



A Comparison of Resident-Completed and Preceptor-Completed Formative Workplace-Based Assessments in a Competency-Based Medical Education Program

Jonathon R. Lee, PhD; Shelley Ross, PhD

BACKGROUND AND OBJECTIVES: In competency-based medical education (CBME), should resident self-assessments be included in the array of evidence upon which summative progress decisions are made? We examined the congruence between self-assessments and preceptor assessments of residents using assessment data collected in a 2-year Canadian family medicine residency program that uses programmatic assessment as part of their approach to CBME.

METHODS: This was a retrospective observational cohort study using a learning analytics approach. The data source was archived formative workplace-based assessment forms (fieldnotes) stored in an online portfolio by family medicine residents and preceptors. Data came from three academic teaching sites over 3 academic years (2015-2016, 2016-2017, 2017-2018), and were analyzed in aggregate using nonparametric tests to evaluate differences in progress levels selected both within and between groups.

RESULTS: In aggregate, first-year residents' self-reported progress was consistent with that indicated by preceptors. Progress level rating on fieldnotes improved over training in both groups. Second-year residents tended to assign themselves higher ratings on self-entered assessments compared with those assigned by preceptors; however, the effect sizes associated with these findings were small.

CONCLUSIONS: Although we found differences in the progress level selected between preceptor-entered and resident-entered fieldnotes, small effect sizes suggest these differences may have little practical significance. Reasonable consistency between resident self-assessments and preceptor assessments suggests that benefits of guided self-assessment (eg, support of self-regulated learning, program efficacy monitoring) remain appealing despite potential risks.

(Fam Med. 2022;54(8):599-605.)

doi: 10.22454/FamMed.2022.854689

Should resident self-assessments be included as part of the array of evidence that is considered by clinical competence committees when making summative progress decisions? The answer

to this question is not entirely clear, and more investigation is needed. Learner self-assessment, including self-reflection on progress in clinical skill development, is recommended as part of programmatic

assessment in competency-based medical education (CBME).^{1,2} These self-assessments are meant to be low-stakes—they are primarily formative, but may be included in the evidence used to make summative (high-stakes) decisions. However, the use of self-assessments has been questioned given the poor ability of learners (and physicians in general) to not only accurately identify weaknesses in their own performance, but to rectify deficiencies when they are found.³⁻⁶ Some authors suggest that self-assessment may even be harmful to learning if it is uninformed and uncalibrated by an external assessor.^{7,8} Others have suggested that the inconsistencies between learner self-assessment and teacher judgements of competence make it inadvisable to use learner self-assessments in decision-making about competence.⁹ For these reasons, it is important that learner self-assessments of progress be monitored for both accuracy (ie, consistency with assessors) and change over the course of training.

Despite these cautions, there is value in including learner self-assessments as part of the evidence of competence in programmatic

From the Faculty of Medicine & Dentistry (Dr Lee) and Department of Family Medicine (Dr Ross), University of Alberta, Edmonton, Alberta, Canada.

assessment in CBME programs. Documenting and making judgements about their own competence during training can provide an opportunity for residents to build skills in self-regulated learning—particularly when embedded in a program that incorporates guided self-assessment. In guided or directed self-assessment,¹⁰ residents are taught to use external evidence to calibrate self-assessment of their own strengths and knowledge gaps. Instruction in these skills is delivered through both explicit means, such as comparing their own performance against assessment rubrics,¹¹ and implicit means, such as role-modeling by clinical teachers.^{12,13} As physicians need to be effective lifelong learners, being able to accurately self-assess is a valuable skill in knowing what continuing professional education to pursue to best serve patients and communities.

The conflicting views about the accuracy of resident self-assessment pose a challenge to educators when coupled with the seemingly contrary recommendations about including such assessments in programmatic assessment frameworks in CBME. Adding further confusion is the lack of published research examining accuracy of self-assessments in formative contexts. There is a pressing need for research that specifically compares self-assessment data to external assessor data in the context of programmatic assessment in CBME. Our residency program offers a unique opportunity to meet this need: as with all family medicine residency programs in Canada, training lasts for 2 years, facilitating examination of data across the full training program. Further, our program is large, accepting 75-85 residents each year, which results in large amounts of available data. Finally, the programmatic assessment framework used in our program (called CBAS¹⁴; Figure 1) intentionally includes resident self-entered formative assessments as part of the overall assessment data set. All data are collected electronically (in an online portfolio called eCBAS),

facilitating learning analytics. Residents and preceptors are provided with resources and tips about best practices in formative feedback and self-assessment (although uptake varies widely).

