



## SOMOSAT: Utility of a web-based self-assessment tool in undergraduate medical education

David E. Leaf, Joseph Leo, David E. Leaf, Joseph Leo, Phillip R. Smith, Herman Yee, Arnold Stern, Pamela B. Rosenthal, Eileen B. Cahill-Gallant & Michael H. Pillinger

**To cite this article:** David E. Leaf, Joseph Leo, David E. Leaf, Joseph Leo, Phillip R. Smith, Herman Yee, Arnold Stern, Pamela B. Rosenthal, Eileen B. Cahill-Gallant & Michael H. Pillinger (2009) SOMOSAT: Utility of a web-based self-assessment tool in undergraduate medical education, Medical Teacher, 31:5, e211-e219, DOI: [10.1080/01421590802650118](https://doi.org/10.1080/01421590802650118)

**To link to this article:** <https://doi.org/10.1080/01421590802650118>



Published online: 21 Jul 2009.



Submit your article to this journal [↗](#)



Article views: 1343



View related articles [↗](#)



Citing articles: 1 View citing articles [↗](#)

## WEB PAPER

# SOMOSAT: Utility of a web-based self-assessment tool in undergraduate medical education

DAVID E. LEAF\*, JOSEPH LEO\*, PHILLIP R. SMITH, HERMAN YEE, ARNOLD STERN, PAMELA B. ROSENTHAL, EILEEN B. CAHILL-GALLANT & MICHAEL H. PILLINGER

New York University School of Medicine, USA

## Abstract

**Background:** Relatively few studies have rigorously assessed the effectiveness of computer-based self-assessment in medical education.

**Aim:** To assess whether an online self-assessment tool can be an effective adjunct to a traditional curriculum for second-year medical students.

**Methods:** The NYU School of Medicine Online Self-Assessment Tool (SOMOSAT) consists of >450 multiple-choice questions spanning disciplines of internal medicine, administered as separate modules focused on individual organ systems. Questions are coded on multiple dimensions, permitting second-year medical students to receive low-stakes, highly specific feedback regarding their knowledge and performance. Students can also review their answers to guide future study. We employed data collected during SOMOSAT operation to assess its utility and effectiveness.

**Results:** Overall, SOMOSAT accurately predicted student performance on future exams. SOMOSAT participants generally performed better than non-participants on subsequent graded course examinations ( $p < 0.05$ ). Students using SOMOSAT subsequently experienced greater improvement in areas in which they initially performed poorly, compared with those in which they initially performed well. Students reported that SOMOSAT was most helpful in filling knowledge gaps, and providing opportunities to practice exam-style questions.

**Conclusion:** The ability of SOMOSAT to enhance learning and exam performance suggests that web-based self-assessment tools can be effective adjuncts to traditional educational methods.

## Introduction

Web-based learning (WBL) is utilized with increasing frequency as a tool for medical education. The potential advantages in using internet-based educational materials are numerous and have been well documented (Cohen 1995; Chueh & Barnett 1997; Mackenzie & Greenes 1997; Zucker et al. 1998; Grundman et al. 2003; Cook et al. 2005). However, most studies evaluating the effectiveness of medical WBL have been directed at the resident level or higher, with relatively few studies assessing the utility of such an approach at the medical school level (Campbell & Johnson 1999).

One approach to stimulate learning through WBL is through the use of self-assessment questions with feedback (Brown & Manogue 2001). Surveys indicate that the question format is often preferred among both college (Gao & Lehman 2003) and medical students (Swagerty et al. 2000; Cook et al. 2001), and a recent randomized controlled study evaluating the effectiveness of online self-assessment questions among internal medicine residents showed that the question format led to higher test scores and greater user satisfaction when compared to WBL modules without questions (Cook et al. 2001).

## Practice points

The use of a web-based self-assessment tool for second-year medical students:

- Demonstrated predictive power for assessing potential performance on actual graded exams.
- Permitted medical students to identify areas of strength and weakness and provided opportunity for directed study.
- Provided students with the opportunity to practice their examination skills in a semi-controlled setting.
- Resulted in improved medical knowledge, and improved graded exam performance, compared with students who did not use the tool.
- Provided an efficient and effective method for medical students to improve their knowledge and performance.

The New York University School of Medicine curriculum for the second year of medical school includes a 7-month course in the mechanisms of disease, organized by organ systems and co-taught in an integrated manner by the

\*These authors contributed equally to the manuscript.

Correspondence: Michael H. Pillinger, Division of Rheumatology, NYU Hospital for Joint Diseases, 301 East 17th Street, New York, NY 10003, USA. Tel: 212-598-6119; fax: 212-951-3328; email: michael.pillinger@med.nyu.edu

Departments of Medicine, Pathology, Pharmacology and Obstetrics and Gynaecology. The course consists of more than 100 lectures by individual faculty experts in their specific topics, as well as small group sessions also directed by specialists in the fields being taught. While overall student evaluations of the course are typically excellent, students have frequently commented that they would prefer more opportunity to receive feedback on their performance in the course. Students have also commented that the course was not fully exploiting opportunities for computer-based learning.

To address these deficits, in 2006–2007 we created School Of Medicine Online Self-Assessment Tool (SOMOSAT), a series of online, organ-system-based self-assessment modules to provide medical students with an opportunity to assess their own knowledge and insight into specific medical areas, as well as to provide additional resources for study and learning. Here, we present a description of SOMOSAT, an assessment of its utility in improving student knowledge and performance, and the impression of the students themselves as to the usefulness of the program.

