

OPEN

Improving the Clinical Skills Performance of Graduating Medical Students Using “WISE OnCall,” a Multimedia Educational Module

Demian Szyld, MD, EdM;

Kristen Uquillas, MD;

Brad R. Green, MD;

Steven D. Yavner, PhD;

Hyuksoon Song, MA, MS, PhD;

Michael W. Nick, MA, MS;

Grace M. Ng, PhD(c), MS, CNM, RN;

Martin V. Pusic, MD, PhD;

Thomas S. Riles, MD;

Adina Kalet, MD, MPH

Introduction: “Transitions to residency” programs are designed to maximize quality and safety of patient care, as medical students become residents. However, best instructional or readiness assessment practices are not yet established. We sought to study the impact of a screen-based interactive curriculum designed to prepare interns to address common clinical coverage issues (WISE OnCall) on the clinical skills demonstrated in simulation and hypothesize that performance would improve after completing the module.

Methods: Senior medical students were recruited to participate in this single group prestudy/poststudy. Students responded to a call from a standardized nurse (SN) and assessed a standardized patient (SP) with low urine output, interacted with a 45-minute WISE OnCall module on the assessment and management of oliguria, and then evaluated a different SP with low urine output of a different underlying cause. Standardized patients assessed clinical skills with a 37-item, behaviorally anchored checklist measuring clinical skills (intraclass correlation coefficient [ICC], 0.55–0.81). Standardized nurses rated care quality and safety and collaboration and interprofessional communication using a 33-item literature-based, anchored checklist (ICC, 0.47–0.52). Standardized patient and SN ratings of the same student performance were correlated (r , 0.37–0.62; $P < 0.01$). Physicians assessed clinical reasoning quality based on the students’ patient encounter note (ICC, 0.55–0.68), ratings that did not correlate with SP and SN ratings. We compared pre-post clinical skills performance and clinical reasoning. Fifty-two medical students (31%) completed this institutional review board–approved study.

Results: Performance as measured by the SPs, SNs, and the postencounter note all showed improvement with mostly moderate to large effect sizes (range of Cohen’s d , 0.30–1.88; $P < 0.05$) after completion of the online module. Unexpectedly, professionalism as rated by the SP was poorer after the module (Cohen’s d , -0.93 ; $P = 0.000$).

Discussion: A brief computer-based educational intervention significantly improved graduating medical students’ clinical skills needed to be ready for residency.

(*Sim Healthcare* 12:385–392, 2017)

Key Words: Transition to residency, Medical student education, Clinical skills, Assessment, Standardized nurse, Standardized patient, OSCE, Clinical reasoning.

From the Department of Emergency Medicine (D.S.), Brigham and Women’s Hospital; The Center for Medical Simulation (D.S.), Boston, MA; New York Simulation Center for the Health Sciences (G.N., T.S.R., A.K.), New York; Institute for Innovations in Medical Education (M.V.P., T.S.R., A.K.), NYU School of Medicine, New York, NY; Department of Obstetrics and Gynecology (K.U.), University of Southern California, Los Angeles, CA; Department of Emergency Medicine (B.R.G.), The Ohio State University, Cleveland, OH; Program for Medical Education and Technology (M.W.N., T.S.R., A.K.), NYU School of Medicine, New York, NY; Department of Journalism (S.D.Y.), Central Connecticut State University; Department of Education (H.S.), Georgian Court University, Lakewood, NJ; Department of Surgery (T.S.R., A.K.), NYU School of Medicine, New York, NY; Department of Medicine (A.K.), Division of General Internal Medicine, Research on Medical Education Outcomes (ROME) Unit, and Program for Medical Education Innovation and Research (PrMEIR) (A.K.), NYU School of Medicine, New York, NY.

Reprints: Demian Szyld, MD, EdM, Center for Medical Simulation, 100 First Ave, Suite 400, Boston, MA 02129 (e-mail: dszyld@bwh.harvard.edu).

WISE OnCall was developed at NYU School of Medicine. D.S. is the Executive Director for both WISE-MD, which produces and distributes WISE OnCall, and the New York Simulation Center. Both are not-for-profit entities. M.W.N. is a member of the Program for Medical Education and Technology at the NYU School of Medicine and Technical Director for WISE-MD and WISE OnCall. The other authors declare no conflict of interests.

D.S. is an Emergency Medicine physician at the Brigham and Women’s Hospital Emergency Department and is the Senior Director of the Institute for Medical Simulation at the Center for Medical Simulation in Boston. At the time of the study, he was an Assistant Professor in the Department of Emergency Medicine at NYU School of Medicine and the Medical Director and Fellowship Director at The New York Simulation Center for Health Sciences (NYSIM). K.U. at the time of the study was a Fellow in Simulation and Education at NYU and NYSIM and is currently a Fellow in Maternal Fetal Medicine at the University of Southern California. B.R.G. at the time of the study was a Fellow in Simulation and Education at NYU and NYSIM and is currently an Assistant Professor in

