Original Article

The efficacy and acceptance of online learning *vs.* offline learning in medical student education: a systematic review and meta-analysis

Meng Gao[1,2#], Yu Cui[1,2#], Hequn Chen[1,2], Huimin Zeng[1,2], Zewu Zhu[1,2], Xiongbing Zu[1,2^]

[1]Department of Urology, Xiangya Hospital, Central South University, Changsha, China; [2]National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, China

Contributions: (I) Conception and design: X Zu; (II) Administrative support: H Chen; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: M Gao, Y Cui; (V) Data analysis and interpretation: H Zeng, Z Zhu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Xiongbing Zu. No. 87, Xiangya Road, Changsha 410008, China. Email: zuxbxy@126.com.

Background: To maintain the continuity of medical education during the COVID-19 epidemic, online learning has replaced traditional face-to-face learning. But the efficacy and acceptance of online learning for medical education remains unknown. This meta-analysis aimed to assess whether online learning improves learning outcomes and is more acceptable to medical students compared to offline learning.

Methods: Four databases were searched for randomized controlled trials (RCTs) and comparative studies (non-RCTs) involving online learning published from January 1900 to October 2020. A total of twenty-seven studies comparing online and offline learning in medical students were included. The Grading of Recommendations, Assessment, Development and Evaluations (GRADE) framework and Newcastle-Ottawa Scale (NOS) were used to assess the methodological quality of RCTs and non-RCTs respectively. The data of knowledge and skills scores and course satisfaction were synthesized using a random effects model for the meta-analysis.

Results: Twenty-one RCTs that were judged to be of high quality according to the GRADE framework and six non-RCTs studies which ranged from 6 to 8 (NOS) and can be considered high-quality were included in this meta-analysis. The revealed that the online learning group had significantly higher post-test scores (SMD =0.58, 95% CI: 0.25 to 0.91; P=0.0006) and pre- and post-test score gains than the offline group (SMD =1.12, 95% CI: 0.14 to 2.11, P=0.02). In addition, online education was more satisfactory to participants than the offline learning (OR: 2.02; 95% CI: 1.16 to 3.52; P=0.01). Subgroup analysis was performed on knowledge and skill scores at the post-test level. The selected factors included study outcome, study design and type, participants, course type and country. No significant factors were observed in the subgroup analysis except for course type subgroup analysis.

Discussion: Online learning in medical education could lead to higher post-test knowledge and skill scores than offline learning. It also has higher satisfaction ratings than offline education. In conclusion, online learning can be considered as a potential educational method during the COVID-19 pandemic. However, given the risk of bias of included studies such as the inclusion of non-randomized comparative studies, the conclusion should be made with cautions.

Trial Registration: CRD42020220295.

Keywords: Online learning; offline learning; meta-analysis; undergraduate medical education

^ ORCID: 0000-0002-3873-251X.

Received: 20 January 2022; Accepted: 19 May 2022; Published: 30 June 2022.
doi: 10.21037/jxym-22-3
View this article at: https://dx.doi.org/10.21037/jxym-22-3

Introduction

The World Health Organization classified COVID-19 as a pandemic on March 11, 2020, and the number of people infected with COVID-19 worldwide has been increasing sharply. Many educational institutions in the world, including schools and hospitals, had to suspend teaching activities. To maintain the continuity of medical education during the COVID-19 epidemic, online learning has replaced traditional face-to-face learning (1) because online technologies allow medical students to work at home between face-to-face classes and academic practices. Online learning is the act of teaching and learning through digital technology. As the core of online learning, digital technology has also become a strategy for improving the education and training of health workers (2) due to its wide application and continuous development and progress in various fields in recent years. Online learning is a general term for a variety of education approaches, concepts, methods and technologies that are constantly changing (3). It can include but is not limited to online computer-based digital education, large-scale open online courses, virtual reality (VR), virtual patients, mobile learning and basic conversion of content into a digital format (for example, PDF or HTML format for books) (3). Online learning can be used flexibly and unlimitedly with traditional methods (such as role-playing with standardized patients) so that students can practice their skills interchangeably. For educators, this educational approach can save time, effort, and space; automatically assess and record student learning progress; and obtain feedback from students (4). A series of studies have compared the effectiveness and feasibility of online and offline education for medical students, but the effect of online education is not particularly clear. Pei *et al.* (5) selected 16 published articles for meta-analysis and suggested that compared with offline learning, online learning has advantages in enhancing the knowledge and skills of medical students. However, He *et al.* (6) pointed out that online learning was not significantly different from traditional education in the effectiveness of knowledge and skills. The main reason for these inconsistent findings may be because the populations included in the two meta-analyses were different.

To provide further evidence for the efficacy and acceptance of online teaching, the current meta-analysis aims to provide new perspectives for comparing the effects of online learning and offline learning interventions. Therefore, we designed this meta-analysis to further compare the effects of online learning and offline learning for medical students including clinical, nursing and pharmacy and to identify the factors that may lead to differences in the effectiveness of the two teaching methods. We present the following article in accordance with the PRISMA reporting checklist (available at https://jxym.amegroups.com/article/view/10.21037/jxym-22-3/rc)

Methods

Search strategy

We developed comprehensive search strategies for the PubMed, Web of Science, Cochrane Controlled Trials Central Registry (CENTRAL) and Embase databases to identify research related to online learning. The search time of the database was from January 1, 1990, to October 2020; 1990 was chosen as the start year of the search because before that, the use of computers was limited to basic tasks (3). The search strategies were as follows: (“online learning” OR “digital education” OR “distance education” OR “Internet-Based Learning” OR “virtual education” AND “offline learning OR traditional education OR face-to-face learning OR classroom education OR usual teaching)”. The “Related Articles” function was also used to expand the search scope and supplement the computer search by manually searching all retrieved studies, reference lists of reference articles and conference abstracts. After completing all searches, we identified all potentially relevant articles, used Endnote X9 (reference management software) without language restrictions, and deleted duplicate studies. Two independent reviewers scanned the title, abstract, and even the full text of all records to identify potentially relevant studies.

Selection of studies

This meta-analysis has been registered at PROSPERO: CRD42020220295. According to the Preferred Reporting Items for Systematic Reviews and Meta-analysis and

Meta-analysis of Observational Studies in Epidemiology recommendations for study reporting (7), the selection of the article was conducted independently by two reviewers. The inclusion criteria were as follows: all available randomized controlled trials (RCTs) and retrospective comparative studies (cohort or case-control studies) that compared any form of online learning online learning with offline learning (traditional learning) to medical students from all over the world, and that had at least one of the following outcomes: knowledge and skill outcomes measured by objective assessment tools. In addition, studies on blended learning models (online + offline learning) were excluded.

In addition, the included studies should meet the following criteria in adherence to the participant, intervention, comparison and outcome (PICO) search in the field of evidence-based medicine:Participants: medical undergraduate students including clinical, nursing and pharmacy.

Interventions: online computer-based digital education, large-scale open online courses, VR, virtual patients, mobile learning and basic conversion of content into a digital format (for example, PDF or HTML format for books). Comparisons: offline learning, especially referring to face-to-face teaching in a classroom, seminars, and reading text-based documents or books only. Outcomes: knowledge and skill outcomes measured by objective assessment instruments. The mean score and standard deviations of post-test, pre- and post-test gains.

Data extraction and assessment

The full texts of the included studies were screened twice, and data from these studies were also separately extracted by two authors in a standardized format. No duplicate publications were found during the data extraction process. The main outcomes were the knowledge and skill scores at post-test. The secondary outcomes were pre- and post-test gains (improvement), retention test scores and students' overall satisfaction with the course format.

Randomized controlled trials were judged to be of high quality according to the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) framework (8), which specifies four levels of evidence: high, moderate, low, and very low quality evidence. The methodological quality of RCTs was assessed by the Cochrane risk of bias tool, which included the following domains (I) random sequence generation, (II) allocation concealment, (III) blinding of participants and personnel, (IV) blinding of outcome assessment, (V) incomplete outcome data, (VI) selective reporting, and (VII) any other source of bias (8). The Newcastle-Ottawa Scale (NOS) was used to assess the methodological quality of those nonrandomized studies (9). The scores range from 0 to 9, and the scale includes: selection of patients, comparability of the study groups, exposure (Case Control Studies) or outcome (Cohort Studies).

Statistical analysis

All meta-analyses were performed using Windows Version 5.3 Review Manager (Cochrane Collaboration, Oxford, England) and STATA 12.0 (Stata Corp LP, University Town, Texas, USA). A random effects model was used due to differences in the expected population and course diversity (10). Standard mean differences (SMDs) were used for continuous parameter data, and odds ratios (ORs) were used for the dichotomous variables, with both types of data reported with 95% confidence intervals (CIs). For some studies that only reported continuous data as the means, 95% confidence interval, range and sample size, the standard deviations were converted using the technique described by Hozo *et al.* (11). The statistical heterogeneity between studies was evaluated using the χ^2 test, and the significance was set to P=0.1, and I^2 statistics were used to evaluate statistical heterogeneity ($I^2 \geq 50\%$ indicating there is heterogeneity) (12). The Z test was used to determine the pooled effects, and a P value <0.05 indicated the presence of statistical significance (13). Data are presented as forest plots, and a funnel plot was routinely constructed to assess publication bias (14).

Results

Results of the search

We searched a total of 2,172 records in four databases: twenty-seven studies including 2,308 participants (1,191 participants for online learning and 1,117 participants for offline learning) met the final inclusion criteria and were full-text articles (*Figure 1*). Seven hundred fifty-seven records were excluded after screening the title and abstract, and 241 studies were excluded after reading the full text (*Figure 1*).