In this study, we used deidentified data from the eCBAS portfolio to examine formative assessment forms that were either preceptor-entered or resident-entered. We conducted our comparison at an aggregate level as our goal for this study was to explore resident self-assessment trends as compared to their preceptors, rather than individual resident self-assessment behavior. Specifically, at an aggregate level, we examined the judgments of progress level selected by residents on formative assessment forms in comparison to preceptor judgements of progress level over the course of residency training.

Methods

Setting and Participants

Family medicine residency training in Canada is a 2-year postgraduate program that follows immediately after the completion of undergraduate medical education. In the subject program, residents are assigned to a teaching site that serves as their home clinic for the duration of the program. Residents also have clinical training experiences at multiple other clinical sites.

We extracted data for this retrospective observational study from a database of archived formative workplace-based assessments called fieldnotes. Fieldnotes are brief captures of feedback about resident performance in a clinical interaction with a patient.¹⁵ This information includes a judgment of the competence demonstrated by the resident and an indication of the competency that was observed (called a Sentinel Habit in the CBAS framework). Fieldnotes can be entered into an online electronic portfolio (eCBAS) by either a resident (ie, self-assessment) or a preceptor.¹⁴ The residency program expects that each resident receive at least one preceptor-entered fieldnote each week that the resident trains

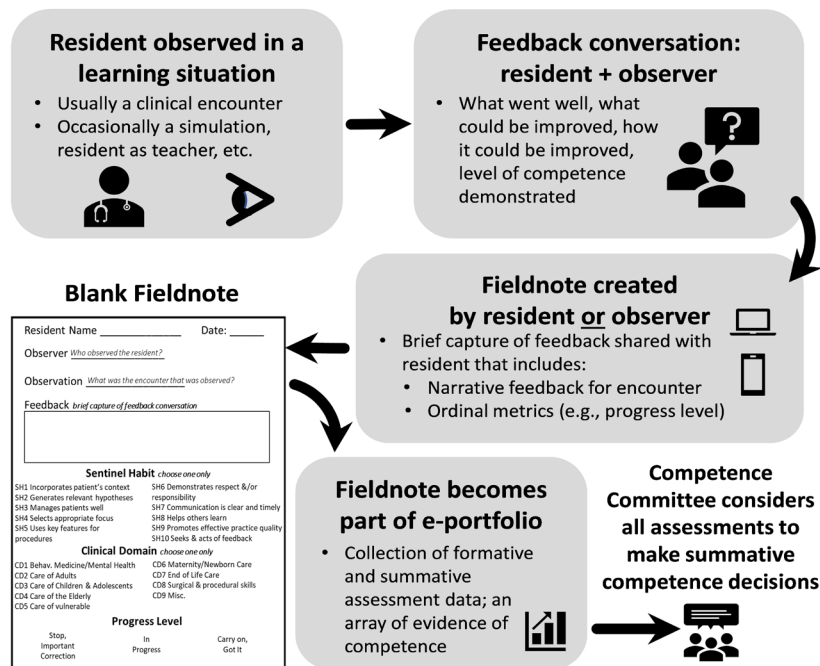
in family medicine clinics (approximately 40 per year); residents are encouraged to enter a similar number of self-assessment fieldnotes during their training. Fieldnotes can also be created by staff, allied health professionals, or patients (via preceptors).¹⁴ Archived fieldnotes indicate the date of entry and who entered a fieldnote, allowing for categorization as self-assessment (ie, resident-created) or preceptor-entered. In this study, we analyzed deidentified fieldnotes (all names and identifying information removed) that were uploaded to eCBAS by preceptors and residents across three academic teaching sites over 3 academic years (2015-2016, 2016-2017, 2017-2018).

This research was approved by the institutional Human Research Ethics Board.

