Core Competencies or a Competent Core? A Scoping Review and Realist Synthesis of Invasive Bedside Procedural Skills Training in Internal Medicine

Ryan Brydges, PhD, Lynfa Stroud, MD, MEd, Brian M. Wong, MD, Eric S. Holmboe, MD, Kevin Imrie, MD, and Rose Hatala, MD, MSc

Abstract

Purpose
Invasive bedside procedures are core competencies for internal medicine, yet no formal training guidelines exist. The authors conducted a scoping review and realist synthesis to characterize current training for lumbar puncture, arthrocentesis, paracentesis, thoracentesis, and central venous catheterization. They aimed to collate how educators justify using specific interventions, establish which interventions have the best evidence, and offer directions for future research and training.

Method
The authors systematically searched Medline, Embase, the Cochrane Library, and ERIC through April 2015. Studies were screened in three phases; all reviews were performed independently and in duplicate. The authors extracted information on learner and patient demographics, study design and methodological quality, and details of training interventions and measured outcomes. A three-step realist synthesis was performed to synthesize findings on each study's context, mechanism, and outcome, and to identify a foundational training model.

Results
From an initial 6,671 studies, 149 studies were further reduced to 67 (45%) reporting sufficient information for realist synthesis. Analysis yielded four types of procedural skills training interventions. There was relative consistency across contexts and significant differences in mechanisms and outcomes across the four intervention types. The medical procedural service was identified as an adaptable foundational training model.

Conclusions
The observed heterogeneity in procedural skills training implies that programs are not consistently developing residents who are competent in core procedures. The findings suggest that researchers in education and quality improvement will need to collaborate to design training that develops a "competent core" of proceduralists using simulation and clinical rotations.

The American Board of Internal Medicine (ABIM)[1] and the Royal College of Physicians and Surgeons of Canada (RCPSC)[2] mandate that internal medicine (IM) physicians be competent in core invasive bedside procedures, yet neither provides formal guidelines for how to train learners to be procedurally competent. The absence of evidence-based standards for training and assessment is a critical gap because performing such procedures is not without risk. Procedural errors and complications can result in increased patient discomfort, longer hospital stays, and higher costs.[3] Procedural complications are also a leading cause of adverse events identified in most national adverse event studies.[4–9] A recent systematic review suggests that procedural complications result in 6.7% to 9.7% of hospital-wide adverse events and that nearly half of these events are considered preventable.[10] IM programs require further guidance on the training model (or models) that develops internists who perform procedures competently.

With the competing demands of trainees needing opportunities to acquire skills and of patients expecting high-quality, safe health care, a delicate balance exists between medical education and patient safety. The majority of IM residents report performing fewer than five invasive bedside procedures during their undergraduate medical training,[11] insufficient exposure to procedural skills during residency training,[12] a lack of proficient faculty to supervise procedures,[13] and low levels of comfort and confidence when performing procedures.[14] These reports raise significant concern that IM programs currently offer inconsistent procedural skills training experiences that may lead to incompetent trainees and put patients at unnecessary risk. Supporting that concern, one survey showed that 70% of Canadian IM program directors agree that a national standard for assessing procedural competence would be beneficial.[15] Improving procedural skills training can have broad implications for the health care system; for example, training interventions have been associated with reduced incidence of complications and preventable adverse events.[16–18]

To date, syntheses of the evidence on invasive bedside procedural skills training in IM include narrow systematic reviews on the use of simulation-based training for central venous catheterization[19,20]; a broad systematic review and meta-analysis of studies that are "heterogeneous and of varying quality and rigour"[21]; and a nonsystematic, critical synthesis presenting a broad conceptual framework for procedural skills training.[22] We suggest there is a need for a comprehensive review using a systematic search paired with a technique

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

Correspondence should be addressed to Ryan Brydges, St. Michael's Hospital and Department of Medicine, University of Toronto, 209 Victoria St, Rm 5-86, Toronto, ON M5B 1T8, Canada; telephone: (416) 864-6060, ext 77530; e-mail: ryan.brydges@utoronto.ca.

Acad Med. 2017;92:1632–1643.
First published online May 9, 2017
doi: 10.1097/ACM.0000000000001726

Supplemental digital content for this article is available at http://links.lww.com/ACADMED/A449.

that is designed for synthesizing data from heterogeneous studies.

Accordingly, we conducted a scoping review and realist synthesis of this literature and aimed to characterize current procedural skills training interventions, collate how educators justify the interventions used in their programs, establish which interventions have the best evidence, and offer directions for future research and training. Our research question was: What can we learn from previous interventions designed to establish competence in five invasive bedside procedures (lumbar puncture, arthrocentesis, paracentesis, thoracentesis, and central venous catheterization [hereafter, the five invasive bedside procedures]) that are considered core competencies for internists in the United States[1] and Canada?[2]

Method

We combined two complementary knowledge synthesis techniques: scoping review and realist synthesis. A scoping review is used to address "an exploratory research question aimed at mapping key concepts, types of evidence, and gaps in research related to a defined area or field by systematically searching, selecting, and synthesizing existing knowledge."[23] To provide an analysis beyond only describing the included studies, we used realist synthesis, which requires researchers to "unpack the context-mechanism-outcome relationship [associated with a training intervention], thereby explaining examples of success, failure, and various eventualities in between."[24]

One explicit purpose of realist synthesis is to compare official expectations versus actual practice.[24] In IM training, we suggest that the official expectation is that all residents develop competence in the five invasive bedside procedures. By contrast, we expected that our synthesis would reveal great variability among actual practices including training contexts, authors' proposed educational mechanisms for why a training intervention would work, and the outcomes used to measure the intervention's success.

We planned, conducted, and reported our scoping review and realist synthesis to adhere to the methodological steps for scoping reviews,[23,25] publication standards for realist synthesis,[26] and the STORIES statement for health care education evidence synthesis.[27] Given that our review did not involve human participants, it was exempt from ethical review at all of the associated institutions.

Data sources, searches, and inclusion criteria

An experienced information specialist designed a peer-reviewed strategy (Supplemental Digital Appendix 1 at http://links.lww.com/ACADMED/A449)[28] for us to use to systematically search the Medline, Embase, Cochrane Library, and ERIC (Education Resources Information Center) databases from inception to December 13, 2013. Search terms included medical subject headings and terms related to learner populations (e.g., medical student, resident, hospitalist, rheumatologist, fellows), to a training or assessment focus (e.g., certification, licensure, assessment), and to the procedures of interest (e.g., paracentesis, arthrocentesis). We updated our search on April 13, 2015, using the original terms, with searches in Medline and Embase only, as these two databases returned the greatest number and most specific results in our original search. We supplemented both searches by hand-searching the reference lists of published reviews and relevant journals, searching our own files, and consulting with colleagues who publish in the domain.

We included studies published in English language only. For participants, we included studies on any health care professional to capture procedural experience in multiple clinical settings (e.g., paracentesis in gastroenterology). We included studies across nonclinical (e.g., simulation laboratory) and clinical (e.g., academic hospitals) settings. We included training interventions focused on procedural skills training and/or on reorganizing clinical practices associated with invasive procedures. We chose the five invasive bedside procedures mentioned above because all are core competencies for internists in the United States[1] and Canada,[2] are performed frequently, and are linked to patient complications. We limited our focus to these five procedures to ensure synthesis feasibility. We included studies with any outcome of interest, such as measures of procedural competence, quality of care, and patient safety. With respect to study design, we included studies using any method (qualitative or quantitative), and both full-text articles and conference abstracts (collectively referred to as studies).

