Journal of Contemporary Medical Education
2021 VOL 11, NO. 1, PAGE 6-26

ORIGINAL RESEARCH

Open Acess

CASE: BASED VIRTUAL PATIENT SIMULATION IMPACT on MEDICAL and GRADUATE MEDICAL EDUCATIONAL COMPETENCIES and MILESTONES: A META-ANALYSIS

Eric W Lam[1], Ivan C Lam[1], Miriam E Schwartz[1], Deborah Woo, Douglas E Paull[2*]
[1]Department of Medicine, University of California Los Angeles, California
[2]Department of Medicine, Georgetown University School of Medicine, Washington DC, USA

ABSTRACT

Objective: Virtual strategies have assumed an even greater role in medical education and graduate medical education during the recent pandemic. The objective of this study was to assess the impact of case-based, virtual patient simulation on medical education and graduate medical education.
Methods: A literature search of years 2000 through 2019 discovered 2,285 potential articles on virtual patient simulation. Fifty-four articles meeting the following criteria were included in the meta-analysis: involved medical education or graduate medical education participants; utilized case-based virtual patient simulations; and contained enough data to calculate an effect size (number of participants in each study arm, mean, and standard deviation of at least one measured learning outcome).
Results: Virtual patient simulation had a large overall impact on medical education and graduate medical education learning outcomes, 0.88 (0.64-1.12), z=7.36, p<0.001. Effect sizes by competencies were patient care, 0.95 (0.62-1.28); medical knowledge, 0.69 (-0.06-1.44); interpersonal and communication skills, 0.52 (0.01-1.05); professionalism, 1.32 (0.29-2.35); and systems-based practice, 0.71 (0.21-1.20). There was, however, a high level of heterogeneity between studies (I2=92.6%) lessening the certainty of the effect size summary.
Conclusions: The current study reinforces the results of previous meta-analyses demonstrating the moderate to large effect of virtual patient simulation interventions upon learning outcomes. Additionally, it highlights the effectiveness of virtual patient simulation for medical education and graduate medical education competencies beyond patient care, medical knowledge, and communication to include systems-based practice and professionalism. Virtual patient simulation is well-suited to address current challenges facing medical education and graduate medical education.

ARTICLE HISTORY
Received: December 15, 2021
Accepted: January 8, 2021
Published: January 19, 2021

KEYWORDS
Learning outcomes; Medical education; Graduate medical education; Virtual patient simulation

Introduction

Health care education is facing changes in recent times. With rapidly advancing technologies, increasing complexities of patient care, and numerous challenges to health care systems, a variety of education methods are needed to address the evolving needs of medical students, graduate medical trainees, and clinicians [1]. Simulation-based education, especially Virtual Patient Simulation (VPS), can help address emerging challenges and gaps by reaching a greater number of learners, supplementing the number and types of patients that learners experience (including rare clinical cases), addressing schedules and the lack of time for both learners and clinical educators, avoiding the medico-legal and ethical implications of skill acquisition from actual patients, and promoting patient safety concepts [2-4].

The current COVID-19 pandemic presents additional challenges to health care processes, team training, and education [5,6]. Given the limited number of persons allowed in gatherings, postponed or canceled surgeries, and limited rotations between training sites, Chick et al. proposed innovative

Coressponding Author: Douglas E Paull dpaull78@comcast.net Department of Medicine, Georgetown University School of Medicine, Washington DC, USA Tel: 734-478-7858

methods of learning such as technology-enhanced modalities, including “the flipped classroom model, online practice questions, teleconferencing in place of in-person lectures, involving residents in telemedicine clinics, procedural simulation, and the facilitated use of surgical videos [5].” Kiely et al. recommended strategies for obstetrics team training and simulation-based education while maintaining social distancing; these recommended modalities include spatial, temporal, video-recording, video-conferencing, and virtual learning opportunities to effectively engage obstetrics team members during the COVID-19 pandemic [6]. Virtual strategies, though not a substitute for hands-on direct patient care and procedures, are an example of innovative modalities of healthcare education that can help address the need for rigorous educational learning during this time of the COVID-19 pandemic [5,6].

“Digital education (also known as electronic learning or digital learning) is the act of teaching and learning by means of digital technologies. It is an overarching term for an evolving multitude of educational approaches, concepts, methods, and technologies [7].” Examples of digital learning include mobile education, serious games, virtual patients, and virtual reality environments [2,7]. VPS or screen-based simulation is an example of digital education. “Virtual patient” simulation has been defined as “a specific type of computer program that simulates real-life clinical scenarios; learners emulate the roles of health care providers to obtain history, conduct a physical exam, and make diagnostic and therapeutic decisions [1,8].” A number of narrative reviews, systematic reviews, and meta-analyses regarding VPS have been published in the preceding years. Cook and Triola (2009) conducted a narrative literature review on VPS and proposed that VPS has the “unique and cost-effective function to facilitate and assess the development of clinical reasoning [1].” They also recommended more research in instructional design, curricular integration, and using VPS to enhance clinical reasoning [1]. In a subsequent systematic analysis of 48 studies, Cook et al. (2010) concluded that virtual patients were “consistently associated with higher learning outcomes” when compared against studies with no intervention(s) [9].

Consorti et al. (2012) compared VPS to more traditional learning methods, either as an alternative learning method or as an addition to the usual curriculum [10]. They reported that VPS was not only effective as an educational tool for clinical reasoning and in clinical data gathering and interpretation, but also useful in preparing students for further education in the domains of communication skills and ethical reasoning [10]. Kononowicz et al. and the Digital Health Education Collaboration Group (2019) performed a systematic review and meta-analysis of 51 randomized controlled trials (RCT) of VPS in health professions education [2]. The participants in the studies were from a variety of disciplines. Skills that were improved in the studies included clinical reasoning, procedural, and team skills.

Objectives

The primary objective of this meta-analysis was to assess the impact of case-based, VPS on medical education and graduate medical education (GME). A secondary objective was to analyze those VPS design factors that may contribute to positive learning outcomes. We believe that the results of such a meta-analysis will be a timely and helpful tool for medical educators in developing, implementing, and evaluating digital education experiences, especially during current world challenges. The framework for the analysis was core competencies and milestones for medical education and GME described by the Accreditation Council for Graduate Medical Education (ACGME) and the American Board of Medical Specialties (ABMS) [11-13].

Methods

Data sources and searches

The Georgetown University Institutional Review Board (IRB) determined that the study was not considered research involving human subjects and therefore IRB approval was not required. A search strategy was formulated to collect articles pertaining to the effectiveness of screen-based VPS in medical education and GME. The primary search strategy employed Boolean search with a combination of the following terms: healthcare screen based simulation, high fidelity medical simulation, high fidelity simulation, patient modeling, patient simulation, patient simulation training, screen based simulation, screen based virtual patient education, screen based virtual patient simulation, simulation medical education, virtual patient, virtual patient education, virtual patient medical education, virtual patient safety, virtual patient simulation, virtual patient simulation education, virtual patient simulator, virtual patient surgery, virtual patient training, virtual patients, and virtual simulation. The following databases were searched: Cochrane, Google Scholar, JSTOR, Microsoft Academic, OvidMD, PubMed, and Science Direct (Figure 1). The search period was restricted to years 2000 through 2019 to maintain modern relevancy. Search results were compiled in Microsoft Excel. After the removal of duplicates, a total of 1,521 studies were found using this initial approach (Figure 2).

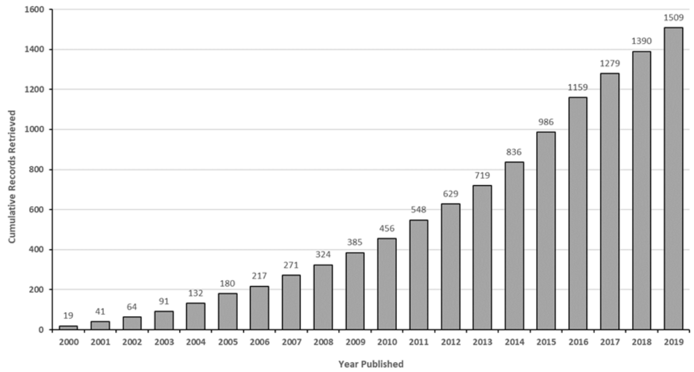


Figure 1. Literature search. Cumulative number of records retrieved from search of databases for years 2000-2019. Databases searched included Cochrane, Google Scholar, JSTOR, Microsoft Academic, OvidMD, PubMed, and Science Direct. Search terms consisted of: healthcare screen based simulation, high fidelity medical simulation, high fidelity simulation, patient modeling, patient simulation, patient simulation training, screen based simulation, screen based virtual patient education, screen based virtual patient simulation, simulation medical education, virtual patient, virtual patient education, virtual patient medical education, virtual patient safety, virtual patient simulation, virtual patient simulation education, virtual patient simulator, virtual patient surgery, virtual patient training, virtual patients, and virtual simulation.

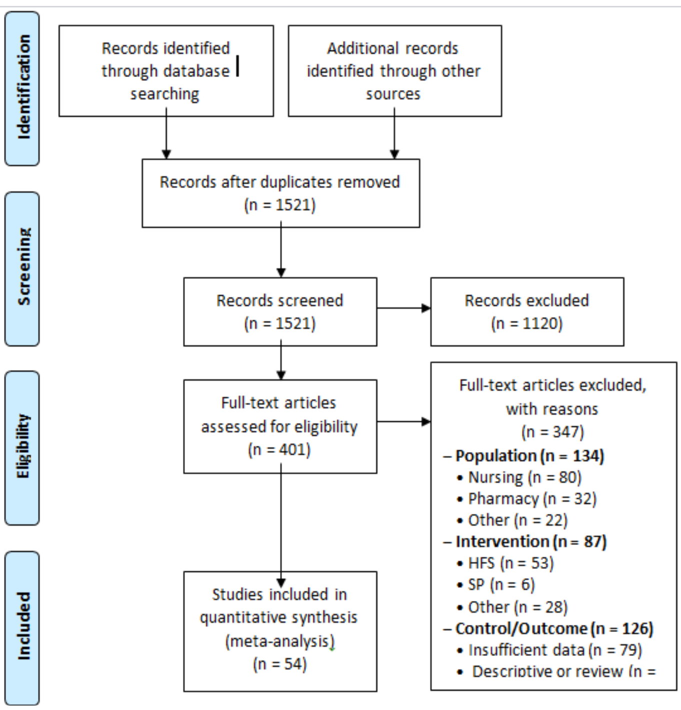


Figure 2. Flowchart. Flowchart of record search

strategy and selection of articles for full review leading to final 54 studies included in the meta-analysis based on inclusion criteria: 1) Population-include medical students and/or residents/fellow; 2) Intervention-case-based, screen-based VPS; 3) Comparison-RCT, pre-post, or experimental-control study design; and 4) Outcome-at least one learning outcome measured in two groups or two instances in time with sufficient data [i.e., mean (M), standard deviation (SD), number of participants (n)] to allow calculation of an effect size (ES). [Add note about Overlapping Publications.]

