Xie et al. BMC Medical Education (2025) 25:826
https://doi.org/10.1186/s12909-025-07187-7

BMC Medical Education

RESEARCH

Open Access

The role of three-dimensional printing models in medical education: a systematic review and meta-analysis of randomized controlled trials

Guanli Xie[1†], Tao Wang[2†], Han Fu[2], Duo Liu[2], Li Deng[2], Xia Zheng[2], Lvyu Li[2*] and Jianglong Liao[2*]

Abstract

Objective The objective of this study is to evaluate the role of three-dimensional printing models (3DPMs) in the medical education for undergrade students.

Method A comprehensive search was performed across three online databases including Medline, EMBASE, CINAHL, Web of SCI, and Scopus spanning from their inception to October 30, 2024. Studies that satisfied the predefined inclusion criteria were incorporated into the analysis. Data analysis was executed utilizing RevMan 5.4.1. Subgroup analyses were conducted based on various models, and overall effects were estimated using either the fixed effects model or the random effects model. The quality of evidence was evaluated using the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) framework.

Results A total of 33 studies were included in this study, involving 2716 medical undergraduates. The findings indicate that 3DPMs demonstrated significant advantages over the control group in theory test of the skeletal system with a moderate effect size (N = 646, $P < 0.00001$, $I^2 = 80\%$, SMD = 0.56, 95% CI 0.20—0.93, Random effect model). Moreover, 3DPMs showed a moderate effect size advantage over the control group in laboratory tests with moderate effect size (N = 299, $P < 0.00001$, $I^2 = 0\%$, SMD = 0.57, 95% CI 0.34 – 0.80, Fixed effect model). Additionally, 3DP showed advantage over the control group in total tests with small effect size (N = 832, $P = 0.20$, $I^2 = 84\%$, SMD = 0.26, 95% CI -0.14–0.66, Random effect model).

Conclusion 3DPMs serve as a valuable adjunct to traditional teaching methodologies and have the potential to enhance both the theoretical understanding and practical laboratory skills of medical students. Nevertheless, caution must be exercised in interpreting the current findings due to variations in model types, low quality of included studies, and the limited number of studies with small sample sizes.

Keywords Three-dimensional printing models, Medical education, Systematic review, Meta-analysis

[†]Guanli Xie and Tao Wang contributed equally to this work.

*Correspondence:
Lvyu Li
lilvyu@QQ.com
Jianglong Liao
liao13988902077@126.com
Full list of author information is available at the end of the article

Introduction

Undergraduate medical education is a complex and demanding field, characterized by a vast scope of knowledge, intricate concepts, and the need for practical skills development [1]. The current curriculum encompasses a wide range of disciplines, from basic sciences to clinical medicine, which necessitating a comprehensive understanding of human anatomy, physiology, pathology, and disease processes [2, 3]. Medical students face numerous challenges in navigating this extensive curriculum. The sheer volume of information can be overwhelming, and the complex interplay of physiological systems requires a deep understanding of relationships and interactions [3, 4]. Additionally, the transition from theoretical knowledge to practical application can be challenging, particularly when it comes to developing clinical skills and decision-making abilities [5]. While traditional teaching methods, such as lectures and textbooks, provide foundational knowledge, they may not fully engage students or facilitate a deep understanding of complex concepts. The static nature of textbooks and 2D images can make it difficult for students to visualize and conceptualize three-dimensional anatomical structures and relationships [6, 7]. Furthermore, the lack of hands-on experience can hinder the development of practical skills and clinical reasoning [8, 9]. These challenges underscore the need for innovative teaching strategies and tools that can enhance learning and facilitate the acquisition of essential knowledge and skills in medical education [10].

In response to the challenges faced in medical education, a variety of innovative technologies and methodologies have been developed and implemented, which has garnered significant attention. The exponential growth of 3DP technology has heralded a novel era in medical education, introducing an innovative teaching and learning resource [11–13]. It is widely utilized in the instruction of human anatomy, providing an interactive and visual learning experience [14], particularly beneficial for understanding rarer and more complex clinical conditions [15–18]. 3DP facilitates the production of models with high anatomical accuracy, thereby enhancing surgical training by offering realistic [19], hands-on experience [20, 21]. This hands-on approach can significantly improve the understanding and retention of complex concepts [22]. Three-dimensional printing (3DP) also allows for the creation of customized models that can be tailored to specific learning objectives or patient cases. This enables students to focus on specific areas of interest or explore unique pathologies, enhancing their learning experience. Manipulating 3DPMs can help students develop spatial reasoning skills and gain a better understanding of anatomical relationships. The 3DPMs enable students to operate independently. This not only significantly improves classroom interaction and students' learning motivation but also enhances teaching quality. Moreover, it provides an effective training method, thus making up for the shortcomings of traditional teaching, which often fails to offer students sufficient clinical practice opportunities [23]. Therefore, 3DP offers a promising solution to the challenges faced by undergraduate medical education. By providing a more effective and engaging learning experience, 3DPMs have the potential to enhance students' understanding and retention of knowledge, ultimately preparing them to become competent and compassionate healthcare professionals. Novel, 3D printed learning tools were favored by medical students and elicited positive responses regarding improved understanding and recognition of various anatomical structures [24].

A growing body of research has been focused on the role of 3DPMs in medical education, yet the findings remain inconsistent. The study reported in 2019 included 60 s-year medical undergraduates to compare the effects of 3DPMs and 3D images(3DIs) on ventricular anatomy learning, and found that both 3DPMs and 3DIs could improve learners' test scores, but there was no difference between them [25]. Two systematic reviews published in 2020 and one systematic review published in 2022 also focused on the educational applications of 3DPMs in medicine, but there remain some problems worthy of attention [26–28]. Firstly, it is of paramount importance to note that the relatively small sample size employed in the existing studies presents a significant limitation. Such a restricted sample may fail to adequately represent the extensive and diverse population of undergraduate medical students. This lack of representativeness inherently diminishes the statistical power and generalizability of the obtained results, casting doubts on the extent to which these findings can be extrapolated to the broader educational context. Secondly, a major concern lies in the methodological approach adopted by most of the included studies. The predominant use of self-reporting and satisfaction reports as the primary outcome measures is inherently flawed. These types of measures, while providing some insights, often lack the depth and objectivity required to comprehensively evaluate the educational impact. Thirdly, the heterogeneity among the studies cannot be overlooked. The inclusion of different types of research designs, such as cohort studies and randomized control trials, along with a diverse range of participants spanning medical students, residents, and surgeons, introduces significant variability. This diversity, rather than enriching the overall understanding, leads to a dilution of the results' effectiveness. The lack of consistency in study design and participant characteristics makes it arduous to draw definitive conclusions or make reliable comparisons across studies. Finally, since the publication of the aforementioned articles, there has been a prolific outpouring of research in this domain. The recent spate of publications, as referenced by [15, 23, 29–34], holds the potential to offer novel and valuable evidence regarding the application of 3DPMs in medical education. These new studies may address some of the existing shortcomings, introduce innovative methodologies, or present hitherto unexplored perspectives. In light of these considerations, our research endeavors to conduct a systematic evaluation of the role of 3DPMs in undergraduate medical education.

Consequently, the present study adopted a more refined approach by exclusively incorporating randomized controlled trials (RCTs) that specifically targeted undergraduate medical students as the research participants. This strategic selection aimed to enhance the internal validity and relevance of the study, as undergraduate medical students constitute a crucial and homogeneous subgroup within the medical education domain. Moreover, to address the issue of substantial heterogeneity that had marred previous research efforts, subgroup analysis was meticulously conducted. This analysis was based on the diverse 3DPMs employed across different systems in medical education, allowing for a more nuanced exploration of the impact and effectiveness of these models. By disentangling the effects associated with different model types, we sought to provide more precise and actionable evidence to inform educational practice. In addition, this study took advantage of the burgeoning research landscape in this area. It systematically updated and incorporated a multitude of newly published RCT studies, thereby capitalizing on the latest advancements and insights. These novel contributions not only augmented the existing body of knowledge but also infused fresh and compelling evidence regarding the application of 3DPMs in medical teaching. This comprehensive update ensured that our research remained at the forefront of the field, capable of offering the most relevant and up-to-date guidance for educators and practitioners alike.

Methods

Eligibility criteria

The inclusion criteria were based on the preestablished PICOS framework. 1) Participants: Studies involving only medical undergraduates were included. 2) Intervention: Studies that used 3DPMs as an intervention for the experimental group were included. 3) Comparison: Studies employing traditional teaching methods such as lectures, 2D images, 3D images, videos, and cadaver specimens, were included. 4) Outcome: Theoretical tests, clinical skills tests, learner satisfaction, accuracy, and learning time were considered as outcomes; 5) Study design: Only randomized controlled studies were included. In the case of a crossover design, only the data from the first phase was included.

Exclusion criteria

The exclusion criteria were as follows: 1) Studies involving non-undergraduate medical students, including medical residents, residents in standardized training, and surgeons, were excluded. 2) Intervention methods other than 3DPMs were excluded. 3) Those non-randomized controlled studies were also excluded.

Data sources

Five online databases, namely Medline (via PubMed), Embase (via Embase), CINAHL (via EBSCOhost), Web of Science and Scopus, were searched for all literature, regardless of language and publication region, from their inception to October 30, 2024. Other studies were screened by reading references that had been included in the study. A comprehensive search strategy was developed using medical subject headings (MeSH) and free-text terms related to 3DP, medical education, and study design. All search strategies for each database have been peer-reviewed by informatics experts and were shown in Appendix 1.

Study selection

All retrieved records were imported into the Document Manager (EndNote X7). Following the removal of duplicate entries, the titles and abstracts of literature records were screened by two authors (GLX and TW) based on the preestablished inclusion criteria of the PICOS framework independently in order to exclude clearly irrelevant records. Subsequently, the full texts of the remaining records were examined in detail to select those that met the inclusion criteria. All stages of this process were meticulously documented. If the event of divergence during the literature selection process, the final decision was reached through consultation with the third author (JLL). The collaborative methodology employed guaranteed a thorough and stringent literature selection procedure.

Data extraction

Data extraction was independently carried out by two authors (LD and FH) from included studies using a standardized data extraction form. The extracted data encompassed study characteristics (e.g., study design, sample size, and duration), participant characteristics (e.g., age, gender, learning experience, and pre—existing knowledge level), details of intervention (e.g., method, form, dosage and duration), details of comparison intervention, and outcome data (e.g., theoretical tests results, clinical skills tests scores, learner satisfaction level, accuracy rates, and learning time). Throughout the data extraction process, in the event that any missing data were identified, the original author was contacted via electronic email to retrieve the missing information.

