A critical review of simulation-based medical education research: 2003–2009

William C McGaghie,[1] S Barry Issenberg,[2] Emil R Petrusa[3] & Ross J Scalese[2]

OBJECTIVES This article reviews and critically evaluates historical and contemporary research on simulation-based medical education (SBME). It also presents and discusses 12 features and best practices of SBME that teachers should know in order to use medical simulation technology to maximum educational benefit.

METHODS This qualitative synthesis of SBME research and scholarship was carried out in two stages. Firstly, we summarised the results of three SBME research reviews covering the years 1969–2003. Secondly, we performed a selective, critical review of SBME research and scholarship published during 2003–2009.

RESULTS The historical and contemporary research synthesis is reported to inform the medical education community about 12 features and best practices of SBME: (i) feedback; (ii) deliberate practice; (iii) curriculum integration; (iv) outcome measurement; (v) simulation fidelity; (vi) skill acquisition and maintenance; (vii) mastery learning; (viii) transfer to practice; (ix) team training; (x) high-stakes testing; (xi) instructor training, and (xii) educational and professional context. Each of these is discussed in the light of available evidence. The scientific quality of contemporary SBME research is much improved compared with the historical record.

CONCLUSIONS Development of and research into SBME have grown and matured over the past 40 years on substantive and methodological grounds. We believe the impact and educational utility of SBME are likely to increase in the future. More thematic programmes of research are needed. Simulation-based medical education is a complex service intervention that needs to be planned and practised with attention to organisational contexts.

*Medical Education 2010: **44**: 50–63*
doi:10.1111/j.1365-2923.2009.03547.x

[1]Augusta Webster, MD, Office of Medical Education and Faculty Development, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA
[2]Gordon Center for Research in Medical Education, Miller School of Medicine, University of Miami, Miami, Florida, USA
[3]Office for Teaching and Learning in Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA

Correspondence: William C McGaghie, Augusta Webster, MD, Office of Medical Education and Faculty Development, Northwestern University Feinberg School of Medicine, 1-003 Ward Building, 303 East Chicago Avenue, Chicago, Illinois 60611-3008, USA. Tel: 00 1 312 503 0174; Fax: 00 1 312 503 0840; E-mail: wcmc@northwestern.edu

INTRODUCTION

Medical education technology shapes and channels medical education policy as research advancements inform new ways to educate and evaluate doctors. Educational technology and policy coalesce with emphases on effectiveness, efficiency, and trainee and teacher morale as new models of medical teaching and testing are introduced. This is true of simulation-based medical education (SBME) in 2010.

The educational legacy of SBME originates at least from 17th century France, where birthing manikins were used,[1] is referred to in the Flexner Report[2] of the early 20th century, which also cites the use of obstetric manikins, and extends to the widespread contemporary use of simulation for teaching and testing doctors and many other health care professionals.[3] Simulation technology is now a central thread in the fabric of medical education.

This article has two goals. The first is to summarise results from three reviews published in 1999,[4] 2005[5] and 2006[6] on SBME covering research reports published from 1969 to 2003. The second is to combine and reflect critically on selected SBME research and scholarship published from 2003 to 2009. The review and reflection prompted us to identify and discuss a set of 12 features and best practices of SBME. We conclude with suggestions about how the features and best practices can be implemented in medical education.

METHODS

This is a qualitative synthesis of SBME research that spans four decades. We begin by distilling the results of three SBME research reviews carried out by our group and covering the 35 years from 1969 to 2003.[4–6] We proceed to review critically selected research and scholarship on SBME features and operational rules that address key educational goals published from 2003 to 2009. This 6-year review is deliberately selective and critical, rather than exhaustive. It relies on Norman and Eva's 'critical review' approach to literature synthesis,[7,8] combined with the 'realist review' approach espoused by Pawson *et al.*[9,10] Eva argues: 'A good educational research literature review … is one that presents a critical synthesis of a variety of literatures, identifies knowledge that is well established, highlights gaps in understanding, and provides some guidance regarding what remains to be understood. The result should give a new perspective of an old problem… The author … should feel bound by a moral code to try to represent the literature (and the various perspectives therein) fairly, but need not adopt a guise of absolute systematicity.'[8] Pawson *et al.* agree by stating: '…the review question must be carefully articulated so as to prioritise which aspects of which interventions will be examined.'[9]

Consequently, this work relies on our group's judgements about recent SBME research quality and utility to spotlight key features of SBME that we believe have power to advance the field. The result is a focused set of 12 features and best practices, which every SBME teacher should know in order to use medical simulation to maximum educational benefit. We conclude by calling for thematic, sustained and cumulative programmes of SBME research.

RESULTS

Historical perspective

Table 1 presents a narrative summary of three reviews completed by our group involving SBME research reports published during a 35-year time span (1969–2003). The reviews highlight the features of medical simulation and mechanisms for its implementation and use that lead to effective learning. Thus the three reviews address a much broader and more important educational issue than simple comparisons of SBME outcomes with results produced by other instructional methods. Another observation from these historical summaries is that much of the early SBME research lacks methodological rigour. This makes it difficult to reach firm conclusions about aggregate research outcomes and to identify SBME best practices. Despite such flaws, the evidence reported in Table 1 represents a comprehensive summary of SBME research up to 2003.

