8-13-2020

Best Approaches to Evaluation and Feedback in Post-Graduate Medical Education

Sara Q. Perkins

Ali A. Dabaja

Humphrey Atiemo

Current Urology Reports (2020) 21:36
https://doi.org/10.1007/s11934-020-00991-2

EDUCATION (G BADALATO, SECTION EDITOR)

Best Approaches to Evaluation and Feedback in Post-Graduate Medical Education

Sara Q. Perkins[1] · Ali Dabaja[1] · Humphrey Atiemo[1]

Abstract

Purpose of Review The objectives of this literature review are to appraise current approaches and assess new technologies that have been utilized for evaluation and feedback of residents, with focus on surgical trainees.

Recent Findings In 1999, the Accreditation Council for Graduate Medical Education introduced the Milestone system as a tool for summative evaluation. The organization allows individual program autonomy on how evaluation and feedback are performed. In the past, questionnaire evaluations and informal verbal feedback were employed. However, with the advent of technology, they have taken a different shape in the form of crowdsourcing, mobile platforms, and simulation. Limited data is available on new methods but studies show promise citing low cost and positive impact on resident education.

Summary No one "best approach" exists for evaluation and feedback. However, it is apparent that a multimodal approach that is based on the ACGME Milestones can be effective and aid in guiding programs.

Keywords Formative evaluation · Summative evaluation · Feedback · Medical education · Technology · Best practices

Introduction

Assessing residents' competency at the national board certification level relies heavily on oral and written examinations; however, this process does not assess the ability to perform procedures, technical skill, or interpersonal communication style in critical situations [1]. Therefore, evaluation of performance as a clinician and technician in these areas are the responsibility of residency programs. Evaluation describes the determination of the value or quality of something or someone [2]. In graduate medical education, it often refers to the gathering of information about what the learner should know, be able to do, or what they should work towards, and is associated with measurements [3]. Two types of evaluation exist, formative and summative. Formative evaluation is used to identify areas for improvement, refine goals, and evolve strategies for achieving them [3, 4]. Summative evaluation is often more formal, it is used to assess whether the results of the object being evaluated met the stated goal, and it tends to be more numeric and quantified [3, 4].

To build on evaluation, feedback is often utilized. Feedback is the transmission of an evaluation or corrective information about an action, event, or process to the original performing source [5]. In order to have appropriate feedback, it must be based on an accurate evaluation [3, 6]. Evaluation and feedback are intimately related although they are separate entities. Both are important factors in medical education as they allow trainees to take an assessment of their performance and learn from previous mistakes, giving them opportunity for improvement of their clinical skills in future practice [7, 8]. The Accreditation Council for Graduate Medical Education (ACGME) has put great emphasis on the importance of evaluation and feedback by annually surveying programs and inquiring the depth of evaluation and feedback the residents feel they receive from their programs [7, 8].

The advent of mobile devices, live streaming, and high-tech simulation models have allowed residencies, especially surgical, to creatively evaluate things such as surgical technique and communication skills. Different technologies have also been researched to make administering feedback more convenient to faculty and more accessible to residents. Currently, there is a lack of consensus regarding the best methods of evaluation and feedback with no formal

This article is part of the Topical Collection on *Education*

✉ Humphrey Atiemo
hatiemo1@hfhs.org

[1] Henry Ford Health System, 2799 W Grand Blvd, K9, Detroit, MI 48202, USA

Published online: 13 August 2020

endorsement by educational bodies on any particular methods. In this article, we review the current approaches as well as new technology that have been used for evaluation and feedback of residents, with focus on surgical trainees.

ACGME Milestones

The ACGME is a not-for-profit organization dedicated to accrediting and setting standards for US residency and fellowship programs [9]. The organization developed the ACGME Milestones in 1999 and they were introduced to the first seven specialties in 2013 as a tool for evaluation. The Milestones system provides a framework for assessing resident development in 6 core competencies, which include the following: medical knowledge; patient care; professionalism; interpersonal and communication skills; practice-based learning and improvement; and systems-based practice [8]. The core competencies are common across all specialty groups; however, they are further broken down into sub-competencies that have been tailored to address specialty-specific skills.

Milestones are arranged from level 1 to level 5; they do not correlate to post-graduation level, but represent the progression from novice to expert in each category. There is no predetermined timing that residents must progress, as the milestones are dynamic. Although reaching level 4 is not a graduation requirement, it is suggested that residents score at level 4 in the majority of categories at the time of graduation. The Milestone system is used to demonstrate accountability of the effectiveness of graduate medical education within accredited programs. The Milestones have been used as a summative evaluation tool of residents, as it assesses if residents have obtained a certain level of skill in each subcategory [8].

