


BMJ Open Cohort study of hospitalists' procedural skills: baseline competence and durability after simulation-based training

Caleb Hale,^{1,2} Jonathan Crocker,^{1,2} Anita Vanka,^{1,2} Daniel N Ricotta,^{1,2} Jakob I McSparron,³ Grace C Huang ^{1,2}

To cite: Hale C, Crocker J, Vanka A, *et al*. Cohort study of hospitalists' procedural skills: baseline competence and durability after simulation-based training. *BMJ Open* 2021;**11**:e045600. doi:10.1136/bmjopen-2020-045600

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-045600>).

CH and JC contributed equally.

Received 06 October 2020
Accepted 31 July 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Department of Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA

²Harvard Medical School, Boston, Massachusetts, USA

³Department of Medicine, University of Michigan Health System, Ann Arbor, Michigan, USA

Correspondence to

Dr Grace C Huang;
ghuang@bidmc.harvard.edu

ABSTRACT

Objectives Hospitalists are expected to be competent in performing bedside procedures, which are associated with significant morbidity and mortality. A national decline in procedures performed by hospitalists has prompted questions about their procedural competency. Additionally, though simulation-based mastery learning (SBML) has been shown to be effective among trainees whether this approach has enduring benefits for independent practitioners who already have experience is unknown. We aimed to assess the baseline procedural skill of hospitalists already credentialed to perform procedures. We hypothesised that simulation-based training of hospitalists would result in durable skill gains after several months.

Design Prospective cohort study with pretraining and post-training measurements.

Setting Single, large, urban academic medical centre in the USA.

Participants Twenty-two out of 38 eligible participants defined as hospitalists working on teaching services where they would supervise trainees performing procedures.

Interventions One-on-one, 60 min SBML of lumbar puncture (LP) and abdominal paracentesis (AP).

Primary and secondary outcome measures Our primary outcome was the percentage of hospitalists obtaining minimum passing scores (MPS) on LP and AP checklists; our secondary outcomes were average checklist scores and self-reported confidence.

Results At baseline, only 16% hospitalists met or exceeded the MPS for LP and 32% for AP. Immediately after SBML, 100% of hospitalists reached this threshold. Reassessment an average of 7 months later revealed that only 40% of hospitalists achieved the MPS. Confidence increased initially after training but declined over time.

Conclusions Hospitalists may be performing invasive bedside procedures without demonstration of adequate skill. A single evidence-based training intervention was insufficient to sustain skills for the majority of hospitalists over a short period of time. More stringent practices for certifying hospitalists who perform risky procedures are warranted, as well as mechanisms to support skill maintenance, such as periodic simulation-based training and assessment.

Strengths and limitations of this study

- The investigators relied on procedural checklists that have been used in multiple publications and have undergone formal standard setting procedures to determine competence thresholds.
- The raters were blinded to pretraining and post-training status.
- The sample size was small.
- There was significant participant attrition at time of delayed reassessment.

BACKGROUND

Internists who care for hospitalised patients, known as hospitalists, are ideally competent in bedside procedures. Accordingly, the Society for Hospital Medicine has defined performance standards for five invasive procedures: arthrocentesis, lumbar puncture (LP), abdominal paracentesis (AP), thoracentesis and central venous catheterisation.¹ However, how hospitalists should achieve and maintain procedural competency remains ill defined,² despite the significant morbidity and mortality associated with procedures.^{3–5} For instance, 2%–5% of APs are complicated by bleeding,⁴ bowel injury⁶ or persistent ascitic leakage,⁴ whereas about 20% of LPs can be complicated by minor complications such as postprocedural headache,⁷ or major complications in 2%–7% of cases, such as paraparesis, severe back pain, secondary meningitis or bleeding leading to spinal haematoma.^{7 8} Additionally, studies showing two decades of declining volume among internists (which included hospitalists and primary care physicians)^{9 10} contribute to concerns about their procedural competency.^{11 12} In contrast to a growing literature on metrics used to assess hospitalist performance in non-procedural realms,^{13–16} there has been no previous objective skills assessment of hospitalists who are already permitted to perform procedures, an

assertion supported by a recent literature review on the topic.¹⁷

Simulation-based mastery learning (SBML) is a recognised procedural training paradigm^{18–21} whereby clinicians practice and are coached to the point of proficiency on simulators without risk to patients. This approach can help address performance deficits among physicians and has the potential to augment skill gained primarily through experience. However, most research examines SBML's impact in trainees,^{22–23} who are generally true novices and tend to produce large effect sizes since they function at the steep portion of the learning curve.^{24–25} These studies do not include independently practising physicians who have already had prior experience with these procedures and for which training could be equated to booster training. Additionally, despite the well-established phenomenon of technical skill decay,^{26–27} no study has assessed skill retention in this population, especially in the setting of low procedural volumes. The limited available research on the natural history of technical skills among attending physicians has focused on difficult airway management among clinicians who have a significant procedural component to their usual practice.^{28–29}

