

# Using cloud-based mobile technology for assessment of competencies among medical students

Gary S. Ferenchick and David Solomon

Department of Internal Medicine, Division of General Medicine, Michigan State University, East Lansing, MI, USA

## ABSTRACT

Valid, direct observation of medical student competency in clinical settings remains challenging and limits the opportunity to promote performance-based student advancement. The rationale for direct observation is to ascertain that students have acquired the core clinical competencies needed to care for patients. Too often student observation results in highly variable evaluations which are skewed by factors other than the student's actual performance. Among the barriers to effective direct observation and assessment include the lack of effective tools and strategies for assuring that transparent standards are used for judging clinical competency in authentic clinical settings. We developed a web-based content management system under the name, *Just in Time Medicine* (JIT), to address many of these issues. The goals of JIT were fourfold: First, to create a self-service interface allowing faculty with average computing skills to author customizable content and criterion-based assessment tools displayable on internet enabled devices, including mobile devices; second, to create an assessment and feedback tool capable of capturing learner progress related to hundreds of clinical skills; third, to enable easy access and utilization of these tools by faculty for learner assessment in authentic clinical settings as a means of *just in time* faculty development; fourth, to create a permanent record of the trainees' observed skills useful for both learner and program evaluation. From July 2010 through October 2012, we implemented a JIT enabled clinical evaluation exercise (CEX) among 367 third year internal medicine students. Observers (attending physicians and residents) performed CEX assessments using JIT to guide and document their observations, record their time observing and providing feedback to the students, and their overall satisfaction. Inter-rater reliability and validity were assessed with 17 observers who viewed six videotaped student-patient encounters and by measuring the correlation between student CEX scores and their scores on subsequent standardized-patient OSCE exams. A total of 3567 CEXs were completed by 516 observers. The average number of evaluations per student was 9.7 ( $\pm 1.8$  SD) and the average number of CEXs completed per observer was 6.9 ( $\pm 15.8$  SD). Observers spent less than 10 min on 43–50% of the CEXs and 68.6% on feedback sessions. A majority of observers (92%) reported satisfaction with the CEX. Inter-rater reliability was measured at 0.69 among all observers viewing the videotapes and these ratings adequately discriminated competent from non-competent performance. The measured CEX grades correlated with subsequent student performance on an end-of-year OSCE. We conclude that the use of JIT is feasible in capturing discrete

Submitted 26 May 2013  
Accepted 29 August 2013  
Published 17 September 2013

Corresponding author  
Gary S. Ferenchick,  
Gary.ferenchick@ht.msu.edu

Academic editor  
Harry Hochheiser

Additional Information and  
Declarations can be found on  
page 12

DOI 10.7717/peerj.164

© Copyright  
2013 Ferenchick and Solomon

Distributed under  
Creative Commons CC-BY 3.0

**OPEN ACCESS**

clinical performance data with a high degree of user satisfaction. Our embedded checklists had adequate inter-rater reliability and concurrent and predictive validity.

**Subjects** Internal Medicine, Human-Computer Interaction, Science and Medical Education

**Keywords** Educational technology, Educational measurement, Medical students, Smart phones, Competency based assessment, Direct observation, Medical faculty, Clinical competence, iPhone, miniCEX

## INTRODUCTION

The assessment of the clinical competence of a medical student is challenging. A competency is, "...an *observable ability* of a health professional related to a specific activity that integrates knowledge, skills, values, and attitudes. Since they are observable, *they can be measured and assessed*". Although seemingly straightforward, competency based education is of limited usefulness in guiding the design and implementation of educational experiences if they are not tied to specific learning objectives ([Whitcomb, 2004](#)). Additionally, learning objectives are of limited usefulness if they are not available to students and faculty when interacting with patients. Finally, observation and assessment help neither students nor patients if they are not captured and documented in a way that facilitates learner specific plans for improvement and excellence. We present a generalizable initiative that makes national curricula functional in local learning environments and improves, and simplifies, observation based assessments and performance-based data tracking for faculty and learners.

## MATERIALS & METHODS

### Content manager

We developed a mobile, cloud-based application called *Just in Time Medicine* (JIT) that functions effectively on smart phones, tablets and laptop computers. The mobile application is supported by a self-service web-based content management system designed with the explicit aim of enabling users with average computing skills to build their own customizable content, including criterion-based checklists that can then be delivered to any internet enabled device such as a smart phone or tablet.

For this project, we utilized nineteen core training problems from the nationally validated Clerkship Directors in Internal Medicine (CDIM) curriculum and combined these training problems with the observable competencies of communication skills, history taking and physical examination to create problem and task specific checklists. For each assessment, the software calculates the students' performance by determining the percentage of all potential items performed correctly, and an algorithm generated grade of "not done/unsatisfactory", "needs improvement" or "well done" is calculated depending on the percentage of items performed correctly. In general, if a student achieved 80% of the expected items correctly they received a "well done" grade; performing 30–79% of the expected items resulted in a "needs improvement" grade, and <30% an "unsatisfactory".

