

The Quality Of Evidence In Preclinical Medical Education Literature: A Systematic Review

This article was published in the following Dove Press journal:
Advances in Medical Education and Practice

Marilyn Leif¹
Natali Semarad²
Vaishnavi Ganesan¹
Kevin Selting¹
Justin Burr¹
Austin Svec¹
Peggy Clements³
Geoffrey Talmon⁴

¹University of Nebraska Medical Center, Omaha, NE, USA; ²Creighton University in Omaha, Omaha, NE, USA; ³Wofford College, Spartanburg, SC, USA; ⁴Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE, USA

Introduction: To practice effective evidence-based teaching, the need for well-designed studies that describe outcomes related to educational interventions is critical. The quality of the literature in basic science disciplines is unknown. The study objective was to conduct a systematic review of the literature to assess study design in articles describing innovations in preclinical medical education.

Method: The authors searched PubMed for all articles published in English between 2000 and 2017 describing interventions in preclinical medical education related to anatomy, physiology, and biochemistry. Articles were scored using a modification of the Medical Education Research Study Quality Instrument.

Results: Of the 817 articles identified, 177 met final inclusion criteria (75 anatomy, 86 physiology, and 16 biochemistry). Laboratory, student-driven, and online activities were the most frequently reported. The average score for all papers was 15.7 (27 points possible). More than 80% reported experiences with one cohort of students and >97% involved only one institution. Only 25–49% of reports utilized a comparison (control) group. Proper statistical models for analysis of results were used in only 44–62% of papers.

Conclusion: Manuscripts had a strong tendency toward single institutional studies that involved one cohort of students. The use of a control/comparison group when assessing effectiveness was seen in <50% and nearly all reported outcomes solely in the form of student satisfaction or factual recall/skill performance.

Keywords: evidence-based teaching, preclinical, anatomy, physiology, biochemistry, study design

Introduction

The traditional paradigm of US medical education is changing. Medical education is beginning to adopt a more andragogic approach to teaching with an understanding that lecture-based education may be less effective,¹ especially for Millennials.² New standardized evaluation tools that measure competency in specific domains such as Entrustable Professional Activities, both created within individual medical schools and developed by societies such as the American Association of Medical Colleges, are being more widely used.³ In addition, many institutions have begun to change their curricula to include instruction on communication, teamwork, health systems, and population health.^{4,5}

In tandem with this has been the explosion of computer-based teaching which has altered the way medical curricula are delivered. Online learning, simulation, and technologies that replace or augment traditional lecture or laboratory experiences that formerly made up the bulk of preclinical experiences are now becoming

Correspondence: Geoffrey Talmon
Department of Pathology and Microbiology, University of Nebraska Medical Center, 983135 Nebraska Medical Center, Omaha, NE 68198-3135, USA
Tel +1(402) 559-4793
Fax +1(402) 559-6018
Email gtalmon@unmc.edu

more common.^{6,7} The use of these new methods is becoming attractive as competition for time increases and resources become scarcer. This is particularly true in the realm of preclinical and basic science education; new additions to curricula often come at the expense of contact hours and laboratory time.

When educators innovate, they tend to share with colleagues in person and via formal methods such as professional meetings and articles. Descriptions of these new instructional tools are prevalent in the literature with almost all discussing how they impacted the training of that institution's students. As educational leaders and instructors contemplate new instructional methods and tools that appear in journals and at conferences, a common question that arises is "What actually will work for us?" In other words, is a new instructional method or tool effective in helping an educator's own learners meet outcomes?

