Clinical Reasoning Web-based Prototypic Module for Tutors Teaching 5th Grade Medical Students: A Pilot Randomized Study

Gerald H. Stein, MD, Hironobu Tokunaga, MD, Hirotaka Ando, MD, Mikako Obika, MD, PhD, Tomoko Miyoshi, MD, Yasuharu Tokuda, MD, MPH, Yoshinori Noguchi, MD, MPH, Mitsuyo Kinjo, MD, MPH, Shun Kohsaka, MD, Hitoshi Honda, MD, Yuka Kitano, MD, Hidetaka Kitazono, MD, Hitomi Kataoka, MD, PhD, and Hidekazu Terasawa, MD

Background: At present clinical reasoning skills are not systematically taught in Japanese medical universities. We developed a prototypic preliminary module for clinical tutors to introduce clinical reasoning to Japanese medical students. We hypothesized that tutored medical students would outperform self-study students.

Author for correspondence: Gerald H. Stein, MD
Department of Medicine, College of Medicine, University of Florida, Gainesville, Florida, USA
E-mail: jerrydoc@ufl.edu
Received for publication 18 September 2014 and accepted in revised form 26 November 2014
© 2015 The Japan Primary Care Association
**Method:** Using the web-based Sequential Question and Answer test that rewarded history and differential diagnosis as proxies for clinical reasoning, we compared the pre and posttest scores of 12 randomized fifth grade tutored students at two universities during four tutor-led 1.5-hour web-based seminars using a structured syllabus to 12 randomized self-study students.

**Results:** The tutored and self-study groups’ pretest scores were statistically similar at about 40 out of 100 weighted correct points. The tutored students’ posttest scores were 62 points, significantly greater ($p = 0.007$) than the pretest mean 42 points, compared to the self-study students’ posttest scores of 52 points, significantly greater ($p = 0.012$) than pretest mean 40 points. The difference between the two posttest groups was of borderline statistical significance ($p = 0.08$).

**Conclusions:** We successfully assessed a prototypic module for tutors to introduce clinical reasoning to Japanese medical students. The tutored students achieved higher scores than the self-study students. Further research is needed to exploit the potential of our modular clinical reasoning system.

**Keywords:** clinical reasoning, clinical tutors, medical education

**INTRODUCTION**
Learning clinical reasoning is difficult for all medical students. Teaching clinical reasoning is difficult also for clinical instructors. For Japanese medical students (MS), the task is even more challenging given their many years of the traditional mode of learning by lecture, the memorization way of education from primary through university education. Silence and listening is rewarded over speaking and questioning. As such, Japanese MS enter their clinical clerkships without a strong and long emphasis in general problem-solving skills. Furthermore, Japanese MS enter residencies frequently lacking confidence in their clinical skills. Moreover, many Japanese medical educators are aware lectures do not improve clinical reasoning. In addition, their clinical instructors are hampered by a similar lack of exposure to problem solving skills; hence the instructors fail to tutor the students in problem solving skills. Most Western and Japanese medical schools have been using clinical problem based learning (PBL) exercises or its modifications for over 20 years. Japanese medical universities tend to restrict PBL to the introductory pre-clinical curriculum. In 2004, the Japanese Ministry of Health, Labor and Welfare instituted several initiatives to improve clinical training, such as a two years postgraduate residency for medical licensure and an American style ‘match’ residency selection process. However, implementation of these directives has been slow. In addition a report of Japanese malpractice cases highlighted the need for leaders of medical schools and teaching hospitals to improve their teaching of clinical reasoning skills; the majority of these claims were adjudicated against physicians for their cognitive and judgment errors. To accelerate these improvements, a group of international medical educators developed a demonstration pilot module which is the subject of this report. The goals of this project were: 1) to produce a prototypical module for Japanese medical educators, and by implication, other Asian medical educator, to introduce clinical reasoning to MS entering their clinical clerkships; 2) and to compare the module’s usefulness by assessing the change of fifth year Japanese MS tutored by Japanese United States trained general internists, compared to a self-study group. This is the first report of an international team to improve clinical reasoning with Japanese MS entering their clinical training using Web-based seminars (webinars).