Risk assessment for inclusion studies

The assessment of the risk of bias within the included studies was conducted by two independent authors (XZ and DL), employing the Medical Education Research Study Quality Instrument (MERSQI) [35] and the Risk of Bias in Systematic Reviews (ROB) tool [36]. The MERSQI was developed in 2007 within the context of a study investigating the relationship between funding and quality of medical education research. It was specifically designed to assess the methodological quality of experimental, quasi-experimental, and observational studies [37]. The content domains and specific items were developed from literature concerning study quality and then iteratively revised. The definitive version of the MERSQI encompasses ten items grouped into six domains, namely study design, sampling institutions and response rate, type of data, data analysis, validity, and outcomes [35, 38]. The ROB tool, which is prevalently utilized in systematic reviews of randomized controlled trials, evaluates the bias risk across six aspects via seven evaluative. Selection bias is assessed via random sequence generation and allocation concealment to ensure proper participant allocation and hidden assignment. Performance bias is examined by implementing blinding procedures for both participants and personnel involved in the interventions. Detection bias focuses on blinding outcome assessment. Attrition bias relates to incomplete outcome data resulting from participant dropouts. Reporting bias checks for selective reporting. Other sources of bias, encompassing flaws in the study design, excessively small sample sizes, and the conspicuous omission of blinding implementation, were taken into account. Each item is judged as "low risk of bias", "high risk of bias", or "unclear", a tripartite classification that is of paramount importance for evaluating the quality and reliability of research within the context of systematic reviews. Any discrepancies were reconciled through deliberation or consultation with a third reviewer (LYL).

Data analysis

All the data analyses were carried out using Revman 5.4. For continuous data, the mean or mean difference (MD) and standard, along with 95% confidence interval (CI), were used. In accordance with the Cochrane criteria, an absolute standardized mean difference (SMD) value of approximately 0.2 signifies a small effect size, a value around 0.5 denotes a moderate effect size, and a value of 0.8 or greater indicates a large effect size. Based on the results of the I^2 statistic and chi-square test, either the random effects model or the fixed effects model was selected to evaluate the effect size. If the I^2 statistic is greater than 50% and the result of chi-square test yielded a P less than 0.1, significant heterogeneity was considered to be present. The sources of heterogeneity were explored using subgroup analysis. The influence of heterogeneity on the results is discussed in the discussion section.

Subgroup analysis

During the research protocol design stage, an initial literature search was conducted to optimize the protocol. The consideration for subgroup analysis was primarily based on the following factors: 1. Different anatomical structures exhibit distinct characteristics, which might lead to varying teaching outcomes. 2. The complexity levels of different human anatomical systems vary, and this as well could result in diverse teaching effects. Therefore, subgroup analyses were performed based on different parts or structures printed in the included studies, such as the skeletal system, nervous system, cardiovascular system, visceral system, oral system, fractures, congenital malformations, and pathological conditions.

Sensitivity analysis

During the critical phases of this study, encompassing literature collection, data extraction, and particularly data analysis, sensitivity analysis was employed to ascertain the robustness of the research findings. In the data analysis phase, robustness was assessed by altering the model of the combined effect size, specifically through the transformation between the random effects model and the fixed effects model. Furthermore, Stata software was utilized to systematically exclude each included study, recalculating the effect size after each exclusion to evaluate the extent of change and assess the robustness of the results. The comprehensive procedures and methodologies employed are detailed in Appendix 1(Appendix Table 1).

Publish bias

According to the established guidelines, the results of a publication bias analysis are deemed reliable only when more than ten studies report the same outcome measure. In this investigation, only eight studies reported theoretical test results of the specified outcome measure. Consequently, an analysis of publication bias analysis was not conducted.

Evaluation of evidence quality

The Grading of Recommendations Assessment, Development, and Evaluation (GRADE) framework was used to evaluate the quality of the evidence. Evidence was classified into four quality grades: high, medium, low, and very low, according to five categorization factors, including risk of bias, imprecision, inconsistency, indirectness, and publication bias.

Results

Characteristic of studies of included

A total of 33 studies were included in the present study, involving 2716 medical undergraduates. All 33 articles included in the present study were initially part of the qualitative analysis. Of these, 11 articles reported outcome measures suitable for quantitative analysis and were therefore incorporated into the quantitative analysis as well. Conversely, the remaining 22 articles were limited to qualitative analysis due to the absence of data amenable to quantitative analysis. Attempts were made to contact the original authors of these 22 articles via email. By the time of submission, responses were received from ten authors, who confirmed that they had not collected the necessary data for quantitative analysis. The remaining 12 authors did not respond. The detailed literature screening process is illustrated in Fig. 1.

The basic characteristics of the included studies are shown in Table 1. Seven studies included first-year medical students [39–45] four studies included second-year medical students [25, 30, 46, 47], three studies included third-year students [32, 33, 48], five studies included fourth-year medical students [14, 23, 29, 31, 34], two studies included five-year medical students and interns [24, 49]. One study included 6-year students [50], one

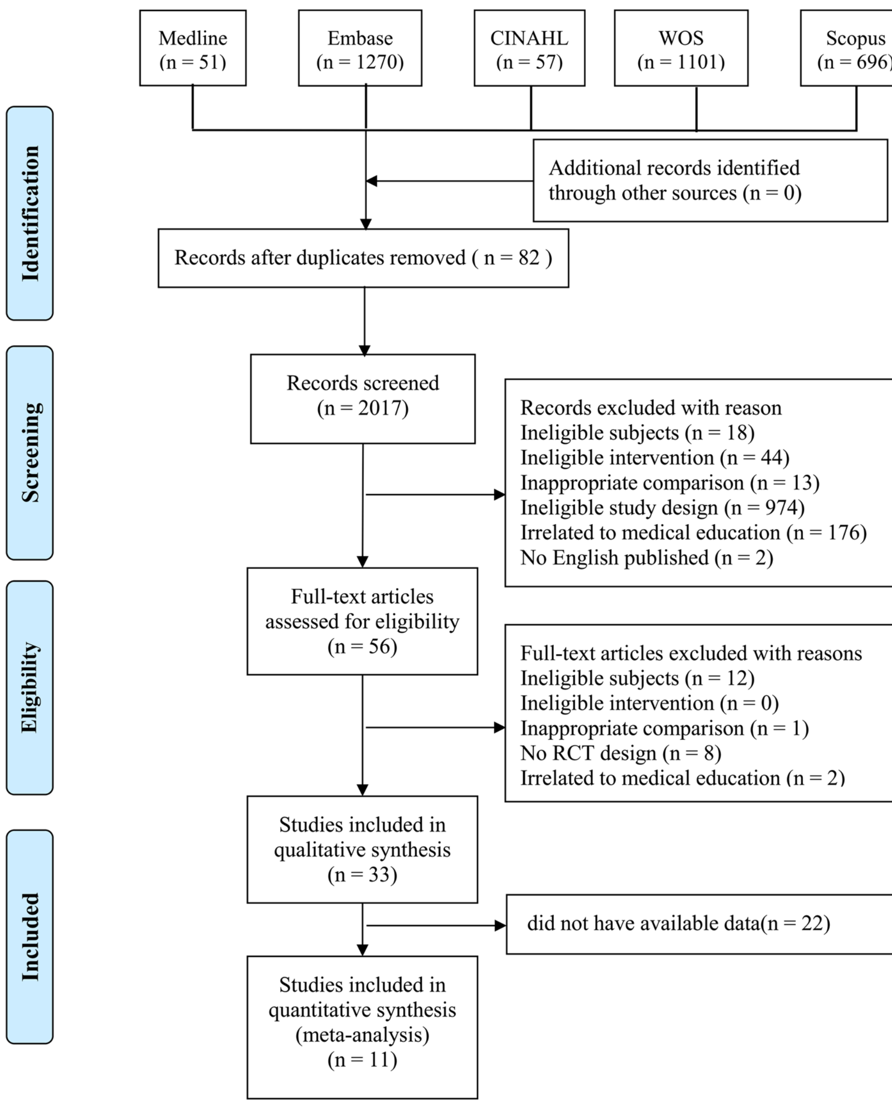

Fig. 1 Flow diagram of the study selection process

Table 1 Characteristics of included studies in this systematic review

Study ID	Participants/Number	Experimental intervention	Part and type of 3D printed model	Control Intervention	Outcomes and Instruments
AlAli 2018 [52]	Medical students/67	PPT + 3DPMs	CLP models	PPT presentation-based educational seminar	Objective tests
Al-Badri 2022 [50]	All 6 th-year undergraduate medical students/85	3DPMs	Craniosynostosis model	2DI	The long-term retention of learning information, satisfaction levels
Banwell 2021 [46]	Second year of the course in 2019 and 2020/27	3DPMs + 1 h training	Foot ulcer models, Callused foot model,	Standardised teaching + 1 h training session	Confidence and anxiety in second years who were using a scalpel for the first time, Competitive State Anxiety Inventory- 2, Task self-efficacy, Purpose-built questionnaire
Brumpt 2024 [33]	Third-year medical students/24	3DPMs	Ear models	Cadaveric model	Students' feelings, Experts' feelings, Experts' satisfaction
Cai 2019 [39]	First year medical students/35	3DPMs	Knee Joint	Didactic lecture	Objective tests
Chen 2017 [48]	Third-year medical students/79	3DPMs	Skull models	Group1: cadaveric skull Group2: the 2D atlas	Theory tests; Lab tests
Chen 2020 [24]	The interns/47	3DPMs	Henle trunk modes	2D image + surgical video	Test; Student Satisfaction;
Cheng 2020 [49]	Clinical interns (fifth grade)/124	2-month clinical education + oral teaching + 3DPMs from the 4 samples	Cerebral model	Oral teaching + imaging data (CT and MRI) + clinical pictures presentation	Self-assessment
Dalgalı 2024 [32]	Third-year dental students/117	3DPMs	CLP model	Group1: Traditional Education Group2: E-Learning-Supported Education	Examination
Fan 2019 [40]	first-year medical student/45	3DPMs	Bladder model	CT images	Training test 20 multiple-choice questions
Feng 2024 [31]	Fourth-year undergraduate/120	CBL + 3DPMs	Hip joint model	CBL	Theoretical knowledge score, Practical skill score, Ability to diagnose cases and plan treatment, Interesting teaching content, Willingness to communicate with lead teachers, Teaching satisfaction
Jiang 2022 [59]	Medical interns/269	3DPMs + PBL	Cerebrovascular disease model	Traditional education	Theoretical and clinical practice skills
Kiesel 2022 [8]	Medical students/84	3DPMs	Gynecological pelvic model	Theoretical introduction + skills training + directly commenced the practical training by working with the commercial models	Students' satisfaction, Gained knowledge
Kong 2016 [45]	medical freshmen who had completed the human anatomy course/61	3DPMs	Hepatic Segment model	Group1: traditional anatomical atlas Group2: 3DV	Objective tests

Table 1 (continued)