the Department of Emergency Medicine at The Ohio State University. S.D.Y. is an Assistant Professor of Journalism at Central Connecticut State University, a PhD Candidate in the Educational Communication and Technology program at NYU/Steinhardt, and a Research Assistant at the NYU School of Medicine. H.S. is an Assistant Professor of Education at Georgian Court University, NJ, and a consultant of WISE OnCall project at NYU School of Medicine. M.W.N. is a member of the Program for Medical Education and Technology at the NYU School of Medicine and Technical Director for the WISE MD initiative. G.N. is the Nursing Director and the NYSIM-Hearst Foundations Nursing and Allied Health Fellowship Director at The New York Simulation Center for the Health Sciences. M.V.P. is the Director of the Division of Learning Analytics in the NYU School of Medicine Institute for Innovations in Medical Education and Assistant Professor of Emergency Medicine. T.S.R. is the Frank C. Spencer Professor of Surgery and currently Associate Dean, Medical Education and Technology, Director of the Office On Continuing Medical Education, New York University School of Medicine, Executive Director of WISE-MD and WISE OnCall, and the Executive Director of the New York Simulation Center for the Health Sciences. A.K. is the Arnold P. Gold Professor of Humanism and Professionalism, Professor of Medicine and Surgery and Director of Research on Medical Education and Outcomes (ROME) unit of the Division of General Internal Medicine and Clinical Innovations (DGIMCI), and Director Research Program on Medical Education and Technology (PMET).

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal’s Web site (www.simulationinhealthcare.com).

Copyright © 2017 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the Society for Simulation in Healthcare. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal. DOI: 10.1097/SIH.0000000000000254

Within weeks of graduating from medical school, newly minted graduates in all specialties must be ready to perform as interns. This transition can be challenging and daunting. A task force of the American College of Surgeons surveyed 12% of US interns in surgery programs in 2009 and found that they feel “only moderately prepared to perform most of the common intern level responsibilities in the domains of medical knowledge and patient care.”¹ In the same study, “clinical management of (predictable) postoperative conditions” seemed to be a source of anxiety during intern year. Program directors and hospital leadership are equally concerned and have invested heavily in orientation programs and increased supervision to mitigate the “July effect”—in which patient care quality and safety are endangered as new trainees acclimate to their new roles.^{2,3} July of internship may be particularly challenging for graduates who do not have rigorous clinical rotations scheduled throughout their final year of medical school and focus instead on nonclinical activities. Regaining forgotten or eroded knowledge, skills, and attitudes, which had been previously acquired, has been the focus of “intern orientation.”

Recently, medical school programs have attempted to better prepare their students by implementing transitions to residency programs or boot camps.^{4–6} These programs are typically 1 week to 1 month in duration; focus on skills, attitudes, and formative assessment; and are usually face-to-face and resource intensive. Generally, they do not include baseline or summative assessments of their trainees' readiness such as performance-based assessment with documented feedback to the candidate or an educational handoff to future program directors. Barriers to conducting performance assessments at the end of medical school include lack of measures as well as costly resources such as simulation center space and time and faculty availability. Despite these efforts, leaders in the academic and hospital-based community are asking: are medical school graduates as ready as possible for residency?

In the interest of patient safety and quality of care, policy makers have advocated for new national standards to assure that residents are prepared to manage critical situations. The Accreditation Council for Graduate Medical Education guidelines, for example, stipulate that direct supervision is required for interns managing common conditions (eg, low urine output,

hypertension, change in neurologic status, etc) until competence is demonstrated and documented by the program director.⁷ The Association of American Medical Colleges similarly stipulated a set of common standards through their document, the core Entrustable Professional Activities for Entering Residency.⁸ In truth, little is known about the capacity of graduating medical students,² to function safely in residency. It is important to understand this gap quantitatively and qualitatively to inform transition to residency programs going forward.

For this study, we designed a program to assess and address readiness for residency of graduating medical students and focused on a single clinical coverage issue. We defined readiness for residency as demonstration of the following clinical skills—responding to a page while on call, assessing a patient complaint, and documenting the encounter—based on sound clinical reasoning, as reflected by the content of the medical record. We addressed readiness for residency using WISE OnCall.

WISE OnCall

We developed “WISE OnCall,” a series of instructional case-based multimedia modules designed to address readiness for internship. The theoretical framework for WISE OnCall is cognitive apprenticeship,^{9,10} an instructional model in which an expert guides an “apprentice” to be able to experience common medical/surgical “coverage” situations, through modeling, coaching, scaffolding, and ultimately fading of instructional guidance. This design of the modules is based on our experience building and researching the use of WISE MD for clinical clerkship education.^{11–16} To date, 8 WISE OnCall modules have been completed with plans to build at least 5 more in the next 2 years (<https://www.wisemed.org/wise-oncall/>). Each module provides a sequential approach to a common clinical issue. First, there is an overview of the problem including underlying diagnostic and management approaches. This is followed by a video simulation illustrating the assessment of a clinical situation emphasizing collection of information from the medical record, the nursing staff, the patient, and medical tests. The case reinforces the importance of effective interprofessional communication and professionalism. The final section includes 3 practice cases that enable the learner to interact by applying diagnostic skills and obtaining feedback. This digital strategy is