Characteristics and quality of the included studies

The main characteristics of the 27 included studies, such

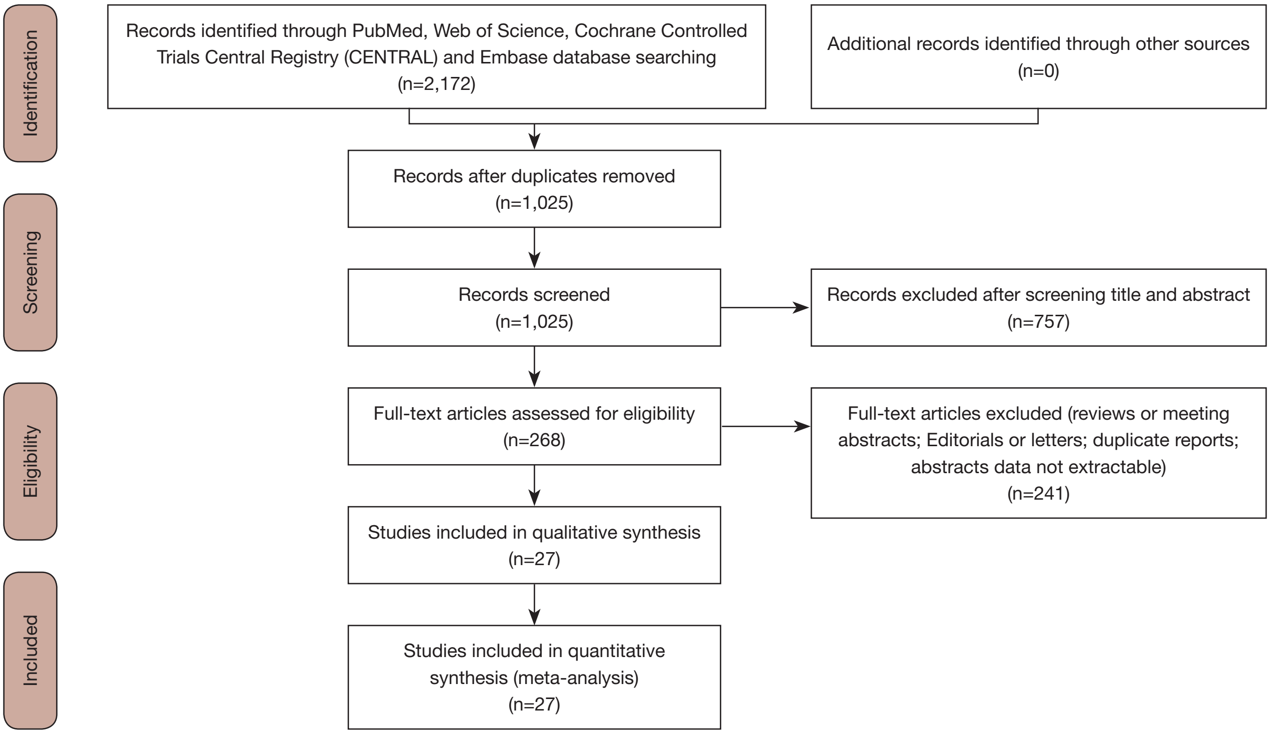

Figure 1 Flow diagram of study selection.

as participants, comparison, course and outcome are shown in *Table 1*. Except for 6 studies (15-20) that were nonrandomized controlled studies, the remaining 21 (21-41) studies were all RCTs that were judged to be of high quality according to the GRADE framework. All articles compared posttest scores; 16 articles compared both posttest scores and pre-test and posttest score gains on the same sample, but only 5 studies had sufficient pre- and posttest score gains for meta-analysis. One study compared retention test scores 22 days after the intervention, and 7 articles compared students' overall satisfaction with the way they attended classes. Most studies were conducted in developed countries, and five studies were conducted in developing countries. The overall risk of bias assessed according to the Cochrane risk-of-bias tool for all included RCTs is shown in *Figure 2*. The framework of the Cochrane bias risk tool contains the seven abovementioned areas mentioned above. Most studies described the randomization process in detail, but few articles could achieve the true blinding of participants and outcome assessment. Only Phadtare *et al.* (23) achieved participant blinding by placing group assignments in sealed envelopes and revealing after participants had signed informed consent and Porter *et al.* (24) performed lecturing teacher blinding. For the rest of the nonrandomized studies, their scores ranged from 6 to 8 on the NOS, which can be considered high-quality. The assessments of detail were shown in *Table 2*.

Outcomes

Knowledge and skill score at the post-test level

Data on knowledge or skill scores were available for all 27 studies, with a total sample size of 2,308 reported. The pooled results showed that the online learning group had significantly higher scores than the offline group (SMD =0.58, 95% CI, 0.25 to 0.91; P=0.0006) (*Figure 3*).

Pre- and post-test score gains

Five studies (20,25,26,31,41) including 278 students provided data on pre- and post-test score gains. There was a significant difference in the pre- and post-test score gains between the two groups (SMD =1.12, 95% CI, 0.14 to 2.11, P=0.02) (*Figure 4*). High heterogeneity was found, and a random-effects model was used (I^2=92%).

Overall satisfaction

Overall satisfaction was reported in 7 eligible articles, but only three studies had suitable data for meta-analysis. A

Table 1 Characteristics of included studies

Author, year, country	Comparison	Samples (T/C, n)	Participants	Course	Study design	Assessment strategies	Outcome	Design [score]
Brettle *et al.*, 2013, UK	Online *vs.* face	70 (35/35)	Undergraduate nurse	Information literacy skill	Pretest/ post-test	Skill test	Pre- and post-session search skills score, follow up skill score	RCT
Hu *et al.*, 2016, USA	3D computer *vs.* text	100 (49/51)	Medical students	Laryngeal anatomy	Post-test only	Knowledge test	Laryngeal anatomy test score and instructional materials motivation survey	RCT
Phadtare *et al.*, 2009, USA	Online *vs.* standard	48 (24/24)	Second- and third-year medical student	Scientific writing	Post-test only	Skill test	Manuscript quality and self-reported participant satisfaction	RCT
Porter *et al.*, 2014, USA	Online *vs.* classroom	140 (71/69)	Second- and third-year medical student	Immunization course	Post-test only	Knowledge test	Grades and evaluation and assessment of course	RCT
Subramanian *et al.*, 2012, USA	Software *vs.* traditional	30 (15/15)	third-year medical student	Arrhythmia	Pretest/ post-test	Knowledge test	Post-test score, improvement and long-term retention	RCT
Bjarne *et al.*, 2013, Denmark	e-learning *vs.* face	42 (21/21)	Anesthesiology nurse	Respiratory and pulmonary physiology	Pretest/ post-test	Knowledge test	Pre- and post-test score and improvement	RCT
Bowdish *et al.*, 2003, USA	Virtual *vs.* text	112 (56/56)	First-year medical students	Human physiology	Post-test only	Knowledge test	Teaching and Learning environment Questionnaire score and student achievement	Quasi-experimental [8]
Chittenden *et al.*, 2013	Web *vs.* written	74 (41/33)	Third-year medical student	Code status discussions	Post-test only	Skill test	Student performance in conducting code status discussions and communication skills	RCT
Soloman *et al.*, 2014	Digital *vs.* live	29 (17/12)	Third-year medical student	CAD and renal failure	Post-test only	Knowledge test	Exam score and feedback on the digital lecture format	RCT
Moazami *et al.*, 2014	Virtual *vs.* traditional	35 (15/20)	Dental medical students	Rotary instrumentation of root canals	Post-test only	Knowledge test	Knowledge acquisition and its retention	RCT
Alemánr *et al.*, 2011	Computer *vs.* convention	41 (15/26)	Second-year nurse student	Medical-surgical nursing	Pretest/ post-test	Skills and knowledge test	Pre- and post-test score, evaluation of the students' experience	RCT
Portero *et al.*, 2013	Virtual *vs.* convention	114 (71/43)	Third-year medical student	Radiology	Post-test only	Knowledge test	Final oral examination and evaluation on image interpretation	Case control [7]
Pusponegoro *et al.*, 2015	Online *vs.* live	75 (39/36)	Fifth-year medical student	Gross motor screening method in infants	Pretest/ post-test	Knowledge test	Pre- and post-test score, improvement and satisfaction	RCT

Table 1 (*continued*)

Table 1 (*continued*)

Author, year, country	Comparison	Samples (T/C, n)	Participants	Course	Study design	Assessment strategies	Outcome	Design [score]
Bhatti *et al.*, 2011	e-learning *vs.* face	148 (75/73)	Third-year medical student	Hemorrhoids	Pretest/ post-test	Knowledge test	Pre and post-test score, improvement and usefulness of website	RCT
Dennis *et al.*, 2003	Online *vs.* face	34 (17/17)	Second-year medical student	Problem-based learning	Post-test only	Knowledge test	Learning outcomes, time on-task and generation of LIs	RCT
Yeung *et al.*, 2012	Computer *vs.* tradition	78 (43/35)	Second-year medical student	Cranial nerve anatomy	Post-test only	Knowledge test	Post-test score and evaluation of participants' experience	RCT
Kaltman *et al.*, 2018	Video *vs.* usual	99 (60/39)	First-year medical student	Communication	Post-test only	Skill test	Simulation experience, OSCE communication behaviors and self-efficacy	RCT
Morente *et al.*, 2013	e-learning *vs.* tradition	73 (30/43)	Undergraduate nursing student	Pressure ulcer	Pretest/ post-test	Knowledge test	Pre- and post-test score and improvement	RCT
Peine *et al.*, 2016	e-learning *vs.* lecture	116 (61/55)	Third-year medical student	Modernized medical curricula	Pretest/ post-test	Knowledge test	Pre- and post-test score and self-assessment	RCT
Nicklen *et al.*, 2017	Online *vs.* face	38 (19/19)	Third-year medical student	Case-based learning	Post-test only	Knowledge test	Learning and self-assessed perception of learning, satisfaction	RCT
Clement *et al.*, 2012	DVD *vs.* lecture	130 (71/59)	Graduate nursing student	Stigma and mental health	Post-test only	Knowledge test	Knowledge, attitudes (cognitive and emotional) and behaviour	RCT
Chao *et al.*, 2012	Online *vs.* Lecture	167 (111/56)	Fourth-year medical student	Delirium	Pretest/ post-test	Skill test	Pre- and post-test score and improvement	Case control [6]
Farahmand *et al.*, 2016	Distance *vs.* Tradition	120 (60/60)	Senior medical students	Initial assessment of trauma	Post-test only	Knowledge and skill test	Post-test score	Quasi-experimental [8]
Taradi *et al.*, 2005	WPBL *vs.* face	121 (37/84)	Second-year medical student	Acid-base physiology	Post-test only	Knowledge test	Test scores and satisfaction survey results	Case control [7]
Assadi *et al.*, 2003	Video *vs.* traditional	81 (41/40)	Undergraduate intern	Basic life support instruction	Pretest/ post-test	Knowledge and skill test	Pre- and post-test score and satisfaction	Prospective research [7]
Raupach *et al.*, 2009	WPBL *vs.* face	143 (72/71)	Fourth-year medical student	Clinical reasoning skills	Post-test only	Knowledge test	Post-test score, student activity and evaluation	RCT
Alnabelsi *et al.*, 2015	e-learning *vs.* face	50 (25/25)	Fourth- and fifth-year medical student	ENT	Post-test only	Knowledge test	Pre- and post-test score, improvement and satisfaction	RCT

T/C, test group/control group; RCT, randomized controlled trial; LIs, a key product that facilitates self-directed learning during the tutorial process; ENT, Ear, Nose and Throat.

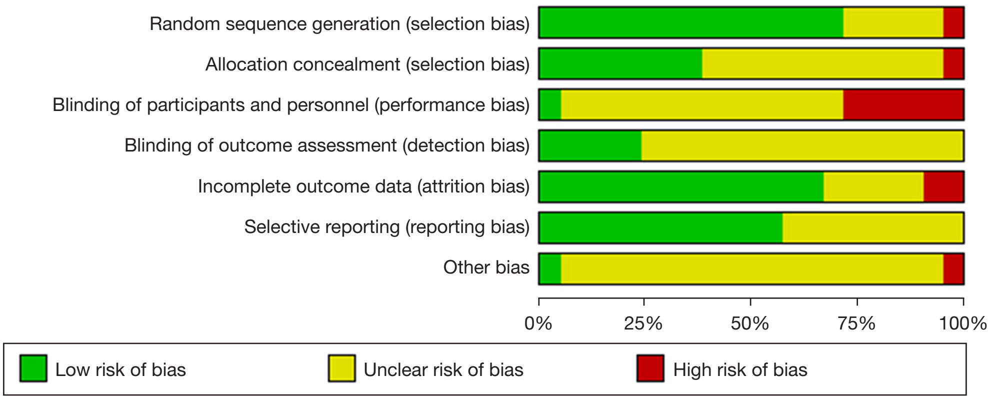

Figure 2 The overall risk of bias for included RCTs. RCTs, randomized controlled trials.