Study ID	Participants/Number	Experimental intervention	Part and type of 3D printed model	Control Intervention	Outcomes and Instruments
Láinez Ramos-Bossini 2024 [30]	Second university semester/40	10-h conventional theory program + 10-hour practice (using traditional 2D illustrations and interactive 3D models) + imaging examinations (CT and MRI) + 3DPMs	Spine model	10-h conventional theory program + 10-hour practice (using traditional 2D illustrations + ideal anatomical models)	Test
Lane 2020 [47]	second year medical student/44	3DPMs	Plastic surgery model (Model of sagittal synostosis)	PPT presentation	Quiz
Li 2015 [53]	Medical students/120	3DPMs	Spine fracture models	Group1: a lecture using PPT Group2: 3D image	Sum scores of correct answers and time spent, Overall time (s) required for answering the 10 questions, Answers to the evaluated questions
Lim 2016 [41]	first year medical students/52	3DPMs	External Cardiac Model	Cadaveric materials	Learning objectives tasks
Mogali 2022 [42]	Year 1 students/63	3DPMs	Cardiac model	Plastinated specimens	Objective test
Nicot 2022 [54]	undergraduate medical students/55	3DPMs	Craniofacial Trauma model	Two-dimensional	Multiple-choice question evaluation questionnaire
O'Brien 2021 [51]	Medical students from years one to three/31	Self-study + 3DPMs	Tracheobronchial tree model	Self-study + 2DI	Test
Ochoa 2019 [43]	first-year medical student/20	3DPMs	Cardiac Models	Traditional lecture	Written exam
Radzi 2022 [44]	Year 1 LKC Medicine MBBS students/96	3DPMs	Cardiac (one full and one cross-sectional heart)	Cardiac plastinated specimens	Survey
Su 2018 [55]	Medical students/68	3DPMs	Congenital heart disease Model	Traditional teaching	Objective test, Likert-type questionnaires
Wang 2017 [57]	Medical students/34	3DPMs	Cardiac Model	Traditional Model	Objective test, Satisfaction, understanding and preferred teaching methods
Wang 2020 [56]	Medical students/45	3DPMs	Tooth Model	Traditional teaching	Objective test
Wu 2018 [58]	Medical students/90	3DPMs	Bone spatial anatomy and fractures models	Traditional radiographic image	Test, Student Satisfaction
Yan M 2023 [14]	Fourth-year clinical medical students/100	3DPMs + X-rays + 2D CT + 3DI	Pelvic fracture model	3DPM + X-rays + 2D CT	Rate of accuracy; Overall time spent
Yan X 2023 [34]	Fourth-year clinical medical students/422	3DPMs + PBL + CBL	Respiratory diseases Models	Conventional teaching methodologies	Self-evaluation and satisfaction questionnaires
Yi 2019 [25]	Second-year medical students/60	3DPMs	Ventricular System Models	Group1: PPT + 3DI Group2: PPT + 2DI	Theory test, Practice test

Table 1 (continued)

Study ID	Participants/Number	Experimental intervention	Part and type of 3D printed model	Control Intervention	Outcomes and Instruments
Zeng 2024 [29]	Fourth-year undergraduate/28	Coloured 3DPMs	Tooth models	Plain tooth models	Clinical confidence, operability, a feel of the right preparations, overall learning process, the preparation quality
Zhang 2024 [23]	Fourth year undergraduate/104	3DPMs with pictures + X-rays + CT	Tibiofibular fracture	Class lecture + PPT	Knowledge, satisfaction
Zhao 2024 [15]	Senior undergraduate students/60	CBL + 3DPMs	Tetralogy of fallot Models	CBL (Case-based learning)	Theoretical exam scores

Abbreviations: *3DPMs* Three-dimensional printed Models, *PPT* Power Point, *CLP* Cleft lip and palate, *3DV* Three-dimensional Visualization, *CT* Computed tomography, *MRI* Magnetic resonance imaging, *3DI* Three-dimensional image, *2DI* Two-Three-dimensional, *CBL* Case-Based Learning, *PBL* Problem-Based Learning study included medical students from years 1 to 3 [51] and one study included senior undergraduate students [15]. Eight studies included medical students but did not report the grade of medical students [8, 52–58].

The intervention methods of the experimental group included 3DPMs alone and 3DPMs combined with other teaching methods, as follows: Twenty-one studies used 3DPMs alone as the intervention for the experimental group [8, 24, 25, 32, 33, 39–45, 47, 48, 50, 53–58], two studies used 3DPMs plus CBL [15, 31], one study used 3DPMs combined with CBL and PBL [34]. Other interventions included 3DP plus PPT [52], 3DP plus oral teaching [49], 3DPMs plus CT or MRI images [14, 23]. 3DPMs encompass a wide range of anatomical and pathological representations, including CLP models, craniosynostosis models, foot ulcer models, ear models, knee joint models, skull models, Henle trunk models, cerebral models, bladder models, hip joint models, cerebrovascular disease models, gynecological pelvic models, hepatic segment models, spine models, plastic surgery models, sagittal synostosis models, spine fracture models, pelvic fracture models, tracheobronchial tree models, cardiac models, tooth models, fracture models, respiratory system models, ventricular system models, and tetralogy of Fallot models. These models primarily pertain to the skeletal, cardiovascular, respiratory, plastic surgery, nervous, and oral systems. The model data of 10 studies were derived from the imaging data of real cases, of which 8 studies used the imaging data of patients to construct the model [14, 15, 24, 31, 49, 53, 57, 58], and 2 studies used the data of normal healthy people to construct the model [25, 45]. Of all the studies, only four studies reported expert evaluation of the 3DPMs before their use [25, 45, 49, 59]. Five studies reported the time cost of the 3DPMs, ranging from 4–6 h to 30 h [25, 30, 48, 52, 59]. One study reported that the processing time usually takes 3 to 7 days from the data extraction of CT or MRI scans to the model completion [49]. Five studies reported the material cost of the 3DPMs from 10 to 20 Euro or $14 to $281.61 [25, 30, 47, 48, 53], one of which mentioned that although the material was cheap, 3DP machines are expensive [43]. Two studies reported students' evaluation of 3DPMs [44, 47], including whether they should be recommended to others, whether it can improve academic performance and other evaluation content, and one study reported that polymer-based 3DPMs were deemed less life-like and less realistic, especially regarding the finer details —for example the margin of the fossa ovalis was not prominent in the cardiac 3DPMs compared to its plastinate [44].

The intervention methods of the control group included solely traditional teaching methods, traditional teaching methods combined with MRI or CT, cadaveric models, lectures, 2DIs, 3DIs, PBL, self-study, CBL, three-dimensional visualization(3DV), plastinated specimens, and E-Learning-Supported Education.

Outcome measures used included theoretical tests, clinical skills tests, learner satisfaction, accuracy, and learning time. The question types of theory tests mainly include single choice questions, multiple choice questions, short answer questions, noun interpretations, and medical record analysis and so on. Clinical skills tests (lab tests) included labeled structures to be recognized, history taking, medical record writing, imaging reading, choosing the surgical plan, and preparing an inlay. Total tests used in included studies included the theoretical score plus the practical score, and some studies also include the learner's satisfaction and learning confidence.

Among the included studies, only seven studies observed the follow-up effects of 3DPMs [32, 34, 43, 45, 46, 50, 51], which were followed up for 5 days, 2 weeks, 1 month, 6 weeks, and 3 months, while the other studies only observed the short-term effects of 3DP, that is, the immediate effects after the completion of learning.

The quality of the studies included

According to the MERSQI, the quality of the studies included was expressed by the actual score stacking plot (Fig. 2A, Table 2) and the percentage stacking plot (Fig. 2B). All studies included were designed using RCT. For the entry of sampling: institutions (Number of institutions refers to the origin of study participants), only two studies involved participants from two institutions [24, 52], considered as 1 point. The remaining 31 studies involved participants from a single institution. The proportion of those enrolled who completed the intervention evaluation was more than 75% in all studies included. The type of data in 29 studies was considered objective [8, 14, 24, 25, 29–34, 39–43, 45, 47, 48, 50, 51, 53–59], and 4 studies that used only participants' assessment data were not considered objective [44, 46, 49, 52]. Four studies did not measure a psychological construct and there was no instrument to rate (e.g., gender as the sole outcome), so the term "not applicable" was used to evaluate the "Validity evidence for evaluation instrument scores" [24, 39, 45, 51]. In terms of data analysis methods, only 1 study used only descriptive statistical reports [44], and the other studies all used statistical inference tests. All studies were considered as "Data analysis appropriate for study design and type of data". Five studies used satisfaction, attitudes, perceptions, opinions, general facts as outcomes [33, 44, 46, 49, 50] and 27 studies used knowledge/skills (paper, computer, simulation, or patients in a nonauthentic setting) as outcomes.

In accordance with the ROB tool, within the included studies, only 11 studies reported the methods of random

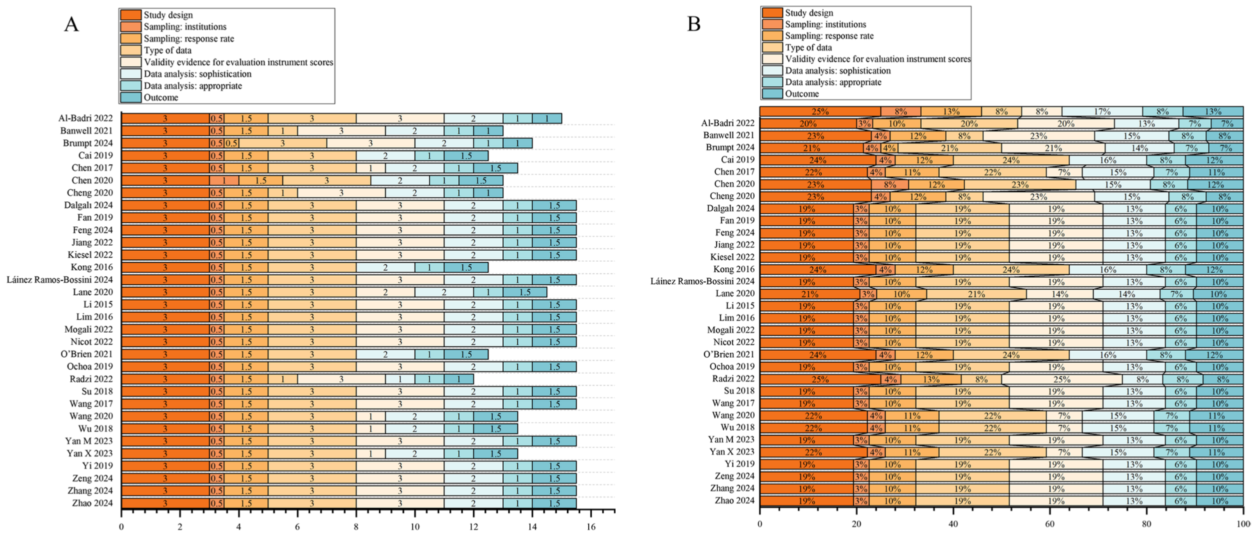

Fig. 2 **A** Medical Education Research Study Quality Instrument (MERSQI) scores for each study. **B** Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies

sequence generation [8, 23, 25, 29, 31, 42, 44, 53–55, 59]. These methods encompassed the generation of random numbers via computer algorithms, the use of random number tables, and the practice of drawing lots. Consequently, these studies were classified as presenting a low risk with respect to Random sequence generation. Regarding allocation concealment, only 2 studies provided relevant reports, and were therefore assigned a low risk in this entry [29, 48]. In terms of blinding of participants and personnel, only 5 studies reported subject and intervenor blindness in the study [34, 40, 42, 44, 48], which were considered as a low risk for this aspect. When addressing blinding of outcome assessment, only 2 studies reported the outcome assessment blindness [42, 48], which were considered as a low risk for the blinding of outcome assessors. Regarding the items lacking explicit reports, all of them were considered assigned a rating of "unclear". In the domain of other biases, the majority of studies were regarded as at high risk. This was mainly due to their small sample sizes and substandard methodological approaches. A comprehensive visualization of these detailed results can be found in Fig. 3.

