What do quantitative ratings and qualitative comments tell us about general surgery residents’ progress toward independent practice? Evidence from a 5-year longitudinal cohort

Ara Tekian a, *, Martin Borhanib b, Sarette Tiltonc, Eric Abasoloc, Yoon Soo Park a

a Department of Medical Education, University of Illinois at Chicago College of Medicine, Chicago, IL, USA
b Department of Surgery, University of Illinois at Chicago College of Medicine, Chicago, IL, USA
c University of Illinois at Chicago College of Pharmacy, Chicago, IL, USA

Abstract

Background: This study examines the alignment of quantitative and qualitative assessment data in end-of-rotation evaluations using longitudinal cohorts of residents progressing throughout the five-year general surgery residency.

Methods: Rotation evaluation data were extracted for 171 residents who trained between July 2011 and July 2016. Data included 6069 rotation evaluations forms completed by 38 faculty members and 164 peer-residents. Qualitative comments mapped to general surgery milestones were coded for positive/negative feedback and relevance.

Results: Quantitative evaluation scores were significantly correlated with positive/negative feedback, \( r = 0.52 \) and relevance, \( r = -0.20 \), \( p < .001 \). Themes included feedback on leadership, teaching contribution, medical knowledge, work ethic, patient-care, and ability to work in a team-based setting. Faculty comments focused on technical and clinical abilities; comments from peers focused on professionalism and interpersonal relationships.

Conclusions: We found differences in themes emphasized as residents progressed. These findings underscore improving our understanding of how faculty synthesize assessment data.

© 2018 Elsevier Inc. All rights reserved.

Introduction

In graduate medical education, decisions to promote and remediate learners are often based on combining information from multiple assessments.1 In this respect, faculty use assessment data consisting of quantitative ratings and qualitative comments to provide feedback to learners and to inform the clinical competency committee (CCC) on progress toward independent practice.2-6

End-of-rotation evaluations (also known as in-training evaluation report [ITER] in Canada) require raters (faculty, fellows, or peer residents) to assign quantitative competency ratings based on predetermined anchors and write qualitative narrative comments.7 Prior studies have noted concerns over rotation evaluation scores as not being reliable assessments of learner performance, despite their prevalent use in making promotion decisions.8-13 However, recent studies have shown that rotation evaluations scores can be reliable and demonstrate validity evidence, provided that evaluations from multiple raters are collected and aggregated over a sufficient period of time.7,14-21 This study focuses on synthesizing both quantitative and qualitative data from rotation evaluations, in the context of identifying competency-based developmental levels.

Analysis of qualitative data, captured as narrative comments on rotation evaluation forms, have shown to provide useful information that allow better discrimination of learner performance. For example, studies have found that narrative comments can accompany quantitative scores; moreover, studies have also identified them as more reliable in reflecting performance than quantitative data.22,23 In addition, qualitative comments have demonstrated predictive utility in identifying residents’ future performances.24 These implications reveal potential use of qualitative data in contributing to making decisions of learner progress. Yet, the relationship between qualitative and quantitative data in rotation evaluation forms has not been examined within the context of the Next Accreditation System (NAS) by the Accreditation Council for...
Graduate Medical Education (ACGME),25,26 leaving Clinical Competency Committees and faculty responsible for making supervision decisions of learners unsure how to incorporate qualitative data and themes into their assessment decision process. The NAS promotes ongoing assessment of learners throughout their residency training and the use of multiple assessment data to inform promotion decisions.

Since July 2014, General Surgery postgraduate programs in the United States began tracking learners based on milestones and reporting them to ACGME every six months. As such, improving our understanding of qualitative narrative comments and incorporating them to inform promotion decisions and learner feedback have become increasingly important.

This study examines the alignment of quantitative and qualitative assessment data using longitudinal cohorts of residents progressing throughout the five-year general surgery residency. We studied the longitudinal progression of learners using both quantitative ratings and qualitative narrative comments. The following questions were examined:

1. Do qualitative comments align with quantitative scores for trainees from entry to graduation?
2. Do faculty provide rich qualitative comments that are useful to provide relevant and actionable feedback to trainees?