We aimed to address two gaps in the literature. We wished to assess baseline procedural competence of hospitalists with prior experience in LP and AP who have been credentialed with the expectation that they can safely perform procedures. Second, we aimed to measure the impact of simulation-based training on the durability of hospitalists' skills to perform these procedures. We hypothesised that we would observe a sustained performance among the participants, as defined by >50% still exceeding the Minimum Passing Score (MPS).

METHODS

Study setting

We conducted a prospective cohort study of hospitalists in a large urban academic medical centre from July 2013 to March 2015. We chose to focus on LP and AP, which were two of the three invasive procedures expected of hospitalists at our institution (the other being joint arthrocentesis). To maintain procedural privileges, hospitalists must perform two LPs and two APs within each 2-year credentialing cycle, which is verified by billing data and procedural notes in the electronic medical record.

Patient and public involvement

As this was not a clinical research study, we did not involve patients or the public in the design and execution of the study.

Study participants

The primary inclusion criterion was hospitalists at our urban-based, academic medical centre who spent any amount of time attending on the teaching service, during which period they would have responsibility supervising

residents performing procedures. We were intentionally inclusive to give all hospitalists the opportunity to receive training. The only exclusion criterion was not being able to participate in the initial simulation-based training.

Study measures

Our primary outcome was the percentage of hospitalists obtaining passing scores on LP and AP checklists. Our secondary outcomes were average checklist scores and self-reported confidence. We relied on published checklists for which MPS had been defined for both procedures using robust standard-setting procedures; validity evidence for both checklists and the MPS had been substantiated through their use across multiple studies. The LP checklist had a maximal score of 21 and a predetermined MPS of 18²¹; the AP checklist had a maximal score of 25 and an MPS of 21.¹⁸ We conducted a sample size calculation based on the prior studies on trainees but lowered our expectations of the pre–post effect size, given that our participants were not true novices; no literature among hospitalists' procedural skill was available to support this adjustment and thus our estimate was relatively arbitrary and intentionally conservative. Our calculation showed that 18 participants entering this study would have 80% power to detect a difference in checklist scores at a two-sided 0.05 significance level, if the true difference between pretraining and post-training were 4 points (approximately 50% of the increase in checklist scores among novices) and the assumption that within-subject SD was 4 points (corresponding to a large effect size of 1.0).

To ensure participation across a spectrum of confidence levels among our faculty, we designed a survey and piloted the instrument among three hospital medicine colleagues who elected to not participate in the study. The primary questions were: 'Please rate your confidence in performing an LP on a scale of 1–5 (not at all, minimally, somewhat, fairly, very)' and 'Please rate your confidence in supervising an LP on a scale of 1–5.' We repeated the question for AP.

Intervention

We recruited volunteers among our group of 38 hospitalists via email and an in-person staff meeting, describing the project aims, processes and benefits/risks of participation. After obtaining written consent, we administered the survey to study participants in-person before training. We also collected demographic data (age, gender, years since residency) and the number of procedures each clinician recalled performing in the prior year. We then oriented participants to the procedural simulators that we would use both for training and assessment during the study. We videotaped participants performing LP and AP on these simulators to assess pretraining performance, which we considered to represent their baseline competence. One of the coauthors was physically present but did not intervene to give guidance or feedback on the procedure. The video camera was also intentionally

angled to avoid capturing faces so that participants would remain anonymous to raters.