The Milestones can also be fashioned into questionnaire format; the organization provides diagrams of how evaluation of Milestones can be structured. Each Milestone level has a description of what skills would qualify for that individual level; in addition, there are responses boxes on the line in between levels that indicate performance in between levels [10••, 11]. Institutions can base their evaluation structures on the examples provided by the ACGME. Studies have shown that rotation evaluation scores using the Milestones demonstrate validity and reliability; however, evaluations must be aggregated over time with multiple raters [12, 13, 14•, 15, 16].

The Clinical Competency Committee (CCC) of each program is required to review all residents' progress on the core competencies. CCCs are a required part of earning accreditation. Each CCC should be composed of at least three, actively teaching faculty members, and the faculty should have clear descriptions of their roles. The CCC is responsible for synthesizing resident evaluations, case volumes, and other measures to assess resident progression on the Milestones. The CCC can then use their assessment as a basis to provide formal feedback to residents on their progression. However, structuring of when meetings occur, how they synthesize performance information, and how feedback is provided to residents is decided by each individual committee [17••].

The ACGME suggests that educational leaders not only evaluate residents on the Milestones for their given specialty but also create an assessment program. The assessment program must be built so raters have a clear understanding of how to use and complete the assessment tools. They note that there is not a single evaluation tool or method that can encompass all competencies, hence the interest in elucidating best practices. The importance of feedback is also emphasized by the ACGME, as research has shown that feedback is one of the most effective educational tools for improvement [8].

They suggest that feedback be given with the five basic features of high-quality feedback: timeliness, specificity, balance reinforcing (positive) and corrective (negative) feedback, learner reaction and reflection, and action plans [8, 18]. The ACGME also gives two models that may be helpful for feedback: ADAPT (Ask-Discuss-Ask-Plan Together) model and the R2C2 (Rapport building, explore Reaction, Explore Content, Coach for performance change) [8, 19]. Since no standard exists, research on the best approaches to evaluation and feedback of surgical residents has become an increasingly important in the medical community, particularly in consideration of the ACGME Milestones and the introduction of new technologies.

Best Practices and Innovative Ideas for Evaluation and Feedback

Utility of Milestone Evaluation

The implementation and utility of ACGME Milestones have been assessed by a number of research studies. One study looking at the implementation of the Milestones in a general surgery program analyzed evaluations of 44 residents. They found significant increases of Milestone scoring between most post-graduate year (PGY) levels ($p < 0.05$), the interrater reliability, meaning the agreement among raters, for the total score and 6 competency domains was very high (ICC: 0.87–0.98 and α: 0.84–0.97). The general consensus of the cohort was that the Milestones and CCC meetings increased opportunities for residents to reflect on their performance and increase faculty participation in the educational process [20]. Another study found a trend in increased score in the Milestones by the level of PGY training. They compared the traditional evaluation system (generic 16-item survey using a 5-point Likert scale ranging from 1 to 5, and a free-text comments section) to the Milestone system. Using the Milestone system, they found the median score was 2.69 (range: 1.5–3.7, $p < 0.01$); the median score differed across PGYs, and

increased by PGY level of training ($p < 0.01$). The study found no meaningful difference in median score by PGY year in the traditional evaluation system [21].

In contrast, a study by Sebesta et al. surveyed 133 urology program directors on their perception of the Milestone system. They had 88 responses (66% response rate) with 48% of the cohort answered that Milestones were very or somewhat unhelpful. The study also found that 30% of respondents felt neutral about the use of Milestones based evaluations leading to better resident feedback. 49% of respondents felt Milestones were not predictive of passing board rates, and 58% of respondents felt that the Milestones were completely or somewhat uncorrelated with in-service examination scores [22•]. A separate study conducted over 4 years examining the association between Milestone evaluations and American Board of Surgery In-Training Examination Scores (ABSITE) found that neither annual nor subset evaluation scores were significantly associated with passing the ABSITE ($n = 102$; for annual evaluation, odds ratio = 0.949; 95% CI, 0.884–1.019; $p = 0.15$). There was no difference in mean evaluation score between those who passed versus failed the ABSITE (mean [SD] evaluation score, 91.77 [5.10] versus 93.04 [4.80], respectively; $p = 0.14$) [23].

Of note, the ACGME is currently developing Milestones 2.0 as concerns were addressed over the existing model. The structure of the Milestones are very similar; they will be broken up into competencies and sub-competencies that are being developed by specialty-specific development groups [24•]. However, the language for each developmental level description will be less complex to ensure that they are clear and easy to understand. Currently, specialties are in various stages of development and all are expected to have initiated the revision process by the end of 2020. The revision of the Milestone system will provide critical improvement and long-term research will be needed to assess the revision [24•]. General Surgery Milestones will take effect in 2020 and will reflect the Milestone 2.0 changes, while the most recent Urology Milestones were devised in 2016 and are currently undergoing revision [10••, 11].