Just In Time Medicine
Assessments
Logs
Curriculum

## Assessment Admin - Groups

Organizational Unit:

|        | Problem               | Competency            | Combo Name        |        |
|--------|-----------------------|-----------------------|-------------------|--------|
| Select | Abdominal Pain        | Communication Skills  | Case Pres         | Delete |
| Select | Abdominal Pain        | History Taking Skills | Abd Pain: History | Delete |
| Select | Abdominal Pain        | Physical Exam Skills  | Physical Exam     | Delete |
| Select | Altered Mental Status | Communication Skills  | Case Pres         | Delete |
| Select | Altered Mental Status | History Taking Skills | History           | Delete |
| Select | Altered Mental Status | Physical Exam Skills  | Physical Exam     | Delete |
| Select | Anemia                | Communication Skills  | Case Pres         | Delete |
| Select | Anemia                | History Taking Skills | History           | Delete |
| Select | Anemia                | Physical Exam Skills  | Abnormalities     | Delete |
| Select | Anemia                | Physical Exam Skills  | Physical          | Delete |
| Select | Back pain             | Communication Skills  | Case Pres         | Delete |
| Select | Back pain             | Communication Skills  | Communication     | Delete |
| Select | Back pain             | History Taking Skills | History           | Delete |
| Select | Back pain             | Physical Exam Skills  | Abnormalities     | Delete |
| Select | Back pain             | Physical Exam Skills  | Physical Exam     | Delete |

Add New Combo
| 2 3 4 5

**Figure 1 Step 1. Content manager for development of assessment tools.** Using simple interfaces, faculty adds content (e.g., the problem *altered mental status*) and the specific competency to be assessed (e.g., *history taking*).

Figures 1 and 2 present screen shots for the process of building checklists using our content manager for the problem *altered mental status* and the competency *history taking*. Additionally, Figs. 3 and 4 show how the assessment tools are displayed on the user's device. Figures 5–7 show the permanent cloud-based reporting options associated with individual assessments. A fully functional version of JIT can be accessed at: [www.justintimedicine.com/mobile](http://www.justintimedicine.com/mobile); log in username is `testuser@journal.com`, and the password is `test`. To access examples of Cloud-based performance reporting, go to [www.justintimedicine.com](http://www.justintimedicine.com); username: `testadministrator@journal.com` and password: `test`.

### Goals and hypotheses

In introducing JIT in our clerkship, we hypothesized that JIT would: (1) facilitate the direct observation and provision of feedback to trainees on their clinical competencies; (2) generally be accepted by faculty; (3) provide a means for recording the observations of trainee performance, and (4) possess adequate reliability and validity.

The screenshot displays a web-based content management system interface. At the top, there are tabs for 'Details', 'Line Items', 'Grading Tiers', and 'Taxonomies'. Below these is a table with the following data:

|          | Line Item Name                | Ordinal |        |
|----------|-------------------------------|---------|--------|
| Selected | Dr-Pt Comm / Characterizes Sx | 1       | Delete |
| Select   | Assoc Sx's & Delirium         | 2       | Delete |
| Select   | Past Hx / Habits              | 3       | Delete |
| Select   | Medications                   | 4       | Delete |

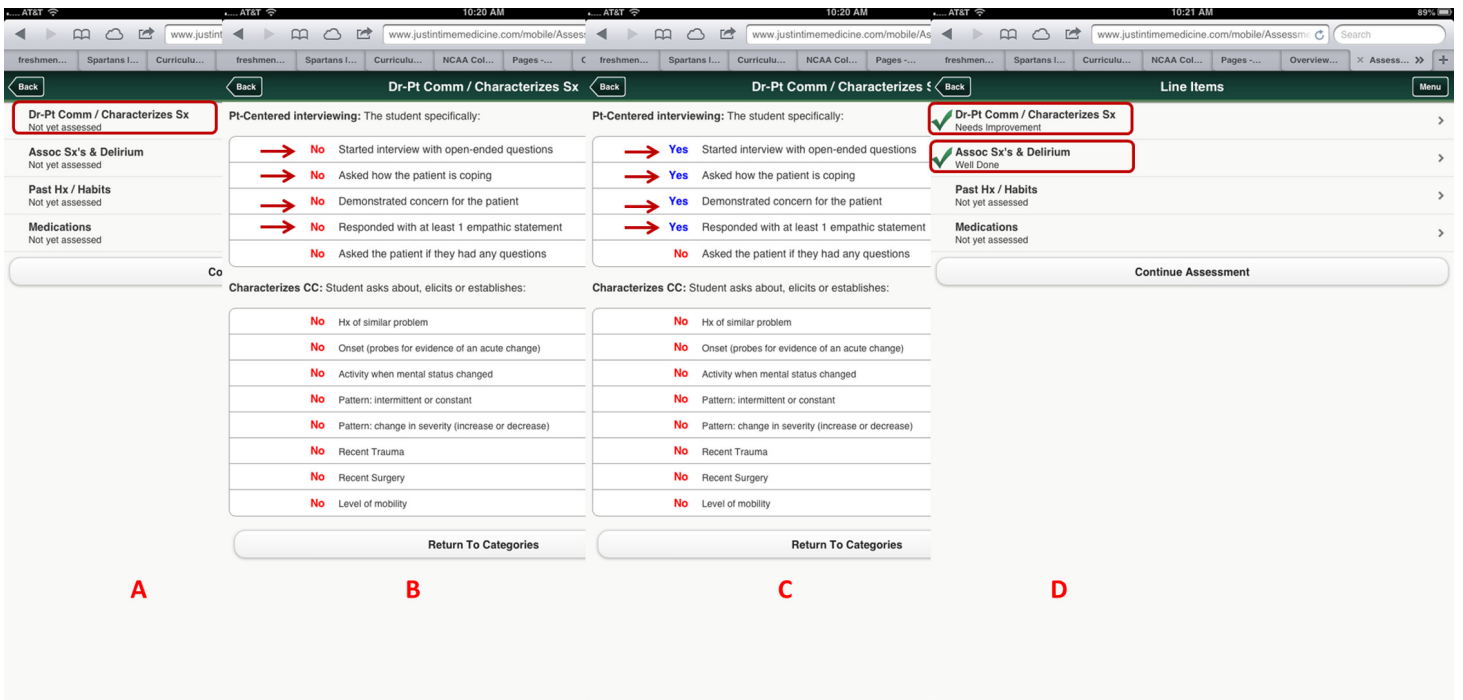
Below the table is a button labeled 'Add New Line Item'. The 'Details' view for the selected item shows the following information:

- Line Item Name:** Dr-Pt Comm / Characterizes Sx
- Ordinal:** 1
- Checkmark Type:** Yes/No
- Description:**
  - Pt-Centered interviewing:** The student specifically:
    - Started interview with open-ended questions
    - Asked how the patient is coping
    - Demonstrated concern for the patient
    - Responded with at least 1 empathic statement
      - Asked the patient if they had any questions
  - Characterizes CC:** Student asks about, elicits or establishes:
    - Hx of similar problem
    - Onset (probes for evidence of an acute change)
    - Activity when mental status changed
    - Pattern: intermittent or constant
    - Pattern: change in severity (increase or decrease)
    - Recent Trauma

**Figure 2** Step 2. Content manager for development of assessment tools. Using the self-service web-based content management system, faculty then adds assessment questions reflecting specific criterion-based outcomes (e.g., *The student started the interview with open-ended questions*).

### Setting

The College of Human Medicine (CHM) at Michigan State University is a community-based medical school with clinical training in 7 communities throughout Michigan. Between July 2010 and October 2012 we implemented JIT as an integral part of the Internal Medicine Clerkship among 367 students. Each student was required to complete ten directly observed clinical evaluation exercises (CEXs) with real patients in authentic clinical settings. A CEX is a short (generally <20 min) directly observed trainee – patient interaction (e.g., history-taking, examination, counseling, etc.); faculty observes, rates,



**Figure 3** Criterion-based assessment for altered mental status and history-taking as displayed on the mobile device for use anytime and anywhere. A displays how the specific checklist is accessed on the device; B displays the criterion-based tasks, which are defaulted to No and change to Yes (C) once the task is completed by the learner. D displays the algorithm generated grade.

and provides written comments on the interaction. Students received an orientation to the CEX application and were required to become familiar with the software. Evaluators (attending faculty and residents) received an email on the importance of direct observation and the basic functionality of the CEX application.

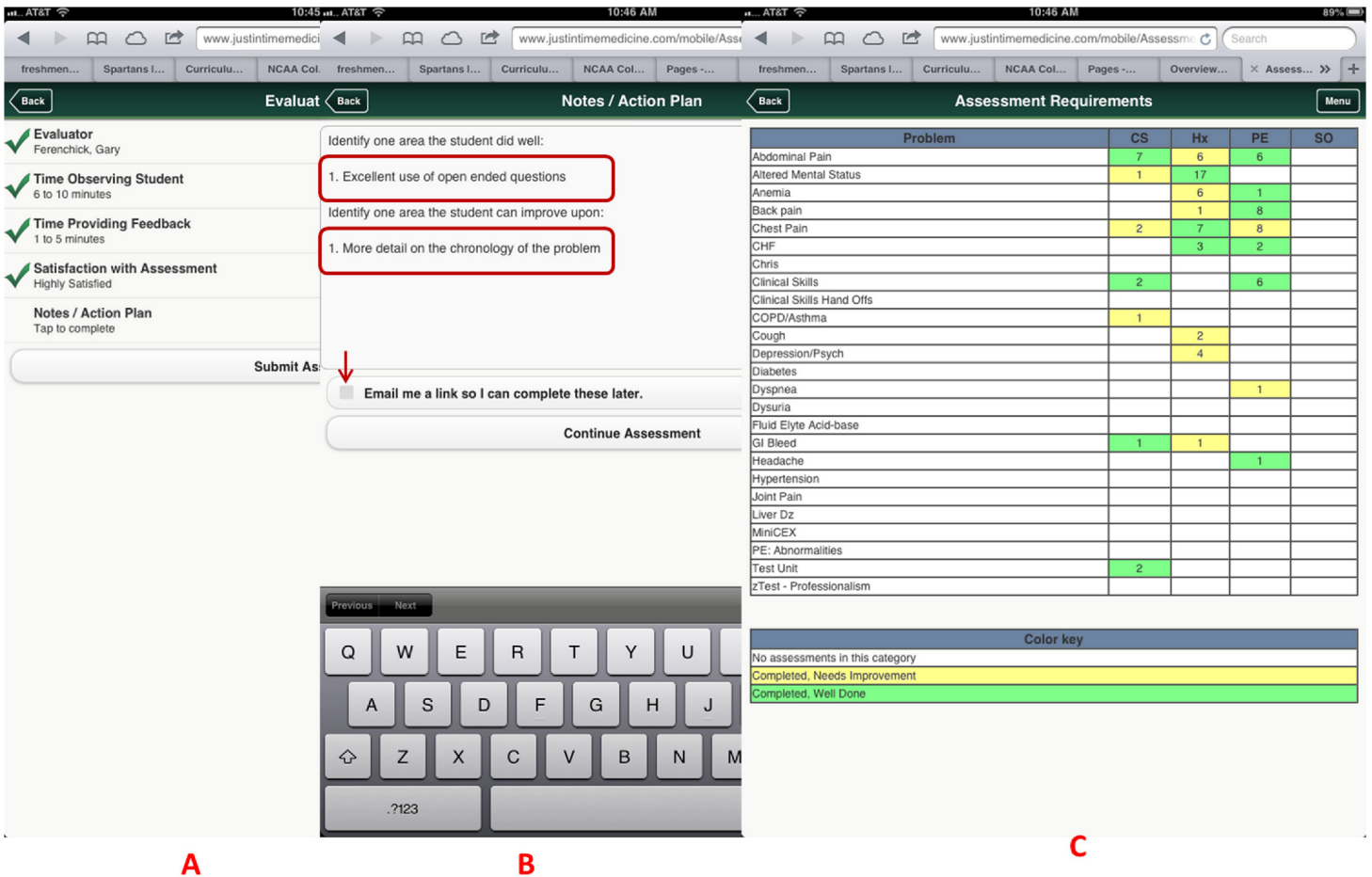
In general, students chose the patient, problem and competency upon which to be assessed. At the time of the assessment, students handed their mobile device, with the checklists displayed, for evaluator use during the assessed interaction. A total of 516 evaluators subsequently used JIT to guide their observations and assessments of students interacting with patients.