Study selection

We removed duplicate studies initially using EndNote software (Version X7.7.1, Thomson Reuters, New York, NY). We conducted study screening in three phases using DistillerSR (Web-based software, Evidence Partners, Ottawa, Ontario, Canada), performing all reviews independently and in duplicate. First, several of us (R.B., R.H., L.S., B.M.W.) pilot-reviewed 15 random study abstracts, developed consensus on the operational criteria for judging study inclusion, and discussed the adequacy of the search strategy. We then reviewed the abstracts of all studies and resolved conflicts by consensus (weighted kappa = 0.46, moderate agreement). Second, we reviewed all of the full-text articles, resolving conflicts by consensus. Interrater agreement (kappa) for the full-text review was 0.84 (original research reporting empirical data), 0.63 (focus on IM procedures), and 0.65 (focus on training). Third, we reviewed all of the studies, judging whether the context–mechanism–outcome linkage was a good representation or useful refutation of the proposed training intervention. This analysis resulted in our excluding many studies—for example, one in which the authors described the rationale as "a need to utilize novel technology" (context), designed a simulation-based intervention to improve physicians' procedural competence (mechanism), and measured impact as teachers' and learners' perceived utility of training (outcome). We excluded this study on the basis of the superficial rationale, and the misalignment between the goal of training (i.e., improved trainee skills) and the performance outcome (i.e., self-report). In making these judgments, our interrater agreement was moderate (kappa = 0.47), and for the studies with disagreement between the two reviewers, we resolved conflicts via discussions with a third reviewer.

Quality assessment

Independently and in duplicate, we (R.B., R.H., L.S., B.M.W.) extracted information on learner and patient demographics, study design and methodological quality, and details of the training interventions and measured outcomes. To evaluate study quality of the full-text articles (but not the conference abstracts), we used the MERSQI (Medical Education Research Study Quality Instrument).[29] We resolved all conflicts via group consensus.

Data extraction

For studies with clearly described and well-aligned context–mechanism–outcome linkages, we used a structured form (Supplemental Digital Appendix 2 at http://links.lww.com/ACADMED/A449) to extract data for each component, which we operationalized as follows. Context included which cultural, institutional, societal, and/or regulatory drivers prompted the training intervention and details about the setting (or settings) in which the intervention was delivered. Mechanism included the theory or conceptual framework authors associated with their interventions and the process by which the intervention was expected to affect procedural skills competence. Outcome included authors' proposed links between the outcomes they measured and the contexts and mechanisms they used, as well as the intended and unintended consequences associated with the intervention.

Data synthesis and analysis

After iterative team discussions, we (all authors) decided to aggregate studies according to mechanisms because organizing them this way better aligned with regulatory requirements from the RCPSC and ABIM, which do not attend to either context or outcome (i.e., they are usually focused on the training process), and resulted in a parsimonious set of four types of training interventions. We then performed our realist synthesis in three steps (outlined in Figure 1):

Step 1. In step 1, we synthesized findings for context, mechanism, and outcome separately for the studies categorized in each of the four intervention types. We followed the realist principle of seeking demiregularities[26] (in this case, details that were consistently present or absent in authors' descriptions) in how authors described their context, mechanism, and outcome.

Step 2. In step 2, we analyzed the results of step 1 and produced syntheses of the findings for context, mechanism, and outcome separately across all four intervention types.

Step 3. In step 3, we analyzed the collective syntheses from step 2, with the perspective that realist syntheses should "build explanations across interventions … that share similar underlying 'theories of change' as to why they work (or do not work) … in particular contexts."[30] Here, we aimed to identify a foundational training model that could be adapted to accommodate the lessons learned from all three steps of our realist synthesis.

Finally, we (all authors) presented preliminary results to a group of relevant stakeholders (26 IM program directors, residents, and researchers) during a conference in fall 2014 and discussed how the foundational training model we identified would interface with current training policies.[26,31] After iterative rounds of consensus building, we produced a list of key components of future IM procedural skills training and developed a list of key lines of inquiry for future IM curriculum design and research.

Results

Study characteristics

Our initial and updated searches yielded 6,671 relevant studies, from which we identified 149 for full data extraction (91 full-text articles and 58 conference abstracts). Data on the characteristics of these 149 studies are included in Supplemental Digital Appendixes 3 and 4 (at http://links.lww.com/ACADMED/A449). Of these, we found that only 67 (45%) studies reported sufficient information about the context, mechanism, and outcome associated with the training intervention for realist synthesis. Figure 2 shows the flow of study inclusion or exclusion, and Table 1 provides a summary of study characteristics for all 67 included studies (63 full-text articles and 4 conference abstracts).[16–18,32–95]

Realist synthesis findings

Our analysis yielded four types of procedural skills training interventions: "see one, do one"[17,32–52]; educational-theory-informed (divided into mastery learning[16,53–69] and other, including self-regulated learning and cognitive,

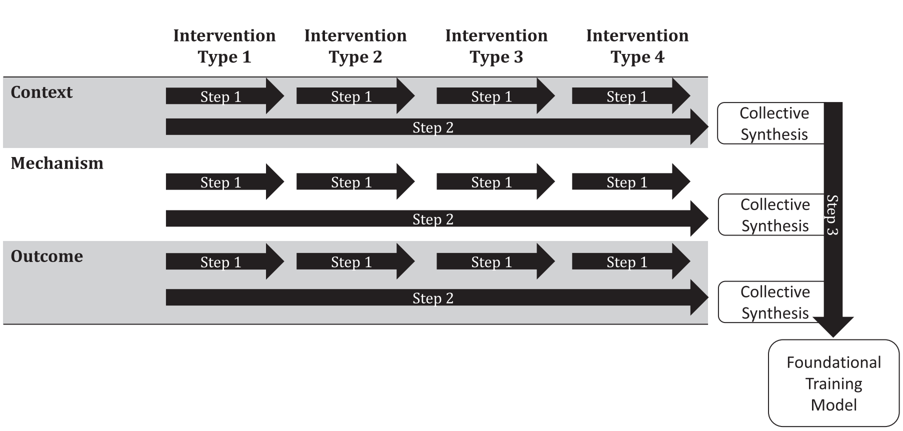

Figure 1 Diagram of the authors' realist synthesis process, used in a 2015 scoping review and realist synthesis of the literature on invasive bedside procedural skills training in internal medicine. In step 1, the authors synthesized findings for context, mechanism, and outcome separately for the studies categorized in each of the four intervention types. In step 2, the authors produced collective syntheses from the step 1 results for context, mechanism, and outcome separately across all four intervention types. In step 3, the authors analyzed the collective syntheses from step 2 to identify a foundational training model that could be adapted to accommodate the lessons learned from this process.