Methods

Participants

Residents. Retrospective data from 171 residents, who were part of the General Surgery residency program at the University of Illinois at Chicago (UIC) College of Medicine were used (five years: July 1, 2011–June 30, 2016).

Raters. Rotation evaluation forms were rated by General Surgery faculty (n = 38) and peer residents (n = 164, including categorical General Surgery residents and rotating residents from other specialties).

Assessment data

End-Of-Rotation Evaluation. End-of-rotation evaluation forms from the General Surgery residency program at the University of Illinois at Chicago (UIC) College of Medicine were used. The end-of-rotation evaluation form has 30 items (13 items in faculty form, 17 items in peer form). Each item was mapped to one or more of the 16 General Surgery competencies which are derived from the six ACGME Core Competencies (3 subcompetencies in Patient Care [PC], 2 subcompetencies in Medical Knowledge [MK], 2 subcompetencies in Systems-Based Practice [SBP], 3 subcompetencies in Practice-Based Learning and Improvement [PBLI], 3 subcompetencies in Professionalism [PROF], and 3 subcompetencies in Interpersonal Communication Skills [ICS]). Each item (mapped to a General Surgery subcompetency) was rated on a 9-point scale, corresponding to milestone anchors (“Level 1” = 1, “Level 2” = 3, “Level 3” = 5, “Level 4” = 7; “Level 5” = 9 points). Even-numbered points corresponded to scores between milestone levels. Data were collected using the New Innovations (NI) database and extracted for analysis.

Qualitative analysis

Two types of coding were conducted: (1) extracting themes and subthemes and (2) coding comments for positive/negative feedback and relevance of feedback. Coding was completed by two trained researchers. A sample of 50 comments were initially coded extracting themes and subthemes; discrepancies were discussed until agreement was reached. Interrater agreement kappa for the initial coding was 0.85. A third reviewer further refined and confirmed the themes and the subthemes. In the second type of coding, comments were coded on a 4-point scale: (1) positive/negative feedback, and (2) relevance of feedback (Table 1). The basis for this coding was rooted in prior work by Hatala et al.12 and Ginsburg et al.13 who previously examined the alignment between qualitative and quantitative assessment data, focusing on positive/negative feedback. We added a second dimension on relevance to further examine whether the coding was actionable and diagnostic, to add meaning to the quality of feedback provided. Interrater agreement kappa for positive/negative and relevance were 0.79 and 0.72, respectively. Disagreement was resolved through consensus.

Coding positive/negative and relevance of feedback. Qualitative comments were coded to measure their association with quantitative evaluation scores assigned by raters. We wanted to examine whether a resident with high evaluation scores also had positive comments, and conversely whether low evaluation scores corresponded with negative comments. In addition, we wanted to examine the relationship between evaluation scores and the relevance of feedback provided in the qualitative comments. For example, a comment such as “excellent job” does not provide meaningful or actionable information; on the other hand, a comment such as “[resident] developed a positive rapport with his patients which was evidenced on my rounds” provides more relevant information about the learner. In the first dimension (positive/negative feedback), coding for 1 to 4 anchors corresponded to “Highly Critical”, “Critical”, “Modest Praise”, and “High Praise”, respectively. In the second dimension (relevance of feedback), coding for 1 to 4 anchors corresponded to “Highly Irrelevant”, “Irrelevant”, “Relevant” and “Highly Relevant” (see Table 1).

Themes. Comments were also analyzed for themes following a descriptive study using content analysis and generating categories.27 Two reviewers extracted themes from the comments, by identifying themes and sub-themes, stratifying them by PGY level. As member check, results were compared and re-categorized into different themes and sub-themes, until agreement was reached by the two reviewers and confirmed by a surgeon-educator. Eight themes were identified and remained consistent throughout the five years (Table 2). An additional 46 sub-themes were also identified. Saturation was reached after analyzing 400 comments. The two reviewers were able to independently identify each unique idea represented.