Given this historical view, we now address SBME research and scholarship published during a recent 6-year time span, 2003–2009. The methodological quality and rigour of research published in this period is much improved. The new research, combined with the historical record, allows us to identify and discuss 12 features and best practices of SBME that medical educators should know and use. The features and best practices are listed in an order that starts with the five items named in one of the earlier historical reviews,[5] followed by the seven that are evident from recent research and practice. These are:

Table 1 Summary of three simulation-based medical education (SBME) research reviews: 1969–2003

Issenberg *et al.* (1999)[4] (selective, narrative review)	Issenberg *et al.* (2005)[5] (qualitative, systematic review spanning 35 years, included 670 peer-reviewed journal articles)	McGaghie *et al.* (2006)[6] (quantitative synthesis of a subset of 31 journal articles referring to 32 research studies drawn from the previous qualitative, systematic review [Issenberg *et al.* 2005[5]])
1 Simulation technology provides a means for learners to engage in acquisition and practice of clinical skills without using live patients 2 Simulation users should weigh the benefits and costs of the technology 3 'The key element in the successful use of simulators is that they become integrated throughout the entire curriculum so that deliberate practice to acquire expertise over time is possible' 4 Simulation allows learners to receive 'professional feedback with opportunities for repetition and feedback'	'The weight of the best available evidence suggests that high-fidelity medical simulations facilitate learning under the right conditions. These include the following' (in order of importance): 1 Feedback is provided during the learning experience 2 Learners engage in repetitive practice 3 The simulator is integrated into an overall curriculum 4 Learners practise with increasing levels of difficulty 5 The simulator is adaptable to multiple learning strategies 6 The simulator captures clinical variation 7 The simulator is embedded in a controlled environment 8 The simulator permits individualised learning 9 Learning outcomes are clearly defined and measured 10 The simulator is a valid (high-fidelity) approximation of clinical practice	'Two principal findings emerge from this study' 1 'The evidence is clear … that repetitive practice involving medical simulations is associated with improved learner outcomes. Simulation-based practice in medical education appears to approximate a dose–response relationship in terms of achieving desired outcomes: more practice yields better results' 2 'Few published journal articles on the effectiveness of high-fidelity simulations in medical education have been performed with enough quality and rigour to yield useful results. Only 5% of research publications in this field (31/670) meet or exceed the minimum quality standards used for this study'

1 feedback;
2 deliberate practice;
3 curriculum integration;
4 outcome measurement;
5 simulation fidelity;
6 skill acquisition and maintenance;
7 mastery learning;
8 transfer to practice;
9 team training;
10 high-stakes testing;
11 instructor training, and
12 educational and professional context.

The list of features and best practices does *not* include interdisciplinary education because the research foundation for this activity is not yet well developed.[11] Detailed research agenda setting for SBME is not listed because that topic has been addressed elsewhere recently.[12] The simulation features and research-based best practices are presented in Table 2, along with gaps in understanding that warrant more study.

Twelve features and best practices

Feedback

In the historical review, feedback is the most important and frequently cited variable about the use of SBME to promote effective learning.[5] Contemporary research amplifies the importance of educational feedback to shape learning by isolating

Table 2 Medical simulation features and best practices, and gaps in understanding

Simulation features	Well-established knowledge, 'best practices'	Gaps in understanding
1 Feedback	Essential role in SBME Core elements: varieties, sources, impact Team debriefing	What model of feedback? What dose of feedback? How to gauge quality of feedback? Feedback adaptation to educational goal
2 Deliberate practice	Essential role in SBME Learner-centred Apparent dose–response relationship	Verify dose–response relationship Verify value of distributed practice versus massed practice
3 Curriculum integration	Integrate with other learning events Focus on educational objectives SBME *complements* clinical education	What is the best mix of learning modalities? How and when to best integrate with other modalities?
4 Outcome measurement	Reliable data → valid actions: feedback, personnel decisions, research inferences Methods: observer ratings, trainee responses (selected, constructed), haptics	Historical problem Narrow bandwidth versus complex professional practice Multiple measures: convergence–divergence, method variance, generalisability analyses
5 Simulation fidelity	Goals–tools match Multi-modal simulation uses manikins, task trainers, and SPs Attention to learning context	How much fidelity is enough or too much? Conditions of training: target outcomes, timeframe, resources How does trainee readiness shape simulation use?
6 Skill acquisition and maintenance	Procedural, professional, cognitive and group skills Maintenance versus decay Aptitude and readiness: cognitive, proprioceptive	What are the mechanism(s) of skill maintenance? Determine conditions of skill decay: person, context, tasks
7 Mastery learning	Rigorous approach to competency-based education All learners master educational goals at a high achievement level with little or no outcome variation Time needed for learning varies	What are the sources of variation in time to mastery standard: cognitive aptitude, motor skill, professional experience? Level of resources needed Is mastery case-specific or generalisable?
8 Transfer to practice	Highest level of Kirkpatrick hierarchy Stretch measurement endpoint from simulation lab to hospital or clinic Translational science	Pathway: simulation laboratory → health care clinic Cascaded inference Study designs: difficult to formulate and execute
9 Team training	Patient care [can be] a 'team sport' Health care team training principles are evidence-based	Determine approaches for clinical team composition and assembly Team skill maintenance Are team members interchangeable?
10 High-stakes testing	Research advances drive new test applications Highly reliable data → valid decisions	Creation and pilot studies of test mechanisms Just because we can, should we?
11 Instructor training	Effective SBME is not easy or intuitive Clinical experience is *not* a proxy for simulation instructor effectiveness Instructor and learner need not be from the same health care profession	Should simulation instructors be certified for various devices? What are appropriate mastery learning models for simulation instructors? Specific to simulation or general teaching skills?

Table 2 (Continued)

Simulation features	Well-established knowledge, 'best practices'	Gaps in understanding
12 Educational and professional context	Context authenticity is critical for SBME teaching and evaluation Context is changing, adaptive	How to break down barriers and overcome inertia? Reinforcement of SBME outcomes in professional contexts What is the effect of local context for success of SBME interventions? How to acknowledge cultural differences among the health care professions?

SBME = simulation-based medical education; SP = simulated patient

three of its core elements: varieties; sources, and impact.