Outcomes Measured

“Progress level” in CBAS fieldnotes refers to resident progress in the development of skills and behaviors (competencies) important for a practicing physician, and is measured on a three-point scale: (1) “Stop, Important Correction” is selected when urgent corrections to resident performance in clinical interactions are indicated; (2) “In Progress” is selected when residents are considered to be in development but not quite competent; and (3) “Carry On, Got It” is selected when residents are thought to have demonstrated competence. Following a resident-patient interaction, the resident and observing preceptor debrief and discuss the resident’s performance. The discussion may then be captured as a fieldnote by either the preceptor (assessment of the resident) or the resident (self-assessment) and will include the selection of a judgement rating or “Progress Level” (Figure 1). Physician competencies in the CBAS framework, referred to as “Sentinel Habits,” are actionable statements aligned with the roles described in the CanMEDS-FM framework.¹⁶ CanMEDS-FM is the competency framework developed by the College of Family Physicians of Canada. It describes the expected competencies for family

Figure 1: Fieldnotes in the Competency-Based Achievement System



A fieldnote can be entered into the e-portfolio by either the resident or the preceptor. Each fieldnote is about a single observation of the resident and serves as a capture of the feedback conversation between the observer and the resident about what was observed. In addition, each fieldnote includes a point-in-time judgement of how well the resident did during the observed event or encounter. Fieldnotes are intended to be formative assessments; however, all fieldnotes in the electronic portfolio may be considered as part of the accumulated evidence of competence of a resident.

physicians in Canada, across the continuum of training and practice.

Statistical Analysis

We performed all statistical analyses using SPSS 26 (IBM, Armonk, NY). Reported progress levels contained in resident-created and preceptor-created fieldnotes were analyzed in aggregate and compared using non-parametric tests (Pearson χ^2). Where appropriate, we performed post hoc comparisons with Bonferroni correction. For Figure 3, we performed statistical tests on raw progress level frequencies. Data are expressed in graphics as frequencies relative to within-group totals. Our sample size of 6,863 fieldnotes produced an observed statistical power of 0.96 with an α of 0.05.

Results

Over 3 academic years (2015-2016, 2016-2017, 2017-2018), 7,167 fieldnotes were entered at the three

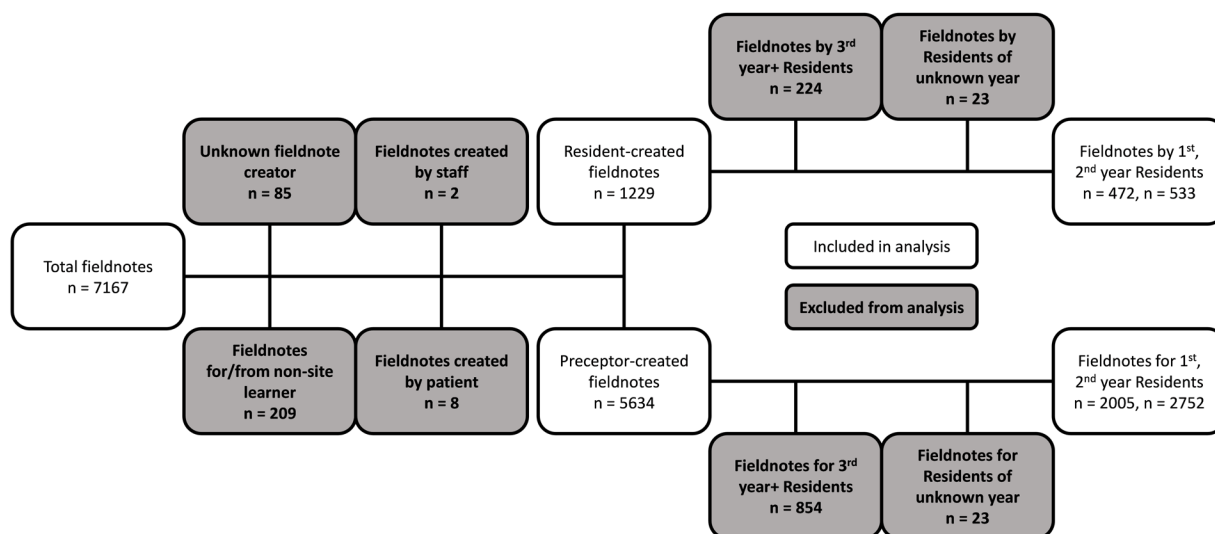
teaching sites included in this study (site 1 $n=2,255$, site 2 $n=2,359$, site 3 $n=2,532$). Of the total data set, 6,863 fieldnotes were examined for comparisons of resident-entered ($n=1,229$) and preceptor-entered ($n=5,634$) fieldnotes. A summary of fieldnote exclusions is shown in Figure 2.