The effective of 3DPMs for Theoretical tests

Eight studies reported the effectiveness of 3DPMs for theoretical tests [15, 25, 31, 42, 47, 48, 56, 58]. According to the different models of 3DP, a subgroup analysis was conducted based on different 3DPMs. The results showed that 3DPMs group had advantages over the control group in the learning of the anatomical structure and pathological structure of the skeletal system with moderate effect size (such as skull structure, spine, upper limb, lower limb, pelvis, hip joint) (N =646, $P<$ 0.00001, $I^2=$ 80%, SMD =0.56, 95% CI 0.20—0.93, random effects model), while there was no difference between 3DPMs and the control group in the learning of orthopedic models, cardiovascular structure and dental structures (Fig. 4A).

The effect 3DPMs for lab tests

Four studies reported the effectiveness of 3DPMs for lab tests [29, 31, 48, 56]. The results showed that 3DPMs group had advantages over the control group for lab tests with moderate effect size (N =299, $P<$ 0.00001, $I^2=$ 0%, SMD =0.57, 95% CI 0.34–0.80, fixed effects model, Fig. 5A).

The effect 3DP for total tests

Six studies reported the effectiveness of 3DPMs for total tests [24, 25, 31, 48, 55, 57]. The results showed that there was insufficient evidence to suggest that 3DPMs improved the performance of ungraduated medical students with a small effect size (N =832, $P=$ 0.20, $I^2=$ 84%, SMD =0.26, 95% CI − 0.14–0.66, random effects model, Fig. 6A).

Sensitivity analysis

The detailed procedures and measures implemented are presented in Appendix 1 Table 1. In the data analysis phase, the robustness of the findings was evaluated using Stata software (version 15.0) to reassess the effect size after the exclusion of individual studies. This procedure aimed to ascertain whether the exclusion resulted in significant alterations in the results. A significant alteration would imply that the excluded study exerted a

Table 2 Medical Education Research Study Quality Instrument (MERSQI) Scores for each study

Study ID	Study design	Sampling: institutions	Sampling: response rate	Type of data	Validity evidence for evaluation instrument scores	Data analysis: sophistication	Data analysis: appropriate	Outcome	Total
AlAli 2018 [52]	3	1	1.5	1	1	2	1	1.5	12
Al-Badri 2022 [50]	3	0.5	1.5	3	3	2	1	1	15
Banwell 2021 [46]	3	0.5	1.5	1	3	2	1	1	13
Brumpt 2024 [33]	3	0.5	0.5	3	3	2	1	1	14
Cai 2019 [39]	3	0.5	1.5	3	NA	2	1	1.5	12.5
Chen 2017 [48]	3	0.5	1.5	3	1	2	1	1.5	13.5
Chen 2020 [24]	3	1	1.5	3	NA	2	1	1.5	13
Cheng 2020 [49]	3	0.5	1.5	1	3	2	1	1	13
Dalgalı 2024 [32]	3	0.5	1.5	3	3	2	1	1.5	15.5
Fan 2019 [40]	3	0.5	1.5	3	3	2	1	1.5	15.5
Feng 2024 [31]	3	0.5	1.5	3	3	2	1	1.5	15.5
Jiang 2022 [59]	3	0.5	1.5	3	3	2	1	1.5	15.5
Kiesel 2022 [8]	3	0.5	1.5	3	3	2	1	1.5	15.5
Kong 2016 [45]	3	0.5	1.5	3	NA	2	1	1.5	12.5
Láinez Ramos-Bossini 2024 [30]	3	0.5	1.5	3	3	2	1	1.5	15.5
Lane 2020 [47]	3	0.5	1.5	3	2	2	1	1.5	14.5
Li 2015 [53]	3	0.5	1.5	3	3	2	1	1.5	15.5
Lim 2016 [41]	3	0.5	1.5	3	3	2	1	1.5	15.5
Mogali 2022 [42]	3	0.5	1.5	3	3	2	1	1.5	15.5
Nicot 2022 [54]	3	0.5	1.5	3	3	2	1	1.5	15.5
O'Brien 2021 [51]	3	0.5	1.5	3	NA	2	1	1.5	12.5
Ochoa 2019 [43]	3	0.5	1.5	3	3	2	1	1.5	15.5
Radzi 2022 [44]	3	0.5	1.5	1	3	1	1	1	12
Su 2018 [55]	3	0.5	1.5	3	3	2	1	1.5	15.5
Wang 2017 [57]	3	0.5	1.5	3	3	2	1	1.5	15.5
Wang 2020 [56]	3	0.5	1.5	3	1	2	1	1.5	13.5
Wu 2018 [58]	3	0.5	1.5	3	1	2	1	1.5	13.5
Yan M 2023 [14]	3	0.5	1.5	3	3	2	1	1.5	15.5
Yan X 2023 [34]	3	0.5	1.5	3	1	2	1	1.5	13.5
Yi 2019 [25]	3	0.5	1.5	3	3	2	1	1.5	15.5
Zeng 2024 [29]	3	0.5	1.5	3	3	2	1	1.5	15.5
Zhang 2024 [23]	3	0.5	1.5	3	3	2	1	1.5	15.5
Zhao 2024 [15]	3	0.5	1.5	3	3	2	1	1.5	15.5

considerable impact on the findings, whereas the absence of such change would indicate robustness. The analysis revealed that the exclusion of any single study did not result in significant alterations to the results, thereby confirming their robustness (Fig. 4B, Fig. 5B, Fig. 6B).

Publication bias

According to the guidelines, publication bias results are reliable only if more than ten studies report the same outcome measure. In present study, only eight studies reported a theoretical test of the outcome measure, so publication bias analysis was not performed. The research team concluded that there was publication bias, which was mainly due to the small number of included studies.

Quality level of evidence

Owing to the suboptimal quality of the studies included in the analysis, the evidence supporting the assertion that 3DPMs of orthopedic models may moderately improve the theoretical performance of medical students has been downgraded to moderate level. While 3DPMs exhibited

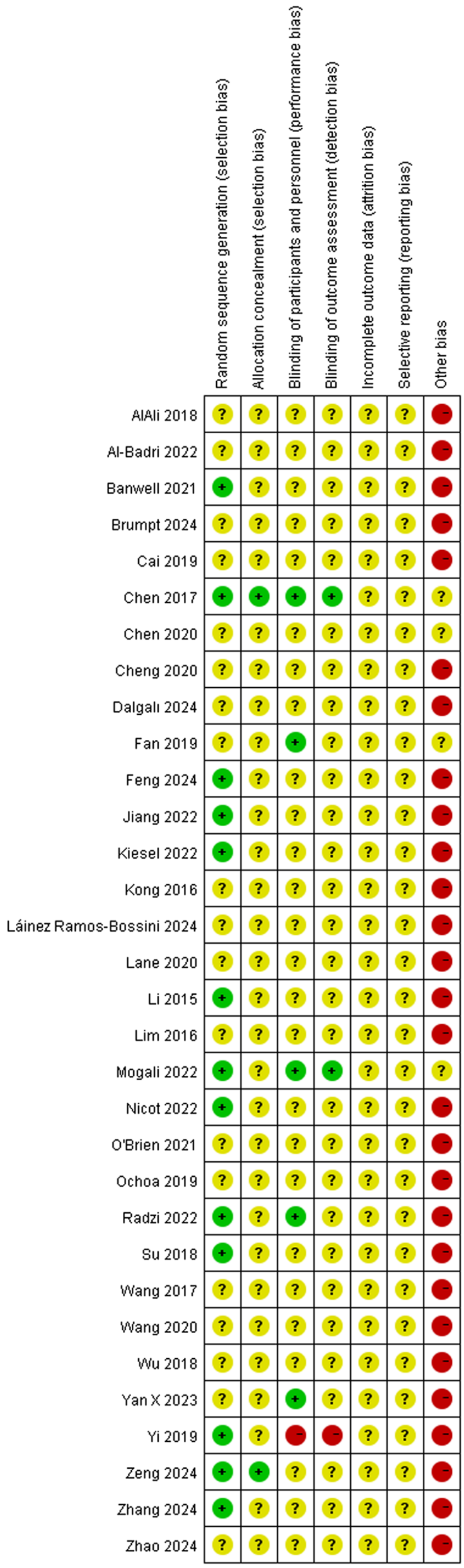

Fig. 3 Risk of bias summary: review authors' judgements about each risk of bias item for each included study

a moderate effect on enhancing both laboratory test and overall test performance among medical students, the overall quality of evidence was further downgraded by two levels to low. This downgrade is attributed to the poor quality of the included studies and the limited sample sizes.

Discussion

This study provides a systematic review and meta-analysis of the application of 3DPMs in medical student education, with the results showing that 3DPMs may promote the theoretical test performance and lab test performance of medical students with a moderate effect size. The results will provide theoretical basis, new ideas and directions for the application of 3DPMs in medical education, as well as directions for improvement for 3DP technology developers. However, due to the small number of studies included in quantitative analysis, their poor quality, and small sample size, caution should be exercised when interpreting the results.

The results of this study show that for theoretical performance, 3DPMs of the skeletal system can significantly improve learners' theoretical performance; however, there was considerable variability in the results. The reasons considered are as follows: First, among the included studies, a notable degree of heterogeneity was observed among the subjects. Given that students at different academic levels possess distinct knowledge reserves, this variation inevitably gives rise to heterogeneity in the research outcomes. The differences in prior knowledge can significantly influence how students engage with the study materials and respond to the interventions, thereby affecting the comparability and generalizability of the results. Second, there were disparities in the 3DP materials and printing accuracy leading to variations in the realism and precision of the models. These factors, combined with the differences in the control groups, contribute to the heterogeneity of the statistical results. The lack of standardization in these aspects makes it challenging to draw consistent and reliable conclusions from the collective body of research. Finally, different studies adopted diverse outcome indicators to assess students' academic performance. Some studies utilized multiple-choice questions, while others relied on single-choice questions or open-ended questions. Each type of assessment tool has its own strengths and limitations, and the choice of outcome indicator can significantly affect the measurement of students' learning outcomes. This divergence in outcome indicators also results in heterogeneity in the research results, as different assessment methods may capture different aspects of students' knowledge and skills.