## Methods

### Description of the question bank

At present, SOMOSAT consists of 454 multiple-choice questions addressing the pathophysiology, pathology and pharmacotherapy of organ-system disease. Most of the questions were composed in a collaboration between several medical students and faculty members with the goal of creating a standardized set of questions by adhering to the following guidelines: (1) questions should each have five answer choices but only one correct answer, (2) answer choices could not include 'multiple multiples' (choices allowing the student to select more than one option (e.g. 'A and C only', 'all of the above, etc.)), (3) questions phrased in the negative (e.g. 'Which of the following is NOT...') should be avoided wherever possible, (4) questions should incorporate a clinical vignette/scenario wherever possible and test both knowledge and clinical reasoning, (5) questions should be written such that an informed student should theoretically be able to answer the question asked without recourse to the answer choices. These criteria for question writing are similar to those promulgated by the American Board of Internal Medicine and the National Board of Medical Examiners (NBME 2007). Question writers were asked to provide brief explanations for all questions, including explanations of the incorrect answer choices. In some cases, questions from a pre-existing bank of old examination questions were also included. Although these questions were vetted prior to inclusion, they did not always rigorously meet the formal criteria described above.

The SOMOSAT is structured into eight modules, each module consisting of 50–60 questions (a single module remains incomplete, with 26 questions). The SOMOSAT modules are organ-system based (circulation (heart), respiration (lungs), excretion (kidneys), digestion (gastrointestinal tract), the endocrine system, the reproductive system (including neonatal disease), the musculoskeletal system (including bone, joint and rheumatologic diseases) and the

hematologic system, and correspond directly to the individual organ systems modules taught in the NYU Mechanisms of Disease course.

Authors aimed to produce an adequate distribution of question themes within a given module. Each new question was reviewed carefully by the faculty members responsible for that topic and edited to improve it. Before entering questions into the database they were coded on multiple axes in order to provide students with informative categories for which to receive feedback. Specifically, questions in each module were concurrently coded by subtopic and discipline. For example, a question in the circulation module would be further categorized to a subtopic such as congestive heart failure, cardiomyopathy or myocardial infarction, and would also be categorized by discipline as addressing primarily pathophysiology, pathology or pharmacology. The subtopics and disciplines employed correspond directly with similar axes applied to questions in the database of graded course examinations, permitting the creators of SOMOSAT to compare performance in specific areas of the online modules with those on the graded course examinations.

### Instructions to students

Students were informed that the purpose of SOMOSAT was to provide them with a self-assessment of their strengths and weaknesses in specific areas within each organ system module in order to facilitate additional learning. A secondary goal was to provide students with the opportunity to practice answering integrative questions in an exam-style setting. An additional aim was to provide students with an opportunity to study within the body of knowledge that defined each module, including an opportunity to review the correct answers to the questions posed, as well as explanations for most of the correct as well as incorrect answer choices on the modules. However, students were explicitly instructed that the questions included in each SOMOSAT module would have no direct relation to the questions on any graded course examination, and that since one goal of each SOMOSAT module was to provide a mixture of questions addressing both core concepts and specific skills and details, mastery of the material covered by SOMOSAT should not be presumed to provide adequate preparation, in and of itself, for any graded course examination. Since the goal of SOMOSAT was to provide feedback opportunities in a 'low-stakes' environment, students were assured that all SOMOSAT scores would be kept strictly confidential, would not directly affect their grades, and would not be shared with course directors. Students were informed that SOMOSAT participation was not required, but their participation was encouraged both for their own benefit, as well as to provide the developers with sufficient feedback to permit assessment of the usefulness of the system. This study was unfunded, and designated as exempt from review by the Institutional Review Board and the Office of Clinical Trials of New York University School of Medicine.

Individual SOMOSAT modules were made available to students at, or just prior to the end of each teaching unit, and remained available until the relevant graded course examination. Since each graded course examination was a combined

assessment of two modules, as a practical matter this meant that individual SOMOSAT modules were made available to students either for several weeks, or for several days prior to their graded course examinations.

To encourage students to employ SOMOSAT as a tool for self-assessment and not merely as a question bank, the SOMOSAT software was configured such that the first time students logged onto a given module they were required to complete it in 'Quiz Mode', a rough simulation of graded exam conditions. In Quiz Mode, students were required to complete all questions within a limited period of time (2.5 min per question, or ~2.5 h for a typical module) before having access to answers or explanations (analyses discussed in this manuscript will generally refer to performance under 'Quiz Mode' conditions). While taking the quiz, students had the option of 'flagging' challenging questions that they wanted to be sure to review subsequently. Upon completion of Quiz Mode, students received a web page indicating their overall score, as well as their performance breakdown by subtopic and discipline.

The students' second login to SOMOSAT (Initial Review Mode) allowed them to pass through the questions without significant time constraint. During this second pass-through, selection of an answer choice resulted in display of the correct answer, a report of whether they had answered it correctly in 'Quiz Mode', and in most cases an explanation of the correct and incorrect answers. Questions the students had flagged were highlighted to be sure they did not skip past them. Upon a third and final pass-through (Final Review Mode), students were able to review those questions they had never answered correctly in the previous two passes. Details related to the creation of the software used to power SOMOSAT are published elsewhere (Smith 2003, 2007).