Of the 6,863 fieldnotes analyzed, 17.9% (1,229 of 6,863) were self-assessments generated by residents. Of these, 38.4% (472 of 1,229) were entered by residents in their first year, 43.4% (533 of 1,229) in their second year, and 18.2% (224 of 1,229) in their third or higher year of residency (training may be extended for a variety of reasons including parental leave). While residents overall reported themselves as competent, selecting "Carry On, Got It" in self-assessment fieldnotes more often than "In Progress," second-year residents reported "Carry On, Got It" more often (60.6%; 323 of 533) than

first-year residents (44.5%; 210 of 472). Of the 5,634 preceptor-entered fieldnotes, 35.6% (2,005 of 5,634) were created for first-year residents, 48.8% (2,752 of 5,634) for second-year residents, and 15.6% (877 of 5,634) for residents in their third or higher year. Preceptor assessments of residents were more evenly split, with nearly half of fieldnotes indicating "Carry On, Got It" (48.6%; 2,739 of 5,634) and "In Progress" (48.7%; 2,744 of 5,634). Only 49% (1,349 of 2,752) of preceptor fieldnotes indicated "Carry On: Got It" for second-year residents, up from 43.5% (873 of 2,005) for first-year residents. Overall, preceptors more frequently flagged progress by reporting "Stop: Important Correction."

Figure 3a shows a significant but very weak association between fieldnote creator and the selected progress level (Pearson $\chi^2 [2]=21.04$, $\phi_c=0.05537$, $P<.001$) where "Stop: Important Correction" (post hoc,

Figure 2: Competency-Based Achievement System Fieldnotes Included in Analysis



Fieldnotes were excluded if one or more of the following applied: the role of the fieldnote creator was unknown (n=85); fieldnote was created for residents not assigned to one of the three teaching sites (n=209), and; fieldnote was created by staff (n=2) or patients (n=8). For comparison of first- and second-year residents, fieldnotes that were created by (n=224) or for (n=854) residents in their third or later year (ie, residents in extension of program) were also excluded, as were residents whose year of residency was not recorded (n=23).

$\chi^2 [1]=14.05, P<.001$) and “In Progress” (post hoc, $\chi^2 [1]=10.21, P<.01$) were chosen more often in preceptor-entered fieldnotes, while “Carry On, Got It,” was selected significantly more often in resident-entered fieldnotes (post hoc, $\chi^2 [1]=7.682, P<.01$). Stratifying by year of residency (Figure 3b1 and b2) revealed a weak partial association between fieldnote creator and progress level for both first-year residents (Pearson $\chi^2 [2]=6.348, \phi_c=0.03257, P<.05$) and second-year residents (Pearson $\chi^2 [2]=53.76, \phi_c=0.08124, P<.001$). We found no significant pairwise associations for first-year residents, while second-year residents selected ‘In Progress’ significantly less frequently (post hoc, $\chi^2 [1]=43.86, P<.001$), and “Carry On, Got It” significantly more frequently than was seen on preceptor-entered fieldnotes (post hoc, $\chi^2 [1]=51.21, P<.001$).

Figures 3c1 and c2 illustrate that progress level selected on preceptor-entered fieldnotes differed with resident year of training. A very weak partial association was found (Pearson $\chi^2 [2]=18.84, \phi_c=0.06293, P<.001$) where preceptors selected “In Progress” more often (post-hoc, $\chi^2 [2]=18.15, P<.001$) and “Carry

On, Got It” less often (post hoc, $\chi^2 [2]=13.98, P<.001$) when assessing residents in their first year. Residents displayed a similar—and more sizable—shift in the progress levels they selected on self-entered fieldnotes (Pearson $\chi^2 [2]=28.75, \phi_c=0.1691, P<.001$). As with preceptor-entered fieldnotes, first-year residents chose “In Progress” significantly more often (post hoc, $\chi^2 [2]=28.40, P<.001$) and “Stop, Important Correction” significantly less often (post hoc, $\chi^2 [2]=26.08, P<.001$) than did their second-year counterparts.