**METHODS**
**Study design**
We designed the study to measure the usefulness of structured four tutored webinars (web-based seminars)
with geographically diverse fifth grade MS, led by American trained Japanese tutors; we made no attempt to study the use of the Internet in this project. Additionally, we contrasted these tutored students with a self-study MS group, the traditional way of learning in Japanese medical universities. We administered via the Internet a pretest and a posttest, four-five weeks later, the Sequential Question and Answer (SQA), to all medical students. We coached the tutors to use the slideshow syllabus as a script, written to promote group discussions at each of the 1.5-hour sessions. We made no attempt to ensure the tutors followed the traditional Socratic teach, question and discuss method. Because this study was a pilot project to demonstrate the feasibility of an innovative teaching module, we did not estimate validity or reliability for the SQA tests and syllabus.

Participants

Medical students

From 07/2011 to 09/2011 inclusive, study coordinators at two Japanese national medical universities disseminated information about the project to all 100 fifth year MS at each of the two participating national Japanese universities, designated university 1 and university 2. All MS were eligible for this randomized comparison-group, non-blinded prospective four-week study. The authors instituted a multi-level randomization process during the four weeks of 09/2011. Specifically, every 5th MS was offered the opportunity to join this protocol from the Japanese alphabetical class roll of 100 MS. Of these 20 randomly selected students, the first 12, again in alphabetical order, were invited; all invited MS volunteered freely to participate with none refusing. Lastly, using alphabetically order, the 12 MS, six from each of the two universities, were randomized, to the tutored group, and 12 MS, six from each of the two universities, were randomized to the self-study comparison group.

Tutors

The authors invited two Japanese-English bilingual native Japanese physicians as volunteer tutors. They completed at least three years of United States general internal medicine residency and were involved in clinical education in their different Japanese medical universities. The authors met once with the tutors in a virtual, that is, an online classroom to review ways to encourage student discussion and review the syllabus before any virtual student tutoring started. Study coordinators developed and implemented the randomization process, coordinated the tutors’ and MS virtual classroom assignments, as well as the pre and post tutoring SQA tests and questionnaire.

Settings

Students participated in their own apartments or university-provided classrooms, the later chosen for better Internet access. Tutors conducted their four sessions in their own homes with each session lasting 1.5 hours.

Technical aspects

The chief study coordinator contracted with the Japanese affiliate of WebEx Communications Inc. for the tutoring platform, a Web-based proprietary secure browser [www.webex.co.jp], which combined real time audio-video interaction, ‘white board,’ instant message texting, and multi-media viewing, all simultaneously available and recordable for post analysis. Each of the two virtual classrooms consisted of one tutor and six medical students, three of each from the two universities. All seven persons in each classroom could see and hear one another simultaneously.

Instruments

Measures

SQA tests

The actively practicing clinician authors developed an instrument, the SQA, to measure the change within the two protocol groups, tutored and self-study students, and between these two randomized groups, respectively. Its details have been described previously. Briefly, the Japanese language pre and posttest SQA consisted of eight questions from a common pneumonia case narrative, with clinical information, followed by a question requiring short text written answers, with immediate correct answers appearing, more clinical information and the subsequent question, progressing in the usual clinical narrative from the clinical history through to the physical examination, laboratory data to the differential diagnosis (Figure 1). The two SQA pre
Figure 1. Browser images of SQA question 1 (Q.1), remaining time 23:53", correct answers in grey, and MS answers (A.1) in black. Questions 2 (Q.2) and 3 (Q.3), remaining time 17:01", correct answers in grey, the MS answers (A.2, A.3) in black. New data generates question 3. Not shown is scrollability to review data. For illustrative purposes only; actual SQA tests were in Japanese language and in color. Reproduced with permission of Imagine Labo, Japan.
and posttests and cases were modest variations of common bacterial pneumonias: community acquired pneumonia from *Streptococcus pneumoniae* (the SQA pretest), and cancerous-related lymph node obstructive pneumonia (the SQA posttest). The tutors led the 3rd and 4th sessions with the syllabus having a third variation of a common pneumonia, namely, reactive pulmonary tuberculosis.