Questionnaire Based, a Chance for Summative and Formative Feedback

Evaluation of surgical residents is primarily performed utilizing questionnaires completed by faculty. Evaluations are based on a combination of resident knowledge and observation of procedural conduct [25]. For technical skills, tools such as Ottawa Surgical Competency OR Evaluation (O-SCORE), Objective Structured Assessment of Technical Skills (OSATS), the Resident Report Card (RRC), and single-item versus multiple-item operative performance ratings (OPRs) have been investigated [1, 26, 27]. Most utilize survey tools, such as Likert scales, to rate residents' performances in cases, allowing surgical residents to identify areas for improvement. Additionally, other tools have been conceptualized to assess clinical competence and communication skills, such as The Ottawa Clinic Assessment Tool and the Communication Assessment Tool, respectively [28, 29]. All of these methods represent summative evaluation tools in terms of the numerical- or rating-based evaluations. However, many of these platforms allow for evaluators to use comments to further elaborate what learners need more improvement on or what steps they performed well as a form of feedback.

Simulation—an Opportunity for Summative and Formative Evaluation

Interpersonal communication skills and professionalism are core competencies evaluated by the Milestone system in each specialty. Traditionally, this was accomplished by faculty observation of residents' interactions with patients. However, recently, standardized patients (SPs) in Mini-Clinical Evaluation Exercises (mini-CEX) or objective structured clinical examinations (OSCE) have been utilized [30–32]. SPs in simulation-based training (SBT) have been used to aid in the evaluation of surgical residents in situations such as delivering bad news, performing history and physical exam, and dealing with cultural differences [30, 32–35].

A study by Gee et al. used SP simulation to evaluate resident management and approach to right upper quadrant pain. They compared milestone evaluations from SPs, faculty, and the CCC and found they did not correlate. Mean Milestone scores from SPs were significantly higher than from faculty or the CCC. They proposed that SP evaluations on Milestones would not suffice alone [33]. However, they found that SP scenarios are ideal to assess strengths and weaknesses and to provide individualized feedback [33]. Many SP curricula have been devised to incorporate formative and summative evaluation from SPs and faculty with the use of questionnaires, as well as comment boxes. Additionally, curricula include the opportunity for formal feedback and debriefing after SP exercises. Evidence has shown improvement of surgical resident professionalism and interpersonal and communication skills over time with institution of SP based curricula [34, 36, 37]. Additionally, studies have shown that learners report satisfaction with the exercises conducted and the feedback gleaned from them [31, 33].

New technology has presented an opportunity to build on SBT evaluation and feedback. Examples of advanced simulation trainers include the SimMan3G®, Cook® Medical Ureteroscopy model, and dynamic haptic robot trainers [38, 39•, 40]. Utilization of simulation trainers aimed at the development of skills for specific procedures has become increasingly popular. Use of the new technology allows trainees to practice and fail without harm and the opportunity to learn from mistakes [38, 41]. SBT using validated Cook®

Medical URS model showed a significant difference between pre and post-task completion times (15.2 min versus 9.1 min, $p < 0.001$) [39•]. Another study of urology residents demonstrated increased familiarity with indications (mean difference = + 0.59) and preparation for fusion biopsy (mean difference = + 1.16), as well as the method of MRI to TRUS image registration (mean difference = + 1.38) with SBT [40] Formative evaluation during SBT takes place as faculty can directly observe and give active evaluation while the resident is performing these tasks; additionally, it allows for timely feedback.

Furthermore, SBT has been used to evaluate learners' communication skills on interdisciplinary teams and their function in high-risk situations [38, 41]. A systematic review and meta-analysis revealed that in comparison with no intervention, simulation was associated with large effects for outcomes of skills, behaviors, and knowledge [42]. For surgical residents who often perform various procedures or communicate in high-risk situations, it is evident that SBT offers many benefits. It can employ both summative and formative evaluations, as well as provide timely feedback to residents as many of these SBT utilize debriefing after activities from evaluators as well as peers. However, SBT availability to surgical residents is limited by cost, equipment, and time constraints [43].

Coaching-Based Formative Evaluation and Feedback

The surgical technique is traditionally gained in the operating room (OR); however, with modern changes in resident education, such as duty hour, concerns about decreased operative time have arisen [44–46]. Coaching in surgery has been investigated as a tool to give learners more in-depth feedback and advance technical skills [47••, 48]. One model developed by Hu et al. filmed surgeons completing operations, then held 1-h coaching sessions focused on operative techniques; this is an example of formative evaluation; the feedback was also offered during coaching sessions. Video review was received well by participants, especially when identifying failures to progress during surgery or troubleshooting alternate approaches [49].