Table 2 Risk of bias assessment for included non-RCT trials

Study	Selection				Comparability		Exposure/outcome			Score
	a	b	c	d	e	f	g	h	i	
Bowdish *et al.*, 2003	★	★	★	★	★		★	★	★	8
Portero *et al.*, 2013	★	★	★		★		★	★	★	7
Chao *et al.*, 2012	★	★	★		★		★		★	6
Farahmand *et al.*, 2016	★	★	★	★	★		★	★	★	8
Taradi *et al.*, 2005,	★	★		★	★		★	★	★	7
Assadi *et al.*, 2003	★	★	★	★	★		★		★	7

a, adequate case definition; b, representativeness of the cases; c, selection of controls; d, definition of controls; e, study controls for the most important factor; f, study controls for any additional factor; g, ascertainment of exposure; h, some methods of ascertainment for cases and controls; I, non-response rate. ★ , a qualified identification, no special instructions are required. RCT, randomized controlled trial.

meta-analysis of these 3 studies (24,31,41) showed that online education was more satisfactory to participants than offline learning (OR: 2.02; 95% CI, 1.16 to 3.52; P=0.01). There was a moderate degree of heterogeneity, and a fixed effects model was used (P=0.12, I^2=53%) (*Figure 5*). A summary of the outcomes and the results of the meta-analysis are shown in *Table 3*.

Subgroup analysis

Subgroup analysis was performed on knowledge and skill scores at the post-test level (*Table 4*). The selected factors included study outcome, study design and type, participants, course type and country. There was a significant difference in course type subgroup analysis (*Figure 6*) compared with the original analysis (P=0.006), foundation course group analysis (SMD =0.07, 95% CI: −0.11 to 0.25, P=0.44) and other course group analysis (SMD =0.09, 95% CI: −1.10 to 1.28, P=0.88) were different from clinical course group (SMD =0.86, 95% CI: 0.41 to 1.31, P=0.0002) and original analysis (SMD =0.58, 95% CI: 0.25 to 0.91, P=0.0006). For the other selected factor subgroups, there was no significant difference between these subgroups (Figures S1-S5).

Publication bias

The research funnel chart (*Figure 7*) included in the meta-analysis was used to assess the publication bias in the knowledge and skill score at the post-test level. Most studies lay inside the 95% CIs, with a small number of studies having an uneven distribution, indicating that there was slight asymmetry.

Sensitivity analysis

Twenty-one RCTs and 6 CCTs that scored six or more on

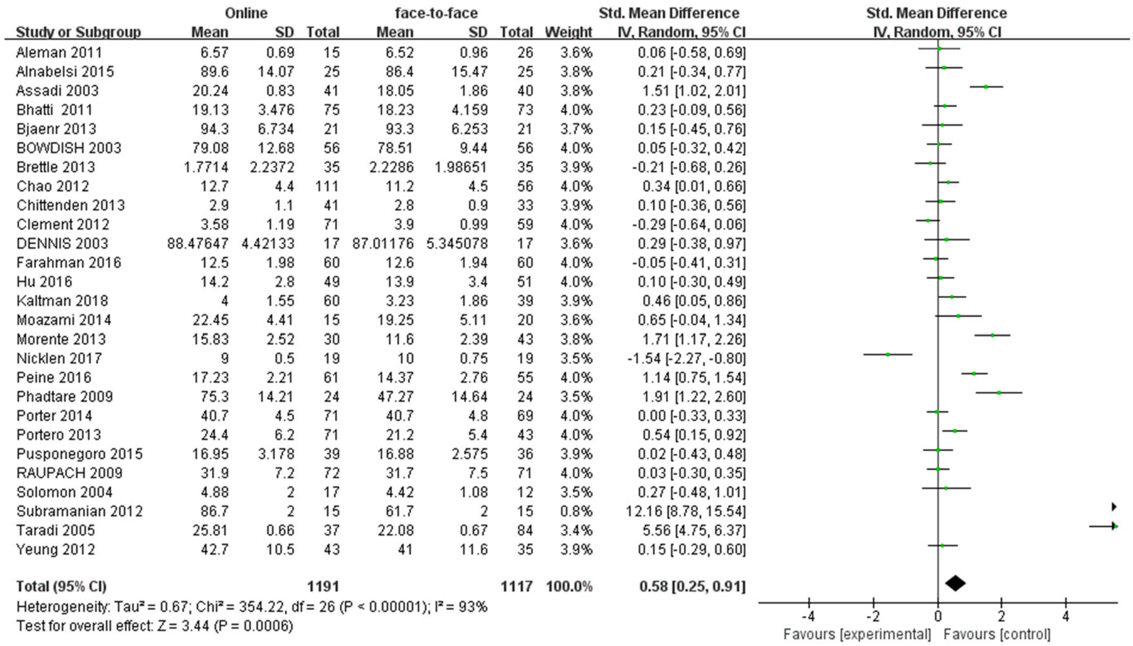

Figure 3 Forest plot for knowledge and skill score at the post-test level.

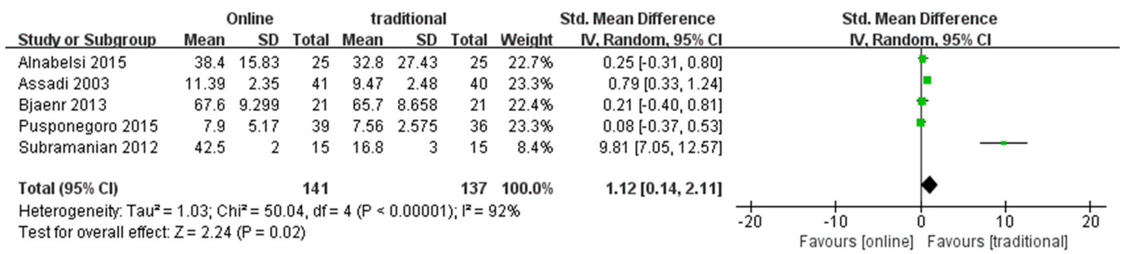

Figure 4 Forest plot for pre- and post-test score gains.

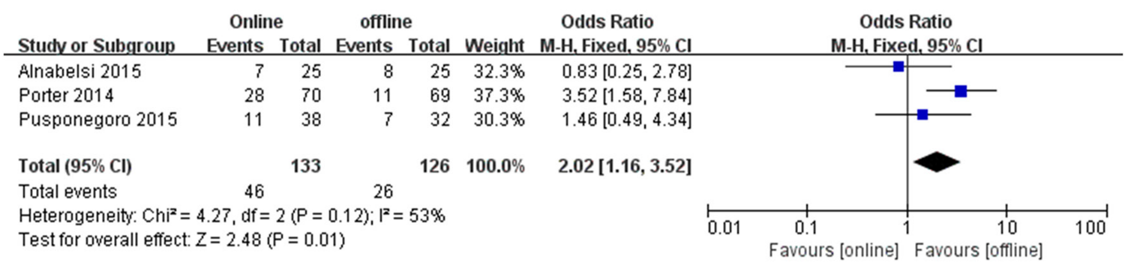

Figure 5 Forest plot for overall satisfaction at the post-test level.

the Newcastle-Ottawa scale were included in the sensitivity analysis. Leave-one-out cross validation was used in the sensitivity analysis to assess the stability of the meta-analysis results. There was no change in the significance of any of the outcomes except for overall satisfaction, which indicated that these meta-results were stable (*Figures 8,9*). When removing the article reported by Porter *et al.* (24), the result was no longer statistically significant (*Figure 10*)

Table 3 Results of meta-analysis comparison of online and offline learning

Outcome	Studies No.	Online group No.	Offline group No.	SMD/OR (95% CI)	P value	Study heterogeneity			
						χ^2	df	I^2 (%)	P value
Knowledge and skills (post-test)	27	1,191	1,117	0.58 (0.25 to 0.91)	0.0006	354.22	26	93	<0.00001
Knowledge gains (pretest/post-test)	5	141	137	1.12 (0.14 to 2.11)	0.02	50.04	4	92	<0.00001
Overall satisfaction	3	133	126	2.02 (1.16. to 3.52)	0.01	4.27	2	53	0.12

SMD/OR, standard mean deviance/odds ratio; df, degrees of freedom; CI, confidence interval.

Table 4 Subgroup analyses of online *vs.* offline education on knowledge and skills acquisitions at the post-test levels

Subgroup	Studies No.	Participants No.	SMD/OR (95% CI)	Study heterogeneity				P value
				χ^2	df	I^2 (%)	P value	
All intervention	27	2,308	0.58 (0.25 to 0.91)	354.22	26	93	<0.00001	0.0006
Study outcome								0.76
Knowledge	23	1,928	0.63 (0.26 to 1.00)	314.58	22	93	<0.00001	0.001
Skills	5	444	0.77 (–0.05 to 1.59)	65.60	4	94	<0.00001	0.07
Study design								0.46
Post-test only	16	1,415	0.47 (0.03 to 0.92)	232.13	15	94	<0.00001	0.04
Pretest/post-test	11	893	0.73 (0.23 to 1.23)	111.68	10	91	<0.00001	0.004
Study type								0.09
RCT	21	1,593	0.35 (0.05 to 0.66)	161.68	20	88	<0.00001	0.02
Non-RCT	6	715	1.27 (0.25 to 2.28)	180.53	5	97	<0.00001	0.01
Participants								0.63
Medical students	20	1,764	0.64 (0.23 to 1.04)	286.80	19	93	<0.00001	0.002
Nurse students	5	356	0.27 (–0.43 to 0.98)	39.86	4	90	<0.00001	0.45
Others	2	188	0.93 (–0.94 to 2.80)	23.82	1	96	<0.00001	0.33
Country								0.14
Developed	22	1,876	0.34 (0.07 to 0.61)	163.03	21	87	<0.00001	0.01
Developing	5	432	1.51 (–0.01 to 3.03)	173.46	4	98	<0.00001	0.05
Course type								0.006
Clinical	18	1,586	0.86 (0.41 to 1.31)	280.74	17	94	<0.00001	0.0002
Foundation	5	472	0.07 (–0.11 to 0.25)	0.4	4	0	0.98	0.44
Other	4	250	0.09 (–1.10 to 1.28)	49.39	3	93	<0.00001	0.88

SMD/OR, standard mean deviance/odds ratio; df, degrees of freedom; CI, confidence interval.