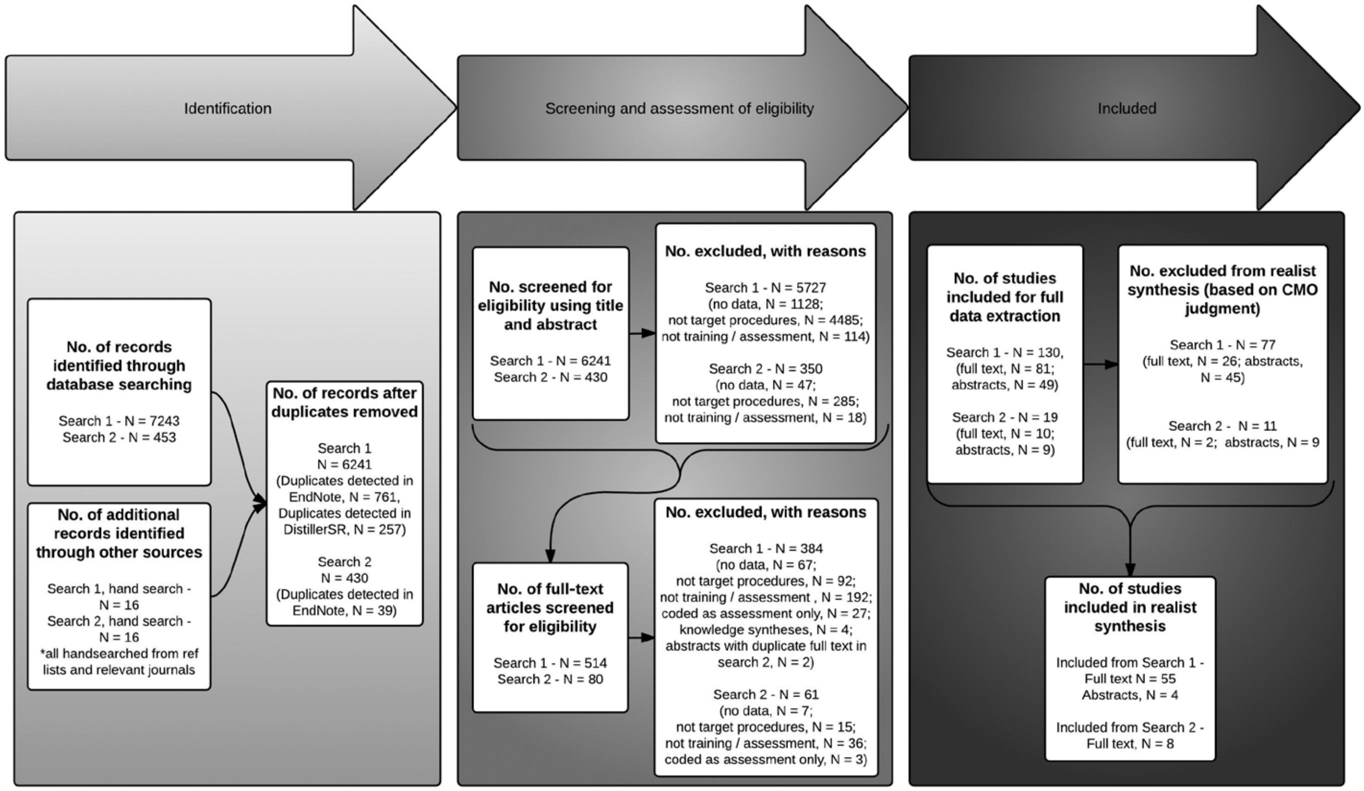

Figure 2 Flow diagram of study inclusion or exclusion, used in a 2015 scoping review and realist synthesis of the literature on invasive bedside procedural skills training in internal medicine. Search 1 was completed on December 13, 2013, and search 2 was completed on April 13, 2015. Abbreviation: CMO indicates context–mechanism–outcome.

theories[70–77]); medical procedural services (MPSs)[78–86]; and multifaceted quality improvement/patient safety (QI/PS) interventions.[18,87–95] These four intervention types involved delivering procedural skills training in variable ways, and even within a single intervention type, studies described heterogeneous approaches to training. We describe each intervention type and the characteristics of the aggregated studies within each type in Chart 1.

Below, we outline the results of our syntheses from step 1 (within each intervention type) and step 2 (across all four intervention types) for context, mechanism, and outcome. We also provide a summary of our syntheses from step 3, which outlines our identified foundational training model.

Synthesis of context themes.

Step 1 (within each intervention type). For all four intervention types, there was notable consistency across the contexts in which procedural skills training was initiated.

Authors used educational technologies, especially simulation, as the training modality in all intervention types, except for QI/PS interventions, which used in-service presentations and workshops grounded in clinical practice. Rationales for using simulation included to capitalize on new educational technologies (see one, do one); to adhere to ABIM recommendations that simulation-based training should precede clinical practice (see one, do one); to move initial or early training away from patients, where harm may occur (see one, do one and educational-theory-informed); to evaluate the impact of educational designs, like competency-based education, on learning outcomes (educational-theory-informed); and to respond to the perceived decline of exposure to procedures during clinical training (see one, do one and MPS). When simulation was used, authors mostly delivered training in simulation centers, with some "just in time, just in place" use of simulation in the clinical setting.

When procedural skills training took place in clinical settings, authors described a need to increase the quality of supervision from staff (MPS), as well as a need to avoid financial penalties—for example, from the Centers for Medicare and Medicaid Services, related to high infection rates (QI/PS).

Step 2 (across all four intervention types). Collectively, authors valued simulation as a safe training option for facilitating increased exposure to the repetitive practice of procedures. They positioned simulation as a precursor to, rather than a replacement of, clinical training and as a modality with which trainees can commit and learn from errors so that adverse events are minimized in clinical practice.

Synthesis of mechanism themes.

Step 1 (within each intervention type). Despite the similarity in contexts, where authors' rationales and study settings overlapped greatly, we found significant differences in how authors' designed procedural training and in how they rationalized the underlying mechanism of training. For most see one, do one studies, authors suggested that novice trainees would experience increased comfort with and exposure to procedures via the active or experiential learning that technology-enhanced learning provides; yet, authors did not cite the use of any educational principles in the design of training interventions. For educational-theory-informed studies, the mechanisms depended on authors' chosen theory; for example, authors studying mastery learning proposed that baseline testing, deliberate practice with feedback, and a final assessment with a minimum passing standard would combine to ensure procedural competency. For most MPS studies, authors proposed that trainees would benefit from experiencing an integrated curriculum combining simulation-based training and clinical exposure on a two- or four-week procedural rotation. For most QI/PS studies, authors emphasized hospital-based rather than educational components, such as administrative and clinical champions (nurses were a common target group) who provided oversight, designed training interventions according to quality improvement principles (e.g., in-service presentations in the workplace), and served as drivers for accountability.

Step 2 (across all four intervention types). Collectively, all studies emphasized active learning during training, yet differed in how such active learning was accomplished: It was assumed in see one, do one studies, designed in educational-theory-informed studies, integrated in MPS studies, and situated in QI/PS studies.

Synthesis of outcome themes.

Step 1 (within each intervention type). For each of the four intervention types, there were significant differences in outcomes. For most interventions in nonclinical, simulation settings, authors assessed performance using individual-level outcomes such as self-reported confidence or direct observation of procedural skills. For most interventions in clinical settings, authors assessed performance using self-reported procedural success and group-level infection or complication rates, rather than direct observation of performance or chart audit. Often, authors did not report favorable validity evidence (e.g., reliability metrics) to demonstrate that patient outcomes were sensitive to the training interventions.

For see one, do one studies, the experimental groups' procedural competence improved from baseline or

Table 1

Characteristics of the Studies Included After Realist Synthesis in a 2015 Scoping Review and Realist Synthesis of the Literature on Invasive Bedside Procedural Skills Training in Internal Medicine

Study characteristic	No. of studies (n = 67)	No. of participants (n = 3,684)
Study design		
Single-group, posttest only	2	289
Single-group pre/posttest	26	1,714
Nonrandomized, multiple groups	18	661
Qualitative	1	11
Randomized controlled trial	16	948
Conference abstract or could not tell	4	61
Geographical location		
United States	55	3,339
Canada	6	201
United Kingdom	1	30
Asia	2	74
Oceania	1	Undefined
Conference abstract or could not tell	2	40
Study setting		
Nonclinical setting (including simulation, computer-based, etc.)	37	1,932
Clinical setting (hospital, field, clinic)	8	219
Both	22	1,533
Procedural skill studied[a]		
Lumbar puncture	23	1,255
Arthrocentesis	6	196
Paracentesis	10	279
Thoracentesis	12	273
Central venous catheterization	43	2,889
Ultrasound guidance used in training?		
Yes	30	1,294
No	36	2,325
Conference abstract or could not tell	1	65
Patient demographics reported?		
Yes	20	1,177
No	25	1,579
Not applicable	22	928
Participants[a]		
Medical students	8	404
Physicians, postgraduate training	47[b]	2,697
Physicians in practice	6	229
Nurses in practice	3	87
Other	1	46
Mix of health care professionals, without a breakdown	13[c]	44
Clinical specialty[a]		
Anesthesia	6	514
Critical care	14	1,230
Emergency medicine	16	950
Internal medicine, general or subspecialty	35	2,259
Family medicine	2	44

(*Table continues*)

Table 1
(Continued)

