



## Evaluation of e-learning for medical education in low- and middle-income countries: A systematic review

Sandra Barteit<sup>a,\*</sup>, Dorota Guzek<sup>a</sup>, Albrecht Jahn<sup>a</sup>, Till Bärnighausen<sup>a,b,c</sup>, Margarida Mendes Jorge<sup>a</sup>, Florian Neuhaus<sup>a</sup>

<sup>a</sup> Heidelberg Institute of Global Health, Heidelberg University Hospital, Heidelberg University, Germany

<sup>b</sup> Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Boston, USA

<sup>c</sup> Africa Health Research Institute (AHRI), KwaZulu-Natal, South Africa

### ARTICLE INFO

#### Keywords:

e-learning  
Country-specific developments  
Evaluation methodologies  
Evaluation of CAL systems  
Post-secondary education  
Medical education

### ABSTRACT

In low- and middle-income countries (LMICs), e-learning for medical education may alleviate the burden of severe health worker shortages and deliver affordable access to high quality medical education. However, diverse challenges in infrastructure and adoption are encountered when implementing e-learning within medical education in particular. Understanding what constitutes successful e-learning is an important first step for determining its effectiveness. The objective of this study was to systematically review e-learning interventions for medical education in LMICs, focusing on their evaluation and assessment methods.

Nine databases were searched for publications from January 2007 to June 2017. We included 52 studies with a total of 12,294 participants. Most e-learning interventions were pilot studies (73%), which mainly employed summative assessments of study participants (83%) and evaluated the e-learning intervention with questionnaires (45%). Study designs, evaluation and assessment methods showed considerable variation, as did the study quality, evaluation periods, outcome and effectiveness measures. Included studies mainly utilized subjective measures and custom-built evaluation frameworks, which resulted in both low comparability and poor validity. The majority of studies self-concluded that they had had an effective e-learning intervention, thus indicating potential benefits of e-learning for LMICs. However, MERSQI and NOS ratings revealed the low quality of the studies' evidence for comparability, evaluation instrument validity, study outcomes and participant blinding. Many e-learning interventions were small-scale and conducted as short-termed pilots. More rigorous evaluation methods for e-learning implementations in LMICs are needed to understand the strengths and shortcomings of e-learning for medical education in low-resource contexts. Valid and reliable evaluations are the foundation to guide and improve e-learning interventions, increase their sustainability, alleviate shortages in health care workers and improve the quality of medical care in LMICs.

### 1. Introduction

To achieve the 2030 Sustainable Development Goals (SDGs) for health—specifically, goal 3 (Good Health and Well-Being) and goal 4

\* Corresponding author. Heidelberg University, Heidelberg Institute of Global Health, Im Neuenheimer Feld 130.3, 69120 Heidelberg, Germany  
E-mail address: [barteit@uni-heidelberg.de](mailto:barteit@uni-heidelberg.de) (S. Barteit).

(Quality Education)—low- and middle-income countries (LMICs) need effective and affordable education strategies to address critical health worker shortages (Al-Shorbaji, Majeed, & Wheeler, 2015, p. 156; Bollinger et al., 2013). This is particularly crucial considering the significant challenges most LMICs face with limited health care budgets (World Health Organization, 2014), insufficient qualified health workers (Mills, 2014), limited access to equipment and infrastructure, and a lack of continuing professional development and continuing medical education programs (Mack, Golnik, Murray, & Filipe, 2017). Training facilities and resources for health workers in LMICs are insufficient, and a substantial number of the few trained professionals are subject to “brain drain” when regularly recruited to high-income countries (Aluttis, Bishaw, & Frank, 2014). E-learning interventions for medical education could be of great benefit for targeting good health and well-being and quality education, especially in low-resource settings (Al-Shorbaji, Atun, Car, Majeed, & Wheeler, 2015; Dawson, Walker, Campbell, & Egede, 2016; Liu et al., 2016; World Health Organization, 2016a,b). Current technological progress in hardware, software and affordable internet connectivity enable a broader technology usage in low-resource settings than a few years ago (Mack et al., 2017). The question seems to be not “if”, but rather “how” technology will be most beneficial in LMICs. The Lancet Global Commissioners for the Education of Health Professionals for the 21st Century stated, “The effect of electronic learning (e-learning) is likely to be revolutionary, although how precisely it will revamp professional education remains unknown” (Frenk et al., 2010, p. 1944). One potential target for medical e-learning could be to strengthen practical training for health workers and continuing medical education, since both are currently inadequate in quantity and quality in LMICs (Al-Shorbaji, R, J, Majeed, & Wheeler, 2015, p. 156; Asongu & Le Roux, 2017; Childs, Blenkinsopp, Hall, & Walton, 2005; Mack et al., 2017; World Health Organization, 2016b). Currently, low-resource contexts are faced with a lack of infrastructure, i.e., insufficient classrooms, lecture halls and dorms, and medical educators, which leads to a very limited capacity to enroll and educate students (Frenk et al., 2010). Infrastructure upscaling needs substantial investment, and staff upscaling requires a substantial investment in training. E-learning offers the possibility to quickly scale up training without the need for simultaneous resource-intensive infrastructure upscaling. E-learning may bridge distances in rural areas for accessibility to up-to date information in medical education and reduce the need to teach theory-based, on-site classes where there are a limited number of medical teachers. E-learning could strengthen the quantity and quality of medical education in LMICs through blended learning approaches wherein e-learning covers the theoretical and procedural training, and face-to-face training covers practical skills, for example. Since e-learning is adaptable and expandable to new technological advancements, such as artificial intelligence, once implemented, the potential of e-learning could be groundbreaking for training health care workers in LMICs. Yet, sound scientific evidence on the effect of e-learning for medical education in LMICs, including its strengths and shortcomings, is limited (Frehywot et al., 2013).

## 2. Background

Current systematic reviews of e-learning for medical education mainly focus on high-income countries (Al-Shorbaji et al., 2015, p. 156; Liu et al., 2016; Nicoll, MacRury, Van Woerden, & Smyth, 2018). A systematic review by the World Health Organization (WHO) (Al-Shorbaji et al., 2015, p. 156) examined global e-learning and blended learning methods and their effect on knowledge, skills, attitudes and satisfaction in comparison to other methods (traditional, alternative e-learning/blended learning, no intervention). Al-Shorbaji et al. (2015, p. 156) highlighted that only five studies have been conducted in LMICs, making their findings potentially inapplicable for low-resource contexts because the students’ educational backgrounds may differ and have an impact on their perceptions and overall e-learning experiences.

Frehywot et al. (2013) focused their review on medical education in general in LMICs (not e-learning specifically) and summarized the current scope of literature for low-resource contexts. They described respective approaches to incorporating e-learning in a low-resource context and subsequent evaluation, and found significant variation, but a majority of blended learning approaches (Frehywot et al., 2013). The effectiveness of e-learning was mainly evaluated by comparing e-learning to other learning approaches, such as traditional teaching methods, or evaluating students and teachers’ attitudes (Frehywot et al., 2013). A more recent systematic review by Nicoll et al. (2018) (not incorporating low-resource contexts), reviewed evaluation and assessment tools as well as frameworks of technology-enhanced learning approaches (specifically, digitally mediated and blended approaches), for continuing education of healthcare professionals. Overall, the authors identified a lack of standard evaluation tools for technology-enhanced health care education programs, thus pointing to the limited validity of the evaluation results (Nicoll et al., 2018). Focusing specifically on the assessment of knowledge, Liu et al. (2016) conducted a meta-analysis comparing the effect of blended learning on health professional knowledge acquisition and found that blended learning may have a neutral to positive effect compared to non-blended instruction (pure e-learning), as well as a largely positive effect compared to no intervention in this context. Based on the results of the two studies, there are indications that e-learning may be beneficial for medical education, but there is very limited evidence for its impact in LMICs (Al-Shorbaji et al., 2015, p. 156; Frehywot et al., 2013; Frenk et al., 2010). Furthermore, the validity of e-learning evaluation results is limited due to the heterogeneity of e-learning interventions and evaluation approaches employed, making comparability and meta-synthesis difficult (Clunie, Morris, Joynes, & Pickering, 2017; Nicoll et al., 2018). Currently, there are a limited number of available studies that consider the context of LMICs, where e-learning could be an enabling factor to ameliorate the current critical scarcity of health workers. Tailored evaluation methods for assessing e-learning in the context of medical education in LMICs have not been addressed in current literature. Therefore, this review addresses the knowledge gap in evaluation methods for e-learning interventions for medical education in low-resource contexts. We systematically reviewed e-learning evaluation approaches for medical education, especially in low-resource contexts, to document their current e-learning implementation and effectiveness, and to distill lessons learned. In particular, we systematically reviewed evaluation methods of asynchronous medical e-learning since poor internet connections in these settings do not allow for synchronous e-learning and outcome measures to be used in the evaluation of e-learning interventions.

Guiding questions for this study were: (a) “How has medical e-learning been implemented in low-resource contexts?”, (b) “What e-learning evaluation and assessment methods have been employed in the included studies?”, (c) “Which outcome measures have been evaluated?”, and (d) “How has success for e-learning interventions been defined?”

### 3. Methods

We conducted a systematic literature review “to collate all evidence that fits pre-specified eligibility criteria” on e-learning in low-resourced settings (Higgins & Green, 2011, p. 1:1). Systematic reviews aim to “minimize bias by using explicit, systematic methods” (Higgins & Green, 2011, p. 1:1) and by gathering, synthesizing and reviewing study findings (Jesson, Matheson, & Lacey, 2011). We conducted the systematic review in line with the Cochrane Handbook for Systematic Reviews of Interventions (Higgins & Green, 2011). The review was undertaken in distinct stages: identifying inclusion and exclusion criteria, choosing data sources and search strategies, performing quality assessments of included studies, coding data and summarizing results, and conducting data analysis. The details of these stages are described in the following sub-sections.

Furthermore, the PICOS (population, intervention, comparison, outcome, study design) framework guided search term development for the systematic review and was the foundation for study inclusion and exclusion criteria (see Table 1) (Moher et al., 2009). The systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist for the standardized reporting of this systematic review (Moher et al., 2009) (see Appendix A).

#### 3.1. Data sources

Peer-reviewed literature and grey literature were searched in the following databases: Cochrane Library, ERIC, PsycInfo, PubMed, Science Direct, Web of Science, Scie, Cabi, Eldis, Global Health Index, Global Index Medicus, GreyLit.org, World Bank IMF, Medbox, OIASTER WORLDCAT, opengrey, Popline, SciDevNet, CRD and USAID. The search strategy and terms followed the PICOS framework (see Appendix B). Only results from the following nine databases were included since the others yielded no results: Pubmed, Google Scholar, Science Direct, Web of Science, PsycInfo, ERIC, Cochrane, Global health Index, Popline (see Fig. 1).

#### 3.2. Search strategy

Peer-reviewed studies and grey literature were included only if they were available in English; participants were medical students and/or medical doctors (excluding dentists, nurses and other health workers); e-learning was asynchronous; and conducted in LMICs according to May 2017 World Bank classifications (World Bank, 2017). The selected population was limited to medical doctors and medical students since they provide the highest level of care in national health systems and have been found to be the predominant groups that currently utilize e-learning for medical education (Frehywot et al., 2013). The search was not restricted to a particular study design since the literature pre-screening indicated a heterogeneity of study designs including randomized controlled-trials (RCTs), case studies, cross-sectional studies, and observational studies. Asynchronous e-learning was defined as content delivered offline, e.g., self-paced e-learning modules, and did not include live video conferencing or live virtual classrooms, which are examples of synchronous e-learning and not included in this review (Ruggeri, Farrington, & Brayne, 2013). We only included studies conducted after 2007 to allow for comparable e-learning technologies used in medical education. Commentaries, editorial notes, systematic reviews, meta-analyses and opinion articles were all excluded.