Study or Subgroup	Online Mean	Online SD	Online Total	face-to-face Mean	face-to-face SD	face-to-face Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.5.1 clinical								
Aleman 2011	6.57	0.69	15	6.52	0.96	26	3.6%	0.06 [-0.58, 0.69]
Alnabelsi 2015	89.6	14.07	25	86.4	15.47	25	3.8%	0.21 [-0.34, 0.77]
Assadi 2003	20.24	0.83	41	18.05	1.86	40	3.8%	1.51 [1.02, 2.01]
Bhatti 2011	19.13	3.476	75	18.23	4.159	73	4.0%	0.23 [-0.09, 0.56]
Brettle 2013	1.7714	2.2372	35	2.2286	1.98651	35	3.9%	-0.21 [-0.68, 0.26]
Chao 2012	12.7	4.4	111	11.2	4.5	56	4.0%	0.34 [0.01, 0.66]
Chittenden 2013	2.9	1.1	41	2.8	0.9	33	3.9%	0.10 [-0.36, 0.56]
Farahman 2016	12.5	1.98	60	12.6	1.94	60	4.0%	-0.05 [-0.41, 0.31]
Kaltman 2018	4	1.55	60	3.23	1.86	39	3.9%	0.46 [0.05, 0.86]
Moazami 2014	22.45	4.41	15	19.25	5.11	20	3.6%	0.65 [-0.04, 1.34]
Morente 2013	15.83	2.52	30	11.6	2.39	43	3.8%	1.71 [1.17, 2.26]
Peine 2016	17.23	2.21	61	14.37	2.76	55	4.0%	1.14 [0.75, 1.54]
Portero 2013	24.4	6.2	71	21.2	5.4	43	4.0%	0.54 [0.15, 0.92]
Pusponegoro 2015	16.95	3.178	39	16.88	2.575	36	3.9%	0.02 [-0.43, 0.48]
RAUPACH 2009	31.9	7.2	72	31.7	7.5	71	4.0%	0.03 [-0.30, 0.35]
Solomon 2004	4.88	2	17	4.42	1.08	12	3.5%	0.27 [-0.48, 1.01]
Subramanian 2012	86.7	2	15	61.7	2	15	0.8%	12.16 [8.78, 15.54]
Taradi 2005	25.81	0.66	37	22.08	0.67	84	3.4%	5.56 [4.75, 6.37]
Subtotal (95% CI)			**820**			**766**	**65.8%**	**0.86 [0.41, 1.31]**
Heterogeneity: Tau² = 0.84; Chi² = 280.74, df = 17 (P < 0.00001); I² = 94%								
Test for overall effect: Z = 3.76 (P = 0.0002)								
1.5.2 foundation								
Bjaenr 2013	94.3	6.734	21	93.3	6.253	21	3.7%	0.15 [-0.45, 0.76]
BOWDISH 2003	79.08	12.68	56	78.51	9.44	56	4.0%	0.05 [-0.32, 0.42]
Hu 2016	14.2	2.8	49	13.9	3.4	51	4.0%	0.10 [-0.30, 0.49]
Porter 2014	40.7	4.5	71	40.7	4.8	69	4.0%	0.00 [-0.33, 0.33]
Yeung 2012	42.7	10.5	43	41	11.6	35	3.9%	0.15 [-0.29, 0.60]
Subtotal (95% CI)			**240**			**232**	**19.6%**	**0.07 [-0.11, 0.25]**
Heterogeneity: Tau² = 0.00; Chi² = 0.40, df = 4 (P = 0.98); I² = 0%								
Test for overall effect: Z = 0.77 (P = 0.44)								
1.5.3 other								
Clement 2012	3.58	1.19	71	3.9	0.99	59	4.0%	-0.29 [-0.64, 0.06]
DENNIS 2003	88.47647	4.42133	17	87.01176	5.345078	17	3.6%	0.29 [-0.38, 0.97]
Nicklen 2017	9	0.5	19	10	0.75	19	3.5%	-1.54 [-2.27, -0.80]
Phadtare 2009	75.3	14.21	24	47.27	14.64	24	3.5%	1.91 [1.22, 2.60]
Subtotal (95% CI)			**131**			**119**	**14.6%**	**0.09 [-1.10, 1.28]**
Heterogeneity: Tau² = 1.37; Chi² = 49.39, df = 3 (P < 0.00001); I² = 94%								
Test for overall effect: Z = 0.15 (P = 0.88)								
Total (95% CI)			**1191**			**1117**	**100.0%**	**0.58 [0.25, 0.91]**
Heterogeneity: Tau² = 0.67; Chi² = 354.22, df = 26 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 3.44 (P = 0.0006)								
Test for subgroup differences: Chi² = 10.31, df = 2 (P = 0.006), I² = 80.6%								

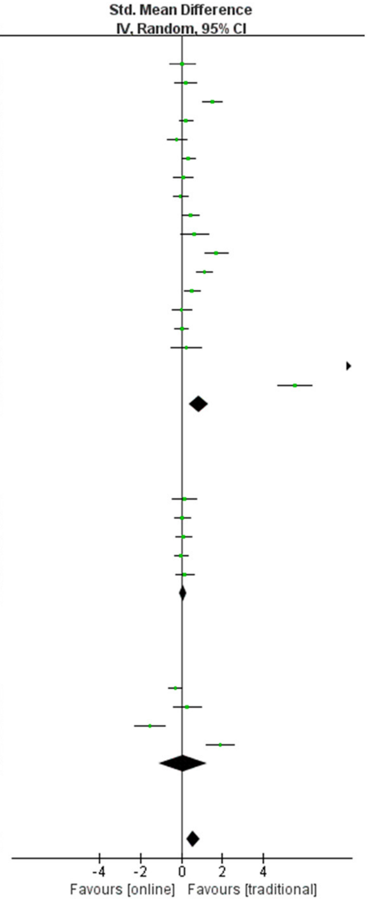

Figure 6 Subgroup analysis of online *vs.* offline education on knowledge and skills acquisitions for course type at the post-test levels.

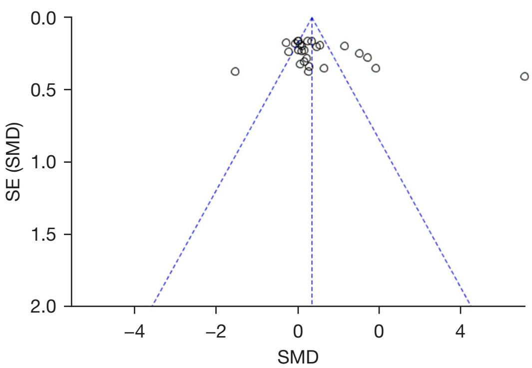

Figure 7 Funnel plots illustrating meta-analysis of knowledge and skills acquisitions.

compared with the original meta-analysis (OR: 1.13; 95% CI: 0.51 to 2.53; P=0.77). This may be caused by a small sample and the forms of online learning and courses of learning were different for each study, there was heterogeneity between the included studies, which may influence the results of the meta-analysis.

Discussion

This meta-analysis of 21 RCTs and 6 CCTs including 2,308 students comparing the efficacy of online learning and offline learning showed that online learning was more effective for undergraduate medical students on post-test scores, pre- and post-test score improvement and overall satisfaction. No factors that significantly impacted the overall results were observed through subgroup analysis. Because the experimental design of the included articles was very different in participants, courses, examination format, and outcome measurement methods, there was considerable heterogeneity among the included studies. However, our sensitivity analysis showed that the results of the meta-

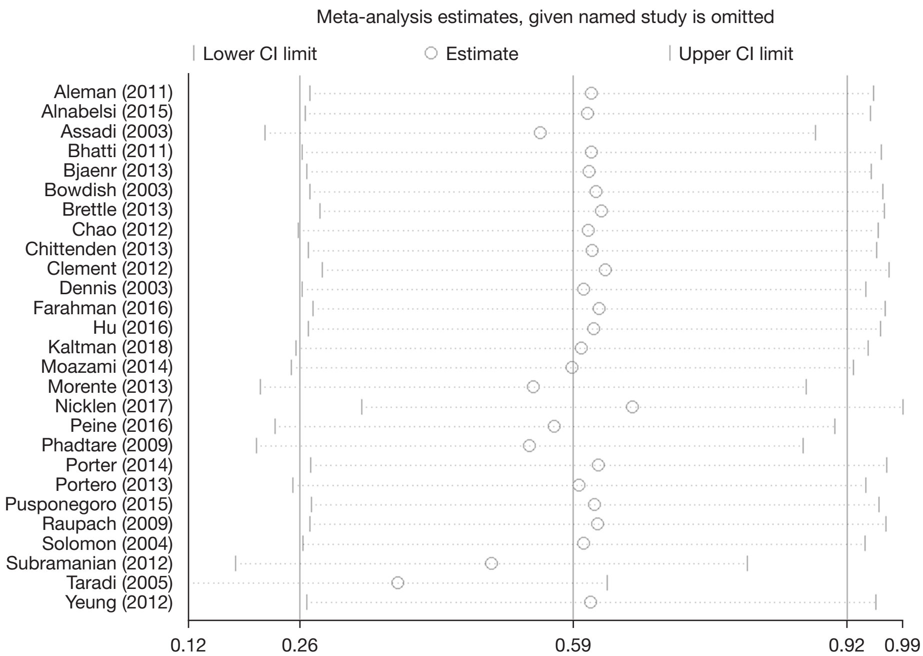

Figure 8 Sensitivity analysis of knowledge and skill score at the post-test level.

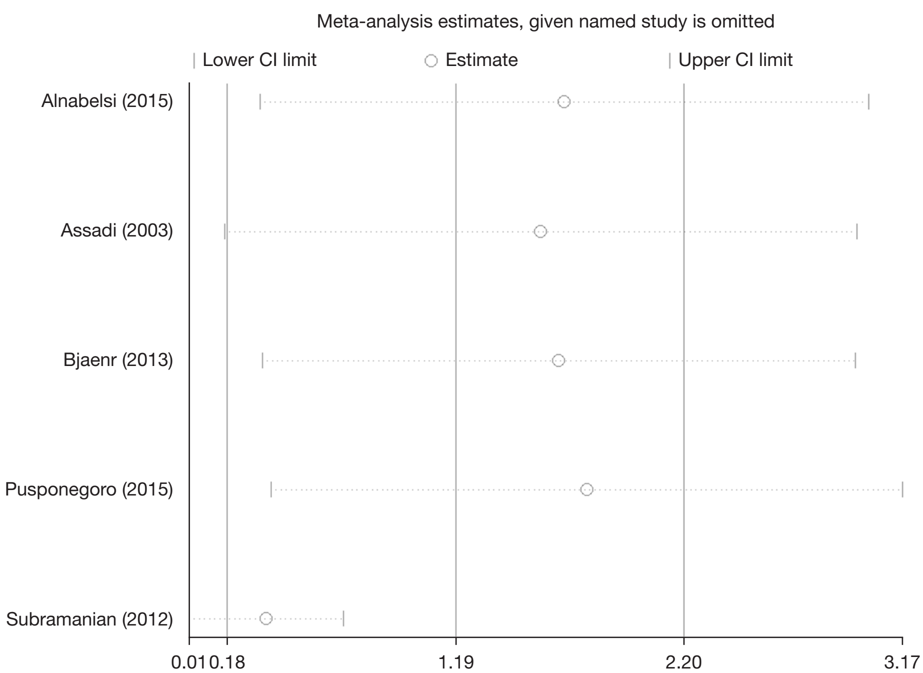

Figure 9 Sensitivity analysis of pre- and post-test score gains.

analysis were robust.

