Quality of randomised controlled trials in medical education reported between 2012 and 2013: A systematic review protocol.

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-005155
Article Type:	Protocol
Date Submitted by the Author:	28-Feb-2014
Complete List of Authors:	Tolsgaard, Martin; Copenhagen University Hospital Rigshospitalet, Capital Region of Denmark, Centre for Clinical Education and the Juliane Marie Centre Ku, Cheryl; University of Toronto and University Health Network, The Wilson Centre Woods, Nicole; University of Toronto and University Health Network, Department of Surgery and The Wilson Centre Kulasegaram, Kulamakan; University of Toronto and University Health Network, Department of Family and Community Medicine and The Wilson Centre Brydges, Ryan; University of Toronto and University Health Network, Department of Medicine and The Wilson Centre Ringsted, Charlotte; University Health Network, The Wilson Centre
Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Medical publishing and peer review
Keywords:	EDUCATION & TRAINING (see Medical Education & Training), EPIDEMIOLOGY, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS

Quality of randomised controlled trials in medical education reported between 2012 and 2013: A systematic review protocol.

Martin G. Tolsgaard, MD, PhD[1], Cheryl Ku, MSci[2], Nicole N. Woods, PhD[3], Kulamakan Mahan Kulasegaram, PhD[4], Ryan Brydges, PhD[5], Charlotte Ringsted, MD, PhD, MHPE[6].

[1] Centre for Clinical Education and the Juliane Marie Centre, Copenhagen University Hospital Rigshospitalet, Denmark.

[2] The Wilson Centre, University of Toronto and University Health Network, Canada.

[3] Department of Surgery and The Wilson Centre, University of Toronto and University Health Network, Canada

[4] Department of Family and Community Medicine and The Wilson Centre, University of Toronto and University Health Network, Canada

[5] Department of Medicine and The Wilson Centre, University of Toronto and University Health Network, Canada

[6] Department of Anesthesia and The Wilson Centre, University of Toronto and University Health Network, Canada

Running head: Quality of randomised controlled trials in medical education.

Corresponding author:
Martin G. Tolsgaard, MD, PhD
Postdoctoral research fellow, Centre for Clinical Education and the Juliane Marie Center, Copenhagen University Hospital Rigshospitalet
9 Blegdamsvej, 2100 Copenhagen-O
Denmark
Tel: +456103072
martintolsgaard@gmail.com

Word count: XXX

Keywords: Randomised controlled trials, medical education

Abstract

Introduction: Research in medical education has increased in volume over the last decades but concerns have been raised regarding the quality of trials conducted within this field. Randomised controlled trials (RCTs) involving educational interventions that are reported in biomedical journals have been criticised for insufficient conceptual, theoretical framework. RCTs published in journals dedicated to medical education, on the other hand, have been questioned regarding their methodological rigor.

The aim of this study is therefore to assess the quality of RCTs of educational interventions reported in 2012 and 2013 in journals dedicated to medical education compared to biomedical journals with respect to objective quality criteria.

Methods and analysis: RCTs published between 1 January 2012 and 31 December 2013 in English are included. The search strategy is developed by the help of experienced librarians to search online databases for key terms. All of the identified RCTs are screened based on their titles and abstracts individually by the authors and then compared in pairs to assess agreement. Data is extracted from the included RCTs by independently scoring each RCT using a data collection form. The data collection form consists of four steps. Step 1 includes confirmation of RCT eligibility; Step 2 consists of the CONSORT checklist; Step 3 consists of the MERSQI framework; Step 4 consists of a Medical Education Extension (MEdEx) to the CONSORT checklist. The MEdEx includes the following elements: Description of scientific background, explanation of rationale, quality of research questions and hypotheses, clarity in the description of the use of the intervention and control as well as interpretation of results.

Ethics and Dissemination: This review is the first to systematically examine the quality of RCTs conducted in medical education. We plan to disseminate the results through publications and presentation at relevant conferences. Ethical approval is not sought for this review.

Article summary

Strengths and Limitations of this study

Strengths:

- The first systematic review of the quality of randomised controlled trials in medical education.
- The use of duplicate, independent, and reproducible data coding of quality measures pertaining to research methodology and reporting.

Limitations:

- To provide a current state of evidence on trial quality, only studies reported from 2012 to 2013 are included in this review.
- Only articles in English are included in this systematic review.

Introduction

Medical education as a field has grown during the past twenty years. It has become a billion dollar industry accounting for about US$100 billion per year worldwide[1] and increasing awareness of linking education to patient outcomes has brought focus on evidence-based medical education.[2] The growing interest is reflected in the rise in number of publications within this area over the past several decades.[3] However, this is not unproblematic, as several scholars have warned that the medical education research lacks methodological rigor.[3] In a study of randomised controlled trials (RCTs) published between 2000 and 2003, a large proportion fell short of the criteria developed by the International Committee of Medical Journal Editors for reporting RCTs.[4] Meanwhile, some argue that judging the quality of research being performed in medical education by any 'objective' checklist is insufficient.[5] Instead, quality of medical education research should be based on how well the theoretical understanding of a problem becomes, rather than on how well a particular research methodology has been adopted. [5] Other viewpoints state that whatever method is used should comply with the highest standards of practice for that design.[6] Thus, two discourses of evaluating quality have been promoted. One is assessing quality against 'gold-standards' such as the checklists and guidelines, and another is judging the advancement of theory.

In clinical epidemiological research, RCTs take on a central role when evaluating health care interventions. Since 2000, the CONSORT group has provided guidelines to improve the transparency and rigor when reporting randomised trials within biomedicine (ref). Although the CONSORT statement does not

include recommendations for designing, conducting, and analysing trials, it indirectly affects design and conduct as transparent reporting may expose deficiencies in research if they exist[7] Furthermore, CONSORT is informed by methodological theorists and practitioners in clinical epidemiology as well as biostatistics. Assessing quality of RCTs in medical education using the CONSORT statement may, however, not capture advancement of theory. Insufficient use of a conceptual theoretical framework may lead to failure to identify the active component of training interventions and poor description of context of the study as well as trainee characteristics limit the external validity in terms of generalizability to other settings and populations. Reporting should therefore also relate the study to a relevant theoretical context to justify how it uses and advances existing theory[5] including thorough descriptions of context, educational intervention and control circumstances, and trainee characteristics.[8] However, these aspects are not assessed using the CONSORT statement and other measures to evaluate study quality within medical education research may be warranted. To further evolve our understanding of quality of RCTs conducted in medical education, we aim to explore the adherence to standardised quality criteria as well as the use of theory in recent literature. The research question of this review is:

In randomised controlled trials in medical education reported between 2012 and 2013, what characterises the quality of papers published in journals dedicated to medical education compared to papers published in biomedical journals with respect to objective quality criteria?