After completion of the pretest by all 24 MS, the coordinators announced the clinical topic of fever, infection and pneumonia to all 24 MS participants and encouraged all students to study fever, infection and pneumonia from standard Japanese textbooks written for Japanese MS. At this time the students were randomized to tutored and self-study groups. Only the tutored MS were shown the syllabus. At the conclusion of the four-week tutoring sessions, all MS took the posttest SQA; the posttest case was a slightly different pneumonia narrative. The two tutors were informed that the MS were required to take the SQA pre and posttests; however, they were blinded to its content for the duration of the entire study.

**Tutored syllabus**

The actively practicing clinician authors developed a four-session slideshow syllabus as a script, containing basic clinical knowledge of fever, infection, and pneumonia, planned for the first two sessions, and a clinical simple pneumonia case with usual clinical narrative of medical history, physical examination, laboratory data, imaging, and differential diagnosis for the final two sessions (Figure 2). The 48 slides contained color-coded and italicized fonts which prompted the tutors to stimulate students’ discussion. These authors developed and wrote the four-session syllabus, initially in the English language and then translated into the Japanese language with many revisions until consensus was reached. The final Japanese version contained animations and cartoons; the animation consisted of questions for tutor-led group discussion followed by the correct answers.

**Analysis**

**Data analysis**

The authors analyzed the participants’ SQA pre and posttest answers to the eight questions after translated from Japanese into English. The authors developed a manual master scoring sheet, based on a weighted scoring system totaling 100 points, expressed as percent of correct answers (Table 1). More weighted points were awarded for the one medical history answers (24 points) and two differential diagnoses answers (42 points). The authors completed the scoring without access to the identity of the participants and universities. Differences were resolved by consensus.

**Statistical analysis**

Changes in pre- and posttest scores were calculated from mean values. If the distributions of scores between the groups were similar, we presented the raw data by using mean, standard deviation and range. Continuous variables were compared using paired Student’s t-test for paired data and unpaired Student’s t-test for unpaired data if the data could be considered normally distributed based on the Shapiro-Wilk W test for normal data. Dichotomous variables were compared using Fisher’s exact test. Two-tailed p-value less than 0.05 was considered as statistical significance. STATA version 16 (College Station, TX, USA) was used for all analyses.

**Ethical considerations**

Ethical approval for educational studies was obtained from the hospital ethics committees of the participating facilities of Okayama University Medical School, Fukui University, Faculty of Medical Sciences, and Tsukuba University, School of Medicine. The study adhered to national Japanese medical student research participants’ requirements and to the International Ethical Guidelines of the Declaration of Helsinki. All 24 MS signed the approved Japanese language uniform informed consent form with a study coordinator who answered all students’ questions. The consent form stated the objectives of the protocol, that they were not identifiable from the data, that they were volunteering, without coercion, could withdraw any time without prejudice and that every session was recorded for analysis.

**RESULTS**

The SQA program captured all of the participants’ SQA text answers on a downloadable spreadsheet for data retrieval and analysis. All 24 MS completed the
Japanese language web-based pretest program, the SQA pretest, during the first and second weeks of November, 2011; all 24 MS completed the SQA posttest during the first and second weeks of December, 2011. No student withdrew. Their age range was 24–27 years old. Nine participants (38%) were female students. The two groups’ ages and sexes were matched by the randomization without intention. The tutored and self-study groups’ pretest scores were similar, achieving a mean 40 points each (Figure 3). Shapiro-Wilk W test for normal data indicated that pre and post scores were normally distributed with p-values of 0.69 and 0.71, respectively. With the two universities’ scores combined, the tutored MS posttest scores achieved a mean 62 points which was significantly greater (p = 0.007) than pretest mean 42 points. The self-study MS achieved posttest scores with mean 52 points which was significantly greater (p = 0.0121) than pretest mean 40 points. The difference between the two posttest groups did not attain statistical significance (p = 0.08).