There are two broad varieties of performance feedback: formative and summative. Most SBME feedback or debriefing is formative because its purpose is to improve trainee clinical performance rather than to present summative judgements (e.g. pass, fail). A recent example of debriefing as formative assessment in a medical simulation setting is the four-step model presented by Rudolph *et al.*: 'The steps are to: (i) note salient performance gaps related to predetermined objectives; (ii) provide feedback describing the gap; (iii) investigate the basis for the gap by exploring the frames and emotions contributing to the current performance level, and (iv) help close the performance gap through discussion or targeted instruction about principles and skills relevant to performance.'[13] The four-step model has a long empirical and experiential history. It is grounded in 'evidence and theory from education research, the social and cognitive sciences, experience drawn from conducting over 3000 debriefings, and teaching debriefing to approximately 1000 clinicians worldwide'.[13]

Another recent example addressing varieties of feedback in medical education appears in a discussion about debriefing medical teams. Salas *et al.*[14] present 12 evidence-based best practices and tips for team debriefing for use after critical incidents or recurring clinical events. The 12 debriefing best practices are directly applicable to giving medical trainees feedback in the SBME context. Salas *et al.*[14] list their evidence-based best practices as follows:

1 Debriefs must be diagnostic.
2 Ensure that the organisation creates a supportive learning environment for debriefs.
3 Encourage team leaders and team members to be attentive of teamwork processes during performance episodes.
4 Educate team leaders on the art and science of leading team debriefs.
5 Ensure that team members feel comfortable during debriefs.
6 Focus on a few critical performance issues during the debriefing process.
7 Describe specific teamwork interactions and processes that were involved in the team's performance.
8 Support feedback with objective indicators of performance.
9 Provide outcome feedback later and less frequently than process feedback.
10 Provide both individual and team-oriented feedback, but know when each is most appropriate.
11 Shorten the delay between task performance and feedback as much as possible.
12 Record conclusions made and goals set during the debrief to facilitate feedback during future debriefs.[14]

Using a sample or all 12 of Salas *et al.*'s[14] best practices is likely to boost the quality and utility of trainee feedback in SBME. These ideas are reinforced in scholarly argument by van de Ridder *et al.*[15]

Fanning and Gaba also address the role of debriefing in simulation-based learning.[16] Their essay points out that feedback in debriefing sessions can come from several potential sources, including a trained facilitator, the simulation device (e.g. a manikin), and video or digital recordings. Each feedback source has strengths and limits and thus their use in combination is likely to yield greater educational results.

The impact of feedback in SBME has been addressed by several research groups. An Australian research group, Domuracki *et al.*,[17] studied medical student learning of cricoid pressure during positive pressure ventilation cardiopulmonary resuscitation (CPR) and during anaesthesia with patients at risk of regurgitation. In a randomised trial, medical students and nursing staff received cricoid pressure simulator training with or without force feedback. Research outcomes show that simulation training with force feedback produced significantly better student performance than the no feedback strategy. These results transferred directly to the clinical setting. In the USA, Edelson *et al.*[18] studied the impact of feedback about in-hospital CPR performance using a novel protocol, resuscitation with actual performance integrated debriefing (RAPID), enhanced by objective data from a CPR-sensing and feedback-enabled defibrillator. The CPR performance of simulator-trained residents was compared with the performance of a historical resident cohort. The simulator-trained group displayed significantly better CPR performance than the historical cohort on a variety of clinically meaningful measures (e.g. return of spontaneous circulation). In these illustrations, SBME with potent feedback has a clear impact on trainee clinical behaviour.

Despite this evidence, several questions remain regarding specific feedback methods. What model and dose of feedback are needed for a particular outcome? Do some methods prove more efficient, require fewer resources and yield longer-lasting effects? Feedback standards and guidelines need to be developed so that instructor competence can be measured for this critical SBME skill.

Deliberate practice

Deliberate practice (DP) is an important property of powerful[19] SBME interventions used to shape, refine and maintain trainee knowledge, skills and attitudes. Deliberate practice is very demanding of learners. Originated by psychologist K Anders Ericsson, the DP model is grounded in information processing and behavioural theories of skill acquisition and maintenance.[20] Deliberate practice has at least nine features or requirements when used to achieve medical education goals.[21] It relies on:

1 highly motivated learners with good concentration (e.g. medical trainees);
2 engagement with a well-defined learning objective or task, at an
3 appropriate level of difficulty, with
4 focused, repetitive practice, that leads to
5 rigorous, precise measurements, that yield
6 informative feedback from educational sources (e.g. simulators, teachers), and where
7 trainees also monitor their learning experiences and correct strategies, errors and levels of understanding, engage in more DP, and continue with
8 evaluation to reach a mastery standard, and then
9 advance to another task or unit.

Research that documents the power of DP-based educational interventions is available from the quantitative review cited earlier[6] and from original research on skill acquisition among medical learners in advanced cardiac life support (ACLS),[22,23] thoracentesis[24] and catheter insertion.[25,26]

The value of DP as an educational variable was noted by internists Richard Cabot and Edwin Locke more than a century ago, in 1905.[27] These medical educators were prescient in the observation: 'Learning medicine is not fundamentally different from learning anything else. If one had 100 hours in which to learn to ride a horse or speak in public, one might profitably spend perhaps an hour (in divided doses) in being told how to do it, 4 hours in watching a teacher do it, and the remaining 95 hours in practice, at first with close supervision, later under general oversight.'

Questions still remain about differences between distributed DP over a long time span versus massed DP during a short time period. This has important implications for the integration and implementation of SBME into existing curricula and training programmes.

Curriculum integration

A third principle of sound SBME is that simulated events and simulator practice should be curriculum features that are carefully integrated with other educational events, including clinical experience, lectures, reading, laboratory work, problem-based learning (PBL) and many others. This means that SBME education and evaluation events must be planned, scheduled, required and carried out thoughtfully in the context of a wider medical curriculum. Simulation-based medical education is one of many educational approaches that is used most powerfully and effectively to achieve learning objectives in concert with other educational methods. It *complements* clinical education but cannot substitute for training grounded in patient care in real clinical settings.[3,28] This is reinforced by Kneebone's argument that '[education in] procedural skills should not be divorced from their clinical context and that oversimplification of a complex process can interfere with deep understanding'.[29]

Inertia and organisational barriers can hinder SBME curriculum integration. For example, trainee scheduling is a common problem. The pressure of clinical duties, overwork, ingrained habits and perceptions that SBME is less valuable than clinical experience can sabotage scheduled training sessions, reduce SBME practice time, and deliver a less powerful educational 'dose' than intended. This is manifest in empirical SBME research studies as treatment-by-occasion statistical interactions where intended outcomes are delayed and weaker than expected.[30,31]

There are practical issues of concern such as establishing the best approach to integrate SBME into existing curricula and the impact of this introduction on faculty and administrative resources. Research should also address the impact of combining SBME with other educational models, such as using simulations as the clinical trigger and context for PBL cases.

