Moderator: James Nixon, MD
Discussant: Kimberly Ephgrave, MD

Point-of-Care Assessment of Medical Trainee Competence for Independent Clinical Work

Tara J. T. Kennedy, Glenn Regehr, G. Ross Baker, and Lorelei Lingard

Abstract

Background

Clinical supervisors make frequent assessments of medical trainees' competence so they can provide appropriate opportunities for trainees to experience clinical independence. This study explored context-specific assessments of trainees' competence for independent clinical work.

Method

In Phase One, 88 teaching team members from internal and emergency medicine were observed during clinical activities (216 hours), and 65 participants completed brief interviews. In Phase Two, 36 in-depth interviews were conducted using video vignettes. Data collection and analysis employed grounded theory methodology.

Results

Supervisors' assessments of trainee trustworthiness for independent clinical work involved consideration of four dimensions: knowledge/skill, discernment of limitations, truthfulness, and conscientiousness. Supervisors' reliance on language cues as a source of trustworthiness data was revealed.

Conclusions

This study provides an initial exploration of context-specific competence assessments, which affect both patient safety and education, and provides a novel framework for study of the links between language use and competence.

Acad Med. 2008;83(10 Suppl):S89–S92.

Throughout North America, the active participation of medical trainees in patient care is considered to be critical to their learning process.[1,2] Maintaining an appropriate balance between independence for learning and supervision for safety is an ongoing process for the clinical supervisors of medical trainees. In an earlier stage of our research program, we described how supervising physicians consider a number of factors when making decisions about how much clinical oversight to provide to trainees.[3] Identified triggers for the provision of more intensive oversight included acute or severe clinical situations, issues raised by nurses or family members, and concerns about a trainee's competence to handle a specific clinical situation.[3] The present study was intended as an in-depth exploration of supervising physicians' assessments of this case-specific competence for independent clinical work.

Although the process of formal evaluation of medical trainee competence has been the subject of much empirical study,[4] the process of assessment of trainees' competence to provide independent care for a given patient or in a specific clinical context has not been described. This "point-of-care" competence assessment (i.e., occurring at the time and in the setting of clinical care) arguably has much more practical impact on patient care and trainee education than does any formal evaluation process, because it guides decisions about the nature of the day-to-day monitoring of trainees' clinical activities provided by supervising physicians. As part of an ongoing study of clinical supervision practices, this study aimed to explore supervising physicians' assessments of trainees' competence to provide independent clinical care, and the process employed to make these assessments.

Method

The study was designed using grounded theory methodology.[5,6] The study took place in three teaching hospitals affiliated with an urban Canadian medical school. Institutional review board approval was obtained. Study settings included the emergency medicine (EM) department and the general internal medicine (GIM) inpatient teaching wards. These areas were chosen because of their heavy involvement in clinical teaching and because they employ different clinical supervisory structures (in EM, trainees report to the attending physician, whereas in GIM, senior trainees supervise junior trainees in a "hierarchical" supervisory structure).

Study participants were clinical teaching team members in GIM and EM, including attending physicians (AP), junior and senior residents (JR and SR), and medical students (MS). Participants at different levels of experience and of both genders were recruited through purposeful sampling.[5] Saturation of the data (the point at which further sampling ceases to yield any new analytic concepts)[7] was the final determinant of sample size.

The study design involved two phases. Phase One involved nonparticipant observations and brief, on-site interviews. Phase Two employed in-depth interviews using video vignette prompts.