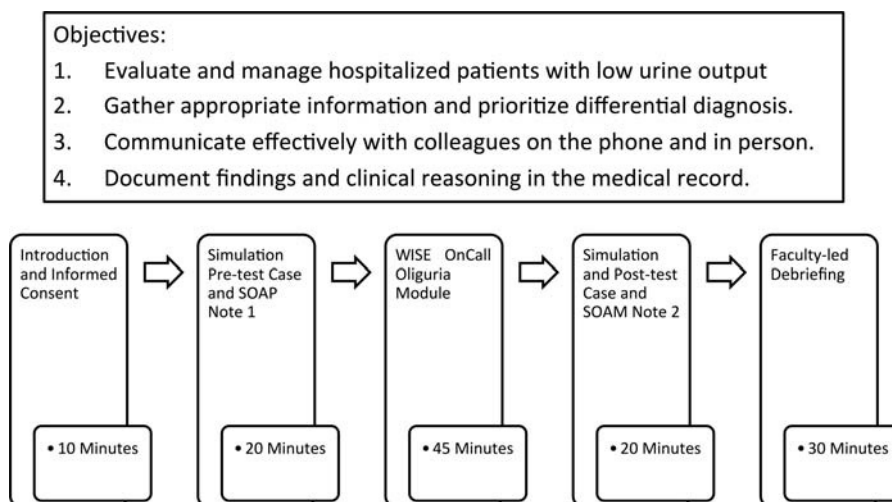


FIGURE 1. Objectives, sequence, and timeline of the readiness for internship experience.

designed to help transitioning medical students or novice residents develop more advanced diagnostic schemas or cognitive illness scripts, which then become the basis of the development of their clinical reasoning.¹³

For this study, we chose the oliguria (low urine output) module because it is a topic that all students are likely to have basic familiarity with by the end of medical school and it is a condition interns in medical and surgical disciplines can expect to encounter during a typical night on hospital call. The objectives for the module are listed in Figure 1. We hypothesized that on average, the clinical skills required to safely care for a patient with low urine output would be demonstrably strengthened through interacting with WISE OnCall.

METHODS

Subjects and Setting

All senior-year medical students ($n = 164$) at our medical school in the North Eastern United States received an e-mail toward the second half of the academic year from the principal investigators inviting them to participate in a 3-hour “readiness for internship” event held in our simulation center. Our students have extensive experience with performance-based assessment involving standardized patients (SPs) and are also accustomed to the assessment framework used in this study. There are 18 formative and summative required interactions spread throughout the 4-year curriculum. In the letter, we explained that 1 purpose of this event was to study the effectiveness of WISE OnCall—a series of Web-based modules designed to prepare them to care for patients with common clinical coverage issues—and that they would also be able to debrief their experience with a clinical faculty member. Fifty-two senior medical students out of a class of 164 (31%) volunteered to participate, and all completed the study protocol. They received 100 US dollars for completing all study materials. No academic credit was given for participation in this activity, and it was not required for graduation. The institutional review board at our medical school reviewed and approved this study after a full protocol review. We choose not to collect any demographic information to ensure the confidentiality of our students, some of whom would remain at our institution for residency training.

Standardized Patient/Nurse Scenarios

We assessed ability to evaluate and manage oliguria in the acute care setting before and after they completed the WISE OnCall module using the SP/standardized nurse (SN) simulation cases. We developed 2 interprofessional simulated scenarios, both involving an SP and an SN. In both cases, the intern was asked to evaluate an otherwise clinically stable, hospitalized patient with low urine output. The cases were designed to be of equivalent difficulty for senior medical students who were familiar with the differential diagnosis of new onset oliguria. Accompanying instruments were developed to assess students' clinical reasoning and clinical skills after standard case design methodology.¹⁷

Procedures

Figure 1 illustrates the sequence of activities and timing for the readiness of internship experience. After a brief orientation and informed consent discussion, participants entered a

simulated encounter in which they participated as the covering intern on a surgical service. The scenario began with a page from an SN calling them to assess a postoperative patient with oliguria (simulation pretest/case no. 1).

Upon entering the examination room, the students were expected to engage with both an SP and the SN to assess the clinical situation and to make a plan for management and follow-up. The students were told that they could request laboratory results, electrocardiograms, and imaging results and order and receive the results from simple bedside procedures (eg, Foley catheter insertion, bladder scan with ultrasound, intravenous fluid administration) from the SN.

After the 15-minute encounter, the participants exited the room and were given 5 minutes to type a “coverage” note summarizing their findings (subjective and objective), clinical reasoning (assessment), and plan in the common subjective, objective, assessment, plan (SOAP) format into a word processing program on a computer.

The primary intervention is the WISE OnCall oliguria module. Participants were given 45 minutes to view the module in a private space with a proctor. After viewing the module, participants had a second simulation encounter (simulation posttest/case no. 2). The encounter lasted 15 minutes, and they had 5 minutes to complete the associated SOAP note. To maintain realism and decrease rater bias, a different SP and SN pair was used for the second case. The characteristics of the case and level of difficulty were similar, but the root cause of oliguria was different (Table 1).