Inclusion criteria and study selection

Titles, abstracts, and text of articles were screened and reviewed by three reviewers (DEP, EWL, MES). Discussion and consensus were used to resolve disagreements. Screen-based VPS design was part of the key inclusion criteria. Articles forwarded for study inclusion shared the following criteria derived from our PICO statement: 1) Population - include medical students and/or residents/fellows; 2) Intervention - case-based, screen-based VPS; 3) Comparison - RCT, pre/post, or experimental-control study design; and 4) Outcome - at least one learning outcome measured in two groups or two instances in time with sufficient data [i.e., mean (M), standard deviation (SD), number of participants (n)] to allow calculation of an effect size (ES). A total of 54 articles were ultimately selected for inclusion in the meta-analysis under these criteria and methods.

Data extraction and items

Full text of selected articles was collected, and data extraction occurred independently and in duplicate. Extracted data variables included study design, number of participants, learner level (medical students or residents/fellows), specialty (e.g., medicine, surgery), type of control group (e.g., didactic, standardized patient), VPS intervention features (e.g., feedback, practice), number of unique VPS scenarios, Kirkpatrick level (e.g., knowledge, performance) assessed, and data (i.e., M, SD, n) for each study and control group. The competencies assessed in each article were compared to the 18 milestones among five competencies outlined by ACGME including: patient care, medical knowledge, interpersonal and communication skills, professionalism, and system-based practice [14]. A sixth competency, practice-based learning and improvement was not included because all the articles and study interventions could be viewed as addressing this competency. The competency assigned to each article by three reviewers (DEP, EWL, MES) demonstrated moderate-good agreement, Fleiss's Kappa 0.52. Disagreements between reviewers were resolved by consensus.

Data synthesis

ESs (Hedge's g) were calculated using pooled SDs from comparison groups from each included article. A pre-determined algorithm was used to determine which ES was entered for those studies reporting more than one summary, evaluable learning outcome: 1) the highest Kirkpatrick level learning outcome; and 2) if more than one method of assessment was present for the highest level of learning outcome, the ESs were averaged following previously published guidelines [15]. A random effects model was utilized to develop a forest plot and ES summary with 95% confidence intervals for the included VPS articles [16,17]. To assess for and address any publication bias, a funnel plot was utilized for the synthesis and an Egger's regression applied to the dataset [18]. An adjusted random effects summary and confidence interval were then calculated and reported using an R0 trim and fill technique. Sub-group comparisons were pre-determined and related to the study questions regarding factors impacting learning outcomes for a VPS intervention [19]. Random effects models were used for subgroup analyses and statistical comparisons conducted with ANOVA.

Quality assessment

Heterogeneity between studies was evaluated by Q and I2 statistics [20,21]. The quality of studies was assessed using the 18-point Medical Education Research Study Quality Instrument (MERSQI) checklist tool [22]. There was substantial agreement between the two observers' (DEP, EWL) who scored the MERSQI, Cohen's Kappa of 0.56. Bias within individual studies was evaluated using the Cochrane handbook [23]. The risks for selection, attrition, detection, performance, and reporting biases were rated for each study as low (green), high (red), or unclear (yellow). Inter-rater reliability between the three reviewers (DEP, EWL, MES) rating bias for individual studies was substantial, Fleiss's Kappa=0.62. Bias across all studies was calculated as percentages of low, high, and unclear ratings for each type of bias using a stacked bar approach.

Results

Study selection

Fifty-four studies were included in the final analysis, flow diagram, (Figure 2) [24-77]. Each study involved a VPS intervention involving medical students and/or residents/fellows that was based on case scenario(s) and included sufficient pre/post or study group/control data to calculate an ES.

Study characteristics

Study characteristics are summarized in Table 1. Twenty-three studies (23/54, 43%) employed randomization; 35 studies (35/54, 65%) had a control group; and 33 studies

Study (Year)	Study Design	n	Learner Level	Special-ty	Control	VPS Fea-tures	No. of Scenar-ios	GME Mile-stone	MERSQI Score
Chon (2019)	PP	140	MS	SURG	NO	INT, FB, PRAC	4	MK-1	14
Sezer (2019)	R, PP, C	88	MS	FM	SP	INT, FB, PRAC	1	ICS-1	14.5
Fleiszer (2018)	C	90	MS	SURG	VPS	INT, FB	1	PC-3	14.5
Taekman (2017)	PP	48	OTHER	OB GYN	NO	INT, FB	1	SBP-1	7.5
Dankbaar (2017)	PP, C	103	MS	MED	DIDACT	INT, FB, PRAC	1	SBP-2	14.5
Tolsgaard (2016)	R, C	45	MS	MED	VPS	INT, PRAC	4	PC-1	15
McKendy (2016)	PP	29	RES	SURG	NO	FB, PRAC	18	MK-1	11
Sullivan (2016)	PP	98	MS	SURG	NO	INT, FB, PRAC	2	MK-1	13

Paull (2016)	C	212	RES	MED	HFS	INT, FB	1	SBP-2	11
Foster (2016)	R, C	70	MS	PSYCH	VPS	FB	3	PROF-1	16
Edelbring (2016)	C	190	MS	MED	VPS	INT, FB, PRAC	1	MK-1	12
Elledge (2016)	PP	29	MS	SURG	NO	INT, FB, PRAC	5	PC-2	10.5
Kleinert (a) (2015)	PP	25	MS	SURG	NO	INT, FB, PRAC	4	PC-1	12
Kleinert (b) (2015)	PP	62	MS	SURG	NO	INT, FB, PRAC	3	PC-2	9
Close (2015)	PP	71	RES	SURG	NO	INT, FB, PRAC	20	PC-2	12
Kleinsmith (2015)	C	73	MS	MED	SP	NO	4	PROF-1	12.5
Foster (2015)	R, PP, C	67	MS	PSYCH	OTHER	INT, FB	1	ICS-1	14
Johnson (2015)	PP	52	RES	PEDS	NO	NO	2	SBP-4	10
Woodham (2015)	C	119	MS	SURG	VPS	INT, FB	1	SBP-2	11
Pantziaras (a) (2015)	PP	32	RES	PSYCH	NO	NO	1	MK-1	12
Pantziaras (b) (2015)	PP	32	RES	PSYCH	NO	FB	1	PROF-3	9.5
Leung (2015)	PP, C	130	MS	ANESTH	VPS	INT, FB, PRAC	6	PC-2	13.5
Sperl-Hillen (2014)	R, PP, C	341	RES	FM	DIDACT	FB, PRAC	18	PC-2	16
Poulton (2014)	R, C	81	MS	MED	VPS	INT, FB	5	PC-2	12.5
Bediang (2013)	R, C	20	MS	FM	VPS	NO	2	PC-1	12
Harris (2013)	R, C	120	RES	FM	VPS	INT, FB, PRAC	5	PC-2	12
Funke (2013)	PP	116	MS	SURG	NO	INT, FB, PRAC	6	SBP-3	11
Tan (2013)	PP	137	MS	FM	NO	INT, PRAC	1	SBP-1	11.5
Courteille (2013)	R, C	82	RES	SURG	DIDACT	INT, FB	1	MK-1	11.5
Yang (2013)	PP, C	31	MS	SURG	OTHER	FB, PRAC	3	PC-3	11.5
Kononowicz (2012)	R, PP, C	226	MS	MED	DIDACT	INT, FB, PRAC	6	MK-2	13
Lin (2012)	C	66	MS	PSYCH	DIDACT	INT, FB	1	PC-1	10
Lunney (2012)	PP	91	RES	PEDS	NO	INT, PRAC	10	PROF-1	13
Persky (2011)	R, C	76	MS	FM	VPS	FB	2	PROF-3	13

Oliven (2011)	R, C	262	MS	FM	SP	FB, PRAC	5	PC-1	13
Botezatu (a) (2010)	R, C	49	MS	MED	DIDACT	INT, FB	6	PC-1	14
Botezatu (b) (2010)	R, C	216	MS	MED	DIDACT	INT, FB	6	PC-1	14
Andreatta (2010)	R, PP, C	15	RES	MED	SP	INT, FB, PRAC	1	SBP-4	15
Gucwa (2010)	PP	16	RES	FM	NO	NO	1	PROF-1	9.5
Kandasamy (2009)	R, PP, C	62	MS	SURG	OTHER	INT	5	PC-2	11
Deladisma (2009)	R, PP, C	29	MS	SURG	DIDACT	INT	1	PC-1	10
Vu-kanovic-Criley (2008)	PP, C	82	MS	MED	DIDACT	INT, FB, PRAC	1	PC-4	13
Youngblood (2008)	R, PP, C	30	OTHER	SURG	HFS	FB, PRAC	6	SBP-1	14.5
Boyd (2008)	PP	101	RES	OB GYN	NO	INT, FB	1	PROF-3	12.5
Deladisma (2007)	R, C	84	MS	FM	SP	NO	1	PROF-1	13.5
Sijistermans (2007)	PP	134	MS	PSYCH	NO	FB, PRAC	10	SBP-4	10
Vash (2007)	R, C	48	MS	SURG	DIDACT	INT, FB, PRAC	14	PC-1	12.5
Nendaz (2006)	C	6	MS	FM	SP	NO	3	MK-1	14.5
Stevens (2006)	C	20	MS	FM	VPS	NO	1	PROF-1	8
Ferguson (2006)	PP	30	RES	PEDS	NO	INT, FB, PRAC	1	PROF-1	12.5
Triola (2006)	R, PP, C	55	OTHER	MED	SP	INT, FB, PRAC	4	ICS-1	15
Dickerson (2006)	R, C	17	MS	MED	VPS	NO	1	PC-1	11
Bearman (a) (2001)	PP, C	212	MS	FM	VPS	INT, FB	2	ICS-1	14
Bearman (b) (2001)	R, C	167	MS	FM	VPS	INT, FB	2	ICS-1	10
communication skills; INT=Integrated into curriculum; MED=Medicine; MERSQI=Medical Education Research Study Quality Instrument; MK=Medical knowledge; MS=Medical students; n=Number of participants; NO=None; OB GYN=Obstetrics and Gynecology; OTHER=interprofessional participants that include residents or fellows; PC=Patient care; PEDS=Pediatrics; PP=Pre and post; PRAC=Practice; PROF=Professional; PSYCH=Psychiatry; R=Randomization; RES=Residents; SBP=Systems-based practice; SP=Standardized patient; SURG=Surgery; VPS=virtual patient simulation.									