Study characteristic	No. of studies (n = 67)	No. of participants (n = 3,684)
Lab medicine, pathology, or radiology	2	430
Obstetrics–gynecology	1	242
Pediatrics	17	701
Surgery	9	448
Only specified as a trainee class as captured above	4	164
Could not tell	2	74
Outcome measures[a,d]		
Self-report, reaction, or confidence	27	1,492
Knowledge	23	1,231
Skill, time	4	158
Skill, nontime	25	1,344
Skill, product	3	122
Behavior, time	5	222
Behavior, nontime	27	2,263
Patient outcome	39	2,629
Study quality[e]		
All studies, mean = 12.7 (SD = 2.1)	62[f]	3,612
MERSQI score > 12.7	33	2,680

Abbreviations: SD indicates standard deviation; MERSQI, Medical Education Research Study Quality Instrument.
[a]The number of studies and participants in this category may add up to more than the number for all studies because several studies included more than one procedure or learner group, fit within more than one clinical specialty, or reported multiple outcomes.
[b]Note that of the 47 studies, 8 did not define the sample size of the learner population.
[c]Note that of the 13 studies, 12 did not define the sample size of the learner population.
[d]Note that skills are measured in the nonclinical setting, whereas behaviors are measured in the clinical setting. Nontime refers to skills and behaviors that are assessed based on observable processes (e.g., a Likert scale score assessing technical skill performance), while time refers to skills and behaviors that are assessed based on duration. For more information, see Chart 1. Skill, product refers to outcomes of performance in a nonclinical setting (e.g., successful task completion, or assessment of the quality of the product of task performance).
[e]Study quality was calculated using the MERSQI (total possible score of 18). To facilitate interpretation, note that a previous systematic review of 289 studies on simulation-based training[99] found a mean MERSQI of 12.3 (SD = 1.8).
[f]Five full-text articles did not report sufficient information to calculate a MERSQI score.

as compared with control groups (who either had no training or traditional training). Most educational-theory-informed studies of mastery learning groups found that they outperformed control groups, though two large trials showed that a single mastery session did not improve future lumbar puncture success in pediatric patients.[64,96] For other educational-theory-informed studies, authors applied most educational principles successfully (e.g., group conformity). All MPS groups improved from baseline or outperformed control groups, though authors commented that despite the observed benefits, the MPS was often assigned the most challenging patients, which may have implications for procedural success rates. For QI/PS studies, all showed improved outcomes related to the multifaceted approach, though none could specify which facet (or facets) led to the observed benefit, and none identified education as a key factor.

Step 2 (across all four intervention types). Although most authors labeled their training interventions as successful, our synthesis suggests that this was likely a function of their using weak comparator groups (e.g., nonintervention controls) and outcomes without sufficient evidence supporting their use as sensitive metrics. A clear demiregularity was authors' use of group-level assessments in clinical settings at the expense of individual-level assessments, like direct observation, which were used often in nonclinical, simulation settings.

Step 3: Collective syntheses summary and identified foundational training model. Together, our syntheses suggest a list of key components for future IM procedural skills training—namely, the need to design training that gives trainees the opportunity for active learning in a curriculum that tightly integrates simulation-based training with interventions in the clinical setting. The two intervention types that best aligned with these principles were the MPS and QI/PS. We suggest that studies of MPSs provide the most robust foundation for future procedural skills training curricula, particularly because most QI/PS studies did not describe the educational component of their multifaceted interventions adequately. Our analysis suggests that the MPS model will be adaptable to most institutional settings and can be customized to local settings using the lessons from our syntheses. For example, the MPS model can be adapted to address key context demiregularities by increasing both the volume and variability of training (e.g., training that varies relevant to situational or patient factors). Another adaptation is that the MPS model can be designed according to educational-theory-informed mechanisms to prompt active learning during simulation-based training. Finally, the MPS model can be adapted to include the notable practices identified in QI/PS studies, such as involving nursing and other health professions and identifying champions across clinical specialties. We consider the implications of such adaptations to future MPS interventions and generate related key lines of inquiry below.

Discussion

We synthesized a heterogeneous literature to help stakeholders establish the key components of rigorous, evidence-based training for core invasive bedside procedures in IM. From 67 studies, we identified four intervention types, which we synthesized to identify key considerations for future IM procedural skills training curricula. The observed heterogeneity in how procedural skills training interventions are designed (mechanism) and in how competence is assessed (outcome) suggests that the official expectation that all residents develop competence in the five invasive bedside procedures is likely not fulfilled consistently. Our synthesis suggests that the most robust foundational model would be an adaptable MPS; this finding aligns with recent perspectives on procedural competence.[97] After first

Chart 1

Select Characteristics of the Studies Aggregated Under Each of the Four Intervention Types in a 2015 Scoping Review and Realist Synthesis of the Literature on Invasive Bedside Procedural Skills Training in Internal Medicine

Characteristic	Intervention type: "See one, do one" (n = 22)	Educational-theory-informed (n = 26)	MPS (n = 9)	Multifaceted QI/PS (n = 10)
Description	Interventions essentially replicated the see one, do one approach that is the hallmark of the clinical apprenticeship model. Interventions were structured in a typical sequence: (1) lecture, (2) instructor demonstration, and (3) a short period of hands-on practice; very few ended with a performance-based assessment. Interventions used educational (Web-based [n = 4], simulation-based [n = 15]) or medical technologies (ultrasound guidance [n = 3]). Most author groups did not cite educational evidence or theory as contributing to the instructional design.	Interventions designed using principles from theories in education, including mastery learning (n = 18), self-regulated learning (n = 2), and a variety of cognitive theories (n = 6, e.g., cognitive task analysis [n = 2]). <u>Mastery learning</u>: The key component, according to all authors, was ensuring that all trainees train to a minimum passing standard, though the instructional design differed widely. Most (n = 12) were conducted by author groups from two institutions that adhered to comprehensive definitions of mastery learning. <u>Other theories</u>: Interventions were specific to the cited educational theory. Please refer to each individual study for more information.	All authors defined the MPS as a two- or four-week rotation staffed by physicians from various specialties (e.g., hospitalist, pulmonary, critical care). Eight interventions combined that clinical exposure with an integrated curricular approach that included lectures and simulation-based training with trained instructors.	Interventions combined an educational module with other activities aimed at improving the quality of patient care. Most (n = 8) did not include an adequate description of the educational module. Interventions included two or more components, such as interprofessional collaboration and training, mandatory nurse-led bedside checklists, mandated use of ultrasound guidance, and bundles for infection or complication prevention.
Study design	Designs included two-group nonrandomized (n = 8), single-group pre/posttest (n = 7), RCT (n = 5), qualitative (n = 1), and conference abstract (n = 1).	<u>Mastery learning</u>: Designs included single-group pre/posttest (n = 7), two-group nonrandomized (n = 6), RCT (n = 3), two-group posttest only (n = 1), and conference abstract (n = 1). <u>Other theories</u>: Designs included RCT (n = 7) and two-group posttest only (n = 1).	Designs included single-group pre/posttest (n = 3), two-group nonrandomized (n = 3), RCT (n = 1), and conference abstract (n = 2).	Designs included single-group pre/posttest (n = 9) and two-group nonrandomized (n = 1).
MERSQI study quality score, mean (SD)[a]	12.7 (2.4), n = 20	<u>Mastery learning</u>: 13.2 (1.1), n = 17 <u>Other theories</u>: 13.6 (2.4), n = 8	11.6 (2.4), n = 7	12.1 (2.0), n = 10
Outcome measures[b]	Measures included trainee's self-reported reactions to training or confidence (n = 13), tests of knowledge (n = 12), and patient outcomes (n = 11). Patient outcomes were measured via trainees' self-reported procedural success (n = 4), observed performance or chart audit (n = 4), and group-level infection or complication rates (n = 3).	<u>Mastery learning</u>: Measures included patient outcomes (n = 10), participants' skills on the simulator (n = 10), and participants' behaviors in the clinical context (n = 8). Patient outcomes were measured via trainee's self-reported procedural success (n = 6) and group-level infection or complication rates (n = 4). <u>Other theories</u>: Measures included participants' skills on the simulator (n = 6) and patient outcomes (n = 2), as measured via direct observation (n = 1) and group-level infection or complication rates (n =1).	Measures included patient outcomes (n = 6), participants' behaviors when performing procedures on patients (n = 5), and participants' self-reported reactions to training (n = 4). Patient outcomes were measured using group-level infection or complication rates (n = 4), direct observation using checklists (n = 2), and trainees' self-reported procedural success (n = 1).	Measures included patient outcomes (n = 10) and participants' behaviors when performing procedures on patients (n = 4). Patient outcomes were measured using group-level infection or complication rates (n = 7), trainees' self-reported procedural success (n = 2), and x-ray records (n = 1).