Finally, students participated in a 30-minute debriefing session with a member of the clinical faculty where all aspects of the exercise were discussed. The main purpose of this debriefing was to answer students' questions. To protect the privacy of students who may have been matched for residency within our institution, the faculty member did not view the student's actual performance. Therefore, the debriefing addressed only student self-reported concerns about their performance.

Simulation Cases and Clinical Skills Assessment

Experienced SPs were recruited and underwent 3 hours of case portrayal and rater training to prepare them to reliably assess clinical skills performance using a 37-item behaviorally anchored checklist measuring communication skills in 3 subdomains—data gathering (6 items), rapport building (6 items), patient education (3 items), history gathering (10 items), and physical examination skills (9 items) (interrater intraclass correlation coefficient [ICC], 0.55–0.81) previously validated for use as a measure of core clinical skills.¹⁷ Experienced SPs were recruited to participate as SNs and were trained by an experienced nurse educator (G.N.) to portray a nurse in the simulation and to rate the student's competence. The SN used a 33-item literature-based, behaviorally anchored checklist to rate the student's performance.¹⁸ The checklist was modified from an instrument developed for OSCE-based assessment of interprofessional communication including aspects of care quality and safety, collaboration and teamwork, role clarity, and relationship development tailored to the specific clinical case requirements.¹⁹ After the behaviorally anchored items, at the end of the SP and SN checklists, raters completed 2 global items each on a 4-point scale. They were asked separately if they would

TABLE 1. Simulation Case Descriptors

		Simulation Pretest/Case 1	Simulation Posttest/Case 2
Name		Robert Jackson	Micah Taylor
Vital signs	Temperature	98.6°F	97.8°F
	Pulse	86	114
	Respiratory rate	16	16
	Blood pressure	141/86	134/82
	Oxygen saturation	100%	100%
Situation	Nurse phone script	“I’m calling about Mr. Jackson. His urine output has dropped and he has mild abdominal pain”	“I’m calling about Mr. Taylor. His urine output has dropped and he remains without chest pain”
	Chief complaint	Lower abdominal discomfort and fullness	No complaints
Background	Age, y	64	57
	Sex	Male	Male
	Past medical history	<ul style="list-style-type: none"> • Hypertension • Hypercholesterolemia • BPH • AAA 	<ul style="list-style-type: none"> • Coronary artery disease • Angioplasty and stent • Atrial fibrillation
	Reason for hospitalization	Elective endovascular repair of AAA, postoperative day 3	Observation to rule out acute cardiac ischemia and pulmonary embolism
	Social history	Retired	
	Medications	Hydrochlorothiazide 25 mg daily Lipitor 25 mg daily Lisinopril 20 mg daily Tamsulosin 0.4 mg daily	Aspirin 81 mg daily Metoprolol 50 mg daily Lipitor 25 mg daily Lisinopril 20 mg daily
Assessment	Urine output after Foley/color	1.1 L Light yellow	300 mL Dark brown
	Prerenal	Sepsis, hemorrhage, hematoma, dehydration	Myocardial infarction, arrhythmia, dehydration
	Renal	Contrast-induced nephropathy	Contrast-induced nephropathy
	Postrenal	Urinary obstruction and retention (BPH, opioids, held tamsulosin)	Urinary obstruction
Recommendation and diagnostic plan	Bedside	Foley catheter or ultrasound of urinary bladder	Foley catheter or ultrasound of urinary bladder
	Laboratory tests	Normal	Normal except elevated BUN and creatinine
	Electrocardiogram	Normal	Sinus tachycardia
	Management plan	Maintain Foley Hold hydrochlorothiazide Reevaluate within 1 H	Isotonic fluid bolus and infusion Hold hydrochlorothiazide Reevaluate within 1 H
	Supervision	Contact resident or attending to discuss case	Contact resident or attending to discuss case

AAA, abdominal aortic aneurism; BPH, benign prostatic hypertrophy; BUN, blood urea nitrogen.

recommend this student based on medical competence (not recommend, recommend with reservations, recommend, or highly recommend), and then they were asked to rate the student’s professionalism (not at all, somewhat, professional, and highly professional). Although both SP and SN were in the same room viewing the same student performance, they completed their checklists independently without conferring with each other. A 10% random sample of interactions was reviewed and rerated on videotape by a member of the study team (G.N.) (ICC, 0.47–0.52). Standardized patient and SN checklists for both cases are included as supplemental materials (see PDF, Supplementary Digital Content #1, <http://links.lww.com/SIH/A340>; SP Checklist Oliguria Case 1), (see PDF, Supplementary Digital Content #2, <http://links.lww.com/SIH/A341>; SP Checklist Oliguria Case 2), (see PDF, Supplementary Digital Content #3, <http://links.lww.com/SIH/A342>; SN Checklist Case 1), (see PDF, Supplementary Digital Content #4, <http://links.lww.com/SIH/A343>; SN Checklist Case 2).