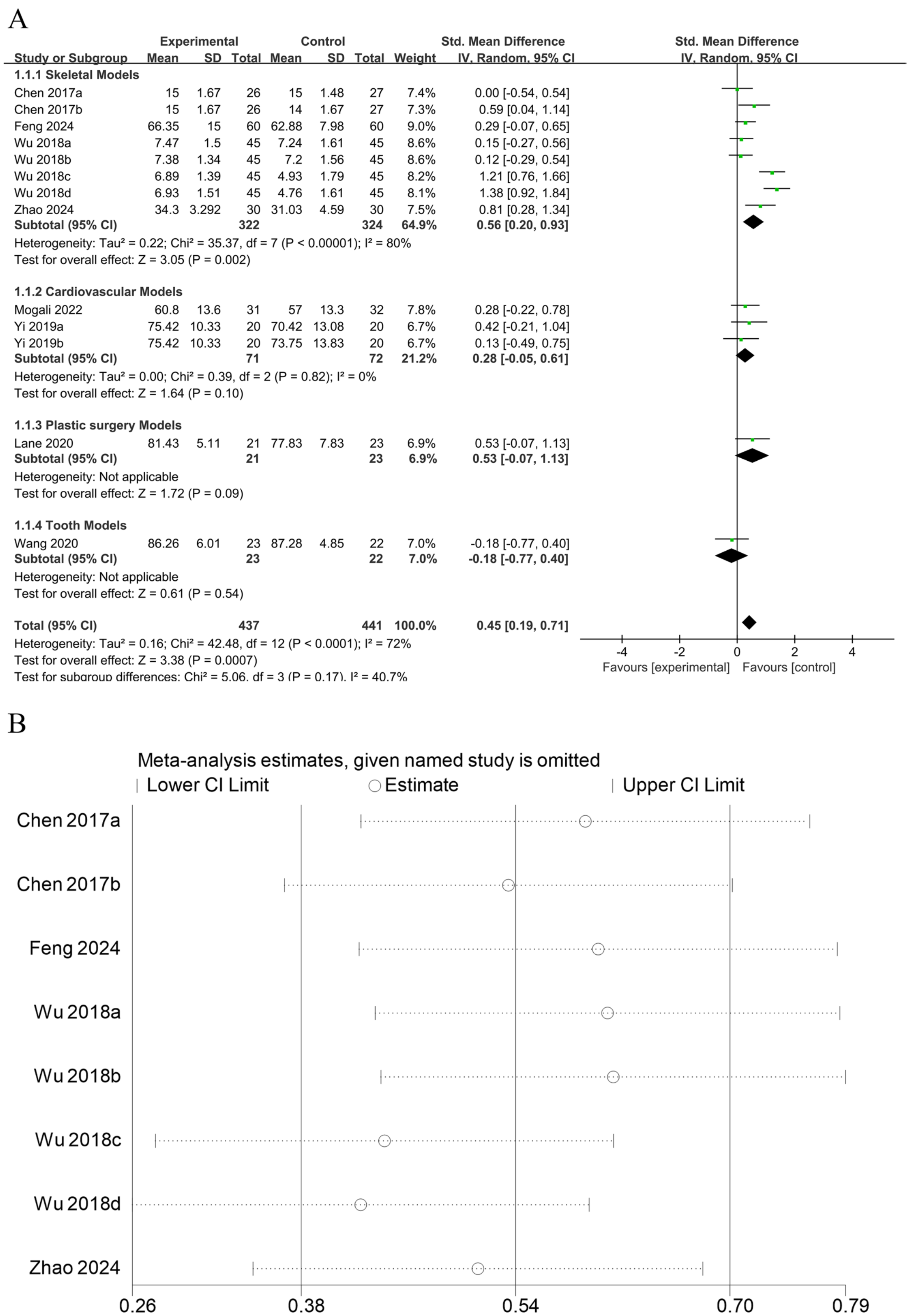

Fig. 4 **A** Meta-analysis of theory test of 3DPMs compared with control and **B** Sensitivity analysis of theory test of 3DPMs compared with control for Skeletal Models

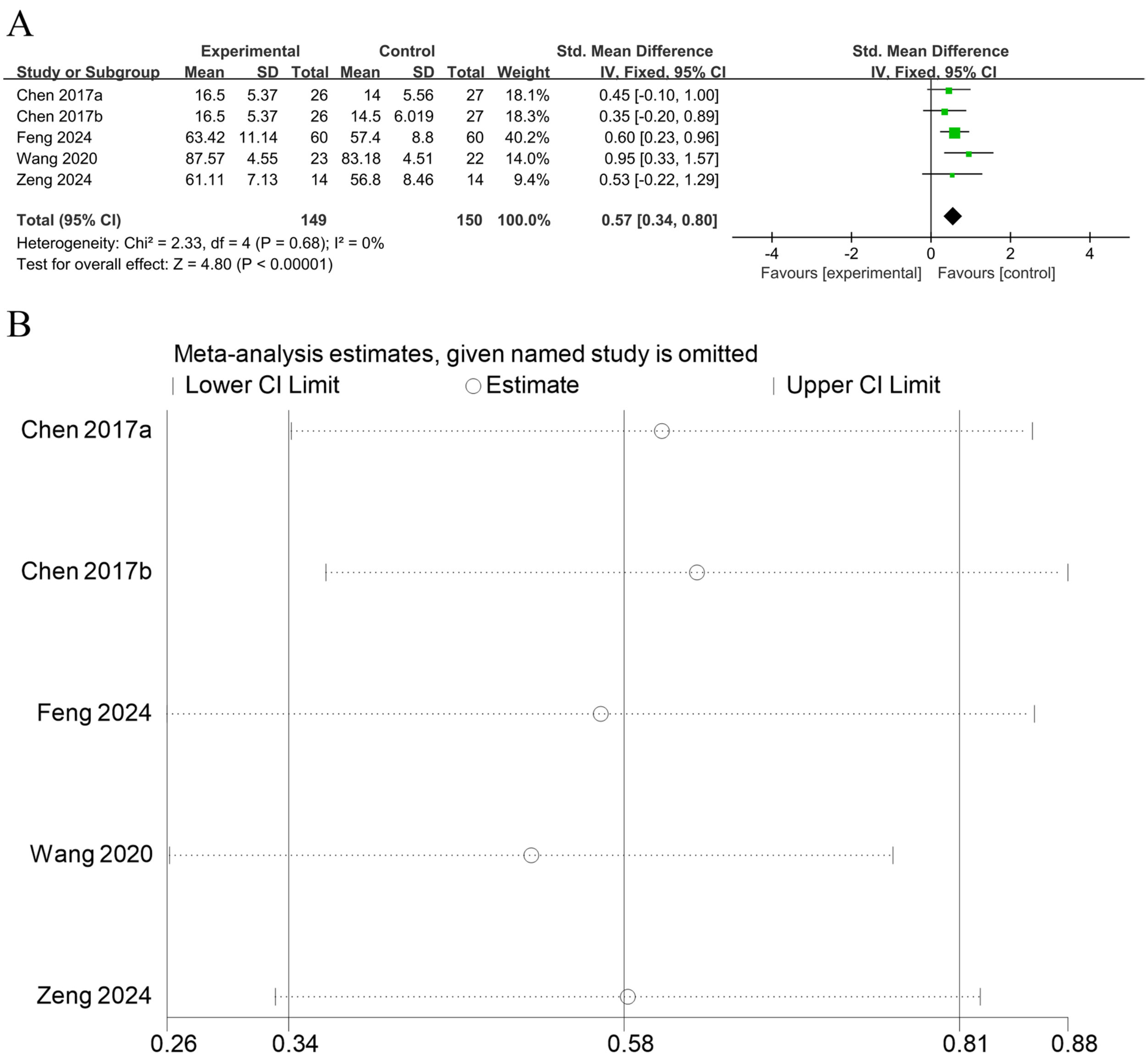

Fig. 5 **A** Meta-analysis of lab test of 3DPMs compared with control and **B** Sensitivity analysis of lab test of 3DPMs compared with control

Nevertheless, there is a lack of evidence indicating that cardiovascular, plastic, and dental 3DPMs significantly enhance learners' academic performance. The reasons considered are as following: First, four studies used bone models with a total sample size of 646, while the above models were only used in one or two studies. Therefore, the small sample size may be the reason for the negative results. Second, structural complexity of different 3DPMs between studies may also lead to different results [60]. Prior research indicates that anatomical complexity can modify the effectiveness of educational tools on learning and performance outcomes [60–62]. In this study, the 3DPMs used were diverse and varied in their complexity (bone models, heart models, plastic surgery models and tooth models), which may also have led to inconsistent results. It was reported that 3DPMs of the upper and lower limbs did not significantly improve learners' performance compared to the control group, while 3DPMs of the pelvis and spine significantly improved learners' performance on theory tests [58]. Furthermore, Studies have shown that for medical students, the learning difficulty of normal human anatomy is much lower than the pathological state of clinical patients [15, 53, 58]. In the results of this quantitative analysis, only 8 studies used medical cases as the data source for 3DPMs, which may may have results in insufficient power of the results.

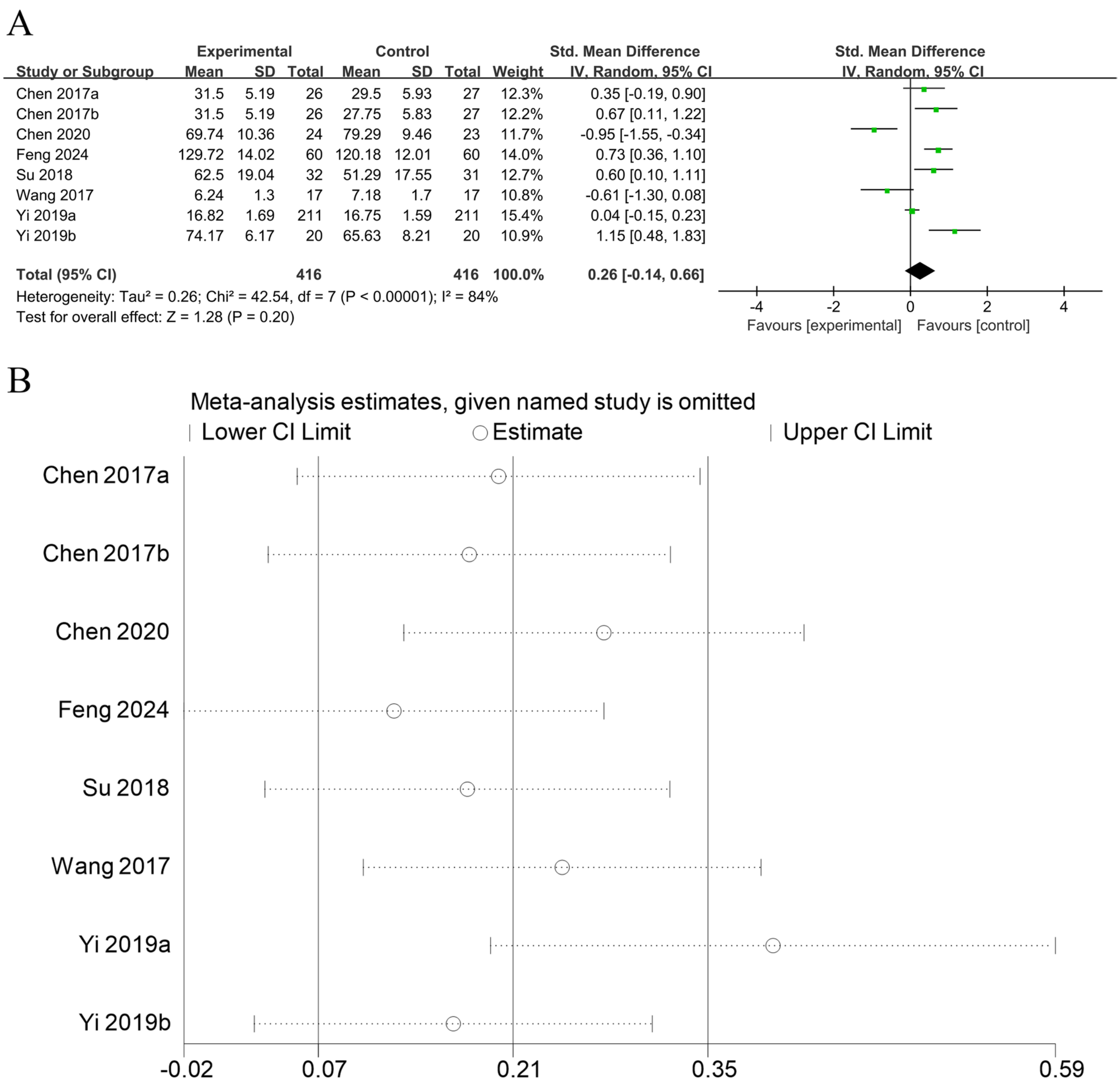

Fig. 6 **A** Meta-analysis of total test of 3DPMs compared with control **B** Sensitivity analysis of total test of 3DPMs compared with control

Finally, the situation of different control groups may also lead to this result. As the previous research results show, there is no difference between 3DPMs and cadaveric skull models, 2D atlases or 2DIs, but there are obvious differences between 3DPMs and lectures. In the results of this quantitative analysis, the control intervention included cadaveric skull, the 2D atlases, CBL, Traditional radiographic image, PPT +3DIs, plastinated specimens, traditional teaching, and PPT, which led to inconsistent results.