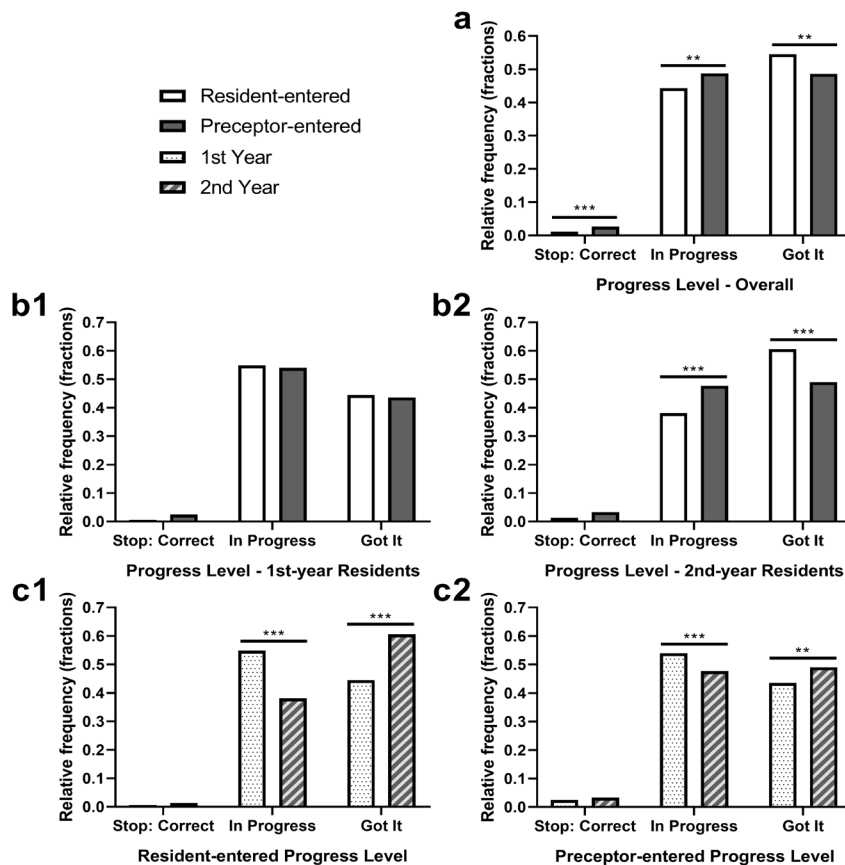
Discussion

This exploratory study contributes to our understanding of the value of formative resident self-assessments in CBME. The findings indicate that (1) progress levels selected in self-assessments were more positive than those selected in preceptor assessments of residents in aggregate data, but that (2) first-year residents’ self-reporting of progress was consistent with preceptors, and that (3) this discrepancy was attributed to second-year residents. Additionally, the data show a within-group shift in recorded progress from “In

Progress” to “Carry On, Got It” for both resident- and preceptor-entered assessments. Notably, effect sizes in our study ranged from very weak ($\phi_c<0.1$) to weak ($0.1<\phi_c<0.3$), suggesting that self-assessment by these residents may not be as subject to bias as suggested in the literature. Despite this, the minor differences that were found between residents and preceptors in progress levels selected may provide some insight into resident-preceptor dynamics in formative assessment in the clinical training environment.

Given the importance of accurate self-assessment,¹² the previously-observed lack of congruency^{3,17–19} between learner self-assessments and preceptor judgements in medical education contexts is concerning. Available data from CBME programs are limited but are generally more mixed in this regard. One study found that in the first postgraduate years, residents in clinic overestimated their performance on self-assessments, a problem that was heightened among the youngest of those physicians.²⁰ Similarly, studies in postgraduate emergency medicine²¹ and ophthalmology²² CBME programs revealed that residents

Figure 3: Reported Progress Level in Competency-Based Achievement System Fieldnotes



Relative frequency of progress levels reported by residents in self-assessments and preceptors in appraisals of residents. (a) Resident-entered progress level compared with preceptor-entered progress level in aggregate data. (b1) Comparison of reported progress between residents and preceptors in fieldnotes for first-year residents and (b2) second-year residents. (c1) Comparison of reported progress in fieldnotes generated by first- and second-year residents and by preceptors for first- and (c2) second-year residents.

overestimated their abilities relative to instructors. Contrary to these findings, resident self-assessments and preceptor assessments were found to be closely correlated in an anesthesiology residency program,²³ while a multicenter study in postgraduate general surgery found minimal disparity between resident self-assessments and faculty appraisals.²⁴ An important observation about most of these studies is that they looked at summative self-assessments of competence—either performance in examination contexts,^{18,19} or self-reported competence on milestones or benchmarks.²¹⁻²⁴ What appears to be missing from the literature are data

about learner self-assessment behavior in formative assessment contexts.