Clinical Reasoning Assessment

Clinical reasoning was assessed based on the SOAP note and scored by 2 experienced clinicians (K.U., D.S.) using a rubric that we developed based on Pangaro’s Reporter-Interpreter-Manager-Educator framework.²⁰ We selected this framework because it is how we teach students clinical reasoning and patient note writing and is reinforced in the WISE OnCall Module (Table 2). Each student’s notes were scored on 4 characteristics—reporter, interpreter, manager, and overall clinical reasoning—each on a 4-point scale using a rubric with detailed case specific anchors. Twenty percent of the SOAP notes were rated independently by both raters (ICC, 0.55–0.68).

Data Analysis

Standardized patients and SNs rated students on 3-level behaviorally anchored (not done, partially done, well done) checklist items. For analysis, we reduce item scores to a dichotomous variable (not done/partially done, well done) and

TABLE 2. Scoring Rubric for Assessing Oliguria SOAP Note

Reporting Level	Poor Reporter (1)	Beginning Reporter (2)	Competent Reporter (3)	Strong Reporter (4)
In this section, the student was asked to describe the history, physical examination, and relevant laboratory findings he or she obtained from this patient and the nurse.	Markedly disorganized with subjective and objective illogically mixed and/or content insufficient to adequately support the assessment and plan, missing a “problem statement” or chief complaint and critical pertinent +/- (from Hx and/or PE Labs), and lack of semantic competence	At least somewhat organized with most S&O info in right places, content positively supports assessment and plan but no evidence of ruling out other causes, includes a “problem statement” or chief complaint, missing critical pertinent +/- (from Hx and/or PE Labs), or lacking semantic competence	Clinically logical and organized S&O, content positively supports the assessment and plan with some but not all pertinent negatives, includes “problem statement” and all critical pertinent +, missing pertinent negatives	Well organized with a full set of clinically relevant S&O data, content supports several reasonable diagnoses, includes efficient “problem statement” with semantic modifiers and all critical pertinent +/-, includes pertinent negatives, and is semantically sophisticated (clinical words used accurately “sounds like a great intern”)
Interpretation Level	Poor Interpreter (1)	Beginning Interpreter (2)	Competent Interpreter (3)	Strong Interpreter (4)
In this section, we assess the student’s interpretation and management skills. For instance, the assessment section may begin with a more sophisticated “problem statement” summarizing all the pertinent data—score this as interpretation (not as reporting)	No summary of pertinent S&O data, no clear statement of most likely cause of oliguria (no commitment)	Summary of pertinent S&O data (with key missing features), no clear statement of most likely cause of oliguria (no commitment)	Complete summary of pertinent S&O data (may be too long or repetitive), accurate, clear statement of most likely cause of oliguria but no clear plan to reconsider in the short term (eg, follow-up time frame)	A complete, efficiently articulated (“few words as needed”) problem statement with semantic modifiers (eg, acute, chronic, progressive, etc), accurate, clear statement of most likely cause of oliguria and clear plan to assess for alternative explanations (eg, follow-up time frame, seeking refutational information—eg, assess for post op bleeding, rule out congestive heart failure)
Managing Level	Poor Manager (1)	Beginning Manager (2)	Competent Manager (3)	Strong Manager (4)
This should be based on the plan (SOAP), which includes diagnostic tests, orders for nursing care or reporting, recommended follow-up (eg, “trending” ECGs, troponins, restricted activity, fluids, Foley catheter placement or monitoring of in & out maneuvers, laboratory tests, imaging, etc)	No appropriate plan	Incomplete or illogical (“shotgun”) list of appropriate interventions, nursing orders, diagnostic studies and follow-up, and/or some inappropriate studies	Comprehensive but vague lists of interventions (eg, “fluid challenge,” Foley), additional laboratory tests (eg, UA, trending ECGs, or BMP/cr/BUN), nursing call orders (I’s&O’s, change in status), appropriate diagnostic studies and follow-up plan, and comments on seeking supervision	Comprehensive explicit/specific lists of interventions (eg, “fluid challenge,” Foley), additional laboratory tests (eg, UA, trending ECGs, or BMP/cr/BUN), nursing “call when” orders (I’s&O’s, change in status), appropriate diagnostic studies and specific follow-up plan, and comments on seeking supervision
Please rate the overall quality of clinical reasoning evidenced in this note:	Poor (1)	Beginning (2)	Competent (3)	Strong (4)

BUN indicates blood urea nitrogen; S&O, subjective and objective; UA, urinalysis; ECG, electrocardiogram; cr, creatinine; BMP, basic metabolic panel.

summarize these for each subdomain as percent well done (% WD). This is the SP checklist data analysis and reporting strategy we have developed over 15 years for both educational and research purposes because we obtain the highest inter- and intrarater reliability, have established consistently high checklist internal validity (Cronbach α for subdomains range, 0.79–0.90), pass/fail cutoffs (which are determined based on multiple factors), and hover around 2 SDs below the mean, which is typically between 55% and 60%. We find this approach to be most transparent and educationally meaningful to students and educators. This is because it reflects the percentage of items in

each subdomain the individual student did well, standardizing the scores across subdomains with uneven number of items. When we are working individually with a student to understand the subtleties of his or her performance, we use the data from the “partially done” category.