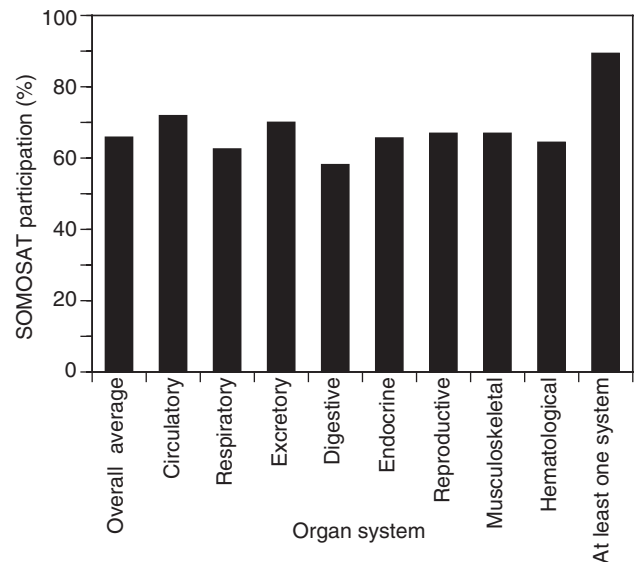
## Results

### Student participation

Overall, the average participation in any given SOMOSAT module was 66% of the class. Approximately 90% of the class participated in at least one SOMOSAT module. Participation appeared to vary little with time, with little or no drop off in overall use as the term progressed (Figure 1). For each module, student scores were observed to follow a Gaussian distribution (not shown), suggesting that SOMOSAT was able to distinguish between higher- and lower-performing students.

### The SOMOSAT performance: Correlation with graded course examination scores

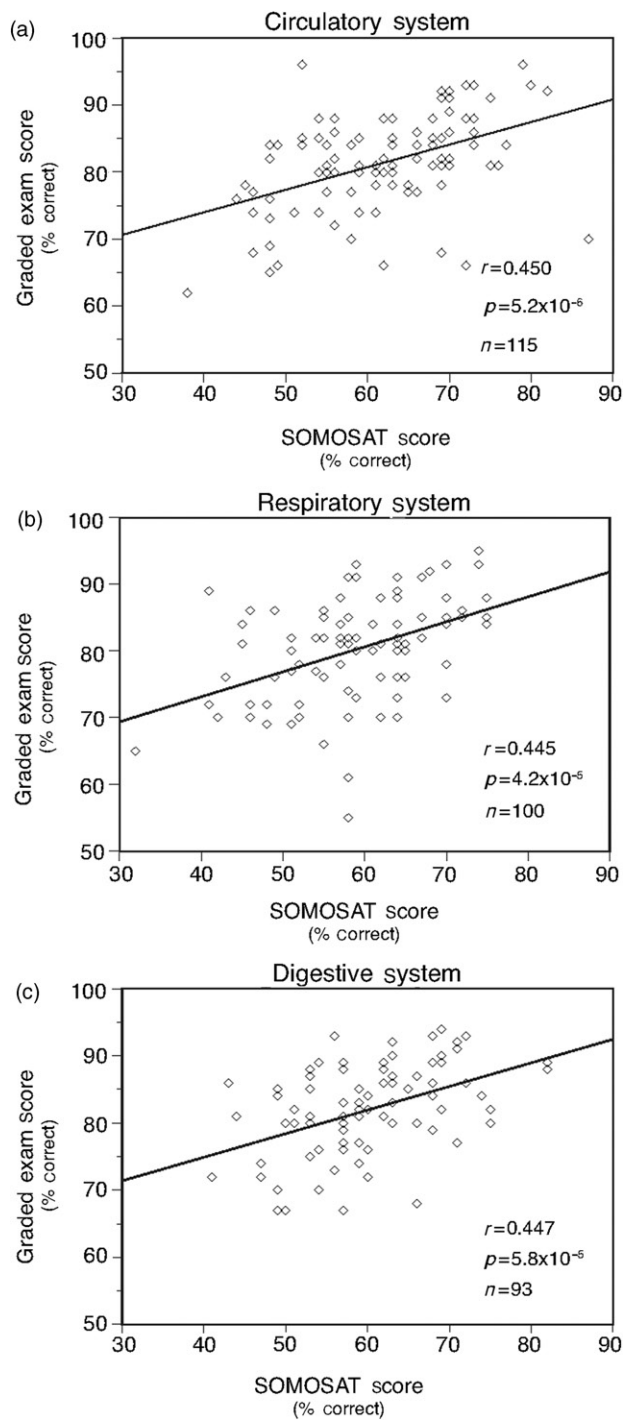
To assess the ability of the SOMOSAT modules to predict examination performance, we compared students' scores on specific SOMOSAT modules with their subsequent performance on corresponding graded course examinations. Overall, SOMOSAT scores corresponded well ( $0.4 < r < 0.5$ ) with subsequent examination scores. Figure 2 shows representative plots of graded course examination scores *versus* SOMOSAT quiz scores for the Circulation (A), Respiration (B) and Digestion (C) modules. Regression analyses of the least-



**Figure 1.** The SOMOSAT participation (% of the class) by organ system. Organ systems are presented in the chronological order in which the modules became available to students (i.e. beginning with the circulatory and ending with the haematological system).

