The Effectiveness of Using Virtual Patient Educational Tools to Improve Medical Students' Clinical Reasoning Skills: A Systematic

Ruth Plackett (ruth.plackett.15@ucl.ac.uk)
UCL
Angelos P. Kassianos
UCL
Sophie Mylan
London School of Hygiene & Tropical Medicine
Maria Kambouri
Institute of Education, UCL
Rosalind Raine
UCL
Jessica Sheringham
UCL

Research Article

Keywords: Computer simulation, virtual patient, computer-assisted instruction, educational technology, medical education, clinical decision-making, clinical reasoning, clinical skills, review, medical students

Posted Date: December 20th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1117083/v1

Abstract

Background

Use of virtual patient educational tools could fill the current gap in the teaching of clinical reasoning skills. However, there is a limited understanding of their effectiveness. The aim of this study was to synthesise the evidence to understand the effectiveness of virtual patient tools aimed at improving undergraduate medical students' clinical reasoning skills.

Methods

We searched MEDLINE, EMBASE, CINAHL, ERIC, Scopus, Web of Science and PsycINFO from 1990 to October 2020, to identify all experimental articles testing the effectiveness of virtual patient educational tools on medical students' clinical reasoning skills. Quality of the articles was assessed using an adapted form of the MERSQI and the Newcastle-Ottawa Scale. A narrative synthesis summarised intervention features, how virtual patient tools were evaluated and reported effectiveness.

Results

The search revealed 7,290 articles, with 20 articles meeting the inclusion criteria. Average study quality was moderate (M=7.1, SD=2.5), with around a third not reporting any measurement of validity or reliability for their clinical reasoning outcome measure (7/20, 35%). Eleven articles found a positive effect of virtual patient tools on reasoning (11/20, 55%). Seven (7/20, 35%) reported no significant effect or mixed effects and two found a significantly negative effect (2/20, 10%). Several domains of clinical reasoning were evaluated. Data gathering, ideas about diagnosis and patient management were more often found to improve after virtual patient use (27/46 analyses, 59%) than knowledge, flexibility in thinking, problem-solving, and critical thinking (4/10 analyses, 40%).

Conclusions

Using virtual patient tools could effectively complement current teaching especially if opportunities for face-to-face teaching or other methods are limited, as there was some evidence that virtual patient educational tools can improve undergraduate medical students' clinical reasoning skills. Evaluations that measured more case specific clinical reasoning domains, such as data gathering, showed more consistent improvement than general measures like problem-solving. Case specific measures might be more sensitive to change given the context dependent nature of clinical reasoning. Consistent use of validated clinical reasoning measures is needed to enable a meta-analysis to estimate effectiveness.

Background

It has been recommended that more explicit training should be provided in undergraduate medical education on applying clinical reasoning skills, to reduce the impact of future diagnostic errors and potential patient harm[1–4]. Clinical reasoning refers to the thought processes and steps involved in making a clinical judgement[2, 5]. Clinical reasoning requires several complex cognitive skills and is a context dependent skill[2]. It is an evolving and cyclical process that involves applying medical knowledge, gathering necessary information from patients and other sources, interpreting (or reinterpreting) that information and problem formulation (or reformulation)[2, 5]. To be proficient in clinical reasoning, clinicians need to also acquire the requisite knowledge and skills in reflective enquiry[2].

Currently, teaching of clinical reasoning in most medical schools in the UK remains a largely implicit component of small group tutorials, problem-based learning, clinical communication skills sessions, and clinical placements[3]. Making the teaching of these skills more explicit may help students to reflect on their skills, which many models of learning suggest is essential for improving skills[6, 7]. Virtual patient educational tools are becoming increasingly popular in medical education and have been used to explicitly teach clinical reasoning skills[5, 8]. They are defined as "A specific type of computer-based program that simulates real-life clinical scenarios; learners emulate the roles of health care providers to obtain a history, conduct a physical exam, and make diagnostic and therapeutic decisions"[9]. They allow students to practise clinical reasoning with realistic patients, in a safe environment[5, 9–11]. They may also be particularly suited to providing training on clinical reasoning skills that require deliberate practice with a wide variety and large number of clinical cases. Indeed, many students may have limited contact with patients, where it is also not possible to pre-determine what range of presentations and problems students will meet[5]. Educational and cognitive theories, and empirical research also suggest that virtual patient educational tools could provide an ideal platform for developing clinical reasoning skills if they incorporate best practice features for simulation-based educational tools, in particular providing opportunities for feedback and reflection[6, 7, 10, 12].

Previous systematic reviews and meta-analyses conducted between 2008 and 2012 have indicated that online learning, including virtual patient tools, can significantly improve the learning outcomes of both health professionals and students[13–15]. However, since these reviews were conducted, online learning technologies and the place of online learning in medical education has changed substantially. Furthermore, there was limited information in these reviews about whether best practice features for simulation-based educational tools were incorporated into the virtual patient tools and how they might relate to effectiveness. There were also no sub-group analyses to show the specific effect of these interventions on undergraduate medical students, who have different training needs and ways of learning compared to professionals. Previous reviews have also not explored the effectiveness of virtual patient tools at improving clinical reasoning skills as a specific outcome[13–15]. Thus, there is insufficient evidence for educators to understand the impacts of virtual patient educational tools on clinical reasoning skills[14, 16]. This review, therefore, aims to address the question "How effective are virtual patient educational tools at improving the clinical reasoning skills of undergraduate medical students?". Other objectives of this review were to:

a) identify the use of empirically and theoretically informed intervention features in virtual patient tools, such as reflection;

b) identify the outcome measures used to assess clinical reasoning skills.

Methods

This systematic review was conducted following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and the PRISMA checklist is available as Additional File 1; the review protocol was presented in RP's doctoral thesis[17, 18].

Inclusion and exclusion criteria

Table 1 describes in detail the inclusion and criteria for this review.

Table 1
Inclusion and exclusion criteria

Key Concepts	Criteria
Population	Undergraduate medical students. Excluded: health professionals, postgraduate students, other health students.
Intervention	Interventions that describe an educational method that is distributed and facilitated online and simulates a real-life clinical scenario between a 'physician' and 'patient'. The student should emulate the role of a clinician by gathering data from the patient, interpreting information, and making diagnostic decisions. Excluded: high fidelity simulators, manikins, standardised patients, and decision support tools.
Comparator	Teaching as usual, an alternative instructional method e.g., face-to-face, or paper-based instruction. Excluded: alternative formats e.g., comparing different types of patient cases.
Outcome	Clinical reasoning skills are the thought processes required to identify likely diagnoses, formulate appropriate questions and reach clinical decisions[2]. Interventions that provided sufficient detail to establish whether it improved clinical reasoning skills in a written, oral, or practical test. Commonly used synonyms for clinical reasoning were accepted e.g., clinical decision-making, clinical reasoning, problem-solving, critical thinking, and clinical judgement skills.
Study type(s)	RCTs, crossover trials, quasi-experimental studies, and observational studies. Excluded: qualitative designs.
Publication type(s)	Peer reviewed articles including theses. Excluded: conference papers, editorials letters, notes, comments, and meeting abstracts. Articles not in English.
Time	Articles from the year 1990, as this was when online learning was beginning to be described[15].