The greatest concerns for medical students' online learning were knowledge acquisition and skill training. It is well known that undergraduate medical courses mainly focus on basic knowledge and skills. In this review, posttest knowledge and skill scores were reported differently

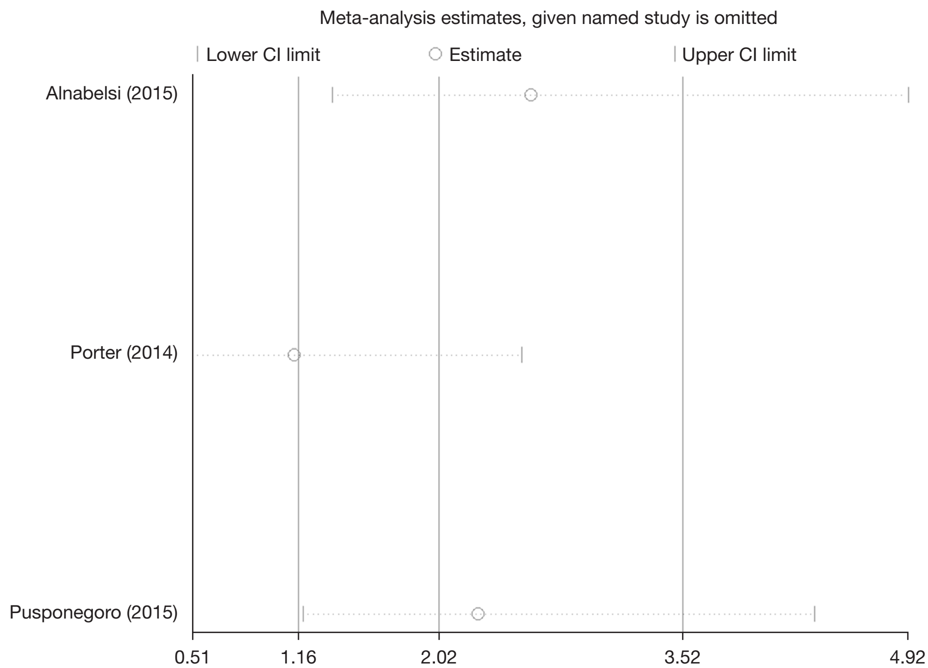

Figure 10 Sensitivity analysis of overall satisfaction.

in each included study. Therefore, we compared these two outcomes between the online and offline groups and found that the posttest scores of the online learning group were significantly higher. Considering prior knowledge or skill levels, the difference between the pre-intervention and post-intervention test scores for each student was calculated and designated as "improvement". The pooled data of improvement included five studies that also showed that online learning students had a significantly higher improvement score. Subramanian *et al.* (25) reported that the average improvement score of the online group was nearly three times that of the offline group and demonstrated that not only was online learning an effective way of learning for medical students compared with the offline format, but it can also promote long-term retention. In most of the studies we included, multiple-choice questions (MCQs) were used as the posttest. The MCQ can not only objectively evaluate students' test scores but also predict objective structured clinical examination (OSCE) scores, which in turn is a powerful predictor of clinical performance (42). The reasons why online learning works better are as follows. First, students can learn about medical knowledge and skills without participating in traditional classroom learning because they can access the information as many times as needed. Second, in addition to the same teaching materials used in online learning, good educational cases, such as representative patients, were also provided. This can prevent certain patients from being suitable for students due to ethical considerations, and there is no need to consider patients who refuse student participation in their care (25,43). In addition, as a novel instructional method, online learning can simulate and practice different clinical situations (experiential learning) (44). However, online learning also has some shortcomings and limitations, and technical problems have made students feel frustrated, so they need technical support related to learning (30). Hence, most of the studies we included were conducted in developed countries, and only five articles (18-20,29,31) were performed in developing countries. Additional problems included having no teacher present, learner isolation, and a lack of peer support and competition (45). These concerns are exacerbated when online methods are used to develop interpersonal and high-level clinical skills, where contextual clinical reasoning is the basis of competence (46).

In addition, although the included studies included medical students of all grades, the knowledge and skills taught in these studies actually only cover a small part of the learning objectives in medical education. Therefore, it is difficult to say that online learning is better than offline learning for topics that have yet to be studied. For online learning mainly composed of static and non-interactive learning resources, these learning resources are similar to offline learning to a large extent; usually, no significant difference was found when compared to offline learning (5). A study conducted by Nesterowicz *et al.* (47) reported that 92% of the subjects believed that online-learning was effective and that the subject of the course was the most important aspect.

In terms of subjective evaluation, contemporary medical students grew up in the Internet era. They are accustomed to the constant stimulation of e-mail, text, and social media, and their experiences affect their behaviour in the classroom. They prefer to listen to podcasts at twice the speed instead of attending lectures to use their time more effectively. They would rather choose a self-paced online training module learning method than using a rigorous 12-week course (22). Our meta-analysis of three studies also showed that the online learning group had a higher rate of overall satisfaction than the offline learning group. In addition to these three studies, Taradi (19) and Phadtare *et al.* (23) gained student satisfaction by surveys and showed that there was a statistically significant difference in the overall satisfaction with the course between the two groups; the online group had a higher overall satisfaction score. However, Raupach *et al.* (40) found that the overall satisfaction score with an online module was low; Nicklen *et al.* (38) also surveyed student satisfaction and showed that 63 percent of those in the intervention group reported a perception that online learning negatively impacted their learning. This variation in student satisfaction may be a result of the different online learning methods, and more similar studies are needed for further confirmation. When students encounter difficulties in using the online learning system, they need technical assistance and learn many things before they are able to use the system, which consumes their learning time and energy.

Currently, the number of people infected with COVID-19 disease is still rising sharply worldwide and there is no vaccine that can effectively prevent the infection of the virus. The global educational centre had not to force to close their classrooms and quickly make changes in medical education to ensure that all students still receive the absolute best level of education possible (48). Moreover, the world is changing, and the causes of education interruptions are not limited to epidemics; wars, regional conflicts, and various types of natural disasters are issues that should be kept on the future agenda as potential sources of interruption (49). Online learning has been the best choice to maintain regular teaching and learning (1). This review further confirms that online learning is more effective than offline learning in undergraduate medical education.

Despite the valuable conclusions drawn, the meta-analysis still has some limitations. First, our study included controlled clinical trials (CCTs), which may not be adequately powered. Second, educators who achieve good results with online learning tend to publish their results, which may result in potential publication bias. Third, because the forms of online learning and courses of learning were different for each study, there was heterogeneity between the included studies, which may influence the results of the meta-analysis. Random effects model can only address statistical heterogeneity but the heterogeneity caused by different ways of online learning cannot be addressed via statistical analysis. Last, the included studies in our review were not conducted under the circumstance of the COVID-19 pandemic. Therefore, it is difficult to conclude that online learning is more effective than offline learning for those courses influenced by COVID-19. More comparative studies conducted in the context of the COVID-19 pandemic are needed.

Conclusions

In summary, our meta-analysis demonstrates that online learning methods in medical education could achieve higher knowledge and skill scores at the posttest level than offline learning methods. In addition, it also has higher satisfaction ratings than offline education, indicating that contemporary medical students prefer this education mode. Through subgroup analysis, no significant factors were observed except the subject of the course, which indicates that not all courses are suitable for online learning.

Acknowledgments

The authors are grateful to the staff of Xiangya Hospital Central South University and all those who actively participated in this study.

Funding: The study was supported by Education Reform Research Project of Central South University (No. 2022JGB051).

Footnote

Reporting Checklist: The authors have completed the

PRISMA reporting checklist. Available at https://jxym.amegroups.com/article/view/10.21037/jxym-22-3/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://jxym.amegroups.com/article/view/10.21037/jxym-22-3/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Al-Balas M, Al-Balas HI, Jaber HM, et al. Distance learning in clinical medical education amid COVID-19 pandemic in Jordan: current situation, challenges, and perspectives. BMC Med Educ 2020;20:341.
2. Crisp N, Gawanas B, Sharp I, et al. Training the health workforce: scaling up, saving lives. Lancet 2008;371:689-91.
3. Car J, Carlstedt-Duke J, Tudor Car L, et al. Digital Education in Health Professions: The Need for Overarching Evidence Synthesis. J Med Internet Res 2019;21:e12913.
4. Choules AP. The use of elearning in medical education: a review of the current situation. Postgrad Med J 2007;83:212-6.
5. Pei L, Wu H. Does online learning work better than offline learning in undergraduate medical education? A systematic review and meta-analysis. Med Educ Online 2019;24:1666538.
6. He L, Yang N, Xu L, et al. Synchronous distance education vs traditional education for health science students: A systematic review and meta-analysis. Med Educ 2021;55:293-308.
7. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009;339:b2700.
8. Shuster JJ. Review: Cochrane handbook for systematic reviews for interventions, Version 5.1.0, published 3/2011. Julian P.T. Higgins and Sally Green, editors. Research Synthesis Methods 2011;2:126-30.
9. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. Eur J Epidemiol 2010;25:603-5.
10. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ 2003;327:557-60.
11. Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol 2005;5:13.
12. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539-58.
13. Zaykin DV. Optimally weighted Z-test is a powerful method for combining probabilities in meta-analysis. J Evol Biol 2011;24:1836-41.
14. Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629-34.
15. Bowdish BE, Chauvin SW, Kreisman N, et al. GTravels towards Problem Based Learning in Medical Education (VPBL). Instructional Science 2003;31:231-53.
16. Sendra-Portero F, Torales-Chaparro OE, Ruiz-Gómez MJ, et al. A pilot study to evaluate the use of virtual lectures for undergraduate radiology teaching. Eur J Radiol 2013;82:888-93.
17. Chao SH, Brett B, Wiecha JM, et al. Use of an online curriculum to teach delirium to fourth-year medical students: a comparison with lecture format. J Am Geriatr Soc 2012;60:1328-32.
18. Farahmand S, Jalili E, Arbab M, et al. Distance Learning Can Be as Effective as Traditional Learning for Medical Students in the Initial Assessment of Trauma Patients. Acta Med Iran 2016;54:600-4.
19. Taradi SK, Taradi M, Radic K, et al. Blending problem-based learning with Web technology positively impacts student learning outcomes in acid-base physiology. Adv Physiol Educ 2005;29:35-9.
20. Assadi T, Mofidi M, Rezai M, et al. The Comparison between two Methods of Basic Life Support Instruction: Video Self-Instruction versus Traditional Method. Hong Kong Journal of Emergency Medicine 2015;22:291-6.
21. Brettle A, Raynor M. Developing information literacy