Methods

This systematic review is designed according to the seven-step approach recommended for conducting systematic reviews in medical education[2] and reported according to the PRISMA statement.[9]

Study eligibility

Broad inclusion criteria are used to obtain a broad range of randomised trials in medical education. Studies published between 1 January 2012 and 31 December 2013 in English are included. This period is chosen as new guidelines for reporting randomised trials were published in June 2010 and previous studies argued that the evaluation of reporting guidelines should first be evaluated 18-24 months following publication.[10-11] All research papers in medical education using randomised designs are included. Medical education research is defined as "any original research study pertaining to medical students, residents, fellows, faculty development, or continuing medical education for physicians."[12] Using this definition, studies on veterinary-, nursing-, pharmacist-, physiotherapist- and dentistry education research are not eligible. Parallel group studies, cross-over studies, non-inferiority and equivalence studies are all included whereas pseudo-randomised studies are not.

Search

The search strategy is developed by the help of experienced librarians to search MEDLINE, EMBASE, CINAHL, PsychINFO, ERIC, Web of Science, and Scopus for key terms. These terms include truncated search on random* and MeSH terms

relating to medical education (e.g. "Education, Professional"). Related domains are also included in the search to account for research not categorised under medical education (e.g. health professions education, simulation, undergraduate medical education, technology-enhanced education, clinical reasoning, skills assessment, education professional, student health occupation, internship and residency, curriculum planning, instructional method, self-directed learning etc.). The search is supplemented by adding the reference-lists of recent reviews in simulation-based medical education and with authors' records of studies published in the period of interest. The authors' records are used to refine the search strategy in an iterative way so that as many relevant randomised studies as possible are included in the online search.

Study selection

All of the identified studies are screened based on their titles and abstracts individually and compared in pairs to assess agreement so that all studies have been screened by two authors. Potential disagreement is solved by discussion until consensus is reached. If the title or abstract is insufficient for determining eligibility, the full text is reviewed. The agreement between the raters is determined using intra-class correlation coefficients (ICCs).

Data collection process

Data is extracted from included studies by duplicate and independent scoring of each study using a data collection form. The data collection form consists of four steps. Step 1 includes confirmation of study eligibility; Step 2 consists of the

CONSORT checklist; Step 3 consists of the Medical Education Research Study Quality Instrument (MERSQI) framework; Step 4 consists of a Medical Education Extension to the CONSORT statement developed by the review group.

Step 1. The first step includes confirmation of study eligibility, extraction of Study ID (created by review author) as well as name and focus of journal (medical education/biomedical).

Step 2. The dichotomous CONSORT checklist is completed by ticking off each item when either present (=1), absent (=0), or not applicable (N/A). The CONSORT statement recommends that researchers provide a scientific background for the study as well as present specific objectives and hypothesis, thoroughly describe the intervention and control conditions, randomisation procedure, data analysis and interpretation of results.

Step 3. The Medical Education Research Study Quality Instrument (MERSQI)[12] is used to provide an established measure of study quality and scores are compared in pairs and discussed until consensus. Evidence of validity of the MERSQI framework has been established in a previous study.[12] The MERSQI framework provides a measure of trial size (single- or multiple institutions), validity of assessment instruments used, and the Kirkpatrick level of outcome measures used (a taxonomy for classifying training programmes). Hence, studies of a certain size and focusing on patient outcomes would receive higher scores than single-institution studies that assess the impact of interventions on health care professionals' knowledge or behaviour in a simulated setting.

Step 4. The Medical Education Extension (MEdEx) is developed by the study group through a literature review of relevant quality research in medical

education. To further advance our understanding of the use of theory in the scientific background of the RCTs, the reporting of specific hypotheses, clarity of description of interventions[13] and controls, and the use of theory in the interpretation of the observed results, we chose to include these factors in a medical education extension (MEdEx) to the CONSORT checklist. In step 4, the following items are therefore included: 1) Scientific background, 2) Explanation of rationale, 3) Objectives or research question, 4) Hypotheses, 5) Description of the intervention and control circumstances, 6) Interpretation of results (see Appendix).

Statistical analysis

Inter-rater reliability is calculated using Intra-Class Correlation Coefficients. In the event of disagreement, the assessments will be solved by consensus. The CONSORT-scores, MERSQI-scores, and MEdEx-scores are correlated to each other, to journal impact factor, and compared across papers published in journals dedicated to medical education and biomedical journals using parametric statistics if the conditions are met.

Discussion and dissemination

In parallel with the rise in publications in medical education over the last decades, increasing interest is being paid to systematically evaluate the quality of research conducted within this field. We chose to include three different quality measures in the data collection form for this systematic review. To evaluate the quality of reporting, the CONSORT checklist is included as a measure of the

degree to which current medical education research adhere to the guidelines endorsed by the *World Association of Medical Editors, the International Committee of Medical Journal Editors (ICMJE), and the Council of Science Editors*.[14] The second quality measure includes the MERSQI framework, which has been used extensively in several recent reviews.[15-17] Although MERSQI-scores have been shown to correlate with journal impact factor,[12] it provides limited information on the use of theory or clarity in the description of interventions. Hence, to account for the use of conceptual theoretical frameworks in medical education RCTs, we include a third quality measure, the MEdEx,

We hypothesise that this review may demonstrate differences between different quality measures in RCTs reported in biomedical journals compared to those published in journals dedicated to medical education. We expect that RCTs reported in biomedical journals adhere more strictly to the CONSORT statement and use outcome measures that relate to the upper Kirkpatrick levels than RCTs reported in medical education journals. Finally, we hypothesise that RCTs published in medical education journals use theory in the rationale for their research question, methods, and in their interpretation of the results, whereas this may be missing in research published in biomedical or clinical journals.

The review results will be submitted for publication in a peer-reviewed general medical journal and will be disseminated through relevant international conferences.

The results of this review will help clarify the state of quality of education research by using common quality standards. The comparative analysis with clinical epidemiology will provide feedback for the medical education

community and contribute to raising the quality of research and improve the reporting of studies within this field.

Funding statement: This review has been supported by the Laerdal Foundation.

Competing interest statement: None disclosed.

Contributors: All authors contributed to the design of the review and all authors contributed and approved the final manuscript.

References

1) Frenk J, Chen L, Bhutta ZA, Cohen J, Crisp N, Evans T, Fineberg H, Garcia P, Ke Y, Kelley P, Kistnasamy B, Meleis A, Naylor D, Pablos-Mendez A, Reddy S, Scrimshaw S, Sepulveda J, Serwadda D, Zurayk H. Health professionals for a new century: transforming education to strengthen health systems in an interdependent world.Lancet. 2010 Dec 4;376(9756):1923-58.

2) Cook DA, West CP. Conducting systematic reviews in medical education: a stepwise approach.Med Educ. 2012 Oct;46(10):943-52. doi: 10.1111/j.1365-2923.2012.04328.x.

3) Lee K, Whelan JS, Tannery NH, Kanter SL, Peters AS. 50 years of publication in the field of medical education.Med Teach. 2013 Apr 22

4) Todres M, Stephenson A, Jones R. Medical education research remains the poor relation. BMJ 2007;335:335–5.