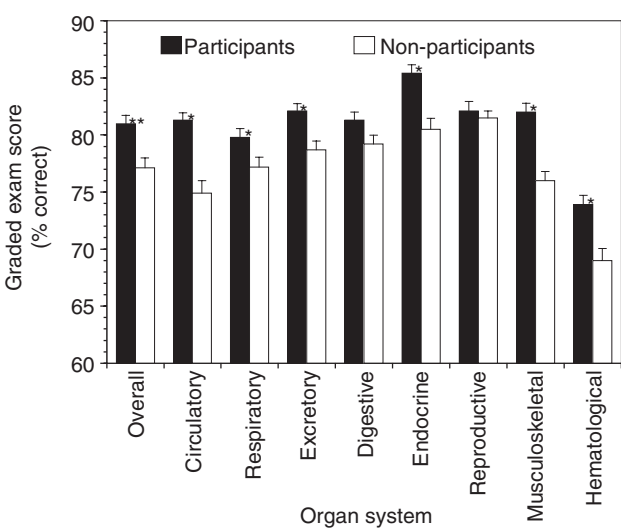
squared lines in Figure 2 indicate that the correlation between graded course examination scores and SOMOSAT scores was statistically significant ( $p < 0.001$ ) for each module. These data indicate that SOMOSAT scores provided students with reasonable insight into their overall mastery of a given organ system, and their potential performance on the graded course examination. SOMOSAT appeared to be most useful when utilized at least several days in advance of the graded course examination; students seeking to access SOMOSAT for the first time within 24 h of the graded course examination for the circulatory system unit achieved lower graded examination scores relative to students utilizing SOMOSAT earlier (not shown). Because of a concern that last-minute SOMOSAT access might be diverting students from more fundamental methods of study, on subsequent modules students were required to log into Quiz Mode at least 24 h before the graded exam. Students who fulfilled this requirement were subsequently permitted access to review the questions and answers without restriction.

The ability of SOMOSAT to provide a reasonable prediction of performance on individual subtopics and disciplines was also assessed. The circulation SOMOSAT was selected as a representative module, since that module had both the highest student participation ( $n = 115$ ) as well as the largest number of questions ( $n = 70$ ), thus conferring greater statistical power to detect differences. Regarding subtopics, regression analyses indicated that SOMOSAT performance was a significant ( $p < 0.05$ ) predictor of graded course examination performance on three out of six subtopics (congestive heart failure, myocardial infarction and valvular heart disease), and did not reach significance as a predictor of examination performance on the remaining three subtopics (arrhythmias, congenital heart disease and coronary artery disease). Regarding the individual disciplines, regression analyses



**Figure 2.** Correlations between SOMOSAT score (% correct), and performance on the graded course examinations, for the circulatory (a), respiratory (b) and digestive (c) system modules.

indicated that SOMOSAT performance was a significant ( $p<0.001$ ) predictor of examination performance for pharmacology, trended towards significance ( $p=0.09$ ) for pathophysiology, and was not significant ( $p=0.13$ ) for pathology. Overall, these data indicate that with our current banked questions SOMOSAT provides students with reasonable insight into their performance in some, but not all, subtopics and disciplines of knowledge and competence. Since the subtopics and disciplines in any SOMOSAT module constituted smaller



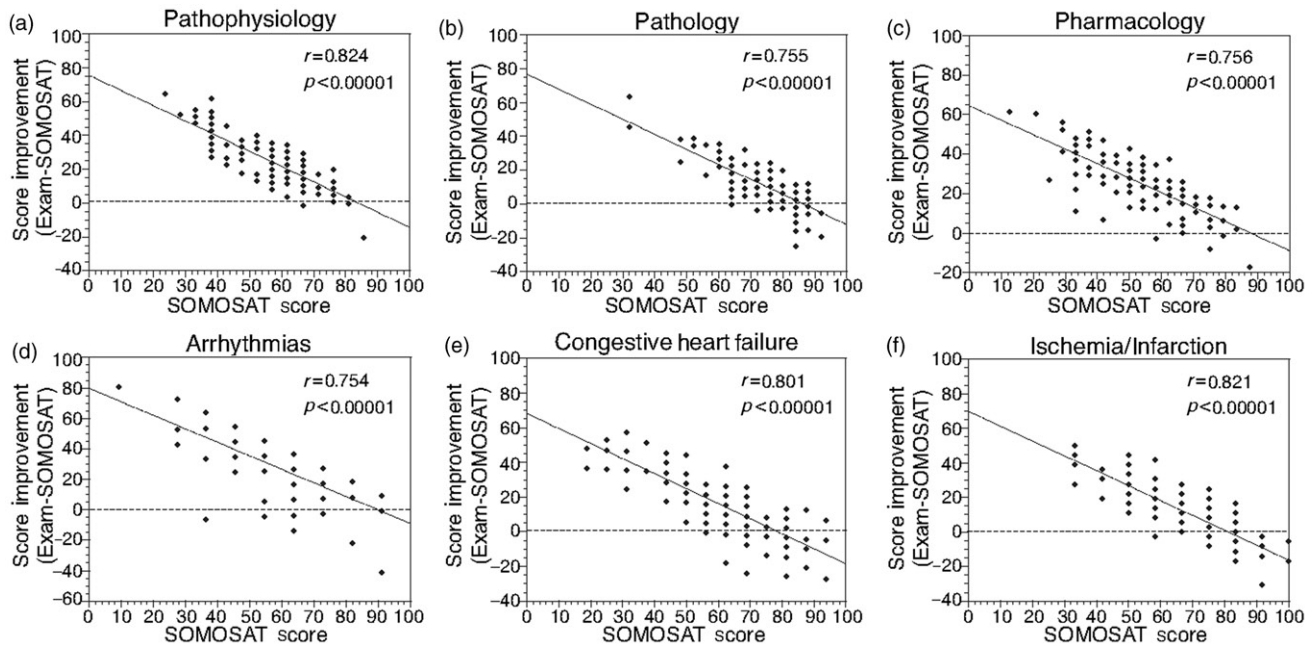
**Figure 3.** Graded Course Examination Scores (% correct) for SOMOSAT participants *vs.* non-participants for all organ system modules. (\* $p<0.05$  *vs.* SOMOSAT non-participants; \*\* $p<0.001$  *vs.* SOMOSAT non-participants overall).

numbers of questions (typically around 10) than the overall test, it is likely that more correlations would have achieved statistical significance had a larger overall test been employed.