Abbreviations: MPS indicates medical procedure services; QI/PS, quality improvement/patient safety; RCT, randomized controlled trial; MERSQI, Medical Education Research Study Quality Instrument; SD, standard deviation.

[a]Note that a MERSQI score could not be computed for five included full-text articles. To facilitate interpretation, note that a previous systematic review of 289 studies on simulation-based training[99] found a mean MERSQI score (out of a possible 18) of 12.3 (SD = 1.8).

[b]The number of studies in this category may add up to more than the number of all studies because several studies reported multiple outcomes.

comparing our findings with those of previous reviews, we describe and consider the implications of three interrelated lines of inquiry for studying IM invasive bedside procedural skills training in the future.

Comparison with previous research

A recent systematic review on procedural skills training[21] ended with recommendations to use simulation where possible; to use strong research designs like randomized controlled trials, especially when examining differences between instructional methods; and to teach using "competency-based methods such as mastery learning and deliberate practice." While our review supports these suggestions, below we provide more specific recommendations for how simulation and clinical training can be integrated in a systems-based approach that aims to increase the volume and variability of procedural skills training opportunities and that emphasizes assessing evidence-based educational and clinical outcomes. Notably, our findings suggest that not all mastery learning interventions are successful, and thus future research will need to clarify the mechanisms for when mastery learning is successful or not, as well as the mechanisms of other promising instructional designs (e.g., directed self-regulated learning[71]).

Line of inquiry 1

Based on the need to ensure accountability to patients and, by extension, regulatory bodies, as well as the limited training resources, and performance and observation of procedures in clinical practice, researchers need to test whether any adapted *MPS model will be feasible for training all IM trainees.*

Assuming the current level of resources and funding allocated to IM procedural skills training remains static,[98] combined with IM residents' limited clinical exposure to procedures, program directors will likely be challenged to implement any adapted MPS training model. If that assumption holds true, then policy makers may need to make the difficult decision to recommend targeted training of a smaller group of trainees, who have been identified as needing to develop and maintain procedural competence throughout their careers. A reinvestment of resources and training opportunities to smaller groups of trainees would mark a shift from expecting core competence in all trainees to training a *competent core* with a specialization in procedures. In such a system, for example, all IM residents could be expected to achieve cognitive competence (i.e., understand the indications, limitations, contraindications, and complications of procedures), as presently required by the ABIM. Beyond this cognitive competence, though, a proceduralist selection system would need to be implemented, based on trainee interest and a career path requiring procedural competence (e.g., plan to practice IM in community settings or in academic centers with a responsibility for training and assessment), to ensure a core set of clinicians who are procedurally competent. We acknowledge that this proposal would require large-scale changes in the procedure service-delivery models of hospitals that currently rely on all IM residents to perform procedures, as well as a philosophical shift in the professional identity and scope of practice of general internists.

A 2009 study provides a practical example of how programs might use criteria to decide privileges for performing procedures.[90] When pulmonologists working at an outpatient pulmonary clinic learned that they had a higher frequency of iatrogenic pneumothorax compared with a nearby radiology practice, they imposed numerous practice changes including required retraining on thoracentesis skills to competency standards. The clinic did not allow pulmonologists who did not meet the standard to perform thoracentesis on patients.[90] The authors reported a significant decline in pneumothorax rates, which held constant for two years post intervention. This example demonstrates the potential of investing in a core group of trainees, which could be a prudent resource allocation strategy that helps to address the pressing factors of system accountability, patient safety, and the rising costs of clinical errors. Research will be needed to determine whether this approach to training is appropriate for all invasive bedside procedures or whether trainee competence in some procedures might be realistically achieved in core training.

Line of inquiry 2

To build on and optimize implementation of adaptable MPS models, researchers will need to study how best to integrate the instructional designs of educational-theory-informed researchers and the systems-level thinking of quality improvement researchers.

While the MPS studies did use some notable practices of instructional design (i.e., integrating simulation with clinical training[99]), they did not cite or use notable practices from QI/PS interventions (i.e., appointing champions and emphasizing accountability).[100] A shortcoming of many QI/PS studies, however, was that they did not use simulation, which has been shown to be a common component of most procedural skills training interventions.[21] Additionally, we found that authors of educational-theory-informed and QI/PS studies largely responded to different contextual drivers, emphasized different educational mechanisms, and generated different outcome measures, all while pursuing the same goal of ensuring that bedside procedures are performed competently. Hence, we agree with recent calls for a better alignment of efforts between these two research domains and believe that such alignment would produce optimized MPS models.[101] Specifically, educational-theory-informed researchers should include systems-level QI/PS experts as team members in future studies, and hospital-based quality improvement teams should include education experts as members on their committees; both groups should work to align the design, implementation, and evaluation of procedural skills training that integrates the simulation and clinical settings.

Line of inquiry 3

Research is needed to evaluate validity evidence for outcomes measured in the nonclinical, simulation-based and clinical settings. Research generating evidence for relationships between patient and health care system outcomes and more accessible educational outcomes (i.e., educational surrogates) will be particularly important.

A 2013 article calls for research programs that establish evidence for links between outcome measures collected in the nonclinical setting with those collected in the clinical setting.[102] For example, a 2015 meta-analysis examined the relationship between simulation-based assessments and clinical assessments and found that tools requiring raters to observe individual performance directly (e.g., global ratings of a procedure) showed the highest correlations between the two settings.[103] We suggest that the benefit of direct observation might result because assessing at the individual level helps avoid unit-of-analysis errors, which arise when outcomes are measured at a group level (i.e., collapsing infection or complication rates for an entire intensive care unit likely masks multiple data points from high or low performers, reducing the specificity of the measurement). Although there are approaches available to analyze such nested data, like hierarchical generalized linear models,[104] none of the studies included in our review adjusted for such nesting using these techniques.

Researchers will need to collect a wide array of validity evidence to clarify "pathways that link training interventions to patient health outcomes."[105] Rather than using outcomes that are low-hanging fruit and for which there is little validity evidence, such as self-reported procedural success and group infection or complication rates, researchers will need to identify educationally sensitive outcomes in the clinical setting, especially those involving direct observation,[106,107] and establish chains of evidence between outcomes measured in the nonclinical and clinical settings.[102,103] Given the validity evidence supporting the use of global rating scales (with or without checklists) in the simulation setting,[108–110] adapting these scales to the clinical setting is likely a fruitful research direction.

Limitations

The primary literature on IM bedside procedural skills training had several limitations which impacted our review. Authors reported nearly universal success and few failures of their training interventions, which implies that there may be an issue of publication bias of positive studies in our dataset. Some procedures were studied more extensively than others, and nearly all studies emphasized the procedures' technical components and excluded components such as judging whether a procedure needs to be performed, obtaining informed consent, coordinating care, and documenting the procedure.[111] All but one study[90] evaluated how training affects the development of procedural competence rather than the *maintenance of competence*. Although we judged the context–mechanism–outcome linkage independently and in duplicate, our evaluations remain subjective; however, that only 45% of studies met our standard for sufficient information on the context, mechanism, and outcome suggests that there are important gaps in how research on procedural skills training has been conducted and reported. By using a realist synthesis approach, we excluded many studies, some of which might have unearthed additional themes. We did not conduct meta-analyses, particularly because we believe that knowledge synthesis methods supported by qualitative research paradigms, like realist synthesis, provide more targeted answers regarding gaps in research, as well as potential solutions and next steps.

