

Workplace-Based Assessment of Internal Medicine Resident Diagnostic Accuracy

MICHAEL D. JAIN, MD, PhD
 GEORGE A. TOMLINSON, PhD
 DANICA LAM, MD
 JESSICA LIU, MD
 DEEPTI DAMARAJU, MD
 ALLAN S. DETSKY, MD, PhD
 LUKE A. DEVINE, MD, MHPE

Abstract

Background Making an accurate diagnosis is a core skill residents must develop. Assessments of this skill and decisions to grant residents clinical independence often are based on global impressions. A workplace-based assessment of diagnostic accuracy could be a useful part of a competency-based assessment program and could inform decisions about granting residents independence.

Innovation We developed a method for measuring diagnostic accuracy that was integrated into the workflow of internal medicine residents and attending physicians.

Methods Four senior medical residents and 6 attending physicians working in the internal medicine clinical teaching unit of a tertiary hospital participated in this study. To determine their diagnostic accuracy, residents documented a leading diagnosis for each patient they evaluated in the emergency department. After reviewing each case with the resident and after

examining the patient, the resident's attending physician documented the diagnosis. Discharge diagnosis was determined by retrospective chart review to allow determination of resident and attending physician diagnostic accuracy. Data were collected for 240 consecutive patients referred for a medicine consultation.

Results Resident diagnostic accuracy was 66% (95% CI 60–72), whereas attending physician accuracy was significantly higher at 79% (95% CI 74–84, $P < .001$). By logistic regression, the accuracy of the attending physician was found to be influenced by the accuracy of the resident. Participants felt this process motivated them to improve their clinical reasoning.

Conclusions Measuring resident diagnostic accuracy provides information that could be used in a competency-based assessment program to provide feedback and motivation to stimulate performance improvement.

Michael D. Jain, MD, PhD, is a Resident, Department of Medicine, University of Toronto; **George A. Tomlinson, PhD**, is Associate Professor, Institute of Health Policy, Management and Evaluation, University of Toronto; **Danica Lam, MD**, is a Resident, Department of Medicine, University of Toronto; **Jessica Liu, MD**, is a Resident, Department of Medicine, University of Toronto; **Deepti Damaraju, MD**, is a Resident, Department of Medicine, University of Toronto; **Allan S. Detsky MD, PhD**, is Professor, Institute of Health Policy, Management, and Evaluation, University of Toronto, Department of Medicine, University of Toronto, and Departments of Medicine, Mount Sinai Hospital and University Health Network; and **Luke A. Devine, MD, MHPE**, is a Lecturer, Department of Medicine, University of Toronto, and Departments of Medicine, Mount Sinai Hospital and University Health Network.

Funding: The authors report no external funding source for this study.

Conflict of Interest: The authors declare they have no competing interests.

The authors would like to thank T. Borschel, S. Gandhi, M. Otremba, Y. Patel, and H. Vellend for their participation in this study and D. Panisko for comments on earlier versions of the manuscript.

Corresponding author: Luke A. Devine, MD, MHPE, University of Toronto, 600 University Avenue, Suite 427, Toronto, ON, Canada M5G 1X5, ldevine@mtsinai.on.ca

Received November 27, 2013; revision received March 12, 2014; accepted March 24, 2014.

DOI: <http://dx.doi.org/10.4300/JGME-D-13-00431.1>

Introduction

A successful internist possesses many attributes, but the ability to make an accurate diagnosis is paramount. Diagnostic acumen emerges as a major theme when master clinicians are interviewed about clinical practice,¹ and the public expects that physicians make accurate diagnoses.² Currently, most measures of diagnostic excellence are anecdotal; masterful clinicians are identified not by any metric but by peer recognition and reputation. A practical means of measuring diagnostic accuracy in clinical practice is lacking.

In the conventional clinical apprenticeship model, a trainee is granted progressive independence in patient care responsibilities over time, rather than after achieving a prespecified level of competence.^{3,4} Independence for residents is often significantly increased during on-call shifts, when they receive less direct supervision. The degree of independence afforded is determined by clinical supervisors and by the trainees themselves. This can be problematic because supervisors are often insufficiently

critical of trainees they know well and therefore trust more than they should,^{4,5} and trainees may not ask for help when necessary because of inaccurate self-assessment of their diagnostic accuracy⁶ or a desire to maintain independence and reputation.³ Measurement of trainees' diagnostic accuracy could be a valuable tool in the assessment of competence. As part of the competency-based education movement, there has been a focus on developing methods to assess what physicians do in actual clinical practice.⁷ Arriving at an accurate diagnosis requires integration of several of the Accreditation Council for Graduate Medical Education competencies, including medical knowledge, patient care, and interpersonal and communication skills,⁸ whereas the exercise of measuring, reflecting, and improving on patient care practices embodies practice-based learning and improvement.⁹ A workplace-based assessment of diagnostic accuracy could be a helpful component of a competency-based assessment program and could inform decisions about progressive independence.