### Utility of SOMOSAT in improving knowledge/competence

When compared with SOMOSAT scores, student scores on graded course examinations tended to be higher (Figure 2), raising the possibility that SOMOSAT might be useful in improving students' knowledge of, and competence in, specific areas. To assess this possibility, we compared the graded course examination performance of students who did, *versus* those who did not participate in SOMOSAT. Figure 3 shows mean examination scores for each module. Compared with non-participants, SOMOSAT participants performed better overall across all units, as well as on every individual unit, with statistical significance ( $p<0.05$ ) achieved for all units except the digestion and reproduction systems. To assess for the presence of any intrinsic defects in those two modules, we compared a series of performance indices for digestion and reproduction with those of the other modules. We observed no consistent differences between digestion and reproduction *versus* the other modules with regard to difficulty indices (measure of the relative difficulty of the quizzes), discrimination indices or, point biserials (measures of the ability of individual questions to predict overall performance on the quizzes), or Kuder–Richardson 20 scores (measure of internal consistency of performance across the front and back halves of each quiz). However, we did observe that, compared with the other modules, these two contained a disproportionate share of questions drawn from our older question bank, suggesting that while these modules may have performed well psychometrically, they may not have accurately assessed or communicated the current content of the course. In subsequent SOMOSAT administrations, we intend to substitute these older questions with newer ones specifically created for SOMOSAT.





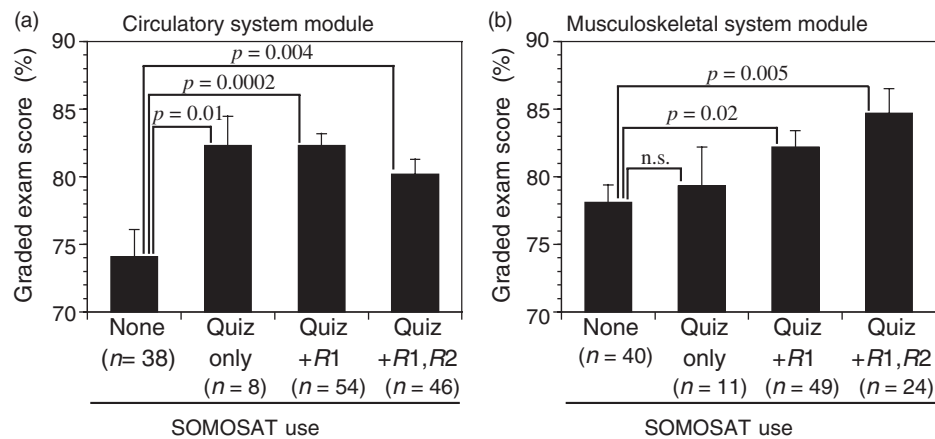
**Figure 4.** Inverse correlation between score improvement (Graded Course Examination Score minus SOMOSAT Score) *vs.* SOMOSAT score in the circulatory system module for the three disciplines of pathophysiology (a), pathology (b) and pharmacology (c) and three representative circulatory system subtopics: Arrhythmias (d), Congestive Heart Failure (e) and Ischemia/Infarction (f). ( $n=115$  for all panels, however, multiple data points are overlapping due to students receiving similar combinations of scores).

The improved performance on graded course examinations that was observed among SOMOSAT users might have been due to SOMOSAT use, but might also have been due to selection bias (i.e. students who elected to use SOMOSAT might have been a self-selected group with stronger performance characteristics). We therefore assessed the possibility of selection bias in several of the modules (respiratory, musculoskeletal and circulatory modules) in which significant differences were observed between the SOMOSAT users and the non-users. To do so, we compared the performances of SOMOSAT participant and non-participant groups on examinations in a major academic course taught earlier in the second year of NYU Medical School, the Host Defence Block. We found no statistically significant differences in Host Defence performance between the SOMOSAT users and non-users for eight of the nine modules. The only module in which a statistically significant difference was found was the circulatory module, in which the circulatory SOMOSAT users had higher host defence scores as compared with non-users ( $p=0.01$ ). However, the difference between the host defence exam scores for the two groups was relatively small ( $81 \pm 0.5\%$  vs.  $78 \pm 1.0\%$ ). These data suggest that SOMOSAT user and non-user groups were comparable in academic performance prior to SOMOSAT administration, and that any pre-existing differences between the groups were insufficient to account for differences in performance on the subsequent graded course examinations.

If SOMOSAT was capable of identifying student weaknesses, and permitting students to improve their knowledge and competence, we hypothesized that participation in SOMOSAT would result in greater improvement in subtopics and/or disciplines in which students initially performed poorly

compared with subtopics/disciplines in which they initially performed well. We therefore examined the degree of score improvement (defined as graded course examination score minus SOMOSAT score) *vs.* SOMOSAT score, for disciplines and subtopics within the circulation module. As shown in Figure 4, students experienced a significantly greater degree of improvement in graded course examination performance in those areas in which they initially performed poorly on SOMOSAT, as compared to those areas in which they initially performed well.