5) Eva KW. Med Educ. Broadening the debate about quality in medical education research. 2009 Apr;43(4):294-6. doi: 10.1111/j.1365-2923.2009.03342.x.

6) Stephenson A, Todres M, Jones R. Reply to Dornan et al.'s 'On evidence'. Med Educ. 2009 Apr;43(4):390-1. doi: 10.1111/j.1365-2923.2009.03306.x.

7) Schulz KF, Altman DG, Moher D; CONSORT Group .CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med. 2010 Jun 1;152(11):726-32. doi: 10.1059/0003-4819-152-11-201006010-00232. Epub 2010 Mar 24.

8) Frambach JM, van der Vleuten CP, Durning SJ. AM last page. Quality criteria in qualitative and quantitative research. Acad Med. 2013 Apr;88(4):552. doi: 10.1097/ACM.0b013e31828abf7f.

9) Prisma statement, available at www.prisma-statement.org. Accessed 5 January 2014.

10) Prady SL, Richmond SJ, Morton VM, et al. A Systematic evaluation of the impact of STRICTA and CONSORT recommendations on quality of reporting for acupuncture trials. PLoS One 2008;3:e1577.

11) Fuller T, Pearson M, Peters JL, Anderson R. Evaluating the impact and use of Transparent Reporting of Evaluations with Non-randomised Designs (TREND) reporting guidelines. BMJ Open. 2012 Dec 19;2(6). pii: e002073. doi: 10.1136/bmjopen-2012-002073. Print 2012.

12) Reed DA, Cook DA, Beckman TJ, Levine RB, Kern DE, Wright SM. Association between funding and quality of published medical education research. JAMA. 2007 Sep 5;298(9):1002-9.

13) Hoffmann TC, Erueti C, Glasziou PP. Poor description of non-pharmacological interventions: analysis of consecutive sample of randomised trials. BMJ. 2013 Sep 10;347:f3755. doi: 10.1136/bmj.f3755

14) Altman DG. Endorsement of the CONSORT statement by high impact medical journals: survey of instructions for authors. BMJ. 2005 May 7;330(7499):1056-7.

15) Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, Erwin PJ, Hamstra SJ.Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA. 2011 Sep 7;306(9):978-88. doi: 10.1001/jama.2011.1234. Review.

16) Cook DA, Brydges R, Zendejas B, Hamstra SJ, Hatala R. Mastery learning for health professionals using technology-enhanced simulation: a systematic review and meta-analysis. Acad Med. 2013 Aug;88(8):1178-86. doi: 10.1097/ACM.0b013e31829a365d. Review.

17) Cook DA, Levinson AJ, Garside S. Method and reporting quality in health professions education research: a systematic review. Med Educ. 2011 Mar;45(3):227-38. doi: 10.1111/j.1365-2923.2010.03890.x.

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	4
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	N/A – not subject for registration with PROSPERO as this is a methodological review.
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	6
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	8

PRISMA 2009 Checklist

Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	8
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	8
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	N/A

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	8
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	N/A
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	N/A
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	N/A
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	N/A
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	N/A
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	N/A
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	N/A
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	N/A
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	N/A
FUNDING			

PRISMA 2009 Checklist

Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	10

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

Page 2 of 2

BMJ Open

Quality of randomised controlled trials in medical education reported between 2012 and 2013: A systematic review protocol.

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-005155.R1
Article Type:	Protocol
Date Submitted by the Author:	14-Jul-2014
Complete List of Authors:	Tolsgaard, Martin; Copenhagen University Hospital Rigshospitalet, Capital Region of Denmark, Centre for Clinical Education and the Juliane Marie Centre Ku, Cheryl; University of Toronto and University Health Network, The Wilson Centre Woods, Nicole; University of Toronto and University Health Network, Department of Surgery and The Wilson Centre Kulasegaram, Kulamakan; University of Toronto and University Health Network, Department of Family and Community Medicine and The Wilson Centre Brydges, Ryan; University of Toronto and University Health Network, Department of Medicine and The Wilson Centre Ringsted, Charlotte; University Health Network, The Wilson Centre
Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Medical publishing and peer review
Keywords:	EDUCATION & TRAINING (see Medical Education & Training), EPIDEMIOLOGY, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS

Quality of randomised controlled trials in medical education reported between 2012 and 2013: A systematic review protocol.

Martin G. Tolsgaard, MD, PhD[1], Cheryl Ku, MSci[2], Nicole N. Woods, PhD[3], Kulamakan Mahan Kulasegaram, PhD[4], Ryan Brydges, PhD[5], Charlotte Ringsted, MD, PhD, MHPE[7].

[1] Centre for Clinical Education and the Juliane Marie Centre, Copenhagen University Hospital Rigshospitalet, Denmark.

[2] The Wilson Centre, University of Toronto and University Health Network, Canada.

[3] Department of Surgery and The Wilson Centre, University of Toronto and University Health Network, Canada

[4] Department of Family and Community Medicine and The Wilson Centre, University of Toronto and University Health Network, Canada

[5] Department of Medicine and The Wilson Centre, University of Toronto and University Health Network, Canada

[6] Surgical Divisions, Neuroscience, & Medical Education, Health Sciences Library, University Health Network, Toronto General Hospital, Canada.

[7] Department of Anesthesia and The Wilson Centre, University of Toronto and University Health Network, Canada

Running head: Quality of randomised controlled trials in medical education.

Conflicts of interest: None of the authors report any conflicts of interest.

Corresponding author:
Martin G. Tolsgaard, MD, PhD
Postdoctoral research fellow, Centre for Clinical Education and the Juliane Marie Center, Copenhagen University Hospital Rigshospitalet
9 Blegdamsvej, 2100 Copenhagen-O
Denmark
Tel: +456103072
martintolsgaard@gmail.com

Word count: XXX

Keywords: Randomised controlled trials, medical education

Abstract

Introduction: Research in medical education has increased in volume over the last decades but concerns have been raised regarding the quality of trials conducted within this field. Randomised controlled trials (RCTs) involving educational interventions that are reported in biomedical journals have been criticised for insufficient conceptual, theoretical framework. RCTs published in journals dedicated to medical education, on the other hand, have been questioned regarding their methodological rigor.

The aim of this study is therefore to assess the quality of RCTs of educational interventions reported in 2012 and 2013 in journals dedicated to medical education compared to biomedical journals with respect to objective quality criteria.

Methods and analysis: RCTs published between 1 January 2012 and 31 December 2013 in English are included. The search strategy is developed by the help of experienced librarians to search online databases for key terms. All of the identified RCTs are screened based on their titles and abstracts individually by the authors and then compared in pairs to assess agreement. Data is extracted from the included RCTs by independently scoring each RCT using a data collection form. The data collection form consists of four steps. Step 1 includes confirmation of RCT eligibility; Step 2 consists of the CONSORT checklist; Step 3 consists of the MERSQI framework; Step 4 consists of a Medical Education Extension (MEdEx) to the CONSORT checklist. The MEdEx includes the following elements: Description of scientific background, explanation of rationale, quality of research questions and hypotheses, clarity in the description of the use of the intervention and control as well as interpretation of results.

