

Use of Resident-Sensitive Quality Measure Data in Entrustment Decision Making

Citation for published version (APA):

Schumacher, D. J., Martini, A., Sobolewski, B., Carraccio, C., Holmboe, E., Busari, J., Poynter, S., van der Vleuten, C., & Lingard, L. (2020). Use of Resident-Sensitive Quality Measure Data in Entrustment Decision Making: A Qualitative Study of Clinical Competency Committee Members at One Pediatric Residency. *Academic Medicine*, 95(11), 1726-1735. <https://doi.org/10.1097/ACM.0000000000003435>

Document status and date:

Published: 01/11/2020

DOI:

[10.1097/ACM.0000000000003435](https://doi.org/10.1097/ACM.0000000000003435)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Use of Resident-Sensitive Quality Measure Data in Entrustment Decision Making: A Qualitative Study of Clinical Competency Committee Members at One Pediatric Residency

Daniel J. Schumacher, MD, PhD, Abigail Martini, Brad Sobolewski, MD, MEd, Carol Carraccio, MD, MA, Eric Holmboe, MD, Jamiu Busari, MD, PhD, MHPE, Sue Poynter, MD, MEd, Cees van der Vleuten, PhD, and Lorelei Lingard, PhD

Abstract

Purpose

Resident-sensitive quality measures (RSQMs) are quality measures that are likely performed by an individual resident and are important to care quality for a given illness of interest. This study sought to explore how individual clinical competency committee (CCC) members interpret, use, and prioritize RSQMs alongside traditional assessment data when making a summative entrustment decision.

Method

In this constructivist grounded theory study, 19 members of the pediatric residency CCC at Cincinnati Children's Hospital Medical Center were purposively and theoretically sampled between February and July 2019. Participants

were provided a deidentified resident assessment portfolio with traditional assessment data (milestone and/or entrustable professional activity ratings as well as narrative comments from 5 rotations) and RSQM performance data for 3 acute, common diagnoses in the pediatric emergency department (asthma, bronchiolitis, and closed head injury) from the emergency medicine rotation. Data collection consisted of 2 phases: (1) observation and think out loud while participants reviewed the portfolio and (2) semistructured interviews to probe participants' reviews. Analysis moved from close readings to coding and theme development, followed by the creation of a model illustrating theme interaction. Data collection and analysis were iterative.

Results

Five dimensions for how participants interpret, use, and prioritize RSQMs were identified: (1) ability to orient to RSQMs: confusing to self-explanatory, (2) propensity to use RSQMs: reluctant to enthusiastic, (3) RSQM interpretation: requires contextualization to self-evident, (4) RSQMs for assessment decisions: not sticky to sticky, and (5) expectations for residents: potentially unfair to fair to use RSQMs. The interactions among these dimensions generated 3 RSQM data user profiles: eager incorporation, willing incorporation, and disinclined incorporation.

Conclusions

Participants used RSQMs to varying extents in their review of resident data and found such data helpful to varying degrees, supporting the inclusion of RSQMs as resident assessment data for CCC review.

Contemporary clinical competency committees (CCCs) in graduate medical education have the foundational task of reviewing resident assessment data.^{1–3} Literature to date on CCCs has largely focused on 2 areas: (1) detailing how to design a CCC^{4–9} and (2) potential best practices for CCCs based on expert opinion, evidence from fields other than medicine, and personal experiences at

individual residency programs.^{7,10–20} However, extending research on CCCs beyond their structure and process to seek an understanding of how CCCs interpret, use, and prioritize different types of assessment data in their decision making is needed.²⁰ This can help inform the types of assessment data that are provided to CCCs. Preliminary work in this area suggests that individuals may prioritize qualitative assessment data.^{21,22} Furthermore, data that are difficult to interpret can lead to lengthy, labored discussions in an effort to understand it.²³ However, to our knowledge, no study has considered the relative value that CCC members ascribe to different types of assessment data.

Health care quality measures are rarely used in assessment, despite recent calls and efforts to determine how quality measures can be used in graduate medical education.^{24–27} Such measures that are also sensitive to the work that residents perform not only enhance

meaningful involvement of residents in quality improvement work but can also serve as a patient-centered assessment approach, indicating the quality of care provided by residents.²⁸ To this end, Warm and Mathis²⁹ have suggested that Accreditation Council for Graduate Medical Education requirements for residency training should include measuring and improving patient care outcomes. Furthermore, Jaffe and colleagues³⁰ have advocated for educators partnering with institutional quality improvement experts to design resident quality and safety reports. However, many traditional quality metrics may not align well with the work that residents perform.³¹ Paradoxically, quality measures sometimes attributed to supervising physicians may be attributable to residents as well.³² These challenges led our research team to develop resident-sensitive quality measures (RSQMs), which are quality measures that meet 2 criteria: (1) they are likely performed by an individual resident and not another

Please see the end of this article for information about the authors.

Correspondence should be addressed to Daniel J. Schumacher, Division of Emergency Medicine, Cincinnati Children's Hospital Medical Center, MLC 2008, 3333 Burnet Ave., Cincinnati, OH 45226; telephone: (513) 803-2639; email: daniel.schumacher@cchmc.org.

Acad Med. 2020;95:1726–1735.

First published online April 21, 2020

doi: 10.1097/ACM.0000000000003435

Copyright © 2020 by the Association of American Medical Colleges

Supplemental digital content for this article is available at <http://links.lww.com/ACADMED/A925> and <http://links.lww.com/ACADMED/A926>.

member of the team or by the team collectively and (2) they are important to the care of a patient with a given illness of interest.^{27,33}

While initial implementation evidence for RSQMs has been established,^{34,35} how these measures may be used for assessment purposes has not been examined. To probe this area, we explored how individual CCC members interpret, use, and prioritize RSQMs alongside traditional assessment data when making a summative entrustment decision.