Conclusion

We found that actual practices in procedural skills training in IM are highly variable. Such variability is not surprising considering that regulatory organizations mandate procedural competence, yet do not provide guidelines for program directors to follow when implementing training programs. We have identified the MPS as a foundational training model and provided a list of potential key components that educators can incorporate into future procedural training curricula, which researchers can study and test systematically. In an era where evidence shows that high-quality training translates into high-quality care,[101,104,112] the imperative to design the best educational experience for our trainees has never been stronger.

Acknowledgments: The authors wish to thank Arija Birze and Judy Tran for providing research support, and Laure Perrier and Marina Englesakis for developing and conducting the literature search strategies.

Funding/Support: The authors are thankful for study funding provided by the Canadian Institutes of Health Research via Knowledge Synthesis Grant no. 201304KRS-132055.

Other disclosures: None reported.

Ethical approval: As this study did not involve human subjects, it was exempt from ethical review at all associated institutions.

R. Brydges is assistant professor, Department of Medicine, University of Toronto, and scientist, Wilson Centre, University Health Network, Toronto, Ontario, Canada.

L. Stroud is assistant professor, Department of Medicine, University of Toronto, Toronto, Ontario, Canada.

B.M. Wong is associate professor, Department of Medicine, University of Toronto, Toronto, Ontario, Canada.

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

K. Imrie is immediate past president, Royal College of Physicians and Surgeons of Canada, Ottawa, Ontario, Canada.

R. Hatala is associate professor, Department of Medicine, University of British Columbia, Vancouver, British Columbia, Canada.

References

1 American Board of Internal Medicine. Internal medicine policies. http://www.abim.org/certification/policies/imss/im.aspx. Accessed March 1, 2017.
2 Royal College of Physicians and Surgeons of Canada. Objectives of training in the specialty of general internal medicine. http://www.royalcollege.ca/cs/groups/public/documents/document/y2vk/mdaw/~edisp/tztest3rcpsced000901.pdf. Published 2012. Accessed March 1, 2017.
3 Reynolds MR, Cohen DJ, Kugelmass AD, et al. The frequency and incremental cost of major complications among Medicare beneficiaries receiving implantable cardioverter–defibrillators. J Am Coll Cardiol. 2006;47:2493–2497.
4 Baker GR, Norton PG, Flintoft V, et al. The Canadian Adverse Events Study: The incidence of adverse events among hospital patients in Canada. CMAJ. 2004;170:1678–1686.
5 Brennan TA, Hebert LE, Laird NM, et al. Hospital characteristics associated with adverse events and substandard care. JAMA. 1991;265:3265–3269.
6 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;324:377–384.
7 Neale G, Woloshynowych M, Vincent C. Exploring the causes of adverse events in NHS hospital practice. J R Soc Med. 2001;94:322–330.
8 Thomas EJ, Orav EJ, Brennan TA. Hospital ownership and preventable adverse events. Int J Health Serv. 2000;30:211–219.
9 Wilson RM, Harrison BT, Gibberd RW, Hamilton JD. An analysis of the causes of adverse events from the Quality in Australian Health Care Study. Med J Aust. 1999;170:411–415.
10 de Vries EN, Ramrattan MA, Smorenburg SM, Gouma DJ, Boermeester MA. The incidence and nature of in-hospital adverse events: A systematic review. Qual Saf Health Care. 2008;17:216–223.
11 Promes SB, Chudgar SM, Grochowski CO, et al. Gaps in procedural experience and competency in medical school graduates. Acad Emerg Med. 2009;16(suppl 2):S58–S62.
12 Wolf KS, Dooley-Hash S. Emergency medicine procedures: Examination of trends