Outcome measurement

Outcome measurement that yields reliable data is essential to SBME and all other approaches to medical education. *Reliable* data have a high signal : noise ratio, where the signal refers to information about trainee competence and noise represents useless random error. Reliable data are the foundation needed for educators to reach *valid* decisions, judgements or inferences about trainees.[32–34] Reliable data are vital for, firstly, providing accurate feedback to learners about educational progress and, secondly, making arguments for valid research results.

Recent SBME research amplifies a 50-year historical legacy[35] by acknowledging that measures of clinical competence cover a very narrow bandwidth, whereas effective medical practice involves a broad and deep repertoire too complex to capture fully with today's evaluations.[5,33] Measurement development is a high-priority issue in SBME.

Today, there are three primary sources of SBME evaluation and research data, all of which are imperfect. The first and most common are observational ratings of trainee performance. Despite their ubiquity, observational ratings are subject to many sources of potential bias (unreliability) unless they are conducted under controlled conditions with much rater training and calibration.[36] A second source of SBME outcome data is trainee responses, which are either *selected* (as in multiple-choice questions [MCQs]) or *constructed* (e.g. when the candidate is instructed to write a patient note or respond to a simulated patient [SP] question).[37] The reliability of trainee response data measured directly is usually higher than the reliability of data from observational ratings.[32] A third source of SBME outcome data is represented by haptic sensors. Here simulators capture and record trainee 'touch' in terms of location and depth of pressure at specific anatomical sites. The pioneering research with haptic measurement in women's health care simulation carried out by Mackel *et al.*[38] and Pugh *et al.*[39] is noteworthy. Reliability estimation of haptic data is now in its infancy and much more work is needed.

The historical record and recent research show that SBME outcome measurement is one of the greatest challenges now facing the field. Progress in SBME outcome measurement research – multiple measures, convergence–divergence, generalisability analyses – is needed to advance medical education in general and SBME effectiveness specifically.

Simulation fidelity (low to high, multi-mode)

A key principle of SBME is that educational goals must dictate decisions about the acquisition and use of simulation technology for teaching and testing.[28] Effective use of medical simulation depends on a close match of education goals with simulation tools. Education in basic procedural skills like suturing, intubation and lumbar puncture can be delivered using simple task trainers, devices that mimic body parts or regions (e.g. the arms, pelvis, torso). Complex clinical events such as team responses to simulated hospital 'codes' require training on much more sophisticated medical simulators. These are lifelike full-body manikins that have computer-driven physiological features (e.g. heart rate, blood pressure), respond to physical interventions like chest compression, respond to drug administration and drug interactions, record clinical events in real time and simulate many other parameters. Virtual reality (VR) simulators are now in use to educate surgeons and medical subspecialists (e.g. invasive cardiologists) in complex procedures that are too dangerous to practise on live patients. However, decisions about the use of these and other SBME technologies should consider the match between goals and tools.[37]

Recent work by Kneebone *et al.*[40] uses multi-mode educational simulation. These investigators combine ‘inanimate models attached to simulated patients [to] provide a convincing learning environment’. Clinical skills including suturing a wound and urinary catheter insertion are taught and evaluated coincidentally with attention to doctor–patient interaction, patient comfort and patient privacy. This work unites the best features of inanimate simulation with animate standardised patients to present realistic clinical challenges for education and evaluation.[29,40]

Skill acquisition and maintenance

Clinical skill acquisition is the most common learning objective of SBME. Procedural skill acquisition accounts for the most research attention in SBME, whereas other skills and attributes of professionalism needed for clinical competence, such as communication skills, cultural sensitivity and patient ‘handover’ abilities, have received comparatively less research emphasis. Examples of high-quality clinical skill acquisition studies include the work of Murray *et al.*[41] on acute care skills in anaesthesiology and that of Wayne *et al.*,[22–26] which has focused on skill acquisition in internal medicine.

A growing number of new studies are being performed to evaluate the maintenance or decay over time of skills acquired in SBME settings. The results are mixed. The Wayne research group has demonstrated that ACLS skills acquired by internal medicine residents in a simulation laboratory do not decay at 6 and 14 months post-training.[42] This finding is reinforced by Crofts *et al.*[43] in obstetrics, who have shown that acquired skill at managing shoulder dystocia is largely maintained at 6 and 12 months post-SBME training among midwives and doctors in the UK. Contrary findings come from Sinha *et al.*,[44] whose data indicate some laparoscopic surgical skills decay after 6 months without added practice, especially for fine motor skills. Lammers[45] also reports significant skill decay after 3 months without follow-up practice among emergency medicine and family practice residents who earlier learned posterior epistaxis management using an oronasopharyngeal simulator. Thus it appears that skill decay depends on the specific skill acquired, the degree of skill learning (or overlearning) and the time allowed to elapse between learning and follow-up measurement. More research is clearly needed here.

Mastery learning

Mastery learning is an especially rigorous approach to competency-based education that dovetails closely with educational interventions featuring DP. In brief, mastery learning has seven complementary features:[21]

1. baseline (i.e. diagnostic) testing;
2. clear learning objectives, sequenced as units ordered by increasing difficulty;
3. engagement in educational activities (e.g. skills practice, data interpretation, reading) that are focused on reaching the objectives;
4. establishment of a minimum passing standard (e.g. test score, checklist score) for each educational unit;[46]
5. formative testing to gauge unit completion at a preset minimum passing *mastery* standard;
6. advancement to the next educational unit given measured achievement at or above the mastery standard, or
7. continued practice or study on an educational unit until the mastery standard is reached.

The goal of mastery learning is to ensure that *all* learners accomplish *all* educational objectives with little or no outcome variation. However, the amount of time needed to reach mastery standards for a unit’s educational objectives varies among learners. This represents a paradigm shift from the way simulation-based and many other educational activities are currently carried out. The mastery learning model will have significant impact on programme design, implementation and resource use.