Table 2 shows the MS average points by each university. Distributions of scores were similar based on the score ranges. The tutored MS posttest scores significantly increased at university 2 (p = 0.005). The self-study MS scores also increased scores in university

<table>
<thead>
<tr>
<th>Problem List from History, Physical Examination, Laboratory, Imaging</th>
<th>Assessment - Diagnosis “Do Not Miss” Danger</th>
<th>Plan Studies</th>
<th>Plan treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Cough, wet(sputum)</td>
<td>*R/O pneumonia, CAP, bacterial, atypical pneumonia</td>
<td>Sputum &amp; blood cultures, Gram stain</td>
<td>Antibiotics after cultures, stains</td>
</tr>
<tr>
<td>2 Night sweats Fever</td>
<td>R/O sepsis, Tb</td>
<td>*AFB stain, sputum culture, QuantiFERON</td>
<td>Pending AFB stains</td>
</tr>
<tr>
<td>3 Smoking</td>
<td>R/O COPD, Lung cancer</td>
<td>Sputum cytology</td>
<td>Pending cytology, anti-smoking campaign</td>
</tr>
<tr>
<td>4 Involuntary weight loss</td>
<td>R/O Lung cancer</td>
<td>Sputum cytology, diet review</td>
<td>Pending cytology</td>
</tr>
<tr>
<td>5 *SIRS ↑ BT! ↑ HR! ↑ RR</td>
<td>R/O pneumonia, CAP, sepsis, AMI, pulmonary infarction</td>
<td>ECG, Trop, CK-MB, repeat 4-6 hours; D-dimer, Leg Doppler</td>
<td>Pending studies and repeat vital signs</td>
</tr>
<tr>
<td>6 ↓ SpO2</td>
<td>Respiratory failure, R/O *ACS</td>
<td>See # 5 above</td>
<td>Oxygen, aspirin, consider heparin</td>
</tr>
<tr>
<td>7 ↓ Hb/Hct/platelets</td>
<td>Anemia,thrombocytopenia S/O pancytopenia S/O from pneumonia S/O from Tb</td>
<td>Repeat CBC, pending bone marrow aspirate, biopsy, culture</td>
<td>Pending studies</td>
</tr>
<tr>
<td>8 Chest x ray right upper lobe patchy infiltration</td>
<td>R/O usual Gram + cocci, Klebsiella, Tb</td>
<td>See # 1 &amp; 2 above</td>
<td>See # 1 &amp; 2 above</td>
</tr>
</tbody>
</table>

*R/O Rule Out, CAP Community Acquired Pneumonia, AFB Acid Fast Bacilli, SIRS Systemic Inflammatory Response Syndrome BT Body Temperature, HR Heart Rate, RR Respiratory Rate, S/O Suspicion Of, ACS Acute Coronary Syndrome

- Problem Based Learning: List Every Problem
- Link Every Problem to an Assessment/Diagnosis
- Every Assessment/Diagnosis Must Have a Plan
- Every Plan Must Have an Assessment/Diagnosis

Figure 2. This is image of slide # 44, the summary table of the clinical case. Syllabus consisted of 48 slides which were the script for tutors to lead webinars. For illustrative purposes only; actual syllabus in Japanese language and in color.
Table 1a. Scoring table of the Sequential Question and Answer (SQA) pretest and posttest for the 24 participating students.