Quantitative Data Analysis. Rotation evaluation scores were examined using descriptive statistics. Unit of analysis was the learner by year. Mean scores were created for the six core competencies, by taking the average across subcompetencies. Evaluation scores were correlated with coded values of positive/negative feedback and relevance of feedback from qualitative comments, to examine the alignment between qualitative and quantitative scores in rotation evaluation forms. Generalizability theory was used to examine variance components in rotation evaluation scores by PGY level and to estimate reliability (Table 3). A [rater (r): person (p)] x [subcompetency (s): competency (c)] design was used, following variance components structure from previous analysis.11,22

Data compilation and analyses were conducted using Stata 14 (Stata Corp, College Station, TX). The UIC institutional review board approved this study.

Results

Descriptive statistics

Evaluation records and qualitative comments. A total of 6069
Table 1

Categorization of narrative comments: Positive/negative and relevance of feedback.

<table>
<thead>
<tr>
<th>Dimension 1: Positive/Negative Feedback</th>
<th>Dimension 2: Relevance of Feedback</th>
<th>Scale</th>
<th>Sample Comment</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Relevant</td>
<td>4</td>
<td>Very strong overall performance. Resident is hard working and anxious to improve his clinical performance. He has developed a great rapport with the patients. He does an excellent job of prioritizing his work and taking the initiative to learn new skills. Typical of an intern, I am confident that he will continue to improve.</td>
<td>Resident</td>
</tr>
<tr>
<td>Positive</td>
<td>Irrelevant</td>
<td>3</td>
<td>Resident is smart and hard working. He seemed a little overwhelmed on the trauma service early in this rotation, but seemed more comfortable as time went by and his confidence level increased.</td>
<td>Resident</td>
</tr>
<tr>
<td>Positive</td>
<td>Relevant</td>
<td>2</td>
<td>Needs to become more engaged in the care of patients. Very efficient resident, but seems to lack some self-awareness concerning his approach.</td>
<td>Resident</td>
</tr>
<tr>
<td>Positive</td>
<td>Highly Relevant</td>
<td>1</td>
<td>Highly critical comments on residents’ clinical performance. He is up to the same level as his peers. He has difficulty with case sentiment and I am recommending that he be monitored closely during his next year.</td>
<td>Resident</td>
</tr>
</tbody>
</table>

Note: Coding for all narrative comments were conducted by two reviewers. Discrepancies were resolved through discussion until consensus was achieved for all comments.

Quantitative scoring of qualitative comments. Majority of narrative comments were very positive (71% from faculty and 84% from peer evaluations). However, for relevance, 56% of faculty and 69% of peer evaluations had relevant or very relevant comments. Between training years, there were no differences in quality of comments from peers, $p = .487$. However for faculty qualitative comments, there were significant differences in the quality of comments by training year, $p < .001$. Across training years, there were 28%, 31%, 41%, 58%, and 33% of comments coded as highly relevant from PGY1 to PGY5 (see Table 4). Between positive/negative and relevance coding, there was negative association, $p < .001$.

Alignment between rotation evaluation scores and qualitative comments. The reliability of evaluation ratings was good, $\Phi$-coefficient = 0.72. Quantitative evaluation scores were significantly correlated with positive/negative feedback indicating alignment between quantitative and qualitative feedback, $r = 0.52$. However, when residents received higher quantitative ratings, the relevance of comment was significantly lower, $r = -0.20$, $p < .001$. These findings indicate alignment between quantitative and qualitative comments. However, when qualitative comments were more negative, there was greater quality of diagnostic and actionable feedback for residents.