To this end, there is a need for strong experimental design when evaluating preclinical educational activities. In undergraduate medical education, particularly preclinical/basic science education, studies as a rule are not well designed. The outcomes reported in the vast majority of reports of novel educational programs, technologies, or approaches are usually immediate pre/posttests performance (factual recall), satisfaction surveys, or learner confidence assessments. Unfortunately, outcomes described from novel basic science or preclinical activities are not robust or long term, nor is it often that results are compared with a well-matched (and in some instances any) control group. One does not need to look far in the literature to find evidence of this recurring limitation of generalizability.⁸⁻¹²

One of the first steps in calling for such a change is to assess the true "state-of-affairs." While rare reviews, like that performed by Chen et al⁸ examining the effectiveness medical education interventions as a whole exist, assessment of the quality of evidence of interventions viewed in the lens of preclinical fields (anatomy, biochemistry, and physiology) is lacking. A search of four literature databases (PubMed, ERIC, Google Scholar, CINAHL) for such systematic reviews related to each of these disciplines revealed no pertinent results.

We undertook a systematic review of the literature for studies that describes an intervention in preclinical, basic science education (anatomy, and physiology, biochemistry), defined as a change made to an established teaching method and/or content. The fundamental question that we asked was: how many descriptions of preclinical educational intervention studies are presented in form that speaks to general applicability to other institutions? To this end, we analyzed

relevant articles using a modification of the Medical Education Research Study Quality Instrument (MERSQI).¹³

Methods

The study was reviewed by the institutional review boards of UNMC and Johns Hopkins University. The PubMed (Medline) database was searched for all works describing an intervention in preclinical anatomy, physiology, and biochemistry medical education. The search for and attainment of articles/abstracts were facilitated by librarians at the author's institution, using the following queries:

Anatomy

(((((anatomy[ti]))) AND ((education[ti] OR educational[ti] OR instruct[ti] OR instructed[ti] OR instruction[ti] OR teach[ti] OR teacher[ti] OR teaching[ti] OR train[ti] OR training[ti] OR trained[ti] OR trainer[ti] OR taught[ti] OR pedagogy[ti] OR learn[ti] OR learning[ti] OR learner[ti] OR curriculum[ti]))) AND ((("medical student"[ti]))) OR (((("Anatomy/education"[Mesh]) AND (((("Teaching Materials"[Mesh]) OR "Teaching/methods"[Mesh]) OR "Education/methods"[Mesh])) AND ((("Students, Medical"[Mesh]) OR "Education, Medical"[Mesh]))

Biochemistry

(((((biochemistry[ti] OR "molecular biology"[ti] OR neurochemistry[ti] OR biochemical[ti] OR biochemics[ti] OR "cellular biology"[ti] OR metabolism[ti] OR metabolic[ti]))) AND ((education[ti] OR educational[ti] OR instruct[ti] OR instructed[ti] OR instruction[ti] OR teach[ti] OR teacher[ti] OR teaching[ti] OR train[ti] OR training[ti] OR trained[ti] OR trainer[ti] OR taught[ti] OR pedagogy[ti] OR learn[ti] OR learning[ti] OR learner[ti] OR curriculum[ti]))) AND ((("medical student"[tiab] OR "medical students"[tiab] OR undergraduate[tiab] OR undergraduates[tiab] OR resident[tiab] OR residents[tiab] OR residency[tiab] OR intern[tiab] OR interns[tiab] OR internship[tiab]))) OR (((("Biochemistry/education"[Mesh] AND ("Teaching Materials"[Mesh] OR "Teaching/methods"[Mesh] OR "Education/methods"[Mesh])) AND ((("Students, Medical"[Mesh] OR "Education, Medical"[Mesh]))

Physiology

(((((physiology[ti] OR electrophysiology OR endocrinology[ti] OR neurophysiology[ti] OR neuroendocrinology[ti] OR physiological[ti]))) AND ((education[ti] OR educational[ti] OR instruct[ti] OR instructed[ti] OR instruction[ti] OR teach[ti] OR teacher[ti] OR teaching[ti] OR

train[ti] OR training[ti] OR trained[ti] OR trainer[ti] OR taught[ti] OR pedagogy[ti] OR learn[ti] OR learning[ti] OR learner[ti] OR curriculum[ti])) AND (((“medical student”[tiab] OR “medical students”[tiab] OR undergraduate[tiab] OR undergraduates[tiab] OR resident[tiab] OR residents[tiab] OR residency[tiab] OR intern[tiab] OR interns[tiab] OR internship[tiab])) NOT medline[sb])) OR (((“Physiology/education”[Mesh]) AND (((“Teaching Materials”[Mesh]) OR “Teaching/methods”[Mesh]) OR “Education/methods”[Majr])) AND (“Students, Medical”[Mesh]) OR “Education, Medical”[Majr]))