in procedures performed by emergency medicine residents. Acad Emerg Med. 2012;19:S171.
13 Ma IW, Teteris E, Roberts JM, Bacchus M. Who is teaching and supervising our junior residents' central venous catheterizations? BMC Med Educ. 2011;11:16.
14 Huang GC, Smith CC, Gordon CE, et al. Beyond the comfort zone: Residents assess their comfort performing inpatient medical procedures. Am J Med. 2006;119:71.e17–71.e24.
15 Pugh C, Plachta S, Auyang E, Pryor A, Hungness E. Outcome measures for surgical simulators: Is the focus on technical skills the best approach? Surgery. 2010;147:646–654.
16 Cohen ER, Feinglass J, Barsuk JH, et al. Cost savings from reduced catheter-related bloodstream infection after simulation-based education for residents in a medical intensive care unit. Simul Healthc. 2010;5:98–102.
17 Sherertz RJ, Ely EW, Westbrook DM, et al. Education of physicians-in-training can decrease the risk for vascular catheter infection. Ann Intern Med. 2000;132: 641–648.
18 Berenholtz SM, Pronovost PJ, Lipsett PA, et al. Eliminating catheter-related bloodstream infections in the intensive care unit. Crit Care Med. 2004;32:2014–2020.
19 Ma IW, Brindle ME, Ronksley PE, Lorenzetti DL, Sauve RS, Ghali WA. Use of simulation-based education to improve outcomes of central venous catheterization: A systematic review and meta-analysis. Acad Med. 2011;86:1137–1147.
20 Madenci AL, Solis CV, de Moya MA. Central venous access by trainees: A systematic review and meta-analysis of the use of simulation to improve success rate on patients. Simul Healthc. 2014;9:7–14.
21 Huang GC, McSparron JI, Balk EM, et al. Procedural instruction in invasive bedside procedures: A systematic review and meta-analysis of effective teaching approaches. BMJ Qual Saf. 2016;25:281–294.
22 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;90:1025–1033.
23 Colquhoun HL, Levac D, O'Brien KK, et al. Scoping reviews: Time for clarity in definition, methods, and reporting. J Clin Epidemiol. 2014;67:1291–1294.
24 Wong G, Greenhalgh T, Westhorp G, Pawson R. Realist methods in medical education research: What are they and what can they contribute? Med Educ. 2012;46:89–96.
25 Arskey H, O'Malley L. Scoping studies: Towards a methodological framework. Int J Soc Res Methodol. 2005;8:19–32.
26 Wong G, Greenhalgh T, Westhorp G, Buckingham J, Pawson R. RAMESES publication standards: Realist syntheses. BMC Med. 2013;11:21.
27 Gordon M, Gibbs T. STORIES statement: Publication standards for healthcare education evidence synthesis. BMC Med. 2014;12:143.
28 Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C. An evidence-based practice guideline for the peer review of electronic search strategies. J Clin Epidemiol. 2009;62:944–952.
29 Reed DA, Cook DA, Beckman TJ, Levine RB, Kern DE, Wright SM. Association between funding and quality of published medical education research. JAMA. 2007;298: 1002–1009.
30 Tricco AC, Soobiah C, Antony J, et al. A scoping review identifies multiple emerging knowledge synthesis methods, but few studies operationalize the method. J Clin Epidemiol. 2016;73:19–28.
31 McCormack B, Rycroft-Malone J, Decorby K, et al. A realist review of interventions and strategies to promote evidence-informed healthcare: A focus on change agency. Implement Sci. 2013;8:107.
32 Britt RC, Novosel TJ, Britt LD, Sullivan M. The impact of central line simulation before the ICU experience. Am J Surg. 2009;197:533–536.
33 Chenkin J, Lee S, Huynh T, Bandiera G. Procedures can be learned on the Web: A randomized study of ultrasound-guided vascular access training. Acad Emerg Med. 2008;15:949–954.
34 Miller AH, Roth BA, Mills TJ, Woody JR, Longmoor CE, Foster B. Ultrasound guidance versus the landmark technique for the placement of central venous catheters in the emergency department. Acad Emerg Med. 2002;9:800–805.
35 Sekiguchi H, Tokita JE, Minami T, Eisen LA, Mayo PH, Narasimhan M. A prerotational, simulation-based workshop improves the safety of central venous catheter insertion: Results of a successful internal medicine house staff training program. Chest. 2011;140:652–658.
36 Froehlich CD, Rigby MR, Rosenberg ES, et al. Ultrasound-guided central venous catheter placement decreases complications and decreases placement attempts compared with the landmark technique in patients in a pediatric intensive care unit. Crit Care Med. 2009;37:1090–1096.
37 Gaies MG, Morris SA, Hafler JP, et al. Reforming procedural skills training for pediatric residents: A randomized, interventional trial. Pediatrics. 2009;124:610–619.
38 Griswold-Theodorson S, Hannan H, Handly N, et al. Improving patient safety with ultrasonography guidance during internal jugular central venous catheter placement by novice practitioners. Simul Healthc. 2009;4:212–216.
39 Grover S, Currier PF, Elinoff JM, Katz JT, McMahon GT. Improving residents' knowledge of arterial and central line placement with a Web-based curriculum. J Grad Med Educ. 2010;2:548–554.
40 Kamdar G, Kessler DO, Tilt L, et al. Qualitative evaluation of just-in-time simulation-based learning: The learners' perspective. Simul Healthc. 2013;8:43–48.
41 Kilbane BJ, Adler MD, Trainor JL. Pediatric residents' ability to perform a lumbar puncture: Evaluation of an educational intervention. Pediatr Emerg Care. 2010;26:558–562.
42 Hasley P, Preisner R, Anish E, Bulova P, Collin T, Kim Y. Is doing superior to knowing? Simulation training improves junior faculty confidence to teach joint aspiration and injection. J Gen Intern Med. 2010;25:S447–S448.
43 Lenchus J, Issenberg SB, Murphy D, et al. A blended approach to invasive bedside procedural instruction. Med Teach. 2011;33:116–123.
44 Ma IW, Chapelsky S, Bhavsar S, et al. Procedural certification program: Enhancing resident procedural teaching skills. Med Teach. 2013;35:524.
45 Miranda JA, Trick WE, Evans AT, Charles-Damte M, Reilly BM, Clarke P. Firm-based trial to improve central venous catheter insertion practices. J Hosp Med. 2007;2:135–142.
46 Smith CC, Huang GC, Newman LR, et al. Simulation training and its effect on long-term resident performance in central venous catheterization. Simul Healthc. 2010;5:146–151.
47 Srivastava G, Roddy M, Langsam D, Agrawal D. An educational video improves technique in performance of pediatric lumbar punctures. Pediatr Emerg Care. 2012;28:12–16.
48 Thomas SM, Burch W, Kuehnle SE, Flood RG, Scalzo AJ, Gerard JM. Simulation training for pediatric residents on central venous catheter placement: A pilot study. Pediatr Crit Care Med. 2013;14:e416–e423.
49 Vogelgesang SA, Karplus TM, Kreiter CD. An instructional program to facilitate teaching joint/soft-tissue injection and aspiration. J Gen Intern Med. 2002;17:441–445.
50 Wayne DB, Cohen ER, Singer BD, et al. Progress toward improving medical school graduates' skills via a "boot camp" curriculum. Simul Healthc. 2014;9:33–39.
51 White ML, Jones R, Zinkan L, Tofil NM. Transfer of simulated lumbar puncture training to the clinical setting. Pediatr Emerg Care. 2012;28:1009–1012.
52 Xiao Y, Seagull FJ, Bochicchio GV, et al. Video-based training increases sterile-technique compliance during central venous catheter insertion. Crit Care Med. 2007;35:1302–1306.
53 Barsuk JH, Cohen ER, Caprio T, McGaghie WC, Simuni T, Wayne DB. Simulation-based education with mastery learning improves residents' lumbar puncture skills. Neurology. 2012;79:132–137.
54 Barsuk JH, Cohen ER, Feinglass J, McGaghie WC, Wayne DB. Use of simulation-based education to reduce catheter-related bloodstream infections. Arch Intern Med. 2009;169:1420–1423.
55 Barsuk JH, Cohen ER, McGaghie WC, Wayne DB. Long-term retention of central venous catheter insertion skills after simulation-based mastery learning. Acad Med. 2010;85(10 suppl):S9–S12.
56 Barsuk JH, Cohen ER, Vozenilek JA, O'Connor LM, McGaghie WC, Wayne DB. Simulation-based education with mastery learning improves paracentesis skills. J Grad Med Educ. 2012;4:23–27.
57 Barsuk JH, McGaghie WC, Cohen ER, O'Leary KJ, Wayne DB. Simulation-based mastery learning reduces complications during central venous catheter insertion in a medical intensive care unit. Crit Care Med. 2009;37:2697–2701.
58 Cohen ER, Barsuk JH, Moazed F, et al. Making July safer: Simulation-based mastery learning during intern boot camp. Acad Med. 2013;88:233–239.
59 Conroy SM, Bond WF, Pheasant KS, Ceccacci N. Competence and retention in performance of the lumbar puncture