Clinical reasoning was summarized as mean scores on the 4 Reporter-Interpreter-Manager-Educator characteristics on a 1 to 4 scale and SD. We examine whether WISE OnCall improves the participants’ clinical performance and reasoning using paired samples *t* tests. A priori α level was set as 0.05 for all estimations and $\beta = 0.20$. The magnitude of the effect of the

TABLE 3. Clinical Skills and Clinical Reasoning Performance Before and After WISE OnCall Module

Task	Category	Range	Simulation Pretest		Simulation Posttest		ES	P
			Mean	SD	Mean	SD		
SP checklist	Communication (%WD, 17 items)	0–100	73.30	27.53	84.05	21.36	0.30	0.037
	History (%WD, 12 items)	0–100	39.69	20.91	60.04	17.08	0.81	0.000
	Physical examination (%WD, 8 items)	0–100	51.77	15.02	72.84	12.80	1.06	0.000
	Recommendation–medical competence*	1–4	2.16	0.62	2.78	0.42	0.85	0.000
	Professionalism*	1–4	2.57	0.61	2.00	0.00	–0.93	0.000
SN checklist	Interprofessional communication (%WD, 30 items)	0–100	36.23	14.45	45.40	15.68	0.53	0.000
	Recommendation–medical competence*	1–4	2.00	0.69	2.31	0.67	0.35	0.014
	Professionalism*	1–4	2.25	0.67	2.53	0.51	0.37	0.026
Patient note	Reporting*	0–3	1.31	0.55	1.69	0.62	0.49	0.001
	Interpretation*	0–3	1.00	0.58	1.44	0.74	0.52	0.001
	Managing*	0–3	1.02	0.33	1.13	0.44	0.19	0.200
	Overall*	0–3	0.94	0.52	1.35	0.76	0.48	0.002

*Single items that were rated on a 4-point scale.

The effect size calculated as Cohen's *d* is interpreted as follows: 0.2 to 0.3 as a “small” effect, around 0.5 a “medium” effect, and 0.8 to infinity as a “large” effect. ES indicates effect size.

module on clinical skills is measured by Cohen's *d*, which is the difference between the means divided by the pooled SD. Cohen's *d* is a unit-less ratio allowing comparison of intervention effects across domains measured in different ways and across studies. It is interpreted qualitatively. In general, an effect size of 0.2 to 0.3 is interpreted as a “small” effect, around 0.5 a “medium” effect, and 0.8 to infinity a “large” effect.^{21,22} Interrater agreement on each measure was calculated as an ICC for pairs of raters within domains.

To assess the concurrent validity of our measures in this small sample, we hypothesized that observers of the same performance would correlate, and, based on our experience measuring clinical skills in this manner, the clinical reasoning as reflected in writing the patient note would not correlate with observed skills performance. We generated a correlation matrix among the comparable domains on the SP (communication skills, recommendation-medical competence, recommendation-professionalism) and SN checklists (interprofessional communication, recommendation-medical competence, recommendation-professionalism), and the faculty member assessed overall clinical reasoning based on the patient notes by each SP/SN case. Because some of the variables are ordinal, we conducted both Pearson and Spearman correlation coefficients, which yielded nearly the same results. We report Pearson *r* for ease of interpretation.

RESULTS

Group Performance and Clinical Reasoning at Baseline and After WISE OnCall

Table 3 shows the potential range, mean scores, and SDs on the categories of simulation pretest and posttest. As hypothesized, all categories measured improved a moderate to large amount after viewing the module, except for professionalism as rated by the SP, which declined, and the “managing” item based on the patient note, which remained the same from the pre- to postsimulation. Comparable SP and SN ratings within the same case were correlated (case no. 1: all 9, $r = 0.50$ [$P < 0.001$]; case no. 2: range for 9, $r = 0.37$ – 0.62 [$P < 0.01$]),

and none of the SP and SN ratings correlated with the overall clinical reasoning.

DISCUSSION

WISE OnCall significantly improved the ability to perform key clinical skills expected of a covering intern. We observed moderate to large gains in all clinical competency areas measured supporting both the immediate educational impact of WISE OnCall on performance and the sensitivity and responsiveness of our measures to instruction. The decrease in the global professionalism rating by the SPs after the module is intriguing. It is possible that some students, after being reminded by the WISE OnCall module to be more clinically systematic and thorough, did this at the expense of the other aspects of professional communication, such as rapport building in this time-bounded patient care task. We will be exploring this potentially unintended negative phenomenon in future studies.

It is not surprising that medical students learned in response to instruction. However, given the consistency and magnitude of the improvement and the fact that these were senior medical students, it is likely that this reflects that WISE OnCall's benefit is in reminding students of previous knowledge and ability rather than de novo learning.^{23,24} Medical students develop basic cognitive schema for oliguria during both preclinical and clinical training. With regular practice during the clinical clerkship curriculum, we assume that students automate more sophisticated thinking patterns and store complex schemas in long-term memory. However, during the fourth and final year of medical school, this material is probably not frequently accessed because our students spend 12 weeks on a scholarly activity, complete only 6 required clinical weeks (critical care and subinternship), and spend a significant amount of time focused on the residency selection process.^{25–27} We believe that this supports the importance of focused learning experiences, such as WISE OnCall, in critical competency domains as part of transition to internship.