Despite these considerations, a small but growing number of published research reports document the feasibility of mastery learning in SBME skill acquisition studies. These studies also use some form of DP to power the educational intervention. Examples include the studies of mastery learning of ACLS, thoracentesis and catheter insertion skills among internal medicine residents reported by Wayne *et al.*[23–26] The Lammers study on acquisition of posterior epistaxis management skills among emergency medicine and family practice residents employed a ‘pause-and-perfect’ training model, which is a close approximation to mastery learning.[45]

Transfer to practice

Transfer to practice demonstrates that skills acquired in SBME laboratory settings generalise to real clinical settings. This is the highest level of the Kirkpatrick hierarchy that is used widely to classify training programme outcomes.[47] Research into SBME that demonstrates its results transfer from the learning laboratory to real patient care settings and improved patient care 'stretches the endpoint'.[12] Studies that achieve these goals are also very hard to design and execute. Such work qualifies as 'translational science' because results from laboratory research are brought to the public in terms of, firstly, more skilful behaviour in clinical settings, secondly, improved patient care and, thirdly, improved patient outcomes.[48]

Several recent illustrations of SBME research have documented transfer of training to patient care settings. One report shows that simulation-trained internal medicine residents respond as teams to real hospital 'codes' (cardiac arrest events) with much greater compliance to established treatment protocols than more educationally advanced teams of residents who were not simulator-trained.[49] A second study involving internal medicine residents shows that trainees who have mastered central venous catheter (CVC) insertion in a simulation laboratory experience significantly fewer procedural complications (e.g. arterial puncture) in an intensive care unit (ICU) than residents who are not simulation-trained.[50] Patients in the ICU receiving care from CVC mastery residents also experience significantly lower rates of catheter-related bloodstream infections than patients receiving care from other residents.[51] In surgery, Seymour[52] has published convincing evidence that VR simulation training transfers directly to patient care by improving surgeons' operating room performance. In obstetrics, Draycott *et al.*[53] have published extensive research demonstrating improved neonatal outcomes of births complicated by shoulder dystocia after implementation of simulation-based training. Previously cited research reports by Domuracki *et al.*[17] and Edelson *et al.*[18] provide more evidence about the transfer of SBME learning to clinical practice.

The generalisability and utility of SBME research findings are likely to be demonstrated further as larger experimental and quasi-experimental studies report clinical outcome data. These studies are very difficult to design and conduct rigorously.

Team training

Psychologist Eduardo Salas and his colleagues[54] argue that 'patient care is a team sport'. These investigators cite evidence that one marker of team behaviour, communication, is the root cause of nearly 70% of errors (sentinel events) in clinical practice. Other signs of ineffective teamwork in clinical practice, including lack of shared goals, situation awareness, role clarity, leadership, coordination, mutual respect and debriefing, have been linked to such adverse clinical patient outcomes as nosocomial infections, adverse drug events and risk-adjusted mortality.[55] Health care team training has recently achieved recognition as an important educational goal. The Salas research team points out that 'training also provides opportunities to practise (when used with simulation) both task- and team-related skills in a ''consequence-free'' environment, where errors truly are opportunities for learning and providers receive feedback that is constructive, focused on improvement, and non-judgemental'.[54]

Salas and colleagues perceived a need to identify and describe key principles of team training in health care that can be embodied in curricula and taught using simulation technology.[54] They performed a quantitative and qualitative review of available literature including a 'content analysis of team training in health care'. The result is a set of 'eight evidence-based principles for effective planning, implementation, and evaluation of team training programmes specific to health care'. The 'eight critical principles are:

1. identify critical teamwork competencies and use these as a focus for training content;
2. emphasise teamwork over task work, design teamwork to improve team processes;
3. one size does not fit all … let the team-based learning outcomes desired, and organisational resources, guide the process;
4. task exposure is not enough … provide guided, hands-on practice;
5. the power of simulation … ensure training relevance to transfer environment;
6. feedback matters … it must be descriptive, timely and relevant;
7. go beyond reaction data … evaluate clinical outcomes, learning, and behaviours on the job, and
8. reinforce desired teamwork behaviours … sustain through coaching and performance evaluation.'[54]

The bottom line message from this scholarship is that team training works in carefully designed curricula which allow opportunities for the DP of teamwork skills in an SBME environment. The Salas research team has also published 11 'best practices' for measuring team performance in simulation-based training in a companion journal article.[56]

High-stakes testing

The standardisation, fidelity and reproducibility of medical simulation make the technology well suited to formative and summative evaluations of clinical competence. Formative evaluations are for practice and feedback, but summative evaluations are for 'high-stakes' decisions, such as those that involve the candidate passing a programme or course of study, or gaining certification or licensure. High-stakes testing demands highly reliable data that permit valid inferences about the competence of medical candidates. We anticipate increasing use of simulation in high-stakes medical testing as the technology advances and matures and as SBME measurement methods become more precise.[57]

Recent research and scholarship, chiefly in the procedural specialties, have demonstrated the utility of medical simulation in high-stakes testing. A prominent illustration is carotid stenting – typically performed by cardiologists, radiologists and vascular surgeons – in which simulation-based training and certification are now required for professional practice.[58]

The use and acceptance of simulation technology in training and high-stakes testing in anaesthesiology is growing. Berkenstadt *et al.*[59,60] have designed a research and development programme and implemented a simulation-based objective structured clinical examination (OSCE) into the Israeli national board examination in anaesthesiology. The OSCE was crafted carefully by a team of clinicians, simulation experts and testing specialists to include: 'three steps: (i) definition of clinical conditions that residents are required to handle competently; (ii) definition of tasks pertaining to each of the conditions, and (iii) incorporation of the tasks into hands-on simulation-based examination stations in the OSCE format including [1] trauma management, [2] resuscitation, [3] crisis management in the operating room, [4] regional anaesthesia, and [5] mechanical ventilation.' This high-stakes certification examination has yielded reliable data, is acceptable to candidates and practising anaesthesiologists, and will undergo continuous refinement and quality improvement.