Since SOMOSAT users had the option of passing through the system in three distinct phases (Quiz Mode, Initial Review Mode and Final Review Mode), we further analyzed student performance in the two modules demonstrating the largest score increases on the graded course examinations (circulatory and musculoskeletal systems), to investigate which components of SOMOSAT might be most helpful in improving mastery of the material. In performing this analysis, students who failed to complete all of the Quiz Mode questions before proceeding onto Review Mode were excluded. For the circulatory module (Figure 5a), students who participated only in the Quiz Mode experienced significantly improved performance on the graded course examination compared with students who did not use any component of the circulatory SOMOSAT module. Students who used both the Quiz Mode and the Initial Review Mode or who availed themselves of all three phases of the system, continued to demonstrate improved examination scores relative to the non-users, although their scores were not significantly improved as compared to the Quiz Mode-only users. In contrast, for the Musculoskeletal SOMOSAT module (Figure 5b), students who availed themselves of the Quiz Mode only showed no

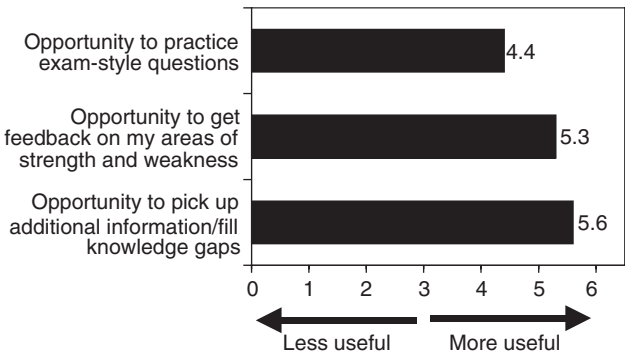


**Figure 5.** Graded Course Examination Scores as a function of the degree of participation in SOMOSAT for the circulatory (a) and musculoskeletal system (b) modules. Quiz + R1 = Quiz Mode and Initial Review Mode only; Quiz + R1 + R2 = Quiz Mode, Initial Review Mode and Final Review Mode.

significant difference over students who did not take SOMOSAT. However, students who accessed both the Quiz Mode and the Initial Review Mode, as well as students who accessed all three phases of the system, showed statistically significant increases in graded examination performance as compared to the non-users. Thus, it would appear that both the opportunity to take the Quiz and receive feedback, as well as the opportunity to subsequently review the didactic material, have the potential to improve students' knowledge and/or examination performance.

Student feedback

Finally, we were interested in analyzing feedback from students regarding the usefulness of SOMOSAT as a study tool. Survey questions were prepared and disseminated through the web-based course assessment tool already in place in the school. Response rates to individual survey questions ranged from 50% to 60% of the class. All questions were coded on a 7-point Likert scale, with results expressed as the group mean  $\pm$  SD. In each case, 1 represented 'Not At All', and 7 represented 'A Great Deal'. On balance, students responded favourably when asked whether SOMOSAT was useful in preparing for graded course examinations ( $3.9 \pm 1.4$ ), and felt that the SOMOSAT questions provided a reasonable insight into the nature of the exam-style questions ( $5.2 \pm 1.8$ ). When asked whether they used their SOMOSAT results to target areas in which they needed additional effort, students offered mixed but predominantly positive responses ( $3.92 \pm 1.5$ ). Students expressed only limited enthusiasm for the rigid structure of SOMOSAT (requiring students to progress from Quiz Mode to final review mode) ( $3.2 \pm 1.9$ ). Indeed, a number of students felt that the requirement to first utilize SOMOSAT in the Quiz Mode represented an infringement on their autonomy. When asked about the specific aspects of SOMOSAT that were most helpful in learning and studying, students were more positive, and were most enthusiastic about the opportunity to pick up additional information and fill in knowledge gaps (Figure 6). Concordant with the observed data, students felt that the most useful time to access SOMOSAT was several days before the graded course



**Figure 6.** Student feedback about the specific aspects of SOMOSAT that were most helpful in learning and studying. Responses are recorded on a Likert scale.

examination, rather than earlier or immediately before the examination date. Overall, student written comments regarding SOMOSAT were considered to be positive, with negative comments mainly reflecting frustration with the rigid structure of the SOMOSAT process. Selected representative students comments, both favourable and unfavourable, are provided in Table 1.

Discussion

The principle findings of this study are: (1) the majority of the class (roughly two-thirds in any given module) participated in SOMOSAT, (2) performance on SOMOSAT gave students reasonable insight into their potential overall performance on the upcoming graded course examinations, (3) performance on at least some SOMOSAT subtopics and disciplines gave students reasonable insight into their performance on the corresponding areas on the graded course examination, (4) SOMOSAT participants performed better overall on graded course examinations as compared with SOMOSAT non-participants, (5) students who utilized SOMOSAT demonstrated significantly greater improvement in those subtopics and disciplines in which their baseline SOMOSAT performance was initially poor, suggesting that SOMOSAT may provide

**Table 1.** Student comments regarding SOMOSAT.

Favourable	Unfavourable
<ul style="list-style-type: none"> <li>● I loved the SOMOSAT. I think the quiz mode should be required BEFORE review mode at it is.</li> <li>● I found the SOMOSAT extremely useful. It forced you to think critically about topics that you had only been attempting to memorize up until that point, which usually helped me establish whether I had any real clue as to what was going on.</li> <li>● I liked having test-style questions to practice on.</li> <li>● The breakdown by area of strengths and weaknesses was helpful.</li> </ul>	<ul style="list-style-type: none"> <li>● I don't like the format in requiring us to progress from quiz mode to final review mode. I would rather that answer and explanations be provided on the first-run (immediate review mode).</li> <li>● I did not like that you had to access SOMOSAT by a certain time or else lost privileges to use it.</li> <li>● I would have liked to have been able to access the quizzes more than three times.</li> <li>● More detailed explanations would be nice.</li> <li>● It was very frustrating when there were computer glitches and I couldn't log on anymore to go over the questions.</li> </ul>

students with feedback that permits them to focus on areas most needing improvement, (6) both the use of Quiz Mode alone as well as the use of the Review Mode options had the potential to result in improved mastery of the material and (7) students found SOMOSAT to be most helpful in filling in knowledge gaps and in providing an opportunity to practice exam-style questions, but some expressed frustration with the rigorous structure of SOMOSAT.