### Data collection

We collected the following data: the specific training problems and competencies observed and assessed by the evaluators, the grades associated with the observation and descriptive data from faculty on the use of JIT. Descriptive data was collected from the faculty via “pull-down” menus located on the last screen of each assessment. A screen shot of the interface is displayed in Fig. 4.

### Reliability and validity assessments

A group of 17 evaluators, 9 internal medicine residents and 8 general internist faculty members viewed and rated six scripted videotaped encounters using JIT. Each case was scripted for both satisfactory and unsatisfactory performance. These cases have been



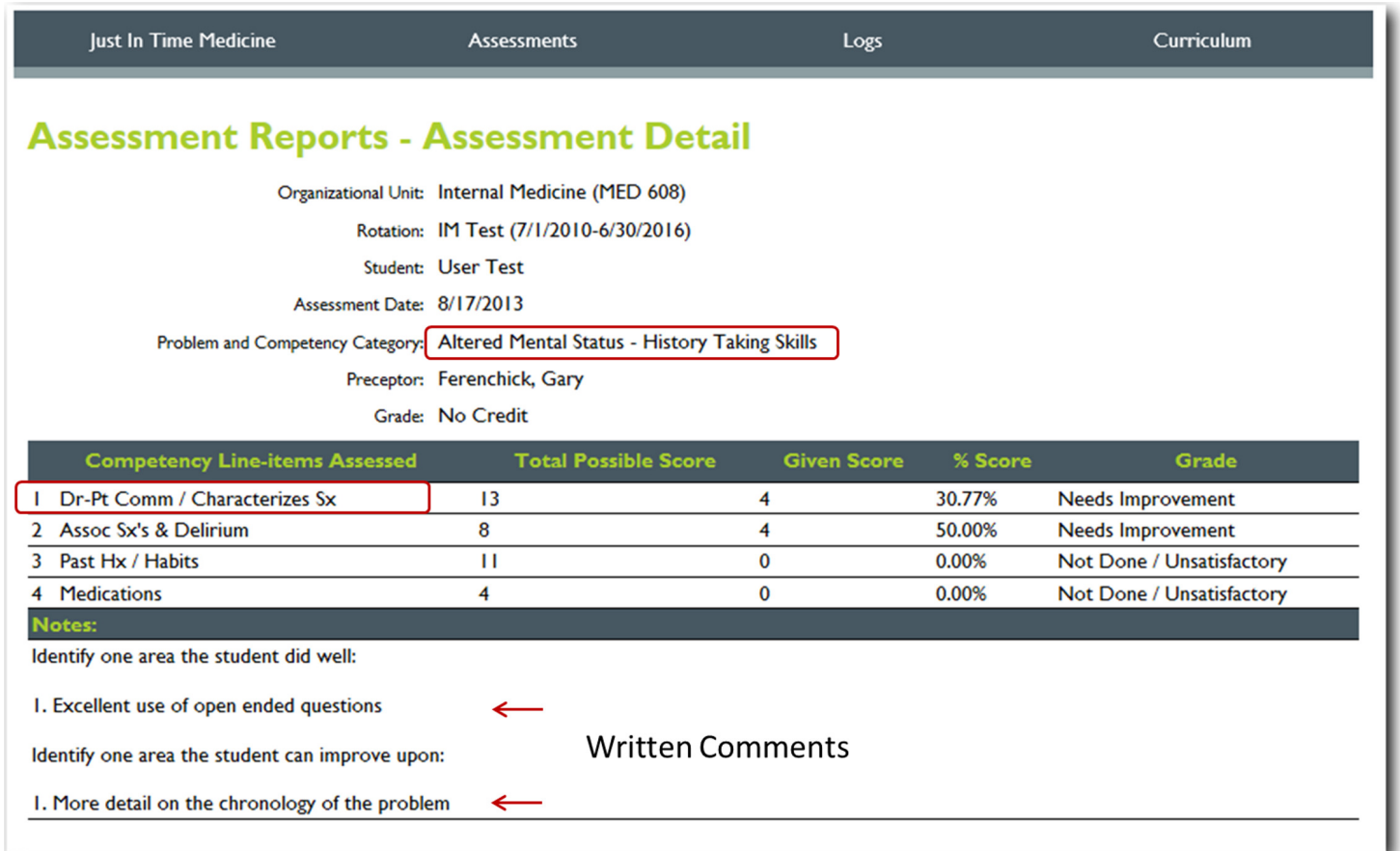
**Figure 4** Evaluator information is collected using simple interfaces on the device after the assessment is completed, including open-ended qualitative comments. Faculty enters information concerning their observation (A), and their feedback and action plans (B). A color coded competency registry is displayed on the learner's device (C). Note in B, the evaluator has the option to have an email link sent to him/her to complete the qualitative assessment at a later time. All evaluations become part of the learner's cloud-based permanent record.

previously validated by Holmboe as representing levels of competence which range from unequivocally poor to satisfactory (Holmboe, Hawkins & Huot, 2004). The sample of raters reflected the number we could reasonably obtain given our small general internal medicine faculty and residency program. We felt it was adequate to provide a stable estimate of the inter-rater reliability of the assessment process. We calculated the inter-rater reliability using a formula developed by Ebel and implemented using software developed by one of the authors (Ebel, 1951; Solomon, 2004). All other statistical analyses were performed with SPSS version 21.