Immediately after the baseline assessment, participants received one-on-one, 60 min simulation-based training for each procedure by one of 3 instructors (CH, JC and AV), employing Peyton's four-step model of procedure training as an educational framework.³⁰ Training began with participants viewing published LP and AP videos.^{31 32} The instructor then demonstrated the procedure on the simulator, verbalising each of the steps as listed on the checklist and in that specific order. The instructor then allowed the participant to practice on the simulator, while providing direct, specific feedback on the steps. We allowed the participants to practice as many times as they desired, either on particular steps or on the entire procedure from start to finish. In accordance with mastery learning practices,²² once a participant reported that they had had sufficient practice, the instructor assessed their performance against the checklist. If they did not achieve the MPS (ie, 18 points for LP, 21 points for AP), they then iteratively received direct feedback and targeted practice on the steps that were not achieved until they were able to independently complete the entire procedure with achievement of the MPS. They completed the survey in-person again immediately after training concluded.

We sent individual emails with multiple reminders to all participants to schedule reassessments 6 months after initial SBML. The subset of participants who responded were scheduled for delayed reassessment. At that time, participants performed LP and AP on simulators while being videotaped. They also repeated the self-confidence survey.

Two trained raters (DNR and JIM) independently reviewed all the videos at the end of the study, namely after completion of both baseline and delayed assessments. They scored participants' performance using the checklists. Thus, they were blinded to participant identity and timing (baseline vs delayed assessment). We calculated inter-rater reliability using inter-rater agreement and kappa statistics.

Analysis

We tabulated participant characteristics. We used Wilcoxon rank sum tests to compare mean checklist scores at baseline and delayed reassessment, and we used χ^2 or Fisher's exact tests to compare proportions of hospitalists meeting the MPS. Inferential statistics were performed using Stata V.12.0.

RESULTS

Baseline performance

Twenty-two of 38 eligible hospitalists participated. The average age was 36.5 (31–46) years. Participants reported an average of 5.9 (range 1–14) years in practice. At baseline, only 3 (16%) and 6 (32%) hospitalists achieved MPS for LP and AP, respectively. Participants scored on average 15 out of 21 items on the LP checklist and 17 out

Table 1 Baseline characteristics of hospitalists participating in simulation-based procedural training

| | Lumbar puncture | Paracentesis |
|---|-----------------|-----------------|
| | n (%) | n (%) |
| No of participants | 19 (100) | 22 (100) |
| Female participants | 13 (68) | 14 (64) |
| | Mean SD, range | Mean SD, range |
| Age | 36.5 5.1, 31–46 | 36.6 5.4, 31–46 |
| Years out of residency | 5.9 4.5, 1–14 | 5.6 4.0, 1–14 |
| Self-reported procedural experience in the past 12 months | 2.8 2.8, 0–10 | 2.2 2.9, 0–10 |
| Comfort performing procedure (no reporting 'fairly' or 'very comfortable') | 10 (53) | 11 (50) |
| Comfort supervising procedure (no reporting 'fairly' or 'very comfortable') | 11 (58) | 11 (50) |

of 25 on the AP checklist. The cohort cited only moderate levels of comfort performing and supervising LP and AP, respectively (range 50%–58% reporting 'fairly' or 'very comfortable') (table 1) with a mean response of 'somewhat' comfortable for each procedure, respectively.

Delayed reassessment

Skills reassessment skills occurred at a mean of 7 months post-training (table 2). We were only able to schedule 10 hospitalists given limited availability in the simulation centre. The number of hospitalists achieving passing scores was 40% and 40% for LP and AP, respectively. They attained average scores of 16 and 20 for LP and AP, respectively.

Procedural confidence

Confidence performing and supervising LPs and APs increased immediately after SBML and declined by the time of delayed reassessment, but not fully to baseline (figure 1). The inter-rater agreement and inter-rater reliability via the kappa statistic for overall checklist scores were 0.87 and 0.62 for LP and 0.81 and 0.29 for AP, respectively.

DISCUSSION

In our cohort study of academic hospitalists undergoing SBML training for bedside procedures, most did not meet predefined thresholds for LP and AP competence at baseline. Fewer than half maintained procedural competency several months after simulation-based training. Confidence increased immediately after training but then diminished. Our work is the first to document low levels of procedural competence among academic hospitalists already credentialed to perform procedures on patients

Table 2 Hospitalists' technical performance on lumbar puncture and paracentesis simulators, at baseline and several months after simulation-based training

| | Lumbar puncture | Paracentesis |
|---|-------------------|-------------------|
| Baseline performance | | |
| No of hospitalists with passing scores, n (%) | 3 (16) | 6 (32) |
| Mean scores (SD, range) | 15.7 (2.1, 11–19) | 17.9 (4.6, 9–25) |
| Immediate post-training performance | | |
| No of hospitalists with passing scores, n (%) | 19 (100) | 22 (100) |
| Several months after simulation-based training (n=10) | | |
| No of hospitalists with passing scores, n (%) | 4 (40) | 4 (40) |
| Mean scores (SD, range) | 16.4 (2.7, 13–20) | 20.1 (2.3, 16–24) |
| P-values | | |
| Percent passing, pre vs delayed post | 0.193 | 0.698 |
| Scores, pre vs delayed post | 0.472 | 0.268 |

The minimum passing scores for lumbar puncture and abdominal paracentesis were 18 out of 21 and 21 out of 25, respectively.

and to show the limited durability of SBML in the setting of low procedural volumes.