#### 3.3. Study selection

Of the initial 8,585 identified references, 5,407 unique references remained after duplicate removal (see Fig. 1 for PRISMA flowchart showing detailed numbers of studies available at each stage). Three researchers (SB, DG, MMJ) independently screened all titles after randomly dividing the articles into three groups. If inclusion was uncertain, papers were left for the title or abstract review stages. An additional 5052 records were excluded after title and abstract screening. References were randomly distributed for abstract screening among the aforementioned three researchers. If there was disagreement or insufficient information during abstract screening, we left the respective references for full-text screening and randomly distributed them again among the three researchers. After examining the full-texts of the remaining 355 studies, 52 eligible full-text studies were identified and retained for the review. No additional relevant articles were found after reviewing the references of the included full-text records (included studies are listed in

**Table 1**  
Inclusion criteria according to the PICOS framework.

Parameter	Description
Population	Health care professionals comprising medical students and medical doctors taking part in an educational endeavor taking place in LMICs.
Intervention	Studies that have evaluated asynchronous e-learning for medical education.
Comparison	Studies may have no comparison or comparator.
Outcome	Studies must include (i) evaluation methods used for asynchronous e-learning for medical education and (ii) outcome measures within their e-learning intervention
Studies	All study types that were published in English between 2007 and 2018 from peer-reviewed journals, conference papers. Opinion papers, commentaries, editorial notes, systematic reviews and meta-analysis were excluded.

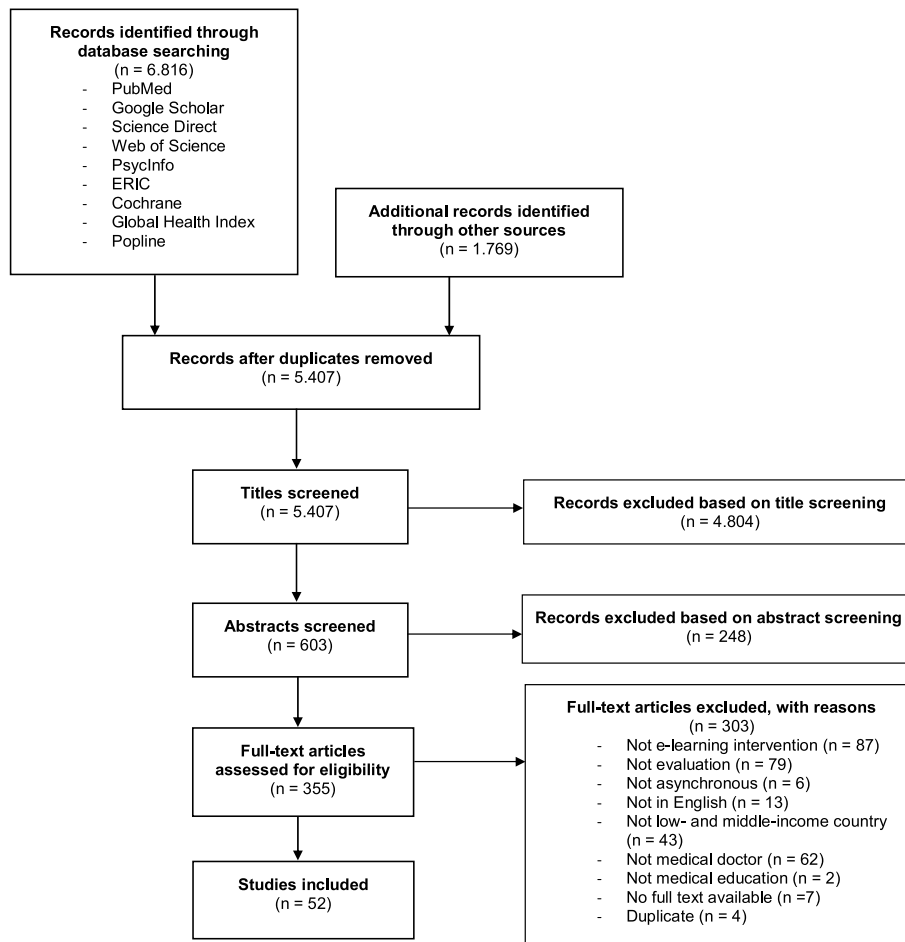


Fig. 1. PRISMA flowchart (Moher et al., 2009) showing study selection process.

Appendix C). Conflicts that arose during study selection were discussed within the research group until mutual consensus was reached.

### 3.4. Quality assessment

We reviewed and rated included studies on three quality scores: Medical Education Research Study Quality Instrument (MERSQI) (Cook & Reed, 2015; Reed, Beckman, et al., 2008), Newcastle-Ottawa Scale (NOS) (Stang, 2010), and Newcastle-Ottawa Scale-Education (NOS-E) (Cook & Reed, 2015). The objective was to use these frameworks to determine the quality of the included studies, their various study designs (credibility), and compare the quality measures in order to rate the adequacy of the e-learning intervention (see Appendix C). Domain items of the NOS-E framework include the representativeness of the intervention group, selection and comparability of the comparison group, study retention and assessment blinding. The NOS framework items include the representativeness of exposed cohort, selection of the non-exposed cohort, ascertainment of exposure, demonstration that an outcome of interest was not present at the start of the study, comparability of cohorts based on the design or analysis, outcome assessment, adequate follow-up for outcomes to occur and adequate cohort follow up. The MERSQI framework items include study design, sampling, type of data, the validity of evaluation instruments scores, data analysis, and outcome. The median was the threshold for categorizing the studies as low or high quality (Cook et al., 2011). We calculated interrater reliability based on the intraclass correlation coefficient (ICC) employing a two-way mixed effects model with mean of  $k$  ( $k = 3$ ) raters according to absolute agreement (see Appendix H for ICC ratings). ICCs and their respective confident intervals were calculated for MERSQI, NOS, and NOS-E scores and followed the ICC categorization scheme as described in Landis & Koch (1977).

### 3.5. Data coding

Included studies were randomly split among three researchers (SB, DG, MMJ) for data coding and extraction. Pertinent data from studies were coded using the qualitative data analysis computer software NVivo and later cleaned using a cloud-based spreadsheet to allow for simultaneous document edits. The researchers independently extracted data and information according to 29 aspects: author

name, title of publication, year of publication, country of intervention, institution, objective, study design, study cohort, study limitations, stage of training, type of evaluation, type of intervention, timing of evaluation, duration of evaluation, framework used, evaluation methods, course content, course delivery method, sampling size for evaluation, response rate to the evaluation, units of outcome measurement, data analysis methods, outcome measures, challenges, overall results, conclusions, success reported (by studies), outlook, and recommendations.

### 3.6. Data analysis

Studies were classified as quantitative, qualitative, or mixed methods and whether they utilized subjective or objective measures based on definitions by [David and Sutton \(2004\)](#) (see [Table 2](#)). Furthermore, evaluation methods of included studies were categorized into six groups, following the main evaluation methods for e-learning as discussed in [Phillips, McNaught, and Kennedy \(2012\)](#): (a) questionnaire, (b) document review, (c) interviews, (d) focus groups, (e) system log data and (f) knowledge testing (see [Appendix D](#) for classification criteria, and [Appendix E](#) for assessment and evaluation methods).

[Attwell \(2006\)](#) five levels were used to categorize the effectiveness of e-learning implementations.

- individual learner variables (physical characteristics, learning history, learner attitude, and learner motivation);
- learning environment variables (immediate learning environment, organizational or institutional environment, and the subject environment);
- contextual variables (socio-economic factors, political context, cultural background, and geographic location);
- technology variables (hardware, software, connectivity, media, and mode of delivery);
- pedagogic variables (level and nature of learner support systems, accessibility issues, methodologies, flexibility, learner autonomy, selection and recruitment, assessment and examination, accreditation, and certification).

An intervention was determined to be effective based on the study's description and was categorized as (a) successful, (b) not successful or (c) non-applicable (no report on the success of the intervention).

Furthermore, studies were classified as cross-sectional if they gathered information from participants at one point in time (i.e., surveys), or experimental if they included non-controlled studies, one self-controlled group, or two-groups or more with a control and intervention group. Finally, the included studies were categorized as pilots or a fully implemented studies.

## 4. Results

### 4.1. Synthesis of findings

A total of 52 articles published from January 2007 to June 2017 were included in this review (see summary of extracted data in

**Table 2**  
Overview of terminology definitions.

Term	Definition
assessment	The systematic collection of individual learner achievements, also of groups, with regard to knowledge, ability and advancement, usually measured with marks or percentages ( <a href="#">Gibbs, Brigden, &amp; Hellenberg, 2006</a> )
asynchronous	Content delivery that occurs at a different time than the student receives it and thus may be carried out while the student is offline ( <a href="#">Wagner, Hassanein, &amp; Head, 2008</a> )
blended learning	E-learning supplemented by face-to-face training ( <a href="#">Moore et al., 2011</a> ), or vice versa
content type	Category of media used for an e-learning intervention (e.g. PowerPoints, videos, audio)
delivery method	E-learning delivery techniques and characteristics (e.g., via mobile devices, computer-lab in school)
diagnostic assessment	Pre-assessment to determine user knowledge and skills prior to the intervention ( <a href="#">Black, 1993</a> )
distance learning	Provision of access to learning for those who are located geographically separated from the instructor ( <a href="#">Moore et al., 2011</a> )
document review	Evaluation method using audit and analysis of design documents or documents produced by learners as they use an e-learning environment ( <a href="#">Phillips et al., 2012</a> )
e-learning	The use of computer technology to deliver training including technology-supported learning either online, offline or both ( <a href="#">Moore et al., 2011</a> )
evaluation	The measurement of educational processes, content and adequacy for interventions, activities, programs, curricula ( <a href="#">Gibbs et al., 2006</a> )
evaluation timing	Point in time when the e-learning evaluation is completed (prior, during, or post intervention)
formative assessment	The continuous, often informal communication between teacher and learner/s to provide feedback on strengths, weaknesses and potential improvements, generally more descriptive and qualitative ( <a href="#">Gibbs et al., 2006</a> )
interactive content	E-learning content that makes use of multimedia content, such as text, graphics, audio and/or video and provides options for navigation, interaction and communication ( <a href="#">Hofstetter, 2001</a> )
podcasts	"A video and/or audio file made available in digital format for download over the Internet" ( <a href="#">Merriam-Webster, 2018</a> )
static content	E-learning content with limited interactivity and content engagement for learners ( <a href="#">Zhang &amp; Zhou, 2003</a> )
summative assessment	The end-point of an assessment, often formal, whose character is more numeric and quantitative ( <a href="#">Gibbs et al., 2006</a> )
units of measurement	The manner or scale in which the intervention outcome is measured, such as a Likert-scale

Appendix F). The number of participants within the studies ranged from 3 (Abdollahi, Sheikhabahaei, Meysamie, Bakhshandeh, & Hosseinzadeh, 2014) to 2,897 (Kotb, Elmahdy, Khalifa, El-Deen, & Lotfi, 2015) respectively. The duration of the interventions ranged from five days (Florescu, Mullen, Nguyen, Sanders, & Vu, 2015) to six months, (Walsh & De Villiers, 2015), and the average e-learning intervention was 12 weeks. Detailed study characteristics and features of the employed e-learning system and evaluation methods are shown in Table 3 and Table 4 respectively, and are presented in more detail in the following paragraphs.