## RESULTS

### Number and types of evaluations

Five hundred sixteen evaluators used the application to assess 367 students for a total of 3567 separate assessments. The number of CEX's completed per student was 9.7 ( $\pm 1.8$ ) and the average number of CEX's completed per faculty was 6.9 ( $\pm 15.8$ ). The average



**Figure 5 Detailed cloud-based reporting options.** One of the web-based permanent records of the students' performance; displaying the item(s) assessed, the percentage of potential items correctly performed, and algorithm generated grade and evaluators written comments on the learners performance (note all of these features are editable, based upon the users' needs).

number of training problems a student was assessed on was 6.7; of the three competency domains of communication skills, history taking, and physical examination 68% of the students had at least one evaluation in each of the three categories.

In terms of the grades, time variables and satisfaction, ~83% of the encounters were associated with a "well done" grade, and on average students were credited with performing ~86% of the items correctly (Fig. 8). Between 43–50% of the CEX assessments took <10 min as estimated by the faculty, and in ~69% of the encounters feedback was estimated to occur in less than 10 min. In 92% of the encounters, faculty rated that they were either satisfied or highly satisfied with the CEX.

The estimated inter-rater reliability of a single rater observing the videotaped encounters was 0.69 (slightly higher for faculty at 0.74 vs. residents at .64). In judging the same simulated patient case scripted to be satisfactory and non-satisfactory, the residents and faculty using JIT discriminated between the satisfactory and non-satisfactory performance. The mean number of items checked for the videotapes scripted for unsatisfactory

| Organizational Unit: Internal Medicine (MED 608)  |                      |             |         |                           |
|---|----------------------|-------------|---------|---------------------------|
| Rotation: IM Test (7/1/2010-6/30/2016)  |                      |             |         |                           |
| Student: User Test  |                      |             |         |                           |
| Assessment Date: 8/17/2013  |                      |             |         |                           |
| Problem and Competency Category: <b>Altered Mental Status - History Taking Skills</b>   |                      |             |         |                           |
| Preceptor: Ferenchick, Gary   |                      |             |         |                           |
| Grade: No Credit  |                      |             |         |                           |
| Competency Line-items Assessed  | Total Possible Score | Given Score | % Score | Grade                     |
| 1 Dr-Pt Comm / Characterizes Sx   | 13                   | 4           | 30.77%  | Needs Improvement         |
| <ul style="list-style-type: none"> <li>Yes Started interview with open-ended questions</li> <li>Yes Asked how the patient is coping</li> <li>Yes Demonstrated concern for the patient</li> <li>Yes Responded with at least 1 empathic statement</li> <li>No Asked the patient if they had any questions</li> <li>No Hx of similar problem</li> <li>No Onset (probes for evidence of an acute change)</li> <li>No Activity when mental status changed</li> <li>No Pattern: intermittent or constant</li> <li>No Pattern: change in severity (increase or decrease)</li> <li>No Recent Trauma</li> <li>No Recent Surgery</li> <li>No Level of mobility</li> </ul> |                      |             |         |                           |
| 2 Assoc Sx's & Delirium   | 8                    | 4           | 50.00%  | Needs Improvement         |
| 3 Past Hx / Habits  | 11                   | 0           | 0.00%   | Not Done / Unsatisfactory |
| 4 Medications   | 4                    | 0           | 0.00%   | Not Done / Unsatisfactory |
| <b>Notes:</b>   |                      |             |         |                           |
| Identify one area the student did well:   |                      |             |         |                           |
| 1. Excellent use of open ended questions  |                      |             |         |                           |
| Identify one area the student can improve upon:   |                      |             |         |                           |
| 1. More detail on the chronology of the problem   |                      |             |         |                           |

**Figure 6** JIT detailed cloud-based reporting options. With the click of a hyperlink, a detailed list of all the items that were either performed or not by the student is displayed.

performance was 35% vs. 59% for those scripted for more satisfactory performance. We believe this provides evidence supporting the construct validity of JIT.

To assess predictive validity, we calculated a Pearson product moment correlation between “gateway” performance assessment examinations taken by 282 students at the end of their third year required clerkships with the CEX assessments obtained by JIT. There was a small (but statistically significant  $0.144, p = .008$ ) correlation between students’ CEX scores and communications skills in the gateway performance assessment exam.

## DISCUSSION

Although national learning objectives have been published for all core clerkships, their usefulness for assessing learning outcomes has been limited. As an example, the core



## Taxonomy Assessment Reports - Summary

Organizational Unit: IM GME Test

Taxonomy: IM Milestone &gt; Milestone Subcompetency Taxonomy

Rotation: Test (12/8/2012-12/8/2014)

Trainee: Test Resident9

Hide Rows Without Scores: 