The initial article review was performed by four of the authors (ML, VG, JB, AS) to evaluate which articles should be excluded. Works not related to medical student education, not describing an intervention, or published in a language other than English, and those published before 2000 and after 2017 were not included. The articles were randomly assigned to the reviewers, with each being reviewed by two individuals. The senior author (GT) refereed discrepancies and reviewed 5% of all articles for quality control. Prior to the proceeding, complete agreement between all reviews was achieved.

A second review of the publications meeting inclusion criteria was performed by a single co-author (NS). These articles were categorized by the type of intervention described and scored using a modification of the MERSQI, a previously validated tool developed to evaluate the quality of medical education research (13). This modified scoring system evaluated papers on the basis of several domains, including number of learners, cohorts, and institutions included in the study. In addition, it included the response rate when surveys were used as an outcome measurement, the type of outcome data reported, sophistication of data analysis, appropriateness of statistical analysis (i.e., use of the correct statistical model for the type of data presented), and Kirkpatrick level of outcome reported (Table 1).¹⁴ The MESQRI domains related to the validity of assessment were not utilized for this analysis.

Those papers in which at least two domains could not be evaluated were further excluded from analysis. Each remaining paper was assigned a score based on these above domains, with possible scores ranging from 2 to 27. An identical refereeing and quality control process was undertaken as in the first review with 5% of papers being reviewed by the senior author for quality control. Statistical analysis was performed using ANOVA, where applicable.

Results

The literature search yielded a total of 817 articles as noted in Figure 1. After exclusion of works that did not solely relate to preclinical medical student education or describe an intervention (e.g., commentary), 293 articles remained. Removal of those that did not provide enough information to allow scoring in at least two domains yielded 177 articles for analysis: 75 of which discussed interventions in anatomy, 86 in physiology, and 16 related to biochemistry.

The types of interventions described in the remaining articles are outlined in Figure 2. Overall, laboratory, online, and student-driven activities were the most commonly reported. More than half of the anatomy papers discussed new or modified experiences in laboratories. Student-driven/led activities represented the most frequent intervention in those works related to both preclinical physiology and biochemistry teaching. The reports in 31 anatomy (41%), 20 physiology (23%), and 6 biochemistry (38%) papers involved more than one type of intervention.

The overall mean score for all of the analyzed articles was 15.7 (range 2–23), with the highest score possible being 27. The mean overall score for anatomy manuscripts was the highest. Biochemistry papers had the least variation in scores (range 13–20), although substantially fewer articles from this discipline were reviewed (Table 2).

The percent of studies attaining points in each individual domain is shown in Table 3. In total, the mean scores for each of the analyzed domains did not vary by greater than 1 point among all three disciplines. The largest variation was seen in the score for appropriateness of data analysis (0.72 points).

At least 80% of reports in all three disciplines either only described experiences with a single group of students or this could not be determined from the information provided, with significantly more anatomy works not reporting the number of cohorts studied. Most reports outlined outcomes of cohorts larger than 100 students. More than 97% of manuscripts reported the experience of only one institution; only one physiology paper involved the experience at two.