procedure in a task trainer model. Simul Healthc. 2010;5:133–138.
60 Diederich E, Rigler S, Mahnken J, Dong L, Williamson T, Sharpe M. The effect of model fidelity on learning outcomes of a simulation-based education program for central venous catheter insertion. Chest. 2012;142(4 suppl):535A.
61 Dodge KL, Lynch CA, Moore CL, Biroscak BJ, Evans LV. Use of ultrasound guidance improves central venous catheter insertion success rates among junior residents. J Ultrasound Med. 2012;31:1519–1526.
62 Evans LV, Dodge KL, Shah TD, et al. Simulation training in central venous catheter insertion: Improved performance in clinical practice. Acad Med. 2010;85:1462–1469.
63 Jiang G, Chen H, Wang S, et al. Learning curves and long-term outcome of simulation-based thoracentesis training for medical students. BMC Med Educ. 2011;11:39.
64 Kessler DO, Arteaga G, Ching K, et al. Interns' success with clinical procedures in infants after simulation training. Pediatrics. 2013;131:e811–e820.
65 Kessler DO, Auerbach M, Pusic M, Tunik MG, Foltin JC. A randomized trial of simulation-based deliberate practice for infant lumbar puncture skills. Simul Healthc. 2011;6:197–203.
66 Taitz J, Wyeth B, Lennon R, et al. Effect of the introduction of a lumbar-puncture sticker and teaching manikin on junior staff documentation and performance of paediatric lumbar punctures. Qual Saf Health Care. 2006;15:325–328.
67 Wayne DB, Barsuk JH, O'Leary KJ, Fudala MJ, McGaghie WC. Mastery learning of thoracentesis skills by internal medicine residents using simulation technology and deliberate practice. J Hosp Med. 2008;3:48–54.
68 Kessler D, Pusic M, Chang TP, et al; INSPIRE LP Investigators. Impact of just-in-time and just-in-place simulation on intern success with infant lumbar puncture. Pediatrics. 2015;135:e1237–e1246.
69 Barsuk JH, Cohen ER, Potts S, et al. Dissemination of a simulation-based mastery learning intervention reduces central line-associated bloodstream infections. BMJ Qual Saf. 2014;23:749–756.
70 Beran TN, McLaughlin K, Al Ansari A, Kassam A. Conformity of behaviors among medical students: Impact on performance of knee arthrocentesis in simulation. Adv Health Sci Educ Theory Pract. 2013;18:589–596.
71 Brydges R, Nair P, Ma I, Shanks D, Hatala R. Directed self-regulated learning versus instructor-regulated learning in simulation training. Med Educ. 2012;46:648–656.
72 Duncan JR, Henderson K, Street M, et al. Creating and evaluating a data-driven curriculum for central venous catheter placement. J Grad Med Educ. 2010;2: 389–397.
73 Murphy MA, Neequaye S, Kreckler S, Hands LJ. Should we train the trainers? Results of a randomized trial. J Am Coll Surg. 2008;207:185–190.
74 Shanks D, Brydges R, den Brok W, Nair P, Hatala R. Are two heads better than one? Comparing dyad and self-regulated learning in simulation training. Med Educ. 2013;47:1215–1222.
75 Velmahos GC, Toutouzas KG, Sillin LF, et al. Cognitive task analysis for teaching technical skills in an inanimate surgical skills laboratory. Am J Surg. 2004;187:114–119.
76 Craft C, Feldon DF, Brown EA. Instructional design affects the efficacy of simulation-based training in central venous catheterization. Am J Surg. 2014;207:782–789.
77 Chan A, Singh S, Dubrowski A, et al. Part versus whole: A randomized trial of central venous catheterization education. Adv Health Sci Educ Theory Pract. 2015;20:1061–1071.
78 Chang W, Popa A, DeKorte M. A medical invasive procedure service and resident procedure training elective. J Hosp Med. 2012;7:S120.
79 Lenchus J, De Moraes AG, Garg M, Kalidindi V, Soto A, Pavon A. Impact of a standardized curriculum on reducing thoracentesis-induced pneumothorax. J Hosp Med. 2011;6:S69.
80 Lenchus JD. End of the "see one, do one, teach one" era: The next generation of invasive bedside procedural instruction. J Am Osteopath Assoc. 2010;110:340–346.
81 Lenhard A, Moallem M, Marrie RA, Becker J, Garland A. An intervention to improve procedure education for internal medicine residents. J Gen Intern Med. 2008;23: 288–293.
82 Mourad M, Ranji S, Sliwka D. A randomized controlled trial of the impact of a teaching procedure service on the training of internal medicine residents. J Grad Med Educ. 2012;4:170–175.
83 Ramakrishna G, Higano ST, McDonald FS, Schultz HJ. A curricular initiative for internal medicine residents to enhance proficiency in internal jugular central venous line placement. Mayo Clin Proc. 2005;80:212–218.
84 Tolbert T, Haines L, Dickman E, MacArthur L, Terentiev V, Likourezos A. Central venous catheter location changes and complication rates after the institution of an emergency ultrasonography division. Ann Emerg Med. 2012;1:S7.
85 Lenchus JD, Carvalho CM, Ferreri K, et al. Filling the void: Defining invasive bedside procedural competency for internal medicine residents. J Grad Med Educ. 2013;5:605–612.
86 Tukey MH, Wiener RS. The impact of a medical procedure service on patient safety, procedure quality and resident training opportunities. J Gen Intern Med. 2014;29:485–490.
87 Burden AR, Torjman MC, Dy GE, et al. Prevention of central venous catheter-related bloodstream infections: Is it time to add simulation training to the prevention bundle? J Clin Anesth. 2012;24:555–560.
88 Cherry RA, West CE, Hamilton MC, Rafferty CM, Hollenbeak CS, Caputo GM. Reduction of central venous catheter associated blood stream infections following implementation of a resident oversight and credentialing policy. Patient Saf Surg. 2011;5:15.
89 Costello J, Livett M, Stride PJ, West M, Premaratne M, Thacker D. The seamless transition from student to intern: From theory to practice. Intern Med J. 2010;40:728–731.
90 Duncan DR, Morgenthaler TI, Ryu JH, Daniels CE. Reducing iatrogenic risk in thoracentesis: Establishing best practice via experiential training in a zero-risk environment. Chest. 2009;135:1315–1320.
91 McMullan C, Propper G, Schuhmacher C, et al. A multidisciplinary approach to reduce central line-associated bloodstream infections. Jt Comm J Qual Patient Saf. 2013;39:61–69.
92 McKee C, Berkowitz I, Cosgrove SE, et al. Reduction of catheter-associated bloodstream infections in pediatric patients: Experimentation and reality. Pediatr Crit Care Med. 2008;9:40–46.
93 Coopersmith CM, Rebmann TL, Zack JE, et al. Effect of an education program on decreasing catheter-related bloodstream infections in the surgical intensive care unit. Crit Care Med. 2002;30:59–64.
94 Wall RJ, Ely EW, Elasy TA, et al. Using real time process measurements to reduce catheter related bloodstream infections in the intensive care unit. Qual Saf Health Care. 2005;14:295–302.
95 See KC, Jamil K, Chua AP, Phua J, Khoo KL, Lim TK. Effect of a pleural checklist on patient safety in the ultrasound era. Respirology. 2013;18:534–539.
96 Kessler DO, Fein D, Chang TP, et al. Impact of a simulator based just-in-time refresher training for interns on their clinical success rate with infant lumbar puncture. Pediatr Emerg Care. 2011;27:1002.
97 Vaisman A, Cram P. Procedural competence among faculty in academic health centers: Challenges and future directions. Acad Med. 2017;92:31–34.
98 Mullan F, Salsberg E, Weider K. Why a GME squeeze is unlikely. N Engl J Med. 2015;373:2397–2399.
99 Cook DA, Hamstra SJ, Brydges R, et al. Comparative effectiveness of instructional design features in simulation-based education: Systematic review and meta-analysis. Med Teach. 2013;35:e867–e898.
100 Pronovost P, Needham D, Berenholtz S, et al. An intervention to decrease catheter-related bloodstream infections in the ICU. N Engl J Med. 2006;355:2725–2732.
101 Wong BM, Holmboe ES. Transforming the academic faculty perspective in graduate medical education to better align educational and clinical outcomes. Acad Med. 2016;91:473–479.
102 Cook DA, West CP. Perspective: Reconsidering the focus on "outcomes research" in medical education: A cautionary note. Acad Med. 2013;88:162–167.
103 Brydges R, Hatala R, Zendejas B, Erwin PJ, Cook DA. Linking simulation-based educational assessments and patient-related outcomes: A systematic review and meta-analysis. Acad Med. 2015;90:246–256.
104 Bansal N, Simmons KD, Epstein AJ, Morris JB, Kelz RR. Using patient outcomes to evaluate general surgery residency program performance. JAMA Surg. 2015;151:111–119.
105 Kalet AL, Gillespie CC, Schwartz MD, et al. New measures to establish the evidence base for medical education: Identifying educationally sensitive patient outcomes. Acad Med. 2010;85:844–851.
106 Kogan JR, Holmboe ES, Hauer KE. Tools for direct observation and assessment of clinical skills of medical trainees: A systematic review. JAMA. 2009;302:1316–1326.
107 Carraccio C, Englander R, Holmboe ES, Kogan JR. Driving care quality: Aligning

trainee assessment and supervision through practical application of entrustable professional activities, competencies, and milestones. Acad Med. 2016;91:199–203.

108 Walzak A, Bacchus M, Schaefer JP, et al. Diagnosing technical competence in six bedside procedures: Comparing checklists and a global rating scale in the assessment of resident performance. Acad Med. 2015;90:1100–1108.

109 Ilgen JS, Ma IW, Hatala R, Cook DA. A systematic review of validity evidence for checklists versus global rating scales in simulation-based assessment. Med Educ. 2015;49:161–173.

110 Hatala R, Cook DA, Brydges R, Hawkins R. Constructing a validity argument for the Objective Structured Assessment of Technical Skills (OSATS): A systematic review of validity evidence. Adv Health Sci Educ Theory Pract. 2015;20: 1149–1175.

111 MacMillan TE, Wu RC, Morra D. Quality of bedside procedures performed on general internal medicine in-patients: Can we do better? Can J Gen Intern Med. 2014;9:17–20.

112 Asch DA, Nicholson S, Srinivas S, Herrin J, Epstein AJ. Evaluating obstetrical residency programs using patient outcomes. JAMA. 2009;302:1277–1283.