Weller *et al.*[61] report a similar experience in Australia and New Zealand during the development and testing of a college-accredited simulation-based crisis management course for anaesthesia education. These scientists assert, 'Exposure to the concepts of crisis management is now widespread in the anaesthetic community in the region and should contribute to improved patient safety.'[61]

Simulation technology has also been applied to high-stakes testing in internal medicine. Hatala *et al.*[62,63] report Canadian studies that require candidates for board certification to examine an SP and then identify related clinical findings using a simulation of a patient abnormality. The OSCE stations measure candidate skills in the domains of cardiology and neurology. These SP encounters make a valuable contribution to the Canadian board examination in internal medicine and will probably grow in number with experience and improvement.

A final illustration of the use of medical simulation in high-stakes testing is drawn from research outside the procedural specialties. Instead, it involves work by educational scientists at the Educational Commission for Foreign Medical Graduates (ECFMG) who designed and evaluated a clinical skills assessment (CSA) for doctors who aspire to become certified to practise in the USA. van Zanten *et al.*[64] have published research that demonstrates how medical simulation in the form of SPs yields reliable evaluation data about candidates' interpersonal skills that allow for valid decisions about their professional competence. Medical simulation can be an effective tool for evaluating candidates' personal qualities and attributes, not just their procedural skills.

Instructor training

With regard to the effectiveness of SBME, the role of the instructor in facilitating, guiding and motivating learners is shrouded in mystery. There is a great unmet need for a uniform mechanism to educate, evaluate and certify simulation instructors for the health care professions. Evaluation research is lacking, but observation and experience teach several valuable lessons: effective SBME is *not* easy or intuitive; clinical experience alone is *not* a proxy for simulation instructor effectiveness, and simulation instructors and learners need *not* be from the same health care profession.

Many commercial vendors of medical simulation technology offer training courses for buyers and users of their equipment. Simulation instructor courses are increasingly available from schools and colleges of health professions education and from professional associations. Several descriptions of simulation instructor training courses have been published.[65–67] However, the short- and long-term value and utility of these educational opportunities are unknown without trustworthy data from evaluation research studies. Much more work is needed here.

Educational and professional context

Contexts of education and professional practice have profound effects on the substance and quality of learning outcomes and on how professional competence is expressed clinically. Roger Kneebone's work with authentic, multi-mode simulation provides visible testimony to the importance of context on learning and practice.[29,40] Schuwirth and van der Vleuten[68] argue that: 'Authenticity should have a high priority when programmes for the assessment of professional competence are being designed. This means that situations in which a candidate's competence is assessed should resemble the situation in which the competence will actually have to be used.' Simulation-based medical education that ignores its educational and professional context for teaching, evaluation or application in clinical practice is misdirected.

We are also reminded by the work of Pawson *et al.*[9,10] that SBME is a complex service intervention whose introduction in a medical education environment will not be smooth or easy. This group asserts that such interventions have a variety of key elements, including a long implementation chain, features that mutate as a result of refinement and adaptation to local circumstances, and represent open systems that feed back on themselves: 'As interventions are implemented, they change the conditions that made them work in the first place.'[9] In the words of the Greek philosopher Heraclitus, 'You cannot step twice into the same river.' The introduction and maintenance of SBME innovations will reshape the goals and practices of medical education programmes.

We believe this is the area of greatest need for additional research to inform SBME. Technical features of simulation devices have marginal influence on studies that support or refute the benefit and impact of SBME. Instead, features of the educational and professional contexts in which SBME is embedded have powerful influence on the process and delivery of training. Faculty expertise in training with these devices, their motivation to succeed, the local reward system, and institutional support contribute significantly to the success or failure of SBME. Such contextual features warrant detailed study and understanding so they can be shaped as needed to improve educational results.

DISCUSSION

This brief review is a distillate of our research and scholarly experience with SBME that covers a 40-year time span. The list of 12 features and best practices that we propose and amplify reflects our judgements about how the field has grown, matured, reached its current state and is likely to advance in the future. We acknowledge that this work may be biased from our sampling of the published literature and from our perspective as authors. In the spirit of preparing a critical review,[8] our aim was to 'represent various perspectives fairly'. No doubt other authors will have different views.

We are encouraged that productive SBME research groups are emerging in many medical specialties, including anaesthesiology, emergency medicine, internal medicine, obstetrics and gynaecology, paediatrics and surgery. Research programmes produce most valuable results when studies are thematic, sustained and cumulative.

There is no doubt that simulation technology can produce substantial educational benefits. However, informed and effective use of SBME technology requires knowledge of best practices, perseverance and attention to the values and priorities at play in one's local setting.

Contributors: all authors contributed to the conception and design of this review. WCM prepared the first draft of the manuscript. SBI, ERP and RJS contributed to its critical appraisal for intellectual content and revision to final form. All authors approved the manuscript prior to submission.
Acknowledgements: the authors would like to acknowledge the support given to them by Michael Gordon MD, PhD and the University of Miami Gordon Center for Research in Medical Education. WCM's contribution was supported by the Jacob R Suker MD Professorship in Medical Education at Northwestern University, Chicago, IL.
Funding: this work was supported in part by a grant from the Laerdal Foundation for Acute Medicine.
Conflicts of interest: none.
Ethical approval: not applicable.

REFERENCES

1 Buck GH. Development of simulators in medical education. *Gesnerus* 1991;**48** (Part 1):7–28.
2 Flexner A. *Medical Education in the United States and Canada. Bulletin No. 4.* New York, NY: Carnegie Foundation for the Advancement of Teaching 1910.
3 Kyle RR, Murray WB, eds. *Clinical Simulation: Operations, Engineering, and Management.* Burlington, MA: Academic Press 2008.

4 Issenberg SB, McGaghie WC, Hart IR *et al*. Simulation technology for health care professional skills training and assessment. *JAMA* 1999;**282**:861–6.

5 Issenberg SB, McGaghie WC, Petrusa ER, Gordon DL, Scalese RJ. Features and uses of high-fidelity medical simulation that lead to effective learning: a BEME systematic review. *Med Teach* 2005;**27**:10–28.