#### 4.2. Implementation of E-learning interventions: results of individual studies

##### 4.2.1. Overview of included studies

The majority of the 52 included studies focused on e-learning for medical undergraduates (65%), while others targeted continuing medical education and professional development for physicians (see Table 3 for detailed study characteristics). Almost all of the studies (92%) made use of a summative assessment method, which was most often used in combination with a diagnostic (baseline) assessment prior to the intervention (58%). A few studies (15%) also made use of formative assessments during the intervention. The included studies covered a total of 25 countries, whereby almost two-thirds of the studies were from upper-middle-income countries (UMICs) (63%), followed by studies from lower-middle-income countries (33%) and only one study from a low-income country (LIC). We found that the majority of studies (73%) involved e-learning in pilots, and only 31% of the studies had progressed beyond the pilot stage. Most e-learning approaches were tailored for online access (delivery mode not specified) (73%), while a small number specified mobile online access, such as smartphones, (10%), and a few of those were based on offline approaches (8%).

##### 4.2.2. Study design

Cross-sectional study designs with no control group were the most common (52%). Experimental designs were mainly implemented with two or three groups for the intervention and control (35%) or a single self-controlled group (13%). Non-random assignment to intervention or control group was predominant (79%), although a few studies did randomize group assignments (21%). Most studies (75%) had used quantitative methods, while a few studies used mixed-methods (17%) or a fully qualitative approach (8%). Most of the quantitative studies (79%) were based on subjective measures. In general, the effectiveness of the quantitative methods was measured

**Table 3**

Study characteristics of 52 included studies describing an e-learning program evaluation that was implemented in a low-and middle-income country (Jan 2007–June 2017).

Study Characteristics	Number of Participants, n (%)	Number of Interventions, n (%)
Total	12,294 (100%)	52 (100%)
<b>Study design</b>		
Cross-sectional or not controlled	7,172 (58%)	27 (52%)
Experimental (1 group, self-controlled)	2,388 (19%)	7 (13%)
Experimental (2 or 3 groups, Intervention and control)	2,734 (23%)	18 (35%)
Randomized	938 (8%)	11 (22%)
Non-randomized	11,356 (92%)	41 (78%)
Objective <sup>b</sup>	519 (4%)	11 (21%)
Subjective <sup>b</sup>	2,779 (23%)	19 (37%)
Mix of subjective and objective <sup>b</sup>	8,996 (73%)	22 (42%)
Qualitative	3,106 (25%)	4 (8%)
Quantitative	8,844 (72%)	39 (75%)
Mixed methods	344 (3%)	9 (17%)
<b>Country income groups</b>		
Low-income country	N/A <sup>a</sup>	1 (2%)
Lower-middle-income country	4,200 (34%)	17 (33%)
Upper-middle-income country	7,890 (64%)	33 (63%)
All of the above combined	204 (2%)	1 (2%)
<b>Geographical location<sup>b</sup></b>		
Africa	6,498 (51%)	12 (22%)
Asia	1,726 (14%)	24 (44%)
Europe	204 (2%)	3 (6%)
South America	4,322 (33%)	15 (28%)
<b>Cohort/target group<sup>b</sup></b>		
Medical students (undergraduate)	9,311 (72%)	34 (61%)
Medical students and Residents	71 (1%)	1 (2%)
Medical students, Residents and Physicians	798 (6%)	2 (4%)
Residents (postgraduate)		
Residents and Physicians	310 (2%)	5 (9%)
Physicians (work-experienced)	276 (2%)	3 (5%)
	2,185 (17%)	7 (13%)

Characteristics described in more than one study are indicated. The number of participants and number of e-learning interventions are depicted in absolute numbers and the percentage is based on the absolute numbers of participants and studies for each individual characteristic.

<sup>a</sup> N/A: number of participants not explicitly stated in the study.

<sup>b</sup> More than one option possible for each study.

**Table 4**

Features of the e-learning system, and characteristics of assessment and evaluation methods of 52 included studies on LMICs e-learning program evaluation (Jan 2007–June 2017).

<b>Features of the e-learning system</b>		
<b>Characteristics</b>	<b>Number of Participants, n (%)</b>	<b>Number of E-learning interventions, n (%)</b>
<b>Stage of implementation<sup>a</sup></b>		
pilot	4,690 (34%)	38 (72%)
implemented	9,611 (66%)	16 (28%)
<b>E-Learning access</b>		
online	9,458 (77%)	38 (73%)
offline	300 (2%)	4 (8%)
online and offline	2,536 (20%)	10 (20%)
<b>Format of e-learning delivery</b>		
mobile-based	334 (3%)	5 (9%)
computer-based	9,301 (76%)	39 (75%)
mobile- and computer-based	2,236 (18%)	3 (6%)
unspecified	423 (3%)	5 (10%)
<b>Assessment and evaluation</b>		
<b>Characteristics</b>	<b>Number of Participants, n (%)</b>	<b>Number of E-learning interventions, n (%)</b>
<b>Assessment and evaluation method<sup>a, b</sup></b>		
questionnaire	12,140 (44,6%)	46 (47%)
knowledge testing	8,732 (32%)	35 (35%)
document review	43 (0,2%)	1 (1%)
interviews	110 (0,4%)	3 (3%)
focus groups	228 (0,8%)	6 (6%)
system log data	6,001 (22%)	8 (8%)
<b>Assessment type<sup>a</sup></b>		
Diagnostic	2,187 (14%)	30 (35%)
Formative	2,569 (17%)	8 (9%)
Summative	10,339 (69%)	48 (56%)
<b>Number of assessment and evaluation methods employed</b>		
1 method type	1,133 (9%)	19 (37%)
2 method types	4,495 (37%)	12 (23%)
3 method types	5,933 (48%)	16 (31%)
4 method types	438 (4%)	4 (8%)
5 method types	295 (2%)	1 (2%)
<b>Quality score<sup>c</sup></b>		
MERSQI $\geq 10.625$	2,025 (17%)	26 (50%)
MERSQI $< 10.625$	10,269 (83%)	26 (50%)
NOS-E $\geq 2.5$	6,550 (53%)	33 (64%)
NOS-E $< 2.5$	5,774 (47%)	19 (36%)
NOS $\geq 5$	5,408 (44%)	32 (62%)
NOS $< 5$	6,886 (56%)	20 (38%)
<b>Learning outcomes (assessment)</b>		
<b>Outcome measurement<sup>a, d</sup></b>		
1. knowledge	6,618	31 (16%)
2. satisfaction	6,613	20 (10%)
3. usage	8,680	17 (9%)
4. perceptions	7,681	14 (7%)
5. attitude	963	10 (5%)
6. opinion	2,127	10 (5%)
7. skills	335	9 (5%)
8. usefulness	2,348	8 (4%)
9. demographics	2,513	7 (4%)
10. ease of use	333	4 (2%)
11. usability	156	4 (2%)
12. acceptance	289	4 (2%)
13. experience	330	3 (2%)
14. self-perceived confidence	48	3 (2%)
15. learning relevance	243	3 (2%)
<b>Response rate (to assessment)</b>		
100%	665 (5%)	14 (27%)
80–99%	2,809 (23%)	16 (31%)
50–79%	2,672 (22%)	7 (13%)
below 50%	3,806 (31%)	5 (10%)
NA <sup>e</sup>	2,342 (19%)	10 (19%)
<b>Success reported</b>		
Yes	10,349 (84%)	45 (87%)

(continued on next page)

Table 4 (continued)

No	1,184 (10%)	5 (10%)
NA <sup>e</sup>	761 (6%)	2 (4%)

Characteristics described in more than one study are indicated. The number of participants and the number of e-learning interventions are depicted in absolute numbers and the percentage is based on the absolute numbers of participants and studies for each individual characteristic.

<sup>a</sup> More than one option possible for each study.

<sup>b</sup> Participant numbers separately counted for each option.

<sup>c</sup> Scores between raters averaged.

<sup>d</sup> Measurement outcomes only included if study  $n \geq 3$ , for a detailed list see Appendix I.

<sup>e</sup> N/A: number of participants not explicitly stated in study.

by study participants' subjective self-reported data (perception) (40%), objective data (17%) or a mix of both (25%).

#### 4.2.3. E-learning contents

The scope of the content covered 35 preclinical and clinical subjects (see Appendix F for details). The most frequently taught e-learning topic was pediatrics (10%), followed by anatomy (6%) and emergency medicine (6%).

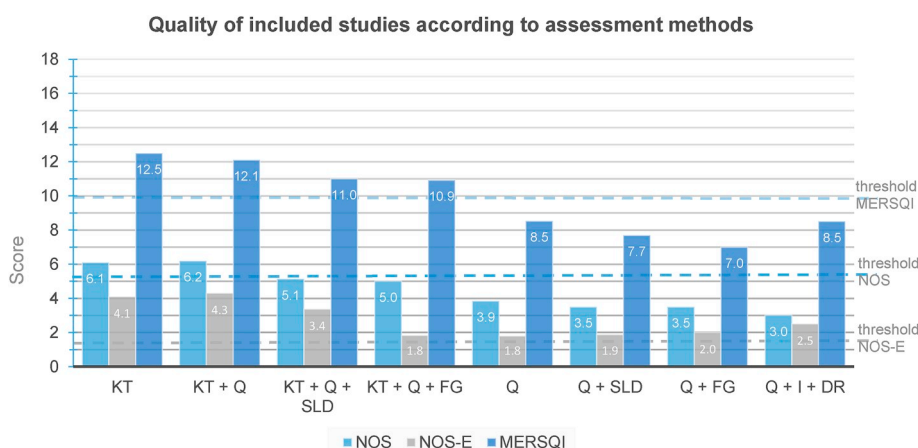
In several studies, static content was offered in PowerPoint presentations alongside video content visualizing medical procedures and examinations, and audio-visual content, such as enhanced podcasts showing presentation slides with recorded audio or animation. Content included digital textbooks, scientific papers, lecture notes, original drawings, and links to relevant websites. The interactive e-learning segments covered assignments with case scenarios and questions, self-tests with peer feedback, quizzes including photo quizzes, crosswords, matching puzzles, online-based question and answer sessions, as well as recorded video-conferencing sessions made available as asynchronous e-learning materials.

#### 4.3. Data analysis

The data analysis of e-learning studies generally included inferential statistics (58%), while 42% of studies included descriptive statistics. Inferential statistics only were reported in just under one-third of the studies (29%) or in combination with descriptive statistics (30%); a solely descriptive approach was rarely reported (12%). Some studies did not describe their data analysis at all or used no statistical tests (29%). The most frequent analysis method was a *t*-test (46%), reported as a two-sample *t*-test (23%), a paired two-sample *t*-test (15%), or an unpaired two-sample *t*-test (8%). Studies mentioned using a *t*-test to compare the performance of two groups to determine the significance of the effect of the intervention (Bediang et al., 2013), or to compare pre-test to post-test scores (Chang et al., 2012; Vieira et al., 2017), grades (Gazibara et al., 2015), cognitive load on both learning environments (Grangeia et al., 2016) or mean scores of experimental and control groups (Tomaz, Mamede, Filho, Roriz Filho, & van der Molen, 2015). A chi-square test was used in a few studies (12%), while other data analysis methods included Wilcoxon-Mann-Whitney test (6%), Levene's test (4%), linear regression (4%), variance analysis (4%), Spearman's rank-order correlation (2%), Bonferroni's correction (2%), and the Kruskal-Wallis test (2%).