B

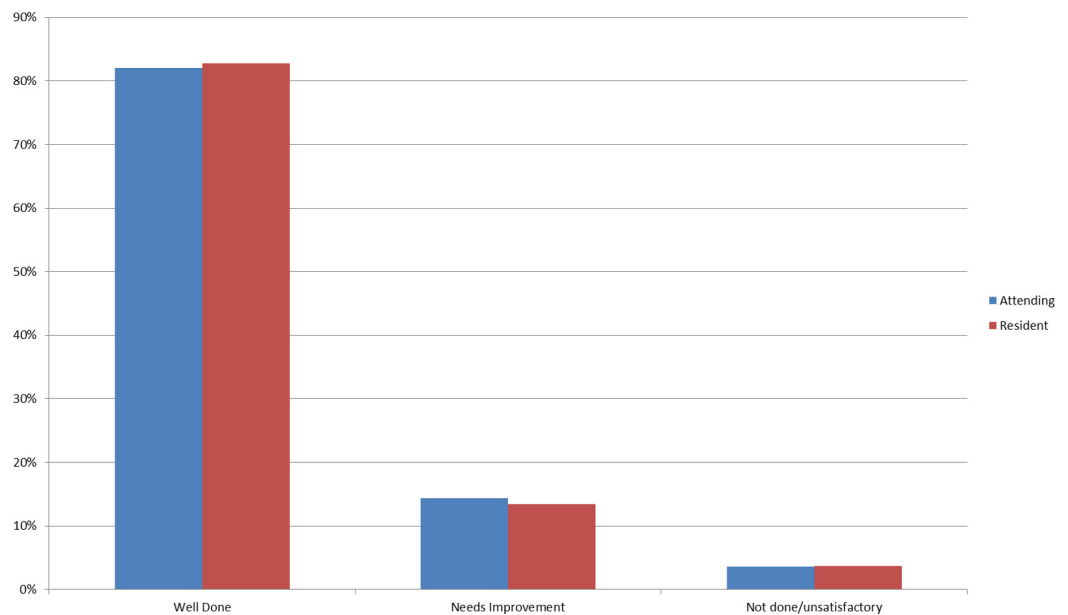
C

| IM Milestones  | IM Milestone Subcompetencies   | Assessments  |    |    |    |    |    |    |  |  |  |  |  |
|--|--|--|----|----|----|----|----|----|--|--|--|--|--|
| Gathers and synthesizes essential and accurate information to define each patients clinic problem(s). (PC1)            | Clinical knowledge 1: Has adequate scientific, socioeconomic and behavioral knowledge for patient care | Ch   | Ch |    |    |    |    |    |  |  |  |  |  |
|  | Patient care 1a: Elicits Secondary Data  | Us   | Us | Us |    |    |    |    |  |  |  |  |  |
|  | Patient care 1a: Hypothesis driven history   | Hi   | Hi | Hi | Hi | Hi |    |    |  |  |  |  |  |
|  | Patient care 1a: Accurate PE   | PE   | PE | PE | PE |    |    |    |  |  |  |  |  |
|  | Patient care 1a: Collects accurate history   | Ba   | Hi | Hi | Hi | Ba | Ba | Ba |  |  |  |  |  |
|  | Patient care 1a: Defines patients central problem  | Id   |    |    |    |    |    |    |  |  |  |  |  |
|  | Patient care 1a: Differential Diagnosis  | Pr   |    |    |    |    |    |    |  |  |  |  |  |
|  | Patient care 2: Develops appropriate care plans  | Hx   |    |    |    |    |    |    |  |  |  |  |  |
|  | Skills in performing procedures. (PC4)   | Patient care 4: Procedures   | Lu |    |    |    |    |    |  |  |  |  |  |
|  | Clinical knowledge (MK1)   | Clinical knowledge 2: Interprets basic diagnostic tests accurately   |    |    |    |    |    |    |  |  |  |  |  |
| Clinical knowledge 2: Understands pretest probability and test characteristics   |  |  |    |    |    |    |    |    |  |  |  |  |  |
| Patient care 1a: Accurate PE   |  | Ab   |    |    |    |    |    |    |  |  |  |  |  |
| Works effectively within an interprofessional team (e.g. peers, consultants, nursing, ancillary professionals). (SBP1) | Communication: Engages collaboratively with  | <div style="border: 1px solid black; padding: 5px;"> <p><b>Shared Decision Making</b><br/>ICS1</p> <p><b>D</b></p> <p><b>Communicates effectively with patients and caregivers</b></p> <p><i>The learner engages patients/advocates in shared decision making</i></p> </div> |    |    |    |    |    |    |  |  |  |  |  |
| Transitions patients effectively within and across health delivery systems. (SBP4)                                     | Communication: Verbal, non-verbal, written collaboration   |  |    |    |    |    |    |    |  |  |  |  |  |
| Accepts responsibility and follows through on tasks. (PROF2)   | Communication: Engages collaboratively with  |  |    |    |    |    |    |    |  |  |  |  |  |
| Responds to each patient's unique characteristics and needs. (PROF3)   | Communication: Incorporates patient specific   |  |    |    |    |    |    |    |  |  |  |  |  |
| Communicates effectively with patients and caregivers. (ICS1)  | Clinical knowledge 1: Has adequate scientific knowledge for patient care                               |  |    |    |    |    |    |    |  |  |  |  |  |
|  | Shared decision making - basic   | Sh   | Co | Co | Sh |    |    |    |  |  |  |  |  |
| Communicates effectively in interprofessional teams. (ICS2)  | Communication: Quickly develops therapeutic relationship   |  |    |    |    |    |    |    |  |  |  |  |  |
|  | Communication: Engages collaboratively with all team members   | Te   | Te | Te | Te |    |    |    |  |  |  |  |  |
| Appropriate utilization and completion of health care records. (ICS3)  | Communication: Health records are organized, comprehensive and accurate                                |  |    |    |    |    |    |    |  |  |  |  |  |
|  | Communication: Health records are succinct, relevant and patient specific                              | Ch   | Ch |    |    |    |    |    |  |  |  |  |  |

**Figure 7 JIT cloud-based reporting options.** Another option for a cloud-based record or registry of the learner's performance. This image represents a milestone based report with the identified milestones (A) the milestone subcompetencies (B) a color-coded table of all of the learners assessments (C) a roll-over option (D) identifies which specific assessment is represented in each cell. This table shows the ACGME competency taxonomy for internal medicine.

competency *gathering essential and accurate information* seems relatively straightforward. However, when applied to a single condition such as chronic obstructive pulmonary disease, there are at least 28 specified clinical tasks related to history taking and performing a physical examination that a student should demonstrate to meet the expected outcomes as defined in the Clerkship Directors in Internal Medicine (CDIM) curricular objectives for that problem. Of these 28, how many will a faculty evaluator remember when assessing the student? More importantly how many can they remember and what level of consistency will there be among preceptors providing feedback to students?