6 McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. Effect of practice on standardised learning outcomes in simulation-based medical education. *Med Educ* 2006;**40**:792–7.

7 Norman G, Eva KW. *Quantitative Research Methods in Medical Education*. Edinburgh: Association for the Study of Medical Education 2008.

8 Eva KW. On the limits of systematicity. *Med Educ* 2008;**42**:852–3.

9 Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review – a new method of systematic review designed for complex policy interventions. *J Health Serv Res Policy* 2005;**10** (Suppl 1):21–34.

10 Pawson R. *Evidence-based Policy: a Realist Perspective*. Thousand Oaks, CA: Sage Publications 2006.

11 Hammick M, Olckers L, Campion-Smith C. Learning in interprofessional teams: AMEE Guide no. 38. *Med Teach* 2009;**31**:1–12.

12 McGaghie WC. Research opportunities in simulation-based medical education using deliberate practice. *Acad Emerg Med* 2008;**15**:995–1001.

13 Rudolph JW, Simon R, Raemer DB, Eppich WJ. Debriefing as formative assessment: closing performance gaps in medical education. *Acad Emerg Med* 2008;**15**:1010–6.

14 Salas E, Klein C, King H, Salisbury M, Augenstein JS, Birnbach DJ, Robinson DW, Upshaw C. Debriefing medical teams: 12 evidence-based best practices and tips. *Jt Comm J Qual Patient Saf* 2008;**34**:518–27.

15 van de Ridder JMM, Stokking KM, McGaghie WC, ten Cate OThJ. What is feedback in clinical education? *Med Educ* 2008;**42**:189–97.

16 Fanning RM, Gaba DM. The role of debriefing in simulation-based learning. *Simul Healthc* 2007;**2**:115–25.

17 Domuracki KJ, Moule CJ, Owen H, Kostandoff G, Plummer JL. Learning on a simulator does transfer to clinical practice. *Resuscitation* 2009;**80**:346–9.

18 Edelson DP, Litzinger B, Arora V, Walsh D, Kim S, Lauderdale DS, van den Hoek TL, Becker LB, Abella BS. Improving in-hospital cardiac arrest process and outcomes with performance debriefing. *Arch Intern Med* 2008;**168**:1063–9.

19 Cordray DS, Pion GM. Treatment strength and integrity: models and methods. In: Bootzin RR, McKnight PE, eds. *Strengthening Research Methodology: Psychological Measurement and Evaluation*. Washington, DC: American Psychological Association 2006;103–24.

20 Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med* 2004;**79** (Suppl 10): 70–81.

21 McGaghie WC, Siddall VJ, Mazmanian PE, Myers J. Lessons for continuing medical education from simulation research in undergraduate and graduate medical education: effectiveness of continuing medical education: American College of Chest Physicians evidence-based educational guidelines. *Chest* 2009;**135** (Suppl 3):62–8.

22 Wayne DB, Butter J, Siddall VJ, Feinglass J, Wade LD, McGaghie WC. Simulation-based training of internal medicine residents in advanced cardiac life support protocols: a randomised trial. *Teach Learn Med* 2005;**17**:210–6.

23 Wayne DB, Butter J, Siddall VJ, Fudala MJ, Wade LD, Feinglass J, McGaghie WC. Mastery learning of advanced cardiac life support skills by internal medicine residents using simulation technology and deliberate practice. *J Gen Intern Med* 2006;**21**:251–6.

24 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.

25 Barsuk JH, Ahya SN, Cohen ER, McGaghie WC, Wayne DB. Mastery learning of temporary haemodialysis catheter insertion skills by nephrology fellows using simulation technology and deliberate practice. *Am J Kidney Dis* 2009;**54**:70–6.

26 Barsuk JH, McGaghie WC, Cohen ER, Balachandran JS, Wayne DB. Use of simulation-based mastery learning to improve the quality of central venous catheter placement in a medical intensive care unit. *J Hosp Med* 2009;**4**:397–403.

27 Cabot RC, Locke EA. The organisation of a department of clinical medicine. *Boston Med Surg J* 1905;**153**:461–5.

28 Issenberg SB. The scope of simulation-based health care education. *Simul Healthc* 2006;**1**:203–8.

29 Kneebone R. Simulation and transformational change: the paradox of expertise. *Acad Med* 2009;**84**:954–7.

30 Butter J, Grant TH, Egan M, Kaye M, Wayne DB, Carrion-Carire V, McGaghie WC. Does ultrasound training boost first year medical student competence and confidence when learning the abdominal examination? *Med Educ* 2007;**41**:843–8.

31 Adler MD, Vozenilek JA, Trainor JL, Eppich WJ, Wang EE, Beaumont JL, Aitchison PR, Erickson T, Edison M, McGaghie WC. Development and evaluation of a simulation-based paediatric emergency medicine curriculum. *Acad Med* 2009;**84**:935–41.

32 Downing SM. Reliability: on the reproducibility of assessment data. *Med Educ* 2004;**38**:1006–12.

33 Downing SM. Validity: on the meaningful interpretation of assessment data. *Med Educ* 2003;**37**:830–7.

34 Kane MT. Validation. In: Brennan RL, ed. *Educational Measurement*, 4th edn. Westport, CT: American Council on Education, Praeger Publishers 2006;17–64.

35 Gagne RM. Training devices and simulators: some research issues. *Am Psychol* 1954;**9**:95–107.

36 Williams RG, Klamen DA, McGaghie WC. Cognitive, social and environmental sources of bias in clinical competence ratings. *Teach Learn Med* 2003;**15**:270–92.

37 Scalese RJ, Issenberg SB. Simulation-based assessment. In: Holmboe ES, Hawkins RE, eds. *Practical Guide to the Evaluation of Clinical Competence*. Philadelphia, PA: Mosby Elsevier 2008;179–200.

38 Mackel TR, Rosen J, Pugh CM. Markov model assessment of subjects' clinical skill using the e-pelvis physical simulator. *IEEE Trans Biomed Eng* 2007;**54**:2133–41.

39 Pugh CM, Domont ZB, Salud LH, Blossfield KM. A simulation-based assessment of clinical breast examination technique: do patient and clinician factors affect clinician approach? *Am J Surg* 2008;**195**:874–80.