Method

Design

We conducted a constructivist grounded theory study to explore how individual CCC members interpret, use, and prioritize RSQM (or quality measure) data when inserted into their usual review processes. Aligning with a constructivist paradigm, we were interested in the views, values, and perspectives of CCC members. Furthermore, we anticipated that a theory of how CCC members interpret, use, and prioritize quality measure data in entrustment decision making could inform next steps in how such information can or should be used for entrustment decision making by CCCs.

We began by purposively sampling members from the pediatric residency CCC at Cincinnati Children's Hospital Medical Center (CCHMC) via email to ensure a representative sample of roles in the program (e.g., associate program directors, faculty without formal leadership roles in the program), gender, generalist versus subspecialist, and perceived leniency versus stringency in making assessment decisions (as determined by program leadership). At the time of this study, the CCHMC pediatric residency CCC had 32 members (22 directors of resident rotations and 10 program leaders, including chief residents).

Many CCCs in pediatrics have each member review portfolios of resident assessment data and bring a summary of findings along with a proposed advancement or entrustment decision to a larger group of CCC members for vetting. The recommendations of individual CCC members may often

stand with little or no additional vetting of these initial decisions for residents who do not have any noted performance concerns. Given these contextual considerations, individual CCC member review decisions play a significant role in summative assessment decisions made for pediatric residents. Therefore, we focused on individual CCC members in this study.

After redacting the resident's name, gender, and training year, we provided participants with a real resident portfolio that contained the traditional assessment data (milestone and/or entrustable professional activity [EPA] ratings as well as narrative comments from 5 rotations) CCC members review before a CCC meeting. The resident portfolio presented to CCC members represented, by the judgment of the research team, a typically developing resident in the middle of training who was generally performing well but had some concerns noted, a few of which were consistent across rotations and the remainder of which were isolated within specific rotations. A synopsis of the resident portfolio is presented in Supplemental Digital Appendix 1 (at <http://links.lww.com/ACADMED/A925>). In addition to the traditional assessment data included, participants were provided RSQM performance data for 3 acute, common diagnoses in the pediatric emergency department (asthma, bronchiolitis, and closed head injury; see Appendix 1 and Supplemental Digital Appendix 2 at <http://links.lww.com/ACADMED/A926>). These novel data are a product of our previous studies^{27,33} and reflect average performance data on these measures from our other previous studies.^{34,35} That is, at the time of this study, these RSQMs had only been used in a study setting. Thus, we were not able to use actual performance data from the resident represented in the assessment portfolio. However, we did use average data from several residents who participated in our previous studies to create a RSQM profile for review that was consistent with the typically developing resident we used in this study.

Data collection

We held sessions with 19 participants (of 21 contacted, 2 individuals did not respond to the invitation to participate) between February and July 2019. Data collection consisted of 2 phases: (1)

observation and think out aloud while participants reviewed the resident portfolio and (2) semistructured interviews to probe participants' reviews. The first phase lasted approximately 7.5–20 minutes, and the second phase lasted approximately 8.5–21 minutes across participant sessions. The entirety of the sessions (both phases) was audiorecorded and professionally transcribed.

Two authors (D.J.S. and A.M.) conducted the sessions and recorded field notes to capture participants' nonverbal nuances, how long participants spent reviewing the portfolio, and key thoughts from participants.

Phase 1: Observation and think out loud. At the outset of each participant session, CCC members were provided the resident portfolio and instructed to review as they normally would, verbalizing their thoughts, ideas, and decisions. We told them we were seeking to understand how they interpret, use, and prioritize assessment data, without specifically mentioning the RSQM data. As such, no time or effort was devoted to orienting the participants to the RSQMs or any of the other data. The decision to not tell participants that we were most interested in how they interpreted, used, and prioritized RSQM data was an effort to avoid participant reactivity,³⁶ that is, participants behaving other than they normally would—in this case, unduly focusing on the RSQMs due to our interests. To provide focus, we asked participants to review the resident portfolio with the goal of providing a summative entrustment decision for the general pediatrics EPA “manage patients with acute, common diagnoses in an ambulatory, emergency, or inpatient setting.”³⁷ We chose this EPA because of its relevance to the RSQMs presented.

Phase 2: Semistructured interview. Once participants completed their review, we conducted a brief, semistructured interview to probe how they interpreted, used, and prioritized the RSQMs in their review and decisions. We also probed how they did or did not differentially use other assessment data to mask our focused interest in the RSQMs and thus minimize response bias.

Both data collection and analysis were iterative—that is, once initial themes

were identified, theoretical sampling was used in subsequent sessions to confirm, disconfirm, and/or elaborate on the evolving results.

Data analysis

Constructivist grounded theory³⁸ posits that meaning is developed through the interaction of participants and researchers. Therefore, we believed that a research team (all authors) made up of individuals with varied experiences with resident education and performance review was important. Our team consisted of clinicians and nonclinicians, current and former residency program leadership, current and former CCC members, a former CCC chair, and education researchers with and without previous scholarly work in the area of CCCs, many of whom have held leadership roles in medical education. Two members of the research team were members of the CCHMC pediatric residency CCC and current CCHMC residency program leaders (S.P., program director, and B.S., assistant program director). These individuals helped to ground our analysis in the context of local nuances when applicable. Our team composition allowed us to negotiate varied perspectives, placing us in a position to recognize the influence of our individual perspectives; to benefit from productive debate when we viewed things differently; and to develop a richer set of insights into the studied phenomena through this dialectic, co-constructive process.

