue for a particular diagnosis, and suggest further diagnostic studies.

For SP cases, faculty observers completed the checklist of history and physical examination requirements as the encounter unfolded. For WB cases, student performance on the history and physical examination items was recorded by the computer. These performances were later reviewed by faculty members and manually converted into a checklist score.

Four days after each session, students in both cohorts received written feedback that included the completed checklist of their performance, a SOAP note answer key, and a study guide listing the differential diagnosis for the chief complaint (i.e., abdominal pain or headache) and the findings commonly associated with each condition.

**Outcome Measures**

Four weeks following the second teaching session, students in both cohorts completed a final examination. The final examination consisted of two observed
SP encounters, one involving abdominal pain and the other involving headache pain. The content of these cases was similar to earlier exposures but with variations to reduce perceived redundancy (e.g., diagnosis, sex, age). The scores on this examination comprised the primary outcome measure of this study: knowledge acquisition in clinical reasoning.

Students had 20 min to complete each SP patient examination. Two trained faculty observers were present and scored student performance on standard checklists. After each encounter, students were given 10 min to complete a SOAP note. Two trained reviewers, who were blind to the students’ cohort allocation, graded the SOAP notes using an answer key created by faculty. We calculated Cronbach’s alpha for both checklist and SOAP note scoring.

After each teaching session, students completed a written evaluation of the experience using a 5-point Likert-type scale ranging from 1 (strongly disagree) to 5 (strongly agree) to assess agreement with seven statements.

We recorded the start-up costs and ongoing costs involved with providing each type of teaching session. To assess the financial costs attributable to faculty and staff labor efforts, we used a conservative estimate of $50 as an hourly wage for both faculty and technical staff.

Statistical Analyses

Analyses followed intention-to-treat principles. All 30 students who consented to having their final examination data included in the analyses received the interventions to which they were randomized. All students who completed the educational sessions also completed the postcourse evaluation of preference for learning format.
We used $t$ tests to compare the mean final examination SOAP note and checklist scores of each group and to compare students’ written evaluations of the teaching sessions. Probability values of less than .05 were considered to represent statistical significance.

**Results**

Cronbach’s alpha was .69 for the checklist and .68 for the SOAP note, which suggested acceptable reliability. A total of 49 students completed both training sessions and the final exam (Figure 1). Of those, 30 students consented for their data to be included in analysis: 16 in Cohort 1 and 14 in Cohort 2 (Figure 1). Each cohort had similar consenting rates. Nonconsenters were more likely than consenters to be male, although the result did not reach statistical significance (74% of nonconsenters were male vs. 47% of consenters; $p = .06$).

**Knowledge Gained**

WB and SP training resulted in similar scores on both the Abdominal Pain checklist and the Headache checklist (Table 1). WB training produced a higher score on the Abdominal Pain SOAP note but not the Headache SOAP note (Table 1).

**Format Preference**

A total of 53 evaluations for SP sessions and 53 evaluations for WB sessions were received. Students consistently preferred the SP format, evaluating all seven rating dimensions significantly higher than the WB format (Table 2).

**Cost Comparison**

We recorded both the start-up and ongoing costs associated with providing each type of training. Creating two SP-format cases required approximately 19 hr of faculty time, $240 of SP training salary, and $1,000 in capital expenses (e.g., otoscopes, sphygmomanometers, gowns). Start-up costs for two equivalent WB cases were 8 hr of faculty time, 2 hr of technical support staff, and $1,800 to purchase the case software. Thus, estimated total start-up costs were $2,190 for SP teaching and $2,250 for WB teaching.

Ongoing costs, attributable to the time involved for faculty, administrative staff, and SPs, were estimated to be $45 per case per student for the SP and $30 for the WB format.

**Discussion**

In this randomized, crossover trial comparing WB and SP formats to teach clinical diagnosis skills to medical students, we found that test scores and start-up costs were comparable, students strongly preferred the SP, and the WB format had lower ongoing costs.