A follow-up study was performed by the same group and compared teaching points made in the OR versus those made in review coaching sessions with the use of independent coders. Higher-level teaching concepts, such as intraoperative decision-making (mean, 9.70 versus 2.77 instances per hour, $p = 0.03$) and failure to progress (mean, 1.20 versus 0.13 instances per hour, $p = 0.04$), both forms of formative evaluation, were more frequently observed in the coaching sessions. In regard to the quality of feedback, it was found that surgeons asked almost 3 times more questions to prompt reflection or critical thinking (9.30 versus 3.32, $p = 0.07$) and set 10 times more goals to facilitate future practice (2.90 versus 0.28, $p = 0.11$) [47••]. Although studies looking at this technique are limited, they have shown that video-based coaching is particularly useful for individualizing formative evaluation, offering constructive feedback, improving technical skills, and teaching higher-level concepts, such as decision-making [47••, 48, 50, 51].

Smartphone-Based

In providing feedback to the twenty-first-century trainee, smartphones and portable devices have started to become one platform investigated as a tool to improve efficiency and efficacy of competency documentation. The System for Improving and Measuring Procedural Learning (SIMPL) is a smartphone-based tool that is capable of making real-time intraoperative performance assessment. SIMPL includes a 3-question performance assessment for both trainees and attending's following any procedure that they perform together; additionally, attending's can provide specific dictated feedback. SIMPL was employed to evaluate residents in 13 general surgery residencies between 2015 and 2016. The study found that the overall response rate was 70%, the dictation rate was 24%, and the mean response time was 12 h. Assessments increased from 357 (September 2015) to 1146 (February 2016) [52]. A study by Williams et al. suggests that at least 2.3 assessments per month are needed to provide a stable estimate of a resident's operative performance ability. 100% (20/20) of the most active residents exceeded the 2.3 operative performance assessments per month benchmark [52, 53].

A follow-up study looking at value and barriers to using the SIMPL tool, performed in 2019, found that frequent users perceived a benefit for both numerical evaluations (76% versus 30%) and dictated feedback (92% versus 30%) compared with low users. Faculty engagement was a major barrier to adaptation, as 30% of faculty sampled noted they were "never users." Suggesting that a mechanism to increase faculty participation, such as automatic notifications, could be used to enhance timeliness of evaluation submission [54].

SIMPL is just one of many web and smartphone-based tools that have been studied. The Resident Report Card (RRC) is a similar smartphone-based platform utilized to garner feedback on urology residents' surgical skills, such as safe conduct, efficient conduct, tissue handling. RRC allowed residents' surgical skill progress to be charted over time with 100% of residents noting that the tool is useful [55]. Other mobile applications studied include Quicknotes, a nameless smartphone-based platform that was instituted at Carolina's Health System, and the Southern Illinois University Operative Performance Rating Scale (OPRS) [56•, 57, 58]. All applications noted similar benefits of smartphone-based evaluation, such as accessibility, low cost, and ability to trend resident progression [52, 54, 55, 56•, 57, 58]. Of note, with the use of these applications, faculty evaluation of residents resulted in residents reporting a positive impact on the feedback

they received [52, 54, 55, 56•, 57, 58]. While evaluation and feedback are separate entities, many studies results on this topic suggest that sharing evaluations with residents is viewed as a form of feedback. However, many projects had similar limitations, such as small cohort size or issues with response rates [56•, 57, 58].

Crowdsourcing—a Tool for Formative Evaluation and Immediate Feedback

Crowdsourcing is an approach to problem solving where a task is completed with input by a large number of decentralized individuals [59]. Studies have demonstrated that a crucial component to developing mastery of surgical technique is immediate and frequent formative evaluation and feedback [18, 60]. With the advent of the Internet, crowdsourcing can provide an efficient source of diverse views on problem solving, offering formative evaluation. A systematic review on the use of crowdsourcing in the setting of evaluating technical skills was performed by Dai et al. [61••]. In the 13 articles reviewed, they found that in almost all studies, feedback was received quicker from crowds (2.83 h–5 days) than it was from experts (26 h–60 days). Crowds were also found to complete summative evaluations quicker than experts (4.8–150.9 times fast) [61••].

Dai et al. found there was considerable concordance between evaluation scores from crowds and experts for almost all tasks examined [61••]. However, poorer agreement was found in studies that involved novice-level technicians [61••, 62, 63]. This may suggest that feedback from crowdsourcing may be more accurate when applied to more skilled surgeons [61••]. A comparison of pooled comments from crowds and experts showcased discussions of similar themes regarding techniques such as tissue handling and efficiency, offering the technician formative evaluation. However, more research is needed to identify how helpful these comments are to the technician [61••]. Cost is also an important consideration; one study found that the cost for 30 evaluations from crowdworkers was around $16.50, while 3 surgeon evaluations ranged from $54–$108 [64]. Free technology, like Facebook Live, live streaming through Instagram, and live tweeting, gives crowdsourcing the potential to cost even less than previous estimates. Although limited studies are available for analysis, crowdsourcing appears promising in terms of delivering immediate formative evaluation and cost-effective feedback for surgical residents.