Ethics and Dissemination: This review is the first to systematically examine the quality of RCTs conducted in medical education. We plan to disseminate the results through publications and presentation at relevant conferences. Ethical approval is not sought for this review.

Article summary

Strengths and Limitations of this study

Strengths:

- The first systematic review of the quality of randomised controlled trials in medical education.
- The use of duplicate, independent, and reproducible data coding of quality measures pertaining to research methodology and reporting.

Limitations:

- To provide a current state of evidence on trial quality, only studies reported from 2012 to 2013 are included in this review.
- Only articles in English are included in this systematic review.

Introduction

Medical education as a field has grown during the past twenty years. It has become a billion dollar industry accounting for about US$100 billion per year worldwide[1] and increasing awareness of linking education to patient outcomes has brought focus on evidence-based medical education.[2] The growing interest is reflected in the rise in number of publications within this area over the past several decades.[3] However, this is not unproblematic, as several scholars have warned that the medical education research lacks methodological rigor.[3] In a study of randomised controlled trials (RCTs) published between 2000 and 2003, a large proportion fell short of the criteria developed by the International Committee of Medical Journal Editors for reporting RCTs.[4] Meanwhile, some argue that judging the quality of research being performed in medical education by any 'objective' checklist is insufficient.[5] Instead, quality of medical education research should be based on the advancement of our theoretical understanding, rather than on how well a particular research methodology has been adopted. [5] Other viewpoints state that whatever method is used should comply with the highest standards of practice for that design.[6] Thus, two discourses of evaluating quality have been promoted. One is assessing quality against 'gold-standards' such as the checklists and guidelines, and another is judging the advancement of theory.

In clinical epidemiological research, RCTs take on a central role when evaluating health care interventions. Since 2000, the CONSORT group has provided guidelines to improve the transparency and rigor when reporting randomised trials within biomedicine.[7] Although the CONSORT statement does not include

recommendations for designing, conducting, and analysing trials, it indirectly affects design and conduct as transparent reporting may expose deficiencies in research if they exist[7] Furthermore, CONSORT is informed by methodological theorists and practitioners in clinical epidemiology as well as biostatistics. Assessing quality of RCTs in medical education using the CONSORT statement may, however, not capture advancement of theory. Insufficient use of a conceptual theoretical framework may lead to failure to identify the active component of training interventions. Furthermore, poor description of context of the study as well as trainee characteristics limit the external validity in terms of generalizability to other settings and populations. Reporting should therefore also relate the study to a relevant theoretical context to justify how it uses and advances existing theory[5] including thorough descriptions of context, educational intervention and control circumstances, and trainee characteristics.[8] However, these aspects are not assessed using the CONSORT statement and other measures to evaluate study quality within medical education research may be warranted. To further evolve our understanding of quality of RCTs conducted in medical education, we aim to explore the adherence to standardised quality criteria as well as the use of theory in recent literature. The research question of this review is:

In randomised controlled trials in medical education reported between 2012 and 2013, what characterises the quality of papers published in journals dedicated to medical education compared to papers published in biomedical journals with respect to objective quality criteria?

Methods

This systematic review is designed according to the seven-step approach recommended for conducting systematic reviews in medical education[2] and reported according to the PRISMA statement.[9]

Study eligibility

Broad inclusion criteria are used to obtain a broad range of randomised trials in medical education. Studies published between 1 January 2012 and 31 December 2013 in English are included. This period is chosen as new guidelines for reporting randomised trials were published in June 2010 and previous studies argued that the evaluation of reporting guidelines should first be evaluated 18-24 months following publication.[10-11] All research papers in medical education using randomised designs are included. Medical education research is defined as "any original research study pertaining to medical students, residents, fellows, faculty development, or continuing medical education for physicians."[12] Using this definition, studies on veterinary-, nursing-, pharmacist-, physiotherapist- and dentistry education research are not eligible. Parallel group studies, cross-over studies, non-inferiority and equivalence studies are all included whereas pseudo-randomised studies are not.

Search

The search strategy is developed by the help of experienced librarians to search MEDLINE, EMBASE, CINAHL, PsychINFO, ERIC, Web of Science, and Scopus for key terms. These terms include truncated search on random* and MeSH terms

relating to medical education (e.g. "Education, Professional"). Related domains are also included in the search to account for research not categorised under medical education (e.g. health professions education, simulation, undergraduate medical education, technology-enhanced education, clinical reasoning, skills assessment, education professional, student health occupation, internship and residency, curriculum planning, instructional method, self-directed learning etc.). The search is supplemented by adding the reference-lists of recent reviews in simulation-based medical education and with authors' records of studies published in the period of interest. The authors' records are used to refine the search strategy in an iterative way so that as many relevant randomised studies as possible are included in the online search.

Study selection

All of the identified studies are screened based on their titles and abstracts individually and compared in pairs to assess agreement so that all studies have been screened by two authors. Potential disagreement is solved by discussion until consensus is reached. If the title or abstract is insufficient for determining eligibility, the full text is reviewed. If consensus cannot be reached by two of the co-authors, the whole author team will decide whether to include the paper or not. The agreement between the raters is determined using intra-class correlation coefficients (ICCs).

Data collection process

Data is extracted from included studies by duplicate and independent scoring of each study using a data collection form. The data collection form consists of four steps. Step 1 includes confirmation of study eligibility; Step 2 consists of the CONSORT checklist; Step 3 consists of the Medical Education Research Study Quality Instrument (MERSQI) framework; Step 4 consists of a Medical Education Extension to the CONSORT statement developed by the review group.

Step 1. The first step includes confirmation of study eligibility, extraction of Study ID (created by review author) as well as name and focus of journal (medical education/biomedical).

Step 2. The dichotomous CONSORT checklist is completed by ticking off each item when either present (=1), absent (=0), or not applicable (N/A). The CONSORT statement recommends that researchers provide a scientific background for the study as well as present specific objectives and hypothesis, thoroughly describe the intervention and control conditions, randomisation procedure, data analysis and interpretation of results.

Step 3. The Jadad Scale for reporting randomised trials[13] is used to assess the methodological rigor of the included studies. The Jadad scale consists of three items pertaining to randomisation procedure, blinding, and participant withdrawal or dropouts.

Step 4. The Medical Education Research Study Quality Instrument (MERSQI)[12] is used to provide an established measure of study quality and scores are compared in pairs and discussed until consensus. Evidence of validity of the MERSQI framework has been established in a previous study.[12] The MERSQI

framework provides a measure of trial size (single- or multiple institutions), validity of assessment instruments used, and the Kirkpatrick level of outcome measures used (a taxonomy for classifying training programmes). Hence, studies of a certain size and focusing on patient outcomes would receive higher scores than single-institution studies that assess the impact of interventions on health care professionals' knowledge or behaviour in a simulated setting.