If we take almost any clinical skill and start to dissect it, we find very quickly that existing human memory is insufficient in recalling all of the explicit steps related to potentially hundreds of conditions that help frame the expected outcomes of a trainee's educational experience and curricula. As the expectations for assessment of discrete competencies



**Figure 8** Bar chart of grade distribution comparing resident to faculty raters.

increases, the evaluation burden for educators, students and administrators becomes progressively more educationally incomplete and logistically unmanageable.

The inability of faculty to remember and accurately assess for outcomes related to potentially hundreds of discrete educational objectives while evaluating trainees in clinical settings is one of the major reasons faculty have a hard time reliably discriminating unsatisfactory from satisfactory performance, as has been noted by many authors over the past decade using paper-based systems (*Holmboe, Hawkins & Huot, 2004; Kogan et al., 2011*). For example, in a study of mini-CEX evaluations among 300 medical students, Hill noted that problems existed, “in trying to ensure that everyone was working to the same or similar standards” (*Hill et al., 2009*). In another study of 400 mini-CEX assessments, Fernando concluded faculty evaluators were unsure of the level of performance expected of the learners (*Fernando et al., 2008*). Hasnain noted that poor agreement among faculty evaluating medical students on a Family Medicine Clerkship was due to the fact that “Standards for judging clinical competence were not explicit” (*Hasnain et al., 2004*). In a randomized trial of a faculty development effort, Holmboe studied the accuracy of faculty ratings by having them view videotaped trainee-patient encounters that were scripted to portray three levels of proficiency; unsatisfactory, marginal or satisfactory. Faculty viewing the exact same encounter varied widely in their assessment of trainee competence, with ratings from unequivocally unsatisfactory (CEX scores of scores 1–3) to unequivocally superior (CEX scores of 7–9), regardless of whether the video was scripted to be unsatisfactory or not. After an intensive 4 day faculty development workshop in which participants were tasked with developing a shared mental model of what specific competencies should look like, problems still existed among faculty in discriminating

satisfactory from unsatisfactory performance in these scripted encounters (*Holmboe, Hawkins & Huot, 2004*).

Kogan noted that in the absence of easily accessible frameworks, faculty evaluators default back to a myriad of highly variable evaluation strategies including such idiosyncratic features as instinct, “gut feelings”, “unsubstantiated assumptions” and the faculty members’ emotional response to providing feedback. What she also noted was that faculty raters commonly *fail to use existing frameworks or external standards* in guiding their evaluations of trainees, thus explaining much of the well-recognized problems with poor validity and inter-rater reliability associated with clinical evaluations (*Kogan et al., 2011*).

Given these realities, it is not surprising that medical trainees commonly do not view the feedback received from faculty as credible nor influential in learning, especially if the feedback was not immediate and tied to the trainees’ clinical work-place performance (*Watling et al., 2012*). Enhancing the effectiveness of clinical assessments, the delivery of feedback related to learning objectives and the creation of better systems for documenting faculty observations are commonly cited needs in medical education (*Hasnain et al., 2004; Howley & Wilson, 2004; Torre et al., 2007; Hauer & Kogan, 2012; Whitcomb, 2002*).

Given these and other trends, systems that are capable of disseminating curricular objectives to students and faculty and which also enable criterion-based assessment have become key educational needs. We believe that cloud-based technology, appropriately applied to maximize efficiency, can contribute to optimizing the learning environment by directly aligning learning objectives from national disciplinary curricula with assessment tools for use by students and faculty anywhere and anytime, especially at the bedside.

In our first feasibility study, we demonstrated our ability to deliver national educational objectives published by the CDIM to electronic hand-held personal digital assistants (PDAs) such as Palm<sup>®</sup> and PocketPC<sup>®</sup> devices (*Ferenchick, Fetters & Carse, 2008*). In a second feasibility study, we subsequently demonstrated that this system could be used to deliver, and successfully implement, competency-based checklists for student assessment related to the CDIM curricular objectives using PDAs (*Ferenchick et al., 2010*). Data from these studies helped us determine that the distribution and use of curricular objectives and related assessment tools by students and faculty in our geographically dispersed medical school could be facilitated with just in time mobile technology. Importantly, we also determined that students and preceptors valued the fact that the content and expected competencies were transparent and such transparency facilitated learner assessment (*Ferenchick et al., 2010*). However, technical issues with PDAs – such as lack of direct internet connection and the requirement to “synchronize” data from PDAs to the web using desktop computers – limited the practicality of PDA based assessment; a process that is not needed with contemporary internet enabled devices such as iPads, iPhones and other smartphones. These devices have become almost ubiquitous in the past four years and we have leveraged this trend to evolve JIT to a platform-neutral Cloud-based system. The displayed assessment tools function like an “application” on mobile devices,

but are device-agnostic in that they function on all internet-enabled devices, including smartphones.

Our study, like most others, has several inherent limitations. First, this is a single institution study and these results may not be generalizable. Future studies should focus on the use of this technology in other settings. Second, establishing the reliability of all of the customized checklists within the CEX application is needed, as is establishing its reliability in real clinical settings such as the hospital wards. Third, we have not established the validity of the electronic grading algorithm. Fourth, like many tools for direct observation, we have not established the effect of this tool on learning nor the transfer of acquired clinical skills to other areas, or the effect that such direct observation has on the most important outcome of patient care.