Conclusion

As the old surgical saying goes “there is more than one way to skin a cat”; based on the reviewed literature, it is clear that there is not only one best approach to performing evaluation and administering feedback to residents. Evidence suggests that a multimodal approach is ideal. In terms of technical skill improvement, use of video-based coaching, simulation with physiologically responsive models, and crowdsourcing are viable options. The use of standardized patients and controlled simulations of high-risk scenarios have aided in evaluation and feedback of resident communication skills. Some studies investigating new technologies are limited by small cohort size and more research is needed to further elucidate the role in feedback and evaluation. Nonetheless, new technology has been met with positivity as they can offer a cost-effective approach to evaluation and more timely feedback to learners, in the face of a changing landscape for surgical trainees.

Evaluation and feedback are areas of growing research as they are important to learner growth and medical education. While evaluation and feedback are often intertwined in the educational process, future investigations could focus on delineating the two processes. Research looking at formative and summative evaluation tools would likely yield clarity for educational faculty desiring to implement these tools. Accordingly, new feedback strategies should be equally transparent to allow the learner to become aware of the feedback session and prepare them to receive the feedback.

Compliance with Ethical Standards

Conflict of Interest Sara Perkins, Humphrey Atiemo, and Ali Dabaja each declare no potential conflicts of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:
- • Of importance
- •• Of major importance

1. Gofton WT, Dudek NL, Wood TJ, Balaa F, Hamstra SJ. The Ottawa Surgical Competency Operating Room Evaluation (O-SCORE): a tool to assess surgical competence. Academic medicine : Journal of the Association of American Medical Colleges. 2012;87(10):1401–7. https://doi.org/10.1097/ACM.0b013e3182677805.
2. Merriam-Webster.com Dictionary: Merriam-Webster Evaluation.
3. Gibbs T, Brigden D, Hellenberg D. Assessment and evaluation in medical education. S Afr Fam Pract. 2006;48(1):5–7. https://doi.org/10.1080/20786204.2006.10873311.
4. Guyot WM. Summative and formative evaluation. The Journal of Business Education. 1978;54(3):127–9. https://doi.org/10.1080/00219444.1978.10534702.
5. Merriam-Webster.com. Dictionary: Merriam-Webster. Feedback.