Search strategy

We applied a search strategy for the following databases: MEDLINE, EMBASE, CINAHL, ERIC, Scopus, Web of Science and PsycINFO, from 1990 to July 2016 and the search was updated to include all articles up to October 2020. Further articles were identified by hand searching the reference lists of included articles. Search terms included a combination of subject headings and key word searches. The full search strategy used in MEDLINE is available as Additional File 2.

Study selection

One author (RP) screened all the articles retrieved from the search by title and abstract for eligibility of inclusion. Another author (APK) double screened a proportion of the abstracts (2.3%, n=116/4977), with moderate agreement (Cohen's Kappa=0.59). Discrepancies were resolved in a consensus meeting and articles were included for full text screening if the abstract lacked enough detail to confirm eligibility. One of the authors (RP) screened all the full text articles and APK double screened a proportion of these articles (43.9%, n=54/123), with substantial agreement (Cohen's Kappa=0.66). Discrepancies were resolved in a consensus meeting with the wider team.

Data extraction

Data on study design, population, setting, delivery of intervention, outcomes, results, and limitations was extracted in an Excel spreadsheet. We also extracted data on the features that were included in the virtual patient tools, such as reflection and feedback. APK and SM piloted the data extraction form with two articles. RP extracted data from 13 articles included in the review, APK extracted data from eight and SM extracted data from one. All extractions were double-checked by either RP, APK and SM; discrepancies were resolved in a consensus meeting.

Quality assessment

Three authors (RP, APK and SM) assessed the quality of the included articles independently. Quality was assessed using a checklist that incorporated items from two previously developed checklists, the Medical Education Research Study Quality Instrument (MERSQI) and an adapted form of the Newcastle-Ottawa Scale (NOS), which have both been used in previous reviews in this area[15, 16, 19, 20]. The two checklists were incorporated as the NOS was designed to identify aspects of quality related to potential biases in the study design and sample selection, and the MERSQI was designed to identify other aspects of quality, such as the validity and reliability of outcome measures. In addition, articles were given a point if they described how theory informed assessment of clinical reasoning skills or used a previously validated measure that was based on theory e.g., key features problems[21]. Articles could receive a score of up to 14, with scores ranging from 0-4 suggesting low quality, scores of 5-9 suggesting moderate quality and scores of 10-14 indicating high quality.

Data analyses

We conducted a narrative synthesis of the included articles to address the review objectives. We summarised the characteristics of the interventions to understand what features were included in virtual patient tools and how they were delivered. The study designs used to evaluate the virtual patient tools and the reported effectiveness of each intervention were also reported; Cohen's *d* effect size was calculated where possible. We also summarised the various clinical reasoning outcome measures used and grouped outcomes measured in each article into specific domains of clinical reasoning informed by the model of clinical reasoning by Higgs et al. [2] and author descriptions of the clinical reasoning outcomes they measured. The analysis of clinical reasoning domains was undertaken at the level of analyses, as articles often reported on more than one domain, and so each domain was included separately in the analysis. In all the articles it was possible to identify at least one domain of clinical reasoning that was measured. Most articles (17/20, 85%) used an aggregate score to represent several domains of clinical reasoning.

Results

Study characteristics

The search strategy identified 7,290 records of which 20 were included in the review. See Figure 1 for the PRISMA flow diagram of the number of articles included at each stage of the review. The most common study locations were Germany (7/20, 35%) and the USA (5/20, 25%). Most of the articles were published since 2010 (16/20, 80%).

Intervention features

Table 2 describes the characteristics of the interventions. There was a great variety of virtual patient tools that were used to improve reasoning; only two - MedU[22, 23] and EMERGE[24, 25] - were evaluated in more than one study. Just under half of the interventions (8/20, 40%) required the students to gather information from the virtual patient, and were more interactive, while half (10/20, 50%) were less interactive and presented patients with the patient history already completed. There was not enough information in two articles to determine interactivity (2/20, 10%)[23, 26]. Most of the interventions (15/20, 75%) required students to work individually rather than in groups. Those that were delivered in groups required students to work together to complete the case and make decisions. The clinical topic of the interventions varied; cardiology followed by paediatrics were the most common topics (6/20, 30% and 3/20, 15% respectively). The number of patient cases within the virtual patient tools ranged from 1-40, with two patient cases being the most common number (5/20, 25%). The duration of the patient cases varied from approximately nine minutes to complete a case[27] to 10 hours to complete one case (over several weeks)[28]. Most commonly students had multiple opportunities to use and complete the patient cases (17/20, 85%).

Most interventions provided feedback to students on their performance (15/20, 75%). They did this in several ways including: providing the correct answers, providing feedback from experts on how they would have completed the case either via text or video, and discussing answers with a facilitator after completing a case. Reflection was explicitly described in one intervention where users were prompted to reflect during each patient case on their decisions and were required to complete open-ended reflection questions at the end of each case[29]. There were two interventions where the use of reflection was implied, but it was unclear from their description whether the activities were explicitly for reflection[30, 31].

Table 2
Characteristics of the interventions

First author (year)	Country	Virtual Patient tool name	Need to gather data	Delivery	Clinical topic	No. cases	Approximate time to complete one case	Delivered on single or multiple occasions	Feed-back used	Reflection used
Aghili et al. 2012	Iran	Not reported	Yes	Solo	Endocrinology	2	Not reported	Multiple	Yes	No
Basu Roy & McMahon 2012	USA	Not reported	No	Group	Endocrinology and reproduction	2	2 hours	Multiple	No	No
Botezatu et al. 2010	Colombia	Web-SP	No	Solo	Haematology and cardiology	6	1 hour	Multiple	Yes	No
Chon et al. 2019	Germany	EMERGE	Yes	Solo	Surgery	4	15 mins	Multiple	Yes	No
Devitt & Palmer 1998	Australia	MEDICI	No	Solo	Liver disease	5	18 mins	Multiple	Yes	No
Isaza-Restrepo et al. 2018	Colombia	The Virtual Patient: Simulator of Clinical Case	Yes	Solo	Gastroenterology	16	2 hours	Multiple	Yes	No
Kahl et al. 2010	Germany	Not reported	No	Group	Psychiatry	Not reported	Not reported	Multiple	No	No
Kalet et al. 2007	USA	WISE-MD	No	Solo	Surgery	Not reported	Not reported	Multiple	No	No
Kamin et al. 2003	USA	Project L.I.V.E.	No	Group	Paediatrics	1	1.5 hours	Multiple	Yes	No
Kim et al. 2018	USA	MedU	No	Solo	Multiple	22 (these were required but access to more)	Not reported	Multiple	Yes	No
Kleinart et al. 2015	Germany	ALICE	Yes	Solo	Cancer	3	Not reported	Single	Yes	No
Lehman et al. 2015	Germany	CAMPUS	No	Solo	Paediatrics	2	1 hour	Multiple	Yes	No
Ludwig et al. 2018	Germany	Not reported	No	Solo	Multiple	30	15 mins	Multiple	No	No
McCoy 2014	USA	Decision Simulation™	Yes	Group	Cardiology	2	20 mins	Single	Yes	No
Middeke et al. 2018	Germany	EMERGE	Yes	Solo	Accident and emergency	40	9 mins	Multiple	Yes	No
Plackett et al. 2020	UK	eCREST	Yes	Solo	Cardio-respiratory	3	13 mins	Multiple	Yes	Yes
Raupach et al. 2009	Germany	Clix ®	No	Group	Cardio-respiratory	1	10 hours	Multiple	No	No
Sobocan et al. 2017	Slovenia	MedU	Not reported	Solo	Internal medicine	Not reported	Not reported	Multiple	No	No