Dedoose, version 8.2.14 (SocioCultural Research Consultants, LLC, Los Angeles, California) was used to aid analysis. Analysis began with individual, close readings of 2 transcripts and field notes from participant sessions by 3 members of the research team (D.J.S., A.M., and L.L.) and then discussing the initial codes, which were then used to analyze another 4 participant sessions. This group (D.J.S., A.M., and L.L.) then met to resolve differences, identify early themes, and determine initial strategies for theoretical sampling. Following the analysis of an additional 7 participant sessions, we coalesced the findings into a preliminary model illustrating how themes interact. When presented to the full research team, we identified additional issues for theoretical sampling. One such issue was the influence of benchmarks on participants' interactions with the RSQM data. Expected performance for

the RSQMs had not been included in the data for the first 16 participant sessions; to theoretically sample this issue, a benchmark of 80% for all measures was provided in the resident portfolio for the final 3 participant sessions.

Iterative data collection (D.J.S. and A.M.) and analysis (D.J.S., A.M., and L.L.) continued in an effort to fully characterize the identified themes (or dimensions), as well as to understand the relationships among them. The latter effort resulted in the articulation of 3 profiles illustrating how the dimensions interact. As our results approached theoretical sufficiency, the analysis was again discussed with the full research team to ensure agreement on final results.

As analysis proceeded following the initial participant sessions, transcripts from subsequent sessions were used to challenge initial codes and themes. We also returned to previous sessions' transcripts repeatedly during analysis to ensure constant comparison.

The CCHMC Institutional Review Board determined this study to be exempt.

Results

Of the 19 participants, 8 (42%) were men and 11 (58%) were women. Nine (47%) were pediatric subspecialists (from separate subspecialties) with the remainder being generalists (general pediatricians, hospitalists, and chief residents; 10 [53%]). Some were in their first year serving as CCC members, and others had served for multiple years.

Participants tended to view and use all types of data (narrative comments from all 5 rotations, milestone and/or EPA ratings from all 5 rotations, and RSQMs from the emergency medicine rotation) substantively in their decision making. However, they demonstrated preferences for the information provided from faculty (versus peer) ratings, narrative comments from all 5 rotations, and sometimes narrative comments from specific rotations (continuity clinic, emergency medicine, and hospital medicine) in particular.

Engagement with the RSQMs

We identified 5 dimensions related to how participants interpret, use, and prioritize RSQMs. We classified these as dimensions because this term has a

notion of measurable extent built into it, and each of the identified dimensions has a spectrum to it.

Ability to orient to RSQMs: Confusing to self-explanatory. While most participants were able to correctly and quickly orient themselves to the RSQMs in an efficient manner, one participant did note: "Well, I barely understand that paper [with the quality measures on it], so I'm probably not going to use it very much to be honest with you." (Participant 9) Some noted they had not seen such data in their previous reviews, but they were still able to interpret the RSQMs after taking some time to process: "Oh man, how do I read this? [a 7-second pause] Okay, these are like the metrics of [whether] they followed certain guidelines of practice. Okay." (Participant 15)

Propensity to use RSQMs: Reluctant to enthusiastic. Some participants were reluctant to use the RSQMs. One barrier to using them was the volume of data they contain:

It's just a lot of information on 1 page for me to parse through ... and try to understand. I think if I was more familiar with it, I would probably find it more helpful. (Participant 9)

Another source of reluctance arose from viewing the RSQMs as simply "checking a box" of giving quality measure feedback to residents. In this case, the RSQMs were not seen as providing sufficient value to be used for assessment purposes. A final source of reluctance arose from viewing the RSQMs as being too specific to one setting (i.e., the emergency department), creating a barrier for giving weight to the quality measure data.

While some participants expressed varying degrees of reluctance, others displayed variable levels of enthusiasm in using the RSQMs, feeling they were helpful in their review and noting qualities such as RSQM data being objective and providing an overall picture of performance: "I think the quality data is really helpful because I think it's just very specific to say in these types of patients this is this resident's typical practice." (Participant 7)

RSQM interpretation: Requires contextualization to self-evident. Some participants felt that properly interpreting the RSQMs required more details about

the context in which they were collected, such as having some assurances that a measure should have been met based on the clinical context and understanding a resident's reason for not meeting it:

When I'm looking through the different measures, I want more context, I guess.... Follows bronchiolitis pathways—6 times they did this, 3 times they did not.... But, I don't know why.... If that was the correct thing to do, why didn't they? Or what was their thinking? (Participant 6)

Having a peer comparison or benchmark for how often RSQMs should be met was also noted as an important piece of additional context for interpreting RSQM performance:

Did not seem like this person consistently hit all those measures. But, I'm not sure that I know how this stacks up to someone else in this comparable situation and, so, I had a hard time putting them into perspective about whether or not they did well or not well. (Participant 14)

However, while some participants thought a benchmark would help interpret RSQMs, only some of the final participants who were presented a benchmark thought it was helpful or necessary.

Not being an emergency medicine physician who understands the importance of some of the RSQMs was also noted as a contextual issue for interpreting RSQM performance. For example, a nephrologist noted, "this is not something that you encounter on the regular floor. So, this is unique for [the] emergency room.... I don't know what is expected." (Participant 16) Additionally, an endocrinologist noted, "if [diabetic ketoacidosis] were on here, I probably would have perked up." (Participant 19)

Documentation of RSQMs was another area where participants sometimes wanted more context. Some participants were inclined to give residents the benefit of the doubt; that is, they believed that even though the resident did not document something they likely did it. This stance was rationalized by the idea that knowing to do something and documenting doing it may be 2 different skills to learn and perhaps a resident needed to learn the latter rather than the former. The trend in the data was for these participants to discard the RSQMs focused on documentation, not allowing

them to influence their performance picture for the resident:

The quality measures paint a very loose picture of this resident [who] did some things by the book and did some things that weren't quite by the book. But that doesn't necessarily mean that they made bad decisions. They just didn't have documented stuff. So that didn't mean as much to me. (Participant 10)

Participants also sometimes noted they would give themselves the benefit of the doubt with regard to performing quality measures even if they had not documented them and so they would also give this same benefit of the doubt to residents.