40 Kneebone RL, Kidd J, Nestel D, Barnet A, Lo B, King R, Yang GZ, Brown R. Blurring the boundaries: scenario-based simulation in a clinical setting. *Med Educ* 2005;**39**:580–7.

41 Murray DJ, Boulet JR, Avidan M, Kras JF, Henrichs B, Woodhouse J. Performance of residents and anaesthesiologists in a simulation-based skill assessment. *Anesthesiology* 2007;**107**:705–13.

42 Wayne DB, Siddall VJ, Butter J, Fudala MJ, Wade LD, Feinglass J, McGaghie WC. A longitudinal study of internal medicine residents' retention of advanced cardiac life support skills. *Acad Med* 2006;**81** (Suppl 10):9–12.

43 Crofts JF, Bartlett C, Ellis D, Hunt LP, Fox R, Draycott TJ. Management of shoulder dystocia: skill retention 6 and 12 months after training. *Obstet Gynecol* 2007;**110**:1069–74.

44 Sinha P, Hogle NJ, Fowler DL. Do the laparoscopic skills of trainees deteriorate over time? *Surg Endosc* 2008;**22**:2018–25.

45 Lammers RL. Learning and retention rates after training in posterior epistaxis management. *Acad Emerg Med* 2008;**15**:1181–9.

46 Downing SM, Tekian A, Yudkowsky R. Procedures for establishing defensible absolute passing scores on performance examinations in health professions education. *Teach Learn Med* 2006;**18**:50–7.

47 Kirkpatrick DL. *Evaluating Training Programs*, 2nd edn. San Francisco, CA: Berrett-Koehler 1998.

48 Woolf SH. The meaning of translational research and why it matters. *JAMA* 2008;**299**:211–3.

49 Wayne DB, Didwania A, Feinglass J, Fudala MJ, Barsuk JH, McGaghie WC. Simulation-based education improves the quality of care during cardiac arrest team responses at an academic teaching hospital: a case-control study. *Chest* 2008;**133**:56–61.

50 Barsuk JH, McGaghie WC, Cohen ER, O'Leary KS, 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.

51 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**:1420–3.

52 Seymour NE. VR to OR: a review of the evidence that virtual reality simulation improves operating room performance. *World J Surg* 2008;**32**:182–8.

53 Draycott T, Crofts JF, Ash JP, Wilson LV, Yard E, Sibanda T, Whitelaw A. Improving neonatal outcome through practical shoulder dystocia training. *Obstet Gynecol* 2008;**112**:14–20.

54 Salas E, DiazGranados D, Weaver SJ, King H. Does team training work? Principles for health care. *Acad Emerg Med* 2008;**11**:1002–9.

55 Sorbero ME, Farley DO, Mattke S, Lovejoy S. *Outcome Measures for Effective Teamwork in Inpatient Care. (RAND Technical Report TR-462-AHRQ)*. Arlington, VA: RAND Corporation 2008.

56 Rosen MA, Salas E, Wilson KA, King HB, Salisbury M, Augenstein JS, Robinson DW, Birnbach DJ. Measuring team performance in simulation-based training: adopting best practices for health care. *Simul Healthc* 2008;**3**:33–41.

57 Pugh CM. Simulation and high-stakes testing. In: Kyle RR, Murray WB, eds. *Clinical Simulation: Operations, Engineering, and Management*. Burlington, MA: Academic Press 2008;655–66.

58 Gallagher AG, Cates CV. Approval of virtual reality training for carotid stenting: what this means for procedural-based medicine. *JAMA* 2004;**292**:3024–6.

59 Berkenstadt H, Ziv A, Gafni N, Sidi A. Incorporating a simulation-based objective structured clinical examination into the Israeli national board examination in anaesthesiology. *Anesth Analg* 2006;**102**:853–8.

60 Berkenstadt H, Ziv A, Gafni N, Sidi A. The validation process of incorporating simulation-based accreditation into the anaesthesiology Israeli national board exams. *Isr Med Assoc J* 2006;**8**:728–33.

61 Weller J, Morris R, Watterson L, Garden A, Flanagan B, Robinson B, Thompson W, Jones R. Effective management of anaesthetic crises: development and evaluation of a college-accredited simulation-based course for anaesthesia education in Australia and New Zealand. *Simul Healthc* 2006;**1**:209–14.

62 Hatala R, Kassen BO, Nishikawa J, Cole G, Issenberg SB. Incorporating simulation technology in a Canadian internal medicine specialty examination: a descriptive report. *Acad Med* 2005;**80**:554–6.

63 Hatala R, Cole G, Kassen BO, Bacchus CM, Issenberg SB. Does physical examination competence correlate with bedside diagnostic acumen? An observational study. *Med Teach* 2007;**29**:199–203.

64 van Zanten M, Boulet JR, McKinley D. Using standardised patients to assess the interpersonal skills of physicians: 6 years experience with a high-stakes certification examination. *Health Commun* 2007;**22**:195–205.

65 Vollmer J, Monk S, Heinrichs W. Staff education for simulation: train-the-trainer concepts. In: Kyle RR, Murray WB, eds. *Clinical Simulation: Operations, Engineering, and Management*. Burlington, MA: Academic Press 2008;625–42.

66 Chow RE, Naik VN. Experiential training for new simulation coordinators. In: Kyle RR, Murray WB, eds. *Clinical Simulation: Operations, Engineering, and*

Management. Burlington, MA: Academic Press 2008;643–6.
67 Dieckmann P, Rall M. Becoming a simulation instructor and learning to facilitate: the Instructor and Facilitation Training (InFacT) Course. In: Kyle RR, Murray WB, eds. *Clinical Simulation: Operations, Engineering, and Management*. Burlington, MA: Academic Press 2008;647–52.
68 Schuwirth LWT, van der Vleuten CPM. The use of clinical simulation in assessment. *Med Educ* 2003;**37** (Suppl 1):65–71.

Received 16 February 2009; editorial comments to authors 18 May 2009; accepted for publication 16 September 2009