Table 1. Study characteristics.

(33/54, 61%) included both pre and post data. The control group was another VPS group (13/35, 37%); didactic or traditional curriculum not involving simulation (10/35, 29%), standardized patient (7/35, 20%), high-fidelity simulation (2/35, 6%), and "other" (3/35, 8%). The studies included n=4,827 participants with M=89.4 participants per study; 37 studies (37/54, 68%) primarily involved medical students and 14 studies (14/54, 26%) residents/fellows. Three (3/54, 6%) studies included medical students or residents/ fellows combined or as part of an interprofessional team. Specialty designations included surgery (16/54, 29%), medicine (13/54, 24%), family medicine (13/54, 24%), psychiatry (6/54, 11%), pediatrics (3/54, 6%), obstetrics gynecology (2/54, 4%), and anesthesia (1/54, 2%).

The studies included n=217 virtual patient simulation scenarios with M=4.0 scenarios per study, (Table 1).

There were 90 evaluable learning outcomes in the 54 studies including learner satisfaction (22/90, 24%), knowledge (41/90, 46%), and performance (27/90, 30%), (Table 2). Eleven studies (11/54, 20%) included enough information to calculate an ES for more than one level of learning outcome, and 19 studies (19/54, 35%) recorded more than one evaluable metric within a single level of outcome. Among the 20 studies (20/54, 37%) assessing one or more measures of performance, assessment was by an embedded VPS scoring system (7/20, 35%); a standardized patient simulation (5/20, 25%); a combination of VPS and SP (5/20, 25%); or another method (3/20, 15%). In 26 studies (26/54, 48%) there was a reported time interval, M=6.2 weeks, between the intervention and the assessment. Fourteen studies (14/54, 26%) measured learning outcomes immediately after the VPS intervention. In 15 studies (15/54, 28%), the evaluation occurred after an unspecified period.

Study (Year)	**Kirk-patrick Level**	**Assess-ment**	**VPS Study Group**	**M**	**n**	**SD**	**Com-par-ison Group**	**M**	**n**	**SD**	**P-Val-ue**
Chon (2019)	K	TEST	VPS POST	76	140	11.6	PRE	60.4	140	16.6	<0.05
Sezer (2019)	K	TEST	VPS POST	80.6	44	9.88	PRE	57.1	44	11.6	<0.05
	K	TEST	VPS POST	80.6	44	9.88	SP POST	77.5	44	11.8	NS
Fleiszer (2018)	P	VPS PROC	VPS SRS POST	59	25	13	VPS JRS POST	51	57	12	<0.05
	P	VPS COMM	VPS SRS POST	31	25	15	VPS JRS POST	28	57	14	NS
Taekman (2017)	S	QUEST	VPS POST	8.95	48	1.42	PRE	7.83	48	1.55	<0.05
Dankbaar (2017)	K	TEST	VPS POST	57.9	34	6.5	DI-DACT POST	52.6	37	7.1	<0.05
	K	TEST	VPS GAME POST	60.1	32	6.7	DI-DACT POST	52.6	37	7.1	<0.05
Tolsgaard (2016)	K	TEST	VPS CONST POST	61.4	20	5.2	PRE	58.9	20	7	NS
	K	TEST	VPS CONST POST	61.4	20	5.2	VPS PR SLV POST	62.6	19	5.7	NS
	P	SP	VPS CONST POST	59.1	20	12.8	VPS PR SLV POST	60.8	19	11.5	NS

McKendy (2016)	K	TEST	VPS POST	55.4	26	6.6	PRE	59.6	29	8.1	<0.05
Sullivan (2016)	P	VPS DIVERT	VPS POST	67.9	76	29	PRE	35.1	98	34.8	<0.05
	P	VPS GI BLEED	VPS POST	82.1	51	19.8	PRE	41.8	78	40.9	<0.05
Paull (2016)	S	QUEST	VPS POST	4.6	108	0.8	HFS POST	4.6	104	0.7	NS
Foster (2016)	P	SP	VPS EMP FB POST	2.91	35	0.16	VPS NO FB POST	2.27	17	0.21	<0.05
Edelbring (2016)	S	QUEST	VPS STUD POST	4.18	58	0.53	VPS TEACH POST	3.6	27	0.51	<0.05
Elledge (2016)	S	VAS	VPS POST	45.7	29	16.6	PRE	29.2	29	19.2	<0.05
	K	TEST	VPS POST	13	29	3.56	PRE	10	29	2.52	<0.05
Kleinert (a) (2015)	K	TEST	VPS POST	8.88	25	0.9	PRE	7.24	25	0.9	<0.05
Kleinert (b) (2015)	K	TEST	VPS POST	7	62	1	PRE	5	62	1	<0.05
Close (2015)	P	VPS	VPS POST	68	71	40	PRE	22	71	40	<0.05
Kleinsmith (2015)	P	VPS + SP	VPS POST	18.7	73	12.7	SP POST	14.5	73	9.1	<0.05
Foster (2015)	P	SP	VPS POST	22.4	34	3.7	VIDEO POST	21.6	33	4.2	NS
Johnson (2016)	S	QUEST	VPS POST	3.72	219	0.76	PRE	3.64	221	0.84	NS
Woodham (2015)	S	QUEST	VPS VIDEO POST	3.82	116	0.9	VPS TEXT POST	3.87	119	0.75	NS
Pantzarias (a) (2015)	K	TEST IMMED	VPS POST	8.47	32	1.65	PRE	7.44	32	0.32	<0.05
	K	TEST 8 WKS	VPS POST	8.38	26	2.02	PRE	7.44	32	0.32	<0.05
Pantzarias (b) (2015)	S	QUEST	VPS POST	4.2	32	0.8	PRE	3.86	32	0.73	<0.05
Leung (2015)	S	QUEST	VPS BRANCH POST	5.19	130	2.1	VPS LINEAR POST	4.5	130	2.9	<0.05
	K	TEST MC	VPS BRANCH POST	85	32	11.5	PRE	66	32	8.6	<0.05
	K	TEST MC	VPS BRANCH POST	85	32	11.5	VPS LINEAR POST	69	32	18.7	<0.05
	K	TEST ESSAY	VPS BRANCH POST	64	32	17.2	PRE	43	32	12.9	<0.05

	K	TEST ESSAY	VPS BRANCH POST	64	32	17.2	VPS LINEAR POST	49	32	17.2	<0.05
	K	TEST-EOY	VPS BRANCH POST	54	32	12.9	PRE	48	32	10	<0.05
	K	TEST-EOY	VPS BRANCH POST	54	32	12.9	VPS LINEAR POST	51	32	12.9	NS
Sperl-Hillen (2014)	K	TEST	VPS POST	5.3	92	1.8	DI-DACT POST	4.1	128	1.6	<0.05
Poulton (2014)	K	TEST	VPS BRANCH POST	8.26	37	1.31	VPS LINEAR POST	6.94	43	1.62	<0.05
Bediang (2013)	P	SP ARF	VPS ARF POST	73.9	20	4.4	VPS POST CON-TROL	63.8	20	11.7	<0.05
	P	SP CSH	VPS CSH POST	66.1	20	6.2	VPS POST CON-TROL	60.2	20	7.9	<0.05
Harris (2013)	S	QUEST	VPS POST	211	32	28.1	DI-DACT POST	211	50	28.4	NS
	S	QUEST	VPS FAC POST	222	30	21.6	VPS RES POST	204	90	28.4	<0.05
Funke (2013)	P	VPS	VPS CASE 6	62.3	116	5.6	VPS CASE 2	53.9	116	5.6	<0.05
Tan (2013)	S	QUEST	VPS POST	3.14	127	0.76	PRE	2.17	127	0.81	<0.05
	K	TEST	VPS POST	10	127	2.39	PRE	7.69	127	2.27	<0.05
Courteille (2013)	K	TEST RES	VPS POST	10	20	1.1	DI-DACT POST	9.9	21	1.1	NS
	K	TEST MS	VPS POST	9	18	1.3	DI-DACT POST	9.4	23	1.4	NS
Yang (2013)	K	TEST NBME	VPS POST	86.5	33	7.4	PRE	83.5	36	9	NS
	P	VPS	VPS POST	70.8	31	25.9	PRE	56.7	27	34.8	NS
Kononowicz (2012)	K	TEST	VPS POST	48.3	47	3.2	PRE	36.9	47	3.4	<0.05
	K	TEST	VPS POST	48.3	47	3.2	DI-DACT POST	45.8	75	3.8	<0.05