This study adds new information to our understanding of resident self-assessment by observing patterns of behavior in formative clinical workplace assessment. Progress levels reported by first-year residents on their self-entered fieldnotes are generally in-line with progress levels indicated by preceptors in their appraisals of residents. Although second-year residents tended to overestimate their performance in comparison to their preceptors' judgements, effect sizes suggest that this is unlikely to be of practical significance—an important consideration given the

formative nature of the assessments. The lack of large differences between resident-entered and preceptor-entered fieldnotes could reflect successful training in a CBME framework where guided self-assessment is explicitly coached.^{2,13,25,26}

However, it is worth discussing other possible interpretations of the finding that differences between resident self-assessments and preceptor judgements of progress diverge more noticeably as residents advance through postgraduate training. For instance, the divergence in the second year could indicate a deficit in the training of resident meta-cognitive skills (eg, reflection) and

reveal an opportunity for improvement to the CBAS regimen, and an early point of intervention for training in self-regulated learning. Notably, resident self-assessment accuracy (ie, consistency with an external measure) may not improve over time solely with feedback on performance.^{27,28} Instead, calibrating judgment skills directly by providing feedback on resident self-assessments in addition to performance could lead to improved self-assessment accuracy.²⁹⁻³¹

The finding that resident self-assessments showed selection of higher progress levels among second-year residents compared with those in first year may be explained by improved resident self-confidence throughout training. High resident self-confidence does not always equate to higher competence and may even be associated with lower performance and diagnostic errors.^{3,32-34} Notably, the shift in appraisal from “In Progress” to “Carry On, Got It” in preceptor-entered fieldnotes from first to second year suggests that preceptors were receptive to resident growth. Thus, while residents and preceptors both indicated higher progress levels on fieldnotes in second year, residents may have viewed themselves as improving at a higher rate than was actually occurring. This is consistent with findings that training can have a positive effect on both self- and external appraisal of clinical skills, but that this change is more pronounced in learner self-appraisal,^{35,36} similar to what was found in a study of senior ophthalmology residents.²² Altogether, it is noteworthy that transition from a time-based (ie, traditional) medical education approach to one that is competency-based may not necessarily translate to a similar shift in the self-reflective mindset of the learner. In other words, feelings of personal progress may advance (at least in part) due to time spent in program (and perhaps the resultant gain in knowledge, experience, and confidence), and not solely due to the achievement of competency milestones.

A final possible explanation to consider is that self-assessments may reflect an authentic improvement in competency, and that residents are more in tune with personal progress than their preceptors, especially in formative contexts. In general, preceptor-derived judgements are implicitly accepted as more correct, though it has been argued that care should be taken in this assumption due to the potential for reliability issues in instructor assessments.³⁷ In other words, there is a possibility that the divergence in resident-indicated progress level and the level of progress selected by preceptors in fieldnotes about second-year residents could reflect poorly-calibrated preceptors rather than delusional residents. Conversely, it is possible that as residents near the end of training and consistently demonstrate competence, preceptors abdicate some of the responsibility of entering fieldnotes, and encourage residents to self-enter fieldnotes, especially given the formative nature of these assessments. This is a possibility that could be explored in future research.

Our study has several limitations and targets for further study. The biggest limitation is that we conducted analyses at the aggregate level in order to examine overall trends in resident formative self-assessment behavior. Resident self-reflection is subject to many inter- and intrapersonal factors such as gender dyad,³⁸ age, and metacognitive ability; none of these were analyzed here. Additionally, some of the more junior preceptors in the subject program received their residency training within a CBME framework, while more senior preceptors did not; it is therefore possible that some of what we observed could be due to variations in availability or uptake of faculty development. It is also important to note that while our sample size provided adequate statistical power, there is a large discrepancy between resident- and preceptor-entered fieldnotes. While our examination of fieldnotes in aggregate improves generalizability and

provides a top-down view of performance in the CBAS program at an institutional level, it did not allow for adequate analysis of many of these other factors. Future research could examine the impact of site and stakeholder characteristics on resident training, as well as individual differences between the residents or preceptors. The possibility always exists that some outliers could have had an impact on the overall results of the analyses. A deeper dive into the data to examine this question is worth pursuing.

Conclusion

Overall, we observed few practical differences between resident-entered and preceptor-entered fieldnotes. Where statistically significant differences were found, the effect sizes were notably small. This could be taken as an encouraging sign for development of accuracy of formative self-assessment, tempered by development of clinical confidence alongside clinical competence. It is entirely possible that the contrast between our findings and expected differences in comparative appraisals of progress based on existing literature about accuracy of self-assessment may be due to comparing accuracy of self-assessment in formative versus summative contexts. The findings from this study provide new information about resident formative self-assessment in CBME. Specifically, we show low levels of self-reporting bias among residents and highlight the utility of resident formative self-assessments as an effective tool for monitoring learner development.

ACKNOWLEDGMENTS: The authors thank Delane Linkiewich for support in preparing data for analyses.