In Phase One,[3] nonparticipant observation[8] of 12 teaching teams was performed (seven teams in GIM and five in EM; total of 88 team members observed). Each team was observed for six 3-hour periods during the course of one month (total 216 hours of observation), and a brief (15 minutes) on-site interview was conducted near the end of the month of observations with 65 participants. Details of Phase One

Correspondence: Tara J. T. Kennedy, MD, PhD, Stan Cassidy Centre for Rehabilitation, 800 Priestman St., Fredericton, NB, E3B 0C7, Canada; e-mail: (tara.kennedy@utoronto.ca).

methodology have been previously published.[3]

Phase Two, conducted in the year after Phase One, was designed to refine and expand the emerging understanding of point-of-care competence assessment through in-depth interviews using video prompts. A series of 10 videotaped vignettes was developed (five set in GIM and five in EM), each crafted to present a dilemma relevant to decisions about supervision. The vignettes were based on events which occurred during Phase One observations (with details altered to render the original participants unidentifiable). For example, one vignette portrayed a resident who had ordered an erroneous investigation without checking with the AP.

The 36 Phase Two participants included APs (n = 19), residents (n = 13), and MSs (n = 4). Although the trainee interviews provided some contextual and confirmatory data which were relevant, the present report is drawn primarily from Phase One data and the AP interviews from Phase Two. During the interviews, participants viewed the relevant videos (in the same sequence) and were asked to discuss their opinion of what they would do in response to the dilemma presented in each vignette. To gain insight into both tacit and explicit influences on supervision decisions, participants were asked to discuss the rationale for their responses and to articulate other possible responses to the vignette dilemma and their reasons for rejecting these (the discourse-based interview method).[9]

Interviews were audiorecorded and transcribed without identifying information. Interview transcripts were analyzed both for emergent themes and for preselected themes that had emerged from the Phase One data. Data collection and analysis proceeded simultaneously in an iterative fashion, in which the results of the ongoing data analysis informed the subsequent data collection, as per grounded theory methods.[5] A preliminary coding structure was developed through a recursive reading of the complete data set by two researchers.[5] The full research team (a clinician–educator, a qualitative education scientist, a cognitive psychologist, and a health policy researcher) then discussed and refined the coding structure and participated in higher-level analysis and theory development.

Results

Dimensions of trustworthiness. Analysis revealed that decisions about how much supervision to provide were based on more than simple assessments of clinical skills. Supervising physicians assessed the "trustworthiness" of trainees to act independently, which involved four dimensions: *knowledge and skill, discernment, conscientiousness,* and *truthfulness* (see Table 1). Although clinical knowledge and skill were important to trustworthiness assessments, supervisors also considered whether trainees had the discernment to identify the limits of their competence. They also considered whether trainees were conscientious in identifying all relevant concerns and following through with treatment plans. Truthfulness during

trainees' interactions with their supervisors was another important concern.

Assessment processes. Given that these four dimensions are central to point-of-care competence assessments, we wondered how supervising physicians were assessing these dimensions. What processes did they use to gain insight into these aspects in their trainees? The observational and interview data revealed that the two most important processes used for assessment of trustworthiness were *double-checks* and *language cues*.

The process of *double-checking* the results of a trainee's clinical assessment was ubiquitously described by our participants as an important method of gauging trustworthiness. Supervising physicians often checked the trainee's findings against their own assessment when they repeated elements of the history or physical examination. For example, an EM physician said,

> I guess you want to see whether the history that they have given you from the first two or three patients is the same as the history you get from the patient . . . so they might slant it that the patient is fine . . . and if you go and get the exact same impression for the first few patients, then you just have more confidence in the resident. (AP11 [EM])

We frequently watched this process of double-checking history and physical examination results during our observation periods, and we noted how the amount of double-checking often decreased as a supervisor gained familiarity with a trainee's abilities. In other cases, supervisors checked the