##### 4.3.1. Study quality

The quality scores of studies varied with the highest MERSQI score of 15.5/18 points, and the lowest 5.75 (MERSQI scores:



KT = knowledge testing; Q = questionnaire; SLD = system log data; FG = focus group; I = interview; DR = document review

**Fig. 2.** Study quality according to NOS, NOS-E and MERSQI quality assessment tools of the various evaluation methods for e-learning intervention in the included studies.



**Table 5**  
Outcome measures categorized by the evaluation and assessment methods.

Outcome measures by evaluation and assessment methods		
Methods	Outcome Measure	Number of e-learning interventions (n)
<b>Evaluation methods</b>		
document review	content fit	1
focus group	needs	2
	utility	1
	opinion	2
	usage	1
	usefulness	1
interview	learning relevance	1
	learning relevance	2
	interaction	1
	attitude	1
	perspective	1
system log data	improvements	1
	usage	6
	technical problems	1
	demographics	1
	participation	2
questionnaire	access date	1
	activity	1
	performance	1
	access date	1
	access speed	1
	access to technology	4
	activity	1
	attitude	6
	cognitive load	1
	computer literacy	1
	content fit	4
	demographics	5
	ease of questions	1
	ease of use	5
	educational environment	1
	effectiveness	3
	efficacy	1
	efficiency	1
	expectations	1
	experience	3
	improvements	1
	infrastructure	1
	interaction	3
	learning approaches	1
	learning relevance	2
	motivation	3
	opinion	7
	participation	2
	perceived benefits	1
	perceived learning gains	1
perceptions	10	
performance	2	
playability	1	
preferences	1	
quality of learning materials	1	
recommendable	1	
satisfaction	11	
self-evaluation	1	
self-regulation	1	
service quality	1	
suitability	1	
teaching quality	1	
teaching strategies	1	
technical problems	1	
technology acceptance	1	
time spent to perform task	1	
usability	3	
usage	9	
usefulness	3	
user-friendliness	1	

(continued on next page)

Table 5 (continued)

Outcome measures by evaluation and assessment methods		
Methods	Outcome Measure	Number of e-learning interventions (n)
<b>Assessment methods</b>		
pre-post-testing	skills	10
	knowledge	29
system log data	knowledge	1
questionnaire	affected behavior	1
	self-assessment	1
	self-perceived skills	1
	self-reported behavior	1

SD = 2.26, FX = 10.67) (see Fig. 2 and Appendix C and J for details). In total, 50% of the studies were below the median MERSQI score of 10.625. The highest NOS score was 8.5/9 points, and the lowest 2.5/9 (NOS scores: SD = 1.57, FX = 5.19), and 38% of the studies had a score below the median. For the NOS-E quality scale for education, three studies achieved the maximum rating of six points and one study the minimum score of 0.5 points (SD = 1.54, FX = 3.22), and 36% of studies were below the NOS-E median of 2.5 points.

Studies that employed knowledge testing for assessment received the higher average quality ratings (MERSQI FX = 12.1, NOS FX = 6.2, and NOS-E FX = 4.3), as well as studies that exclusively assessed knowledge (MERSQI FX = 12.5; NOS FX = 6.1; NOS-E FX = 4.1). The lowest average ratings were found for studies that used only a questionnaire for evaluation of the e-learning (MERSQI FX = 8.53; NOS FX = 3.85; NOS-E FX = 1.8), a questionnaire in combination with a system log data for usage (MERSQI FX = 7.69; NOS FX = 3.5; NOS-E FX = 1.9), or a questionnaire in combination with a focus group (MERSQI FX = 7; NOS FX = 3.5; NOS-E FX = 2).

Overall, NOS-E and NOS scores showed similar distributions since 63% of studies were above the NOS-E and 62% above the NOS quality threshold. ICC estimates and their 95% confident intervals were calculated based on a single-measurement, absolute-agreement, two-way mixed-effects model, since multiple raters assessed the study quality. NOS-E scores (ICC = 0.675, CI 0.49–0.79) were shown to be more reliable with a fair to good agreement (Landis & Koch, 1977), as compared to NOS (ICC = 0.37, CI 0.11–0.58) and MERSQI (ICC = 0.25, CI -0.01–0.49), which was indicative of poor reliability below 0.4 (see Appendix H for detailed ICC estimates).

#### 4.3.2. Technology and e-learning delivery instruments

Most study participants accessed e-learning with computers (75%), although a minority of studies reported on the usage of mobile devices (10%) or both computers and mobile devices (5%). Some studies did not report the e-learning delivery mode (10%). A few studies provided electronic devices for e-learning access to study participants (8%); however, in most studies (92%), students used their own computer or mobile device.

Moodle was the predominant learning management system used in studies (29%), although not all authors were explicit about their learning management system. Other learning management system software applications included social media services such as Facebook (Sabouni et al., 2017), YouTube (Sabouni et al., 2017; Vieira et al., 2017), iTunes U (Florescu et al., 2015), GoogleDocs (Bandhu & Raje, 2014), GoogleSites (Gaikwad & Tankhiwale, 2014) and Blackboard (Martínez & Tuesca, 2014). Some studies custom built web-based e-learning platforms (Adanu et al., 2010; Davids, Chikte, & Halperin, 2011; de Sena, Fabricio, Lopes, & da Silva, 2013; Fontelo, Faustorilla, Gavino, & Marcelo, 2012; Kotb et al., 2015; Varghese, Faith, & Jacob, 2012). Self-developed mobile device apps were also used (Kulier et al., 2012; Martínez & Tuesca, 2014).

### 4.4. Assessment and evaluation methods

#### 4.4.1. Assessment methods

Assessment of included studies largely followed a summative design (92%) (see Table 3, Table 5). The predominant assessment method in studies that reported their assessment method ( $n = 43$ ) was knowledge testing (58%) and skills assessment (33%) using a pre- and post-test design (19%) (see Appendix E for detailed description of assessment and evaluation methods and outcome measures). Overall, the majority of studies (71%) reported a 50% and above response rate for assessments.

#### 4.4.2. Evaluation methods

Most of the studies (90%) used up to three methods (e.g., one method, two methods combined or three methods combined) to evaluate the e-learning implementation, and a minority used four or more evaluation methods (10%). The principal evaluation method was a questionnaire and most questionnaires were either designed by the study team or not specified.

A few studies based their questionnaires on validated frameworks, such as the System Usability Scale (Davids et al., 2011; Diehl, de Souza, Gordan, Esteves, & Coelho, 2015; Pusponogoro, Soebadi, & Surya, 2015), Berlin questionnaire (Sabouni et al., 2017), the course experience questionnaire (CEQ), a teaching performance indicator for higher education (Abdelhai et al., 2012; Ramsden, 1991), or adapted other validated questionnaires (Diehl et al., 2015; Grangeia et al., 2016; Pusponogoro et al., 2015; Rohwer, Young, & Van Schalkwyk, 2013). Other evaluation methods included the following: document review (Rohwer et al., 2013), interviews (De Villiers & Walsh, 2015; Rohwer et al., 2013; Wu, Wang, Johnson, & Grotzer, 2014), focus groups (De Villiers & Walsh, 2015; Gaikwad & Tankhiwale, 2014; Kaliyadan et al., 2010; Newberry, Kennedy, & Greene, 2016; Rajapakse, Fernando, Rubasinghe, & Gurusingham,

2009), or system log data (Heydarpour et al., 2013; Kotb et al., 2015; Masters & Gibbs, 2007; Patil et al., 2016; Pino, Mora, Diaz, Guarnizo, & Jaimes, 2017; Seluakumar, Jusof, Ismail, & Husain, 2011; Kukulja Taradi et al., 2008; Walsh & De Villiers, 2015).

#### 4.4.3. Outcome measures

Outcome measures were defined following evaluation methods. Respectively, we assigned outcome measures for all included studies and the most frequent outcome measures, in order of frequency, included: knowledge ( $n = 31$ ), satisfaction ( $n = 20$ ), usage ( $n = 17$ ), perceptions ( $n = 14$ ), attitude ( $n = 10$ ), opinion ( $n = 10$ ), skills (studies  $n = 9$ ), usefulness ( $n = 8$ ), and demographics ( $n = 7$ ) (see Table 4, Table 5, and Appendix E for detailed overview of outcome measures).

Knowledge assessment was mainly conducted using multiple-choice questions (33%) (Abdelhai et al., 2012; Chang et al., 2012; de Sena et al., 2013; Farokhi, Zarifsanaiy, Haghghi, & Mehrabi, 2016; Gaikwad & Tankhiwale, 2014; Hemmati et al., 2013; Kaliyadan et al., 2010; Khoshbaten et al., 2014; Kulatunga, Marasinghe, Karunathilake, & Dissanayake, 2013; Kulier et al., 2012; Martínez & Tiesca, 2014; Muñoz, Ortiz, González, López, & Blobel, 2010; Pino et al., 2017; Sabouni et al., 2017; Sánchez-Mendiola et al., 2012; Tomaz et al., 2015; Vieira et al., 2017), or multiple-choice questions mixed with other question types, e.g., extended matching questions (Abdelhai et al., 2012). The number of multiple-choice questions ranged from 5 to 108 in de Sena et al. (2013) and Chang et al. (2012), respectively. Other studies using multiple-choice questions had 10 questions (Gaikwad & Tankhiwale, 2014; Muñoz et al., 2010), 15 questions (Sabouni et al., 2017), 20 questions (Hemmati et al., 2013; Kulatunga et al., 2013; Puspongoro et al., 2015), 21 questions (Khoshbaten et al., 2014), 24 questions (Farokhi et al., 2016), 30 questions (Martínez & Tiesca, 2014; Mojtahedzadeh et al., 2014), and 70 questions (Sánchez-Mendiola et al., 2012). Most studies followed a pre- and post-design to assess knowledge ( $n = 21$ ), while a few studies assessed knowledge solely as summative ( $n = 8$ ).

Skills were assessed in 10 studies (19%) (Abdollahi et al., 2014; Bustam, Azhar, Veriah, Arumugam, & Loch, 2014; Heydarpour et al., 2013; Khoshbaten et al., 2014; Kulier et al., 2012; O'Donovan, Ahn, Nelson, Kagan, & Burke, 2016; Sabouni et al., 2017; Sánchez-Mendiola et al., 2012; Seluakumar et al., 2011; Tomaz et al., 2015). Medical skills were assessed with objectively structured clinical examinations (8%) (Grangeia et al., 2016; Kulier et al., 2012; O'Donovan et al., 2016; Tomaz et al., 2015), by blinded experts rating the student's echocardiography images (6%) (Bustam et al., 2014; Florescu et al., 2015), hands-on tissue ratings by an inter-observer agreement according to the Gleason grade score (Abdollahi et al., 2014), or by blinded experts grading a rhomboid flap skin model (de Sena et al., 2013). Skills were assessed after the e-learning intervention (6%) (Bediang et al., 2013; Bustam et al., 2014) or in a pre- and post-test design (13%). Evidence-based practice was evaluated in one study with a validated outcome measure based on a five-point scale covering seven domains: knowledge and learning materials, learner support, general relationships and support, institutional focus on evidence-based medicine (EBM), education training and supervision, EBM application opportunities, and affirmation of EBM environment (Kulier et al., 2012).

The outcome measures of perception, opinion, attitude, demographics, ease of use, satisfaction and learning relevance were predominantly evaluated with questionnaires. Other methods were used sporadically, e.g., focus groups (opinion) (De Villiers & Walsh, 2015; Gaikwad & Tankhiwale, 2014), interviews (attitude) (De Villiers & Walsh, 2015), learning relevance (De Villiers & Walsh, 2015; Wu et al., 2014) and system log data (demographics) (Hornos et al., 2013). Student satisfaction was predominantly evaluated subjectively (21%).