skills in pre-registration nurses: an experimental study of teaching methods. Nurse Educ Today 2013;33:103-9.
22. Hu A, Shewokis PA, Ting K, et al. Motivation in computer-assisted instruction. Laryngoscope 2016;126 Suppl 6:S5-S13.
23. Phadtare A, Bahmani A, Shah A, et al. Scientific writing: a randomized controlled trial comparing standard and on-line instruction. BMC Med Educ 2009;9:27.
24. Porter AL, Pitterle ME, Hayney MS. Comparison of online versus classroom delivery of an immunization elective course. Am J Pharm Educ 2014;78:96.
25. Subramanian A, Timberlake M, Mittakanti H, et al. Novel educational approach for medical students: improved retention rates using interactive medical software compared with traditional lecture-based format. J Surg Educ 2012;69:449-52.
26. Worm BS. Learning from simple ebooks, online cases or classroom teaching when acquiring complex knowledge. A randomized controlled trial in respiratory physiology and pulmonology. PLoS One 2013;8:e73336.
27. Chittenden EH, Anderson WG, Lai CJ, et al. An evaluation of interactive web-based curricula for teaching code status discussions. J Palliat Med 2013;16:1070-3.
28. Solomon DJ, Ferenchick GS, Laird-Fick HS, et al. A randomized trial comparing digital and live lecture formats ISRCTN40455708. BMC Med Educ 2004;4:27.
29. Moazami F, Bahrampour E, Azar MR, et al. Comparing two methods of education (virtual versus traditional) on learning of Iranian dental students: a post-test only design study. BMC Med Educ 2014;14:45.
30. Fernández Alemán JL, Carrillo de Gea JM, Rodríguez Mondéjar JJ. Effects of competitive computer-assisted learning versus conventional teaching methods on the acquisition and retention of knowledge in medical surgical nursing students. Nurse Educ Today 2011;31:866-71.
31. Pusponegoro HD, Soebadi A, Surya R. Web-Based Versus Conventional Training for Medical Students on Infant Gross Motor Screening. Telemed J E Health 2015;21:992-7.
32. Bhatti I, Jones K, Richardson L, et al. E-learning vs lecture: which is the best approach to surgical teaching? Colorectal Dis 2011;13:459-62.
33. Dennis JK. Problem-based learning in online vs. face-to-face environments. Educ Health (Abingdon) 2003;16:198-209.
34. Yeung JC, Fung K, Wilson TD. Prospective evaluation of a web-based three-dimensional cranial nerve simulation. J Otolaryngol Head Neck Surg 2012;41:426-36.
35. Kaltman S, Talisman N, Pennestri S, et al. Using Technology to Enhance Teaching of Patient-Centered Interviewing for Early Medical Students. Simul Healthc 2018;13:188-94.
36. Morente L, Morales-Asencio JM, Veredas FJ. Effectiveness of an e-learning tool for education on pressure ulcer evaluation. J Clin Nurs 2014;23:2043-52.
37. Peine A, Kabino K, Spreckelsen C. Self-directed learning can outperform direct instruction in the course of a modern German medical curriculum - results of a mixed methods trial. BMC Med Educ 2016;16:158.
38. Nicklen P, Keating JL, Paynter S, et al. Remote-online case-based learning: A comparison of remote-online and face-to-face, case-based learning - a randomized controlled trial. Educ Health (Abingdon) 2016;29:195-202.
39. Clement S, van Nieuwenhuizen A, Kassam A, et al. Filmed v. live social contact interventions to reduce stigma: randomised controlled trial. Br J Psychiatry 2012;201:57-64.
40. Raupach T, Muenscher C, Anders S, et al. Web-based collaborative training of clinical reasoning: a randomized trial. Med Teach 2009;31:e431-7.
41. Alnabelsi T, Al-Hussaini A, Owens D. Comparison of traditional face-to-face teaching with synchronous e-learning in otolaryngology emergencies teaching to medical undergraduates: a randomised controlled trial. Eur Arch Otorhinolaryngol 2015;272:759-63.
42. Wilkinson TJ, Frampton CM. Comprehensive undergraduate medical assessments improve prediction of clinical performance. Med Educ 2004;38:1111-6.
43. Zary N, Johnson G, Boberg J, et al. Development, implementation and pilot evaluation of a Web-based Virtual Patient Case Simulation environment--Web-SP. BMC Med Educ 2006;6:10.
44. Ruiz JG, Smith M, van Zuilen MH, et al. The educational impact of a computer-based training tutorial on dementia in long term care for licensed practice nursing students. Gerontol Geriatr Educ 2006;26:67-79.
45. Vaona A, Banzi R, Kwag KH, et al. E-learning for health professionals. Cochrane Database Syst Rev 2018;1:CD011736.
46. Grimmer-Somers K, Milanese S, Chipcase L. Research into Best Practices in e-Learning for Allied Health clinical education and training. Brisbane: Clinical Education and Training Queensland, 2011.
47. Nesterowicz K, Librowski T, Edelbring S. Validating

e-learning in continuing pharmacy education: user acceptance and knowledge change. BMC Med Educ 2014;14:33.
48. Theoret C, Ming X. Our education, our concerns: The impact on medical student education of COVID-19. Med Educ 2020;54:591-2.
49. Bozkurt A, Sharma RC. Emergency remote teaching in a time of global crisis due to CoronaVirus pandemic. Asian Journal of Distance Education 2020;15:1-4.

doi: 10.21037/jxym-22-3
Cite this article as: Gao M, Cui Y, Chen H, Zeng H, Zhu Z, Zu X. The efficacy and acceptance of online learning *vs.* offline learning in medical student education: a systematic review and meta-analysis. J Xiangya Med 2022;7:13.

Study or Subgroup	Online Mean	Online SD	Online Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.4.1 knowledge								
Aleman 2011	6.57	0.69	15	6.52	0.96	26	3.5%	0.06 [-0.58, 0.69]
Alnabelsi 2015	89.6	14.07	25	86.4	15.47	25	3.6%	0.21 [-0.34, 0.77]
Bhatti 2011	19.13	3.476	75	18.23	4.159	73	3.9%	0.23 [-0.09, 0.56]
Bjaenr 2013	94.3	6.734	21	93.3	6.253	21	3.6%	0.15 [-0.45, 0.76]
BOWDISH 2003	79.08	12.68	56	78.51	9.44	56	3.8%	0.05 [-0.32, 0.42]
Chao 2012	12.7	4.4	111	11.2	4.5	56	3.9%	0.34 [0.01, 0.66]
Clement 2012	2.9	1.1	41	2.8	0.9	33	3.7%	0.10 [-0.36, 0.56]
DENNIS 2003	88.47647	4.42133	17	87.01176	5.345078	17	3.5%	0.29 [-0.38, 0.97]
Farahman 2016	12.5	1.98	60	12.6	1.94	60	3.8%	-0.05 [-0.41, 0.31]
Hu 2016	14.2	2.8	49	13.9	3.4	51	3.8%	0.10 [-0.30, 0.49]
Moazami 2014	22.45	4.41	15	19.25	5.11	20	3.4%	0.65 [-0.04, 1.34]
Morente 2013	15.83	2.52	30	11.6	2.39	43	3.6%	1.71 [1.17, 2.26]
Nicklen 2017	9	0.5	19	10	0.75	19	3.4%	-1.54 [-2.27, -0.80]
Peine 2016	17.23	2.21	61	14.37	2.76	55	3.8%	1.14 [0.75, 1.54]
Phadtare 2009	75.3	14.21	24	47.27	14.64	24	3.4%	1.91 [1.22, 2.60]
Porter 2014	40.7	4.5	71	40.7	4.8	69	3.9%	0.00 [-0.33, 0.33]
Portero 2013	24.4	6.2	71	21.2	5.4	43	3.8%	0.54 [0.15, 0.92]
Pusponegoro 2015	16.95	3.178	39	16.88	2.575	36	3.7%	0.02 [-0.43, 0.48]
RAUPACH 2009	31.9	7.2	72	31.7	7.5	71	3.9%	0.03 [-0.30, 0.35]
Solomon 2004	4.88	2	17	4.42	1.08	12	3.4%	0.27 [-0.48, 1.01]
Subramanian 2012	86.7	2	15	61.7	2	15	0.8%	12.16 [8.78, 15.54]
Taradi 2005	25.81	0.66	37	22.08	0.67	84	3.3%	5.56 [4.75, 6.37]
Yeung 2012	42.7	10.5	43	41	11.6	35	3.8%	0.15 [-0.29, 0.60]
Subtotal (95% CI)			**984**			**944**	**81.3%**	**0.63 [0.26, 1.00]**
Heterogeneity: Tau² = 0.73; Chi² = 314.58, df = 22 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 3.30 (P = 0.0010)								
1.4.2 skills								
Assadi 2003	20.24	0.83	41	18.05	1.86	40	3.7%	1.51 [1.02, 2.01]
Brettle 2013	1.7714	2.2372	35	2.2286	1.98651	35	3.7%	-0.21 [-0.68, 0.26]
Chittenden 2013	2.9	1.1	41	2.8	0.9	33	3.7%	0.10 [-0.36, 0.56]
Farahman 2016	16.5	1.96	60	12.3	2.22	60	3.8%	1.99 [1.55, 2.43]
Kaltman 2018	4	1.55	60	3.23	1.86	39	3.8%	0.46 [0.05, 0.86]
Subtotal (95% CI)			**237**			**207**	**18.7%**	**0.77 [-0.05, 1.59]**
Heterogeneity: Tau² = 0.82; Chi² = 65.60, df = 4 (P < 0.00001); I² = 94%								
Test for overall effect: Z = 1.84 (P = 0.07)								
Total (95% CI)			**1221**			**1151**	**100.0%**	**0.66 [0.32, 0.99]**
Heterogeneity: Tau² = 0.74; Chi² = 393.13, df = 27 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 3.80 (P = 0.0001)								
Test for subgroup differences: Chi² = 0.09, df = 1 (P = 0.76), I² = 0%								

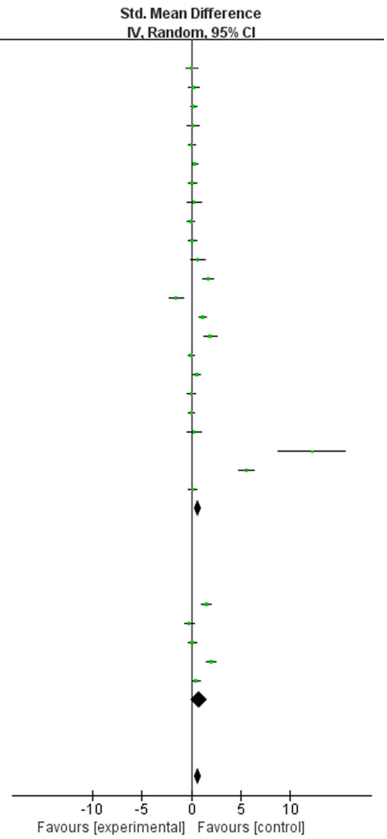

Figure S1 Subgroup analysis on study outcome of online *vs.* offline education on knowledge and skills acquisitions at the post-test levels.