6. Lefroy J, Watling C, Teunissen PW, Brand P. Guidelines: the do's, don'ts and don't knows of feedback for clinical education. Perspect Med Educ. 2015;4(6):284–99. https://doi.org/10.1007/s40037-015-0231-7.
7. Carr BM, O'Neil A, Lohse C, Heller S, Colletti JE. Bridging the gap to effective feedback in residency training: perceptions of trainees and teachers. BMC Med Educ. 2018;18(1):225. https://doi.org/10.1186/s12909-018-1333-9.
8. Holmboe ES, Edgar L, Hamstra S. The Milestone Guidebook2016.
9. Education ACfGM. What we do. 2020. .
10.•• Education ACfGM. Surgery Milestones2019. Defines current surgery Milestones from ACGME**. Explains structure and elaborates on how programs should utilize the Milestones.**
11. Accreditation Council for Graduate Medical Education ABoU. The Urology Milestone Project 2016.
12. Bartlett KW, Whicker SA, Bookman J, Narayan AP, Staples BB, Hering H, et al. Milestone-based assessments are superior to Likert-type assessments in illustrating trainee progression. Journal of graduate medical education. 2015;7(1):75–80. https://doi.org/10.4300/JGME-D-14-00389.1.
13. Beeson MS, Holmboe ES, Korte RC, Nasca TJ, Brigham T, Russ CM, et al. Initial validity analysis of the emergency medicine milestones. Acad Emerg Med. 2015;22(7):838–44. https://doi.org/10.1111/acem.12697.
14.• Goldman RH, Tuomala RE, Bengtson JM, Stagg AR. How effective are new milestones assessments at demonstrating resident growth? 1 year of data. J Surg Educ. 2017;74(1):68–73. https://doi.org/10.1016/j.jsurg.2016.06.009 **Study explores the development of assessment tools based on the milestone system. Authors were able to develop an assessment tool that demonstrated growth of residents overtime and captured differences in skills that are quantifiable between PGY class.**
15. Gardner AK, AbdelFattah K. Comparison of simulation-based assessments and faculty ratings for general surgery resident milestone evaluation: are they telling the same story? Am J Surg. 2017;214(3):547–53. https://doi.org/10.1016/j.amjsurg.2016.07.025.
16.• Tekian A, Borhani M, Tilton S, Abasolo E, Park YS. What do quantitative ratings and qualitative comments tell us about general surgery residents' progress toward independent practice? Evidence from a 5-year longitudinal cohort. Am J Surg. 2019;217(2):288–95. https://doi.org/10.1016/j.amjsurg.2018.09.031 **Quantitative evaluations with area for comments are commonly used for resident feedback and evaluation; this long-term study demonstrates how qualitative feedback correlated with quantitative evaluation of residents.**
17.•• Clinial competency committees: a guidebook for programs. Accredidation Council for Graduate Medical Education2017 September 23, 2017. **Most current ACGME resource on Clinical Competency Committees which are required for accrediation. Explains that this committee is responsible for evaluating each residents progression using the Milestone system. However, it also highlights that each program is responsible for how assessment tools are administered, giving programs more freedom to utilize new techniques.**
18. Bosse HM, Mohr J, Buss B, Krautter M, Weyrich P, Herzog W, et al. The benefit of repetitive skills training and frequency of expert feedback in the early acquisition of procedural skills. BMC Med Educ. 2015;15:22. https://doi.org/10.1186/s12909-015-0286-5.
19. Sargeant J, Lockyer J, Mann K, Holmboe E, Silver I, Armson H, et al. Facilitated reflective performance feedback: developing an evidence- and theory-based model that builds relationship, explores reactions and content, and coaches for performance change (R2C2). Academic medicine : Journal of the Association of American Medical Colleges. 2015;90(12):1698–706. https://doi.org/10.1097/acm.0000000000000809.
20. Meier AH, Gruessner A, Cooney RN. Using the ACGME milestones for resident self-evaluation and faculty engagement. J Surg Educ. 2016;73(6):e150–e7. https://doi.org/10.1016/j.jsurg.2016.09.001.
21. Kuo LE, Hoffman RL, Morris JB, Williams NN, Malachesky M, Huth LE, et al. A milestone-based evaluation system-the cure for grade inflation? J Surg Educ. 2015;72(6):e218–25. https://doi.org/10.1016/j.jsurg.2015.09.012.
22.• Sebesta EM, Cooper KL, Badalato GM. Program director perceptions of usefulness of the accreditation council for graduate medical education milestones system for urology resident evaluation. Urology. 2019;124:28–32. https://doi.org/10.1016/j.urology.2018.10.042 **This study evaluated urology Program Directors' opinions on the ACGME Milestones. 48% of the sample felt that Milestones were somewhat or very unhelpful in evaluation; 58% felt that evaluations were uncorelated with in-service exam scores. This study highlights that the Milestone system is still in its infancy and that program directors perception of its utilization in urology demonstrates room for improvement.**
23. Ray JJ, Sznol JA, Teisch LF, Meizoso JP, Allen CJ, Namias N, et al. Association between American Board of Surgery in-training examination scores and resident performance. JAMA surgery. 2016;151(1):26–31. https://doi.org/10.1001/jamasurg.2015.3088.
24.• Edgar L, Roberts S, Holmboe E. Milestones 2.0: A Step Forward. J Graduate Med Educ. 2018;10(3):367–9. https://doi.org/10.4300/JGME-D-18-00372.1 **Highlights the development of Milestones 2.0 project to better improve the feedback and evaluation of reidents and add clarity to educational goals.**
25. Sidhu RS, Grober ED, Musselman LJ, Reznick RK. Assessing competency in surgery: where to begin? Surgery. 2004;135(1):6–20. https://doi.org/10.1016/s0039-6060(03)00154-5.
26. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. Br J Surg. 1997;84(2):273–8. https://doi.org/10.1046/j.1365-2168.1997.02502.x.
27. Williams RG, Verhulst S, Mellinger JD, Dunnington GL. Is a single-item operative performance rating sufficient? J Surg Educ. 2015;72(6):e212–7. https://doi.org/10.1016/j.jsurg.2015.05.002.
28. Rekman J, Hamstra SJ, Dudek N, Wood T, Seabrook C, Gofton W. A new instrument for assessing resident competence in surgical clinic: the Ottawa Clinic Assessment Tool. J Surg Educ. 2016;73(4):575–82. https://doi.org/10.1016/j.jsurg.2016.02.003.
29. Stausmire JM, Cashen CP, Myerholtz L, Buderer N. Measuring general surgery residents' communication skills from the patient's perspective using the Communication Assessment Tool (CAT). J Surg Educ. 2015;72(1):108–16. https://doi.org/10.1016/j.jsurg.2014.06.021.
30. Falcone JL, Claxton RN, Marshall GT. Communication skills training in surgical residency: a needs assessment and metacognition analysis of a difficult conversation objective structured clinical examination. J Surg Educ. 2014;71(3):309–15. https://doi.org/10.1016/j.jsurg.2013.09.020.
31. Joshi MK, Singh T, Badyal DK. Acceptability and feasibility of mini-clinical evaluation exercise as a formative assessment tool for workplace-based assessment for surgical postgraduate students. J Postgrad Med. 2017;63(2):100–5. https://doi.org/10.4103/0022-3859.201411.
32. Lynch T, Risucci D, Blair P, Sachdeva A. Implementation of the American College of Surgeons (ACS) OSCE to assess the knowledge and competency of entering surgical residents. J Surg Educ - J SURG EDUC. 2008;65:27. https://doi.org/10.1016/j.jsurg.2008.01.045.
33. Gee DW, Phitayakorn R, Khatri A, Butler K, Mullen JT, Petrusa ER. A pilot study to gauge effectiveness of standardized patient