We were not surprised that our clinicians, consisting largely of early-career physicians from diverse programmes, did not perform adequately at baseline. The American Board of Internal Medicine has since 2006 mandated only that residents know how to perform certain procedures, rather than achieve a technical competency standard; most hospitalists who participated in this study completed residency after this policy revision. Of note, a more recent proposal to eliminate identifying the core set of procedures altogether remains underway.³³ These

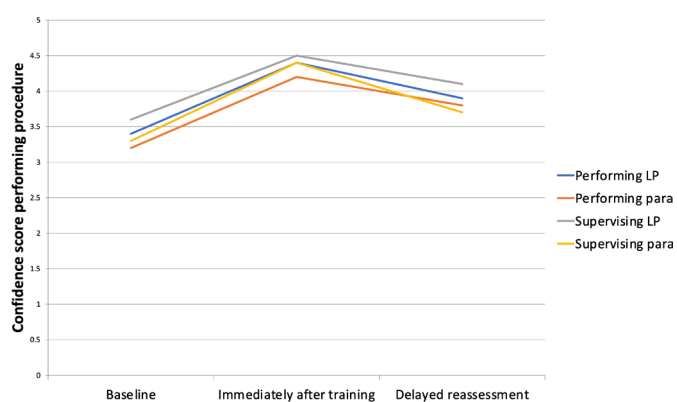


Figure 1 Confidence score (range 1–5) performing and supervising lumbar punctures and abdominal paracentesis, at baseline, after training and at the time of delayed reassessment. LP, lumbar puncture.

regulatory shifts have resulted in residents graduating without the expectation of psychomotor skill as independent providers, creating a pipeline of hospitalists unprepared to perform or teach LP and AP. Additionally, the items most commonly missed on the checklists related to cognitive elements—ordering the appropriate studies and alerting the nurse about the completion of the procedures; these aspects of a diagnostic procedure are essential but may have been neglected in the overemphasis on psychomotor skills.

Additionally, we suspect that hospitalists being unable to sustain procedural competency was likely due to lack of ongoing experience. While we did not audit their procedural volume in the intervening 7 months, the number of procedures our participants self-reported in the year prior to our study were quite low (2.8 for LP, 2.2 for AP), and these numbers are similar in magnitude to an audit we conducted in 2012 (average 2.7 for LP, 1.9 for AP among hospitalists who billed for procedures). If representative of the procedures available to them, hospitalists would have had scarce opportunity to benefit from the booster effect of procedural volume on skill refinement.³⁴ In the current day, our procedural volume continues to be low, averaging less than 1 AP and about 1 LP per hospitalist per year, mirroring the national decline in procedural volume for practising internists.¹⁰ An alternative explanation for the lack of sustained competency could potentially relate to a focus on ‘teaching to the checklist,’ which may have hindered the ability for hospitalists to see the procedure as an integrated whole; alternatively, insufficient adoption of the checklist in clinical practice could have resulted in weak skill retention.^{35 36} However, the SBML approach has extensive research evidence in its support,²² and we abided by its principles closely. We speculate that for low-frequency procedures, procedural experience may play just as large a role in maintaining skill as training itself.