Insights on actual usage were derived with system log data (Heydarpour et al., 2013; Hornos et al., 2013; Kotb et al., 2015; Kukulja; Masters & Gibbs, 2007; Patil et al., 2016; Pino et al., 2017; Seluakumar et al., 2011; Kukulja Taradi et al., 2008; Walsh & De Villiers, 2015), or based on the number of podcast downloads (Heydarpour et al., 2013). Study participants' self-reported usage was documented with questionnaires (Bandhu & Raju, 2014; Carvalho et al., 2012; Kotb et al., 2015; Varghese et al., 2012). In one study, the participants' e-learning usage was discussed in a focus group (De Villiers & Walsh, 2015).

Usability was evaluated using the System Usability Scale in three studies (Diehl et al., 2015; Puspongoro et al., 2015; Rohwer et al., 2013), and one study extended the evaluation by giving a score based on video recordings of the participants' e-learning sessions (Diehl et al., 2015). Another study queried participants with questionnaires about the ease of navigation, esthetic quality of interface and information readability (Muñoz et al., 2010). To evaluate the students' experiences, the CEQ was used as a performance indicator of teaching quality in higher education (Abdelhai et al., 2012).

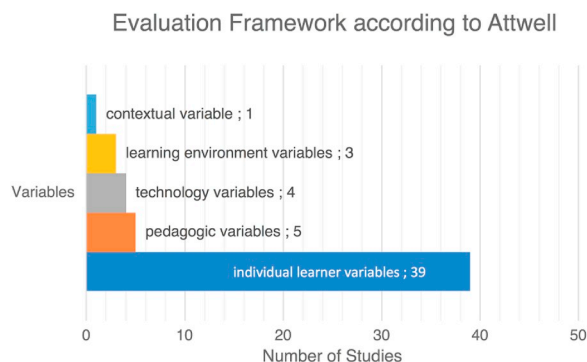


Fig. 3. Outcome measures of included studies according to the variables of Attwell's evaluation framework.

Cognitive load was evaluated with a questionnaire using a seven-point Likert scale from minimum to maximum (Grangeia et al., 2016). Examples of questions included, “How much mental effort was necessary to accomplish virtual rounds?” and “How much mental effort was necessary to accomplish real medical rounds?” These questions were further refined into intrinsic (“How difficult was the online course content?”), extrinsic (“How difficult it was to navigate on our platform?”), and germane load (“How much concentration did you maintain during our course activities?”) (Grangeia et al., 2016).

Most evaluation variables corresponded to individual learner variables ( $n = 47$ ), pedagogic variables ( $n = 5$ ), technology variables ( $n = 2$ ), learning environment variables ( $n = 2$ ), and contextual variables ( $n = 1$ ) (see Fig. 3 and Appendix G for more details) (Attwell, 2006).

#### 4.5. Effectiveness of E-learning

The majority of studies (both quantitative and qualitative) reported effective e-learning (83%) with effectiveness predominantly based on a subjective measure (49%) (Abdelhai et al., 2012; Adanu et al., 2010; Bandhu & Raje, 2014; Davids et al., 2011; de Fátima Wardenski et al., 2012; De Villiers & Walsh, 2015; Diehl et al., 2015; Fontelo et al., 2012; Gazibara et al., 2015; Grangeia et al., 2016; Hemmati et al., 2013; Hornos et al., 2013; Karakuş, Duran, Yavuz, Altintop, & Çalişkan, 2014; Kotb et al., 2015; Kulatunga et al., 2013; Rajapakse et al., 2009; Reis, Ikari, Taha-Neto, Gugliotta, & Denardi, 2015; Sánchez-Mendiola et al., 2012; Tayyab & Naseem, 2013; Varghese et al., 2012; Walsh & De Villiers, 2015; de Fátima Wardenski et al., 2012). Just over one-fifth of the studies (21%) based e-learning effectiveness on objective measures (Abdelhai et al., 2012; Bediang et al., 2013; Bustam et al., 2014; Chang et al., 2012; Farokhi et al., 2016; Kulier et al., 2012; O’Donovan et al., 2016; Suebnukarn & Haddawy, 2007; Vieira et al., 2017). A number of studies (30%) reported effectiveness based on subjective and objective measures (de Sena et al., 2013; Tomaz et al., 2015; Gaikwad & Tankhiwale, 2014; Mojtahedzadeh et al., 2014; Muñoz et al., 2010; Newberry et al., 2016; Pino et al., 2017; Puspongoro et al., 2015; Rohwer et al., 2013; Sabouni et al., 2017; Seluakumaran et al., 2011; Silva, Souza, Silva Filho, Medeiros and Criado, 2011; Tomaz et al., 2015) (see Appendix I for details of categorization by subjective and objective measures).

A few studies (10%) found no difference between e-learning and face-to-face learning with regards to knowledge and skills acquisition (Florescu et al., 2015; Khoshbaten et al., 2014; Martínez & Tuesca, 2014; Kukulja Taradi et al., 2008). One study reported negative results for e-learning since the study participants preferred learning with textbooks rather than mobile devices (Martínez & Tuesca, 2014). Study participants’ positive feedback and attitude was found to contradict their actual low e-learning usage in one study (Patil et al., 2016).

## 5. Discussion

### 5.1. Primary results

We identified and reviewed 52 studies on e-learning for medical education in LMICs. Most of these studies were from UMICs in Asia, South America or Africa, while only a few were from lower-middle-income-countries (only one from an LIC), which was similarly reported by Frehywot et al. (2013). This may indicate that few e-learning studies have been published or conducted over the past 10 years in LICs. Most studies were found to focus primarily on the study participant as the learner, and only used individual learner variables for evaluation of the intervention instead of considering other aspects of a holistic e-learning evaluation, such as pedagogy, technology, context and learning environment (Attwell, 2006).

#### 5.1.1. How was e-learning implemented in interventions?

Some implementations approached medical e-learning innovatively by re-purposing social media platforms for e-learning, and re-using technologies like podcasts for medical education, which may be appropriate for cultures in sub-Saharan Africa (Croft, 2002). Approaches to diversify e-learning contents are worth investigating since they have the potential to foster a more engaging e-learning environment. Podcasts, for example, have a low-entry level and smartphones currently allow for easy recording of high-quality, digital audio. However, in the majority of studies, we found that predominantly online-based e-learning was being implemented. Almost all studies had not accounted for the challenges and costs of online access, nor had they provided potentially costly technical devices for e-learning access. For medically underserved rural areas in LMICs, where there often is no reliable internet connectivity or electricity, offline-based solutions that bridge connectivity challenges are potentially important enabling factors for students to benefit from e-learning (Barteit, Jahn et al., 2019; Barteit, Neuhan et al., 2019). Blended learning that reduces face-to-face contact time may provide an approach for low-resource settings to bridge the lack of infrastructure by partly outsourcing learning as self-directed and outside the actual class room (Graham, Woodfield, & Harrison, 2013). Furthermore, e-learning implementations in low-resource contexts may benefit from minimizing resources needed for e-learning, such as for content development and technical equipment. Especially the development of e-learning materials constitutes a bottleneck since most of the low-resource countries are faced with a lack of medical teachers (Frenk et al., 2010). To decrease the work load of medical teachers for e-learning content development, for example, students may contribute e-learning contents themselves by creating step-by-step, peer-reviewed video guides on medical procedures, like measuring blood pressure. Also, when preparing the step-by-step video guides, students review local guidelines and learn the procedure as per the current guideline. Furthermore, the sustainability of technical devices in low-resource countries is key and “should be carefully balanced within the given setting” (Barteit, Jahn, et al., 2019), as most low-resource countries are constrained in their access to electricity and Internet, and face different climatic environments, like hot temperatures and dust. Another key aspect for low-resource contexts is to take into consideration if the technical device can be repaired in the country and maintained by local ICT

staff, as this may not often be the case, especially with imported technical devices.

#### 5.1.2. What e-learning evaluation and assessment methods were employed in the studies?

A highly prevalent evaluation method was to assess student knowledge with multiple-choice tests, although a small number of studies assessed skills with hands-on assignments. It appears that questionnaire-based evaluations and pre- and post-test assessments of knowledge focusing on the study participants are the go-to methods for e-learning evaluation in LMICs. However, as Attwell (2006) and others (Kirkpatrick, 1975; Lahti, Hätönen, & Välimäki, 2014; Pickering & Joynes, 2016) have argued, effective adaptation of e-learning implementations requires more complex study designs and evaluation methods. For instance, assessment methods should address the actual knowledge and skills transferred via e-learning. Overall, some studies identified shortcomings, e.g., Martínez and Tuesca (2014) who stated they “had several limitations, including the low sample size, the lack of a control group” (p. 404) or Sabouni et al. (2017) who identified the limitation that their e-learning “course did not have a significant effect on students’ attitude toward [EBM] [which] may be explained by our questions not comprehensively assessing [EBM] attitude” (p. 5). Especially in low-resource contexts, many technology-enhanced interventions, like e-learning, never progress past the pilot stage, which has been coined pilotitis (Huang, Blaschke, & Lucas, 2017). A comprehensive evaluation framework (Kirkpatrick, 1975; Lahti et al., 2014; Pickering & Joynes, 2016) may lead to valuable insights into how to best implement e-learning into the local educational infrastructure.

#### 5.1.3. What outcome was evaluated?

We found that a total of five methods (questionnaire, document review, interviews, focus group discussions, system log data) were used to evaluate a wide spectrum of outcome measures ( $n = 52$ ) and most corresponded to an individual level of assessment. This finding highlighted the lack of a standard and common ground for medical education research (Cook, Beckman, & Bordage, 2007). Only very few studies extended beyond the individual level to gain insight into the complex ecosystem that defines effective e-learning implementation in LMICs. For example, Rohwer et al. (2013) discussed tutors’ perceptions of study participants and medical e-learning in groups, thus potentially adding a further layer of insight into their intervention. Taking into consideration the cultural context is a central aspect to avoid frustration and to provide a supportive learning environment (Meier, 2007). Especially when e-learning materials have not been adapted to the respective LMICs context, tutors have shown to help “students adapt to the style of the material and to make a course developed in another country both culturally and pedagogically relevant” (Selinger, 2004, p. 1).

#### 5.1.4. How was success of e-learning defined?

In line with the predominance of subjective evaluation of the learner’s perspective, many studies defined e-learning effectiveness based on subjective measures, such as perception, satisfaction, opinion, and student feedback. Patil et al. (2016) reported that subjective feedback from study participants opposed the actual e-learning usage, since the study participants’ feedback was positive, although their actual usage was quite low. Literature highlights the limitation of using subjective evaluation measures since the “scope of evaluation is typically limited to user enjoyment and satisfaction” (Ruggeri et al., 2013, p. 315).

### 5.2. Quality of included studies

Most studies did not produce scientifically sound evidence to conclusively determine e-learning effectiveness for medical education since only four studies featured RCTs. Other methodological approaches further weakened the significance of the studies’ results, such as not using a comparison group or randomizing the control group to collect baseline data. Other weak studies made use of invalidated methods, such as self-developed questionnaires. We found that almost all of the studies evaluated rather small-scale e-learning pilots. Thus, many studies received ratings below the set threshold of all quality measures (NOS, NOS-E, MERSQI). Poor quality studies in medical education have been previously reported (Cook et al., 2007). Al-Shorbaji et al. (2015, p. 156) found “important methodological flaws, that may have biased the findings” of their systematic review on e-learning and blended learning interventions (p. 18). A potential underlying reason for the overall low quality of included e-learning studies may be the limited available funding for medical education research (Reed et al., 2008).