The results of this study show that 3DPMs may improve the lab test of medical undergraduates with moderate effect size, but there is no evidence that 3DPMs can significantly improve the total test. The inconsistency in results, as we considered, is related to the diversity of total scores and lab test assessment methods. In the studies that included quantitative analysis, the assessment methods of lab scores and total scores are diversified. In terms of the assessment of lab scores, they include marking bone marks, inlay preparations, and completing test items. In the evaluation of the total score, some studies add the theoretical scores and practical scores to get the total score, while some studies not only calculate the theoretical scores and practical scores but also include the learner's satisfaction and the evaluation of the model into the total score. This may lead to the inconsistency of results due to the diversity of forms and contents of score evaluation, and also cause the inconsistency of total results and lab score. Combined with the results of meta-analysis, the P-value and SMD of the 3DPMs group were 0.20 and 0.26 respectively compared with the control group in terms of total scores. There is no evidence that the intervention group is actually better than the control group; therefore, this result needs to be treated with caution, and further research may be needed to confirm the stability and authenticity of this small effect size. For example, increasing the sample size, improving the study design, or conducting subgroup analyses could be done to more accurately assess the effect of the intervention.

The sensitivity analysis showed that after each study was sequentially deleted, there was no impact on the final results. Therefore, the heterogeneity between studies mainly arises from the subjects, intervention methods, controls, outcome indicators and other aspects. First of all, in the studies included in quantitative analysis, although all included medical students, there was obvious heterogeneity among the studies due to the differences in the knowledge reserve of included learners and the educational level in different regions. Secondly, the use of different 3DPMs in the intervention group, different 3D printing techniques and different resolutions also contributed to heterogeneity between studies. Moreover, the intervention methods of the control group were also diverse, such as 3DIs, 2DIs, and cadaver specimens, which also resulted in the existence of heterogeneity. Finally, the studies included in quantitative analysis also have different testing methods for outcomes. From the form of test questions, some only have single choice, while others include single choice, multiple choice, short answer and medical record analysis questions, which were also an important source of heterogeneity.

When using 3DPMs, the accuracy of data sources (such as resolution, layer thickness of CT or MRI images), printing materials, and whether there are clear marks will all affect the effect of 3DPMs on academic performance. In this study, the resolution and layer thickness of pictures are quite different, and the printing materials are also different. This results in less consistency between studies, which may affect the strength of the data results. Learners' basic knowledge and early knowledge reserve will affect the teaching effect. In this study, only 5 studies took some measures to ensure a similar baseline level of knowledge on the subject [25, 31, 39, 42, 52], such as excluding subjects with scores greater than 50% through pre-test [52], only students with no prior knee-related anatomical knowledge at the time of recruitment were selected for the study [39].

Among the included studies, only seven studies had follow-up ranging from 5 days to three months [32, 34, 43, 45, 46, 50, 51]. The findings from these studies are inconsistent. Some studies suggest that 3DPMs enhances short-term learning outcomes compared to traditional teaching methods, although its long-term effects have yet to be confirmed [32, 43]. Conversely, other studies indicate that 3DPMs can improve both short- and long-term learning outcomes [46, 50]. Nonetheless, a consistent finding across these studies is that 3DPMs positively impacts learners' cognitive anxiety and confidence in learning.

There are many confounding factors in the learning process, such as learning motivation, previous learning experience, curiosity about new things, etc. In one study, baseline data were collected on students' previous exposure to 3DPMs and through the assessment of students' spatial representation skills through a mental rotation test [50]. The influence of learners' motivation on learning results should not be ignored. It is believed that students who can actively participate in this kind of research project have a relatively positive learning attitude, and when they are included in the research, they will cherish the learning opportunity [25, 48], which also affects the validity of the results. Besides, anatomical complexity may be a key confounder of learning and performance outcomes, as Wu's research shows that no significant differences were found in the upper limb or lower limb test scores between the 3DPMs group and the traditional radiographic image group; however, the scores on the pelvis and spine test for the traditional radiographic image group were significantly lower than the those for the 3DPMs group [63].

In addition to improving the quality of teaching, we are also concerned about the cost of equipment, materials and time. In this study, only five studies reported the cost of materials and time, and no study reported the cost of equipment. In the studies included in this project, the processing time of 3DPMs was reported to be the as short as 4–5 h and as long as 3–7 days, while the material cost was the as lower as $14 and as high as $281.61. It was previously reported that the 3DPMs was not time-effective and the costs of specialized printing machine and the bioinks were still too high to be affordable, especially for the underdeveloped areas [64, 65]. This also limits the application of 3DPMs in medical education to a certain extent.

It should be emphasized that successful 3DPMs require careful design and conception, especially models that can be assembled or models that add biomechanical elements. Successful design comes from years of clinical work, the accumulation of teaching experience, as well as repeated speculation and reflection. At the same time, the successful production of the model also needs the assistance and guidance of professional and technical personnel. Therefore, we should make it clear that 3DPMs are only an auxiliary teaching means; Its successful application can significantly improve the teaching effect, but we cannot ignore the advantages of traditional teaching.

Strengths and limitations

This study was conducted in strict accordance with the guidelines to ensure the reliability of the results. Firstly, the retrieval strategy was adjusted several times by pre-inspection to ensure the comprehensiveness and accuracy of literature retrieval. Secondly, in the process of literature screening, quality assessment and data extraction, two researchers independently completed the process and consulted with the third experienced researcher to obtain a consensus result. Thirdly, in the process of data analysis, subgroup analysis, sensitivity analysis and other methods were used to explore the factors affecting the results and the sources of heterogeneity, so as to ensure the reliability and validity of the results.

In addition to the above advantages, there are still some problems that need our consideration in this study. First of all, although the subjects included in this study are all medical students, the differences in medical education mode, medical education level and curriculum in each country and region could have affected the stability, validation, and generalization of the final results. Secondly, the different 3DP equipment used in the study, the different materials, the different printing parts, and the different precision of the data images led to the great heterogeneity among the 3DPMs. Also, the printing model and image data did not include analysis of muscles and neurovascular tissue surrounding the model, without which the model cannot completely represent a pathological condition. The precision of 3DPMs will affect the learning effect of learners. In this study, only 2 studies evaluated 3DPMs, which might lead to the existence of heterogeneity. Thirdly, different control teaching methods, including 2DIs, 3DIs, 3D visualization, cadaver specimens, lectures, etc., also led to the existence of outcome heterogeneity. Moreover, although all the studies included in this study were randomized controlled studies, only two studies reported blind methods, and only 2 studies were multi-center studies, resulting in the included studies being mostly identified as high risk in the assessment of risk bias, which affected the reliability of the results. In addition, it should be noted that among the included literatures, only a small number of articles reported the data we desired for quantitative research. By contacting the original authors, we did not obtain all the data required for the quantitative analysis of this study. Therefore, the power of the results of this study needs to be further verified. Finally, it is imperative to emphasize that in the original studies included in this research, most of the outcome evaluation indicators are either self-reported or subjectively judged by the evaluators. The inconsistent cognitive differences and evaluation criteria among different individuals may also lead to some deviations in the results, thereby affecting the accurate assessment of the effects of 3DPMs.

Future research direction

In future research, when using the 3DPMs for medical education, we may pay attention to the following aspects. Firstly, in the practical application context, the time and material costs are the top concerns for medical educators. Nevertheless, only five out of the incorporated studies have addressed the time cost. Precisely, the time span required varies remarkably, stretching from a minimum of 4–6 h to a maximum of 30 h. According to one of the studies, commencing from the acquisition of CT or MRI scan data and concluding with the finalization of the 3DPM suitable for medical instruction, the entire processing pipeline generally demands 3 to 7 days. Regarding the material cost, four studies likewise furnished pertinent information. The expenditure on materials for 3DPMs lies within the interval of 10 to 20 euros or 14 to 281.61 US dollars. Notably, one study accentuated that while the direct material cost for model printing is relatively moderate, the financial outlay for procuring 3DP machines is considerably high, imposing an undeniable economic strain on medical education institutions intending to adopt this technology. Consequently, due to the cost of time and materials, future research can be optimized from the following aspects. On the one hand, the production process of 3DPMs dedicated to medical education can be deeply optimized to minimize the time spent in the process from medical image data extraction to model production, thereby effectively improving the efficiency of teaching preparation and ensuring that 3DPMs can serve medical teaching practice in a more time saving manner. Secondly, make every effort to explore the material replacement strategy that meets the needs of medical education and is more cost-effective. On the basis of fully guaranteeing the quality requirements such as the accuracy and fidelity of 3DPMs, effectively reduce the material procurement cost and the comprehensive cost of equipment acquisition and maintenance, and open up an economic and feasible way for the popularization of 3DP technology in medical education; Thirdly, systematically and deeply explore the impact of different time input and material cost input modes on the final teaching performance of 3DPMs in medical education scenarios, such as the effect of the model on assisting students to understand complex anatomical structures, improving practical operation skills, and actual teaching application effectiveness. Thus, it can provide an accurate and scientific cost–benefit reference basis for medical educators to make optimal decisions when using 3DPMs.