The good news is that at the end of medical school, many students are ready for basic clinical tasks and most are primed

and easily reminded of what they have learned. The WISE OnCall module is potentially an efficient method to both remind students of what they already know and boost their capacity to “put it all together.”

Of course, aggregate data mask important individual patterns of response to this learning experience. Most students had improvements in their pre- to post-WISE OnCall clinical skills and clinical reasoning, but a significant minority demonstrated a deterioration of performance over the course of this study. Because participation in the study was anonymous, we could not interview individual students to explore reasons for this pattern. However, it suggests that this type of readiness for internship experience may be useful in identifying individuals in need of further attention and/or remediation before undertaking clinical responsibilities without supervision as an intern.

The evidence to support the validity of our measurement approach to assessing the impact of the WISE OnCall module on core clinical coverage skills needed in internship includes that (1) the constructs measured were performance based and aligned with expert sources on readiness for residency^{7,8} and the explicit objectives in the WISE OnCall module; (2) the observation and scoring of student skills were made in the rich context of 2 clinical content concordant cases from 3 important perspectives—the patient, nurse, and attending physician—using established rater training and materials, suggesting this was an authentic response process²⁸; (3) experienced and calibrated SPs reported if students collected the available clinical data accurately and completely using appropriate skills, the SN observed and reported on the content and quality of the interprofessional communication and clinical management at the bedside, and an experienced clinician scored the patient note as an ecologically valid manifestation of the students clinical reasoning; and (4) relationship among the measured variables is as hypothesized with correlation between raters observing the same performance and lack of correlation among those direct observers and assessment based on a written note. Unfortunately, rater reliability continues to be a major threat to the internal validity of this study. It remains to be seen if these measures track individual development over time or predict educational or clinical outcomes. We will be exploring our ability to provide meaningful educational handoffs to trainees and residency program directors.

Limitations

The major threat to the validity of our findings is the lack of a control group. We did not include one because we were concerned that we would not be able to recruit enough volunteers for an adequate sample size. Now that we have reason to expect hardy effect sizes, we will be able to randomly assign students to a condition without the WISE OnCall module even with relatively small numbers of participants. It is possible that the findings can be explained simply by the training effect of engaging in 2 sequential simulations. Even if valid, there are several aspects of this study, which limit generalizability of the findings and will be addressed in the future. Subjects were self-selected volunteers and may not reflect the general population of graduating medical students. The performance data are based on a single pre- and postmodule case in a single clinical domain and may not reflect clinical competence in general.

Interrater reliability was moderate for SN and clinical reasoning measure. Although we carefully sought to design 2 equally difficult cases, we did not randomly alternate the cases to ensure our findings were not biased. For purely pragmatic reasons, all students received the same cases in the same order. Therefore, it is possible that, independent of the impact of the module, our findings are explained by the second case being significantly easier than the first.

FUTURE DIRECTIONS

Boosting skills and knowledge at the end of medical school and at the beginning of residency remains important to protect patients during the first few months of residency. Future studies should evaluate various strategies to accomplish this and consider long-term retention of learning, potential negative implications for professionalism, cost, and effectiveness measured by educationally and patient-relevant outcomes. We are in the process of studying the longer-term impact of WISE OnCall by following some of our study subjects into residency, experimenting to identify the best sequence of SP/SN cases and WISE OnCall to enhance readiness for residency, and exploring the case specificity of the learning from a particular WISE OnCall module.

CONCLUSIONS

Not all graduating medical students are ready for assessing and managing low urine output, a common on-call situation. As we have shown, it may be possible to assess and improve their readiness through individualized, clinically meaningful measurements and interval training sessions at the end of medical school. Such efforts may have a significant role in ensuring patient safety, especially early on in the residency year. We are currently engaged in refining this approach and expanding the clinical coverage issues addressed in the hopes that we can produce measures of readiness for internship suitable for meaningful educational handoffs of trainees from undergraduate to graduate medical education settings. This work could lead to both benchmarked, clinically meaningful measures and methods to assess readiness for internship and an efficient educational approach to addressing performance gaps.^{29,30} WISE OnCall was designed to be used in this way addresses predictable, level-appropriate just-in-time learning needs.

REFERENCES

1. Minter RM, Amos KD, Bentz ML, et al. Transition to surgical residency: a multi-institutional study of perceived intern preparedness and the effect of a formal residency preparatory course in the fourth year of medical school. *Acad Med* 2015;90:1116–1124.
2. Lyss-Lerman P, Teherani A, Aagaard E, Loeser H, Cooke M, Harper GM. What training is needed in the fourth year of medical school? Views of residency program directors. *Acad Med* 2009;84:823–829.
3. Sachdeva AK, Loiacono LA, Amiel GE, Blair PG, Friedman M, Roslyn JJ. Variability in the clinical skills of residents entering training programs in surgery. *Surgery* 1995;118:300–308. discussion 8–9.
4. Fisher JW, Thompson BM, Garcia AD. Integrative clinical experience: an innovative program to prepare for internship. *Teach Learn Med* 2007;19:302–307.
5. Laack TA, Newman JS, Goyal DG, Torsher LC. A 1-week simulated internship course helps prepare medical students for transition to residency. *Simul Healthc* 2010;5:127–132.