Qualitative results

Themes extracted were divided into two categories: (1) ACGME-related and (2) non-ACGME related themes. Overall, there were 523 unique faculty evaluations with comments (70%), resulting in 1126 comments extracted for analysis (allowing for duplicates). ACGME-related themes included PC, MK, PROF, ICS, and SBP, with highest number of comments from PC and MK. Non-ACGME related themes were personal attributes and traits, summative judgements, and comparison to level of training. Table 2 summarizes the comments provided by faculty (see Table S1 in Supplementary Material for resident comments). The number of subthemes for the ACGME-related themes varied from one for SBP, to eight for professionalism in both tables. In addition, the number of subthemes were different at each PGY level. For example in Table 2, for professionalism, there was only one subtheme ("Team Player") during PGY3. Under the ACGME-related themes, the highest number of comments were for PC (178) and the lowest was for SBP (4). For non-ACGME related themes, the highest number of comments were for personal attributes and traits (484), and the lowest for comparison of training (64). The total number of comments per PGY level from year one to five were 415, 247, 95, 249, and 120 respectively, indicating a significant drop in the number of comments during PGY3 and an abrupt increase during PGY4. The percentage of evaluations without comments from PGY1 to PGY5 were 32%, 26%, 32%, 25%, and 30% respectively. A sample of quotes by the faculty are presented below highlighting positive (praise) and negative (concerns) areas.

Faculty comments

**Operative Skills.** Faculty frequently commented on residents’
Table 2

<table>
<thead>
<tr>
<th>Theme</th>
<th>PGY1</th>
<th>PGY2</th>
<th>PGY3</th>
<th>PGY4</th>
<th>PGY5</th>
<th>All years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sub-theme C</td>
<td>Sub-theme C</td>
<td>Sub-theme C</td>
<td>Sub-theme C</td>
<td>Sub-theme C</td>
<td>C</td>
</tr>
<tr>
<td>ACGME-related Themes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient Care</td>
<td>1 Clinical Judgment</td>
<td>1 Clinical Judgment</td>
<td>1 Clinical Judgment</td>
<td>1 Clinical Judgment</td>
<td>1 Clinical Judgment</td>
<td></td>
</tr>
<tr>
<td>2 Complex Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Decision-making</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Managing Patient Plans</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Operative Skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Patient Care (General)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical Knowledge Professionalism</td>
<td>1 Clinical Knowledge</td>
<td>1 Clinical Knowledge</td>
<td>1 Clinical Knowledge</td>
<td>1 Clinical Knowledge</td>
<td>1 Clinical Knowledge</td>
<td></td>
</tr>
<tr>
<td>2 Fundamental Knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Level of Professionalism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Punctuality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Respectfulness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Self-awareness for Improvements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Team Player</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Trustworthiness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpersonal/Communication Skills</td>
<td>1 Clinical Judgment</td>
<td>1 Communication Skills</td>
<td>1 Communication Skills</td>
<td>1 Communication Skills</td>
<td>1 Communication Skills</td>
<td></td>
</tr>
<tr>
<td>2 Decision-making</td>
<td></td>
<td>(General)</td>
<td></td>
<td>(General)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Managing Patient Plans</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Patient Care (General)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systems-Based Practice</td>
<td>1 Difficulty with New/Foreign System</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Non-ACGME-related Themes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal Attributes and Traits</td>
<td>1 Ability to Work Independently</td>
<td></td>
<td>1 Ability to Work Independently</td>
<td>1 Ability to Work Independently</td>
<td>1 Ability to incorporate feedback/able to improve upon critique</td>
<td>1 Efficiency</td>
</tr>
<tr>
<td>2 Ability to Work Independently</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Level of Enthusiasm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Level of Self-confidence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Maturity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Motivation to Learn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Personality, Attitude or Demeanor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Sense of Responsibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 Thorough/Attention to Detail</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Willingness to Initiate Action</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued on next page)
operative skills across all five years. As effective operative skills are essential in a surgical resident, faculty members often focused on the residents’ deficiencies.

“... Unfortunately, [he] stands also at the bottom of his class in terms of surgical skills and manual dexterity as assessed by multiple attendings in both operative theater and at the surgical skills lab. He is trying very hard to overcome these deficiencies and we should give credit to him for this genuine and extraordinary effort he is putting in place to perform inside the operating room at the level of his peers and classmates.” (Faculty assessment of PGY2 resident)

Although faculty indicated shortcomings in the residents’ operative skills, they expressed encouragement by giving residents credit for striving to improve, hoping they do well in the future, or stating the resident will continue to develop. This form of constructive criticism is prevalent when commenting on other themes as well.