First author (year)	Country	Virtual Patient tool name	Need to gather data	Delivery	Clinical topic	No. cases	Approximate time to complete one case	Delivered on single or multiple occasions	Feed-back used	Reflection used
Watari et al. 2020	Japan	®Body Interact, Coimbra, Portugal	Not reported	Solo	Cardiology and psychiatry	2	20 mins	Single	No	No
Wu et al. 2014	China	Not reported	Yes	Solo	Nephrology	4	5 hours	Multiple	No	No

Study designs and participants

Table 3 describes the characteristics of the included articles including study design, outcome measures used and reported effectiveness. Just under half of the articles were RCTs (8/20, 40%), one was a feasibility RCT (1/20, 5%)[29] and three were randomised crossover trials (3/20, 15%)[32–34]. A small proportion were non-randomised trials (3/20, 15%)[22, 25, 35] or single group pre-test and post-test design (5/20, 25%). Just under a third of evaluations (6/20, 30%) compared virtual patient tools with tutorials or small group discissions. A quarter (5/20, 25%) compared virtual patient tools to teaching as usual, which included no additional clinical reasoning teaching via any method. Only four evaluations (4/20, 20%)[23, 32, 33, 36] compared virtual to text-based cases. There was a wide variety of year groups that interventions were evaluated with, ranging from those in their 1st year of medical school to those in their 6th year. In most of the evaluations, participants were in their 3rd, 4th, or 6th year (8/20, 40% respectively).

Table 3
Characteristics of included articles ordered by comparator

Authors and reference number	Aim(s) of the study	Research Design	Participants - year group and total n	Domain of clinical reasoning measured	Outcome measure	Main results	Quality (score out of 14)
Comparator: teaching as usual							
Aghili et al. 2012	To evaluate whether virtual patient simulations improve clinical reasoning skills of medical students.	RCT	6th years. N=52 (29 IG, 23 CG)	Data gathering, ideas about patient management	Diagnostic test (using patient cases)	ñ Intervention produced significantly greater improvement in data gathering and ideas about patient management compared to teaching as usual (*d*=1.55).	Moderate (6)
Kalet et al. 2007	To assess the impact of individual WISE-MD modules on clinical reasoning skills.	RCT	Clinical years. N=96 (52 IG, 44 CG)	Data gathering, ideas about patient management	Script concordance test	ñ Intervention produced significantly greater improvement in data gathering and ideas about patient management compared to teaching as usual (*d*=0.25).	Moderate (9)
Lehman et al. 2015	Investigated the effect of Virtual Patients combined with standard simulation-based training on the acquisition of clinical decision-making skills and procedural knowledge, objective skill performance, and self-assessment.	RCT	3rd & 4th years. N=57 (30 IG, 27 CG)	Ideas about diagnoses, ideas about patient management, knowledge	Key feature problems	ñ Intervention produced significantly greater improvement in ideas about diagnoses and patient management, and knowledge compared to teaching as usual (*d*=1.91).	High (13)
Plackett et al. 2020	To assess the feasibility, acceptability and potential effects of eCREST — the electronic Clinical Reasoning Educational Simulation Tool.	Feasibility RCT	5th & 6th years. N=264 (137 IG, 127 CG)	Data gathering, flexibility in thinking about diagnoses (reported separately)*	Virtual patient case & Diagnostic Thinking Inventory (DTI)	ñ Ability to gather essential information (data gathering; *d*=0.19) significantly improved after intervention compared to teaching as usual ó There was no significant difference between groups in relevance of history taking (data gathering; *d*=-0.13) and flexibility in diagnoses (*d*=0.20).	High (11)
Kim et al. 2018	To explore how students use and benefit from virtual patient cases.	Non-randomised trial	3rd years. N=255 (129 IG, 126 CG)	Ideas about diagnoses, knowledge	Standardised patient (actor)	ó Ideas about diagnoses and knowledge did not significantly improve compared to teaching as usual (voluntary access to cases) (*d*=0.09).	Moderate (8)
Comparator: tutorial covering the same case							
Botezatu et al. 2010	To explore possible superior retention results with Virtual Patients versus regular learning activities, by measuring the differences between early and delayed assessment results.	RCT	4th & 6th years. N=49 (25 IG, 24 CG)	Data gathering, ideas about diagnoses, ideas about patient management	Virtual patient cases	ñ Intervention produced significantly greater improvement in data gathering, ideas about diagnoses and patient management compared to tutorial (average effect size across 5 dimensions, *d*=1.57).	Moderate (6)
Devitt & Palmer 1998	To evaluate the intervention by assessing whether it expanded students' knowledge base, improving data-handling abilities and clinical problem-solving skills.	RCT	5th years. N=71 (46 IG, 25 CG)	Problem-solving skills	Multi-step clinical problem (patient case)	ó Intervention produced non-significantly greater improvement in problem-solving skills compared to tutorial (*d*=0.50).	Moderate (6)

* 3 articles reported the impact of the virtual patient tools on each domain of clinical reasoning separately while all others reported an aggregate impact score across several domains of reasoning.

Authors and reference number	Aim(s) of the study	Research Design	Participants - year group and total n	Domain of clinical reasoning measured	Outcome measure	Main results	Quality (score out of 14)
McCoy 2014	This study investigates the utility of Virtual Patients for increasing medical student clinical reasoning skills, collaboration, and engagement.	Randomised crossover trial	1st years. N=108 (54 IG, 54 CG)	Ideas about diagnoses	Diagnostics competency task (using patient cases)	ò Intervention significantly lowered ideas about diagnoses compared to tutorial (*d*=-0.59).	Moderate (9)
Raupach et al. 2009	To explore whether students completing a web based collaborative teaching module show higher performance in a test aimed at clinical reasoning skills than students discussing the same clinical case in a traditional teaching session.	RCT	4th years. N=143 (72 IG, 71 CG)	Data gathering, ideas about diagnoses, Ideas about patient management	Key feature problems	ó Intervention did not significantly improve data gathering, ideas about diagnoses and patient management compared to tutorial (*d*=0.03)	High (10)
Kamin et al. 2003	To determine whether critical thinking differs among groups receiving the same case with the same facilitator in one of three formats.	Non-randomised trial	3rd years. N=65 (25 IG- virtual, 20 – IG video, 20 – CG – text) [iv]	Critical thinking	Students' critical thinking during discussions of patient cases	ñ Intervention produced significantly better critical thinking than the tutorials with either text-based cases (average effect size across 5 dimensions of critical thinking, *d*=2.20) or video modality (average effect size across 5 dimensions of critical thinking, *d*=2.72).	Moderate (6)
Middeke et al. 2018	To compare a Serious Game, the virtual A&E department 'EMERGE' to small-group problem-based learning (PBL) regarding student learning outcome on clinical reasoning in the short term.	Non-randomised trial	5th years, N=112 (78 IG, 34 CG)	Data gathering, ideas about diagnoses, ideas about patient management (reported separately)	Key feature problems & virtual patient cases	ñ Intervention produced significantly better clinical reasoning skills compared to tutorial (*d*=0.47) when measured on key features test and for some domains measured by the virtual patient cases – final diagnosis (ideas about diagnoses), therapeutic interventions (ideas about patient management), physical examination, instrumental examination (data gathering) ó There was no significant difference between groups in history taking (data gathering), laboratory orders and patient transfer (ideas about patient management).	Moderate (6)
Comparator: text-based cases							
Basu Roy & McMahon 2012	To explore video-based cases comparative impact on students' critical thinking.	Randomised crossover trial	2nd years. N=28 (14 IG, 14 CG)	Critical thinking	Proportion of deep utterances and superficial utterances during discussions of patient cases	ò Intervention produced significantly lower odds of deep thinking compared to text-based cases (*d*=-0.23).	Moderate (9)

* 3 articles reported the impact of the virtual patient tools on each domain of clinical reasoning separately while all others reported an aggregate impact score across several domains of reasoning.