Lin (2012)	P	CLIN EVAL	VPS POST	88.2	32	3.1	DI-DACT POST	85.2	34	3.9	<0.05
Lunney (2012)	S	QUEST	VPS POST	10.3	91	2.78	PRE	7.7	91	2.2	<0.05
	K	TEST	VPS POST	12	91	1.55	PRE	8.19	91	1.79	<0.05
Persky (2011)	S	QUEST	VPS OBESE POST	3.85	37	0.86	VPS NON-OB POST	2.6	39	0.72	<0.05
Oliven (2011)	P	VPS + SP	VPS POST	79.3	262	8.9	SP POST	82.3	262	7.9	<0.05
Botezatu (a) (2010)	K	TEST HEMAT	VPS POST	6.2	25	1.9	PRE	4.3	24	1.7	<0.05
	K	TEST CARDI-OL	VPS POST	7.9	25	1.2	PRE	6.1	24	1.7	<0.05
	P	VPS HEMAT	VPS POST	8	25	0.8	PRE	6.5	24	0.4	<0.05
	P	VPS CARDI-OL	VPS POST	8.8	25	0.9	PRE	7.6	24	0.5	<0.05
Botezatu (b) (2010)	K	TEST HEMAT	VPS POST	4.27	25	0.81	PRE	2.93	24	1.07	<0.05
	K	TEST CARDI-OL	VPS POST	4.6	25	0.69	PRE	3.7	24	0.9	<0.05
	P	VPS HEMAT	VPS POST	4.82	25	0.76	PRE	3.61	24	0.43	<0.05
	P	VPS CARDI-OL	VPS POST	5.18	25	0.59	PRE	4.11	24	0.41	<0.05
Andreatta (2010)	K	TEST	VR POST	16.7	7	3.04	PRE	17.1	7	3.63	NS
	K	TEST	VR POST	16.7	7	3.04	SP POST	18.5	8	2.62	NS
	P	VPS + SP	VR POST	3.55	7	1.7	SP POST	3.47	8	0.41	NS
Gucwa (2010)	S	QUEST	VPS POST	3.88	16	0.44	PRE	3.71	16	0.72	NS
Kandasamy (2009)	K	TEST	VPS POST	84.6	28	12.6	PRE	59.1	28	21.4	<0.05
	K	TEST	VPS POST	84.6	28	12.6	DI-DACT POST	74.3	27	15.3	<0.05
Deladisma (2009)	S	QUEST	VPS POST	4.27	15	0.47	DI-DACT POST	3.5	14	0.71	<0.05
Vukanovic-Criley (2008)	K	TEST	VPS POST	73.5	24	8.4	PRE	58.7	24	14	<0.05

	K	TEST	VPS POST	73.5	24	8.4	DI-DACT POST	59.5	42	15.4	<0.05
Youngblood (2008)	P	VPS + HFS	VPS POST	43.1	16	3.94	PRE	23.8	16	4.47	<0.05
	P	VPS + HFS	VPS POST	43.1	16	3.94	HFS POST	44.5	14	5.17	NS
Boyd (2008)	S	QUEST *	VPS POST	20.2	99	5.5	PRE	23.9	99	4.6	<0.05
	K	TEST	VPS POST	12.4	99	2.4	PRE	10.4	99	2.3	<0.05
Deladisma (2007)	P	VPS + SP	VPS POST	4.29	33	1.32	SP POST	3.24	51	1.06	<0.05
Sijistermans (2007)	S	QUEST	VPS POST	3.91	134	0.28	PRE	3.56	134	0.34	<0.05
Vash (2007)	K	TEST	VPS POST	18	23	2.9	DI-DACT POST	13	22	3	<0.05
Nendaz (2006)	P	VPS + SP	VPS POST	61	6	24	SP POST	72	6	23	NS
Stevens (2006)	S	QUEST	VPS VERS 2	7.4	13	0.3	VPS VERS 1	6.4	7	0.2	<0.05
Ferguson (2006)	S	QUEST *	VPS POST	37.5	29	9.17	PRE	51.9	29	10.8	<0.05
	K	TEST	VPS POST	12.4	30	1.58	PRE	9.47	30	1.55	<0.05
Triola (2006)	S	QUEST	VPS POST	2.97	23	0.6	PRE	2.14	23	0.64	<0.05
	S	QUEST	VPS POST	2.97	23	0.6	SP POST	3.19	32	0.65	NS
Dickerson (2006)	P	EX-PERT EVAL	VPS SYN VOICE	4.37	8	1.59	VPS REC VOICE	5	9	1.85	NS
Bearman (a) (2001)	P	SP	VPS NAR POST	38.8	41	4.8	PRE	36.3	55	5.5	<0.05
	P	SP	VPS NAR POST	38.8	41	4.8	VPS PR SLV POST	35.7	38	5.3	<0.05
Bearman (b) (2001)	S	QUEST	VPS NAR POST	33.7	85	5.2	VPS PR SLV POST	32.7	82	5.2	NS

Abbreviations:

ARF=Acute renal failure; BRANCH=Branching design; CARDIOL=Cardiology; CASE 2=Second case of a 6-case curriculum; CASE 6=Sixth case of a 6-case curriculum; CLIN EVAL=Clinical evaluation by instructor; COMM=Communication skills; CONST=Participants construct VPS scenario; CONTROL=Control group; CSH=Chronic subdural hematoma; DIDACT=Didactic or traditional curriculum; DIVERT=Diverticulitis case; EMP FB=Empathetic feedback; EXPERT EVAL=Expert Evaluation of videotapes; FAC=Faculty; GAME=Serious gaming; GI BLEED=Gastrointestinal bleeding case; HEMAT=Hematology; HFS=High-fidelity simulation; JRS=More junior participants; K=Knowledge; LIM=Limited time to complete virtual patient simulation; LINEAR=Linear design; M=Mean; n=Number of participants; NAR=Narrative design; NO FB=No feedback to participant; NON-OB=Non-obese virtual patient; NS=not statistically significant; OBESE=Obese virtual patient; P=Performance; POST=Post intervention whether VPS or control group; PR SLV=Problem-solving design;

PRE=Pre intervention; PROC=Procedural knowledge or skills; QUEST=Question or survey; RES=Resident; S=Satisfaction or self-reported impact; SD=Standard deviation; SP=Standardized patient; SRS=More senior participants; STUD=Student regulated curriculum; TEACH=Teacher regulated curriculum; TEST=Knowledge-based test/examination; TEST 8 WKS=Test 8 weeks after intervention; TEST-EOY=End of year test; TEST ESSAY=Open-ended questions, essay style test; TEST IMMED=Test immediately following intervention; TEST MC=Multiple choice test; TEST MS=Test for medical students; TEST NBME=National Board of Medical Examiners scores; TEXT=Text replacing video in VPS; UNLIM=Unlimited time to complete virtual patient simulation; VAS=Visual analog scale; VERS 1=First version of VPS curriculum; VERS 2=Second version of VPS curriculum after improvements; VIDEO=VPS using videos in place of text; VPS=Virtual patient simulation; VPS REC VOICE=Virtual patient with a recorded voice; VPS SYN VOICE=Virtual patient with a recorded voice; VR=Virtual reality.

Table 2. Results

GME core competencies addressed were: patient care (PC) (21/54, 39%), systems-based practice (SBP) (10/54, 18.5%), medical knowledge (MK) (8/54, 15%), professionalism (PROF) (10/54, 18.5%), and interpersonal and communication skills (ICS) (5/53, 9%), (Table 1). The most frequent milestones for PC, SBP, MK, PROF, and ICS, respectively, were: PC-1 - "Gathers and synthesizes essential and accurate information to define each patient's clinical problem(s)" (10/21, 48%) and PC-2 - "Develops and achieves comprehensive management plan for each patient" (8/21, 38%); SBP-1 - "Works effectively within an interprofessional team" (3/10, 30%), SBP-2 - "Recognizes system error and advocates for system improvement" (3/10, 30%), and SBP-4 - "Transitions patients effectively within and across health delivery systems" (3/10, 30%); MK-1 - "Clinical knowledge" (7/8, 88%); PROF-1 - "Has professional and respectful interactions with patients, caregivers and members of the interprofessional team" (7/10, 70%); and ICS-1 - "Communicates effectively with patients and caregivers" (5/5, 100%).

Study quality and risk of bias within studies

The MERSQI score (maximum score 18 points) for the 53 studies was 12.3 ± 2.0 with a median of 12.0. Possible bias within each of the 54 included studies is shown in Table 3. An aggregated assessment of bias across all included studies is shown in Figure 3. Challenging areas included selection bias (e.g., randomization) and performance bias (e.g., blinding of participants). On the other hand, the majority of included studies was rated as having a small risk of detection bias (e.g., blinding observers/raters) and attrition bias (e.g., high response rates). In 13% (35/270 evaluations) of instances, the level of possible bias could not be determined from the information included in the article.

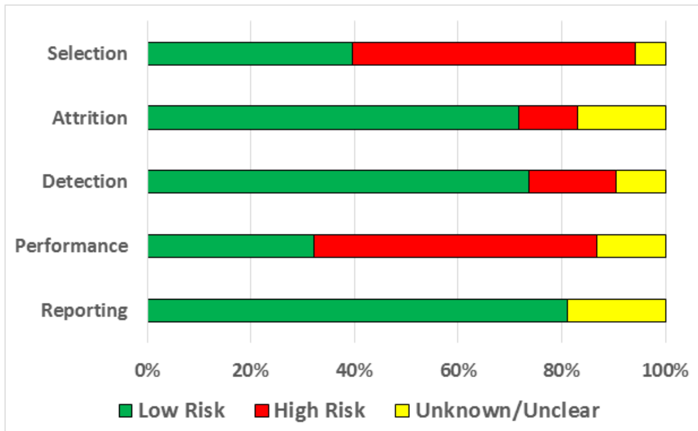


Figure 3. Possible bias across studies. Stacked bar graph showing the percentage of low, high, and unknown/unclear ratings for selection, attrition, detection, performance, and reporting bias across all 54 included studies. Bias ratings were determined by three independent raters, Kappa=0.62.