**Level of Self-Confidence.** Level of self-confidence emerged across all years. Faculty commonly related the residents’ level of confidence to their overall performance.

“... He seemed easily overwhelmed by the work load and was absent on occasion without explanation. He does not seem to function well under stress. He tended to be inefficient in performing the typical work responsibilities of a surgical intern. He did seek assistance appropriately. We cannot gauge his surgical skills since he rarely if ever came to the OR.” (Faculty evaluation of PGY1 resident)

In this case, the faculty member noted that his level of self-confidence negatively impacted his work flow and time spent practicing his operative skills. Other narrative comments also mentioned residents’ level of confidence affecting their efficiency by impeding multitasking.

“[Name] is a good man and tries hard to perform at the level of his Urology colleagues but, unfortunately, has some limitations. He gives the impression of being constantly overwhelmed and has serious difficulty to multi-tasking...” (Faculty evaluation of PGY3 resident)

While some narrative comments related their level of self-confidence to other deficiencies, other comments found the residents were generally proficient apart from their self-confidence. The following two faculty members share similar observations regarding the same resident.

“[Name] is a very serious and hardworking resident. He has sound judgement and demonstrates the ability to apply newly learned principles. He needs to build self-confidence to match his level of performance.” (Faculty evaluation of PGY4 resident)

“I basically like [Name] a lot. He is quiet and effective. Does his work very well without fanfare. But the question is whether he really gets all the credit he deserves. He will have to become a leader as a Chief Resident and you cannot do that by being exceedingly quiet.” (Faculty evaluation of PGY4 resident)

In this comment, the faculty member noted that the resident’s lack of self-confidence would cost him chief resident despite his other positive qualities.

**Observation of Growth.** As faculty follow residents throughout training, they note the residents who have shown improvement...
Note: There were significant growth in scores for PC, MK, PBLI, PROF, and SBP, p < .05.

Table 3
Quantitative ratings by competency and training year: Mean (SD).

<table>
<thead>
<tr>
<th>Competency</th>
<th>PGY-1</th>
<th>PGY-2</th>
<th>PGY-3</th>
<th>PGY-4</th>
<th>PGY-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC</td>
<td>7.65 (1.63)</td>
<td>7.95 (1.56)</td>
<td>7.88 (1.77)</td>
<td>7.73 (1.83)</td>
<td>8.25 (1.66)</td>
</tr>
<tr>
<td>MK</td>
<td>7.21 (1.57)</td>
<td>7.52 (1.45)</td>
<td>7.57 (1.78)</td>
<td>7.27 (1.67)</td>
<td>8.10 (1.64)</td>
</tr>
<tr>
<td>PBLI</td>
<td>7.41 (1.60)</td>
<td>7.47 (1.55)</td>
<td>7.57 (1.73)</td>
<td>7.47 (1.69)</td>
<td>8.00 (1.55)</td>
</tr>
<tr>
<td>ICS</td>
<td>7.99 (1.72)</td>
<td>7.97 (1.68)</td>
<td>8.19 (1.48)</td>
<td>8.02 (1.95)</td>
<td>8.10 (1.67)</td>
</tr>
<tr>
<td>PROF</td>
<td>7.84 (1.59)</td>
<td>7.90 (1.49)</td>
<td>8.17 (1.62)</td>
<td>7.84 (1.81)</td>
<td>8.36 (1.67)</td>
</tr>
<tr>
<td>SBP</td>
<td>7.59 (1.50)</td>
<td>7.75 (1.38)</td>
<td>8.09 (1.50)</td>
<td>7.77 (1.56)</td>
<td>8.30 (1.56)</td>
</tr>
</tbody>
</table>

Note: There were significant differences in the proportion of highly relevant comments for faculty, but not for peer-residents.

and those who strive to improve. Below is an example of two faculty members describing the same resident.

“[Name] makes great efforts to recognize and address his own deficiencies and actively pursues self-improvement. He is respectful, considerate and has outstanding interpersonal skills…” (Faculty evaluation of PGY2 resident)

“[Name] continues to show substantial growth and improvement. He was one of the most engaged senior residents on the service this year…” (Faculty evaluation of PGY2 resident)

At times, observations of growth are most highlighted in the final PGY5 year.