The present study has several limitations. First, while SOMOSAT participants performed better on graded course examinations as compared with non-participants, these groups were self-selected rather than randomized, and therefore we cannot exclude the possibility of a selection bias. For example, students who utilized SOMOSAT may have been a more highly motivated or more highly competent group compared to the non-users. To assess for such a possibility, we examined student performance on a major course taken earlier in the second year of medical school, and generally found no significant differences between SOMOSAT users and non-users at baseline; nonetheless, differences between groups may have still been present. A second limitation relates to the analysis of improvement by subtopic and discipline (Figure 4), where we found that students experienced the greatest subsequent improvement in areas in which SOMOSAT identified pre-existing weakness. To a certain extent, one might expect students to improve more in their areas of weakness, based solely on the fact that there may be more room for improvement in these areas. Similarly, one might also expect lesser improvement in subtopics and disciplines in which the baseline SOMOSAT score was high, owing to the limitations of the scale. While we cannot know the extent to which our analyses were influenced by these 'floor-and-ceiling' effects, the strength of our correlations between score improvement *versus* baseline SOMOSAT score suggest that SOMOSAT itself, by providing students with feedback and/or study opportunities, may indeed have contributed to the improved student performances. A control group of students who took the SOMOSAT quizzes, but were not made privy to their feedback evaluations, would have permitted us to address this question; however, the SOMOSAT project was designed as a teaching tool rather than a research project, and such a control group was therefore not included in our start-up process.

In a subanalysis, we observed that students who completed only the Quiz Mode (with its attendant feedback) for the

circulatory system module experienced as much improvement in their subsequent graded course examination performance as those who also took the opportunity to review the questions. In contrast, students who participated in the musculoskeletal system module only saw significant improvement in their subsequent graded course examination if they completed not only the Quiz Mode, but also at least one pass through the Review Mode. While further study is needed, these data suggest that both the ability to provide objective feedback, and the provision of review questions, may contribute to the improvement of student mastery. Since the Quiz Mode also provided students with the opportunity to practice answering questions under examination conditions, it is likely that this rehearsal process played a role in improving student scores on the circulatory system graded course examination, particularly since this examination was the first one in the course.

The SOMOSAT developers were surprised by the degree to which a portion of the class desired to bypass our self-assessment (Quiz) mode and to have direct access to the SOMOSAT questions and answers. In contrast, the developers felt strongly that the opportunity for self-assessment was the main added value which the on-line system offered, and that students accessing the questions directly would use them to learn a fraction of the material, but not necessarily the material that is included on the graded course examination. Our analyses of student performances on the Circulatory and Musculoskeletal graded course examinations after using Quiz mode only, or Quiz mode plus the additional reviews, would seem to support both points of view. Indeed, in spite of some strong feelings to the contrary, most students seemed to support access in the current manner. After reviewing both the student feedback and our performance analysis data, we have elected to continue administering SOMOSAT largely under the current model. However, we are considering providing some additional flexibility in the administration of the modules, such as giving students the opportunity to retake SOMOSAT modules a second time in Quiz Mode, and providing more flexible options for reviewing the material after the Quiz Mode is complete. We remain cautious, however, about conveying to the students any impression that mastery of the SOMOSAT modules alone is indicative of either adequate preparation for their exams, or – more importantly – an adequate fund of knowledge with which to proceed with clinical training.



Testing medical knowledge and providing useful feedback rank among the most challenging and essential aspects of medical education. WBL has many advantages over more traditional self-assessment tools, chief among them being the ability to provide immediate and detailed feedback. More broadly speaking, the use of e-learning permits the development of an online community of learners, with assessment and self-assessment as core activities (Macdonald 2004). In this regard, it is worth noting that SOMOSAT is actually part of a larger IT-based program (designated 'ALEX') at our medical school, that is building a virtual community to parallel our actual one. An additional advantage of WBL is that it allows not only for testing of factual knowledge but also for testing of cognitive processes and problem-solving, regardless of the field in question (Baker & Mayer 1999). Indeed, the questions on our SOMOSAT modules were not strictly fact-driven, but were designed to test the application of knowledge in reasoning towards appropriate conclusions.

The use of a testing program such as SOMOSAT has obvious applicability to the early years of medical training, in which a large group of students share a common synchronized curriculum. However, self-assessment programs such as SOMOSAT may be even more valuable in later years of medical training, in which individual learning experiences occur at different times and locations, and may vary widely based on available patient populations and the one-on-one nature of clinical teaching. In that setting, the use of on-line teaching and self-assessment programs might serve to enforce a common core curriculum, and alert the medical learner to the extent to which his or her experiences are successfully addressing the expected medical learning, reasoning and skills targets. Review of overall class performance may also give course directors insight into the extent to which their training programs are addressing the learning goals. In our institution, we are currently piloting a WBL and self-assessment program for medical residents on subspecialty rotations, in an attempt to provide a uniform core experience to complement the irreplaceable but inevitably hit-or-miss experience of a brief rotation on a consult service.

## Conclusion

In conclusion, we found that a web-based self-assessment tool for second-year medical students was both well-received, and an effective adjunct to more traditional educational methods. SOMOSAT appeared to enhance knowledge and competence, and allowed students to focus their attention on areas most in need of attention. While additional research is needed, we believe that WBL should be considered for incorporation with greater regularity into medical school education.