However, other participants did not give residents the benefit of the doubt on these measures and believed that if it was not documented, then the resident had not done it. In other words, these measures are self-evident:

I would say: (a) this person either needs to work on their documentation or (b) just needs to work on their history and physical taking. Which, again being objective, it probably has to be the latter versus the former. (Participant 4)

Beyond the RSQMs focused on documentation, viewing the RSQMs as self-evident was manifested in other ways. This included viewing the RSQMs, as a whole, as objective data. One participant noted: "This is very helpful because it tells exactly what they do or do not do." (Participant 7) Sometimes this sense of objectivity was related to the perception that RSQMs arose from care guidelines, which seemed to confer additional weight to those measures: "I think that's very useful to say that you don't meet those guidelines [represented in the RSQMs]. There's nothing more objective than that." (Participant 12)

RSQMs for assessment decisions: Not sticky to sticky. Some participants used the RSQMs in the summative assessment task they were assigned in a manner similar to how they used the other assessment data they were reviewing. For these participants, the RSQMs may have been helpful for them, but they were not "sticky," meaning the data did not stand out, were not prioritized, or both when making entrustment decisions. However, for other participants, the RSQMs were sticky. The characteristic that most often made them sticky was when RSQMs

highlighted something the participant considered a deficiency. Some of these performance gaps concerned areas that participants viewed as care fundamentals that were not necessarily detrimental to patient safety:

Let me look at bronchiolitis here.... Interesting ... as I go more down this list, I see lots of things that I would not have expected.... A bulb suction order for home [not done], ... didn't do documentation, poor feeding as a reason [to return] not in the discharge instructions. I almost wonder if this is a resident that can get a little easily overwhelmed, forgetting to kind of do some of the basic stuff when they get busy. (Participant 1)

However, some of these deficiencies were viewed as more significant gaps in providing safe, appropriate patient care:

I have serious concerns about the safety of the patient because there's not prompt treatment, there's not a good review of when to return to the [emergency department], and, again, clinical assessment; it's just not there. (Participant 12)

Expectations for residents: Potentially unfair to fair to use RSQMs. Participants raised concerns about the potential fairness of using some of the quality measures to assess residents. For example, it was not clear that the level of detail included in some of the RSQMs was important to hold residents to, such as documenting discharge instructions in addition to verbally discussing them. In these instances, some participants noted that they may not include that level of detail in their own work or might not document something they know another provider is expected to document in the chart (e.g., Pediatric Respiratory Assessment Measure score in patients presenting with an asthma exacerbation, which will also be documented by the respiratory therapist). Concerns were also raised about whether residents knew these measures were expected of them. If they did not know this, participants felt it may not be appropriate to hold residents to meeting the measures:

I would never say ignore the measure. I would say: "Are we doing enough to make sure that it is easy to meet this measure?".... I would say: "Maybe nobody ever taught you that but you should be doing it anyway. The reason this is here is because this is good practice." The fact that you didn't know it means, maybe it's my fault we didn't teach you. (Participant 11)

Looking beyond the resident, some participants felt that the RSQMs included measures that should not be expected of individual residents and should instead be automated in the electronic health record. For example, documenting important parts of the history or physical exam could become components of standardized notes and some of the measures pertaining to discharge instructions could become part of standardized discharge instructions.

While concerns were raised with certain aspects of using the RSQMs for performance expectations, participants largely felt that using the RSQMs was both useful and fair. Beyond using them for assessment purposes, participants noted the utility of RSQMs for providing formative feedback to residents about areas for improvement.

Profiles for using RSQMs to make entrustment decisions

The dimensions identified in this study interact with one another. When considering these interactions, we identified 3 RSQM data user profiles: eager incorporation, willing incorporation, and disinclined incorporation. The first 2 profiles were most common, while the final profile was uncommon.

Eager incorporation. Eager incorporation was reflected when a participant's initial reactions were to use the RSQMs in their entrustment decisions. In this profile, there was an enthusiastic propensity to use RSQMs and RSQMs were seen as self-explanatory. Individuals using this profile felt that RSQMs sometimes needed additional context to interpret and sometimes that RSQMs were self-evident, but neither stance precluded them from being an eager adopter. Finally, individuals employing this profile tended to find RSQMs to be sticky and also did not raise any major concerns about the fairness of their use in assessment decisions.

Willing incorporation. Willing incorporation was demonstrated when an individual did not describe the role that RSQMs played in making their entrustment decision, but when subsequently asked specifically about the RSQMs, they described their willingness to, and even the perceived utility of,

using RSQMs. Similar to the first profile, participants employing this profile found the RSQMs to be self-explanatory and had some propensity to use them. They sometimes felt the RSQMs required additional contextualization to interpret and sometimes felt the RSQMs were self-evident. They also sometimes found the RSQMs sticky for summative assessment and other times did not. Finally, some individuals using this profile mentioned that the RSQMs reflected fair expectations for residents while others thought the RSQMs reflected unfair expectations.

Disinclined incorporation. Individuals with a disinclined incorporation profile did not indicate using the RSQMs when making a summative entrustment decision even when specifically probed about the RSQMs. However, they could envision using different RSQMs in the future. Some individuals using a disinclined profile noted the RSQMs were confusing. They were also likely to raise concerns about the fairness of holding residents accountable for meeting RSQMs. Finally, participants employing this profile were reluctant in their propensity to use RSQMs and did not find the RSQMs to be sticky. One participant tended toward the more highly disinclined end of this profile. This person's responses, while close to creating a fourth profile of "dismissive incorporation," did not quite reach this level.