However, there were some studies, such as Kulier et al. (2012), that received overall strong quality ratings since the study design and reporting followed the routine procedures for RCTs. Based on our findings, the facilitation of RCTs seems recommendable, since they may generate significant insights for the design and implementation of e-learning interventions for medical education, especially in LMICs. However, RCTs are challenging to conduct in the highly regulated and applied field of medical education (Davis, Crabb, Rogers, Zamora, & Khan, 2008). Alternatively, strong causal research designs could be used, such as quasi-experimental studies (Briz-Ponce, Juanes-Méndez, García-Peñalvo, & Pereira, 2016; Campbell & Stanley, 1966) or alternately, multidimensional evaluation frameworks (Attwell, 2006; Kirkpatrick, 1975; Lahti et al., 2014; Pickering & Joynes, 2016). A strong study design is key for identifying the weaknesses and strengths of interventions for medical education and adapting e-learning to the realities of the local setting (Frenk et al., 2010).

For most low-resource contexts, the overarching objective of e-learning is to increase medical education opportunities in quantity and quality to address much-needed health worker capacity, while at the same time decreasing the workload of overburdened medical teachers (Frehywot et al., 2013). Our findings underline that the majority of studies included in this review described and evaluated e-learning implementations in a pilot stage. The low number of publications of e-learning interventions in LMICs may impede the development of best practice standards for e-learning implementation in low-resource contexts. This limitation has been already identified as an re-occurring phenomenon in e-learning and has been coined as “pilot-itis” (Barteit, Jahn, et al., 2019; Franz-Vasdeki, Pratt, Newsome, & Germann, 2015; Huang et al., 2017; Oliver, 2014). Although the World Health Organization has reported that the

number of pilot interventions had decreased and large-scale interventions have increased (World Health Organization, 2016a), our findings may be indicative of the continued implementation of pilot projects that are ultimately ceded as unscalable projects, which cannot be translated into sustainable implementation beyond the pilot phase. Yet, sustainability is a pertinent aspect of an e-learning implementation, and specifically in LMICs. Failed e-learning implementation blocks resources and slows down the overall development process and people may simultaneously lose their trust in the usefulness of technologies for medical education and its potentially “revolutionary effect” in low-resource contexts. Main weaknesses for the sustainability of e-learning interventions have been identified and include inadequate training of all involved parties, poor technological support, unmet expectations and inadequate allotment of financial resources (Labrique, Vasudevan, Kochi, Fabricant, & Mehl, 2013). However, in our systematic review we only found a few studies that considered these aspects in their evaluation. In fact, many e-learning interventions do not account for a system-wide approach (Huang et al., 2017), thus decreasing the efficiency and sustainability of e-learning since challenges are not identified in time and potentially result in failure. Sustainable e-learning models that minimize the need for resources are necessary for low-resource contexts in particular (Barteit, Jahn et al., 2019). These include offline-based approaches that reduce the cost of internet usage or bottom-up approaches for content creation, such as student-created e-learning content, which alleviate overburdened senior medical teaching staff and promote student learning and engagement in a pedagogical way.

### 5.3. Limitations

A limitation of this study was the possibility that some sources were missed due to failure to locate articles within the searched databases or an unjustified exclusion during the screening process. However, selection bias was minimized by using a range of keywords to select the studies, a wide date range of included studies, multiple databases, and three researchers for screening and manually searching reference lists.

Due to the heterogeneity of the included studies, a meta-analysis was not conducted.

We did focus our review on medical students and medical doctors and excluded interventions targeting other health workers. However, there were many other studies including nurses, midwives or dentists that may have provided relevant evaluation methods for e-learning medical education in LMICs that were not included in this review. There were evaluation methods for medical education that were not included in the discussion section; however, they may be worthy of consideration during further analysis of the results of this study. Furthermore, we did not investigate the usage and delineation of the terminology of distance learning, blended learning and e-learning used in included studies.

## 6. Conclusions

This systematic review focused on e-learning for medical education implemented in LMICs and summarized the evaluation and assessment methods used to measure e-learning effectiveness. Our findings indicate that e-learning in LMICs has not met its expected potential. Although most studies have praised and implemented e-learning for its many advantages (e.g., adaptability, diversity and economic benefit), there are ongoing limitations to reach its full potential in LMICs. Some reasons for the limited success in e-learning implementation may be due to restricted financial resources and limited evaluation and assessment approaches rooted in an over-emphasis on evaluating individual learner perceptions and knowledge, while not considering other features that may potentially inform and guide a strong e-learning system. E-learning interventions are mainly implemented as small-scale pilot projects, which may indicate that educational institutions and organizations are not giving e-learning a strong mandate as an educational method, but rather see it as a commodity of technological modern days with a short shelf-life. Based on our results, to unfold the potential of e-learning for medical education, institutions and organizations need to embrace and fund e-learning beyond pilot stages and consider a wider picture of e-learning including evaluations that expand their scope beyond the individual learner and provide transparent evaluation designs. Some studies in this review highlighted the potential and effectiveness of e-learning in LMICs settings. To contribute further to the mitigation of the health worker crisis, especially in LICs, innovative and adapted e-learning implementations may increase the quality and quantity of medical education while at the same decrease the workload of involved medical teachers. Thus, medical e-learning has the potential to play a central role in contributing to achieving the SDGs on health for LMICs. Progress to that end requires a combined effort and strong commitment from all involved stakeholders including those on a national level and potential international aid agencies to serve as a foundation to unfold e-learning’s potential and expand the impact of this innovative teaching tool.

### Funding

This work was supported by the Else Kröner-Fresenius-Stiftung [2016\_HA67].

### Declaration of competing interest

None declared.

### Acknowledgements

SB, DG, MMJ conducted the systematic literature review. SB wrote the draft manuscript with support from FN, DG and MMJ. The

manuscript was reviewed by SB, DG, MMJ, TB, AJ and FN.

## Abbreviations

CEQ	course experience questionnaire
ICC	intraclass correlation coefficient
LICs	low-income countries
LMICs	low- and middle-income countries
MERSQI	Medical Education Research Study Quality Instrument
NOS	Newcastle-Ottawa Scale
NOS-E	Newcastle-Ottawa Scale-Education
OSCE	objective structured clinical examination
PICOS	patient, intervention, comparison, outcome and studies
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCTs	randomized-controlled trials
SDGs	sustainable development goals
UMICs	upper-middle-income countries

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.compedu.2019.103726>.

## References

- Abdelhai, R., Yassin, S., Ahmad, M. F., & Fors, U. G. H. H. (2012). An e-learning reproductive health module to support improved student learning and interaction: A prospective interventional study at a medical school in Egypt. *BMC Medical Education*, 12(1), 11. <https://doi.org/10.1186/1472-6920-12-11>.
- Abdollahi, A., Sheikhabahaei, S., Meysamie, A., Bakhshandeh, M., & Hosseinzadeh, H. (2014). Inter-observer reproducibility before and after web-based education in the Gleason grading of the prostate adenocarcinoma among the Iranian pathologists. *Acta Medica Iranica*, 52(5), 370. Retrieved from <http://acta.tums.ac.ir/index.php/acta/article/view/4614>.
- Adanu, R., Adu-Sarkodie, Y., Opare-Sem, O., Nkyekyer, K., Donkor, P., Lawson, A., et al. (2010). Electronic learning and open educational resources in the health sciences in Ghana. *Ghana Medical Journal*, 44(4), 159–162. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/21416051>.
- Al-Shorbaji, N., Atun, R., Car, J., Majeed, A., & Wheeler, E. (2015). *eLearning for undergraduate health professional education. A systematic review informing a radical transformation of health workforce development*. World Health Organisation. Retrieved from <http://whoeducationguidelines.org/sites/default/files/uploads/eLearning-healthprof-report.pdf>.
- Aluttis, C., Bishaw, T., & Frank, M. W. (2014). The workforce for health in a globalized context – global shortages and international migration. *Global Health Action*, 7(1), 23611. <https://doi.org/10.3402/gha.v7.23611>.
- Asongu, S. A., & Le Roux, S. (2017). Enhancing ICT for inclusive human development in Sub-Saharan Africa. *Technological Forecasting and Social Change*, 118, 44–54. <https://doi.org/10.1016/j.techfore.2017.01.026>.
- Attwell, G. (2006). Evaluating e-learning - a guide to the evaluation of e-learning. *Evaluate Europe Handbook Series*, 2, 5–46.
- Bandhu, S. D., & Raje, S. (2014). Experiences with e-learning in ophthalmology. *Indian Journal of Ophthalmology*, 62(7), 792–794. <https://doi.org/10.4103/0301-4738.138297>.
- Barteit, S., Jahn, A., Banda, S. S., Bärnighausen, T., Bowa, A., Chileshe, G., et al. (2019). E-learning for medical education in sub-Saharan Africa and low-resource settings: Viewpoint. *Journal of Medical Internet Research*, 21(1), e12449. <https://doi.org/10.2196/12449>.
- Barteit, S., Neuhaan, F., Bärnighausen, T., Lüders, S., Malunga, G., Chileshe, G., et al. (2019). Perspectives of nonphysician clinical students and medical lecturers on tablet-based health care practice support for medical education in Zambia, Africa: Qualitative study. *JMIR Mhealth Uhealth*, 7(1), e12637. <https://doi.org/10.2196/12637>.
- Bediang, G., Franck, C., Raetz, M.-A., Doell, J., Ba, M., Kamga, Y., et al. (2013). Developing clinical skills using a virtual patient simulator in a resource-limited setting. *Studies in Health Technology and Informatics*, 192, 102–106. <https://doi.org/10.3233/978-1-61499-289-9-102>.
- Black, P. J. (1993). *Formative and summative assessment by teachers*.
- Bollinger, R., Chang, L., Jafari, R., O'Callaghan, T., Ngatia, P., Settle, D., et al. (2013). Leveraging information technology to bridge the health workforce gap. *Bulletin of the World Health Organization*, 91, 890–892. <https://doi.org/10.2471/BLT.13.118737>.
- Briz-Ponce, L., Juanes-Méndez, J. A., García-Peñalvo, F. J., & Pereira, A. (2016). Effects of mobile learning in medical education: A counterfactual evaluation. *Journal of Medical Systems*, 40(6), 136. <https://doi.org/10.1007/s10916-016-0487-4>.
- Bustam, A., Azhar, M. N., Veriah, R. S., Arumugam, K., & Loch, A. (2014). Performance of emergency physicians in point-of-care echocardiography following limited training. *Emergency Medicine Journal*, 31(5), 369–373. <https://doi.org/10.1136/emmermed-2012-201789>.
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Houghton Mifflin Company.
- Carvalho, K. M. de, Iyeyasu, J. N., Castro, S. M., da, C. e, Monteiro, G. B. M., & Zimmermann, A. (2012). Experience with an internet-based course for ophthalmology residents. *Revista Brasileira de Educação Médica*, 36(1), 63–67. <https://doi.org/10.1590/S0100-55022012000100009>.
- Chang, L. W., Kadam, D. B., Sangle, S., Narayanan, S., Borse, R. T., McKenzie-White, J., et al. (2012). Evaluation of a multimodal, distance learning HIV management course for clinical care providers in India. *Journal of the International Association of Physicians in AIDS Care*, 11(5), 277–282. <https://doi.org/10.1177/1545109712451330>.
- Childs, S., Blenkinsopp, E., Hall, A., & Walton, G. (2005). Effective e-learning for health professionals and students—barriers and their solutions. A systematic review of the literature—findings from the HeXL project. *Health Information and Libraries Journal*, 22(s2), 20–32. <https://doi.org/10.1111/j.1470-3327.2005.00614.x>.
- Clunie, L., Morris, N. P., Joyne, V. C. T., & Pickering, J. D. (2017). *How comprehensive are research studies investigating the efficacy of technology-enhanced learning resources in anatomy education? A systematic review* (Vol. 00). Anatomical Sciences Education.
- Cook, D. A., Beckman, T. J., & Bordage, G. (2007). Quality of reporting of experimental studies in medical education: A systematic review. *Medical Education*, 41(8), 737–745. <https://doi.org/10.1111/j.1365-2923.2007.02777.x>.
- Cook, D. A., Hatala, R., Brydges, R., Zengjias, B., Szostek, J. H., Wang, A. T., et al. (2011). Technology-enhanced simulation for health professions education: A systematic review and meta-analysis. *Journal of the American Medical Association*, 306(9), 978–988. <https://doi.org/10.1001/jama.2011.1234>.