Authors and reference number	Aim(s) of the study	Research Design	Participants - year group and total n	Domain of clinical reasoning measured	Outcome measure	Main results	Quality (score out of 14)
Kahl et al. 2010	To explore whether the addition of systematic training in iterative hypothesis testing may add to the quality of the psychiatry course taught to fifth year medical students.	RCT	5th years. N=72 (36 IG, 36 CG)	Ideas about diagnoses	Standardised patient (actor)	ñ Intervention produced significantly greater improvements in ideas about diagnoses compared to using text-based cases with examination of real patients (*d*=1.17).	Moderate (7)
Ludwig et al. 2018	To test the hypothesis that repeated testing with video-based key feature questions produces superior retention of procedural knowledge related to clinical reasoning compared to repeated testing with text-based key feature questions.	Randomised crossover trial	4th years. N=93	Data gathering, ideas about diagnoses, ideas about patient management, knowledge	Key feature problems	ñ Intervention produced significantly greater improvements in data gathering, ideas about diagnoses and patient management and knowledge compared to using text-based cases (*d* not possible to calculate).	Moderate (6)
Sobocan et al. 2017	To determine the educational effects of substituting p-PBL sessions with VP on undergraduate medical students in their internal medicine course.	RCT	3rd years. N=34 (17 IG, 17 CG)	Knowledge and flexibility in thinking	DTI	ó Intervention did not significantly improve knowledge and flexibility in thinking compared to text-based cases (*d*=0.25).	Moderate (7)
Comparator: N/A							
Chon et al. 2019	To test the effect of a serious game simulating an emergency department ("EMERGE") on students' declarative and procedural knowledge	Single group pre & post comparison	Clinical years. N=140	Data gathering, ideas about diagnoses, ideas about patient management, (reported separately)	Patient case	ñ Diagnostic questions (data gathering; *d*=0.77), choosing the correct order of diagnostic procedures (ideas about diagnoses; *d*=0.65) and treatment suggestions improved (ideas about patient management; *d*=0.82) after using intervention. ó There was no significant difference between groups in diagnostic accuracy (ideas about diagnoses; *d*=0.08).	Moderate (5)
Isaza-Restrepo et al. 2018	To present evidence regarding the effectiveness of a low-fidelity simulator: Virtual Patient	Single group pre & post comparison	1st-5th years. N=20	Data gathering, ideas about diagnoses, ideas about patient management	Standardised patient (actor)	ñ Data gathering, ideas about diagnoses and patient management, and presentation of a case significantly improved after using intervention (average effect size across 5 dimensions from 3 evaluators, *d*=1.41).	Moderate (6)
Kleinart et al. 2015	To examine whether the use of ALICE has positive impact on clinical reasoning and is a suitable tool for supporting the clinical teacher.	Single group pre & post comparison	3rd years. N=62	Ideas about diagnoses, ideas about patient management	Patient cases	ñ Ideas about diagnoses and patient management significantly improved after using intervention (*d*=0.92).	Low (3)

* 3 articles reported the impact of the virtual patient tools on each domain of clinical reasoning separately while all others reported an aggregate impact score across several domains of reasoning.

Authors and reference number	Aim(s) of the study	Research Design	Participants - year group and total n	Domain of clinical reasoning measured	Outcome measure	Main results	Quality (score out of 14)
Watari et al. 2020	To clarify the effectiveness of VPSs for improving clinical reasoning skills among medical students, and to compare improvements in knowledge or clinical reasoning skills relevant to specific clinical scenarios.	Single group pre & post comparison	4th years. N=169	Data gathering, ideas about diagnoses, ideas about patient management	Multiple-choice question (MCQ) quiz (using patient cases)	ñ Data gathering, ideas about diagnoses and patient management significantly improved after using intervention (*d*=1.39).	Low (3)
Wu et al. 2014	To examine the effectiveness of a computer-based cognitive representation approach in supporting the learning of clinical reasoning.	Single group pre & post comparison	3rd-5th years. N=50	Problem-solving	Concept maps	ñ Problem-solving significantly improved after using intervention (*d*=1.17).	Moderate (5)
* 3 articles reported the impact of the virtual patient tools on each domain of clinical reasoning separately while all others reported an aggregate impact score across several domains of reasoning.							

Outcome measures

Seven domains of clinical reasoning were identified. Four domains reflected the underlying general cognitive processes required in clinical reasoning and these included: knowledge of the clinical problem derived from theory or experience (4/20, 20%); flexibility in thinking about diagnoses[23, 29]; problem-solving skills[37, 38] and critical thinking[33, 35] (2/20, 10% respectively). One domain reflected more case specific clinical reasoning processes that were measured via data gathering skills, including the relevance of patient examinations (10/20, 50%). Two domains measured the outcomes of the clinical reasoning process in specific cases by measuring the clinical judgements the students made. These included: ideas about diagnoses, including diagnostic accuracy (12/20, 60%), and ideas about patient management, including appropriateness of treatment plans or therapeutic decisions (11/20, 55%).

Half of the evaluations (10/20, 50%) used measures of clinical reasoning that have been previously reported and validated in the wider literature. These included: key features problems[21, 39](4/20, 20%) [25, 28, 32, 40]; Standardised Patients, where an actor simulates a patient (3/20, 15%)[22, 31, 36]; the Script Concordance Test[41] (1/20, 5%)[42] and the Diagnostic Thinking Inventory [43] (DTI; 2/20, 10%)[23, 29]. In six evaluations (6/20, 30%) student performance was assessed using text-based cases that the authors had developed, often followed by open or multiple choice questions regarding history taking, diagnosis and treatment[24, 26, 34, 37, 44, 45], three used additional virtual patient cases (3/20, 15%)[25, 29, 46] and one used concept maps (1/20, 5%) to assess five aspects of performance[38]. Two articles (2/20, 10%) assessed reasoning by assessing critical thinking and evidence of deep thinking in students' discussions about a patient case[33, 35].