Several of the manuscripts used multiple types of study or instrument to measure effectiveness. For those reports, the highest applicable score was recorded. The two main types of study that the reports utilized for physiology and biochemistry involved single groups with a simple satisfaction survey or pretest/posttest as the instrument to

Table 1 Scoring Scheme Used For Analysis Of Articles Meeting Inclusion Criteria (Modified From The Medical Education Research Study Quality Index¹³)

Domain	Attribute	Score
Number of cohorts	Not applicable/could not assess	0
	1	1
	>1	2
Number of learners	Not applicable/could not assess	0
	<10	1
	10–19	2
	20–49	3
	50–100	4
	>100	5
Study design	Not applicable/could not assess	0
	Single group cross sectional or single group posttest only	1
	Single group pretest and posttest	2
	Nonrandomized, 2 groups	3
	Randomized controlled trial	4
Number of institutions	Not applicable/could not assess	0
	1	1
	2	2
	>2	3
Response rate	Not applicable	0
	<50% or not reported	1
	50–74%	2
	At least 75%	3
Type of data	Not applicable/could not assess	0
	Assessment by study participant	1
	Objective measurement	2
Appropriateness of data analysis	Not applicable/could not assess	0
	Inappropriate for study design or type of data	1
	Appropriate for study design or type of data	2
Complexity of analysis	Not applicable/could not assess	0
	Descriptive only	1
	Beyond descriptive analysis	2
Outcomes	Not applicable/could not assess	0
	Satisfaction, attitudes, perceptions, opinions, general facts	1
	Knowledge or skills	2
	Behaviors	3
	Patient/health care system	4

Note: Reproduced with permission from 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(9):1002–1009. Copyright 2007, American Medical Association. All rights reserved.¹³

assess effectiveness. For anatomy manuscripts, non-randomized two group comparisons (i.e., performance 1 year and the prior “control” year) was more significantly common. In each instance, fewer than 10% of studies utilized any form of randomization.

Only among biochemistry reports did more than half of the studies use an objective measurement of the effectiveness of their described intervention (69%). The majority of anatomy and physiology papers solely used student surveys. For those studies that included the results of a student survey as a

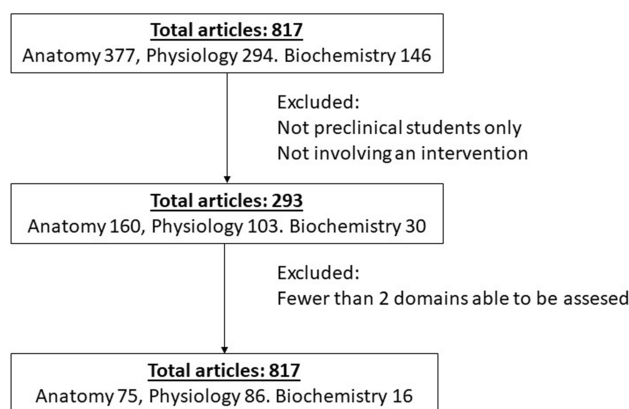


Figure 1 Literature search results and proportion of studies included in the analysis.

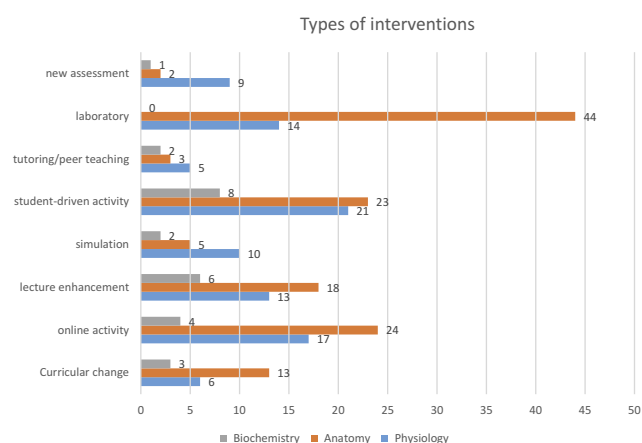


Figure 2 Types of intervention described in analyzed articles.
Note: Multiple interventions were used in 57 manuscripts.

part of outcome data fewer than 60% in each discipline had at least 75% of students responding.

A substantial number of papers from each of the three disciplines reported results beyond a descriptive analysis, ranging from 44% to 62%. When reported, the model used to

assess the statistical significance of any difference in score, perception, etc. was appropriate in 28–68% of papers, with biochemistry having a significantly greater number that did. In approximately 30% of physiology manuscripts, statistical significance was either not reported or the test used was not identified.