Results of individual studies

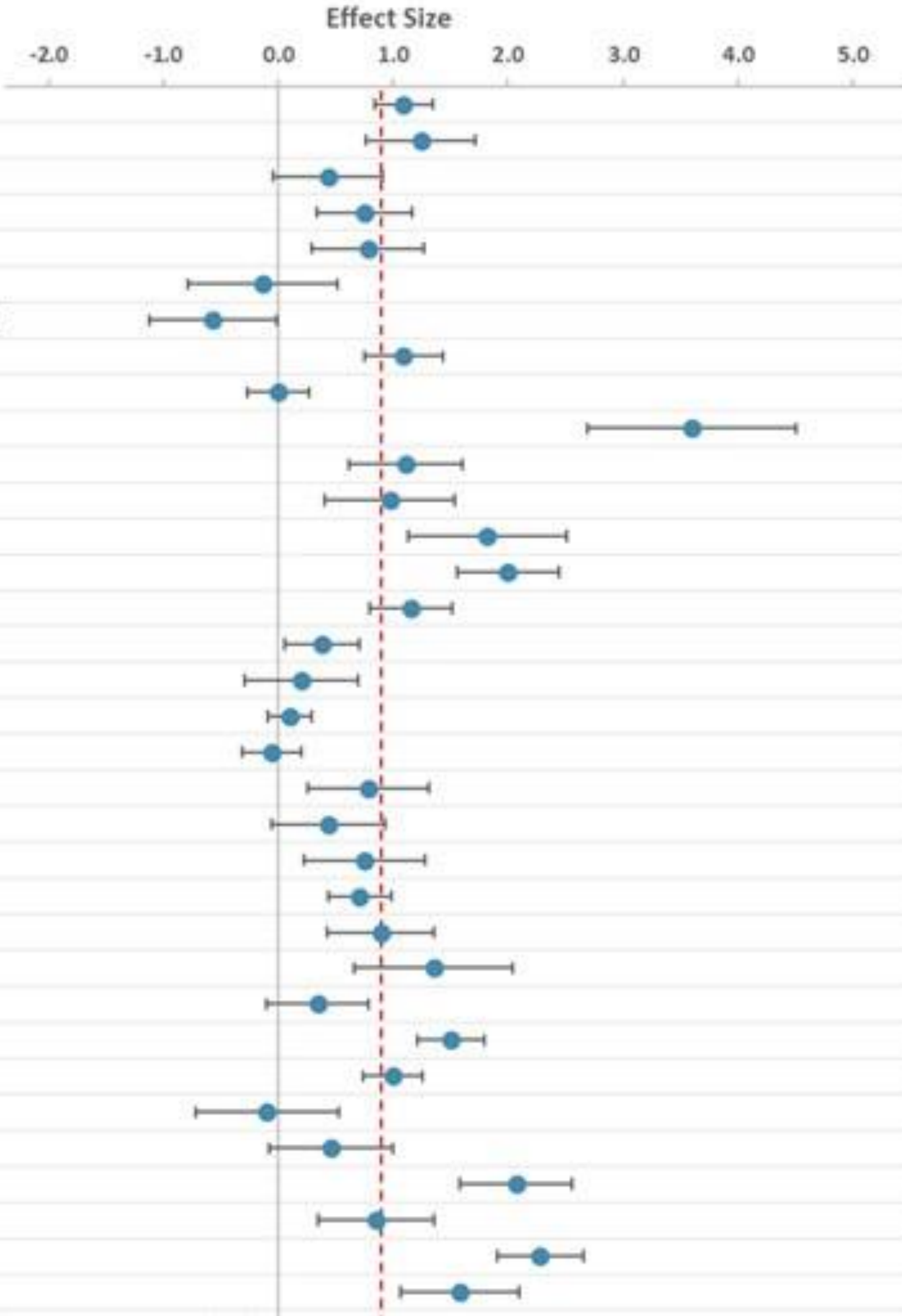
The results of individual study learning outcomes are listed in Table 2. Learner satisfaction showed improvement in 13 of 20 studies (12/19, 65%) including at least one measure of satisfaction, no difference in learner satisfaction between VPS and a baseline or control group in 5 studies (5/20, 25%). Two studies with more than one measure of satisfaction demonstrated mixed results (2/20, 10%). Nineteen of 25 studies (19/25, 76%) with at least one measure of knowledge showed improvement in test scores following a VPS intervention compared to baseline or a control group. Four studies (4/25, 16%) showed either no change or a decrease in learning outcomes associated with the VPS study group compared to baseline or a control group. Two studies with multiple, different measures of knowledge demonstrated mixed results (2/25, 8%). Performance improved in the VPS study group in 12 of 20 studies (60%) in which at least one measure of performance was assessed and was either the same or worse compared to a baseline or control group in 6 studies (6/20, 30%). In two studies (2/20, 10%) where more than one measurement of performance was recorded there were mixed results.

Among the 11 studies that measured more than one level of learning outcome, ten (10/11, 91%) demonstrated concordance between the results (e.g., both showed improvement or no improvement). One study had mixed results (1/11, 9%). For the 19 studies with more than one assessment for a single level of learning outcome 12 demonstrated concordance (12/19, 63%) and 7 (7/19, 37%) discordance in the direction of findings.

Synthesis of results

All VPS studies: The random effects model ES summary for all 54 studies was 0.88 (0.64-1.12), z=7.36, p<0.001, (Figure 4). There was substantial heterogeneity among the studies (I2=92.6%). Funnel plot analysis demonstrated slight asymmetry and the Egger regression was significant (t-test 2.14, p=0.04), suggesting possible publication bias, (Figure 5). After R0 trim and fill, an adjusted ES was similar, 0.85 (0.60-1.10).

Study (Year)	ES	CI [LL, UL]
Chon (2019)	1.09	[0.84, 1.34]
Sezer (2019)	1.24	[0.76, 1.72]
Fleiszer (2018)	0.43	[-0.05, 0.91]
Taekman (2017)	0.75	[0.33, 1.17]
Dankbaar (2017)	0.78	[0.29, 1.27]
Tolsgaard (2016)	-0.14	[-0.79, 0.51]
McKendy (2016)	-0.57	[-1.13, -0.01]
Sullivan (2016)	1.09	[0.75, 1.43]
Paull (2016)	0.00	[-0.27, 0.27]
Foster (2016)	3.60	[2.69, 4.51]
Edelbring (2016)	1.11	[0.62, 1.60]
Elledge (2016)	0.97	[0.40, 1.54]
Kleinert (a) (2015)	1.82	[1.13, 2.51]
Kleinert (b) (2015)	2.00	[1.56, 2.44]
Close (2015)	1.15	[0.79, 1.51]
Kleinsmith (2015)	0.38	[0.05, 0.71]
Foster (2015)	0.20	[-0.29, 0.69]
Johnson (2015)	0.10	[-0.09, 0.29]
Woodham (2015)	-0.06	[-0.32, 0.20]
Pantziaras (a) (2015)	0.78	[0.25, 1.31]
Pantziaras (b) (2015)	0.44	[-0.06, 0.93]
Leung (2015)	0.75	[0.22, 1.28]
Sperl-Hillen (2014)	0.71	[0.43, 0.99]
Poulton (2014)	0.89	[0.42, 1.36]
Bediang (2013)	1.35	[0.66, 2.04]
Harris (2013)	0.34	[-0.10, 0.78]
Funke (2013)	1.50	[1.21, 1.79]
Tan (2013)	1.00	[0.74, 1.26]
Courteille (2013)	-0.10	[-0.72, 0.52]
Yang (2013)	0.46	[-0.08, 1.00]
Kononowicz (2012)	2.07	[1.58, 2.56]
Lin (2012)	0.85	[0.34, 1.36]
Lunney (2012)	2.28	[1.90, 2.66]
Persky (2011)	1.58	[1.06, 2.10]