We sought to develop a method to determine senior on-call internal medicine (IM) residents' diagnostic accuracy that could easily be incorporated into the regular workflow.

Methods

Setting and Participants

This study was conducted between July 2010 and April 2011 at Mount Sinai Hospital, a tertiary care hospital affiliated with the University of Toronto. The study was embedded in the existing on-call system, where a single senior IM resident evaluated patients referred by the emergency department (ED) physician during 24 consecutive hours (8 AM to 8 AM). These referrals were formally reviewed by the attending physician between 6:30 AM and 10 AM on the day following the start of the on-call shift. Infrequently, in the interests of expediency or patient safety, a referred patient's case might have been reviewed prior to that time. The attending physician supervised the same resident for a consecutive 1-month period, during which time the resident would be on call 6 to 8 times.

During the study period, 4 senior (postgraduate year-2) residents and 6 attending physicians volunteered to participate. All attending physicians and 2 residents participated for 1 month, and 2 residents participated for 2 months.

The study was approved by the Mount Sinai Hospital Research Ethics Board.

Program Description

For each patient, the resident was asked to explicitly request the "referring diagnosis" from the ED physician.

What was known

Developing diagnostic skill is a critical aim of resident education.

What is new

Residents' diagnoses were accurate in two-thirds of cases, and faculty diagnostic accuracy was lower when the residents' initial diagnoses were wrong.

Limitations

A single-site study limits generalizability; residents' diagnoses were not independent due to prior diagnosis by emergency department.

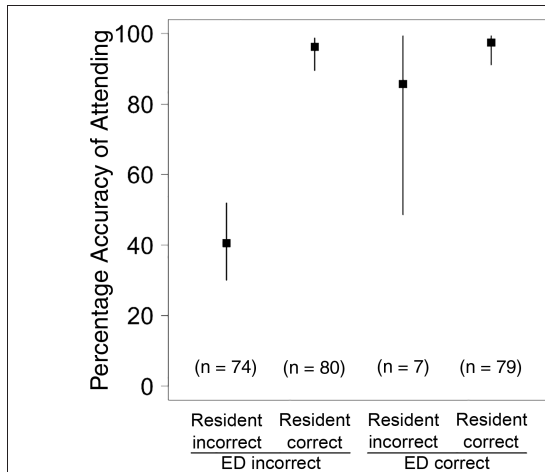
Bottom line

Although the metrics used to assess diagnostic accuracy are not perfect, their use in everyday practice will enhance the quality and safety of care.

The resident documented each referring diagnosis in a log book and then conducted a clinical evaluation, which included taking a history, performing a physical examination, ordering and interpreting investigations, and consulting other services when required. Prior to reviewing the evaluation with the attending physician, the resident documented a specific diagnosis in the log book. If the patient's condition had more than 1 diagnosis, the resident was asked to document the diagnosis most relevant to the patient's presentation at the hospital. Next, the attending physician would review the case with the resident and would round on the patient. After rounds, the attending physician would document the patient's diagnosis in the same log book, having full knowledge of the resident's diagnosis. At the conclusion of the study, informal feedback about the process was sought from residents and attending physicians.

Discharge diagnosis was determined through retrospective chart review. Discharge summaries, radiology and/or pathology reports, blood culture results, and other diagnostic tests were used to confirm or refute the attending physician's admitting diagnosis. When no contrary information was found, the attending physician's admission diagnosis was used as the discharge diagnosis. The term "diagnostic accuracy" was used to denote the percentage of cases in which a physician's initial diagnosis matched the discharge diagnosis. Patients referred to the IM service often have complex medical issues. It was therefore expected that diagnostic accuracy would not approach 100%. Our method was selected as it allowed for a simplified assessment of diagnostic accuracy that could be used to assess and potentially be used to help improve clinical reasoning.