FUNDING STATEMENT: This study was funded by a Social Sciences and Humanities Research Council (SSHRC) Insight Grant.

CORRESPONDING AUTHOR: Address correspondence to Dr Shelley Ross, 6-10 University Terrace, Edmonton, AB, T6G 2T4. 780-248-1264. 780-492-9181.sross@ualberta.ca.

References

1. van der Vleuten CPM, Schuwirth LWT, Driessen EW, et al. A model for programmatic assessment fit for purpose. *Med Teach*. 2012;34(3):205-214. doi:10.3109/0142159X.2012.652239
2. Lawrence K, van der Goes T, Crichton T, Bethune C, Brailovsky C, Donoff M. Continuous Reflective Assessment for Training (CRAFT): a national model for family medicine. *The College of Family Physicians of Canada*; 2018.
3. Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA*. 2006;296(9):1094-1102. doi:10.1001/jama.296.9.1094
4. Eva KW, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. *Acad Med*. 2005;80(10)(suppl):S46-S54. doi:10.1097/00001888-200510001-00015
5. Igen JS, Regehr G, Teunissen PW, Sherbino J, de Bruin ABH. Skeptical self-regulation: resident experiences of uncertainty about uncertainty. *Med Educ*. 2021;55(6):749-757. doi:10.1111/medu.14459
6. Regehr G, Eva K. Self-assessment, self-direction, and the self-regulating professional. *Clin Orthop Relat Res*. 2006;449:34-38. doi:10.1097/01.blo.0000224027.85732.b2
7. Eva KW, Cunnington JPW, Reiter HI, Keane DR, Norman GR. How can I know what I don't know? Poor self assessment in a well-defined domain. *Adv Health Sci Educ Theory Pract*. 2004;9(3):211-224. doi:10.1023/B:AHSE.0000038209.65714.d4
8. Holmboe ES, Sherbino J, Long DM, Swing SR, Frank JR. The role of assessment in competency-based medical education. *Med Teach*. 2010;32(8):676-682. doi:10.3109/0142159X.2010.500704
9. Baxter P, Norman G. Self-assessment or self-deception? A lack of association between nursing students' self-assessment and performance. *J Adv Nurs*. 2011;67(11):2406-2413. doi:10.1111/j.1365-2648.2011.05658.x
10. Sargeant J, Mann K, van der Vleuten C, Metsemakers J. "Directed" self-assessment: practice and feedback within a social context. *J Contin Educ Health Prof*. 2008;28(1):47-54. doi:10.1002/chp.155
11. Andrade HG, Boulay BA; Heidi Goodrich Andrade. Beth A. Boulay. Role of rubric-referenced self-assessment in learning to write. *J Educ Res*. 2003;97(1):21-34. doi:10.1080/00220670309596625
12. Duffy FD, Holmboe ES. Self-assessment in lifelong learning and improving performance in practice: physician know thyself. *JAMA*. 2006;296(9):1137-1139. doi:10.1001/jama.296.9.1137
13. Sargeant J, Armson H, Chesluk B, et al. The processes and dimensions of informed self-assessment: a conceptual model. *Acad Med*. 2010;85(7):1212-1220. doi:10.1097/ACM.0b013e3181d85a4e
14. Ross S, Poth CN, Donoff M, et al. Competency-based achievement system: using formative feedback to teach and assess family medicine residents' skills. *Can Fam Physician*. 2011;57(9):e323-e330.
15. Donoff MG. Field notes: assisting achievement and documenting competence. *Can Fam Physician*. 2009;55(12):1260-1262. e100-e102.
16. Shaw E, Oandasan I, Fowler N. *CanMEDS-Family Medicine 2017: A competency framework for family physicians across the continuum*. The College of Family Physicians of Canada; 2017.
17. Blanch-Hartigan D. Medical students' self-assessment of performance: results from three meta-analyses. *Patient Educ Couns*. 2011;84(1):3-9. doi:10.1016/j.pec.2010.06.037
18. Cleary TJ, Konopasky A, La Rochelle JS, Neubauer BE, Durning SJ, Artino AR Jr. First-year medical students' calibration bias and accuracy across clinical reasoning activities. *Adv Health Sci Educ Theory Pract*. 2019;24(4):767-781. doi:10.1007/s10459-019-09897-2
19. Graves L, Lalla L, Young M. Evaluation of perceived and actual competency in a family medicine objective structured clinical examination. *Can Fam Physician*. 2017;63(4):e238-e243.