The highest reported of outcome of effectiveness for interventions in most biochemistry and physiology papers and all but 3 anatomy manuscripts represented the lower two of Kirkpatrick's levels of evidence. Papers among the three were relatively equally split between student perceptions/attitudes and knowledge with the differences not being statistically significant.

Discussion

This need for well-designed studies in medical education research has been long recognized, particularly as medical educators turn to the literature to guide their teaching practices. The underlying precept is that, much like with clinical interventions, the curricula, activities, and assessment schemes employed in the education of future physicians should have demonstrated effectiveness through robust studies with ample supporting evidence. As such, evidence-based teaching is predicated on the need for sufficient high-quality studies in the body of published literature.

When wearing their clinical “hat,” clinician educators base decisions about patient care interventions on data that in most cases comes from well-designed, large, controlled studies with statistically significant, long-term results. However, when the same individuals attempt to guide their educational activities on the basis of similar data, it is lacking. Compared to clinical trials, reports on the vast majority of educational interventions would be classified as anecdotal, of poor design, often without controls, rich in local confounding factors, and focused on short-term results.

This gap in knowledge is particularly problematic in preclinical disciplines. The example of problem-based learning (PBL) is very apropos. In the 1990s, many medical schools began to adopt PBL in their curricula, with some institutions converting their entire preclinical teaching to a PBL-based format. Much of this trend was based on the multitude of “case reports” that advocated its effectiveness. With more scrutiny and longer follow-up, however, the grand outcomes of PBL in preclinical medical education are at most mixed: “Twenty-two years of research shows that PBL does not impact knowledge acquisition; evidence for other outcomes does not provide unequivocal support for enhanced learning.”¹⁵

Table 2 Overall Scoring Data For Analyzed Manuscripts

	Anatomy	Physiology	Biochemistry	Total
N	75	86	16	177
Mean total score	16.6	14.5	16.1	15.7
Standard deviation	3.9	4.4	1.65	3.3
Minimum	3	2	13	
Maximum	23	21	20	

Table 3 Percentages Of Manuscripts With A Given Score In Each Domain Of Assessment

Criterion	Score	Anatomy	Physiology	Biochemistry	P
Number of cohorts					
Could not assess	0	32.0%	14.0%	12.5%	<0.01
1	1	64.0%	73.3%	75.0%	
>1	2	4.0%	12.8%	12.5%	
Mean score		1.28	0.98	1	
Number of learners					
Could not assess	0	1.3%	8.2%	0%	0.20
<10	1	0.0%	0.0%	0%	
10 to 19	2	5.3%	2.3%	0%	
20 to 49	3	14.8%	9.5%	6.3%	
50 to 99	4	13.3%	23.5%	25.0%	
At least 100	5	65.3%	56.5%	68.8%	
Mean score		4.35	4.09	4.63	
Study design					
Could not assess	0	2.7%	9.4%	0.0%	0.01
Single group cross-sectional or single group posttest only	1	30.7%	43.5%	50.0%	
Single group pre/posttest	2	8.0%	10.6%	18.8%	
Non-randomized, 2 group	3	49.3%	30.6%	25.0%	
Randomized controlled trial	4	9.3%	5.9%	6.3%	
Mean score		2.31	1.79	1.79	
Number of institutions					
Could not assess	0	0.0%	1.2%	0%	0.84
1	1	100.0%	97.6%	100%	
2	2	0.0%	1.2%	0%	
>2	3	0.0%	0.0%	0%	
Mean score		1	0.99	1	
Response rate					
Not applicable	0	22.7%	19.8%	37.5%	0.10
<50% or not reported	1	9.3%	32.6%	18.8%	
50–74%	2	10.7%	10.4%	0%	
At least 75%	3	57.3%	37.2%	43.7%	
Mean score		2.03	1.65	1.5	
Type of data*					
Could not assess	0	10.6%	5.9%	0%	0.69
Assessment by participant	1	42.7%	43.5%	68.8%	
Objective measurement	2	46.7%	50.6%	31.3%	
Mean score		1.36	1.43	1.31	
Appropriateness of data analysis					
Could not assess	0	6.6%	30.2%	0%	<0.01
Inappropriate to study	1	50.7%	41.9%	31.3%	
Appropriate for study	2	42.7%	27.9%	68.8%	
Mean score		1.36	0.98	1.69	