<table>
<thead>
<tr>
<th>Question Category</th>
<th>Score System Correct Answers</th>
<th>Maximum Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Medical History</td>
<td>Chills; chest pain; shortness of breath; color of sputum; some indication of amount of sputum, blood or red color of sputum; amount of tobacco use each day and number of years used; any prior history of Tb or Tb exposure; weight loss, night sweats, prior history of pneumonia, lung disease, impact on daily activities; Rx questions, associated symptoms, others sick home or work, seen by community doctor, better/worse</td>
<td>Total 12 possible answers Weight × 2 = 24 points</td>
</tr>
<tr>
<td>2 Physical Examination-1st step</td>
<td>Vital Signs, Body Temperature (BT), Heart Rate (HR), or pulse rate, Respiratory Rate (RR), Blood Pressure (BP) and SpO2 (Room air or on O2)</td>
<td>Total 6 possible answers Weight × 1 = 6 points</td>
</tr>
<tr>
<td>3 Abnormal Vital Signs</td>
<td>BT, HR, RR, SpO2</td>
<td>Total 4 possible answers Weight × 2 = 8 points</td>
</tr>
<tr>
<td>4 Focused Physical Examination</td>
<td>Chest, lungs, thorax, heart</td>
<td>Total 3 possible answers Weight × 1 = 3 points</td>
</tr>
<tr>
<td>5 Physical examination modalities</td>
<td>Percussion, auscultation, inspection, palpation</td>
<td>Total 4 possible answers Weight × 0.5 = 2 points</td>
</tr>
<tr>
<td>6 Preliminary differential diagnosis</td>
<td>Pneumonia, sepsis, consolidation pneumonia, pleural effusion, tumor, Systemic Inflammatory Response Syndrome (SIRS) organism: bacterial, viral; COPD, Tb, cancer</td>
<td>Total 9 possible answers Weight × 2 = 18 points</td>
</tr>
<tr>
<td>7 Studies</td>
<td>Blood CBC, sugar, electrolytes, renal, liver, blood cultures, sputum smear &amp; culture routine &amp; Tb, Tb skin test, sputum cytology, urinalysis, ABG, ECG, Chest XP</td>
<td>Total 15 possible answers Weight × 1 = 15 points</td>
</tr>
<tr>
<td>8 Final differential diagnosis</td>
<td>Sepsis, Strep. pneumoniae, pneumonia, consolidation/right lung pneumonia, Systemic Inflammatory Response Syndrome (SIRS), COPD</td>
<td>Total 5 possible answers Weight × 4.8 = 24 points</td>
</tr>
</tbody>
</table>

Total Maximum Points 100

Notes for pretest scoring table
1. History: added correct answers from students’ answers not on original correct answers list.
   Minus 1 – point for very wrong answer: Auscultate Lungs
4. Thorax = 3 points
   Minus 1 – point for very wrong answer: JVD, lymph nodes, fingers, and oral cavity
   Note most students failed to understand this question
6. Preliminary differential diagnosis
   Minus 1 – point for every wrong answer: lung edema, heart failure, sarcoid, bronchitis ARDS, DVT Pulmonary emboli, Interstitial pulmonary fibrosis, ACS
7. Studies: If ‘blood test’ = 5 point
   Minus 1 – point for very wrong answers: Chest CT (this is ER patient), ECHO, PFT’s, PET Scan, throat culture, examine fingers, lymph nodes, JVD
8. Final differential diagnosis
   If ‘pneumococcal pneumonia’ was answer = 3 points; if sepsis = 2 points, if COPD with exacerbation = 1 point
Table 1b. Scoring table of the Sequential Question and Answer (SQA) pretest and posttest for the 24 participating students.

<table>
<thead>
<tr>
<th>Question</th>
<th>Category</th>
<th>Score System Correct Answers</th>
<th>Maximum Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Medical History</td>
<td>Chills; chest pain; shortness of breath; color of sputum; some indication of amount of sputum, blood or red color of sputum; amount of tobacco use each day and number of years used; any prior history of Tb or Tb exposure; weight loss, night sweats, prior history of pneumonia, lung disease, impact on daily activities; Rx questions, associated symptoms, others sick home or work, seen by community doctor, better/worse</td>
<td>Total 12 possible answers Weight × 2 = 24 points</td>
</tr>
<tr>
<td>2</td>
<td>Physical Examination-1st step</td>
<td>Vital Signs, Body Temperature (BT), Heart Rate (HR), or pulse rate, Respiratory Rate (RR), Blood Pressure (BP) and SpO2 (Room air or on O2)</td>
<td>Total 6 possible answers Weight × 1 = 6 points</td>
</tr>
<tr>
<td>3</td>
<td>Abnormal Vital Signs</td>
<td>BT, HR, RR, SpO2</td>
<td>Total 4 possible answers Weight × 2 = 8 points</td>
</tr>
<tr>
<td>4</td>
<td>Focused Physical Examination</td>
<td>Chest, lungs, thorax, heart</td>
<td>Total 3 possible answers Weight × 1 = 3 points</td>
</tr>
<tr>
<td>5</td>
<td>Physical examination modalities</td>
<td>Percussion, auscultation, inspection, palpation</td>
<td>Total 4 possible answers Weight × 0.5 = 2 points</td>
</tr>
<tr>
<td>6</td>
<td>Preliminary differential diagnosis</td>
<td>Pneumonia, sepsis, consolidation pneumonia, pleural effusion, tumor, Systemic Inflammatory Response Syndrome (SIRS) organism: bacterial, viral; COPD, Tb, cancer</td>
<td>Total 9 possible answers Weight × 2 = 18 points</td>
</tr>
<tr>
<td>7</td>
<td>Studies</td>
<td>Blood CBC, sugar, electrolytes, renal, liver, blood cultures, sputum smear &amp; culture routine &amp; Tb, Tb skin test, sputum cytology, urinalysis, urine Ag Strep pneum. &amp; Legionella, ABG, ECG, Chest XP</td>
<td>Total 15 possible answers Weight × 1 = 15 points</td>
</tr>
<tr>
<td>8</td>
<td>Final differential diagnosis</td>
<td>Sepsis, Strep. pneumoniae, pneumonia, consolidation/right lung pneumonia, Systemic Inflammatory Response Syndrome (SIRS), COPD, lung cancer, hilar lymph node swelling, post bronchial obstructive pneumonia</td>
<td>Total 5 possible answers Weight × 4.8 = 24 points</td>
</tr>
</tbody>
</table>