Table 1
Dimensions of Trustworthiness

Dimensions and definitions	Representative transcript excerpts
1. Knowledge and skill Refers to a trainee's relevant knowledge and clinical skills	"So if I just ask them very direct questions about 'what are the main causes of x' and if they can't come up with any reasons, then you know that their knowledge base may also be lacking and they may require closer supervision." AP13 (GIM)
2. Discernment Refers to a trainee's awareness of the limits of his or her clinical knowledge and skill	". . . when he doesn't know he comes and asks me for help, and if that's the case . . . you know if the patient's in danger . . . so I trust him." AP13 (EM)
3. Conscientiousness Refers to a trainee's thoroughness in data gathering and dependability in following through with assigned tasks	"In deciding how much autonomy to provide, it matters how diligent they are in terms of follow-up, understanding that you just have to be obsessive–compulsive about all the particulars and, in fact, willing to give time and effort to collect the information." AP7 (GIM)
4. Truthfulness Refers to the absence of deception in a trainee's interactions with the supervisor	"But very infrequently I would say they report that they did something that they didn't really do. . . . It's very rare. Or if it is more common then I haven't caught it. . . . It has happened, and it was almost like a violation. . . . I felt very untrusting of that person from then on. I would reconfirm everything with every patient. Which is a real pain in the patootie because it takes a lot of time." AP3 (EM)

trainee's clinical assessment against another source, like lab test results or nursing notes. Another EM AP said,

> You gain trust in many ways with residents. . . . Usually I'll check the lab work before they've even presented (the case) to me. So I already have a bit of an idea. I'll read the triage notes and notes from EMS and from the nurses and see if it all jives and their presentation is accurate. (AP3 [EM])

Another method for assessment of trustworthiness used by supervising physicians was the interpretation of *language cues*. Language and communication skills used by trainees in case presentations were not assessed as skills unto themselves but as proxy measures for clinical competence. A GIM participant explained,

> I think the reason a strong oral presentation is important is because it reflects the strength of the resident's knowledge, analysis and clinical reasoning. Not because a strong presentation is critical in and of itself. But in general a disorganized confusing presentation reflects a disorganized, confused (trainee). (AP14 [GIM])

Two main aspects of trainees' communication about cases were interpreted by supervisors as indicators of clinical trustworthiness or lack thereof: *structure and delivery*, and *anticipated information*. The *structure and delivery* of the case presentations made by trainees were considered by supervising physicians to be markers of the clinical knowledge dimension of trustworthiness, as described in the following representative excerpt:

> Well it's the standard case presentation . . . but it's more in the level of detail and the skill with which the individual has elicited the history and then presented it in a way that brings it all together. Rather than simply a list of positives and negatives like the way a computer would generate a list of symptoms . . . that's the kind of presentation you get with a (trainee) who doesn't know how to analyze the case. (AP13 [GIM])

Supervising physicians also looked for *anticipated information* when assessing trainee trustworthiness. A trainee's ability to present relevant information spontaneously, before it was solicited by the supervisor, was considered to be a marker of independent clinical judgment. As a GIM AP said,

> Like if the patient comes with atrial fibrillation, then if the trainee identified the potential causes quickly, identified that the patient is quite stable, and actually gave me options that are reasonable to treat the patient *without me prompting* then I'd have a level of comfort that the trainee spontaneously arrived at the reasonable medical decisions. That would be an example of a trainee that might inspire more confidence with the staff and would receive more autonomy in the future. (AP15 [GIM] [emphasis added])

Spontaneously presented clinical information was also considered to be more truthful, as explained by an EM AP:

> Really good (trainees) answer my questions and alleviate my concerns before I even express them and therefore I know that they are not making it up because they volunteered the information . . . whereas if I asked directly "did you listen to the heart sounds," and they had not previously said it, and then they say "yes," I don't know if they are just saying that to please me. So the strong ones and the confident ones have less supervision because I am less worried. (AP12 [EM])

Discussion

This study provides the first exploration of the process of point-of-care assessments of trainees' competence to provide independent patient care. Results suggest that supervising clinicians consider more than clinical knowledge and skill when deciding how much supervision to provide. Rather, they assess a multidimensional construct that we have termed trustworthiness, which includes concepts of discernment, conscientiousness, and truthfulness as well as clinical skill. Supervisors assess trustworthiness through a process of double-checking trainees' clinical assessments against their own assessments and against other information like nursing notes, as well as a process of assessment of the language used in case presentation. Supervising physicians use cues from trainees' use of language to inform assessments of clinical skill and of truthfulness.