(Continued)

Table 3 (Continued).

Criterion	Score	Anatomy	Physiology	Biochemistry	P
Complexity of analysis					
Could not assess	0	8.0%	31.4%	0%	0.01
Descriptive analysis only	1	37.3%	24.4%	37.5%	
Beyond descriptive analysis	2	54.7%	44.2%	62.5%	
Mean score		1.46	1.13	1.63	
Outcomes					
Could not assess	0	6.7%	5.8%	0%	0.96
Satisfaction, attitudes, perceptions, and opinions	1	48.0%	43.0%	56.2%	
Knowledge, skills	2	41.3%	51.2%	43.8%	
Behaviors	3	4.0%	0.0%	0%	
Patient/health care outcomes	4	0.0%	0.0%	0%	
Mean score		1.42	1.45	1.44	

Note: *Some studies included multiple measures of effectiveness, such as a satisfaction survey combined with a pretest/posttest.

While rare studies like that above examining the effectiveness of a type of educational activity exist, assessment of the quality of evidence of interventions viewed through the lens of particular preclinical fields is lacking. This is particularly problematic as more medical schools seek to redesign preclinical curricula with their heavy basic science emphasis with a purported focus on following best practices. To this end, our review (the first of its type to our knowledge) was focused on assessing the quality of study design in manuscripts describing preclinical educational interventions in three core disciplines: anatomy, physiology, and biochemistry.

For the purposes of this analysis, the published MERQSI was modified slightly to include information about the number of learners studied in each work. Additionally, as the majority of works did not specifically describe a novel evaluation tool, scores from this domain were not included. The aim was to “right size” the scoring scheme for basic science manuscripts. Even with this modification, the average score of works in all three disciplines was 16 (59% of the 27 points possible) with scores ranging from 2 to 24. Interestingly, this is similar to that seen in a similar review of educational abstracts in clinical education.¹⁶

One of the primary findings that speaks to the potential gap in high-quality studies relates to the low number of articles from the initial search that were able to be evaluated. Of the 293 works that described an intervention, only 60% presented enough information in their text to be rated in two domains in the MERSQI. The majority of

these papers could be characterized as descriptive in nature. Of note, similar reviews have noted a similar trend in examining the literature related to quality improvement training in medical education.¹⁷

The types of educational interventions discussed in those articles meeting inclusion criteria covered a wide gamut of modalities. Not surprisingly, the majority of anatomy works described novel practices related to laboratory activities. In addition, with the rise of andragogical principles in medical education, student-driven, and online activities also were common sources of innovation.

The key factor to consider is generalizability – what is the likelihood that an intervention that is reported to be effective will have the same success in one’s local environment. The majority of reports in all three disciplines involved only one cohort of student and all but rare anatomy papers were studies that were performed at a single institution. Further, many papers discussed “significant” improvements in student performance following an intervention. However, in 56–72% of instances, the statistical analysis performed was either inappropriate or not listed at all. As such, it may be difficult to determine if any effect in student performance as a result of an intervention would translate to a different institution’s curriculum or student body or whether it was “significant” at all.