## CONCLUSIONS

We have established that just in time Cloud-based mobile technology has great potential in competency-based medical education. Although not an objective of this study, we believe such technology holds great promise for use in authentic clinical settings for measuring student achievement related to educational milestones. Additionally, given the time and cost constraints associated with traditional faculty development efforts, we believe that systems such as JIT have great potential in operationalizing “just in time” faculty development.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This study was funded in part by a grant D56HP05217 from the Division of Medicine, Bureau of Health Professions, Health Services and Resources Administration. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:  
Division of Medicine, Bureau of Health Professions, Health Services and Resources Administration: D56HP05217.

### Competing Interests

Dr David Solomon is an Academic Editor for PeerJ.

### Author Contributions

- Gary S. Ferenchick conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper.
- David Solomon analyzed the data, wrote the paper.

## Human Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

Our medical school has created an “Honest Broker System” for conducting research on student performance data that are collected as a regular part of the students’ educational activities. A designated employee of the medical school, with access to these data, has been designated as the “Honest Broker”. In an honest broker system a person or agency that has access to multiple human subject datasets collected for non-research purposes creates a de-identified dataset that can be used for research purposes without posing risk to the subjects (*Boyd et al., 2007*). This approach has been used in various types of clinical research – though, to the best of our knowledge, it has not been applied in educational research at other institutions.

At Michigan State University this designated individual created the dataset used in this study and made it available to our research team after removing all identifiers. The Social Science/Behavioral/Education Institutional Review Board (SIRB) at Michigan State University granted our study an exempted review.

## Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.164>.

## REFERENCES

- Boyd AD, Hosner C, Hunscher DA, Athey BD, Clauw DJ, Green LA. 2007.** An ‘Honest Broker’ mechanism to maintain privacy for patient care and academic medical research. *International Journal of Medical Informatics* 76(5–6):407–411 DOI 10.1016/j.ijmedinf.2006.09.004.
- Ebel RL. 1951.** Estimation of the reliability of ratings. *Psychometrika* 16:407–424 DOI 10.1007/BF02288803.
- Ferenchick G, Feters M, Carse AM. 2008.** Just in time: technology to disseminate curriculum and manage educational requirements with mobile technology. *Teaching and Learning in Medicine* 20(1):44–52 DOI 10.1080/10401330701542669.
- Ferenchick GS, Foreback J, Towfiq B, Kavanaugh K, Solomon D, Mohmand A. 2010.** The implementation of a mobile problem-specific electronic CEX for assessing directly observed student-patient encounters. *Medical Education Online* 15:4276 DOI 10.3402/meo.v15i0.4276.
- Fernando N, Cleland J, McKenzie H, Cassar K. 2008.** Identifying the factors that determine feedback given to undergraduate medical students following formative mini-CEX assessments. *Medical Education* 42(1):89–95 DOI 10.1111/j.1365-2923.2007.02939.x.
- Hasnain M, Connell KJ, Downing SM, Olthoff A, Yudkowsky R. 2004.** Toward meaningful evaluation of clinical competence: the role of direct observation in clerkship ratings. *Academic Medicine* 79(10 Suppl):S21–S24 DOI 10.1097/00001888-200410001-00007.
- Hauer KE, Kogan JR. 2012.** Realising the potential value of feedback. *Medical Education* 46(2):140–142 DOI 10.1111/j.1365-2923.2011.04180.x.
- Hill F, Kendall K, Galbraith K, Crossley J. 2009.** Implementing the undergraduate mini-CEX: a tailored approach at Southampton University. *Medical Education* 43(4):326–334 DOI 10.1111/j.1365-2923.2008.03275.x.

- Holmboe ES, Hawkins RE, Huot SJ. 2004.** Effects of training in direct observation of medical residents' clinical competence: a randomized trial. *Annals of Internal Medicine* **140**(11):874–881 DOI [10.7326/0003-4819-140-11-200406010-00008](https://doi.org/10.7326/0003-4819-140-11-200406010-00008).
- Howley LD, Wilson WG. 2004.** Direct observation of students during clerkship rotations: a multiyear descriptive study. *Academic Medicine* **79**(3):276–280 DOI [10.1097/00001888-200403000-00017](https://doi.org/10.1097/00001888-200403000-00017).
- Kogan JR, Conforti L, Bernabeo E, Iobst W, Holmboe E. 2011.** Opening the black box of clinical skills assessment via observation: a conceptual model. *Medical Education* **45**(10):1048–1060 DOI [10.1111/j.1365-2923.2011.04025.x](https://doi.org/10.1111/j.1365-2923.2011.04025.x).
- Solomon DJ. 2004.** The rating reliability calculator. *BMC Medical Research Methodology* **4**:11 DOI [10.1186/1471-2288-4-11](https://doi.org/10.1186/1471-2288-4-11).
- Torre DM, Simpson DE, Elnicki DM, Sebastian JL, Holmboe ES. 2007.** Feasibility, reliability and user satisfaction with a PDA-based mini-CEX to evaluate the clinical skills of third-year medical students. *Teaching and Learning in Medicine* **19**(3):271–277 DOI [10.1080/10401330701366622](https://doi.org/10.1080/10401330701366622).
- Watling C, Driessen E, van der Vleuten CP, Lingard L. 2012.** Learning from clinical work: the roles of learning cues and credibility judgements. *Medical Education* **46**(2):192–200 DOI [10.1111/j.1365-2923.2011.04126.x](https://doi.org/10.1111/j.1365-2923.2011.04126.x).
- Whitcomb ME. 2004.** More on competency-based education. *Academic Medicine* **79**(6):493–494 DOI [10.1097/00001888-200406000-00001](https://doi.org/10.1097/00001888-200406000-00001).
- Whitcomb ME. 2002.** Competency-based graduate medical education? Of course! But how should competency be assessed? *Academic Medicine* **77**(5):359–360 DOI [10.1097/00001888-200205000-00001](https://doi.org/10.1097/00001888-200205000-00001).