## Acknowledgements

The authors thank the following individuals for their support in developing, instituting and/or assessing SOMOSAT: Veronica Catanese, Julianne Chase, Mel Rosenfeld, Kenneth Berger, David Goldfarb, David Green, Jeffrey Greenberg, Loren Greene, Melanie Maslow, Khush Mittal, Margaret Nachtigall, Valerie Peck, Michael Poles, Bruce Raphael, Marc

Triola and Paul Tunick. Phillip R. Smith gratefully acknowledges salary support from the NIH/NCRR (RR00096).

**Declaration of interest:** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

## Notes on contributors

EILEEN CAHILL-GALLANT, MPA, is the Manager of the Office of Medical Education at New York University School of Medicine. She has served as coordinator for the second-year basic science curriculum, and is currently involved in student assessment, program evaluation and curriculum review.

DAVID LEAF, MD, a recent graduate of NYU School of Medicine, is currently a resident in internal medicine at Columbia University Medical Center. He served as lead writer and Student Director for the SOMOSAT project and hopes to pursue a career in academic nephrology.

JOSEPH LEO, BA, is a Database Analyst for the NYUSOM Office of Medical Education. He is responsible for the creation and ongoing maintenance of the NYU Exam Database which is used to store exam questions and images, generate exams, and track statistical information. Mr Leo helped to design and deploy the SOMOSAT system.

MICHAEL PILLINGER, MD, is Associate Professor of Medicine and Pharmacology, and served as Course Director for Pathophysiology at NYU. He is Director of Rheumatology Training, and Associate Director of the NYU Masters Program in Translational Investigation. He currently chairs the National Fellows' In-training Examination for the American College of Rheumatology.

PAMELA ROSENTHAL, MD, is Assistant Professor of Medicine at NYU School of Medicine and is the current Course Director for Pathophysiology for the undergraduate medical program. She is also the Site Director for Rheumatology Education at Bellevue Hospital Center.

PHILLIP R. SMITH, PhD, MD, is an Associate Professor of Cell Biology who has been teaching medical students for over 25 years. He has had an ongoing interest in improving the computer-based delivery of teaching and self-evaluation materials and wrote the software used to deliver the quizzes described here.

ARNOLD STERN, MD, PhD, is Professor of Pharmacology, and Assistant Dean for Extramural Education Programs and NYU School of Medicine. He is also the Course Director for Medical Pharmacology. Dr Stern was the Editor of the 8th and 9th editions of Pharmacology: Pre Test Self Assessment and Review.

HERMAN YEE, PhD MD, is Associate Professor of Pathology. He has been actively involved in undergraduate medical education for the 12 years, including serving as the Pathology Course Director. In the evolution from traditional to an integrative curriculum, he was an early adopter of computerized multimedia, including digitization of histopathologic slides.

## References

- Baker EL, Mayer RE. 1999. Computer-based assessment of problem solving. *Comput Hum Behav* 15(3-4):269-282.
- Brown G, Manogue M. 2001. AMEE Medical Education Guide No. 22: Refreshing lecturing: A guide for teachers. *Med Teach* 23:231-244.
- Campbell JK, Johnson C. 1999. Trend spotting: Fashions in medical education. *BMJ* 318:1272-1275.
- Chueh H, Barnett GO. 1997. 'Just-in-time' clinical information. *Acad Med* 72:512-517.
- Cohen J. 1995. Educating physicians in cyberspace. *Acad Med* 70:698.
- Cook D, Dupras DM, Thompson WG, Pankratz VS. 2005. Web-based learning in residents' continuity clinics: A randomized, controlled trial. *Acad Med* 80:90-97.
- Cook D, Thompson WG, Thomas KG, Thomas MR, Pankratz VS. 2001. Impact of self-assessment questions and learning styles in web-based learning: A randomized, controlled, crossover trial. *Acad Med* 81:231-238.

- Gao T, Lehman JD. 2003. The effects of different levels of interaction on the achievement and motivational perceptions of college students in a web-based learning environment. *J Interact Learn Res* 14:367–386.
- Grundman J, Wigton RS, Nickol D. 2003. A controlled trial of an interactive, web-based virtual reality program for teaching physical diagnosis skills to medical students. *Acad Med* 78:205–301.
- Macdonald J. 2004. Developing competent e-learners: The role of assessment. *Assess Eval Higher Educ* 29(2):215–226.
- Mackenzie J, Greenes RA. 1997. The World Wide Web: Redefining medical education. *JAMA* 278:1785–1786.
- National Board of Medical Examiners (NBME). 2007. Constructing Written Test Questions For the Basic and Clinical Sciences. Available online at: <[http://www.nbme.org/PDF/ItemWriting\\_2003/2003IWGwhole.pdf](http://www.nbme.org/PDF/ItemWriting_2003/2003IWGwhole.pdf)> (Accessed 1 July 2008).
- Smith P. 2003. Software implementing web-based study questions for medical students. *Int J E-Learning* April–June:29–34.
- Smith P. 2007. KEYSEARCH. Available online at: <[http://library.med.nyu.edu/epub\\_support/keysearch-manual.html](http://library.med.nyu.edu/epub_support/keysearch-manual.html)> (Accessed 1 July 2008).
- Swagerty Jr D, Studenski S, Laird R, Rigler S. 2000. A case-oriented web-based curriculum in geriatrics for third-year medical students. *J Am Geriatr Soc* 48:1507–1512.
- Zucker S, White JA, Fabri PJ, Konsari LS. 1998. Instructional intranets in graduate medical education. *Acad Med* 73:1072–1075.