Our finding related to hospitalists' skill decay has no analogue in the literature but is supported by studies in other populations that have examined retention of psychomotor performance. Many studies (with one notable exception³⁷ have shown that Advanced Cardiac Life Support (ACLS) skills decay quickly,^{38–40} though arguably ACLS is more a cognitive skill with some technical aspects than it is a complex procedural skill. One published abstract on skill retention after LP training for neurology residents showed that performance scores remained significantly elevated 1 year later but they did not maintain mastery standards¹⁹; neurology residents likely perform LPs at a higher volume and frequency and thus are not readily comparable to hospitalists. Another study of nephrology fellows inserting haemodialysis catheters⁴¹ affirmed that, based on the observation of skill decline at 1 year, a 6-month booster training would be recommended. Most recently, a study of internal medicine residents trained in paracentesis skills showed that skill decay was mitigated if booster training occurred at 3 months rather than 6 months.⁴² These two investigations

shed light on the impact of booster training on procedural skills in nonsurgeons but are imperfect as proxies for hospitalists, who are no longer trainees and have accumulated more experience. Insights from studies on monitoring hospitalist performance in non-procedural skills are also germane in concept, as they focus on collecting and reporting objective individual-level¹⁶ and group-level metrics,¹³ approaches applicable to procedural competence tracking. Ultimately, our study examines a population that has been overlooked in terms of ensuring competence, while highlighting the practical realities that must be overcome for the scalability of this work for other hospital medicine groups.

Limitations of our study include our single institutional setting and underpowered results due to small sample size, including uneven numbers of participants for the two procedures. Our power calculation required making an assumption about skills gain without literature support. Volunteerism may have selected for individuals less comfortable with procedural skills, and we did not compare demographics or procedural experience data from non-participants to quantify the degree of selection bias. Our study was also constrained with respect to the ultimate goal of assessing true competence; though we also measured procedural confidence, this construct is well known not to correlate with true competence. Our simulation-based skills assessment may not be an accurate proxy for performance in real-world settings, and we did not collect data on the actual procedures performed by the study participants after training, such as number of needle passes, time to completion or complication rates. Furthermore, measurement of complication rates would not be a sensitive indicator of competence, given the low volume of procedures performed each year. The inter-rater reliability for AP was unexpectedly low and may have stemmed from challenges we faced having a static video camera capture all angles of the procedure being performed. Additionally, we noted that the kappa statistic was low despite high inter-rater agreement (>0.80 for both procedures); this pattern is characteristic of instruments with low variability, termed the 'Kappa paradox',^{43 44} and in hindsight, we should have measured inter-rater agreement alone. We experienced a significant attrition of participants between the pre-test and the delayed reassessment due to a myriad of factors that speak to challenges for research among attending physicians (eg, simulation centre constraints, hospitalist and instructor scheduling availability), which may have biased the results either toward hospitalists sufficiently confident in their skills to undergo reassessment or towards physicians unsure about their skill level. We did not track procedural volume in the intervening months between training and retesting, which would have disclosed the extent to which hospitalists had occasion to apply their training. We did not systematically teach or assess ultrasound skill, which has become a standard of care for AP. We did not measure performance beyond 7 months to delineate the natural history of procedural skills over time. We did not

specifically investigate the benefit of interval, repeated SBML compared with a single episode of training.

These early-stage findings in an overlooked line of inquiry test the assumption that hospitalists are equipped to perform in-hospital procedures and supports an argument to systematically assess their baseline technical skills as a part of the credentialing process. Recommendations to augment hospitalists' procedural skill sets have included booster training, frequent interval assessment and the development of core 'proceduralists'.^{11 17 45-47} Evidence-based tactics for mitigating skill decay at the time of initial training include using visual or verbal cues for recognition and behavioural feedback.²⁶ A prospective research agenda should include an examination of the contribution of procedural experience to hospitalists' skills over time, a comparison of academic and community hospital settings, where practice scope may differ, and a time-series analysis to determine the optimal frequency of procedural training to maintain competence. It is also important to note that neither credentialing nor fixed-interval assessment addresses the fact that individuals learn psychomotor skills at different rates and that various factors will influence the slope of skill decline. Thus, future initiatives should consider levelling of hospitalists according to frameworks of skill acquisition.⁴⁸

The workforce culture of healthcare is such that physicians graduated from residency are assumed to be proficient practitioners, technical skills being no exception. Institutions (and patients) continue to presume that hospitalists have adequate skill in this domain. Our work, although quite limited in size and scope, challenges this notion by documenting poor procedural skill at baseline and marginal benefit from single-episode training. Overlooking this potential source of medical error may have significant consequences on hospitalised patients, who may not question the technical skill of their bedside physicians.