Step 5. The Medical Education Extension (MEdEx) is developed by the study group through a literature review of relevant quality research in medical education. To further advance our understanding of the use of theory in the scientific background of the RCTs, the reporting of specific hypotheses, clarity of description of interventions[14] and controls, and the use of theory in the interpretation of the observed results, we chose to include these factors in a medical education extension (MEdEx) to the CONSORT checklist. In step 4, the following items are therefore included: 1) Scientific background,[5] 2) Explanation of rationale,[5] 3) Objectives or research question,[4, 6] 4) Hypotheses,[4, 6] 5) Description of the intervention and control circumstances,[6, 8, 14] 6) Interpretation[4, 5, 12] of results (see Appendix).

Statistical analysis

Inter-rater reliability is calculated using Intra-Class Correlation Coefficients. In the event of disagreement, the assessments will be solved by consensus. Descriptive statistics for each of the three different quality measures will be performed. Logistic regression will be performed using journal type as dependent and CONSORT-scores, MERSQI-scores, and MEdEx-scores as

predictor variables. Multiple regression using the same predictor variables and journal impact factor will also be performed to assess the relation between quality measures and journal impact factor.

Discussion and dissemination

In parallel with the rise in publications in medical education over the last decades, increasing interest is being paid to systematically evaluate the quality of research conducted within this field. We chose to include three different quality measures in the data collection form for this systematic review. To evaluate the quality of reporting, the CONSORT checklist is included as a measure of the degree to which current medical education research adhere to the guidelines endorsed by the *World Association of Medical Editors, the International Committee of Medical Journal Editors (ICMJE), and the Council of Science Editors*.[15] The second quality measure includes the MERSQI framework, which has been used extensively in several recent reviews.[16-18] Although MERSQI-scores have been shown to correlate with journal impact factor,[12] it provides limited information on the use of theory or clarity in the description of interventions. Hence, to account for the use of conceptual theoretical frameworks in medical education RCTs, we plan to include a third quality measure in terms of the MEdEx framework.

We hypothesise that this review may demonstrate differences between different quality measures in RCTs reported in biomedical journals compared to those published in journals dedicated to medical education. We expect that RCTs

reported in biomedical journals adhere more strictly to the CONSORT statement and use outcome measures that relate to the upper Kirkpatrick levels than RCTs reported in medical education journals. Finally, we hypothesise that RCTs published in medical education journals use theory in the rationale for their research question, methods, and in their interpretation of the results, whereas this may be missing in research published in biomedical or clinical journals.

The review results will be submitted for publication in a peer-reviewed general medical journal and will be disseminated through relevant international conferences.

The results of this review will help clarify the state of quality of education research by using common quality standards. The comparative analysis with clinical epidemiology will provide feedback for medical education researchers and contribute to raising the quality of research and improve the reporting of studies within this field.

Funding statement: This review has been supported by the Laerdal Foundation.

Competing interest statement: None disclosed.

Contribution statement: All authors contributed to the design of the review and all authors contributed and approved the final manuscript. Everyone listed as an author fulfils all three of the ICMJE guidelines for authorship. MGT and CR were responsible for conception of the review. CK, KMK, RB, NW, MGT, and CR were responsible for designing the review and the MEdEx. MGT drafted the first version of the protocol and CK, RB, NW, KMK, and CR were critically revised the paper.

References

1) Frenk J, Chen L, Bhutta ZA, et al. Health professionals for a new century: transforming education to strengthen health systems in an interdependent world.Lancet. 2010 Dec 4;376(9756):1923-58.

2) Cook DA, West CP. Conducting systematic reviews in medical education: a stepwise approach.Med Educ. 2012 Oct;46(10):943-52. doi: 10.1111/j.1365-2923.2012.04328.x.

3) Lee K, Whelan JS, Tannery NH, et al. 50 years of publication in the field of medical education.Med Teach. 2013 Apr 22

4) Todres M, Stephenson A, Jones R. Medical education research remains the poor relation. BMJ 2007;335:335–5.

5) Eva KW. Med Educ. Broadening the debate about quality in medical education research. 2009 Apr;43(4):294-6. doi: 10.1111/j.1365-2923.2009.03342.x.

6) Stephenson A, Todres M, Jones R. Reply to Dornan et al.'s 'On evidence'. Med Educ. 2009 Apr;43(4):390-1. doi: 10.1111/j.1365-2923.2009.03306.x.

7) Schulz KF, Altman DG, Moher D; CONSORT Group .CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med. 2010 Jun 1;152(11):726-32. doi: 10.1059/0003-4819-152-11-201006010-00232. Epub 2010 Mar 24.

8) Frambach JM, van der Vleuten CP, et al. AM last page. Quality criteria in qualitative and quantitative research. Acad Med. 2013 Apr;88(4):552. doi: 10.1097/ACM.0b013e31828abf7f.

9) Prisma statement, available at www.prisma-statement.org. Accessed 5 January 2014.

10) Prady SL, Richmond SJ, Morton VM, et al. A Systematic evaluation of the impact of STRICTA and CONSORT recommendations on quality of reporting for acupuncture trials. PLoS One 2008;3:e1577.

11) Fuller T, Pearson M, Peters JL, et al. Evaluating the impact and use of Transparent Reporting of Evaluations with Non-randomised Designs (TREND) reporting guidelines. BMJ Open. 2012 Dec 19;2(6). pii: e002073. doi: 10.1136/bmjopen-2012-002073. Print 2012.

12) Reed DA, Cook DA, Beckman TJ, et al. SM. Association between funding and quality of published medical education research. JAMA. 2007 Sep 5;298(9):1002-9.

13) Jadad AR1, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials. 1996;17(1):1-12.

14) Hoffmann TC, Erueti C, Glasziou PP. Poor description of non-pharmacological interventions: analysis of consecutive sample of randomised trials. BMJ. 2013 Sep 10;347:f3755. doi: 10.1136/bmj.f3755

15) Altman DG. Endorsement of the CONSORT statement by high impact medical journals: survey of instructions for authors. BMJ. 2005 May 7;330(7499):1056-7.

16) Cook DA, Hatala R, Brydges R, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA. 2011 Sep 7;306(9):978-88. doi: 10.1001/jama.2011.1234.

17) Cook DA, Brydges R, Zendejas B, et al. Mastery learning for health professionals using technology-enhanced simulation: a systematic review and meta-analysis. Acad Med. 2013 Aug;88(8):1178-86. doi: 10.1097/ACM.0b013e31829a365d.

18) Cook DA, Levinson AJ, Garside S. Method and reporting quality in health professions education research: a systematic review. Med Educ. 2011 Mar;45(3):227-38. doi: 10.1111/j.1365-2923.2010.03890.