scenarios in assessing general surgery milestones. J Surg Educ. 2016;73(6):e1–8. https://doi.org/10.1016/j.jsurg.2016.08.012.
34. Hochberg MS, Berman RS, Kalet AL, Zabar SR, Gillespie C, Pachter HL, et al. The professionalism curriculum as a cultural change agent in surgical residency education. Am J Surg. 2012;203(1):14–20. https://doi.org/10.1016/j.amjsurg.2011.05.007.
35. Posner G, Nakajima A. Assessing residents' communication skills: disclosure of an adverse event to a standardized patient. J Obstet Gynaecol Can. 2011;33(3):262–8. https://doi.org/10.1016/s1701-2163(16)34828-9.
36. Larkin AC, Cahan MA, Whalen G, Hatem D, Starr S, Haley H-L, et al. Human Emotion and Response in Surgery (HEARS): a simulation-based curriculum for communication skills, systems-based practice, and professionalism in surgical residency training. J Am Coll Surg. 2010;211(2):285–92. https://doi.org/10.1016/j.jamcollsurg.2010.04.004.
37. Hochberg MS, Kalet A, Zabar S, Kachur E, Gillespie C, Berman RS. Can professionalism be taught? Encouraging evidence. Am J Surg. 2010;199(1):86–93. https://doi.org/10.1016/j.amjsurg.2009.10.002.
38. Nicksa GA, Anderson C, Fidler R, Stewart L. Innovative approach using interprofessional simulation to educate surgical residents in technical and nontechnical skills in high-risk clinical scenarios. JAMA surgery. 2015;150(3):201–7. https://doi.org/10.1001/jamasurg.2014.2235.
39.• Kim SSY, Blankstein U, Ordon M, Pace KT, RJD H, Lee JY, et al. Evaluation of optimal timing of expert feedback in a simulated flexible ureteroscopy course. J Endourol. 2019;33(6):463–7. https://doi.org/10.1089/end.2018.0732 **Study examining simulation-based training on resident performance of ureteroscopy and optimal timing of feedback. Demonstrates how simulation can improve mean performance scores and that the majority of residents (73%) preferred early feedback from simulation as opposed to delayed feedbak from session.**
40. Setia S, Feng C, Coogan C, Vourganti S, Abern M. Urology residents' experience with simulation: initial evaluation of MRI/US fusion biopsy workshop. Urology. 2019;134:51–5. https://doi.org/10.1016/j.urology.2019.09.004.
41. Kneebone R. Simulation in surgical training: educational issues and practical implications. Med Educ. 2003;37(3):267–77. https://doi.org/10.1046/j.1365-2923.2003.01440.x.
42. Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA. 2011;306(9):978–88. https://doi.org/10.1001/jama.2011.1234.
43. Pena G, Altree M, Babidge W, Field J, Hewett P, Maddern G. Mobile simulation unit: taking simulation to the surgical trainee. ANZ J Surg. 2015;85(5):339–43. https://doi.org/10.1111/ans.12549.
44. Schwartz SI, Galante J, Kaji A, Dolich M, Easter D, Melcher ML, et al. Effect of the 16-hour work limit on general surgery intern operative case volume: a multi-institutional study. JAMA surgery. 2013;148(9):829–33. https://doi.org/10.1001/jamasurg.2013.2677.
45. Kairys JC, McGuire K, Crawford AG, Yeo CJ. Cumulative operative experience is decreasing during general surgery residency: a worrisome trend for surgical trainees? J Am Coll Surg. 2008;206(5):804–13. https://doi.org/10.1016/j.jamcollsurg.2007.12.055.
46. Snyder RA, Phillips SE, Terhune KP. Impact of implementation of a pediatric surgery fellowship on general surgery resident operative volume. Journal of surgical education. 2012;69(6):753–8. https://doi.org/10.1016/j.jsurg.2012.03.007.
47.•• Hu YY, Mazer LM, Yule SJ, Arriaga AF, Greenberg CC, Lipsitz SR, et al. Complementing operating room teaching with video-based coaching. JAMA Surg. 2017;152(4):318–25. https://doi.org/10.1001/jamasurg.2016.4619 **Development of new approaches to evaluation and feedback is important for improvement of medical education. This study demonstrates the development of novel video-based coaching in the OR and its utility in providing residents and faculty opportunity to discuss more complex topics during feedback sessions.**
48. Soucisse ML, Boulva K, Sideris L, Drolet P, Morin M, Dubé P. Video coaching as an efficient teaching method for surgical residents-a randomized controlled trial. Journal of surgical education. 2017;74(2):365–71. https://doi.org/10.1016/j.jsurg.2016.09.002.