Many authors contend that fewer controlled trials have been performed in medical education than should be.^{18,19} Our data show that only 37–59% of reports involved a comparison (control) group and fewer than 10% in each discipline involved randomization. However, randomized,

double-blinded, controlled trial, the gold standard model of clinical trials, likely cannot be adopted wholesale in medical education research. Blinding, one of the cornerstones of experimental medicine, is often not possible (or impractical) when designing educational trials. Randomization often requires longer study periods, and some experts question the ethics of randomization in educational research when the experimental intervention is untried.²⁰ In addition, there may be ethical concerns related to a cohort of students being exposed to an experimental, enhanced learning experience but still being assessed via the same method as their peers.

The ultimate method to assess the effectiveness of basic science teaching is to examine students' ability to apply concepts/facts to the care of patients in clinical rotations and beyond. Currently, published, well-designed assessment tools of this type are lacking. Gillman et al outlined a model for assessing the level of outcomes in medical education in 2001, defining higher-order outcomes at the patient and community level – areas that they related were not traditionally assessed by educators.²¹ Supporting this, Chen et al reported a review of 600 medical education articles in 2004, and only 4 measured these types of end results.⁸

One potential limitation of this review is the use of the MERSQI, predominately designed as an instrument to evaluate educational projects in the clinical arena, to papers describing preclinical educational activities. For example, preclinical manuscripts would likely always score lower in one particular domain, “Outcomes”, by nature of the content and placement in curricula. Not surprisingly, therefore, this was the case in this review. Our data show that more than 90% of papers among all three disciplines only reported lower level outcomes: student perceptions/satisfaction and (often immediate) recall of knowledge or skills performance. Only among a small percentage (<5%) of anatomy papers did any measure of effectiveness include a behavior change, usually related to treatment of cadavers. No work described patient or health-system level outcomes.

While there is universal recognition that an adequate comprehension and ability to apply knowledge of basic science disciplines is a fundamental skill for health care practitioners,²² there are no widely used sets of competencies or metrics for assessing higher level outcomes related to basic science education, especially related to long-term application and effects on distant behaviors and skills. Some specialty societies (such as the Association of

Pathology Chairs) have developed goals and objectives that could be used as a guide.²³ A logical (and much needed) next step is the creation of assessment tools linked to these objectives that could be deployed in the clinical years or residency to determine the ultimate impact of interventions.

In spite of these results, it is not the authors' contention that the papers that either did not score highly or were not included in our analysis are not of potential value. Indeed, even descriptive studies containing no or rudimentary outcome data have the potential to inspire innovation or adjust existing similar practices, providing a springboard for future scholarship. The focus of this review was to assess factors in the reviewed reports that could impact generalizability.

In summary, manuscripts describing preclinical educational interventions had a strong tendency toward single institutional studies that involved one cohort of students. The use of a control/comparison group when assessing effectiveness was seen in <50% of papers and nearly all reported outcomes in the form of student satisfaction or factual recall/skill performance. While its narrow focus is a potential weakness, it is hoped that these results will incite medical educators to begin to consider the more robust study design when evaluating educational interventions.

Disclosure

The authors have no relevant conflicts of interest or financial relationships to disclose in this work.

References

1. Taylor DC, Hamdy H. Adult learning theories: implications for learning and teaching in medical education: AMEE Guide No. 83. *Med Teach*. 2013;35(11):e1561–e1572. doi:10.3109/0142159X.2013.828153
2. Talmon G, Beck-Dallahan G. *Mind the Gap: Generational Differences in Medical Education*. North Syracuse, NY: Gegensatz Press; 2017.
3. Brown DR, Warren JB, Hyderi A, et al. Finding a path to entrustment in undergraduate medical education: a progress report from the AAMCCore entrustable professional activities for entering residency entrustment concept group. *J Gen Intern Med*. 2017;92(6):774–779. (epub ahead of print).
4. Beck M Innovation is sweeping through U.S. Medical Schools. Wall Street Journal. February 26, 2015. Available from: <https://www.wsj.com/articles/innovation-is-sweeping-through-u-s-medical-schools-1424145650>. Accessed July 7, 2017.
5. Shelton PG, Corral I, Kyle B. Advancements in undergraduate medical education: meeting the challenges of an evolving world of education, healthcare, and technology. *Psychiatr Q*. 2017;88(2):225–234. doi:10.1007/s11126-016-9471-x