The finding that supervising physicians double-check trainees' clinical findings to inform assessments of trustworthiness is not unexpected. The finding that supervisors use language assessments as a proxy measure of clinical competence warrants further exploration. AP participants discussed their belief that trainees' language use during case presentations is representative of trainees' clinical skills and clinical reasoning. APs' comfort with competence assessments based on language use was demonstrated in their ability to comment quite specifically on the trustworthiness of the trainees depicted in the video vignettes, for whom they clearly had no access to clinical corroboration. As clinical training curricula are evolving, medical educators are grappling with the practical question of how to decide when to entrust trainees with specific professional activities.[10] An understanding of the link between language use and clinical skills could provide a mechanism for assessment of the context-specific clinical competence of trainees that is not currently being captured in formal evaluation processes.

Theoretical support for the correlation between language use and clinical reasoning can be found in the domain of rhetoric. Rhetorical theory posits that language does not only describe, but also accomplishes action.[11,12] Embedded within the language practices of a professional community are the knowledge, values, and perspectives associated with that community.[13] From a rhetorical perspective, the process of learning to talk like a physician influences the identity of clinical trainees and results in learning to think and feel like a physician. Thus, the language and the thought processes are inextricably linked, and the notion that clinical language use provides a window on clinical thinking processes is a plausible one in the context of rhetorical theory. Preliminary applied studies provide support for the link between case presentation skills and clinical skills, as well. Medical students who provided problem-based case presentations were found to have used more positive history-taking behaviors during their clinical assessments.[14,15] Further study of the association between language markers used in case presentations with other measures of competence will be necessary to understand the practical utility of assessing language as a proxy measure for clinical reasoning.

A cautionary note regarding the use of language cues to inform trustworthiness assessments can be found in the person perception literature in the domain of social psychology. This large body of literature has explored the process of assessing an unfamiliar person's intelligence, which is affected by variables

like responsiveness to a conversation partner and patterns of eye gaze.[16] Similar issues are raised in applied research on oral examinations in medical training. Scores on oral examinations have been shown to be influenced by factors like dress[17] and ethnicity.[18] This research raises the possibility that although APs perceive that they are assessing language cues, they might be influenced by other factors that are less closely linked to the dimensions of trustworthiness than is language. Studies involving experimental manipulation of such factors during case presentations would be useful to tease out the importance of nonlanguage cues to point-of-care trustworthiness assessments.

The use of language assessment to inform decisions about how much supervision to provide could have significant practical implications for clinical education. These implications would be most salient for trainees whose language and clinical skills may not be as closely linked as usual, like international medical graduates (IMGs). IMGs whose language during case presentations does not contain the expected markers of competence might, as a result, receive more intense clinical oversight than their clinical skills warrant. This could lead to a loss of the educational benefit associated with clinical independence, and a relative educational disadvantage. On the other hand, the language markers indicating clinical insecurity might also be difficult to interpret when a trainee's case presentation skills have been learned in a different language context, thus complicating the process of point-of-care assessment of trustworthiness for supervising physicians. The potentially unique impact of point-of-care language assessment on the educational experience of IMGs is an important area for future study.

The issues of observer effect[19] and transferability[20] should be considered when interpreting the results of this analysis. The study design incorporated measures to minimize the impact of observer effect on the results. First, observations were conducted by a consistent researcher for periods of three hours or more so that team members could acclimatize to the observer's presence. Second, participants consented to having all AP–trainee discussions recorded, but they were not made aware of the specific focus on assessment of

competence until after the observations were completed, so that this aspect of their work could not be specifically altered. When considering transferability, it should be noted that these data were collected in teaching hospitals on medical services. These data were collected at three separate institutions, in two distinct clinical settings (GIM and EM), and from a broad range of participants, which enhances the transferability of the analysis. However, it remains to be tested whether or not the theory of trustworthiness assessment will be relevant in different settings like psychiatry clinics or pediatric hospitals.