Quality of included articles

Additional file 3 gives a detailed breakdown of the quality of the included articles. The average quality was moderate (M=7.1, SD=2.5). Only three articles (3/20, 15%) were high quality[28, 29, 40], most were of moderate quality (15/20, 75%) and 2 were of low quality (2/20, 10%)[26, 45]. Nearly three quarters of articles (14/20, 70%) described how theory informed the evaluation, by either describing theoretical frameworks they used to assess clinical reasoning or using previously developed and validated measures of clinical reasoning. Only four articles (4/20, 20%) reported measuring three or more different types of validity and reliability and around a third did not report any measurement of validity or reliability (7/20, 35%). Only one article reported that they selected students from more than one medical school[29]. A quarter of articles (5/20, 25%) reported that the assessor of the outcome was blinded to group allocation. Just under half (8/20, 40%) reported a power calculation, although this was not necessary to calculate for all study designs.

Reported effectiveness

Just over half of the articles (11/20, 55%) reported that virtual patient tools had significantly positive effects on medical students' clinical reasoning skills, 20% (4/20) of articles reported no effect, three articles showed mixed effects (3/20, 15%) and two articles found adverse effects of virtual patient tools on clinical reasoning skills.

Effectiveness by article quality

Of the 3 articles rated as high-quality, 1 found no significant effect of virtual patients on reasoning[28], 1 a positive effect (1/3, 33%)[40], and 1 a mixed effect[29]. Out of the articles that were rated as moderate quality, most reported virtual patient tools had significant benefits (9/15, 60% respectively) than mixed (1/15, 7%)[24], neutral (3/15, 20%)[22, 23, 37] or adverse effects (2/15, 13%)[33, 34]. The 2 articles that were rated as low quality both reported virtual patient tools had significant benefits (2/2, 100%; Figure 2)[26, 45].

Effectiveness by study design

Of the articles that used randomised study designs (12/20, 60%), 50% (6/12) reported that virtual patient tools improved clinical reasoning skills compared with controls [32, 36, 40, 42, 44, 46]. Only two (2/12, 17%) reported that virtual patients have a negative effect on clinical reasoning skills[33, 34]. A third of randomised study designs (4/12, 33%) reported that virtual patient tools had mixed effects or no significant effect on clinical reasoning skills compared to controls[23, 28, 29, 37]. Of the articles that used non-randomised trial study designs (3/20, 15%), one article reported that virtual patient tools improved clinical reasoning skills compared to controls[35], one found mixed effects [25] and one found no significant effects[22]. Of the five articles (5/20, 25%) that used a single group pre and post study design, four articles (4/5, 80%) found a significant improvement in clinical reasoning after using virtual patient tools[26, 31, 38, 45]; only one article (1/5, 20%) reported mixed results (Figure 2)[24].

Effectiveness by comparator

Articles that compared virtual patient tools with teaching as usual (5/20, 25%) reported mostly (3/5, 60%) positive effects on clinical reasoning[40, 44, 47], but some found mixed or no effects on reasoning (2/5, 40%)[22, 29]. Articles that compared virtual patient tools to text-based cases or tutorials (10/20, 50%) had varied effectiveness, as four (4/10, 40%) showed positive effects of virtual patient tools, one showed mixed effects (1/10, 10%)[25], three (3/10, 30%) showed no effect [23, 28, 37] and two (2/10, 20%) showed adverse effects on clinical reasoning (Figure 2)[33, 34].

Effectiveness by domain of clinical reasoning measured and measurement

Data gathering, ideas about diagnoses and patient management were largely found to significantly improve after virtual patient use (27/46 analyses, 59%; Figure 3). Knowledge, flexibility in thinking about diagnoses, problem-solving skills, and critical thinking showed more mixed results, with 40% of analyses showing significant improvement in these skills (4/10 analyses).

Of the 9 articles that used a patient case (text or virtual) and a bespoke measuring rubric to assess clinical reasoning, over three quarters reported positive effects of using virtual patient tools (7/9, 78%), 1 article reported neutral[37] and 1 adverse effects [34]. Articles that used measures of clinical reasoning that have been developed and validated in previous literature, such as the key feature problems, reported mostly significant benefits of using virtual patient tools (7/10, 70%) and under half reported no significant effects (3/10, 30%)[22, 23, 28].

Discussion

This review of published evaluations of virtual patient educational tools found there is some evidence that they can improve medical students' clinical reasoning. Improvements were more consistently reported for domains of clinical reasoning that were more case specific, such as ideas about diagnoses and data gathering, rather than more general reasoning processes, such as problem-solving.

Intervention features

This review illustrates the diversity in design, content, and delivery of virtual patient tools and the clinical context in which they are applied. Most virtual patient educational tools have been designed for individuals to complete. Many of the tools included features that educational theories and empirical research suggests are important to include in simulation-based learning, such as feedback, but relatively few reported how they facilitated reflection[27, 29, 31, 36]. Further consideration of how to facilitate reflection when using virtual patient tools could allow them to be even more effective at developing reasoning skills[7, 30, 48]. There was also variety in the level of interactivity with the virtual patient tools, with half of the tools not requiring students to gather information from the patient. Previous research is inconclusive as to whether greater interactivity produces better learning outcomes[49]. Studies have shown greater interactivity can facilitate deeper learning and more engagement from users, but it can also increase cognitive load, which can interfere with learning[47, 49]. However, virtual patient tools that allow for greater interactivity might be more helpful for educators to observe and assess clinical reasoning skills, as students can demonstrate a broader range of skills in real-time, such as data gathering.

Effectiveness

Our results largely concur with previous reviews that have found virtual patient tools are better than no intervention but might not be superior to other methods of explicitly teaching clinical reasoning, such as problem-based learning tutorials[13–15]. The benefits to using virtual patient tools are that they can be used in circumstances when face-to-face teaching is not possible, e.g., due to a pandemic, or because access to patients is limited. Additionally, once upfront costs are covered, the cost of adapting and scaling up can be low. This review suggests that using virtual patient tools can effectively complement face-to-face teaching. It provides useful evidence for medical educators to guide their decisions about using this technology, which may be especially attractive if there is no other explicit teaching of clinical reasoning skills in the curriculum. Further research is needed to understand the context in which different teaching methods are most effective and the feasibility of implementing into curricula, so that medical educators can make more informed decisions on educational methods.

This review showed some evidence that effectiveness might depend on the domains of clinical reasoning that the virtual patient tools were designed to address and how these were measured. Most articles evaluated the effects of virtual patient tools on domains of data gathering, ideas about diagnoses and patient management and many showed significant improvement in these domains. Knowledge about clinical problems and processes, flexibility in thinking about diagnoses, problem-solving skills, and critical thinking were less commonly measured and showed less consistent improvement after virtual patient use. These findings could be due to issues with measuring different domains of clinical reasoning. Data gathering skills, ideas about diagnoses and patient management are domains that are related to students' judgements on specific cases. Therefore, they are easier to measure using patient cases and measures like the key feature problems, which are case specific and may be more sensitive to change immediately post intervention. In contrast, critical thinking measures may be more related to the underlying cognitive processes of clinical reasoning. These general cognitive skills are less likely to vary over the short-term and measurements, such as the DTI, have not necessarily been designed to be sensitive enough to detect short-term changes in these skills[50, 51]. Case specific outcomes may also be more appropriate for measuring clinical reasoning, as clinical reasoning is a skill that is context dependent[2]. We also found most articles reported aggregated effectiveness over several domains. Future research would benefit from defining the specific domains of clinical reasoning their virtual patient tool aims to improve and provide separate analyses for each aspect. Furthermore, a greater understanding of the psychometric properties of measures of clinical reasoning is needed to identify which domains of reasoning virtual patient tools can effectively teach students and over what timescales.