20. Abadel FT, Hattab AS. How does the medical graduates' self-assessment of their clinical competency differ from experts' assessment? *BMC Med Educ*. 2013;13(1):24. doi:10.1186/1472-6920-13-24
21. Goldflam K, Bod J, Della-Giustina D, Tsyrlunik A. Emergency medicine residents consistently rate themselves higher than attending assessments on ACGME milestones. *West J Emerg Med*. 2015;16(6):931-935. doi:10.5811/westjem.2015.8.27247
22. Srikumaran D, Tian J, Ramulu P, et al. Ability of ophthalmology residents to self-assess their performance through established milestones. *J Surg Educ*. 2019;76(4):1076-1087. doi:10.1016/j.jsurg.2018.12.004
23. Ross FJ, Metro DG, Beaman ST, et al. A first look at the Accreditation Council for Graduate Medical Education anesthesiology milestones: implementation of self-evaluation in a large residency program. *J Clin Anesth*. 2016;32:17-24. doi:10.1016/j.jclinane.2015.12.026
24. Watson RS, Borgert AJ, O Heron CT, et al. A multicenter prospective comparison of the Accreditation Council for Graduate Medical Education Milestones: clinical competency committee vs resident self-assessment. *J Surg Educ*. 2017;74(6):e8-e14. doi:10.1016/j.jsurg.2017.06.009
25. Lockyer J, Carraccio C, Chan M-K, et al; ICBME Collaborators. Core principles of assessment in competency-based medical education. *Med Teach*. 2017;39(6):609-616. doi:10.1080/0142159X.2017.1315082
26. Sargeant J, Eva KW, Armson H, et al. Features of assessment learners use to make informed self-assessments of clinical performance. *Med Educ*. 2011;45(6):636-647. doi:10.1111/j.1365-2923.2010.03888.x
27. Kämmer JE, Hautz WE, März M. Self-monitoring accuracy does not increase throughout undergraduate medical education. *Med Educ*. 2020;54(4):320-327. doi:10.1111/medu.14057
28. Raaijmakers SF, Baars M, Paas F, van Merriënboer JGG, van Gog T. Effects of self-assessment feedback on self-assessment and task-selection accuracy. *Metacognition Learn*. 2019;14(1):21-42. doi:10.1007/s11409-019-09189-5
29. de Bruin ABH, Dunlosky J, Cavalcanti RB. Monitoring and regulation of learning in medical education: the need for predictive cues. *Med Educ*. 2017;51(6):575-584. doi:10.1111/medu.13267
30. Händel M, Harder B, Dresel M. Enhanced monitoring accuracy and test performance: incremental effects of judgment training over and above repeated testing. *Learn Instr*. 2020;65:65. doi:10.1016/j.learninstruc.2019.101245
31. Kostons D, van Gog T, Paas F. Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. *Learn Instr*. 2012;22(2):121-132. doi:10.1016/j.learninstruc.2011.08.004
32. Barnsley L, Lyon PM, Ralston SJ, et al. Clinical skills in junior medical officers: a comparison of self-reported confidence and observed competence. *Med Educ*. 2004;38(4):358-367. doi:10.1046/j.1365-2923.2004.01773.x
33. Berner ES, Graber ML. Overconfidence as a cause of diagnostic error in medicine. *Am J Med*. 2008;121(5)(suppl):S2-S23. doi:10.1016/j.amjmed.2008.01.001
34. Meyer AND, Payne VL, Meeks DW, Rao R, Singh H. Physicians' diagnostic accuracy, confidence, and resource requests: a vignette study. *JAMA Intern Med*. 2013;173(21):1952-1958. doi:10.1001/jamainternmed.2013.10081
35. Arnold L, Willoughby TL, Calkins EV. Self-evaluation in undergraduate medical education: a longitudinal perspective. *J Med Educ*. 1985;60(1):21-28. doi:10.1097/00001888-198501000-00004
36. Fünfer SM, Lesevic H, Rosner S, et al. Improved self- and external assessment of the clinical abilities of medical students through structured improvement measures in an internal medicine bedside course. *GMS J Med Educ*. 2016;33(4):Doc59.
37. Brown GTL, Andrade HL, Chen F. Accuracy in student self-assessment: directions and cautions for research. *Assess Educ*. 2015;22(4):444-457. doi:10.1080/0969594X.2014.996523
38. Loeppky C, Babenko O, Ross S. Examining gender bias in the feedback shared with family medicine residents. *Educ Prim Care*. 2017;28(6):319-324. doi:10.1080/14739879.2017.1362665