Discussion

This study sought to understand how CCC members interpret, use, and prioritize RSQMs alongside traditional assessment data when making an entrustment decision. Participants often gave some types of assessment data more weight than others, such as faculty ratings and narrative comments. These findings corroborate previous research.^{22,39} However, participants tended to view and use all types of data, including RSQMs, substantively in their decision making. We found a range of engagement with RSQMs in 5 dimensions. While for each dimension, participants gravitated toward the end of the spectra that viewed RSQMs more positively, as evidenced by the eager and willing incorporation profiles being more common than the disinclined incorporation profile, examples from the more negative end of the spectra

were present for all dimensions. The defined profiles have similarities to Rogers³⁴⁰ stages of adoption, with eager incorporation aligning with early adopters, willing incorporation corresponding to early majority, and disinclined incorporation perhaps corresponding to late adopters. These findings suggest important implications for the potential future use of RSQMs, as even though this study did not focus time or effort on becoming comfortable with RSQMs, participants still tended toward using them.

While participants largely used the RSQMs in their entrustment decision making, their comments did suggest that RSQMs could be improved for such use in the future. For example, participants' discussions about whether residents should be documenting some of the information represented in several RSQMs warrants exploration. RSQMs were developed and prioritized through consensus group methods (nominal group technique and Delphi method) to meet the criteria of both being done by an individual resident and being important to care for patients with a given illness of interest.^{27,33} This raises the question of whether the consensus group results should be revisited in light of some of the issues raised by participants. The finding that some participants were concerned about the fairness of using some of the RSQMs to assess residents likely underscores the importance of orienting CCC members to and training them on RSQMs (an activity this study intentionally avoided) if the RSQMs are to be used for assessment purposes. To this end, the profiles developed in this study may help program leaders tailor their training efforts for the varying levels of inclination CCC members may have to incorporating RSQMs in their resident assessment data reviews. Such training may also address some of the challenges raised by participants who were unfamiliar with measures that would be important in the pediatric emergency department. Additionally, our participants described giving residents the same benefit of the doubt they would give themselves for the documentation measures, aligning with the findings of Apramian and colleagues,⁴¹ who showed that surgeons' individual procedural preferences may be the largest factor driving their intraoperative assessment of residents.

Some participants noted that the RSQM data felt overwhelming, suggesting that how RSQMs are visually presented to CCC members for review should be considered in the future to minimize the cognitive load experienced when reviewing them. One consideration is whether presenting performance data on all measures is important or if presenting a composite score for all measures and/or subscores for certain types of measures (e.g., documentation, appropriate medications) is sufficient.

Future directions

Some participants were reluctant to use RSQMs because they viewed the measures as too specific to the emergency department and did not want to give undue weight to that setting in their entrustment decisions. Not using these RSQMs because RSQMs for other settings do not exist actually suggests the potential utility of these measures and underscores the value of and need to develop RSQMs for other settings moving forward. As this work is carried out, it will also be important to explore differences between the views of RSQMs among our participants (members of a single pediatric residency CCC) and those of pediatric residency CCC members at other institutions. Furthermore, it will be important to ascertain differences in how CCC members from medical specialties other than pediatrics (e.g., internal medicine, family medicine) as well as surgical (e.g., general surgery, orthopedic surgery) and hospital-based specialties (e.g., radiology, anesthesiology) might use RSQMs.

The findings of this study give insight into what the foundation of CCC member training on RSQMs should include:

- A discussion of how the program would like CCC members to approach their reviews of resident performance data.⁴²
- Adequate orientation to the definition of RSQMs and how they are developed, namely through engaging supervisors and residents with experience in the settings for which the RSQMs have been developed.^{27,33}
- A high-level summary of why the measures apply in the settings for which they have been developed, especially for CCC members reviewing RSQMs from clinical areas that are outside of their expertise.

- Reminders to not give a resident the benefit of the doubt based on information the CCC members do not possess and to not make decisions based on how they would perform if they were in the resident's situation. To help prevent this, context for RSQM performance should be provided to CCC members along with the RSQM performance data.

Such training comes at a cost, most notably in time. However, regular CCC member training is a core component of high-value CCCs.⁴² The visual presentation of quality measure data to CCC members will also be important. For institutions with extensive experience gathering and presenting dashboard data for quality measures, the time and people cost to achieve this may be negligible.

While there are costs associated with implementing RSQM-informed assessment, there may also be costs associated with not implementing this innovation. Namely, patient care that is lower quality given our previous work that demonstrates the wide range of resident performance on these measures, which were determined important for care for the illnesses of interest by residents and supervisors.³⁴

Finally, while the focus of this study was on how CCC members interpret, use, and prioritize RSQMs in completing an assessment task, participants did note the perceived value of presenting RSQM data to residents for formative purposes as well. While a more detailed exploration of this was beyond the scope of this study, future efforts should engage residents to characterize how RSQMs can be used to inform and guide their personal improvement efforts.