Total Points 100

Notes for posttest scoring table
1. History: added correct answers from students’ answers not on original correct answers list.
   Minus 1 – point for very wrong answer: Auscultate Lungs
4. Thorax = 3 points
   Minus 1 – point for very wrong answer: JVD, lymph nodes, fingers, oral cavity
6. Preliminary differential diagnosis
   Minus 1 – point for every wrong answer: lung edema, heart failure, sarcoid, bronchitis ARDS, DVT Pulmonary emboli, Interstitial pulmonary fibrosis, ACS
7. Studies: If ‘blood test’ = 5 point
   Minus 1 – point for very wrong answers: Chest CT (this is ER patient), ECHO, PFT’s, PET Scan, throat culture, examine fingers, lymph nodes, JVD
8. Final differential diagnosis
   If ‘pneumococcal pneumonia’ was answer = 3 points; if sepsis = 2 points, if COPD with exacerbation = 1 point
   Minus 1 – point if cancer omitted since cytology + for cancer cells
(Japanese translation re-confirmed)

General Medicine 2015, vol. 16, no. 1
2 with statistical significance (\(p = 0.01\)). There were no statistical differences between posttest scores for the tutored MS at university 1 and 2 (\(p = 0.53\)), and no statistical differences between the posttest points for self-study MS at university 1 and 2 (\(p = 0.38\)). There was a large statistical difference between the posttest points for tutored MS by tutor 1 compared to tutor 2 (\(p = 0.009\)).

In an analysis of the posttest laboratory results question, 14 of the 24 MS (59%) answered a diagnosis of cancer from the listed cancer cell positive sputum cytology laboratory data. Five of the 12 tutored MS mentioned cancer (42%) and 9 of the 12 self-study MS mentioned cancer (75%); there was no statistical difference between the tutored and self-study groups listing cancer (\(p = 0.214\)).

In a post-hoc analysis using the customary passing score of 70%, none for the 24 MS had a pre test score that reached or surpassed this level. Only four of the 12 tutored students achieved the posttest 70% level. One tutored MS at University 1 (\(n = 6\)) achieved the 70% threshold (17%) and three tutored MS at university 2 (\(n = 6\)) achieved the threshold (50%). None of the self-study MS had a post test score of 70% or greater. For the tutors, tutor 1 had one MS reach the 70% level and tutor 2 had three MS reach this level.

Minor brief technical audio and video distortions occurred during the webinar tutored sessions that had no impact on the tutors or MS. All sessions were recorded for precise content analysis at a later time. The content analysis results have been published previously.16

DISCUSSION

We have developed the first prototypical module to introduce clinical reasoning to Japanese MS. Our module contained an animated syllabus for tutors to teach information and a clinical case to use this information to introduce clinical reasoning, and an innovative evaluation instrument, the SQA; the SQA test provides guidance for organizing the clinical narrative, teaches disease information and assesses clinical reasoning.17

Our data showed that the posttest SQA scores increased within the tutored group and between the tutored and self-study groups. Moreover, the tutored groups’ score increased more than the self-study comparison group. However this increase in tutored MS scores was not as large as anticipated. For example using an arbitrary customary correct weighted posttest answer score of 70%, only 25% of the tutored students achieved this ‘passing score,’ whereas none of the self-study students

Table 2. Mean medical students’ (MS) points for the 2 universities and tutored or self-study groups.