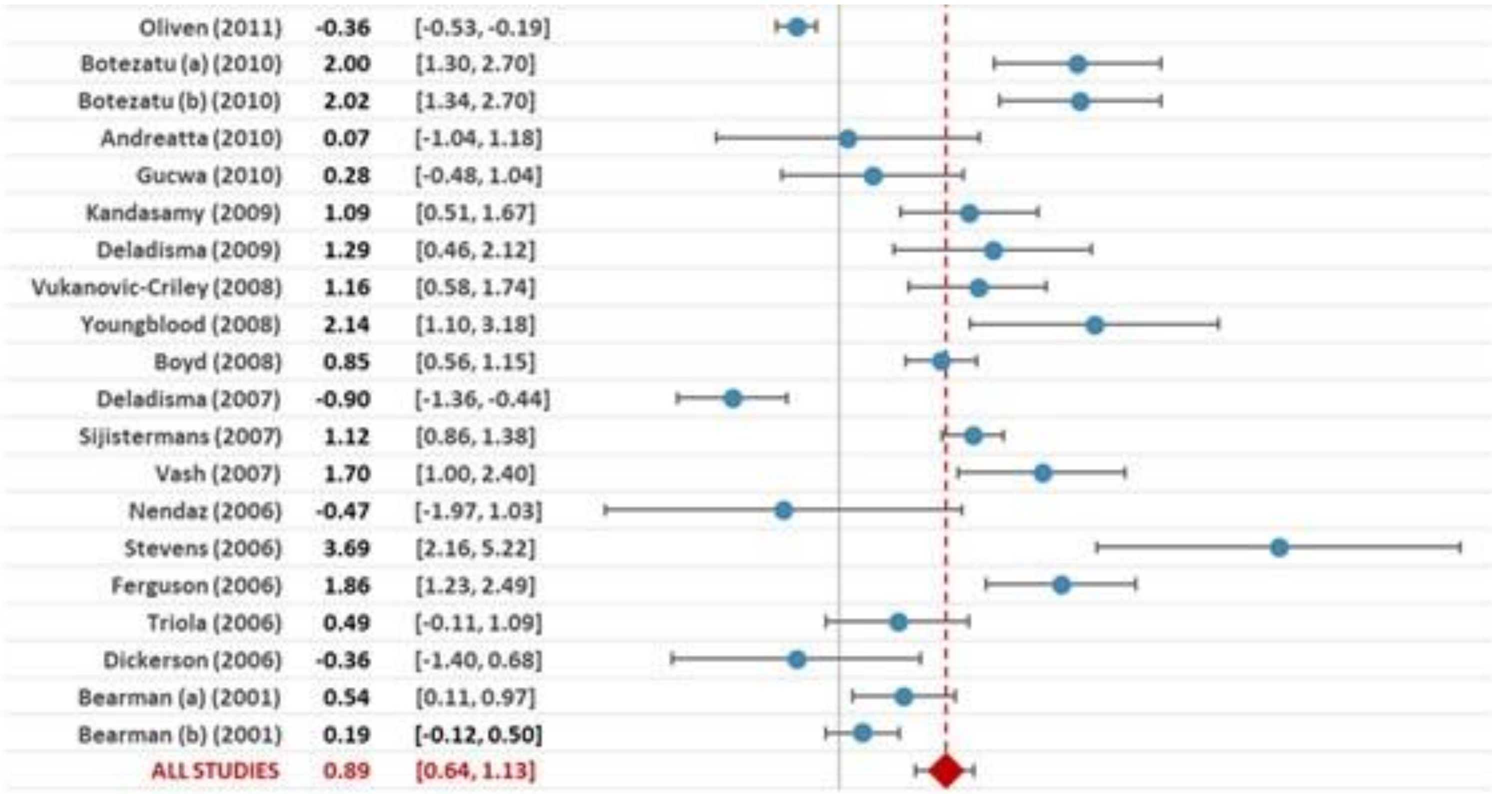


Figure 4. Forest plot. All studies. Forest plot of effect sizes (ES) and 95% confidence interval upper limit (UL) and lower limit (LL) for each of 54 included studies of the impact of VPS upon learning outcomes. Effect size calculated from the highest level of reported learning outcome assessed (satisfaction < knowledge < performance). Random effects model utilized to calculate an effect size summary for all the studies, 0.88 (0.64-1.12), z=7.36, p<0.001. There was substantial heterogeneity among the studies (I2=92.6%).

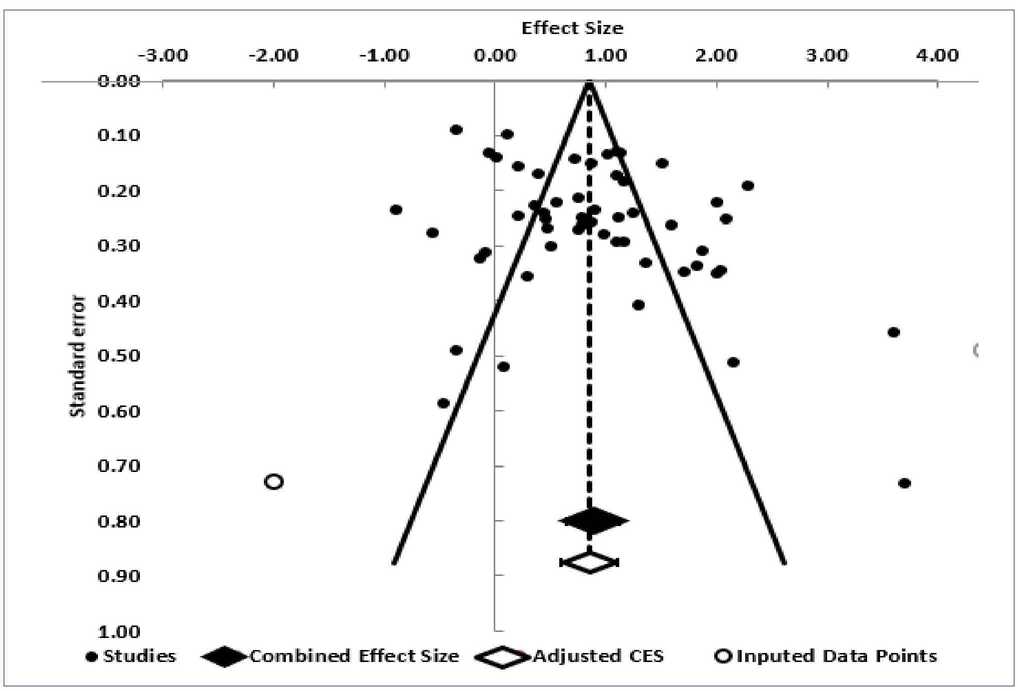


Figure 5. Funnel plot. Funnel plot analysis demonstrated slight asymmetry and the Egger regression was significant (t-test 2.14, p=0.04), suggesting possible publication bias. After R0 trim and fill, an adjusted ES was similar to original random effects model ES summary, 0.85 (0.60-1.10).

Competencies: A random effects subgroup analysis was performed by the competency addressed in each included VPS study. ES for each competency included: PC, 0.95 (0.62-1.28), I2=92.2%; MK, 0.69 (-0.06-1.44),I2=90.1%; ICS, 0.52 (-0.01-1.05), I2=73.1%; PROF, 1.32 (0.29-2.35), I2=95.3%; and SBP, 0.71 (0.21-1.20), I2=93.7%. The differences between subgroups was not statistically significant (Chi2=5.12, df=4, p=0.27). The subgroup of studies addressing "technical" competencies (PC, MK) was compared to the subgroup of studies concentrating on "non-technical competencies" (ICS, PROF, SBP). The technical competency subgroup had an ES of 0.88 (0.59-1.17), which was similar (Chi2=0.01, df=1, p= 0.91) to the non-technical subgroup ES of 0.90 (0.47-1.34).

Additional analyses

Learning outcome: The ES summary for studies, based on the highest evaluable learning outcome for each study, included: satisfaction, 0.74 (0.29 - 1.19), I2=90.0%; knowledge, 1.08 (0.75-1.42), I2=87.6%; and performance, 0.77 (0.26-1.27), I2=93.9% (Chi2 3.01, df=2, p=0.22).

Study design: Subgroup analysis was performed between a subgroup of studies that compared the VPS intervention to either baseline or a traditional curriculum (e.g., lecture) and a subgroup of studies where the VPS intervention was compared to a simulation control (e.g., standardized patient, high fidelity simulation, or different VPS design). There was a significant difference (Chi2=5.49, df=1, p=0.02) in ES between VPS with traditional control subgroup, 1.05 (0.81-1.30), I2=89.4%, and VPS with simulation control subgroup, 0.58 (0.10-1.06), I2=91.7%. RCT studies were compared to all other study designs. ESs were 0.89 (0.44-1.35), I2=93.5%, for RCTs and 0.87 (0.60-1.15), I2=91.4% for all other study designs (Chi2=0.01, df=1, p=0.93).

VPS design: The subgroup of studies where the VPS intervention included all three design characteristics (VPS integrated into a curriculum, feedback to the learner, and the opportunity for repetitive practice) was compared to a subgroup of studies where the VPS intervention included fewer characteristics. ESs were 1.20 (0.93 - 1.46), I2=76.6%, and 0.72 (0.39 - 1.05), I2=92.9%, respectively (Chi2=7.60, df=1, p=0.006). VPS studies with three or more different scenarios were compared to those studies with fewer than three scenarios. ESs were 1.10 (0.72 - 1.49), I2=94.6%, for three or more scenarios and 0.66 (0.38 - 0.95), I2=87.8%, for fewer than three scenarios, (Chi2=4.24, df=1, p=0.04).

Discussion

Summary of evidence

There was a moderate to large [ES=0.88 (0.64-1.12)] overall impact on learning outcomes for case-based VPS interventions in the current meta-analysis when compared to traditional lecture or no intervention. The current study results are consistent in both the positive direction and magnitude of the effect of simulation in general, and VPS specifically, upon learning outcomes [2,78]. In a meta-analysis of 609 studies of technology-enhanced simulation training for healthcare professionals, Cook et al. (2011) reported an ES for performance skills of 1.09 (1.03-1.16); however, that study excluded VPS interventions which only required standard computer equipment [78]. In a recent meta-analysis of 51 RCTs from 1990 to 2018 comparing learning outcomes for VPS and traditional learning formats, Kononowicz (2019) reported an ES of 0.90 (0.49-1.32) favoring VPS [2]. Our subgroup analysis of only RCTs demonstrated an ES of 0.89 (0.44-1.35), similar to the Kononowicz study [2]. As was the case in our study (I2=92.6%), the Cook (2011) and Kononowicz (2019) meta-analyses both reported high levels of heterogeneity between studies, I2 >50% and I2=88%, respectively [2,78]. Similar to the observations of Kononowicz et al., when VPS was compared to other types of simulation such as standardized patients, the ES was significantly smaller in magnitude but in the same positive direction, 0.58 (0.10-1.06) [2]. The cumulative evidence of the current and previous meta-analyses supports the reported increasing trends of medical educators integrating VPS into traditional curricula [79].

The current meta-analysis found a positive impact of VPS across a broad range of medical and GME competencies. Specifically, there was a positive, significant, moderate to large effect on PC, SBP, and PROF. Among the most common milestones addressed in these three competencies were: PC-1 - "Gathers and synthesizes essential and accurate information to define each patient's clinical problem(s)", SBP-1 - "Works effectively within an interprofessional team", and PROF-1 - "Has professional and respectful interactions with patients, caregivers and members of the interprofessional team". Previous meta-analyses have reported that case-based VPS interventions most often target, and have their largest impact on, clinical reasoning skills in PC and teamwork/communication competencies [2,80]. The ES for ICS in our study was 0.52 (-0.01-1.05). The lower ES noted could be due in part to: a smaller proportion of ICS articles in the meta-analysis; that most included ICS studies had another form of simulation as a control group; and the overlap between competencies for SBP and PROF and that for ICS. Nonetheless, when SBP, PROF, and ICS were combined as non-technical skills the ES was 0.90 (0.47-1.34), nearly identical to that for technical skills (PC and

MK). This finding is important for medical education and patient safety leaders addressing teamwork and communication challenges as common contributing factors to patient safety events in healthcare [81]. Increasingly, professionalism and respectful behaviors are recognized as critical competencies affecting patient safety [82]. The current meta-analysis demonstrated a large, significant effect of VPS on professionalism, 1.32 (0.29-2.35).