Limitations

It was not meaningful to conduct a meta-analysis to summarise the overall effectiveness of virtual patient tools on clinical reasoning due to the substantial heterogeneity in the design and content of the virtual patient tools, the measures of clinical reasoning and the characteristics of samples. Many articles developed their own measures of reasoning but with limited validation it was difficult to ascertain what they were measuring and how comparable they were to other measures. The findings of the review were limited by the lack of high-quality articles that were included. The review was updated in October 2020 and by this time the review authors' article on a virtual patient tool was eligible for inclusion. This was rated of high quality, and it is possible the authors were biased in their scoring of their own article. As found in previous reviews, most single group pre-test and post-test evaluations found significant benefits of using virtual patient tools and it is possible there was publication bias with negative findings being unpublished[14, 15]. The review was also limited by the small percentage of abstracts that were double screened for inclusion. However, the agreement between screeners was good and any discrepancies were discussed; abstracts where there was uncertainty of inclusion were included in the full text review to ensure we captured as many relevant articles as possible[52, 53].

Conclusion

Overall, the evidence suggests virtual patient tools could effectively complement current teaching and may be particularly useful if opportunities for face-to-face learning are limited. This research found that evaluations that measured clinical reasoning by measuring case specific domains of clinical reasoning, such as ideas about diagnoses or data gathering, showed more consistent improvement in reasoning than more general measures of reasoning, such as critical thinking. Case specific measures of clinical reasoning may be more sensitive to change following virtual patient cases because they reflect the context dependent nature of clinical reasoning skills. Future evaluations should provide evidence of the validity and reliability of their clinical reasoning outcome measures to aid the comparison of effectiveness between studies. More understanding is needed about how features of virtual patient design and delivery relate to effectiveness.

Declarations

Ethics approval and consent to participate.((Human, Animals, Plants and Source)-

Not applicable.

Consent for publication –

Not applicable.

Availability of data and materials-

The dataset supporting the conclusions of this article is included within this

article and its additional files.

Competing interests-

The authors declare no competing interests.

Funding-

RP was supported by The Health Foundation for her PhD when she undertook this research and is currently supported by the National Institute for Health Research (NIHR) School for Public Health Research (Grant Reference Number PD-SPH-2015). JS is supported by the National Institute for Health Research Applied Research Collaboration (ARC) North Thames. This research was supported by the National Institute for Health Research (NIHR)

Policy Research Programme, conducted through the Policy Research Unit in Cancer Awareness, Screening and Early Diagnosis, 106/0001. The views expressed in this article are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care. The funders had no role in the study design, data collection, analysis, interpretation of data or in writing the manuscript.

Authors' contributions-

RP planned the review and RP, JS, MK, APK and RR shaped the review questions. The literature search was conducted by RP with the assistance of a librarian. RP and APK selected suitable articles which met the inclusion criteria. RP, APK and SM extracted the data from the full text articles. RP, APK and SM critically appraised the articles. RP drafted the manuscript, JS, APK, MK, SM and RR helped revise the paper, contributing intellectual content/commented on drafts of the paper.

Acknowledgments-

The authors would like to acknowledge the University College London Library for their assistance with this literature search.

References

1. Cleland JA, Abe K, Rethans J-J: **The use of simulated patients in medical education: AMEE Guide No 42**. *Medical Teacher* 2009, **31**(6):477-486.
2. Higgs J, Jones MA, Loftus S, Christensen N: **Clinical Reasoning in the Health Professions**: Elsevier; 2008.
3. Page G, Matthan J, Silva A, McLaughlin D: **Mapping the delivery of ‘Clinical Reasoning’in UK undergraduate medical curricula**.
4. Medicine Io: **Improving Diagnosis in Health Care**. Washington, DC: The National Academies Press; 2015.
5. Cook DA, Triola MM: **Virtual patients: a critical literature review and proposed next steps**. *Med Educ* 2009, **43**(4):303-311.
6. Ericsson KA: **Deliberate practice and acquisition of expert performance: a general overview**. *Acad Emerg Med* 2008, **15**(11):988-994.
7. Kolb DA: **Experiential learning : experience as the source of learning and development**. London: Prentice Hall; 1984.
8. Bradley P: **The history of simulation in medical education and possible future directions**. *Medical Education* 2006, **40**(3):254-262.
9. AAMC: **Effective Use of Educational Technology in Medical Education Colloquium on Educational Technology: Recommendations and Guidelines for Medical Educators**. In. Washington, DC: Association of American Medical Colleges; 2007.
10. Barry Issenberg S, Mcgaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ: **Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review**. *Medical teacher* 2005, **27**(1):10-28.
11. Mann JA, Roland D: **What are the measures that can be used to assess performance during in situ Paediatric Emergency Medicine Simulation?** *BMJ Simul Technol Enhanc Learning* 2017.
12. McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ: **Revisiting ‘A critical review of simulation-based medical education research: 2003–2009’**. *Medical Education* 2016, **50**(10):986-991.
13. Consorti F, Mancuso R, Nocioni M, Piccolo A: **Efficacy of virtual patients in medical education: A meta-analysis of randomized studies**. *Computers & Education* 2012, **59**(3):1001-1008.
14. 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.
15. Cook DA, Levinson AJ, Garside S, Dupras D, Erwin P, Montori V: **Internet-based learning in the health professions**. *J Am Med Assoc* 2008, **300**(10):1181-1196.
16. Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, Erwin PJ, Hamstra SJ: **Technology-enhanced simulation for health professions education: a systematic review and meta-analysis**. *Jama* 2011, **306**(9):978-988.
17. Moher D, Liberati A, Tetzlaff J, Altman DG, The PG: **Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement**. *PLOS Medicine* 2009, **6**(7):e1000097.
18. Plackett RL: **Evaluation of an online learning tool to improve medical students' clinical reasoning skills**. UCL (University College London); 2019.
19. 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.
20. **The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses.** [http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp]
21. Page, Bordage, Allen: **Developing key-feature problems and examinations to assess clinical decision-making skills**. *Acad Med* 1995, **70**(3):194-201.
22. Kim S, Willett LR, Pan WJ, Afran J, Walker JA, Shea JA: **Impact of Required Versus Self-Directed Use of Virtual Patient Cases on Clerkship Performance: A Mixed-Methods Study**. *Academic medicine : journal of the Association of American Medical Colleges* 2018, **93**(5):742-749.
23. Sobocan M, Turk N, Dinevski D, Hojs R, Pecovnik Balon B: **Problem-based learning in internal medicine: virtual patients or paper-based problems?** *Internal medicine journal* 2017, **47**(1):99-103.