The McNemar test of equality of paired proportions was used to determine whether the level of diagnostic accuracy differed among ED physicians, residents, and



FIGURE

ATTENDING PHYSICIANS' DIAGNOSTIC ACCURACY

Results are plotted as percentages with 95% CIs. The effects of the emergency department (ED) physician's and resident's diagnoses on the specific diagnosis of the attending physician are each significant ($P < .001$).

attending physicians. Fisher exact tests were used to compare accuracy among the 4 residents and among the 6 attending physicians. Logistic regression was used to determine whether the attending physician's accuracy was related to the accuracy of the ED physician and the resident. All analyses were carried out with R software version 2.15.2 (R Foundation for Statistical Computing), and a P value of $< .05$ was used to determine statistical significance.¹⁰

Results

Over the study period, we collected data for 240 consecutive patients who were clinically evaluated. All patients had been discharged from the hospital before the completion of data analysis. Electronic health records, reviewed to determine discharge diagnosis, took less than 5 minutes per patient, with the exception of the occasional patient with a long and complex hospitalization.

The residents' diagnoses matched the discharge diagnoses in 66% of cases (95% CI 60–72), whereas 79% of diagnoses (95% CI 74–84) matched those of the attending physicians (attending physician versus resident, $P < .001$). There were no statistically significant differences among the performances of individual residents ($P = .79$) or among that of individual attending physicians ($P = .06$). The ED physicians' referring diagnoses matched the discharge diagnosis in 36% of cases (95% CI 30–42).

The influence of the ED physician's and resident's diagnostic accuracy on the attending physician's accuracy is shown in the FIGURE. When the resident was correct, the attending physician made the correct diagnosis in 96% of

cases (95% CI 93–99). When the resident was incorrect, the attending physician was correct in only 44% of cases (95% CI 34–55). When the ED physician and resident were correct, attending physicians were correct in 97% of cases. When the ED physician and resident were incorrect, attending physician accuracy was only 41%. Logistic regression analysis, with attending physician diagnostic accuracy as the dependent variable, showed that the ED physician's accuracy and the resident's accuracy predicted the performance of the attending physician, with the resident's diagnostic accuracy being a much stronger predictor.

All residents and attending physicians found the process to be straightforward and felt it did not significantly impact their efficiency. Residents commented that the method forced them to not simply list the differential diagnosis for a patient's problem when presenting a case but to use clinical reasoning to determine what they felt was the most likely diagnosis, something they were not previously doing. They also felt that the method motivated them to endeavor to improve their performance and made them engage more frequently in discussions regarding clinical reasoning with their attending physicians.

Discussion

To our knowledge, this is the first effort to measure the diagnostic accuracy of IM residents and attending physicians in actual clinical practice. Overall, the diagnostic accuracy of attending physicians was 79%, whereas residents were accurate 66% of the time. We found attending physicians' accuracy to be similar to that of neurology attending physicians in another study that assessed referrals from the ED.¹¹ In our study, the appearance of a lower level of accuracy for the ED physician likely occurred because they often, and appropriately, made the referrals early in the diagnostic process, before much of the relevant clinical data had been gathered.

It was not surprising that the diagnostic accuracy of residents and attending physicians was well below 100%. Many patients admitted to the IM service present complex diagnostic dilemmas and undergo extensive workup after admission, the results of which determine the ultimate discharge diagnosis. Attending physician diagnostic accuracy can be used as a benchmark that residents strive to reach as they develop clinical competence and proficiency. By comparing their diagnoses to those of their attending physicians, residents received explicit feedback on the outcome of their diagnostic reasoning. This often stimulated discussion about the process by which the attending physician arrived at their diagnosis and strategies the resident could use to improve their accuracy. Because it

does not require a substantial investment of resident's or attending physician's time, our method allows for a relatively continuous workplace-based assessment across a wide range of authentic clinical contexts.

This method could also be used to study physician error. We found that when the resident and ED physician made a misdiagnosis, the attending physician was dramatically less accurate (41%) than when they were both correct (97%) or when just the resident was correct (96%). This reduction in accuracy could simply be attributed to the more difficult cases, but bias also is likely to play a role. Early misdiagnoses present a potent combination of anchoring bias and confirmation bias.¹² Measurement of diagnostic accuracy highlighted such cases, potentially producing teaching moments regarding clinical reasoning strategies and pitfalls.