Limitations

This study has limitations to consider. First, it was done at an institution where quality improvement is integrated into the health care system, making it possible that our participants viewed and used RSQMs differently from individuals in other settings would, potentially limiting the transferability of our findings. Second, while we designed this study to mimic authentic resident portfolio reviews by CCC members, it was conducted in a study setting, potentially

influencing participants' responses. Components of our process that did not mirror authentic reviews included using deidentified data without the resident's name, gender, or training year and providing RSQMs in the portfolio. However, the latter was integral to the intended focus of this study. Further, it is known that the presence of a resident's name and/or training year can influence assessment decisions, but we were unable to capture this based on the need for confidentiality.⁴³ Third, we did identify one participant whose RSQM data user profile bordered on dismissive. Future work should seek to determine if cases such as this are more common than it seems from our initial findings. Fourth, RSQM benchmarks were not included initially; however, through theoretical sampling, we were able to demonstrate that they were helpful and important to only some of the final participants. Fifth, we were not able to determine the role that discussion with other members of the CCC might play in how RSQMs are interpreted, used, and prioritized because we focused only on individual CCC members. This should be the focus of future RSQM study. Finally, the dimensions and profiles identified in this study may apply to other types of assessment data that CCCs review as well. Determining this was beyond the scope of this study but should be the focus of future work.

Conclusions

In this study, individual CCC members used RSQMs to varying extents in their review of resident data to make entrustment decisions and found such data helpful to varying degrees. Importantly, participants largely found RSQMs to be of some utility despite receiving no orientation to them before seeing them in this study for the first time. These findings support the inclusion of RSQMs as resident assessment data for CCC review based on our participants' receptivity to them.

Acknowledgments: The authors wish to thank the pediatric resident who allowed them to use his/her assessment data in an anonymized manner to create the resident assessment portfolio used in this study.

Funding/Support: This study was funded by the Division of Emergency Medicine at Cincinnati Children's Hospital Medical Center.

Other disclosures: None reported.

Ethical approval: This study was deemed exempt by the Cincinnati Children's Hospital Medical Center Institutional Review Board (ID: 2018-8541; exempt approval date: January 19, 2019).

D.J. Schumacher is associate professor of pediatrics, Cincinnati Children's Hospital Medical Center and University of Cincinnati College of Medicine, Cincinnati, Ohio.

A. Martini is a clinical research coordinator, Division of Emergency Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio.

B. Sobolewski is associate professor of pediatrics, Cincinnati Children's Hospital Medical Center and University of Cincinnati College of Medicine, Cincinnati, Ohio.

C. Carraccio is vice president of competency-based assessment, American Board of Pediatrics, Chapel Hill, North Carolina.

E. Holmboe is senior vice president for milestones development and evaluation, Accreditation Council for Graduate Medical Education, Chicago, Illinois.

J. Busari is associate professor of medical education, Maastricht University, Maastricht, The Netherlands.

S. Poynter is professor of pediatrics, Cincinnati Children's Hospital Medical Center and University of Cincinnati College of Medicine, Cincinnati, Ohio.

C. van der Vleuten is professor of education, Department of Educational Development and Research, Faculty of Health, Medicine, and Life Sciences, and scientific director, School of Health Professions Education, Maastricht University, Maastricht, The Netherlands.

L. Lingard is professor and scientist, Department of Medicine, and director, Center for Education Research & Innovation, Schulich School of Medicine and Dentistry at Western University, London, Ontario, Canada.