- Cook, D. A., & Reed, D. A. (2015). Appraising the quality of medical education research methods: The medical education research study quality instrument and the Newcastle. *Ottawa Scale-Education*, 90(8), 1067–1076. <https://doi.org/10.1097/ACM.0000000000000786>.
- Croft, A. (2002). Singing under a tree: Does oral culture help lower primary teachers be learner-centred? *International Journal of Educational Development*, 22(3–4), 321–337. [https://doi.org/10.1016/S0738-0593\(01\)00063-3](https://doi.org/10.1016/S0738-0593(01)00063-3).
- Davids, M. R., Chikte, U. M. E., & Halperin, M. L. (2011). Development and evaluation of a multimedia e-learning resource for electrolyte and acid-base disorders. *Advances in Physiology Education*, 35(3), 295–306. <https://doi.org/10.1152/advan.00127.2010>.
- David, M., & Sutton, C. D. (2004). *Social research: The basics*. Sage.
- Davis, J., Crabb, S., Rogers, E., Zamora, J., & Khan, K. (2008). Computer-based teaching is as good as face to face lecture-based teaching of evidence based medicine: A randomized controlled trial. *Medical Teacher*, 30(3), 302–307. <https://doi.org/10.1080/01421590701784349>.
- Dawson, A. Z., Walker, R. J., Campbell, J. A., & Egede, L. E. (2016). Effective strategies for global health training programs: A systematic review of training outcomes in low and middle income countries. *Global Journal of Health Science*, 8(11), 119–138. <https://doi.org/10.5539/gjhs.v7n2p119>.
- De Villiers, M., & Walsh, S. (2015). How podcasts influence medical students' learning—a descriptive qualitative study. *African Journal of Health Professions Education*, 7(1), 130–133.
- Diehl, L. A., de Souza, R. M., Gordan, P. A., Esteves, R. Z., & Coelho, I. C. M. (2015). User assessment of “InsuOnline,” a game to fight clinical inertia in diabetes: A pilot study. *Games for Health Journal*, 4(5), 335–343. <https://doi.org/10.1089/g4h.2014.0111>.
- Farokhi, M. R., Zarifsanayeh, N., Haghighi, F., & Mehrabi, M. (2016). E-Learning or in-person approaches in continuous medical education: A comparative study. *The IIOAB Journal*, 7(2), 472–476.
- de Fátima Wardenski, R., de Espíndola, M. B., Struchiner, M., & Giannella, T. R. (2012). Blended learning in biochemistry education: Analysis of medical students' perceptions. *Biochemistry and Molecular Biology Education*, 40(4), 222–228. <https://doi.org/10.1002/bmb.20618>.
- Florescu, C. C., Mullen, J. A., Nguyen, V. M., Sanders, B. E., & Vu, P. Q. P. (2015). Evaluating didactic methods for training medical students in the use of bedside ultrasound for clinical practice at a faculty of medicine in Romania. *Journal of Ultrasound in Medicine*. <https://doi.org/10.7863/ultra.14.09028>.
- Fontelo, P., Faustorilla, J., Gavino, A., & Marcelo, A. (2012). Digital pathology – implementation challenges in low-resource countries. *Analytical Cellular Pathology*, 35(1), 31–36. <https://doi.org/10.3233/ACP-2011-0024>.
- Franz-Vasdeki, J., Pratt, B. A., Newsome, M., & Germann, S. (2015). Taking mHealth solutions to scale: Enabling environments and successful implementation. *Journal of Mobile Technology in Medicine*, 4(1), 35–38. <https://doi.org/10.7309/jmtm.4.1.8>.
- Frehywot, S., Vovides, Y., Talib, Z., Mikhail, N., Ross, H., Wohltjen, H., et al. (2013). E-learning in medical education in resource constrained low- and middle-income countries. *BMC Human Resources for Health*, 11(4). <https://doi.org/10.1186/1478-4491-11-4>.
- Frenk, J., Chen, L., Bhutta, Z. A., Cohen, J., Crisp, N., Evans, T., et al. (2010). Health professionals for a new century: Transforming education to strengthen health systems in an interdependent world. *The Lancet*, 376, 1923–1958. [https://doi.org/10.1016/S0140-6736\(10\)61854-5](https://doi.org/10.1016/S0140-6736(10)61854-5).
- Gaikwad, N., & Tankhiwale, S. (2014). Interactive e-learning module in pharmacology: A pilot project at a rural medical college in India. *Perspectives on Medical Education*, 3(1), 15–30. <https://doi.org/10.1007/s40037-013-0081-0>.
- Gazibara, T., Marusic, V., Maric, G., Zarić, M., Vujić, I., Kisić-Tepavčević, D., et al. (2015). Introducing e-learning in epidemiology course for undergraduate medical students at the faculty of medicine, University of Belgrade: A pilot study. *Journal of Medical Systems*, 39(10), 121. <https://doi.org/10.1007/s10916-015-0302-7>.
- Gibbs, T., Brigden, D., & Hellenberg, D. (2006). Assessment and evaluation in medical education. *South African Family Practice*, 48(1), 5–7. <https://doi.org/10.1080/20786204.2006.10873311>.
- Graham, C. R., Woodfield, W., & Harrison, J. B. (2013). A framework for institutional adoption and implementation of blended learning in higher education. *Internet and Higher Education*, 18, 4–14. <https://doi.org/10.1016/j.iheduc.2012.09.003>.
- Grangeia, T. de A. G., de Jorge, B., Franci, D., Santos, T. M., Setubal, M. S. V., Schweller, M., et al. (2016). Cognitive load and self-determination theories applied to e-learning: Impact on students' participation and academic performance. *PLoS One*, 11(3), e0152462. <https://doi.org/10.1371/journal.pone.0152462>.
- Hemmati, N. N., Omrani, S., & Hemmati, N. N. (2013). A comparison of internet-based learning and traditional classroom lecture to learn CPR for continuing medical education. *The Turkish Online Journal of Distance Education*, 14(1), 256–265.
- Heydarpour, P., Hafezi-Nejad, N., Khodabakhsh, A., Khosravi, M., Khoshkish, S., Sadeghian, M., et al. (2013). Medical podcasting in Iran: pilot, implementation and attitude evaluation. *Acta Medica Iranica*, 51(1), 59–61.
- Higgins, J. P. T., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions* (Vol. 4). John Wiley & Sons.
- Hofstetter, F. T. (2001). *Multimedia literacy*. McGraw-Hill, Inc.
- Hornos, E. H., Pleguezuelos, E. M., Brailovsky, C. A., Harillo, L. D., Dory, V., & Charlin, B. (2013). The practicum script concordance test: An online continuing professional development format to foster reflection on clinical practice. *Journal of Continuing Education in the Health Professions*, 33(1), 59–66. <https://doi.org/10.1002/chp.21166>.
- Huang, F., Blaschke, S., & Lucas, H. (2017). Beyond pilotitis: Taking digital health interventions to the national level in China and Uganda. *BMC Globalization and Health*, 13(1), 1–11. <https://doi.org/10.1186/s12992-017-0275-z>.
- Jesson, J., Matheson, L., & Lacey, F. M. (2011). *Doing your literature review: Traditional and systematic techniques*. Sage.
- Kaliyadan, F., Manoj, J., Dharmaratnam, A. D., & Sreekanth, G. (2010). Self-learning digital modules in dermatology: A pilot study. *Journal of the European Academy of Dermatology and Venereology*, 24(6), 655–660. <https://doi.org/10.1111/j.1468-3083.2009.03478.x>.
- Karakuş, A., Duran, L., Yavuz, Y., Altıntop, L., & Çalişkan, F. (2014). Computer-based simulation training in emergency medicine designed in the light of malpractice cases. *BMC Medical Education*, 14(1), 155. <https://doi.org/10.1186/1472-6920-14-155>.
- Khoshbaten, M., Soleimanpour, H., Ala, A., Vahdati, S. S., Ebrahimian, K., Safari, S., et al. (2014). Which form of medical training is the best in improving interns' knowledge related to advanced cardiac life support drugs pharmacology? An educational analytical intervention study between electronic learning and lecture-based education. *Anesthesiology and Pain Medicine*, 4(1), e15546. <https://doi.org/10.5812/aapm.15546>.
- Kirkpatrick, D. L. (1975). *Evaluating training programs*. Tata McGraw-Hill Education.
- Kotb, M. A., Elmahdy, H. N., Khalifa, N. E. D. M., El-Deen, M. H. N., & Lotfi, M. A. N. (2015). Pediatric online evidence-based medicine assignment is a novel effective enjoyable undergraduate medical teaching tool: A SQUIRE compliant study. *Medicine*, 94(29), e1178. <https://doi.org/10.1097/MD.0000000000001178>.
- Kukulja Taradi, S., Dogaš, Z., Dabić, M., & Drenjančević Perić, I. (2008). Scaling-up undergraduate medical education: Enabling virtual mobility by online elective courses. *Croatian Medical Journal*, 49(3), 344–351. <https://doi.org/10.3325/cmj.2008.3.344>.
- Kulatunga, G. G. A. K., Marasinghe, R. B., Karunathilake, I. M., & Dissanayake, V. H. W. (2013). Development and implementation of a web-based continuing professional development (CPD) programme on medical genetics. *Journal of Telemedicine and Telecare*, 19(7), 388–392. <https://doi.org/10.1177/1357633X13506525>.
- Kulier, R., Coppus, S. F. P. J., Zamora, J., Hadley, J., Malick, S., Das, K., et al. (2012). Effectiveness of a clinically integrated e-learning course in evidence-based medicine for reproductive health training: A randomized trial. *Journal of the American Medical Association*, 308(21), 2218–2225. <https://doi.org/10.1001/jama.2012.33640>.
- Labrique, A. B., Vasudevan, L., Kochi, E., Fabricant, R., & Mehl, G. (2013). mHealth innovations as health system strengthening tools: 12 common applications and a visual framework. *Global Health: Science and Practice*, 1(2), 160–171. <https://doi.org/10.9745/GHSP-D-13-00031>.
- Lahti, M., Hätönen, H., & Välimäki, M. (2014). Impact of e-learning on nurses' and student nurses knowledge, skills, and satisfaction: A systematic review and meta-analysis. *International Journal of Nursing Studies*, 51(1), 136–149. <https://doi.org/10.1016/j.ijnurstu.2012.12.017>.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>.
- Liu, Q., Peng, W., Zhang, F., Hu, R., Li, Y., & Yan, W. (2016). The effectiveness of blended learning in health professions: Systematic review and meta-analysis. *Journal of Medical Internet Research*, 18(1), e2. <https://doi.org/10.2196/jmir.4807>.
- Mack, H. G., Golinik, K. C., Murray, N., & Filipe, H. P. (2017). Models for implementing continuing professional development programs in low-resource countries. *MedEdPublish*, 6. <https://doi.org/10.15694/mep.2017.000018>.