This study has a number of limitations. It was conducted at a single academic hospital, reducing generalizability. The diagnostic accuracy levels of the ED physician, resident, and attending physician were not independent, as the resident used information provided by the ED physician and the attending physician made a specific diagnosis after receiving the resident's diagnosis. Another limitation is that there is no "gold standard" for determining a patient's diagnosis. Studies in other specialties that have measured diagnostic accuracy in practice have used various comparisons, including operative findings in surgical specialties^{13,14} and attending physician confirmation of residents' diagnoses in radiology.^{15,16} Within a discipline as broad as IM, many diagnoses are based not on a single piece of laboratory, pathologic, or radiographic data but on a constellation of clinical findings. In our study, the admitting diagnosis of the attending physician was used as the final diagnosis if no further information was found to contradict this diagnosis, a method similar to that used in a study examining diagnostic accuracy in neurology.¹¹

Our approach generated useful information about resident diagnostic accuracy, which could be incorporated into the usual clinical practice of residents and attending physicians. The findings show that committing to a specific diagnosis at admission can be a simple way to distill a complex diagnostic process involving multiple clinical competencies into something that can be measured and returned to the resident as performance feedback. Future work should be directed at evaluating this method on a larger scale to determine its ability to discriminate among various levels of resident performance. The impact of diagnostic accuracy on patient management decisions and

outcomes should also be examined. Discussion regarding the expected level of residents' diagnostic accuracy should occur before residents are awarded specific clinical responsibilities.

Conclusion

This method of measuring diagnostic accuracy could be incorporated into a larger program of assessment within a competency-based education framework to help inform decisions regarding when and how much independence should be granted to an on-call resident. These decisions are vital to optimizing patient safety and resident education. While no metric will be perfect, it is time to begin incorporating measures of diagnostic accuracy into everyday practice.

References

- 1 Christmas C, Kravet SJ, Durso SC, Wright SM. Clinical excellence in academia: perspectives from masterful academic clinicians. *Mayo Clin Proc.* 2008;83(9):989–994.
- 2 Detsky AS. What patients really want from healthcare. *JAMA.* 2011;306(22):2500–2501.
- 3 Kennedy TJ, Regehr G, Baker GR, Lingard LA. Progressive independence in clinical training: a tradition worth defending? *Acad Med.* 2005;80(suppl 10):106–111.
- 4 Sterkenburg A, Barach P, Kalkman C, Gielen M, ten Cate O. When do supervising physicians decide to entrust residents with unsupervised tasks? *Acad Med.* 2010;85(9):1408–1417.
- 5 Stroud L, Herold J, Tomlinson G, Cavalcanti RB. Who you know or what you know? Effect of examiner familiarity with residents on OSCE scores. *Acad Med.* 2011;86(suppl 10):8–11.
- 6 Jones R, Panda M, Desbiens N. Internal medicine residents do not accurately assess their medical knowledge. *Adv Health Sci Educ Theory Pract.* 2008;13(4):463–468.
- 7 Van der Vleuten CP, Schuwirth LW, Scheele F, Driessen EW, Hodges B. The assessment of professional competence: building blocks for theory development. *Best Pract Res Clin Obstet Gynaecol.* 2010;24(6):703–719.
- 8 Mylopoulos M, Lohfeld L, Norman GR, Dhaliwal G, Eva KW. Renowned physicians' perceptions of expert diagnostic practice. *Acad Med.* 2012;87(10):1413–1417.
- 9 Ziegelstein RC, Fiebach NH. "The mirror" and "the village": a new method for teaching practice-based learning and improvement and systems-based practice. *Acad Med.* 2004;79(1):83–88.
- 10 R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2012. <http://www.R-project.org/>. Accessed February 22, 2013.
- 11 Moeller JJ, Kurniawan J, Gubitz GJ, Ross JA, Bhan V. Diagnostic accuracy of neurological problems in the emergency department. *Can J Neurol Sci.* 2008;35(3):335–341.
- 12 Croskerry P. Achieving quality in clinical decision making: cognitive strategies and detection of bias. *Acad Emerg Med.* 2002;9(11):1184–1204.
- 13 Jo YH, Kim K, Rhee JE, Kim TY, Lee JH, Kang SB, et al. The accuracy of emergency medicine and surgical residents in the diagnosis of acute appendicitis. *Am J Emerg Med.* 2010;28(7):766–770.
- 14 Kharbada AB, Fishman SJ, Bachur RG. Comparison of pediatric emergency physicians' and surgeons' evaluation and diagnosis of appendicitis. *Acad Emerg Med.* 2008;15(2):119–125.
- 15 Flink CC, Mueller JS, Kiproff PM. Impact of resident call eligibility on major discrepancy rate. *Acad Radiol.* 2010;17(10):1299–1301.
- 16 Ruchman RB, Jaeger J, Wiggins EF III, Seinfeld S, Thakral V, Bolla S, et al. Preliminary radiology resident interpretations versus final attending radiologist interpretations and the impact on patient care in a community hospital. *AJR Am J Roentgenol.* 2007;189(3):523–526.