Study or Subgroup	Online Mean	Online SD	Online Total	face-to-face Mean	face-to-face SD	face-to-face Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.3.1 Developed								
Aleman 2011	6.57	0.69	15	6.52	0.96	26	3.6%	0.06 [-0.58, 0.69]
Alnabelsi 2015	89.6	14.07	25	86.4	15.47	25	3.8%	0.21 [-0.34, 0.77]
Bhatti 2011	19.13	3.476	75	18.23	4.159	73	4.0%	0.23 [-0.09, 0.56]
Bjaenr 2013	94.3	6.734	21	93.3	6.253	21	3.7%	0.15 [-0.45, 0.76]
BOWDISH 2003	79.08	12.68	56	78.51	9.44	56	4.0%	0.05 [-0.32, 0.42]
Brettle 2013	1.7714	2.2372	35	2.2286	1.98651	35	3.9%	-0.21 [-0.68, 0.26]
Chao 2012	12.7	4.4	111	11.2	4.5	56	4.0%	0.34 [0.01, 0.66]
Chittenden 2013	2.9	1.1	41	2.8	0.9	33	3.9%	0.10 [-0.36, 0.56]
Clement 2012	3.58	1.19	71	3.9	0.99	59	4.0%	-0.29 [-0.64, 0.06]
DENNIS 2003	88.47647	4.42133	17	87.01176	5.345078	17	3.6%	0.29 [-0.38, 0.97]
Hu 2016	14.2	2.8	49	13.9	3.4	51	4.0%	0.10 [-0.30, 0.49]
Kaltman 2018	4	1.55	60	3.23	1.86	39	3.9%	0.46 [0.05, 0.86]
Morente 2013	15.83	2.52	30	11.6	2.39	43	3.8%	1.71 [1.17, 2.26]
Nicklen 2017	9	0.5	19	10	0.75	19	3.5%	-1.54 [-2.27, -0.80]
Peine 2016	17.23	2.21	61	14.37	2.76	55	4.0%	1.14 [0.75, 1.54]
Phadtare 2009	75.3	14.21	24	47.27	14.64	24	3.5%	1.91 [1.22, 2.60]
Porter 2014	40.7	4.5	71	40.7	4.8	69	4.0%	0.00 [-0.33, 0.33]
Portero 2013	24.4	6.2	71	21.2	5.4	43	4.0%	0.54 [0.15, 0.92]
RAUPACH 2009	31.9	7.2	72	31.7	7.5	71	4.0%	0.03 [-0.30, 0.35]
Solomon 2004	4.88	2	17	4.42	1.08	12	3.5%	0.27 [-0.48, 1.01]
Subramanian 2012	86.7	2	15	61.7	2	15	0.8%	12.16 [8.78, 15.54]
Yeung 2012	42.7	10.5	43	41	11.6	35	3.9%	0.15 [-0.29, 0.60]
Subtotal (95% CI)			**999**			**877**	**81.4%**	**0.34 [0.07, 0.61]**
Heterogeneity: Tau² = 0.35; Chi² = 163.03, df = 21 (P < 0.00001); I² = 87%								
Test for overall effect: Z = 2.44 (P = 0.01)								
1.3.2 Developing								
Assadi 2003	20.24	0.83	41	18.05	1.86	40	3.8%	1.51 [1.02, 2.01]
Farahman 2016	12.5	1.98	60	12.6	1.94	60	4.0%	-0.05 [-0.41, 0.31]
Moazami 2014	22.45	4.41	15	19.25	5.11	20	3.6%	0.65 [-0.04, 1.34]
Pusponegoro 2015	16.95	3.178	39	16.88	2.575	36	3.9%	0.02 [-0.43, 0.48]
Taradi 2005	25.81	0.66	37	22.08	0.67	84	3.4%	5.56 [4.75, 6.37]
Subtotal (95% CI)			**192**			**240**	**18.6%**	**1.51 [-0.01, 3.03]**
Heterogeneity: Tau² = 2.93; Chi² = 173.46, df = 4 (P < 0.00001); I² = 98%								
Test for overall effect: Z = 1.94 (P = 0.05)								
Total (95% CI)			**1191**			**1117**	**100.0%**	**0.58 [0.25, 0.91]**
Heterogeneity: Tau² = 0.67; Chi² = 354.22, df = 26 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 3.44 (P = 0.0006)								
Test for subgroup differences: Chi² = 2.19, df = 1 (P = 0.14), I² = 54.4%								

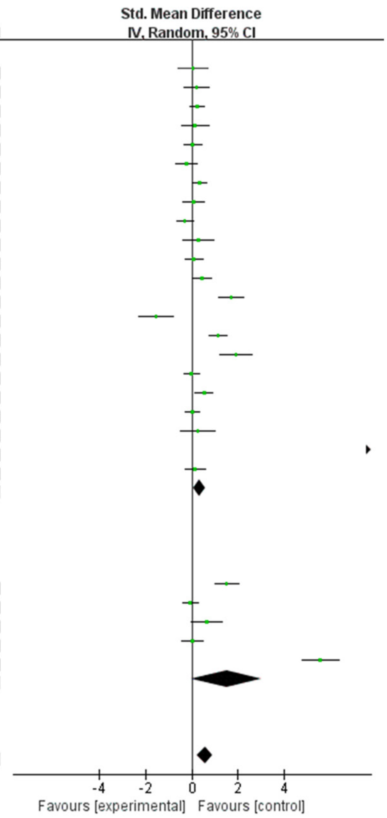

Figure S2 Subgroup analysis on country of online *vs.* offline education on knowledge and skills acquisitions at the post-test levels.

Study or Subgroup	Online Mean	Online SD	Online Total	face-to-face Mean	face-to-face SD	face-to-face Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.2.1 Medical students								
Alnabelsi 2015	89.6	14.07	25	86.4	15.47	25	3.8%	0.21 [-0.34, 0.77]
Assadi 2003	20.24	0.83	41	18.05	1.86	40	3.8%	1.51 [1.02, 2.01]
Bhatti 2011	19.13	3.476	75	18.23	4.159	73	4.0%	0.23 [-0.09, 0.56]
BOWDISH 2003	79.08	12.68	56	78.51	9.44	56	4.0%	0.05 [-0.32, 0.42]
Chao 2012	12.7	4.4	111	11.2	4.5	56	4.0%	0.34 [0.01, 0.66]
Chittenden 2013	2.9	1.1	41	2.8	0.9	33	3.9%	0.10 [-0.36, 0.56]
DENNIS 2003	88.47647	4.42133	17	87.01176	5.345078	17	3.6%	0.29 [-0.38, 0.97]
Farahman 2016	12.5	1.98	60	12.6	1.94	60	4.0%	-0.05 [-0.41, 0.31]
Hu 2016	14.2	2.8	49	13.9	3.4	51	4.0%	0.10 [-0.30, 0.49]
Kaltman 2018	4	1.55	60	3.23	1.86	39	3.9%	0.46 [0.05, 0.86]
Moazami 2014	22.45	4.41	15	19.25	5.11	20	3.6%	0.65 [-0.04, 1.34]
Nicklen 2017	9	0.5	19	10	0.75	19	3.5%	-1.54 [-2.27, -0.80]
Peine 2016	17.23	2.21	61	14.37	2.76	55	4.0%	1.14 [0.75, 1.54]
Portero 2013	24.4	6.2	71	21.2	5.4	43	4.0%	0.54 [0.15, 0.92]
Pusponegoro 2015	16.95	3.178	39	16.88	2.575	36	3.9%	0.02 [-0.43, 0.48]
RAUPACH 2009	31.9	7.2	72	31.7	7.5	71	4.0%	0.03 [-0.30, 0.35]
Solomon 2004	4.88	2	17	4.42	1.08	12	3.5%	0.27 [-0.48, 1.01]
Subramanian 2012	86.7	2	15	61.7	2	15	0.8%	12.16 [8.78, 15.54]
Taradi 2005	25.81	0.66	37	22.08	0.67	84	3.4%	5.56 [4.75, 6.37]
Yeung 2012	42.7	10.5	43	41	11.6	35	3.9%	0.15 [-0.29, 0.60]
Subtotal (95% CI)			**924**			**840**	**73.4%**	**0.64 [0.23, 1.04]**
Heterogeneity: Tau² = 0.74; Chi² = 286.80, df = 19 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 3.10 (P = 0.002)								
1.2.2 Nurse students								
Aleman 2011	6.57	0.69	15	6.52	0.96	26	3.6%	0.06 [-0.58, 0.69]
Bjaenr 2013	94.3	6.734	21	93.3	6.253	21	3.7%	0.15 [-0.45, 0.76]
Brettle 2013	1.7714	2.2372	35	2.2286	1.98651	35	3.9%	-0.21 [-0.68, 0.26]
Clement 2012	3.58	1.19	71	3.9	0.99	59	4.0%	-0.29 [-0.64, 0.06]
Morente 2013	15.83	2.52	30	11.6	2.39	43	3.8%	1.71 [1.17, 2.26]
Subtotal (95% CI)			**172**			**184**	**19.0%**	**0.27 [-0.43, 0.98]**
Heterogeneity: Tau² = 0.58; Chi² = 39.86, df = 4 (P < 0.00001); I² = 90%								
Test for overall effect: Z = 0.76 (P = 0.45)								
1.2.3 Others								
Phadtare 2009	75.3	14.21	24	47.27	14.64	24	3.5%	1.91 [1.22, 2.60]
Porter 2014	40.7	4.5	71	40.7	4.8	69	4.0%	0.00 [-0.33, 0.33]
Subtotal (95% CI)			**95**			**93**	**7.6%**	**0.93 [-0.94, 2.80]**
Heterogeneity: Tau² = 1.75; Chi² = 23.82, df = 1 (P < 0.00001); I² = 96%								
Test for overall effect: Z = 0.97 (P = 0.33)								
Total (95% CI)			**1191**			**1117**	**100.0%**	**0.58 [0.25, 0.91]**
Heterogeneity: Tau² = 0.67; Chi² = 354.22, df = 26 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 3.44 (P = 0.0006)								
Test for subgroup differences: Chi² = 0.91, df = 2 (P = 0.63), I² = 0%								

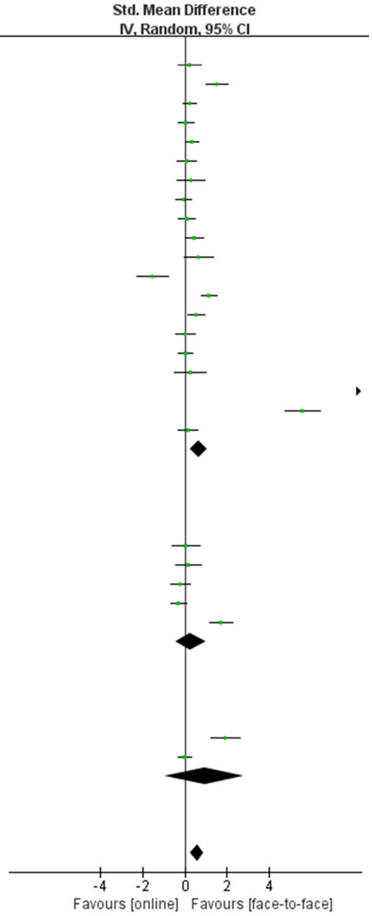

Figure S3 Subgroup analysis on participant of online *vs.* offline education on knowledge and skills acquisitions at the post-test levels.