The second objective in the current meta-analysis was to determine VPS design factors that may be associated with improved learning outcomes. The need for research on VPS design and learning outcomes, as opposed to just comparing VPS to other types of education formats (e.g., lecture), has been emphasized by authors of previous meta-analyses [2,78]. The current study found that VPS interventions that include all three design features (integration, feedback, repetitive practice) had a significantly greater effect upon learning outcomes than those that lack one or more of these features. Improved learning outcomes have been similarly reported in other reviews [83-85]. In perhaps the largest meta-analysis of technology-enhanced simulation in healthcare, instructional design features were not found to be associated with better learning outcomes overall [78]. However, in that study, subgroup analysis demonstrated that integration of simulation into the curriculum was associated with improved knowledge level outcomes [78]. Clinical case diversity has also been reported as a driver of improved learning outcomes for VPS educational interventions. In the current meta-analysis, VPS interventions with three or more scenarios had a significantly greater positive impact on learning outcomes than those interventions with fewer than three scenarios. Even if certain of these design features are not consistently associated with improved quantitative learning outcomes, qualitative studies have demonstrated favorable medical student and resident perceptions of VPS interventions that include these features [31,48]. VPS design features associated with a qualitative positive impact on learning outcomes as determined in focus groups or interviews of medical students, residents, or fellows reported in several included studies were branching as opposed to linear, and narrative over problem-solving designs [45,47,75,76]. Medical educators may wish to consider design factors when making decisions about developing or adopting VPS interventions for their particular clinical learning environment.

Limitations

There was high, unexplained heterogeneity between the included studies within the meta-analysis, with all random effect summary estimates and subgroup analyses demonstrating $I2 > 75\%$. High heterogeneity has accompanied previous meta-analyses on all simulation and VPS ($I2 = 50-83\%$) [2,78]. The inconsistency of individual study results within the meta-analyses makes the validity of the ES summary estimates less certain [19,21]. Bias, both within the individual included studies as well as the meta-analysis itself, is another limitation [23,86]. Though funnel plots and quantitative measures of publication bias have their own limitations, the minimal funnel plot asymmetry and marginally positive Egger's test demonstrated in this meta-analysis would be supportive of a lesser degree of publication bias [18]. Attrition bias also appeared low as the majority of included articles in the meta-analyses had pre to post participant retention rates of >75%. Detection bias also did not appear to be a major contributor to the overall risk of bias in this meta-analysis; most of the highest level learning outcomes in individual studies were determined by objective methods such as multiple choice tests for knowledge or embedded scoring tools within simulation software. Selection bias within the individual included studies (only 43% of studies randomized participants) and reporting (eligible studies may have been missed in the meta-analysis) appear to be the most likely sources of overall bias [87-90].

Conclusions

The current study reinforces the results of previous meta-analyses demonstrating the moderate to large effect of VPS interventions upon learning outcomes. Additionally, it highlights the effectiveness of VPS for medical education and GME competencies beyond PC, MK, and ICS to include SBP and PROF. Medical education and GME have been significantly impacted by current world events. International simulation organizations have called for VPS experiences to address the increased need for online educational opportunities. In addition, there has been an accelerated growth of telehealth, a trend in healthcare that will likely persist into the future. Authors have begun to apply medical education and GME competencies to telehealth. VPS interventions are well-suited components of a curriculum to address telehealth skills. Modern repositories will allow sharing and wide distribution of validated VPS interventions organized and searchable by competency metadata.

References

[1] Cook DA, Triola MM. Virtual patients: A critical literature review and proposed next steps. Med Educ 2009;43(4):303-311.

[2] Kononowicz AA, Woodham LA, Edelbring S. Virtual patient simulations in health professions education: Systematic review and meta-analysis by the digital health education collaboration. J Med Internet Res 2019;21(7):e14676.

[3] Lateef F, Yin S, Suppiah M. The 2019 WACEM expert document on the framework for setting up a simulation centre. J Emerg Trauma Shock 2019;12(4):232-242.

[4] Lateef F. Simulation-based learning: Just like the real thing. J Emerg Trauma Shock 2010;3(4):348-352.

[5] Chick RC, Clifton GT, Peace KM. Using technology to maintain the education of residents during the COVID-19 pandemic. J Surg Educ 2020;77(4):729-732.

[6] Kiely DJ, Posner GD, Sansregret A. Health care team training and simulation-based education in obstetrics during the COVID-19 pandemic. J Obstet Gynaecol Can 2020;42(8):1017-1020.

[7] Car J, Carlstedt-Duke J, Tudor Car L. Digital education in health professions: The need for overarching evidence synthesis. J Med Internet Res 2019;21(2):e12913.

[8] Association of American Medical Colleges (AAMC). Effective use of educational technology in medical education: Summary report of the 2006 AAMC colloquium on educational technology. Washington, DC: AAMC 2007:7.

[9] Cook DA, Erwin PJ, Triola MM. Computerized virtual patients in health professions education: A systematic review and meta-analysis. Acad Med 2010;85(10):1589-1602.

[10] Consorti F, Mancuso R, Nocioni M, Piccolo A. Efficacy of virtual patients in medical education: A meta-analysis of randomized studies. Comput Educ 2012; 59(3):1001-1008.

[11] Eno C, Correa R, Stewart NH. Accreditation council for graduate medical education (ACGME): milestones guidebook for residents and fellows. 2020.

[12] Accreditation council for graduate medical education (ACGME). Frequently asked questions: milestones 2015.

[13] American board of medical specialties (ABMS). Board certification requirements 2020.

[14] Accreditation council for graduate medical education (ACGME) and American board of internal medicine (ABIM). The internal medicine milestone project 2015.

[15] López-López JA, Page MJ, Lipsey MW, Higgins JPT. Dealing with effect size multiplicity in systematic reviews and meta-analyses published online ahead of print, 2018.

[16] Neyeloff JL, Fuchs SC, Moreira LB. Meta-analyses and forest plots using a Microsoft Excel spreadsheet: step-by-step guide focusing on descriptive data analysis. BMC Res Notes 2012;5:52.

[17] Suurmond R, Van Rhee H, Hak T. Introduction, comparison, and validation of meta-essentials: A free and simple tool for meta-analysis. Res Synth Methods 2017;8(4):537-553.

[18] Sterne JA, Sutton AJ, Ioannidis JP. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 2011;343:d4002.

[19] Richardson M, Garner P, Donegan S. Interpretation of subgroup analyses in systematic reviews: A tutorial. Clin Epidemiol Glob Health 2019:7(2):192- 198.

[20] Borenstein M, Hedges LV, Higgins JP, Rothstein HR. Introduction to meta-analysis. Hoboken, NJ: John Wiley & Sons, Ltd 2009.

[21] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-560.

[22] Reed DA, Cook DA, Beckman TJ, Levine RB, Kern DE, Wright SM. Association between funding and quality of published medical education research. JAMA 2007;298(9):1002-1009.

[23] Higgins JPT, Savović J, Page MJ, Elbers RG, Sterne JAC. Chapter 8: Assessing risk of bias in a randomized trial. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane handbook for systematic reviews of interventions version 6.0 (updated July 2019). Cochrane, 2019.

[24] Chon SH, Timmermann F, Dratsch T. Serious

games in surgical medical education: A virtual emergency department as a tool for teaching clinical reasoning to medical students. JMIR Serious Games 2019;7(1):e13028.

[25] Sezer B, Sezer TA. Teaching communication skills with technology: Creating a virtual patient for medical students. Australas J Educ Technol 2019;35:183-198.

[26] Fleiszer D, Hoover ML, Posel N, Razek T, Bergman S. Development and validation of a tool to evaluate the evolution of clinical reasoning in trauma using virtual patients. J Surg Educ 2018;75(3):779-786.

[27] Taekman JM, Foureman MF, Bulamba F. A novel multiplayer screen-based simulation experience for African learners improved confidence in management of postpartum hemorrhage. Front Public Health 2017;5:248.

[28] Dankbaar ME, Richters O, Kalkman CJ. Comparative effectiveness of a serious game and an e-module to support patient safety knowledge and awareness. BMC Med Educ 2017;17(1):30.

[29] Tolsgaard MG, Jepsen RM, Rasmussen MB. The effect of constructing versus solving virtual patient cases on transfer of learning: a randomized trial. Perspect Med Educ 2016;5(1):33-38.

[30] McKendy KM, Posel N, Fleiszer DM, Vassiliou MC. A learner-created virtual patient curriculum for surgical residents: Successes and failures. J Surg Educ 2016;73(4):559-566.

[31] Sullivan SA, Bingman E, O'Rourke A, Pugh CM. Piloting virtual surgical patient cases with 3rd-year medical students during the surgery rotation. Am J Surg 2016;211(4):689-696.e1.

[32] Paull DE, Williams L, Sine DM. Simulation techniques for teaching time-outs: A controlled trial. Patient Saf Qual Healthc 2016:28-37.

[33] Foster A, Chaudhary N, Kim T. Using virtual patients to teach empathy: A randomized controlled study to enhance medical students' empathic communication. Simul Healthc 2016;11(3):181-189.

[34] Edelbring S, Wahlström R. Dynamics of study strategies and teacher regulation in virtual patient learning activities: A cross sectional survey. BMC Med Educ 2016;16:122.

[35] Johnson DP, Zimmerman K, Staples B, McGann KA, Frush K, Turner DA. Multicenter development, implementation, and patient safety impacts of a simulation-based module to teach handovers to pediatric residents. Hosp Pediatr 2015;5(3):154-159.

[36] Woodham LA, Ellaway RH, Round J, Vaughan S, Poulton T, Zary N. Medical student and tutor perceptions of video versus text in an interactive online virtual patient for problem-based learning: A pilot study. J Med Internet Res 2015;17(6):e151.

[37] Kleinert R, Heiermann N, Wahba R, Chang DH, Hölscher AH, Stippel DL. Design, realization, and first validation of an immersive web-based virtual patient simulator for training clinical decisions in surgery. J Surg Educ 2015;72(6):1131-1138.

[38] Kleinert R, Heiermann N, Plum PS. Web-based immersive virtual patient simulators: Positive effect on clinical reasoning in medical education. J Med Internet Res 2015;17(11):e263.