Quality of randomised controlled trials in medical education reported between 2012 and 2013: A systematic review protocol.

Martin G. Tolsgaard, MD, PhD[1], Cheryl Ku, MSci[2], Nicole N. Woods, PhD[3], Kulamakan Mahan Kulasegaram, PhD[4], Ryan Brydges, PhD[5], Charlotte Ringsted, MD, PhD, MHPE[7].

[1] Centre for Clinical Education and the Juliane Marie Centre, Copenhagen University Hospital Rigshospitalet, Denmark.

[2] The Wilson Centre, University of Toronto and University Health Network, Canada.

[3] Department of Surgery and The Wilson Centre, University of Toronto and University Health Network, Canada

[4] Department of Family and Community Medicine and The Wilson Centre, University of Toronto and University Health Network, Canada

[5] Department of Medicine and The Wilson Centre, University of Toronto and University Health Network, Canada

[6] Surgical Divisions, Neuroscience, & Medical Education, Health Sciences Library, University Health Network, Toronto General Hospital, Canada.

[7] Department of Anesthesia and The Wilson Centre, University of Toronto and University Health Network, Canada

Running head: Quality of randomised controlled trials in medical education.

Conflicts of interest: None of the authors report any conflicts of interest.

Corresponding author:
Martin G. Tolsgaard, MD, PhD
Postdoctoral research fellow, Centre for Clinical Education and the Juliane Marie Center, Copenhagen University Hospital Rigshospitalet
9 Blegdamsvej, 2100 Copenhagen-O
Denmark
Tel: +456103072
martintolsgaard@gmail.com

Word count: XXX
Keywords: Randomised controlled trials, medical education

Abstract

Introduction: Research in medical education has increased in volume over the last decades but concerns have been raised regarding the quality of trials conducted within this field. Randomised controlled trials (RCTs) involving educational interventions that are reported in biomedical journals have been criticised for insufficient conceptual, theoretical framework. RCTs published in journals dedicated to medical education, on the other hand, have been questioned regarding their methodological rigor.
The aim of this study is therefore to assess the quality of RCTs of educational interventions reported in 2012 and 2013 in journals dedicated to medical education compared to biomedical journals with respect to objective quality criteria.

Methods and analysis: RCTs published between 1 January 2012 and 31 December 2013 in English are included. The search strategy is developed by the help of experienced librarians to search online databases for key terms. All of the identified RCTs are screened based on their titles and abstracts individually by the authors and then compared in pairs to assess agreement. Data is extracted from the included RCTs by independently scoring each RCT using a data collection form. The data collection form consists of four steps. Step 1 includes confirmation of RCT eligibility; Step 2 consists of the CONSORT checklist; Step 3 consists of the MERSQI framework; Step 4 consists of a Medical Education Extension (MEdEx) to the CONSORT checklist. The MEdEx includes the following elements: Description of scientific background, explanation of rationale, quality of research questions and hypotheses, clarity in the description of the use of the intervention and control as well as interpretation of results.

Ethics and Dissemination: This review is the first to systematically examine the quality of RCTs conducted in medical education. We plan to disseminate the results through publications and presentation at relevant conferences. Ethical approval is not sought for this review.

Article summary

Strengths and Limitations of this study

Strengths:

- The first systematic review of the quality of randomised controlled trials in medical education.
- The use of duplicate, independent, and reproducible data coding of quality measures pertaining to research methodology and reporting.

Limitations:

- To provide a current state of evidence on trial quality, only studies reported from 2012 to 2013 are included in this review.
- Only articles in English are included in this systematic review.

Introduction

Medical education as a field has grown during the past twenty years. It has become a billion dollar industry accounting for about US$100 billion per year worldwide[1] and increasing awareness of linking education to patient outcomes has brought focus on evidence-based medical education.[2] The growing interest is reflected in the rise in number of publications within this area over the past several decades.[3] However, this is not unproblematic, as several scholars have warned that the medical education research lacks methodological rigor.[3] In a study of randomised controlled trials (RCTs) published between 2000 and 2003, a large proportion fell short of the criteria developed by the International Committee of Medical Journal Editors for reporting RCTs.[4] Meanwhile, some argue that judging the quality of research being performed in medical education by any 'objective' checklist is insufficient.[5] Instead, quality of medical education research should be based on ~~how well the~~the advancement of our theoretical understanding~~ of a problem becomes~~, rather than on how well a particular research methodology has been adopted. [5] Other viewpoints state that whatever method is used should comply with the highest standards of practice for that design.[6] Thus, two discourses of evaluating quality have been promoted. One is assessing quality against 'gold-standards' such as the checklists and guidelines, and another is judging the advancement of theory.

In clinical epidemiological research, RCTs take on a central role when evaluating health care interventions. Since 2000, the CONSORT group has provided guidelines to improve the transparency and rigor when reporting randomised trials within biomedicine.[7] Although the CONSORT statement does not include

Formatted: Superscript

recommendations for designing, conducting, and analysing trials, it indirectly affects design and conduct as transparent reporting may expose deficiencies in research if they exist[7] Furthermore, CONSORT is informed by methodological theorists and practitioners in clinical epidemiology as well as biostatistics. Assessing quality of RCTs in medical education using the CONSORT statement may, however, not capture advancement of theory. Insufficient use of a conceptual theoretical framework may lead to failure to identify the active component of training interventions. Furthermore, ~~and~~ poor description of context of the study as well as trainee characteristics limit the external validity in terms of generalizability to other settings and populations. Reporting should therefore also relate the study to a relevant theoretical context to justify how it uses and advances existing theory[5] including thorough descriptions of context, educational intervention and control circumstances, and trainee characteristics.[8] However, these aspects are not assessed using the CONSORT statement and other measures to evaluate study quality within medical education research may be warranted. To further evolve our understanding of quality of RCTs conducted in medical education, we aim to explore the adherence to standardised quality criteria as well as the use of theory in recent literature. The research question of this review is:

In randomised controlled trials in medical education reported between 2012 and 2013, what characterises the quality of papers published in journals dedicated to medical education compared to papers published in biomedical journals with respect to objective quality criteria?

Methods

This systematic review is designed according to the seven-step approach recommended for conducting systematic reviews in medical education[2] and reported according to the PRISMA statement.[9]

Study eligibility

Broad inclusion criteria are used to obtain a broad range of randomised trials in medical education. Studies published between 1 January 2012 and 31 December 2013 in English are included. This period is chosen as new guidelines for reporting randomised trials were published in June 2010 and previous studies argued that the evaluation of reporting guidelines should first be evaluated 18-24 months following publication.[10-11] All research papers in medical education using randomised designs are included. Medical education research is defined as "any original research study pertaining to medical students, residents, fellows, faculty development, or continuing medical education for physicians."[12] Using this definition, studies on veterinary-, nursing-, pharmacist-, physiotherapist- and dentistry education research are not eligible. Parallel group studies, cross-over studies, non-inferiority and equivalence studies are all included whereas pseudo-randomised studies are not.