https://dx.doi.org/10.21037/jxym-22-3

Study or Subgroup	Online Mean	Online SD	Online Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.8.1 post-test only								
BOWDISH 2003	79.08	12.68	56	78.51	9.44	56	4.0%	0.05 [-0.32, 0.42]
Chittenden 2013	2.9	1.1	41	2.8	0.9	33	3.9%	0.10 [-0.36, 0.56]
Clement 2012	3.58	1.19	71	3.9	0.99	59	4.0%	-0.29 [-0.64, 0.06]
DENNIS 2003	88.47647	4.42133	17	87.01176	5.345078	17	3.6%	0.29 [-0.38, 0.97]
Farahman 2016	12.5	1.98	60	12.6	1.94	60	4.0%	-0.05 [-0.41, 0.31]
Hu 2016	14.2	2.8	49	13.9	3.4	51	4.0%	0.10 [-0.30, 0.49]
Kaltman 2018	4	1.55	60	3.23	1.86	39	3.9%	0.46 [0.05, 0.86]
Moazami 2014	22.45	4.41	15	19.25	5.11	20	3.6%	0.65 [-0.04, 1.34]
Nicklen 2017	9	0.5	19	10	0.75	19	3.5%	-1.54 [-2.27, -0.80]
Phadtare 2009	75.3	14.21	24	47.27	14.64	24	3.5%	1.91 [1.22, 2.60]
Porter 2014	40.7	4.5	71	40.7	4.8	69	4.0%	0.00 [-0.33, 0.33]
Portero 2013	24.4	6.2	71	21.2	5.4	43	4.0%	0.54 [0.15, 0.92]
RAUPACH 2009	31.9	7.2	72	31.7	7.5	71	4.0%	0.03 [-0.30, 0.35]
Solomon 2004	4.88	2	17	4.42	1.08	12	3.5%	0.27 [-0.48, 1.01]
Taradi 2005	25.81	0.66	37	22.08	0.67	84	3.4%	5.56 [4.75, 6.37]
Yeung 2012	42.7	10.5	43	41	11.6	35	3.9%	0.15 [-0.29, 0.60]
Subtotal (95% CI)			**723**			**692**	**60.7%**	**0.47 [0.03, 0.92]**
Heterogeneity: Tau² = 0.75; Chi² = 232.13, df = 15 (P < 0.00001); I² = 94%								
Test for overall effect: Z = 2.09 (P = 0.04)								
1.8.2 pretest/post-test								
Aleman 2011	6.57	0.69	15	6.52	0.96	26	3.6%	0.06 [-0.58, 0.69]
Alnabelsi 2015	89.6	14.07	25	86.4	15.47	25	3.8%	0.21 [-0.34, 0.77]
Assadi 2003	20.24	0.83	41	18.05	1.86	40	3.8%	1.51 [1.02, 2.01]
Bhatti 2011	19.13	3.476	75	18.23	4.159	73	4.0%	0.23 [-0.09, 0.56]
Bjaenr 2013	94.3	6.734	21	93.3	6.253	21	3.7%	0.15 [-0.45, 0.76]
Brettle 2013	1.7714	2.2372	35	2.2286	1.98651	35	3.9%	-0.21 [-0.68, 0.26]
Chao 2012	12.7	4.4	111	11.2	4.5	56	4.0%	0.34 [0.01, 0.66]
Morente 2013	15.83	2.52	30	11.6	2.39	43	3.8%	1.71 [1.17, 2.26]
Peine 2016	17.23	2.21	61	14.37	2.76	55	4.0%	1.14 [0.75, 1.54]
Pusponegoro 2015	16.95	3.178	39	16.88	2.575	36	3.9%	0.02 [-0.43, 0.48]
Subramanian 2012	86.7	2	15	61.7	2	15	0.8%	12.16 [8.78, 15.54]
Subtotal (95% CI)			**468**			**425**	**39.3%**	**0.73 [0.23, 1.23]**
Heterogeneity: Tau² = 0.59; Chi² = 111.68, df = 10 (P < 0.00001); I² = 91%								
Test for overall effect: Z = 2.87 (P = 0.004)								
Total (95% CI)			**1191**			**1117**	**100.0%**	**0.58 [0.25, 0.91]**
Heterogeneity: Tau² = 0.67; Chi² = 354.22, df = 26 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 3.44 (P = 0.0006)								
Test for subgroup differences: Chi² = 0.56, df = 1 (P = 0.46), I² = 0%								

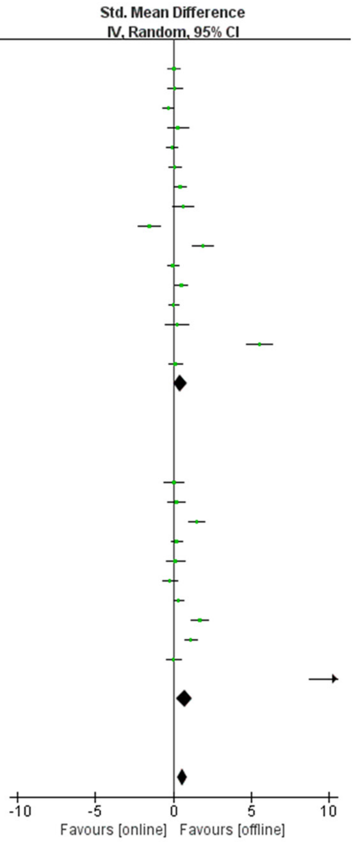

Figure S4 Subgroup analysis on study design of online *vs.* offline education on knowledge and skills acquisitions at the post-test levels.

https://dx.doi.org/10.21037/jxym-22-3

Study or Subgroup	Online Mean	Online SD	Online Total	Offline Mean	Offline SD	Offline Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.9.1 RCT								
Aleman 2011	6.57	0.69	15	6.52	0.96	26	3.6%	0.06 [-0.58, 0.69]
Alnabelsi 2015	89.6	14.07	25	86.4	15.47	25	3.8%	0.21 [-0.34, 0.77]
Bhatti 2011	19.13	3.476	75	18.23	4.159	73	4.0%	0.23 [-0.09, 0.56]
Bjaenr 2013	94.3	6.734	21	93.3	6.253	21	3.7%	0.15 [-0.45, 0.76]
Brettle 2013	1.7714	2.2372	35	2.2286	1.98651	35	3.9%	-0.21 [-0.68, 0.26]
Chittenden 2013	2.9	1.1	41	2.8	0.9	33	3.9%	0.10 [-0.36, 0.56]
Clement 2012	3.58	1.19	71	3.9	0.99	59	4.0%	-0.29 [-0.64, 0.06]
DENNIS 2003	88.47647	4.42133	17	87.01176	5.345078	17	3.6%	0.29 [-0.38, 0.97]
Hu 2016	14.2	2.8	49	13.9	3.4	51	4.0%	0.10 [-0.30, 0.49]
Kaltman 2018	4	1.55	60	3.23	1.86	39	3.9%	0.46 [0.05, 0.86]
Moazami 2014	22.45	4.41	15	19.25	5.11	20	3.6%	0.65 [-0.04, 1.34]
Morente 2013	15.83	2.52	30	11.6	2.39	43	3.8%	1.71 [1.17, 2.26]
Nicklen 2017	9	0.5	19	10	0.75	19	3.5%	-1.54 [-2.27, -0.80]
Peine 2016	17.23	2.21	61	14.37	2.76	55	4.0%	1.14 [0.75, 1.54]
Phadtare 2009	75.3	14.21	24	47.27	14.64	24	3.5%	1.91 [1.22, 2.60]
Porter 2014	40.7	4.5	71	40.7	4.8	69	4.0%	0.00 [-0.33, 0.33]
Pusponegoro 2015	16.95	3.178	39	16.88	2.575	36	3.9%	0.02 [-0.43, 0.48]
RAUPACH 2009	31.9	7.2	72	31.7	7.5	71	4.0%	0.03 [-0.30, 0.35]
Solomon 2004	4.88	2	17	4.42	1.08	12	3.5%	0.27 [-0.48, 1.01]
Subramanian 2012	86.7	2	15	61.7	2	15	0.8%	12.16 [8.78, 15.54]
Yeung 2012	42.7	10.5	43	41	11.6	35	3.9%	0.15 [-0.29, 0.60]
Subtotal (95% CI)			**815**			**778**	**76.8%**	**0.35 [0.05, 0.66]**
Heterogeneity: Tau² = 0.41; Chi² = 161.68, df = 20 (P < 0.00001); I² = 88%								
Test for overall effect: Z = 2.29 (P = 0.02)								
1.9.2 non-RCT								
Assadi 2003	20.24	0.83	41	18.05	1.86	40	3.8%	1.51 [1.02, 2.01]
BOWDISH 2003	79.08	12.68	56	78.51	9.44	56	4.0%	0.05 [-0.32, 0.42]
Chao 2012	12.7	4.4	111	11.2	4.5	56	4.0%	0.34 [0.01, 0.66]
Farahman 2016	12.5	1.98	60	12.6	1.94	60	4.0%	-0.05 [-0.41, 0.31]
Portero 2013	24.4	6.2	71	21.2	5.4	43	4.0%	0.54 [0.15, 0.92]
Taradi 2005	25.81	0.66	37	22.08	0.67	84	3.4%	5.56 [4.75, 6.37]
Subtotal (95% CI)			**376**			**339**	**23.2%**	**1.27 [0.25, 2.28]**
Heterogeneity: Tau² = 1.54; Chi² = 180.53, df = 5 (P < 0.00001); I² = 97%								
Test for overall effect: Z = 2.45 (P = 0.01)								
Total (95% CI)			**1191**			**1117**	**100.0%**	**0.58 [0.25, 0.91]**
Heterogeneity: Tau² = 0.67; Chi² = 354.22, df = 26 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 3.44 (P = 0.0006)								
Test for subgroup differences: Chi² = 2.87, df = 1 (P = 0.09), I² = 65.1%								

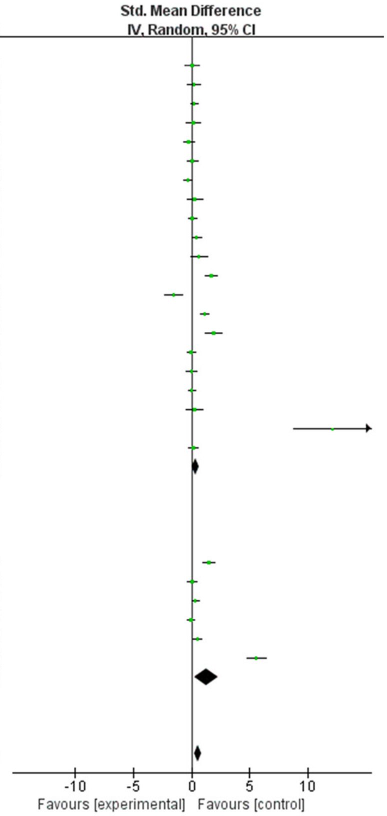

Figure S5 Subgroup analysis on study type of online *vs.* offline education on knowledge and skills acquisitions at the post-test levels.

https://dx.doi.org/10.21037/jxym-22-3