Develop and utilize materials that more closely mimic the mechanical properties of human tissues, such as elasticity, flexibility, and haptic feedback. These materials will make the models more realistic and enhance the learning experience. Explore the use of multi-material 3DP techniques to create models with varying properties, allowing for the representation of different tissues within a single model. The tactile and visual advantages provided by these models allow students to have physical access to anatomical structures, greatly improving their understanding of complex theoretical knowledge [19]. Future advances in implementing 3DP in medical education could include the development of printing devices that allow rapid onsite printing in the teaching hospital and the development of 3DPMs that mimic the haptic characteristics of specific tissue (i.e., nerves, arteries, muscles) [53] It was reported that sight and touch are linked in a cross-modal arrangement in the somatosensory cortices, suggesting that they are mutually enhancing [66],which can improve the learner' performance scores. Furthermore, the integration of these haptic models into a progressive clinical curriculum is likely to enhance comprehension by contextualizing haptic-visual data and facilitating the precise exploration of specific competencies that students are expected to acquire, such as biomechanical concepts [54]. Designing models with interactive elements, such as removable parts or adjustable components, enables exploration and manipulation, thereby potentially increasing engagement and understanding of complex anatomical relationships. It is very important that when applying 3DPMs, their precision should be evaluated so as not to affect the learning performance of learners.

Case-based models such as fracture model, congenital malformation model, vascular disease model, especially for rarer and more complex situations, use cases as a basis and put empty theories into the context of specific cases for exposition. It has been reported that most cadaver specimens and plasticized models have normal anatomy and cannot be used for the study of pathological conditions; however, the use of 3DPMs offers the possibility of contextualizing the course by selecting the desired pathological structure [41]. 3DPMs can be obtained from interesting cases encountered in clinical practice, which can provide diverse pathological changes and strengthen the connection between learners' basic knowledge and clinical practice [67].

This ensures that the models are tailored to the needs of the students and the curriculum. Students' engagement and interest can be increased through the combination of visual and tactile use of 3DPMs and dynamic learning experience in the CBL teaching method. At the same time, instructors can introduce active learning and critical thinking into the traditional classroom, guiding students to analyse problems and explore solutions in a more holistic manner and improving overall teaching and learning outcomes [19, 31].

The confounding factors of learning effect cannot be ignored. In previous studies, students applied and were randomly selected for inclusion in the study. The selected students may cherish this opportunity very much and study hard. When students in the control group realized that they weren't selected to use the 3DPMs, they were likely to work harder to get better scores [25, 48], which leads to a significant reduction in the difference between the two groups. Therefore, when evaluating the role and effect of 3DP in medical education, the influence of learning motivation on the outcome should be considered. Novel interventions usually arouse participants' curiosity and lead to better results [68]. In addition, prior exposure to 3DP printing should be considered in the evaluation of medical education outcomes. In view of the impact of different cognitive loads and anatomical complexity on the teaching effect of 3D printing [61–63]. Future studies will work to identify the impact of 3DP on anatomy learning of a variety of anatomical regions in an effort to identify the ideal niche to maximize learning with 3D print exposure. Additionally, it was also important to note that 3DP involved complex processes and some structures may be lost during printing or post-printing procedures such as removing and cleaning the support materials. Hence, the structural variations and extent of damaged structures in the cadaveric and 3DP materials were noteworthy to consider while planning the trials [48].

Evaluating the effectiveness of 3DPMs in medical education requires a comprehensive approach that considers various aspects of learning and engagement. The Pre- and Post-Testing of the 3DPMs to measure changes in knowledge acquisition and retention. This can be done using written exams, multiple-choice questions, or practical assessments. Evaluate students' practical skills, such as surgical techniques or diagnostic procedures, using 3DPMs. This can be done through simulations, observations, or assessments of model modifications. Case Studies and Problem-Solving Exercises: Assign case studies or problem-solving exercises that require the application of knowledge gained from using 3DPMs. This can assess the ability to transfer knowledge to real-world scenarios. Collect feedback from students through surveys or interviews to assess their engagement, motivation, and satisfaction with the use of 3DPMs. This can provide valuable insights into the impact of the models on the learning experience. Compare the effectiveness of 3DPMs with other teaching methods, such as lectures, textbooks, or computer simulations. This can help determine the unique advantages and limitations of 3DPMs. Conduct follow-up surveys with students after they have completed the course to assess the long-term impact of 3DPMs on their learning and retention of knowledge. Assessment of the potential learning curve of 3D model-based learning would require a longitudinal study of 3D use over time compared to 2D traditional learning [51]. When outcome evaluations are conducted, objective outcome indicators should be strived to be used. Alternatively, when subjective outcome indicators are employed, strict and unified evaluation criteria should be established so that the influence of subjective factors can be minimized and the reliability of the research findings can be enhanced.

Conclusion

3DPMs serve as a valuable adjunct to traditional teaching methodologies and have the potential to enhance both the theoretical understanding and practical laboratory skills of medical students. Nevertheless, caution must be exercised in interpreting the current findings due to variations in model types, low quality of included studies, and the limited number of studies with small sample sizes. To establish more robust evidence regarding the impact of 3DPMs on medical education, there is a need for well-designed, large-scale, multicenter RCTs with adequate follow-up periods.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12909-025-07187-7.

Additional file 1.

Acknowledgements

This study was supported by Health research project of Kunming Health Commission (2023-10-01-001) and the Yunnan Province High-level Chinese medicine talents project (2024).

Authors' contributions

J.L.L. and L.Y.L. conceived and designed the study; G.L.X. and T.W. and made the manuscript preparation and wrote the manuscript; D.L. and H.F. extracted and analyzed data. L.D. and X.Z. evaluated the quality of the included studies. G.L.X., T.W., D.L., H.F., L.D., J.L.L., and L.Y.L. reviewed and edited the manuscript; All authors contributed to drafting the manuscript and have read and approved the final manuscript.

Funding

This study was funded by Health research project of Kunming Health Commission (2023–10 - 01–001) and the Yunnan Province High-level Chinese medicine talents project (2024). The funder had no role in the study design, data collection, analysis, interpretation, or writing of the protocol.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

The formal ethical approval of this systematic review is not applicable because it's a secondary analysis based on literature.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Second Clinical Medical School (Second Affiliated Hospital), Yunnan University of Chinese Medicine, Kunming, Yunnan, China. [2]Kunming Municipal Hospital of Traditional Chinese Medicine, Kunming, Yunnan, China.

Received: 25 January 2025 Accepted: 16 April 2025
Published online: 03 June 2025

References

1. Ogden K, Kilpatrick S, Elmer S. Examining the nexus between medical education and complexity: a systematic review to inform practice and research. BMC Med Educ. 2023;23(1):494.
2. Okorie CL, Bennett R, Aaron DM, Foster-Johnson L, Dagrosa AT. Answering the call to diversify the dermatology curriculum in undergraduate medical education. Clin Exp Dermatol. 2023;48(12):1370–2.
3. Gishen F, Lokugamage A. Diversifying the medical curriculum BMJ (Clinical research ed). 2019;364: l300.
4. Chakrabarti R, Wardle K, Wright T, Bennie T, Gishen F. Approaching an undergraduate medical curriculum map: challenges and expectations. BMC Med Educ. 2021;21(1):341.
5. Seong H, Yun D, Yoon KS, Kwak JS, Koh JC. Development of pre-procedure virtual simulation for challenging interventional procedures: an experimental study with clinical application. The Korean journal of pain. 2022;35(4):403–12.
6. Chew C, O'Dwyer PJ, Young D, Gracie JA. Radiology teaching improves Anatomy scores for medical students. Br J Radiol. 2020;93(1114):20200463.
7. Chytas D, Johnson EO, Piagkou M, Tsakotos G, Babis GC, Nikolaou VS, Markatos K, Natsis K. Three-dimensional printing in anatomy teaching: current evidence. Surgical and radiologic anatomy : SRA. 2020;42(7):835–41.
8. Kiesel M, Beyers I, Kalisz A, Wöckel A, Quenzer A, Schlaiß T, Wulff C, Diessner J. Evaluating the value of a 3D printed model for hands-on training of gynecological pelvic examination. 3D Print Med. 2022;8(1):20.
9. Triepels CPR, Smeets CFA, Notten KJB, Kruitwagen R, Futterer JJ, Vergeldt TFM, Van Kuijk SMJ. Does three-dimensional anatomy improve student understanding? Clinical anatomy (New York, NY). 2020;33(1):25–33.
10. Zhang SL, Ren SJ, Zhu DM, Liu TY, Wang L, Zhao JH, Fan XT, Gong H. Which novel teaching strategy is most recommended in medical education? A systematic review and network meta-analysis. BMC Med Educ. 2024;24(1):1342.
11. Rengier F, Mehndiratta A, von Tengg-Kobligk H, Zechmann CM, Unterhinninghofen R, Kauczor HU, Giesel FL. 3D printing based on imaging data: review of medical applications. Int J Comput Assist Radiol Surg. 2010;5(4):335–41.
12. Knox K, Kerber CW, Singel SA, Bailey MJ, Imbesi SG. Rapid prototyping to create vascular replicas from CT scan data: making tools to teach, rehearse, and choose treatment strategies. Catheterization and cardiovascular interventions : official journal of the Society for Cardiac Angiography & Interventions. 2005;65(1):47–53.
13. Tan H, Huang E, Deng X, Ouyang S. Application of 3D printing technology combined with PBL teaching model in teaching clinical nursing in