[39] Close A, Goldberg A, Helenowski I, Schuller M, DaRosa D, Fryer J. Beta test of web-based virtual patient decision-making exercises for residents demonstrates discriminant validity and learning. J Surg Educ 2015;72(6):e130-e136.

[40] Pantziaras I, Fors U, Ekblad S. Training with virtual patients in transcultural psychiatry: Do the learners actually learn? J Med Internet Res 2015;17(2):e46.

[41] Pantziaras I, Fors U, Ekblad S. Innovative training with virtual patients in transcultural psychiatry: The impact on resident psychiatrists' confidence. PLoS One 2015;10(3):e0119754.

[42] Kleinsmith A, Rivera-Gutierrez D, Finney G, Cendan J, Lok B. Understanding empathy training with virtual patients. Comput Human Behav 2015;52:151-158.

[43] Elledge R, McAleer S, Thakar M, Begum F, Singhota S, Grew N. Use of a virtual learning environment for training in maxillofacial emergencies: Impact on the knowledge and attitudes of staff in accident and emergency departments. Br J Oral Maxillofac Surg 2016;54(2):166-169.

[44] Foster A, Chaudhary N, Murphy J, Lok B, Waller J, Buckley PF. The use of simulation to teach suicide risk assessment to health profession trainees-rationale, methodology, and a proof of concept demonstration with a virtual patient. Acad Psychiatry 2015;39(6):620-629.

[45] Leung JY, Critchley LA, Yung AL, Kumta SM. Evidence of virtual patients as a facilitative learning tool on an anesthesia course. Adv Health Sci Educ Theory Pract 2015;20(4):885-901.

[46] Sperl-Hillen J, O'Connor PJ, Ekstrom HL. Educating resident physicians using virtual case- based simulation improves diabetes management: A randomized controlled trial. Acad Med 2014; 89(12):1664-73.

[47] Poulton T, Ellaway RH, Round J, Jivram T, Kavia S, Hilton S. Exploring the efficacy of replacing linear paper-based patient cases in problem-based learning with dynamic Web-based virtual patients: randomized controlled trial. J Med Internet Res 2014; 16(11):e240.

[48] Funke K, Bonrath E, Mardin WA. Blended learning in surgery using the Inmedea Simulator. Langenbecks Arch Surg 2013; 398(2):335-40.

[49] Bediang G, Franck C, Raetzo MA. Developing clinical skills using a virtual patient simulator in a resource-limited setting. Stud Health Technol Inform 2013; 192:102-6.

[50] Courteille O, Ho J, Fahlstedt M. Face validity of VIS-Ed: a visualization program for teaching medical students and residents the biomechanics of cervical spine trauma. Stud Health Technol Inform 2013; 184:96-102.

[51] Tan A, Ross SP, Duerksen K. Death is not always a failure: outcomes from implementing an online virtual patient clinical case in palliative care for family medicine clerkship. Med Educ Online 2013; 18:22711.

[52] Harris JM Jr, Sun H. A randomized trial of two e-learning strategies for teaching substance abuse management skills to physicians. Acad Med 2013; 88(9):1357-62.

[53] Yang RL, Hashimoto DA, Predina JD. The virtual-patient pilot: Testing a new tool for undergraduate surgical education and assessment. J Surg Educ 2013; 70(3):394-401.

[54] Lin CC, Wu WC, Liaw HT, Liu CC. Effectiveness of a virtual patient program in a psychiatry clerkship. Med Educ 2012; 46(11):1111-12.

[55] Kononowicz AA, Krawczyk P, Cebula G. Effects of introducing a voluntary virtual patient module to a basic life support with an automated external defibrillator course: A randomised trial. BMC Med Educ 2012; 12:41.

[56] Oliven A, Nave R, Gilad D, Barch A. Implementation of a web-based interactive virtual patient case simulation as a training and assessment tool for medical students. Stud Health Technol Inform 2011; 169:233-37.

[57] Persky S, Eccleston CP. Medical student bias and care recommendations for an obese versus non-obese virtual patient. Int J Obes (Lond) 2011; 35(5):728-35.

[58] Lunney CA, Kleinert HL, Ferguson JE 2nd, Campbell L. Effectively training pediatric residents to deliver diagnoses of Down syndrome. Am J Med Genet A 2012; 158A(2):384-90.

[59] Botezatu M, Hult H, Tessma MK, Fors U. Virtual patient simulation: knowledge gain or knowledge loss? Med Teach 2010; 32(7):562-68.

[60] Botezatu M, Hult H, Tessma MK, Fors UG. Virtual patient simulation for learning and assessment: Superior results in comparison with regular course exams. Med Teac 2010; 32(10):845-50.

[61] Andreatta PB, Maslowski E, Petty S. Virtual reality triage training provides a viable solution for disaster-preparedness. Acad Emerg Med 2010; 17(8):870-76.

[62] Gucwa A, Raij A, Kotranza A. Could this be cancer? Self-reflection of emapthetic responses using an immersive "through the eyes of a virtual patient" feedback system. J Surg Res 2010; 158(2):182.

[63] Kandasamy T, Fung K. Interactive Internet-based cases for undergraduate otolaryngology education. Otolaryngol Head Neck Surg 2009; 140(3):398-402.

[64] Deladisma AM, Gupta M, Kotranza A. A pilot study to integrate an immersive virtual patient with a breast complaint and breast examination simulator into a surgery clerkship. Am J Surg 2009; 197(1):102-6.

[65] Youngblood P, Harter PM, Srivastava S, Moffett S, Heinrichs WL, Dev P. Design, development, and evaluation of an online virtual emergency department for training trauma teams. Simul Healthc 2008; 3(3):146-53.

[66] Vukanovic-Criley JM, Boker JR, Criley SR, Rajagopalan S, Criley JM. Using virtual patients to improve cardiac examination competency in medical students. Clin Cardiol 2008; 31(7):334-39.

[67] Boyd SE, Sanders CL, Kleinert HL. Virtual patient training to improve reproductive health care for women with intellectual disabilities. J Midwifery Women's Health 2008; 53(5):453-60.

[68] Sijstermans R, Jaspers MW, Bloemendaal PM, Schoonderwaldt EM. Training inter-physician communication using the Dynamic Patient Simulator. Int J Med Inform 2007; 76(5-6):336-43.

[69] Deladisma AM, Cohen M, Stevens A. Do medical students respond empathetically to a virtual patient? Am J Surg 2007; 193(6):756-60.

[70] Vash JH, Yunesian M, Shariati M, Keshvari A, Harirchi I. Virtual patients in undergraduate surgery education: a randomized controlled study. ANZ J Surg 2007; 77(1-2):54-9.

[71] Stevens A, Hernandez J, Johnsen K. The use of virtual patients to teach medical students history taking and communication skills. Am J Surg 2006; 191(6):806-11.

[72] Nendaz MR, Ponte B, Gut AM. Liveorcomputerized simulation of clinical encounters: do clinicians work up patient cases differently? Med Inform Internet Med 2006; 31(1):1-8.

[73] Triola M, Feldman H, Kalet AL. A randomized trial of teaching clinical skills using virtual and live standardized patients. J Gen Intern Med 2006; 21(5):424-29.

[74] Ferguson JE 2nd, Kleinert HL, Lunney CA, Campbell LR. Resident physicians' competencies and attitudes in delivering a postnatal diagnosis of Down syndrome. Obstet Gynecol 2006; 108(4):898-905.

[75] Bearman M, Cesnik B, Liddell M. Random comparison of 'virtual patient' models in the context of teaching clinical communication skills. Med Educ 2001; 35(9):824-32.

[76] Bearman M, Cesnik B. Comparing student attitudes to different models of the same virtual patient. Stud Health Technol Inform 2001; 84(Pt 2):1004-8.

[77] Dickerson R, Johnsen K, Raij A, Lok B, Stevens A, Bernard T, et al. Virtual patients: assessment of synthesized versus recorded speech. Stud Health Technol Inform 2006; 119:114-9.

[78] Cook DA, Hatala R, Brydges R. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA. 2011; 306(9):978-988.

[79] Lang VJ, Kogan J, Berman N, Torre D. The evolving role of online virtual patients in internal medicine clerkship education nationally. Acad Med 2013; 88(11):1713-18.

[80] Kononowicz AA, Zary N, Edelbring S, Corral J, Hege I. Virtual patients-what are we talking about? A framework to classify the meanings of the term in healthcare education. BMC Med Edu. 2015; 15:11.

[81] Weaver SJ, Dy SM, Rosen MA. Team-training in healthcare: A narrative synthesis of the literature. BMJ Qual Saf 2014; 23(5):359-72.

[82] Leape LL, Shore MF, Dienstag JL. Perspective: A culture of respect, part 2: creating a culture of respect. Acad Med 2012; 87(7):853-58.

[83] Cendan J, Lok B. The use of virtual patients in medical school curricula. Adv Physiol Educ 2012; 36(1):48-53.

[84] Poulton T, Balasubramaniam C. Virtual patients: A year of change. Med Teach 2011; 33(11):933-37.

[85] Posel N, McGee JB, Fleiszer DM. Twelve tips to support the development of clinical reasoning skills using virtual patient cases. Med Teach 2015; 37(9):813-18.

[86] Drucker AM, Fleming P, Chan AW. Research Techniques Made Simple: Assessing Risk of Bias in Systematic Reviews. J Invest Dermatol 2016; 136(11):e109-e114.

[87] Almarzooq ZI, Lopes M, Kochar A. Virtual Learning During the COVID-19 Pandemic: A Disruptive Technology in Graduate Medical Education. J Am Coll

Cardiol 2020;75(20):2635-38.

[88] Foronda C, Armstrong B. Society for Simulation in Healthcare (SSH) and International Nursing Association of Clinical Simulation and Learning (INACSL): Position Statement on Use of Virtual Simulation during the Pandemic 2020.

[89] Hilty D, Chan S, Torous J, Luo J, Boland R. A framework for competencies for the use of mobile technologies in psychiatry and medicine: Scoping review. JMIR Mhealth Uhealth 2020; 8(2):e12229.

[90] Küfner J, Kononowicz AA, Hege I. Virtual patient repositories-A comparative analysis. Stud Health Technol Inform 2014; 205:788-792.