Twitter Grace C Huang @GraceHuangMD

Contributors CH and JC contributed equally to the study design, study execution and the substantive writing of this manuscript. Study design: CH, JC and GH. Study implementation: CH, JC, AV, DR and JM. Data collection: CH, JC, AV, DR and JM. Data analysis: GH. Manuscript preparation: GH, CH, JC, AV, DR and JM.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval The Beth Israel Deaconess Medical Center Committee on Clinical Investigations/Institutional Review Board approved the study under the expedited mechanism.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. The data are available on request to ghuang@bidmc.harvard.edu.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made

indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Grace C Huang <http://orcid.org/0000-0003-2965-0341>

REFERENCES

- The core competencies in hospital medicine: a framework for curriculum development by the Society of hospital medicine. *J Hosp Med* 2006;1 Suppl 1:2–95.
- Jensen TP, Soni NJ, Tierney DM, et al. Hospital Privileging practices for bedside procedures: a survey of hospitalist experts. *J Hosp Med* 2017;12:836–9.
- Leape LL, Brennan TA, Laird N, et al. The nature of adverse events in hospitalized patients. Results of the Harvard medical practice study II. *N Engl J Med* 1991;324:377–84.
- De Gottardi A, Thévenot T, Spahr L, et al. Risk of complications after abdominal paracentesis in cirrhotic patients: a prospective study. *Clin Gastroenterol Hepatol* 2009;7:906–9.
- Ault MJ, Rosen BT, Scher J, et al. Thoracentesis outcomes: a 12-year experience. *Thorax* 2015;70:127–32.
- Runyon BA, Hoefs JC, Canawati HN. Polymicrobial bacterascites. A unique entity in the spectrum of infected ascitic fluid. *Arch Intern Med* 1986;146:2173–5.
- Ruff RL, Dougherty JH. Complications of lumbar puncture followed by anticoagulation. *Stroke* 1981;12:879–81.
- Pitkänen MT, Aromaa U, Cozantitis DA, et al. Serious complications associated with spinal and epidural anaesthesia in Finland from 2000 to 2009. *Acta Anaesthesiol Scand* 2013;57:553–64.
- Wigton RS, Alguire P, American College of Physicians. The declining number and variety of procedures done by general internists: a resurvey of members of the American College of physicians. *Ann Intern Med* 2007;146:355–60.
- Thakkar R, Wright SM, Alguire P, et al. Procedures performed by hospitalist and non-hospitalist General internists. *J Gen Intern Med* 2010;25:448–52.
- Crocker JT, Hale CP, Vanka A, et al. Raising the bar for procedural competency among hospitalists. *Ann Intern Med* 2019;170:654.
- Lucas BP, Asbury JK, Franco-Sadud R. Training future hospitalists with simulators: a needed step toward accessible, expertly performed bedside procedures. *J Hosp Med* 2009;4:395–6.
- Hwa M, Sharpe BA, Wachter RM. Development and implementation of a balanced scorecard in an academic hospitalist group. *J Hosp Med* 2013;8:148–53.
- Rosenthal MA, Sharpe BA, Haber LA. Using peer feedback to promote clinical excellence in hospital medicine. *J Gen Intern Med* 2020;35:3644–9.
- Nelson JR. Assessing individual hospitalist performance: domains and Attribution. *J Hosp Med* 2020;15:639–40.
- Dow AW, Chopski B, Cyrus JW, et al. A STEEEP Hill to Climb: a scoping review of assessments of individual hospitalist performance. *J Hosp Med* 2020;15:599–605.
- Cool JA, Huang GC. Procedural competency among hospitalists: a literature review and future considerations. *J Hosp Med* 2021;16:230–5.
- Barsuk JH, Cohen ER, Vozenilek JA, et al. Simulation-Based education with mastery learning improves paracentesis skills. *J Grad Med Educ* 2012;4:23–7.
- Pressman P, Burroughs B, Gourineni R, et al. Retention of lumbar puncture skills gained by a mastery learning approach using simulation technology and deliberate practice (P07.236). *Neurology* 2012;78:P07.236.
- McGaghie WC, Issenberg SB, Barsuk JH, et al. A critical review of simulation-based mastery learning with translational outcomes. *Med Educ* 2014;48:375–85.
- Barsuk JH, Cohen ER, Caprio T, et al. Simulation-Based education with mastery learning improves residents' lumbar puncture skills. *Neurology* 2012;79:132–7.