6. Ruiz JG, Mintzer MJ, Leipzig R. The impact of E-learning in medical education. *Acad Med.* 2006;81(3):207–212. doi:10.1097/00001888-200603000-00002
7. Jones F, Passos-Neto CE, Braghiroli OFM. Simulation in medical education: brief history and methodology. *Princ Pract Clin Res.* 2015;1(2):56–63.
8. Chen FM, Baucher H, Burstin H. A call for outcomes research in medical education. *Acad Med.* 2004;79:955. doi:10.1097/00001888-200410000-00010
9. Prystowsky JB, Bordage G. An outcomes research perspective on medical education: the predominance of trainee assessment and satisfaction. *Med Educ.* 2001;35:334. doi:10.1046/j.1365-2923.2001.01054.x
10. Sanner MA. Medical students' attitude toward autopsy: how does experience with autopsies influence opinion? *Arch Pathol Lab Med.* 1995;119(9):851–858.
11. Verma SK. Teaching students the value of autopsy. *Acad Med.* 1999;74(8):855. doi:10.1097/00001888-199908000-00005
12. Burton JL. Teaching pathology to medical undergraduates. *Curr Diag Pathol.* 2005;11(5):308–316. doi:10.1016/j.cdip.2005.05.009
13. 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(9):1002–1009. doi:10.1001/jama.298.9.1002
14. Kirkpatrick D. Revisiting Kirkpatrick's four-level-model. *Train Develop.* 1996;1:54–57.
15. Hartling E, Spooner C, Tjosvold L, Oswald A. Problem based learning in preclinical medical education: 22 years of outcome research. *Med Teach.* 2010;35:28–35. doi:10.3109/01421590903200789
16. Smith R, Learman L. A plea for MESQRI: the medical education research study quality instrument. *Obstet Gynecol.* 2017;130(4):686–690. doi:10.1097/AOG.0000000000002091
17. Windish DM, Reed DA, Boonyasai RT, Chakraborti C, Bass EB. Methodological rigor of quality improvement curricula for physician trainees: a systematic review and recommendations for change. *Acad Med.* 2009;84(12):1677–1692. doi:10.1097/ACM.0b013e3181bfa080
18. Sullivan GM. Getting off the “gold-standard”: randomized controlled trials and education research. *J Grad Med Educ.* 2011;3:285–289. doi:10.4300/JGME-D-11-00147.1
19. Torgerson CJ. Educational research and randomised trials. *Med Educ.* 2002;36(11):1002–1003. doi:10.1046/j.1365-2923.2002.01335.x
20. Norman G. RCT = results confounded and trivial: the perils of grand educational experiments. *Med Educ.* 2003;37:582–584. doi:10.1046/j.1365-2923.2003.01586.x
21. Gilman SC, Cullen RJ, Leist JC, Craft CA. Domains-based outcomes assessment of continuing medical education: the VA's model. *Acad Med.* 2002;77:812. doi:10.1097/00001888-200208000-00010
22. Spencer A, Brosenitsch T, Levine AS, Kanter SL. Back to the basic sciences: an innovative approach to teaching senior medical students how best to integrate basic science and clinical medicine. *Acad Med.* 2008;83(7):662–669. doi:10.1097/ACM.0b013e318178356b
23. Pathology Competencies for Medical Education. Association of Pathology Chairs. Available from: <https://journals.sagepub.com/page/apc/pcme>. Accessed December 19, 2018.

Advances in Medical Education and Practice

Publish your work in this journal

Advances in Medical Education and Practice is an international, peer-reviewed, open access journal that aims to present and publish research on Medical Education covering medical, dental, nursing and allied health care professional education. The journal covers undergraduate education, postgraduate training and continuing medical education

including emerging trends and innovative models linking education, research, and health care services. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <http://www.dovepress.com/advances-in-medical-education-and-practice-journal>

Dovepress