24. Chon SH, Timmermann F, Dratsch T, Schuelper N, Plum P, Berlth F, Datta RR, Schramm C, Hander S, Spath MR *et al*: **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).
25. Middeke A, Anders S, Schuelper M, Raupach T, Schuelper N: **Training of clinical reasoning with a Serious Game versus small-group problem-based learning: A prospective study**. *PloS one* 2018, **13**(9):e0203851.
26. Watari T, Tokuda Y, Owada M, Onigata K: **The utility of virtual patient simulations for clinical reasoning education**. *Int J Environ Res Public Health* 2020, **17**(15):1-9.
27. Schubach F, Goos M, Fabry G, Vach W, Boeker M: **Virtual patients in the acquisition of clinical reasoning skills: does presentation mode matter? A quasi-randomized controlled trial**. *BMC medical education* 2017, **17**(1):165.
28. Raupach T, Muenscher C, Anders S, Steinbach R, Pukrop T, Hege I, Tullius M: **Web-based collaborative training of clinical reasoning: a randomized trial**. *Medical Teacher* 2009, **31**(9):e431-437.
29. Plackett R, Kassianos AP, Kambouri M, Kay N, Mylan S, Hopwood J, Schartau P, Gray S, Timmis J, Bennett S *et al*: **Online patient simulation training to improve clinical reasoning: a feasibility randomised controlled trial**. *BMC medical education* 2020, **20**(1):245.
30. Kassier J: **Teaching clinical medicine by iterative hypothesis testing**. *The New England Journal of Medicine* 1983, **309**(15):921-923.
31. Isaza-Restrepo A, Gomez MT, Cifuentes G, Arguello A: **The virtual patient as a learning tool: a mixed quantitative qualitative study**. *BMC medical education* 2018, **18**(1):297.
32. Ludwig S, Schuelper N, Brown J, Anders S, Raupach T: **How can we teach medical students to choose wisely? A randomised controlled cross-over study of video-versus text-based case scenarios**. *Bmc Medicine* 2018, **16**.
33. Basu Roy R, McMahon GT: **Video-based cases disrupt deep critical thinking in problem-based learning**. *Medical Education* 2012, **46**(4):426-435.
34. McCoy L: **Virtual patient simulations for medical education: Increasing clinical reasoning skills through deliberate practice**. Arizona State University; 2014.
35. Kamin C, O'Sullivan P, Deterding R, Younger M: **A comparison of critical thinking in groups of third-year medical students in text, video, and virtual pbl case modalities**. *Acad Med* 2003, **78**(2):204-211.
36. Kahl K, Alte C, Sipos V, Kordon A, Hohagen F, Schweiger U: **A randomized study of iterative hypothesis testing in undergraduate psychiatric education**. *Acta Psychiatrica Scandinavica* 2010, **122**(4):334-338.
37. Devitt P, Palmer E: **Computers in medical education 1: Evaluation of a problem-orientated learning package**. *Australian and New Zealand Journal of Surgery* 1998, **68**(4):284-287.
38. Wu B, Wang M, Johnson JM, Grotzer TA: **Improving the learning of clinical reasoning through computer-based cognitive representation**. *Medical Education Online* 2014, **19**:25940.
39. Bordage G, Page G: **The key-features approach to assess clinical decisions: validity evidence to date**. *Advances in Health Sciences Education* 2018, **23**(5):1005-1036.
40. Lehmann R, Thiessen C, Frick B, Bosse HM, Nikendei C, Hoffmann GF, Tonshoff B, Huwendiek S: **Improving pediatric basic life support performance through blended learning with web-based virtual patients: Randomized controlled trial**. *Journal of Medical Internet Research* 2015, **17**(7):e162.
41. Charlin B, Roy L, Brailovsky C, Goulet F, van der Vleuten C: **The Script Concordance Test: A Tool to Assess the Reflective Clinician**. *Teaching and Learning in Medicine* 2000, **12**(4):189-195.
42. Kalet AL, Coady SH, Hopkins MA, Hochberg MS, Riles TS: **Preliminary evaluation of the Web Initiative for Surgical Education (WISE-MD)**. *American Journal of Surgery* 2007, **194**(1):89-93.
43. Bordage G, Grant J, Marsden P: **Quantitative assessment of diagnostic ability**. *Medical Education* 1990, **24**(5):413-425.
44. Aghili O, Khamseh ME, Taghavinia M, Malek M, Emami Z, Baradaran HR, Mafinejad MK: **Virtual patient simulation: Promotion of clinical reasoning abilities of medical students**. *Knowledge Management and E-Learning* 2012, **4**(4):518-527.
45. Kleinert R, Heiermann N, Plum PS, Wahba R, Chang D-H, Maus M, Chon S-H, Hoelscher AH, Stippel DL: **Web-based immersive virtual patient simulators: Positive effect on clinical reasoning in medical education**. *Journal of Medical Internet Research* 2015, **17**(11):1-11.
46. Botezatu M, Hult H, Tessma MK, Fors U: **Virtual patient simulation: Knowledge gain or knowledge loss?***Medical Teacher* 2010, **32**(7):562-568.
47. Kalet AL, Song HS, Sarpel U, Schwartz R, Brenner J, Ark TK, Plass J: **Just enough, but not too much interactivity leads to better clinical skills performance after a computer assisted learning module**. *Medical Teacher* 2012, **34**(10):833-839.
48. Mamede S, Schmidt HG: **The structure of reflective practice in medicine**. *Medical education* 2004, **38**(12):1302-1308.
49. Homer BD, Plass JL: **Level of interactivity and executive functions as predictors of learning in computer-based chemistry simulations**. *Computers in Human Behavior* 2014, **36**:365-375.
50. Scott JN, Markert RJ, Dunn MM: **Critical thinking: change during medical school and relationship to performance in clinical clerkships**. *Medical Education* 1998, **32**(1):14-18.
51. Niu L, Behar-Horenstein LS, Garvan CW: **Do instructional interventions influence college students' critical thinking skills? A meta-analysis**. *Educational Research Review* 2013, **9**:114-128.

52. McHugh ML: **Interrater reliability: the kappa statistic**. *Biochemia medica: Biochemia medica* 2012, **22**(3):276-282.

53. Viera AJ, Garrett JM: **Understanding interobserver agreement: the kappa statistic**. *Fam Med* 2005, **37**(5):360-363.