Each final examination with an SP had two components: the checklist score and the SOAP note score. The checklist was intended as a measure of clinical thoroughness; the SOAP note was intended as a measure of diagnostic reasoning ability. WB and SP training produced comparable outcomes on both checklist and SOAP note scores for both the abdominal pain and headache final examinations, with one exception: WB trainees scored higher on the abdominal pain SOAP note.

One possible explanation for these findings is that the clinical problem of abdominal pain is better suited to WB training than headache is. The study was not designed to test this hypothesis. Overall, the fact that test scores were comparable on three out of four outcome measures suggests that teaching medical students clinical diagnosis skills by computer is as effective but not substantively more effective than teaching with SPs. This finding is important because, to our knowledge, this is the first study that directly compared these two teaching formats.

These results echo those of earlier studies, which showed no difference in test scores between WB for-

| Table 1. Mean and Standard Deviations of Student Performance Scores Following Teaching Via Standardized Patient and Web-Based Teaching Formats |
|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                  | Standardized Patient$^a$ |                 | Web Based       |                 |                 |                 |                 |                 |                 |                 |                 |
|                  | $M$               | $SD$             | $M$             | $SD$             | $d$             | $p$             | $M$             | $SD$             | $d$             | $p$             |
| Abdominal Pain Checklist | 0.62             | 0.08             | 0.66             | 0.08             | 0.48             | 0.17            |
| Headache Checklist      | 0.63             | 0.12             | 0.56             | 0.08             | 0.07             | 0.07            |
| Abdominal Pain SOAP Note | 0.47             | 0.21             | 0.69             | 0.19             | 0.06             | 0.006           |
| Headache SOAP Note      | 0.67             | 0.28             | 0.69             | 0.31             | 0.07             | 0.85            |

Note: SOAP = subjective–objective assessment plan. $d$ = standardized mean difference, calculated as $(\text{Mean}[\text{Standard Patient}] – \text{Mean}[\text{Web-Based}]) / \text{Pooled SD}$ for each measure. For each condition (abdominal pain or headache), each student’s performance on each measure (checklist or SOAP note) was scored as a proportion of items completed correctly.

$^a\text{n = 14 for the standardized patient abdominal pain checklist and web based headache checklist. n = 16 for the standardized patient headache and Web-based abdominal pain.}$

212
WEB-BASED LEARNING VERSUS STANDARDIZED PATIENTS

Table 2. Mean and Standard Deviation for Student Ratings of Standardized Patient and Web-Based Teaching Formats

<table>
<thead>
<tr>
<th></th>
<th>SP²</th>
<th></th>
<th>WB³</th>
<th></th>
<th>d</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>As a learning tool, this experience was convenient.</td>
<td>3.25</td>
<td>1.25</td>
<td>2.54</td>
<td>1.20</td>
<td>0.58</td>
<td><em>0.004</em></td>
</tr>
<tr>
<td>I feel that this encounter was fairly realistic.</td>
<td>3.79</td>
<td>1.04</td>
<td>2.57</td>
<td>1.28</td>
<td>0.93</td>
<td><em>0.001</em></td>
</tr>
<tr>
<td>This helped me identify skills that need improvement.</td>
<td>4.31</td>
<td>0.67</td>
<td>2.96</td>
<td>1.34</td>
<td>1.07</td>
<td><em>0.001</em></td>
</tr>
<tr>
<td>This case improved my confidence in dealing with headache/abdominal pain.</td>
<td>3.5</td>
<td>0.92</td>
<td>2.45</td>
<td>1.25</td>
<td>0.87</td>
<td><em>0.001</em></td>
</tr>
<tr>
<td>I feel that this experience was good preparation for clinical work.</td>
<td>3.94</td>
<td>1.04</td>
<td>2.57</td>
<td>1.29</td>
<td>1.02</td>
<td><em>0.001</em></td>
</tr>
<tr>
<td>I feel that this experience was good preparation for the USMLE Step 2 Clinical Exam.</td>
<td>3.59</td>
<td>1.02</td>
<td>2.45</td>
<td>.29</td>
<td>0.88</td>
<td><em>0.001</em></td>
</tr>
<tr>
<td>Overall, this was a worthwhile learning experience.</td>
<td>3.94</td>
<td>1.04</td>
<td>2.81</td>
<td>1.30</td>
<td>0.87</td>
<td><em>0.001</em></td>
</tr>
</tbody>
</table>

Notes: SP = standardized patient format; WB = Web-based format; d = standardized mean difference calculated as (Mean[SP]-Mean[WB])/pooled SD for each item. Students responded to the seven statements on a five-point Likert-type scale, with 1 = Strongly Disagree, 2 = Disagree, 3 = Neither, 4 = Agree, 5 = Strongly Agree. 