References

- Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system—Rationale and benefits. *N Engl J Med*. 2012;366:1051–1056.
- Mink RB, Schwartz A, Herman BE, et al; Steering Committee of the Subspecialty Pediatrics Investigator Network (SPIN). Validity of level of supervision scales for assessing pediatric fellows on the common pediatric subspecialty entrustable professional activities. *Acad Med*. 2018;93:283–291.
- Mink RB, Myers AL, Turner DA, Carraccio CL. Competencies, milestones, and a level of supervision scale for entrustable professional activities for scholarship. *Acad Med*. 2018;93:1668–1672.
- Promes SB, Wagner MJ. Starting a clinical competency committee. *J Grad Med Educ*. 2014;6:163–164.
- French JC, Dannefer EF, Colbert CY. A systematic approach toward building a fully operational clinical competency committee. *J Surg Educ*. 2014;71:e22–e27.
- Hauer KE, Kohlwe J, Cornett P, et al. Identifying entrustable professional activities in internal medicine training. *J Grad Med Educ*. 2013;5:54–59.
- Chahine S, Cristancho S, Padgett J, Lingard L. How do small groups make decisions?: A theoretical framework to inform the implementation and study of clinical competency committees. *Perspect Med Educ*. 2017;6:192–198.
- Hauer KE, Chesluk B, Iobst W, et al. Resident competence review in graduate medical education: A qualitative study. *Acad Med*. 2015;90:1084–1092.
- Hauer KE, Cate OT, Boscardin CK, et al. Ensuring resident competence: A narrative review of the literature on group decision making to inform the work of clinical competency committees. *J Grad Med Educ*. 2016;8:156–164.
- 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.
- Sklansky DJ, Frohna JG, Schumacher DJ. Learner-driven synthesis of assessment data: Engaging and motivating residents in their milestone-based assessments. *Med Sci Educ*. 2017;27:417–421.
- Ketteler ER, Auyang ED, Beard KE, et al. Competency champions in the clinical competency committee: A successful strategy to implement milestone evaluations and competency coaching. *J Surg Educ*. 2014;71:36–38.
- Shumway NM, Dacus JJ, Lathrop KI, Hernandez EP, Miller M, Karnad AB. Use of milestones and development of entrustable professional activities in 2 hematology/oncology training programs. *J Grad Med Educ*. 2015;7:101–104.
- Hong R. Observations: We need to stop drowning—A proposal for change in the evaluation process and the role of the clinical competency committee. *J Grad Med Educ*. 2015;7:496–497.
- Mount CA, Short PA, Mount GR, Schofield CM. An end-of-year oral examination for internal medicine residents: An assessment tool for the clinical competency committee. *J Grad Med Educ*. 2014;6:551–554.
- Donato AA, Alweis R, Wenderoth S. Design of a clinical competency committee to maximize formative feedback. *J Community Hosp Intern Med Perspect*. 2016;6(6):33533.
- Schumacher DJ, Sectish TC, Vinci RJ. Optimizing clinical competency committee work through taking advantage of overlap across milestones. *Acad Pediatr*. 2014;14:436–438.
- Johna S, Woodward B. Navigating the Next Accreditation System: A dashboard for the milestones. *Perm J*. 2015;19:61–63.
- Friedman KA, Raimo J, Spielmann K, Chaudhry S. Resident dashboards: Helping your clinical competency committee visualize trainees' key performance indicators. *Med Educ Online*. 2016;21:29838.
- Hauer KE, Chesluk B, Iobst W, et al. Reviewing residents' competence: A qualitative study of the role of clinical competency committees in performance assessment. *Acad Med*. 2015;90:1084–1092.
- Oudkerk Pool A, Govaerts MJB, Jaarsma DADC, Driessen EW. From aggregation to interpretation: How assessors judge complex data in a competency-based portfolio. *Adv Health Sci Educ Theory Pract*. 2018;23:275–287.
- Schumacher DJ, Michelson C, Poynter S, et al; APPD LEARN CCC Study Group. Thresholds and interpretations: How clinical competency committees identify pediatric residents with performance concerns. *Med Teach*. 2018;40:70–79.
- Pack R, Lingard L, Watling CJ, Chahine S, Cristancho SM. Some assembly required: Tracing the interpretative work of clinical competency committees. *Med Educ*. 2019;53:723–734.
- Chahine S, Kulasegaram KM, Wright S, et al. A call to investigate the relationship between education and health outcomes using big data. *Acad Med*. 2018;93:829–832.
- Weinstein DF, Thibault GE. Illuminating graduate medical education outcomes in order to improve them. *Acad Med*. 2018;93:975–978.
- Arora VM. Harnessing the power of big data to improve graduate medical education: Big idea or bust? *Acad Med*. 2018;93:833–834.
- Schumacher DJ, Holmboe ES, van der Vleuten C, Busari JO, Carraccio C. Developing resident-sensitive quality measures: A model from pediatric emergency medicine. *Acad Med*. 2018;93:1071–1078.
- Smirnova A, Sebok-Syer SS, Chahine S, et al. Defining and adopting clinical performance measures in graduate medical education: Where are we now and where are we going? *Acad Med*. 2019;94:671–677.
- Warm EJ, Mathis BR. Ambulatory education: Time to move from process to outcome. *J Grad Med Educ*. 2019;11:143–145.
- Jaffe R, Diemer G, Caruso J, Metzinger M. Creating provider-level quality reports for residents to improve the clinical learning environment. *J Grad Med Educ*. 2017;9:381–382.
- Kalet AL, Gillespie CC, Schwartz MD, et al. New measures to establish the evidence base for medical education: Identifying educationally sensitive patient outcomes. *Acad Med*. 2010;85:844–851.
- Lau BD, Streiff MB, Pronovost PJ, Haider AH, Efron DT, Haut ER. Attending physician performance measure scores and resident physicians' ordering practices. *JAMA Surg*. 2015;150:813–814.
- Schumacher DJ, Martini A, Holmboe E, et al. Developing resident-sensitive quality measures: Engaging stakeholders to inform next steps. *Acad Pediatr*. 2019;19:177–185.
- Schumacher DJ, Martini A, Holmboe E, et al. Initial implementation of resident-sensitive quality measures in the pediatric emergency department: A wide range of performance. *Acad Med*. 2020;95:1248–1255.
- Schumacher DJ, Holmboe E, Carraccio C, et al. Resident-sensitive quality measures in the pediatric emergency department: Exploring relationships with supervisor entrustment and patient acuity and complexity. *Acad Med*. 2020;95:1256–1264.
- Paradis E, Sutkin G. Beyond a good story: From Hawthorne effect to reactivity in health professions education research. *Med Educ*. 2017;51:31–39.
- American Board of Pediatrics. Entrustable professional activities for general pediatrics. <https://www.abp.org/entrustable-professional-activities-epas>. Accessed March 25, 2020.
- Charmaz K. *Constructing Grounded Theory*. London, UK: SAGE Publications Ltd; 2014.

- 39 Watling C, Driessen E, van der Vleuten CP, Lingard L. Learning from clinical work: The roles of learning cues and credibility judgements. *Med Educ.* 2012;46:192–200.
- 40 Rogers E. *Diffusion of Innovations*. 5th ed. New York, NY: Free Press; 1962.
- 41 Apramian T, Cristancho S, Sener A, Lingard L. How do thresholds of principle and preference influence surgeon assessments of learner performance? *Ann Surg.* 2018;268:385–390.
- 42 Kinnear B, Warm EJ, Hauer KE. Twelve tips to maximize the value of a clinical competency committee in postgraduate medical education. *Med Teach.* 2018;40:1110–1115.
- 43 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(10 suppl): S8–S11.