6. Teo AR, Harleman E, O'Sullivan PS, Maa J. The key role of a transition course in preparing medical students for internship. *Acad Med* 2011;86:860–865.
7. Accreditation Council for Graduate Medical Education (ACGME). Specialty-Specific Duty Hour Definitions. Available at: https://www.acgme.org/Portals/0/PDFs/DH_Definitions.pdf. Accessed February 26, 2017.
8. Association of American Medical Colleges (AAMC). The Core Entrustable Professional Activities for Entering Residency. Available at: <https://www.aamc.org/initiatives/coreepas>. Accessed February 26, 2017.
9. Collins A, Brown JS, Newman SE. Cognitive apprenticeship: teaching the crafts of reading, writing, and mathematics, learning and instruction: essays in honor of Robert Glaser. In: Resnick LB, ed. New Jersey: Lawrence Erlbaum Associates; 1989:453–494.
10. Stalmeijer RE, Dolmans DH, Snellen-Balendong HA, van Santen-Hoeufft M, Wolfhagen IH, Scherpbier AJ. Clinical teaching based on principles of cognitive apprenticeship: views of experienced clinical teachers. *Acad Med* 2013;88:861–865.
11. Kalet AL, Coady SH, Hopkins MA, Hochberg MS, Riles TS. Preliminary evaluation of the web initiative for surgical education (wise-md). *Am J Surg* 2007;194:89–93.
12. Kalet AL, Song HS, Sarpel U, et al. Just enough, but not too much interactivity leads to better clinical skills performance after a computer assisted learning module. *Med Teach* 2012;34:833–839.
13. Song HS, Kalet AL, Plass JL. Assessing medical students' self-regulation as aptitude in computer-based learning. *Adv Health Sci Educ Theory Pract* 2011;16:97–107.
14. Song HS, Kalet AL, Plass JL. Interplay of prior knowledge, self-regulation and motivation in complex multimedia learning environments. *J Comput Assist Learn* 2016;32:31–50.
15. Song HS, Pusic M, Nick MW, Sarpel U, Plass JL, Kalet AL. The cognitive impact of interactive design features for learning complex materials in medical education. *Comput Educ* 2014;71:198–205.
16. Yavner SD, Pusic MV, Kalet AL, et al. Twelve tips for improving the effectiveness of web-based multimedia instruction for clinical learners. *Med Teach* 2015;37:239–244.
17. Zabar S, Kachur E, Kalet A, Hanley K. *Objective Structured Clinical Examinations: 10 Steps to Planning and Implementing Osces and Other Standardized Patient Exercises*. New York, NY: Springer Science & Business Media; 2012.
18. Zabar S, Adams J, Kurland S, et al. Charting a key competency domain: understanding resident physician interprofessional collaboration (IPC) skills. *J Gen Intern Med* 2016;31:846–853.
19. Gillespie C, Adams J, Hanley K, et al. 'We might as well be speaking different languages': an innovative interprofessional education tool to teach and assess communication skills critical to patient safety. *BMJ STEL* 2015;1:54–60.
20. Pangaro L. A new vocabulary and other innovations for improving descriptive in-training evaluations. *Acad Med* 1999;74:1203–1207.
21. Cohen J. *Statistical Power Analysis for the Behavioural Sciences*. rev. ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1977.
22. Magnusson K. Interpreting Cohen's d Effect Size: An Interactive Visualization. Available at: <http://rpsychologist.com/d3/cohend/>. Accessed February 27, 2017.
23. Pusic MV, Boutis K, Hatala R, Cook DA. Learning curves in health professions education. *Acad Med* 2015;90:1034–1042.
24. Pusic MV, Kessler D, Szyld D, Kalet A, Pecaric M, Boutis K. Experience curves as an organizing framework for deliberate practice in emergency medicine learning. *Acad Emerg Med* 2012;19:1476–1480.
25. Benson NM, Stickle TR, Raszka WV Jr. Going "fourth" from medical school: fourth-year medical students' perspectives on the fourth year of medical school. *Acad Med* 2015;90:1386–1393.
26. Day CS, Yeh AC, Franko O, Ramirez M, Krupat E. Musculoskeletal medicine: an assessment of the attitudes and knowledge of medical students at Harvard medical school. *Acad Med* 2007;82:452–457.
27. Wu EH, Elnicki DM, Alper EJ, et al. Procedural and interpretive skills of medical students: experiences and attitudes of fourth-year students. *Acad Med* 2008;83:S63–S67.
28. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. *Am J Med* 2006;119:166.e7–166.e16.
29. Ten Cate O. Nuts and bolts of entrustable professional activities. *J Grad Med Educ* 2013;5:157–158.
30. Lypson ML, Frohna JG, Gruppen LD, Woolliscroft JO. Assessing residents' competencies at baseline: identifying the gaps. *Acad Med* 2004;79:564–570.