congenital heart surgery: A case-control study. Medicine (Baltimore). 2021;100(20): e25918.
14. Yan M, Huang J, Ding M, Wang J, Ni J, Wu H, Song D. Three-Dimensional Printing Model Enhances Correct Identification and Understanding of Pelvic Fracture in Medical Students. J Surg Educ. 2023;80(3):331–7.
15. Zhao J, Gong X, Ding J, Xiong K, Zhuang K, Huang R, Li S, Miao H. Integration of case-based learning and three-dimensional printing for tetralogy of fallot instruction in clinical medical undergraduates: a randomized controlled trial. BMC Med Educ. 2024;24(1):571.
16. Sun Z, Wong YH, Yeong CH. Patient-Specific 3D-Printed Low-Cost Models in Medical Education and Clinical Practice. Micromachines. 2023;14(2):464.
17. Asif A, Lee E, Caputo M, Biglino G, Shearn AIU: Role of 3D printing technology in paediatric teaching and training: a systematic review. BMJ paediatr open. 2021, 5(1).
18. Vaz VM, Kumar L. 3D Printing as a Promising Tool in Personalized Medicine. AAPS PharmSciTech. 2021;22(1):49.
19. Montgomery SJ, Kooner SS, Ludwig TE, Schneider PS. Impact of 3D Printed Calcaneal Models on Fracture Understanding and Confidence in Orthopedic Surgery Residents. J Surg Educ. 2020;77(2):472–8.
20. Masada KM, Cristino DM, Dear KA, Hast MW, Mehta S. 3-D Printed Fracture Models Improve Resident Performance and Clinical Outcomes in Operative Fracture Management. J Surg Educ. 2023;80(7):1020–7.
21. Huang Z, Song W, Zhang Y, Zhang Q, Zhou D, Zhou X, He Y: Three-dimensional printing model improves morphological understanding in acetabular fracture learning: A multicenter, randomized, controlled study. PloS one. 2018, 13(1).
22. Tomohisa O, Kamio T, Maeda Y, Tsubosaki K, Kato T, Iwata H. Application of Medical Imaging and 3D Printing Technology in Teaching the Handling of Novel Medicine in Periodontal Surgery. Cureus. 2022;14(9): e29271.
23. Zhang Y, Wang H, Yi J, Zhang H, Dong C, Wang G, Zhu Z, Liu W, Luo W. A novel 3D printed model for educating medical students on limb fractures: a randomized controlled preliminary study. J Orthop Surg Res. 2024;19(1):624.
24. Chen Y, Qian C, Shen R, Wu D, Bian L, Qu H, Fan X, Liu Z, Li Y, Xia J. 3D Printing Technology Improves Medical Interns' Understanding of Anatomy of Gastrocolic Trunk. J Surg Educ. 2020;77(5):1279–84.
25. Yi X, Ding C, Xu H, Huang T, Kang D, Wang D. Three-Dimensional Printed Models in Anatomy Education of the Ventricular System: A Randomized Controlled Study. World neurosurgery. 2019;125:e891–901.
26. Leung G, Pickett AT, Bartellas M, Milin A, Bromwich M, Shorr R, Caulley L. Systematic review and metaanalysis of 3D-printing in otolaryngology education. Int J Pediatr Otorhinolaryngol. 2022;155:111083.
27. Salazar D, Thompson M, Rosen A, Zuniga J. Using 3D Printing to Improve Student Education of Complex Anatomy: a Systematic Review and Meta-analysis. Med Sci Educ. 2022;32(5):1209–18.
28. Fleming C, Sadaghiani MS, Stellon MA, Javan R. Effectiveness of Three-Dimensionally Printed Models in Anatomy Education for Medical Students and Resident Physicians: Systematic Review and Meta-Analysis. J Am Coll Radiol. 2020;17(10):1220–9.
29. Zeng Y, Ji X, Dong B, Zhang L, Zheng Q, Wang Y, Han X, Ye L, Huang D, Wang S. 3D-printed coloured tooth model for inlay preparation in pre-clinical dental education. European journal of dental education : official journal of the Association for Dental Education in Europe. 2024;28(2):481–9.
30. Láinez Ramos-Bossini AJ, López Cornejo D, Redruello Guerrero P, Ruiz Santiago F. The Educational Impact of Radiology in Anatomy Teaching: A Field Study Using Cross-Sectional Imaging and 3D Printing for the Study of the Spine. Acad Radiol. 2024;31(1):329–37.
31. Feng S, Sun YJ, Zhu QR, Shi SF, Zhang YS, Yuan F. Integration of 3D printing and case-based learning in clinical practice for the treatment of developmental dysplasia of the hip. BMC Med Educ. 2024;24(1):986.
32. Dalgalı P, Topsakal KG, Eser Mısır S, Samur Ergüven S, Duran GS, Görgülü S. Evaluating the Impact of Different Education Methods on Cleft Lip and Palate Anatomy Training. Cleft Palate Craniofacial Journal. 2024;61(10):1743–9.
33. Brumpt E, Bertin E, Gabrion X, Coussens C, Tatu L, Louvrier A. Are 3D-printed anatomical models of the ear effective for teaching anatomy? A comparative pilot study versus cadaveric models. Surg Radiol Anat. 2024;46(2):103–15.
34. Yan X, Zhu Y, Fang L, Ding P, Fang S, Zhou J, Wang J. Enhancing medical education in respiratory diseases: efficacy of a 3D printing, problem-based, and case-based learning approach. BMC Med Educ. 2023;23(1):512.
35. Cook DA, Reed DA. Appraising the quality of medical education research methods: the Medical Education Research Study Quality Instrument and the Newcastle-Ottawa Scale-Education. Academic medicine : journal of the Association of American Medical Colleges. 2015;90(8):1067–76.
36. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JA. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ (Clinical research ed). 2011;343: d5928.
37. 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):1002–9.
38. Jaros S, Beck Dallaghan G. Medical education research study quality instrument: an objective instrument susceptible to subjectivity. Med Educ Online. 2024;29(1):2308359.
39. Cai B, Rajendran K, Bay BH, Lee J, Yen CC. The Effects of a Functional Three-dimensional (3D) Printed Knee Joint Simulator in Improving Anatomical Spatial Knowledge. Anat Sci Educ. 2019;12(6):610–8.
40. Fan G, Zhu S, Ye M, Long Y, Liu Z, Ye Y, Han W, Xie Y. Three-dimensional-printed models in bladder radical cystectomy: A valuable tool for surgical training and education. Int J Clin Exp Med. 2019;12(8):10145–50.
41. Lim KH, Loo ZY, Goldie SJ, Adams JW, McMenamin PG. Use of 3D printed models in medical education: A randomized control trial comparing 3D prints versus cadaveric materials for learning external cardiac anatomy. Anat Sci Educ. 2016;9(3):213–21.
42. Mogali SR, Chandrasekaran R, Radzi S, Peh ZK, Tan GJS, Rajalingam P, Yee Yeong W. Investigating the effectiveness of three-dimensionally printed anatomical models compared with plastinated human specimens in learning cardiac and neck anatomy: A randomized crossover study. Anat Sci Educ. 2022;15(6):1007–17.
43. Ochoa S, Segal J, Garcia N, Fischer EA. Three-Dimensional Printed Cardiac Models for Focused Cardiac Ultrasound Instruction. Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine. 2019;38(6):1405–9.
44. Radzi S, Chandrasekaran R, Peh ZK, Rajalingam P, Yeong WY, Mogali SR. Students' learning experiences of three-dimensional printed models and plastinated specimens: a qualitative analysis. BMC Med Educ. 2022;22(1):695.
45. Kong X, Nie L, Zhang H, Wang Z, Ye Q, Tang L, Li J, Huang W. Do Three-dimensional Visualization and Three-dimensional Printing Improve Hepatic Segment Anatomy Teaching? A Randomized Controlled Study. J Surg Educ. 2016;73(2):264–9.
46. Banwell HA, Causby RS, Crozier AJ, Nettle B, Murray C: An exploration of the use of 3D printed foot models and simulated foot lesions to supplement scalpel skill training in undergraduate podiatry students: A multiple method study. PloS one. 2021, 16(12 December).
47. Lane JC, Black JS. Modeling Medical Education: The Impact of Three-Dimensional Printed Models on Medical Student Education in Plastic Surgery. J Craniofac Surg. 2020;31(4):1018–21.
48. Chen S, Pan Z, Wu Y, Gu Z, Li M, Liang Z, Zhu H, Yao Y, Shui W, Shen Z, et al. The role of three-dimensional printed models of skull in anatomy education: a randomized controlled trail. Sci Rep. 2017;7(1):575.
49. Cheng H, Dai X, Wu B, Ye L, Wang H, Li S, Wang W. 3D-printed cranio-cerebral model promotes the interests and education efficiency for the medical interns in neurosurgery department. Int J Clin Exp Med. 2020;13(12):9302–10.
50. Al-Badri N, Touzet-Roumazeille S, Nuytten A, Ferri J, Charkaluk ML, Nicot R. Three-dimensional printing models improve long-term retention in medical education of pathoanatomy: A randomized controlled study. Clinical anatomy (New York, NY). 2022;35(5):609–15.
51. O'Brien C, Souza CA, Sheikh A, Miguel O, Wood T: Use of tracheobronchial tree 3-dimensional printed model: does it improve trainees' understanding of segmentation anatomy? A prospective study. 3D Print Med. 2021, 7(1).
52. AlAli AB, Griffin MF, Calonge WM, Butler PE. Evaluating the Use of Cleft Lip and Palate 3D-Printed Models as a Teaching Aid. J Surg Educ. 2018;75(1):200–8.

53. Li Z, Li Z, Xu R, Li M, Li J, Liu Y, Sui D, Zhang W, Chen Z. Three-dimensional printing models improve understanding of spinal fracture–A randomized controlled study in China. Sci Rep. 2015;5:11570.
54. Nicot R, Druelle C, Chazard E, Roland-Billecart T, Nuytten A, Richard F, Dupré A, Raoul G, Ferri J, Lacroix D, et al. Three-Dimensional Printing Model Enhances Craniofacial Trauma Teaching by Improving Morphologic and Biomechanical Understanding: A Randomized Controlled Study. Plast Reconstr Surg. 2022;149(3):475e–84e.
55. Su W, Xiao Y, He S, Huang P, Deng X. Three-dimensional printing models in congenital heart disease education for medical students: a controlled comparative study. BMC Med Educ. 2018;18(1):178.
56. Wang H, Xu H, Zhang J, Yu S, Wang M, Qiu J, Zhang M. The effect of 3D-printed plastic teeth on scores in a tooth morphology course in a Chinese university. BMC Med Educ. 2020;20(1):469.
57. Wang Z, Liu Y, Luo H, Gao C, Zhang J, Dai Y. Is a three-dimensional printing model better than a traditional cardiac model for medical education? A pilot randomized controlled study. Acta Cardiologica Sinica. 2017;33(6):664–9.
58. Wu AM, Wang K, Wang JS, Chen CH, Yang XD, Ni WF, Hu YZ: The addition of 3D printed models to enhance the teaching and learning of bone spatial anatomy and fractures for undergraduate students: A randomized controlled study. Ann Translat Med. 2018, 6(20).
59. Jiang W, Jiang W, Jin P, Zhang J, Xia J, Wei W, Qing B. Application of 3D printing technology combined with PBL teaching method in clinical teaching of cerebrovascular disease: An observational study. Medicine (United States). 2022;101(47):E31970.
60. Mayer RE. Applying the science of learning to medical education. Med Educ. 2010;44(6):543–9.
61. Young JQ, Van Merrienboer J, Durning S, Ten Cate O: Cognitive Load Theory: implications for medical education: AMEE Guide No. 86. Med Teach. 2014, 36(5):371–384.
62. Mancinetti M, Guttormsen S, Berendonk C. Cognitive load in internal medicine: What every clinical teacher should know about cognitive load theory. Eur J Intern Med. 2019;60:4–8.
63. Haji FA, Rojas D, Childs R, de Ribaupierre S, Dubrowski A. Measuring cognitive load: performance, mental effort and simulation task complexity. Med Educ. 2015;49(8):815–27.
64. Choonara YE, du Toit LC, Kumar P, Kondiah PP, Pillay V. 3D-printing and the effect on medical costs: a new era? Expert Rev Pharmacoecon Outcomes Res. 2016;16(1):23–32.
65. Panesar SS, Magnetta M, Mukherjee D, Abhinav K, Branstetter BF, Gardner PA, Iv M, Fernandez-Miranda JC. Patient-specific 3-dimensionally printed models for neurosurgical planning and education. Neurosurg Focus. 2019;47(6):E12.
66. Hansson T, Nyman T, Björkman A, Lundberg P, Nylander L, Rosén B, Lundborg G. Sights of touching activates the somatosensory cortex in humans. Scand J Plast Reconstr Surg Hand Surg. 2009;43(5):267–9.
67. Nicot R, Druelle C, Schlund M, Roland-Billecart T, Gwénaël R, Ferri J, Gosset D. Use of 3D printed models in student education of craniofacial traumas. Dent Traumatol. 2019;35(4–5):296–9.
68. Cook DA, Beckman TJ. Reflections on experimental research in medical education. Adv Health Sci Educ Theory Pract. 2010;15(3):455–64.

Publisher's Note