Supervising physicians regularly make rapid point-of-care assessments of trainees' competence to deliver independent patient care. In making these decisions about the nature of supervision required in a given clinical situation, supervising physicians assess trainees' trustworthiness, a multidimensional construct which includes clinical skill and knowledge, discernment, conscientiousness, and truthfulness. Two techniques used to rapidly assess trustworthiness include double-checking trainees' clinical findings and assessing trainees' language use, which is considered a proxy measure of clinical competence. Point-of-care assessment of competence has important implications for medical trainee education and patient safety in clinical teaching contexts. This initial exploration of point-of-care trustworthiness assessments provides an important framework for future research to understand the relationship between supervision practices, trainee education, and patient safety.

Acknowledgments

References
1 Association of American Medical Colleges. Compact Between Resident Physicians and Their Teachers. Available at: (http://www.aamc.org/meded/residentcompact). Accessed June 12, 2008.

2 The Royal College of Physicians and Surgeons of Canada. General Standards of Accreditation. Available at: (http://rcpsc.medical.org/residency/accreditation/genstandards_e.pdf). Accessed June 12, 2008.

3 Kennedy TJT, Lingard L, Baker GR, Kitchen L, Regehr G. Clinical oversight: Conceptualizing the relationship between supervision and safety. J Gen Intern Med. 2007;22:1080–1085.

4 Epstein RM. Assessment in medical education. N Engl J Med. 2007;356:387–396.

5 Glaser B, Strauss A. The Discovery of Grounded Theory: Strategies for Qualitative Research. Chicago, Ill: Aldine Pub Co.; 1967.

6 Charmaz K. Grounded theory: Objectivist and constructivist methods. In: Denzin N, Lincoln Y, ed. Handbook of Qualitative Research. 2nd ed. Thousand Oaks, Calif: Sage; 2000:509–535.

7 Morse JM. The significance of saturation. Qual Health Res. 1995;5:147–149.

8 Hammersley M, Atkinson P. What is ethnography? In: Ethnography: Principles in Practice. 2nd ed. London, UK: Routledge; 1995.

9 Odell L, Goswami D, Herrington A. The discourse-based interview: A procedure for exploring the tacit knowledge of writers in nonacademic settings. In: Odell L, Goswami D, ed. Writing in Nonacademic Settings. New York, NY: Guilford Press; 1985.

10 ten Cate O. Trust, competence, and the supervisor's role in postgraduate training. BMJ. 2006;333:748–751.

11 Miller CR. Genre as social action. Q J Speech. 1984;70:151–167.

12 Lingard L, Haber RJ. Teaching and learning communication in medicine: A rhetorical approach. Acad Med. 1999;74:507–510.

13 Burke K. Language as Symbolic Action: Essays on Life, Literature, and Method. Berkeley, Calif: University of California Press; 1966.

14 Chang RW, Bordage G, Connell KJ. The importance of early problem representation during case presentations. Acad Med. 1998; 73(10 suppl):S109–S111.

15 Hasnain M, Bordage G, Connell KJ, Sinacore JM. History-taking behaviors associated with diagnostic competence of clerks: An exploratory study. Acad Med. 2001;76(10 suppl):S14–S17.

16 Murphy NA, Hall JA, Colvin CR. Accurate intelligence assessments in social interactions: Mediators and gender effects. J Pers. 2003;71:465–493.

17 Burchard KW, Rowland-Morin PA, Coe NP, Garb JL. A surgery oral examination: Interrater agreement and the influence of rater characteristics. Acad Med. 1995;70:1044–1046.

18 Roberts C, Sarangi S, Southgate L, Wakeford R, Wass V. Oral examinations—Equal opportunities, ethnicity, and fairness in the MRCGP. BMJ. 2000;320:370–375.

19 Holden JD. Hawthorne effects and research into professional practice. J Eval Clin Pract. 2001;7:65–70.

20 Lincoln Y, Guba EG. Naturalistic Inquiry. Thousand Oaks, Calif: Sage Publications, Inc.; 1985.