49. Hu Y-Y, Peyre SE, Arriaga AF, Osteen RT, Corso KA, Weiser TG, et al. Postgame analysis: using video-based coaching for continuous professional development. J Am Coll Surg. 2012;214(1):115–24. https://doi.org/10.1016/j.jamcollsurg.2011.10.009.
50. Mazer LM, Hu YY, Arriaga AF, Greenberg CC, Lipsitz SR, Gawande AA, et al. Evaluating surgical coaching: a mixed methods approach reveals more than surveys alone. J Surg Educ. 2018;75(6):1520–5. https://doi.org/10.1016/j.jsurg.2018.03.009.
51. Moore MD, Abelson JS, O'Mahoney P, Bagautdinov I, Yeo H, Watkins AC. Using GoPro to give video-assisted operative feedback for surgery residents: a feasibility and utility assessment. J Surg Educ. 2018;75(2):497–502. https://doi.org/10.1016/j.jsurg.2017.07.024.
52. Bohnen JD, George BC, Williams RG, Schuller MC, DaRosa DA, Torbeck L, et al. The feasibility of real-time intraoperative performance assessment with SIMPL (System for Improving and Measuring Procedural Learning): early experience from a multi-institutional trial. J Surg Educ. 2016;73(6):e118–e30. https://doi.org/10.1016/j.jsurg.2016.08.010.
53. Williams RG, Verhulst S, Colliver JA, Sanfey H, Chen X, Dunnington GL. A template for reliable assessment of resident operative performance: assessment intervals, numbers of cases and raters. Surgery. 2012;152(4):517–24; discussion 24-7. https://doi.org/10.1016/j.surg.2012.07.004.
54. Eaton M, Scully R, Schuller M, Yang A, Smink D, Williams RG, et al. Value and barriers to use of the SIMPL tool for resident feedback. Journal of Surgical Education. 2019;76(3):620–7. https://doi.org/10.1016/j.jsurg.2019.01.012.
55. Harriman D, Singla R, Nguan C. The resident report card: a tool for operative feedback and evaluation of technical skills. J Surg Res. 2019;239:261–8. https://doi.org/10.1016/j.jss.2019.02.006.
56.• Hartranft TH, Yandle K, Graham T, Holden C, Chambers LW. Evaluating surgical residents quickly and easily against the milestones using electronic formative feedback. J Surg Educ. 2017;74(2):237–42. https://doi.org/10.1016/j.jsurg.2016.09.006 **This paper describes an institution's implementation of web-based evaluations for surgical residents. This study demonstrates increase in resident and faculty overall satisfaction with web-based evaluation technique.**
57. Green JM. An innovative, no-cost, evidence-based smartphone platform for resident evaluation. J Surg Educ. 2016;73(6):e14–e8. https://doi.org/10.1016/j.jsurg.2016.07.016.
58. Wagner JP, Chen DC, Donahue TR, Quach C, Hines OJ, Hiatt JR, et al. Assessment of resident operative performance using a real-time mobile web system: preparing for the milestone age. J Surg Educ. 2014;71(6):e41–6. https://doi.org/10.1016/j.jsurg.2014.06.008.
59. J H. The rise of crowdsourcing. 2006:1–5.
60. Boyle E, Al-Akash M, Gallagher AG, Traynor O, Hill AD, Neary PC. Optimising surgical training: use of feedback to reduce errors during a simulated surgical procedure. Postgrad Med J. 2011;87(1030):524–8. https://doi.org/10.1136/pgmj.2010.109363.
61.•• Dai JC, Lendvay TS, Sorensen MD. Crowdsourcing in surgical skills acquisition: a developing technology in surgical education. J Graduate Med Educ. 2017;9(6):697–705. https://doi.org/10.4300/

jgme-d-17-00322.1 **Systematic review of crowdsourcing and its utility in surgical training demonstrated that crowdsourced feedback was less expensive than feedback from expert surgeons and was more timely. This paper explores the valididity of crowd sourcings role in evaluation across multiple training levels and skill tasks. This paper provides evidence to support the use this technology in feedback and evaluation of surgical residents.**

62. Vernez SL, Huynh V, Osann K, Okhunov Z, Landman J, Clayman RV. C-SATS: assessing surgical skills among urology residency applicants. J Endourol. 2017;31(S1):S95–s100. https://doi.org/10.1089/end.2016.0569.
63. Aghdasi N, Bly R, White LW, Hannaford B, Moe K, Lendvay TS. Crowd-sourced assessment of surgical skills in cricothyrotomy procedure. J Surg Res. 2015;196(2):302–6. https://doi.org/10.1016/j.jss.2015.03.018.
64. White LW, Kowalewski TM, Dockter RL, Comstock B, Hannaford B, Lendvay TS. Crowd-sourced assessment of technical skill: a valid method for discriminating basic robotic surgery skills. J Endourol. 2015;29(11):1295–301. https://doi.org/10.1089/end.2015.0191.