<table>
<thead>
<tr>
<th>Average MS points</th>
<th>Pretest mean</th>
<th>Pretest SD</th>
<th>Pretest range</th>
<th>Posttest mean</th>
<th>Posttest SD</th>
<th>Posttest range</th>
<th>p values*</th>
</tr>
</thead>
<tbody>
<tr>
<td>University 1 Tutored</td>
<td>42</td>
<td>7.9</td>
<td>29–51</td>
<td>59</td>
<td>14</td>
<td>36–79</td>
<td>0.07</td>
</tr>
<tr>
<td>University 1 Self-study</td>
<td>40</td>
<td>10</td>
<td>25–49</td>
<td>49</td>
<td>9.1</td>
<td>38–59</td>
<td>0.25</td>
</tr>
<tr>
<td>University 2 Tutored</td>
<td>42</td>
<td>8.6</td>
<td>32–57</td>
<td>65</td>
<td>15</td>
<td>41–82</td>
<td>0.005</td>
</tr>
<tr>
<td>University 2 Self-study</td>
<td>40</td>
<td>16</td>
<td>22–65</td>
<td>55</td>
<td>13</td>
<td>33–69</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Notes: Each group \(n = 6\); points are weighted percent correct; p values shown are between pre and posttests; *Paired Student’s t-test, statistical significance \(p < 0.05\). SD = standard deviation.
reached this score.

Our module contains two components for successful implementation of an introductory clinical reasoning curriculum: 1) an animated syllabus designed for student-tutor interactive seminars, to stimulate student discussion and reasoning, containing basic and clinical science information, and a clinical case applying the newly acquired information to an actual clinical patient case to learn the clinical narrative and differential diagnosis, a proxy for clinical reasoning, and 2) the SQA, a teaching and assessment instrument. The SQA has the important elements of teaching as well as assessing. Furthermore, the SQA can be written in most international languages, a feature that enhances its global portability, especially for traditional medical schools struggling to reduce the lecture formats and replace them with modules that train clinical reasoning. Differences between university 1 and university 2 for both tutored and self-study groups did not reach statistical significance implying uniformity between the universities’ educational systems.

We successfully demonstrated the feasibility of Internet technology to link Japanese language tutors with geographically distant, dispersed Japanese language MS for learning clinical problem solving. Our current study expanded our prior reports suggesting that current webinar technology may be broaden to link medical universities with clinical reasoning curriculum to those universities lacking such courses. We developed this module out of the established need for Japanese clinical instructors to have a template to teach MS basic clinical skills in their own language. There are few reports of such Japanese teaching modules. For example, Kita et al reported a Japanese clinical reasoning module for MS based on script theory requiring MS to apply 300 formulas. Their module appeared overly cumbersome compared to our direct methods. We recognized the difficulties of training instructors to teach clinical reasoning. MS entering their first year of residency express their lack of confidence in their clinical skills. Entering residency is not the ideal site for this remediation, given the sudden requirements for patient responsibility and little time for reflective learning. Furthermore, although our use of web-based distant learning technology is not new, our module does permit local facilities and distant affiliated sites access to benefit students, residents and tutors for direct clinical reasoning training. Many progressive medical educators are shifting from lectures to stimulating clinical narratives.

Lastly, our module is suitable for numerous research opportunities especially international studies promoting problem-solving skills in traditional societies. There is an important need to expand studies providing improvements for training clear logical clinical reasoning in a traditional culture.