Appendix 1

RSQM Performance Data Included in a Resident Portfolio That Was Used in a Study of How Clinical Competency Committee Members Interpret, Use, and Prioritize RSQMs When Making Summative Entrustment Decisions, Cincinnati Children's Hospital Medical Center, February–July 2019^{a,b}

Quality measures	No. (%) of encounters in which resident met measure	No. of encounters in which resident did not meet measure	No. of encounters in which measure did not apply ^c
Asthma quality measures (total asthma encounters since last report^d = 11)			
Use asthma order set	8 (72)	3	0
Correct medication dose ordered for albuterol	8 (72)	3	0
Use of dexamethasone as steroid	11 (100)	0	0
Correct medication dose ordered for dexamethasone	11 (100)	0	0
Time from resident assigning self to patient to resident entering steroid order	4 (36)	7	0
Correct medication dose ordered for ipratropium	8 (72)	3	0
Documentation of previous intubation or BIPAP for asthma	3 (27)	8	0
Note the acuity of the patient in documentation	11 (100)	0	0
Documentation of work of breathing	11 (100)	0	0
Documentation of aeration or air exchange	11 (100)	0	0
Presence or absence of wheezing documented	11 (100)	0	0
Ensure at least 3 descriptive words used in respiratory exam documentation	11 (100)	0	0
Resident documents own PRAM score	4 (36)	7	0
Resident-assigned PRAM score matches resident-placed initial medication orders	3 (27)	1	7
Document patient response to intervention	10 (91)	1	0
Documentation of disposition decision	10 (91)	1	1
Use of standardized dosing for discharge medication (i.e., dexamethasone)	7 (64)	0	4
Home dexamethasone instructions documented in written discharge instructions	0 (0)	7	4
State who to follow up with and provide their contact information in discharge papers	5 (45)	2	4
Documentation of needing albuterol more often than every 4 hours as a reason to return in written discharge instructions	0 (0)	7	4
Documentation of worsening respiratory symptoms as a reason to return in written discharge instructions	5 (45)	2	4
Bronchiolitis quality measures (total bronchiolitis encounters since last report^d = 9)			
Follow bronchiolitis pathway	6 (67)	3	0
Document birth history (preemie or not a preemie)	5 (56)	4	0
Day of illness clearly documented	8 (89)	1	0
Documentation of previous wheezing	0 (0)	9	0
Assessment of severity documented	8 (89)	1	0
Effort of breathing documented	9 (100)	0	0
Documented quality of air entry (normal, decreased, etc.)	7 (78)	2	0
Documentation of wheezing	7 (78)	2	0
Documentation of crackles	8 (89)	1	0
Documented presence or absence of subcostal retractions	3 (33)	6	0
Documented presence or absence of intercostal retractions	1 (11)	8	0
Documented presence or absence of suprasternal retractions	0 (0)	9	0
Oxygen saturation clearly documented	0 (0)	9	0
Hydration status clearly documented	0 (0)	9	0
Documentation of patient response to specific therapeutics (i.e., how they responded to suctioning, how they responded to breathing treatment, how they responded to NS bolus, etc.)	5 (56)	4	0
Oral feeding tolerance clearly documented	5 (56)	4	0
Documentation of justification for appropriate disposition (sent home versus admitted)	8 (89)	1	0

(Appendix continues)

Appendix 1

(Continued)

Quality measures	No. (%) of encounters in which resident met measure	No. of encounters in which resident did not meet measure	No. of encounters in which measure did not apply ^c
Documentation of worsening respiratory symptoms as a reason to return in written discharge instructions	6 (67)	3	3
Use standard or premade discharge instructions for bronchiolitis	4 (44)	5	3
Documentation of poor feeding as a reason to return in written discharge instructions	0 (0)	6	3
Bulb suction teaching for home ordered	0 (0)	6	3
State who to follow up with and provide their contact information in discharge papers	5 (56)	1	3
State appropriate number of days to follow up in discharge papers	3 (33)	3	3
Closed head injury quality measures (total closed head injury encounters since last report^d = 5)			
Closed head injury or PECARN ^e best practice advisory used	5 (100)	0	0
Appropriate closed head injury or PECARN ^e pathway used	5 (100)	0	0
Mechanism of injury documented	5 (100)	0	0
Documentation of time of injury	3 (60)	2	0
Documentation of presence or absence of LOC	4 (80)	1	0
Documentation of presence or absence of emesis	5 (100)	0	0
Documentation of whether patient is back to baseline or not	4 (80)	1	0
Documentation of presence or absence of other head injury in the history	2 (40)	3	0
Assessment of severity documented	5 (100)	0	0
Documentation of presence or absence of hematoma on PE (if present, location and size also documented)	3 (60)	2	0
Thorough head exam (head, eyes, skull) documented	0 (0)	5	0
Documentation of GCS score	1 (20)	4	0
Full neurologic exam documented	0 (0)	5	0
Documentation of presence or absence of other nonhead injury on PE	1 (20)	4	0
Appropriate differential diagnosis and medical decision making documented	5 (100)	0	0
Reassessments of patient documented	2 (40)	3	0
Return to school or play recommendations in discharge papers	0 (0)	2	3
Use of standard or premade discharge instructions for diagnosis	5 (100)	0	0
Appropriate follow-up (sports medicine, rehabilitation, neurology, etc.) recommended	3 (60)	2	0

Abbreviations: RSQM, resident-sensitive quality measure; BIPAP, bilevel positive airway pressure; PRAM, Pediatric Respiratory Assessment Measure; NS, normal saline; LOC, loss of consciousness; PE, physical exam; GCS, Glasgow Coma Scale.

^aThe appendix presented here has been minimally edited from what was presented to participants. To see the version presented to participants, please see Supplemental Digital Appendix 2 at <http://links.lww.com/ACADMED/A926>.

^bThe RSQMs are a product of the authors' previous studies^{27,33} and reflect the average resident performance data on these measures from the authors' other previous studies.^{34,35} The benchmark for all measures = 80% (see Method in main text for additional information).

^cFor example, a measure may have been related to discharge, but the patient was admitted in that encounter.

^dTotal since last report indicates the number of new patient encounters for an illness since the last time the clinical competency committee reviewed the resident's portfolio.

^ePECARN is the Pediatric Emergency Care Applied Research Network, which conducted a study that delineated a decision rule for managing closed head injury in Kuppermann N, Holmes JF, Dayan PS, et al. Identification of children at very low risk of clinically-important brain injuries after head trauma: A prospective cohort study. *Lancet*. 2009;374:1160–1170.