Search

The search strategy is developed by the help of experienced librarians to search MEDLINE, EMBASE, CINAHL, PsychINFO, ERIC, Web of Science, and Scopus for key terms. These terms include truncated search on random* and MeSH terms

relating to medical education (e.g. "Education, Professional"). Related domains are also included in the search to account for research not categorised under medical education (e.g. health professions education, simulation, undergraduate medical education, technology-enhanced education, clinical reasoning, skills assessment, education professional, student health occupation, internship and residency, curriculum planning, instructional method, self-directed learning etc.). The search is supplemented by adding the reference-lists of recent reviews in simulation-based medical education and with authors' records of studies published in the period of interest. The authors' records are used to refine the search strategy in an iterative way so that as many relevant randomised studies as possible are included in the online search.

Study selection

All of the identified studies are screened based on their titles and abstracts individually and compared in pairs to assess agreement so that all studies have been screened by two authors. Potential disagreement is solved by discussion until consensus is reached. If the title or abstract is insufficient for determining eligibility, the full text is reviewed. If consensus cannot be reached by two of the co-authors, the whole author team will decide whether to include the paper or not. The agreement between the raters is determined using intra-class correlation coefficients (ICCs).

Data collection process

Data is extracted from included studies by duplicate and independent scoring of each study using a data collection form. The data collection form consists of four steps. Step 1 includes confirmation of study eligibility; Step 2 consists of the CONSORT checklist; Step 3 consists of the Medical Education Research Study Quality Instrument (MERSQI) framework; Step 4 consists of a Medical Education Extension to the CONSORT statement developed by the review group.

Step 1. The first step includes confirmation of study eligibility, extraction of Study ID (created by review author) as well as name and focus of journal (medical education/biomedical).

Step 2. The dichotomous CONSORT checklist is completed by ticking off each item when either present (=1), absent (=0), or not applicable (N/A). The CONSORT statement recommends that researchers provide a scientific background for the study as well as present specific objectives and hypothesis, thoroughly describe the intervention and control conditions, randomisation procedure, data analysis and interpretation of results.

Step 3. The Jadad Scale for reporting randomised trials[13] is used to assess the methodological rigor of the included studies. The Jadad scale consists of three items pertaining to randomisation procedure, blinding, and participant withdrawal or dropouts.

Step ~~4~~3. The Medical Education Research Study Quality Instrument (MERSQI)[12] is used to provide an established measure of study quality and scores are compared in pairs and discussed until consensus. Evidence of validity of the MERSQI framework has been established in a previous study.[12] The MERSQI

Formatted: Superscript

framework provides a measure of trial size (single- or multiple institutions), validity of assessment instruments used, and the Kirkpatrick level of outcome measures used (a taxonomy for classifying training programmes). Hence, studies of a certain size and focusing on patient outcomes would receive higher scores than single-institution studies that assess the impact of interventions on health care professionals' knowledge or behaviour in a simulated setting.

Step 5~~4~~. The Medical Education Extension (MEdEx) is developed by the study group through a literature review of relevant quality research in medical education. To further advance our understanding of the use of theory in the scientific background of the RCTs, the reporting of specific hypotheses, clarity of description of interventions[1~~43~~] and controls, and the use of theory in the interpretation of the observed results, we chose to include these factors in a medical education extension (MEdEx) to the CONSORT checklist. In step 4, the following items are therefore included: 1) Scientific background,[5] 2) Explanation of rationale,[5] 3) Objectives or research question,[4, 6] 4) Hypotheses,[4, 6] 5) Description of the intervention and control circumstances,[6, 8, 14] 6) Interpretation[4, 5, 12] of results (see Appendix).

Formatted: Superscript
Formatted: Superscript
Formatted: Superscript
Formatted: Superscript
Formatted: Superscript
Formatted: Superscript
Formatted: Superscript

Statistical analysis

Inter-rater reliability is calculated using Intra-Class Correlation Coefficients. In the event of disagreement, the assessments will be solved by consensus. Descriptive statistics for each of the three different quality measures will be performed. Logistic regression will be performed using journal type as dependent and ~~The~~ CONSORT-scores, MERSQI-scores, and ~~and~~ MEdEx-scores as

predictor variables.~~-~~ Multiple regression using the same predictor variables and journal impact factor will also be performed to assess the relation between quality measures and journal impact factor. ~~are correlated to each other, to journal impact factor, and compared across papers published in journals dedicated to medical education and biomedical journals using parametric statistics if the conditions are met.~~

Discussion and dissemination

In parallel with the rise in publications in medical education over the last decades, increasing interest is being paid to systematically evaluate the quality of research conducted within this field. We chose to include three different quality measures in the data collection form for this systematic review. To evaluate the quality of reporting, the CONSORT checklist is included as a measure of the degree to which current medical education research adhere to the guidelines endorsed by the *World Association of Medical Editors, the International Committee of Medical Journal Editors (ICMJE), and the Council of Science Editors.*[154] The second quality measure includes the MERSQI framework, which has been used extensively in several recent reviews.[165-187] Although MERSQI-scores have been shown to correlate with journal impact factor,[12] it provides limited information on the use of theory or clarity in the description of interventions. Hence, to account for the use of conceptual theoretical frameworks in medical education RCTs, we plan to include a third quality measure in terms of~~,~~ the MEdEx framework.~~,~~

We hypothesise that this review may demonstrate differences between different quality measures in RCTs reported in biomedical journals compared to those published in journals dedicated to medical education. We expect that RCTs reported in biomedical journals adhere more strictly to the CONSORT statement and use outcome measures that relate to the upper Kirkpatrick levels than RCTs reported in medical education journals. Finally, we hypothesise that RCTs published in medical education journals use theory in the rationale for their research question, methods, and in their interpretation of the results, whereas this may be missing in research published in biomedical or clinical journals.

The review results will be submitted for publication in a peer-reviewed general medical journal and will be disseminated through relevant international conferences.

The results of this review will help clarify the state of quality of education research by using common quality standards. The comparative analysis with clinical epidemiology will provide feedback for ~~the medical education community~~medical education researchers and contribute to raising the quality of research and improve the reporting of studies within this field.

Funding statement: This review has been supported by the Laerdal Foundation.

Competing interest statement: None disclosed.

Contribution statement: All authors contributed to the design of the review and all authors contributed and approved the final manuscript. Everyone listed as an author fulfils all three of the ICMJE guidelines for authorship. MGT and CR were

responsible for conception of the review. CK, KMK, RB, NW, MGT, and CR were responsible for designing the review and the MEdEx. MGT drafted the first version of the protocol and CK, RB, NW, KMK, and CR were critically revised the paper.