Figures

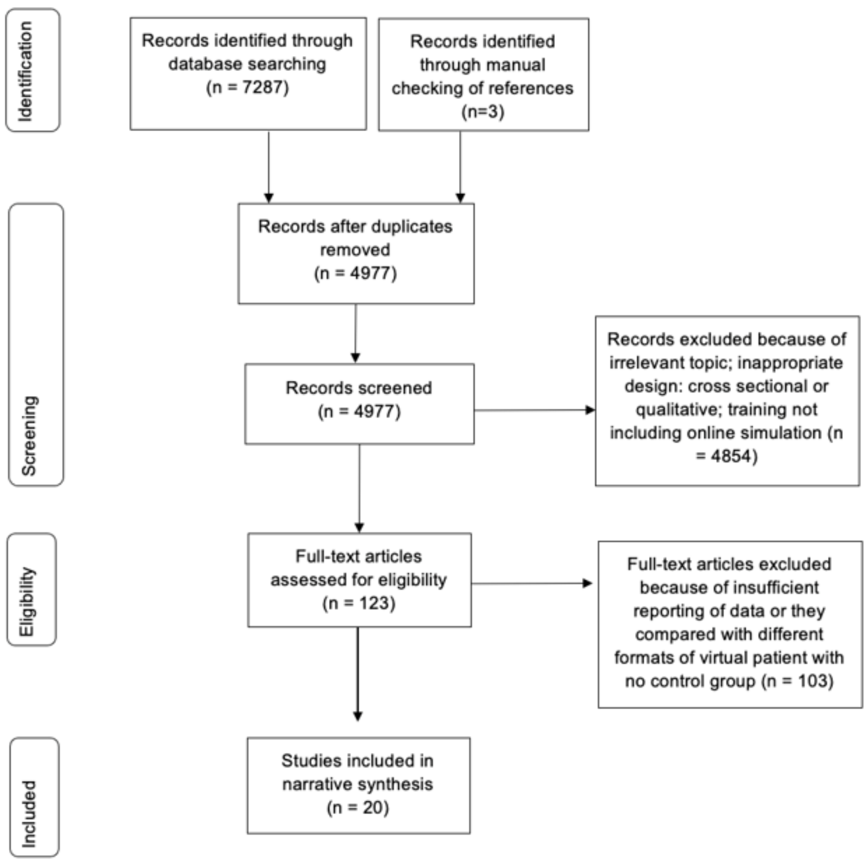
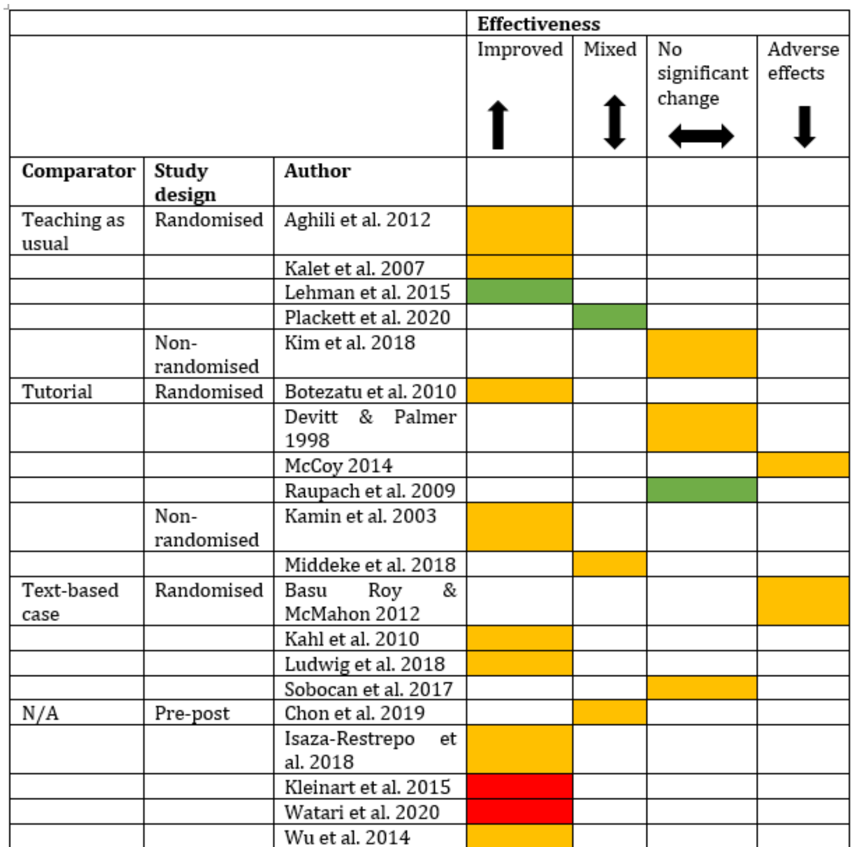


Figure 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension) flow chart for the article search



			Effectiveness			
			Improved ⬆	Mixed ⬍	No significant change ⬌	Adverse effects ⬇
Comparator	**Study design**	**Author**				
Teaching as usual	Randomised	Aghili et al. 2012				
		Kalet et al. 2007				
		Lehman et al. 2015				
		Plackett et al. 2020				
	Non-randomised	Kim et al. 2018				
Tutorial	Randomised	Botezatu et al. 2010				
		Devitt & Palmer 1998				
		McCoy 2014				
		Raupach et al. 2009				
	Non-randomised	Kamin et al. 2003				
		Middeke et al. 2018				
Text-based case	Randomised	Basu Roy & McMahon 2012				
		Kahl et al. 2010				
		Ludwig et al. 2018				
		Sobocan et al. 2017				
N/A	Pre-post	Chon et al. 2019				
		Isaza-Restrepo et al. 2018				
		Kleinart et al. 2015				
		Watari et al. 2020				
		Wu et al. 2014				

Red: low quality articles; Amber: moderate quality articles; Green: high quality articles.

Figure 2

Effectiveness of virtual patient tools by comparator and study design

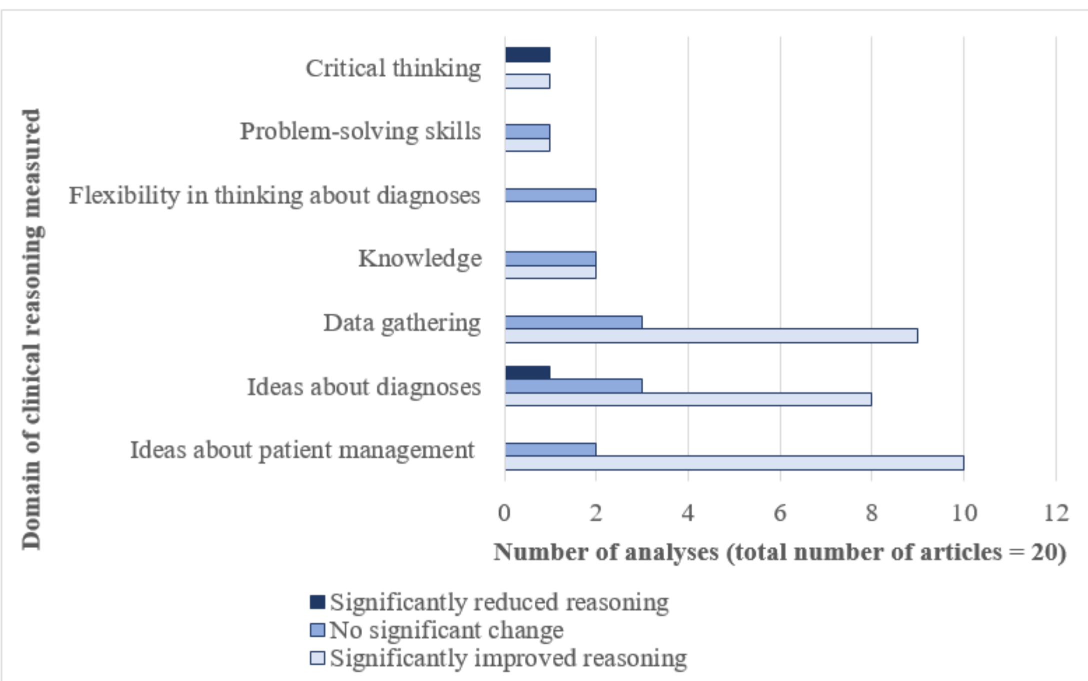


Figure 3

Frequency of analyses that reported different domains of clinical reasoning by effectiveness

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Adddionalfile1PRISMA2020checklist.docx
- Addiitonalfile2searchstrategy.docx
- Addiitonalfile3Quality.docx