“Since the last rotation [Name] has really progressed well and is showing a steep learning curve. She is responsible and a hard worker. Her patient care has continued to improve with more experience. She is reliable and is a terrific member of the team. She is progressing well and I look forward to seeing her continue to improve and shine among her peers. She is outspoken but not to a fault and her eagerness is apparent. Keep up the good work!” (Faculty evaluation of PGY5 resident)

When faculty measured their improvement, they relate it to not only technical skills, but to their interpersonal and communication skills and their personal and professional development.

Maturity. Maturity often emerged in the faculty’s narrative comments. Often statements touching on maturity also touched on comparison level of training.

“[Name] did a great job on his trauma rotation. He is a mature and confident senior resident. He understands his role as a service chief fully. He handled a very difficult situation with an insubordinate junior EM resident with tact and maturity. Clearly, he is one of the best residents in his class.” (Faculty evaluation of PGY1 resident)

Residents’ comments

The total number of narrative comments by residents (4,127) were almost fourfold, compared to the number of faculty comments (1,126). Under the ACCME-related themes, the highest number of comments were for ICS (611) and professionalism (563) in contrast to comments by the faculty, where the highest were for PC (178) and MK (139). For the non-ACGME related themes, the highest number of comments were for personal attributes and traits (1,898), which comprised 46% of the entire comments. Furthermore, the highest number of comments were provided during PGY1 level (39%, 1,628). The total number of comments per PGY level from year one to five were 1628, 811, 559, 619, and 510 respectively, indicating greater number of comments during PGY1 and PGY2. A sample of quotes by residents are presented below highlighting positive (praise) and negative (concerns) areas.

Teaching skills

When junior residents evaluate senior residents on teaching skills, they were able to further elaborate on this topic because they had personal experiences given the nature of the mentor/mentee relationship.

“...He has outstanding empathy and perception, able to realize fine differences between errors that occur from unrealistic expectations versus errors that occur from inexperience. In realizing this, [Name] is able to teach with tactful and judicious consideration, making it an excellent experience to work with and learn from him.” (PGY1 resident assessment of a PGY3 resident)

Team Player. Similar to teaching skills, observations of how one works in a team draws from interactions between collaborating residents.

“Not a team player, very rude to other residents and nurses.” (PGY1 resident assessment of a PGY2 resident)

“...A good team player who ensures all team members are informed and present for learning opportunities. Always willing to help.” (PGY1 resident assessment of a PGY2 resident)

Quotes highlighted ideal and non-ideal traits when working in a team.

Personality, Attitude or Demeanor. Themes around personality, attitude or demeanor emerged from daily interactions between residents.

“Poor quality resident. Rude, selfish, disrespectful and has peculiar personality and attitude. She uses offensive wording all the time when dealing with other residents and when talking about patients. This resident is not helpful to other residents at all, always defer work to other residents to complete … She was described as not helpful by almost all residents working with her and they don’t like working with her given the hostile environment she’s creating…” (PGY2 resident assessment of a PGY4 resident)

Narrative comments showed both favorable and unfavorable characteristics regarding personality, attitude or demeanor. Overall, comments provided by residents compared to faculty were more personal, speaking to nuances of actually working closely with residents.
and combine information conveyed in qualitative and quantitative data that can inform resident progress. Our results highlight alignment in qualitative and quantitative data; however, there was variation in the quality of comments provided by faculty and peer residents, covering different trends throughout their residency training. These findings can be used to inform better synthesis of assessment data for feedback, monitoring of resident progress, and ongoing faculty development, as residents train toward unsupervised practice.

Conflicts of interest

The authors do not have any conflict of interest.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Summary

This study examines the synthesis of qualitative and quantitative data from rotation evaluations collected from longitudinal cohorts of General Surgery residents. Results showed alignment in qualitative and quantitative data, with significant variation in themes and quality of comments as residents progressed toward unsupervised practice.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.amjsurg.2018.09.031.

References