~~*Contributors:* All authors contributed to the design of the review and all authors contributed and approved the final manuscript.~~

References

1) Frenk J, Chen L, Bhutta ZA, Cohen J, Crisp N, Evans T, Fineberg H, Garcia P, Ke Y, Kelley P, Kistnasamy B, Meleis A, Naylor D, Pablos-Mendez A, Reddy S, Scrimshaw S, Sepulveda J, Serwadda D, Zurayk H. Health professionals for a new century: transforming education to strengthen health systems in an interdependent world.Lancet. 2010 Dec 4;376(9756):1923-58.

2) Cook DA, West CP. Conducting systematic reviews in medical education: a stepwise approach.Med Educ. 2012 Oct;46(10):943-52. doi: 10.1111/j.1365-2923.2012.04328.x.

3) Lee K, Whelan JS, Tannery NH, Kanter SL, Peters AS. 50 years of publication in the field of medical education.Med Teach. 2013 Apr 22

4) Todres M, Stephenson A, Jones R. Medical education research remains the poor relation. BMJ 2007;335:335–5.

5) Eva KW. Med Educ. Broadening the debate about quality in medical education research. 2009 Apr;43(4):294-6. doi: 10.1111/j.1365-2923.2009.03342.x.

6) Stephenson A, Todres M, Jones R. Reply to Dornan et al.'s 'On evidence'. Med Educ. 2009 Apr;43(4):390-1. doi: 10.1111/j.1365-2923.2009.03306.x.

7) Schulz KF, Altman DG, Moher D; CONSORT Group .CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med. 2010 Jun 1;152(11):726-32. doi: 10.1059/0003-4819-152-11-201006010-00232. Epub 2010 Mar 24.

8) Frambach JM, van der Vleuten CP, Durning SJ. AM last page. Quality criteria in qualitative and quantitative research. Acad Med. 2013 Apr;88(4):552. doi: 10.1097/ACM.0b013e31828abf7f.

9) Prisma statement, available at www.prisma-statement.org. Accessed 5 January 2014.

10) Prady SL, Richmond SJ, Morton VM, et al. A Systematic evaluation of the impact of STRICTA and CONSORT recommendations on quality of reporting for acupuncture trials. PLoS One 2008;3:e1577.

11) Fuller T, Pearson M, Peters JL, Anderson R. Evaluating the impact and use of Transparent Reporting of Evaluations with Non-randomised Designs (TREND) reporting guidelines. BMJ Open. 2012 Dec 19;2(6). pii: e002073. doi: 10.1136/bmjopen-2012-002073. Print 2012.

122) Reed DA, Cook DA, Beckman TJ, Levine RB, Kern DE, Wright SM. Association between funding and quality of published medical education research. JAMA. 2007 Sep 5;298(9):1002-9.

13) Jadad AR1, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, McQuay HJ. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials. 1996;17(1):1-12.

143) Hoffmann TC, Erueti C, Glasziou PP. Poor description of non-pharmacological interventions: analysis of consecutive sample of randomised trials. BMJ. 2013 Sep 10;347:f3755. doi: 10.1136/bmj.f3755

154) Altman DG. Endorsement of the CONSORT statement by high impact medical journals: survey of instructions for authors. BMJ. 2005 May 7;330(7499):1056-7.

165) Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, Erwin PJ, Hamstra SJ.Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA. 2011 Sep 7;306(9):978-88. doi: 10.1001/jama.2011.1234. ~~Review.~~

176) Cook DA, Brydges R, Zendejas B, Hamstra SJ, Hatala R. Mastery learning for health professionals using technology-enhanced simulation: a systematic review and meta-analysis. Acad Med. 2013 Aug;88(8):1178-86. doi: 10.1097/ACM.0b013e31829a365d. ~~Review.~~

187) Cook DA, Levinson AJ, Garside S. Method and reporting quality in health professions education research: a systematic review. Med Educ. 2011 Mar;45(3):227-38. doi: 10.1111/j.1365-2923.2010.03890.x.

Appendix 1.

Extension to the CONSORT statement

The column on the left-hand side is selected CONSORT elements. The column on the right represents the coding extensions specific for this study. All of these additional items are rated on three-point scales.	
2a Scientific background and explanation of rationale	• Scientific background (maximum score = 3) include the use of 1) Educational instruments. Score = 1.0 *Example: Simulation-based medical education, use of assessment instruments with validity evidence*. 2) Educational concepts. Score = 1.0 *Example: Deliberate practice, self-directed learning*. 3) Educational theories. Score = 1.0 *Example: Cognitive load theory, developmental frameworks*. • Explanation of rationale is the clinical rationale or justification for conducting the study. Maximum score = 3. 1) Clinical background. Score = 1.5. Example: *"laparoscopic surgery has long learning curves and complications occurs more frequently with inexperienced surgeons."* 2) Justification of the use of intervention. Score = 1.5. Example: "Simulation-based training has been shown to be useful for initial training and may therefore reduce the number of complications…"
2b Specific objectives or hypotheses	• Objectives or research question (maximum score = 3) include specifications of 1) Setting and population (Each = 0.5) 2) Intervention and control (Each = 0.5) 3) Outcome measures (Each = 0.1) • Hypotheses are proposed effects or mechanisms of action. 1) Score = 3 if stated clearly as a hypothesis *Example: Our hypothesis was that…* 2) Score 1.5 if potential mechanisms of actions are stated but not explicitly called a hypothesis *Example: "Simulation-based training has previously shown improved operative performances and may therefore also reduce complications…"* 3) Score=0 if no mechanism of action is proposed or no specific hypothesis is suggested.

	Example: Effective communication is difficult. We aimed to explore if a simulated patient programme improved students' confidence in…
5 The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	• <u>Description of the use of the intervention and control</u> (maximum score = 3) include 1) Detailed description of the type of intervention and control conditions. Score = 1 *Example: Type of simulation or type of learner interaction.* 2) Detailed description of instructions/information available to participants. Score = 1 *Example: Verbal or written instructions available prior to and during the intervention and additional resources such as textbooks, web-based learning material etc.* 3) Detailed description of the supervision/ assessment/ feedback provided, the amounts available and the qualifications/training of the persons providing supervision/ assessment /feedback. Score = 1. *Example: How much feedback was provided, how was it provided, by whom and for how much time?*
Interpretations	• <u>Interpretation of results</u> (maximum score = 3) includes 1) Reported consistent with the observed results (Score = 1.5). *Example: "These significantly higher performance-scores suggest that simulation-training of junior surgeons may lead to superior performance in the OR".* 2) Integration of results and interpretation into existing educational theory. (Score = 1.5) *Example: "These results are consistent with cognitive load theory suggesting that…"*

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	4
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	N/A – not subject for registration with PROSPERO as this is a methodological review.
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	6
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	8

PRISMA 2009 Checklist

Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	8
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	8
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	N/A

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	8
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	N/A
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	N/A
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	N/A
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	N/A
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	N/A
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	N/A
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	N/A
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	N/A
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	N/A
FUNDING			

PRISMA 2009 Checklist

Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	10

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.