- Martínez, E. G., & Tucsca, R. (2014). Modified team-based learning strategy to improve human anatomy learning: A pilot study at the Universidad del Norte in Barranquilla, Colombia. *Anatomical Sciences Education*, 7(5), 399–405. <https://doi.org/10.1002/ase.1444>.
- Masters, K., & Gibbs, T. (2007). The spiral curriculum: Implications for online learning. *BMC Medical Education*, 7, 1–10. <https://doi.org/10.1186/1472-6920-7-52>.
- Meier, C. (2007). Enhancing intercultural understanding using e-Learning Strategies. 27(4), 655–672.
- Merriam-Webster. (2018). *Podcast*.
- Mills, A. (2014). Health care systems in low-and middle-income countries. *New England Journal of Medicine*, 370(6), 552–557. <https://doi.org/10.1056/NEJMra1110897>.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>.
- Mojtahedzadeh, R., Mohammadi, A., Emami, A. H., & Rahmani, S. (2014). Comparing live lecture, internet-based & computer-based instruction: A randomized controlled trial. *Medical Journal of the Islamic Republic of Iran*, 28(1), 136.
- Moore, J. L., Dickson-Deane, C., & Galyen, K. (2011). e-Learning, online learning, and distance learning environments: Are they the same? *The Internet and Higher Education*, 14(2), 129–135. <https://doi.org/10.1016/j.iheduc.2010.10.001>.
- Muñoz, D. C., Ortiz, A., González, C., López, D. M., & Blobel, B. (2010). Effective e-learning for health professional and medical students: The experience with SIAS-Intelligent Tutoring System. *Studies in Health Technology and Informatics*, 156, 89–102. <https://doi.org/10.3233/978-1-60750-565-5-89>.
- Newberry, L., Kennedy, N., & Greene, E. A. (2016). Development of a subspecialty cardiology curriculum for paediatric registrars in Malawi: Implementation of a long-distance hybrid model. *Malawi Medical Journal*, 28(2), 57–60.
- Nicoll, P., MacRury, S., Van Woerden, H. C., & Smyth, K. (2018). Evaluation of technology-enhanced learning programs for health care professionals: Systematic review. *Journal of Medical Internet Research*, 20(4), 1–9. <https://doi.org/10.2196/jmir.9085>.
- Oliver, D. (2014). Ineffective pilots are a virus afflicting the NHS. *Health Service Journal*, 124(6405), 20–21. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/25219050>.
- O'Donovan, J., Ahn, R., Nelson, B. D., Kagan, C., & Burke, T. F. (2016). Using low-cost Android tablets and instructional videos to teach clinical skills to medical students in Kenya: A prospective study. *JRSM Open*, 7(8). <https://doi.org/10.1177/2054270416645044>, 2054270416645044.
- Patil, R. N., Almale, B. D., Patil, M., Gujrathi, A., Dhakne-Palwe, S., Patil, A. R., et al. (2016). Attitudes and perceptions of medical undergraduates towards mobile learning (M-learning). *Journal of Clinical and Diagnostic Research*, 10(10), JC06–JC10. <https://doi.org/10.7860/JCDR/2016/20214.8682>.
- Phillips, R., McNaught, C., & Kennedy, G. (2012). *Evaluating e-learning: Guiding research and practice*. Routledge.
- Pickering, J. D., & Joynes, V. C. T. (2016). A holistic model for evaluating the impact of individual technology-enhanced learning resources. *Medical Teacher*, 38(12), 1242–1247. <https://doi.org/10.1080/0142159X.2016.1210112>.
- Pino, S., Mora, C., Diaz, A., Guarnizo, P., & Jaimes, D. (2017). Improving skills in pediatric rheumatology in Colombia: A combined educational strategy supported by ILAR. *Clinical Rheumatology*, 36(7), 1631–1635. <https://doi.org/10.1007/s10067-016-3294-x>.
- Pusponegoro, H. D., Soebadi, A., & Surya, R. (2015). Web-based versus conventional training for medical students on infant gross motor screening. *Telemedicine and E-Health*, 21(12), 992–997. <https://doi.org/10.1089/tmj.2015.0024>.
- Rajapakse, S., Fernando, D., Rubasinghe, N., & Gurusinge, S. (2009). E-Learning in medical education: Guide supplement 32.6 Practical application. *Medical Teacher*, 31(5), 452–453. <https://doi.org/10.1080/01421590902833036>.
- Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The course experience questionnaire. *Studies in Higher Education*, 16(2), 129–150. <https://doi.org/10.1080/03075079112331382944>.
- Reed, D. A., Beckman, T. J., Wright, S. M., Levine, R. B., Kern, D. E., & Cook, D. A. (2008). Predictive validity evidence for medical education research study quality instrument scores: Quality of submissions to JGIM's medical education special issue. *Journal of General Internal Medicine*, 23(7), 903–907. <https://doi.org/10.1007/s11606-008-0664-3>.
- Reed, D. A., Cook, D. A., Beckman, T. J., Levine, R. B., Kern, D. E., & Wright, S. M. (2008). Association between funding and quality of published medical education research. *Journal of the American Medical Association*, 298(9), 1002–1009. <https://doi.org/10.1001/jama.298.9.1002>.
- Reis, L. O., Ikari, O., Taha-Neto, K. A., Gugliotta, A., & Denardi, F. (2015). Delivery of a urology online course using Moodle versus didactic lectures methods. *International Journal of Medical Informatics*, 84(2), 149–154. <https://doi.org/10.1016/j.ijmedinf.2014.11.001>.
- Rohwer, A., Young, T., & Van Schalkwyk, S. (2013). Effective or just practical? An evaluation of an online postgraduate module on evidence-based medicine (EBM). *BMC Medical Education*, 13(1), 77. <https://doi.org/10.1186/1472-6920-13-77>.
- Ruggeri, K., Farrington, C., & Brayne, C. (2013). A global model for effective use and evaluation of e-learning in health. *Telemedicine Journal and e-Health: The Official Journal of the American Telemedicine Association*, 19(4), 312–321. <https://doi.org/10.1089/tmj.2012.0175>.
- Sabouni, A., Bdaiwi, Y., Janoudi, S. L., Namous, L. O., Turk, T., Alkhatib, M., et al. (2017). Multiple strategy peer-taught evidence-based medicine course in a poor resource setting. *BMC Medical Education*, 17(1), 1–6. <https://doi.org/10.1186/s12909-017-0924-1>.
- Sánchez-Mendiola, M., Martínez-Franco, A. I., Rosales-Vega, A., Villamar-Chulín, J., Gatica-Lara, F., García-Durán, R., et al. (2012). Development and implementation of a biomedical informatics course for medical students: Challenges of a large-scale blended-learning program. *Journal of the American Medical Informatics Association*, 20(2), 381–387. <https://doi.org/10.1136/amiajnl-2011-000796>.
- Selinger, M. (2004). Cultural and pedagogical implications of a global e-learning programme. *Cambridge Journal of Education*, 34(2), 223–239. <https://doi.org/10.1080/03057640410001700589>.
- Seluakumaran, K., Jusof, F. F., Ismail, R., & Husain, R. (2011). Integrating an open-source course management system (Moodle) into the teaching of a first-year medical physiology course: A case study. *Advances in Physiology Education*, 35(4), 369–377. <https://doi.org/10.1152/advan.00008.2011>.
- de Sena, D. P., Fabricio, D. D., Lopes, M. H. I., & da Silva, V. D. (2013). Computer-assisted teaching of skin flap surgery: Validation of a mobile platform software for medical students. *PLoS One*, 8(7), e65833. <https://doi.org/10.1371/journal.pone.0065833>.
- Silva, C. S., Souza, M. B., Silva Filho, R. S., Medeiros, L. M. de, & Criado, P. R. (2011). E-learning program for medical students in dermatology. *Clinics*, 66(4), 619–622. <https://doi.org/10.1590/S1807-593220111000400016>.
- Stang, A. (2010). Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *European Journal of Epidemiology*, 25(9), 603–605. <https://doi.org/10.1007/s10654-010-9491-z>.
- Suebnuarn, S., & Haddawy, P. (2007). COMET: A collaborative tutoring system for medical problem-based learning. *IEEE Intelligent Systems*, 22(4), 70–77. <https://doi.org/10.1109/MIS.2007.66>.
- Tayyab, A., & Naseem, K. (2013). Convergence of technologies: The print, web and multimedia united to deliver an integrated basic medical sciences course in a resource limited setting. *Journal of the College of Physicians and Surgeons Pakistan*, 23(8), 607–609. [dx.doi.org/10.2013/JCPS.607609](https://doi.org/10.2013/JCPS.607609).
- Tomaz, J., Mamede, S., Filho, J., Roriz Filho, J. de, & van der Molen, H. (2015). Effectiveness of an online problem-based learning curriculum for training family medical doctors in Brazil. *Education and Health*, 28(3), 187–193. <https://doi.org/10.4103/1357-6283.178605>.
- Varghese, J., Faith, M., & Jacob, M. (2012). Impact of e-resources on learning in biochemistry: First-year medical students perceptions. *BMC Medical Education*, 12(1). <https://doi.org/10.1186/1472-6920-12-21>.
- Vieira, T. C. S. B., Nakamura, M. U., da Silva, I., Torloni, M. R., Ribeiro, M. C., de Tubino Scanavino, M., et al. (2017). Experience of an online course on sexuality during pregnancy for residents. *Sexual & Reproductive Healthcare*, 12, 76–81. <https://doi.org/10.1016/j.srhc.2017.03.004>.
- Wagner, N., Hassanein, K., & Head, M. (2008). Who is responsible for e-learning success in higher education? A stakeholders' analysis. *Journal of Educational Technology & Society*, 11(3).
- Walsh, S., & De Villiers, M. (2015). Enhanced podcasting for medical students: Progression from pilot to e-learning resource. *African Journal of Health Professions Education*, 7(1), 125–129. <https://doi.org/10.7196/AJHP.505>.
- World Bank. (2017). *World Bank country and lending groups 2017*. Retrieved from <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519,%0D>.
- World Health Organization. (2014). *The health of the people: What works - the African regional health report 2014*. Retrieved from <http://www.who.int/bulletin/africanhealth2014/en/#>.

- World Health Organization. (2016). *Atlas of e-Health country profiles*. Retrieved from [https://www.who.int/goe/publications/atlas\\_2015/en/](https://www.who.int/goe/publications/atlas_2015/en/).
- World Health Organization. (2016). *Global strategy on human resources for health: Workforce 2030*. Retrieved from <https://www.who.int/hrh/resources/globstrathrh-2030/en/>.
- Wu, B., Wang, M., Johnson, J. M., & Grotzer, T. A. (2014). Improving the learning of clinical reasoning through computer-based cognitive representation. *Medical Education Online*, 19(1), 25940. <https://doi.org/10.3402/meo.v19.25940>.
- Zhang, D., & Zhou, L. (2003). Enhancing E-learning with interactive multimedia. *Information Resources Management Journal*, 16(4), 1–14. <https://doi.org/10.4018/irmj.2003100101>.