LIMITATIONS

Our study design and module development had several limiting aspects. We performed no systematic study to understand the factors causing the change in posttest scores. Although the tutored MS scores increased more than that of the self-studied students, some of this difference reflects the practice effect of repeating the SQA test; the self-study group’s posttest SQA scores similarly relate to the practice effect. Moreover the low posttest scores for the tutored group, 60% correct, reflects the inefficiency of the tutors’ efforts as documented in our content analysis study: they spoke, that is, lectured, an average of 90% of the time. Improved coaching of tutors may be necessary to change the traditional habit of lecturing to students. But clinically trained coaches are a rarity in Japan. To compensate for the relative paucity of teaching faculty, Hayashi et al reported their successful experience with tutorless sessions. Initial experiences with tutorless clinical session have their usefulness as a supplement for stimulating student interest. However, tutors with experience to teach and with clinical experience offer the highest level of qualities to direct students into the realms of clinical reasoning.

Further, despite showing an abnormal test result of cancer cells in sputum cytology, only 59% of all students correctly used this positive data to suggest a cancer diagnosis. Carelessness or lack of training to use clinical data might be responsible, suggesting easily taught remediation. Incorporation of a detailed problem list into the syllabus might have obviated this diagnostic reasoning deficiency (Figure 2).
Although undefined teaching styles may be a factor in the differences between the increase in the posttest student scores of tutor 1 compared to tutor 2, more likely was the modest increased for students’ response times and their markers of advanced reasoning responses.\textsuperscript{16} We made no attempt to ensure the tutors followed the traditional Socratic teach, question and discuss method. In this regard, content analysis of the tutors’ responses might be useful. We had expected the tutors to apply their American residency educational training to their assigned students. As mentioned above, this did not occur. For our module to generalize to other Japanese or Asian medical universities, much more group discussion coaching of the tutors is needed. Similarly we anticipated that the content analysis of the students’ responses as assessed by the captured four audio and video seminars might reveal trends in clinical reasoning development. However our data failed to show this. We have reported previously the possible reasons for this disappointing outcome with suggested improvements.\textsuperscript{16}

The SQA has not been validated or reliability tested because this project was only a feasibility study. We previously reported the limitations and suggestions to improve the SQA.\textsuperscript{15} We developed the SQA test requiring short text answers to better approximate the clinical reasoning of clinicians; physicians do not think in multiple-choice responses to their patients’ problems. Measurements of clinical reasoning have not been standardized. Our claim about measuring clinical reasoning is based on the students’ listings of their differential diagnosis; a student’s listing of differential diagnosis has been previously reported as a surrogate marker for clinical reasoning since no consensus exists for its definition.\textsuperscript{18} Measurements of clinical reasoning require further studies. We did not assess reflective reasoning or diagnostic justifications to better understand clinical reasoning.\textsuperscript{34,35}

**CONCLUSIONS**

In summary, we demonstrated the feasibility of a new prototypical teaching module for tutors to introduce clinical reasoning to medical students unaccustomed to clinical problem solving themselves. We assessed the usefulness of our unique module via an innovative short text-answers instrument, the SQA, incorporating a problem based clinical narrative. Compared to the pretest scores, the tutored MS achieved higher SQA posttest scores than the self-study students. We encourage clinical tutors in every discipline of medical education, from clinic-oriented ambulatory care to women’s health, from inpatient-oriented ambulatory care to urology, to develop body organ, disease, and symptom related modules, and to use our modular concept for active teaching clinical reasoning.

**Conflict of interest**

The authors declare that they have no conflict of interests.

**Acknowledgements**

The Japanese Ministry of Education, Culture, Sports, Science and Technology, Grant 23590870 funded this project.

We thank the fifth year medical student participants from the University of Fukui, Faculty of Medical Sciences, Japan, and the Okayama University Medical School, Japan, for their participation. Medical student helpers deserve out special thanks: from Fukui University, Yuko Aki, Daiki Yokokawa, and Mizuki Nakano; from Okayama University, Hiroki Kobayashi, Hiroaki Takahashi, and Natsumi Matsuoka; and from Tsukuba University, Ayano Furuya. Makoto Aoki, MD, Joel Branch, MD, Taro Shimizu, MD, MPH, Mitsu-masa Kishimoto, MD, Kuniyuki Endo, MD, Miho Bautista, MD, MPH, Yukari Takata, BA, PhD, and Yumi Okuyama, MD, assisted with several aspects of the project for which we are grateful.

**References**

4 Khoo KH: Implementation of problem-based learning in Asian medical schools and students’
28 Prober CG, Heath C: Lecture halls without