- Cook DA, Brydges R, Zendejas B, et al. Mastery learning for health professionals using technology-enhanced simulation: a systematic review and meta-analysis. *Acad Med* 2013;88:1178–86.
- Chang W, Popa A, DeKorte M. A medical invasive procedure service and resident procedure training elective. *J Hosp Med* 2012;7:S120 <http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L70698186>
- White C, Rodger MWM, Tang T. Current understanding of learning psychomotor skills and the impact on teaching laparoscopic surgical skills. *Obstet Gynecol* 2016;18:53–63.
- Crossman ERFW. A theory of the acquisition of SPEED-SKILL*. *Ergonomics* 1959;2:153–66.
- Arthur Jr. W, Bennett Jr. W, Stanush PL, et al. Factors that influence skill decay and retention: a quantitative review and analysis. *Human Performance* 1998;11:57–101.
- Moser DK, Coleman S. Recommendations for improving cardiopulmonary resuscitation skills retention. *Heart Lung* 1992;21:372–80 <https://pubmed.ncbi.nlm.nih.gov/1629007/>
- Boet S, Borges BCR, Naik VN, et al. Complex procedural skills are retained for a minimum of 1 yr after a single high-fidelity simulation training session. *Br J Anaesth* 2011;107:533–9.
- Kuduvalli PM, Jervis A, Tighe SQM, et al. Unanticipated difficult airway management in anaesthetised patients: a prospective study of the effect of mannequin training on management strategies and skill retention. *Anaesthesia* 2008;63:364–9.
- Walker M, Peyton J. Teaching in Theatre. In: *Teaching and learning in medical practice*. Manticore Publishers Europe, 1998: 171–80.
- Ellenby MS, Tegtmeyer K, Lai S. Videos in clinical medicine. lumbar puncture. *N Engl J Med* 2006;355:e12.
- Thomsen TW, Shaffer RW, White B, et al. Videos in clinical medicine. paracentesis. *N Engl J Med* 2006;355:e21.
- Internal medicine board considers revised procedural requirements. American Board of internal medicine Blog. *Published* 2019 <http://blog.abim.org/internal-medicine-board-considers-revised-procedural-requirements/>
- Barsuk JH, Cohen ER, Feinglass J, et al. Residents' procedural experience does not ensure competence: a research synthesis. *J Grad Med Educ* 2017;9:201–8.
- Levy SM, Senter CE, Hawkins RB, et al. Implementing a surgical checklist: more than checking a box. *Surgery* 2012;152:331–6.
- Eva KW, Regehr G. Self-Assessment in the health professions: a reformulation and research agenda. *Acad Med* 2005;80:S46–54.
- Wayne DB, Siddall VJ, Butter J, et al. A longitudinal study of internal medicine residents' retention of advanced cardiac life support skills. *Acad Med* 2006;81:S9–12.
- Woollard M, Whitfield R, Smith A, et al. Skill acquisition and retention in automated external defibrillator (AED) use and CPR by lay responders: a prospective study. *Resuscitation* 2004;60:17–28.
- Moser DK, Dracup K, Guzy PM, et al. Cardiopulmonary resuscitation skills retention in family members of cardiac patients. *Am J Emerg Med* 1990;8:498–503.
- Hamilton R. Nurses' knowledge and skill retention following cardiopulmonary resuscitation training: a review of the literature. *J Adv Nurs* 2005;51:288–97.
- Ahya SN, Barsuk JH, Cohen ER, et al. Clinical performance and skill retention after simulation-based education for nephrology fellows. *Semin Dial* 2012;25:470–3 <http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L51853897>
- Sall D, Warm EJ, Kinnear B, et al. See one, do one, forget one: early skill decay after paracentesis training. *J Gen Intern Med* 2021;36:1–6.
- Zec S, Soriani N, Comoretto R, et al. High agreement and high prevalence: the paradox of Cohen's kappa. *Open Nurs J* 2017;11:211–8.
- Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990;43:543–9.
- Ault MJ, Rosen BT. Proceduralists--leading patient-safety initiatives. *N Engl J Med* 2007;356:1789–90.
- Ibrahim H, Stadler DJ, Archuleta S, et al. Twelve tips for developing and running a successful women's group in international academic medicine. *Med Teach* 2019;41:1239–44.
- Sawyer T, White M, Zaveri P, et al. Learn, see, practice, prove, do, maintain: an evidence-based pedagogical framework for procedural skill training in medicine. *Acad Med* 2015;90:1025–33.
- Carraccio CL, Benson BJ, Nixon LJ, et al. From the educational bench to the clinical bedside: translating the Dreyfus developmental model to the learning of clinical skills. *Acad Med* 2008;83:761–7.