²N = 53.

³Based on Student’s t test.

mat and paper-based approaches,⁹–¹² although one study suggested superiority of the WB format.⁸ WB training has been shown, however, to be more efficient for students, in the sense that it produces equivalent learning outcomes with less study time.¹¹,¹²

In this study, students consistently preferred the SP format to WB teaching, rating SP higher on all seven rating dimensions. Informal comments from students attributed this preference to the presence of a faculty observer, who offered verbal feedback to the student at the conclusion of the encounter. One student noted, “I really liked the SP better than the WB cases; I enjoyed the immediate feedback. If all of our cases were ‘standardized’ with someone in the room to provide feedback, we would learn much more effectively.” Anecdotally, students also noted that WB teaching suffered from lack of realism, whereas the SP did not. These observations were consistent with other studies that have reported a high level of student satisfaction with SPs.²⁰,²¹

We found that start-up costs of the two formats were comparable but that ongoing costs per case per student were 33% lower with the WB format. Thus, over time and with greater numbers of students, the WB format can offer considerable cost savings. We valued faculty time at $50/hr. This is a conservative estimate of the value of faculty time; higher valuations would tend to favor the WB format because it requires less faculty time than the SP format. We are not aware of any studies that have directly compared the costs of these two formats. Given institutional budget constraints and the greater time and effort involved in recruiting and training SPs, the WB format may offer an important cost advantage.

Our study had limitations. With 24 of 54 students refusing to provide consent to participate, the resultant sample size of 30 students reduced the power of the study to detect statistically significant differences between the two educational formats. For example, with a larger sample size, the 4% absolute difference favoring the WB format in students’ scores on the Abdominal Pain checklist would have been statistically significant; however, such small differences are probably not educationally important. Nonetheless, the fact that only 30 of 54 students agreed to participate deserves further comment. This suboptimal rate of informed consent resulted from a series of unfortunate events. The institutional review board of the University of Nevada Reno provided initial approval of the study protocol prior to commencement of the educational intervention and did not require individual student informed consent. However, during the study, one student expressed concern about the informed consent process and contacted the institutional review board, which, after the completion of data collection, required us to obtain individual informed consent from each student to allow data inclusion for external publication. At that point, only 30 of the 54 students agreed to provide informed consent for the use of their data for external publication, even anonymously and reported in the aggregate. However, an internal analysis showed that the study results did not differ when the data from the nonconsenters was analyzed: The WB training still produced a significantly better results on the SOAP note for the abdominal pain case, but there were no significant differences in the other three scores (Abdominal Pain checklist, Headache checklist, and headache SOAP note). This internal analysis was permitted for program evaluation purposes, but the specific data and details could not be shown in an external research publication.

An additional limitation was the inability to blind students to the nature of the educational intervention. Because this study evaluated only two cases (abdominal pain and headache) and because it was conducted in one medical school, the results may not be generalizable to other settings. In addition, the durability of each learning format remains a question open to further research. This study involved an examination 8 weeks after the abdominal pain teaching session and 4 weeks after the headache session. The finding that the
WB format did produce a higher abdominal pain SOAP note score raises the possibility of greater durability with this intervention. Further research is needed in this area.

The results of this study may help define a rational approach to the appropriation of limited educational resources. SPs remain a proven and well-received method for teaching clinical reasoning to medical students. Future studies should examine the potential advantages of supplementation with computer-based cases (e.g., efficiency, remote access, operational costs per student, reproducibility) in light of this evidence suggesting equivalence in learning outcomes.

References

Received 15 August 2005
Final revision received 19 January 2006