The effectiveness of virtual reality-based technology on anatomy teaching: a meta-analysis of randomized controlled studies

jingjie zhao
Xijing University
xinliang xu
Jining No.1 Peoples Hospital
Hualin Jiang
Xi'an Jiaotong University
yi ding (dingyi.007@163.com)
Xijing Hospital https://orcid.org/0000-0002-2158-3108

Research article

Keywords: Augmented and virtual reality, Improving classroom teaching, Teaching/learning strategies

Posted Date: January 20th, 2020

DOI: https://doi.org/10.21203/rs.2.21254/v1

Version of Record: A version of this preprint was published at BMC Medical Education on April 25th, 2020. See the published version at https://doi.org/10.1186/s12909-020-1994-z.

Abstract

Virtual reality (VR) is an innovation that permits the individual to discover and operate three-dimensional (3D) environment to gain practical understanding instantly. Recently, VR has been advanced as an encouraging tool in the course of clinical college. This research aimed to examine the general efficiency of VR for teaching medical anatomy. We executed a meta-analysis of randomized regulated studies of the performance of VR anatomy education. We browsed 5 databases from the year 1990 to 2019. Ultimately, 15 randomized controlled trials with a teaching outcome measure analysis were included. Tow authors separately chosen studies, extracted information, and examined the risk of bias. The primary outcomes were examination scores of the students. Secondary outcomes were the degree of satisfaction of the students. Random-effects models were used for the pooled evaluations. Standardized mean difference (SMD) was applied to assess the systematic results. The heterogeneity was determined by I2 statistics, and then was investigated by meta-regression and subgroup analyses. In this review, we screened and included fifteen randomized controlled researches (816 students). The pooled analysis of primary outcomes showed that VR improves test scores comparing with other approaches (standardized mean difference [SMD]= 0.53; 95% CI 0.09–0.97; I2= 87.8%). The high homogeneity indicated that the studies were different from each other. Therefore, we carried out meta-regression as well as subgroup analyses using 7 variables (year, country, learners, course, intervention, comparator, and duration). We found that VR improves postintervention test score of anatomy comparing with other types of teaching methods. Although the findings have high internal validity and limited, because of that anatomy teaching in medical universities appears to becoming a dilemma, VR may act as an efficient way to improve the learners' level of anatomy knowledge. Future research should assess other factors like degree of satisfaction, cost-effectiveness, and adverse reactions when evaluating the teaching effectiveness of VR in anatomy.

Introduction

Anatomy is a visual science which is thought an important foundation for medical learning [1]. When studying anatomy, the learners should identify structures and their spatial relationships. Nonetheless, medical students usually below the criteria due to safe medical practice in class. Medical students often commonly experience trouble acquiring adequate understanding of three dimensional (3D) anatomy from graphic images, such as those in textbooks and PowerPoint [2, 3]. So, it has become vital to create modern strategies concentrated on efficient as well as high-quality anatomy education and learning.

With new learning tools developing, clinical college educational programs have started incorporating more interactive media and online materials. The utilization of computer-based 3D models in anatomy education has become favorite over the last years [4]. Notably, VR is a technology that allows exploring and manipulating computer-generated real or artificial 3D multimedia environments in real-time. It allows for a first-person active learning experience through different levels of immersion. The rise of virtual reality technology could be traced back to the 1960's in the entertainment industry. In the next three decades, a dramatic surge of interest in using virtual reality technology beyond the entertainment field was shown in the frontier of professional training.

Recently, increasing interest has been paid on VR in medical educational world, particularly for anatomy teaching and resident surgical training [5, 6]. VR provides students a simulation scene to conceptualize intricate 3D anatomic connections quickly. Currently, VR methods have be established surgeons for preoperative planning and patient education, and even could be used to assess the surgical proficiency of residents for their training [7]. Some studies have compared VR to the other teaching methods for anatomy such as dissection, lectures, 2D images, and blended instruction. For example, Maresky et al. tested the effectiveness of a VR simulation of the heart in medical teaching [8]. They found that students under the VR simulation performed significantly better than the control group in the final test. In recent years, some researchers have conducted a meta-analysis of evaluation of teaching effect of using 3D visualization approaches in educational anatomy [9]. This study demonstrated that digital 3D methods could become a prospective remedy for the situation of insufficient and unsatisfactory anatomy education in some context. However, there is no high level of evidence on how efficient these different VR approaches are when contrasted to various other techniques in randomized controlled studies.

Accordingly, this meta-analysis of randomized controlled studies was to explored the educational effectiveness of VR applied to anatomy education in comparison with conventional or 2D digital methods in class.

Three research questions guided this study:

(1) Are the test scores improved better in VR education as compared to the other types of training?

(2) Are the satisfaction levels higher in VR education as compared to the other teaching methods?

(3) Do year of publication, country of study, subject of learning, intervention, comparator, and duration play a moderating role in the distinction?

Methods

Search strategy

Our study adhered to the PRISMA criteria [10]. Search terms for OVID MEDLINE was firstly performed and after that adjusted for the others: Embase, Cochrane Central Register of Controlled Trials, Web of Science Core Collection, and clinical trial registries. Terms as well as and subheadings such as key terms (anatomy) AND (virtual reality OR virtual learning environment OR mixed reality OR virtual classrooms OR augmented reality OR visualization technologies) AND (educat* OR simulat* OR training). We also searched and screened retrieved studies. Databases were searched from January January 1990 to August 2019. All records were screened on title as well as abstract, as well as suspicious records were analyzed completely in full text.

The search results from various databases were incorporated with Endnote software (EndNote X7, Clarivate Analytics, Philadelphia), and duplications of included studies were eliminated. Two authors (JJZ and YD) separately screened the search results as well as examined full-text research studies for inclusion. Any kind of disputes, for unclear or missing information were settled via conversation between the authors.

Inclusion and exclusion criteria

We included randomized controlled studies on comparing and studying VR intervention with control methods in anatomy teaching. In this review, VR methods including types of interactive 3D models, virtual patient or and surgical simulation could be performed as the single intervention or blended with others [11]. Studies were excluded with the following reasons: not randomized controlled study; not in the field of anatomy education, absence of an intervention; absence of test scores; insufficient data for effect size calculation. Exclusion was conducted by Y.D. and J.J.Z and inconformity was discussed and resolved.

Data extraction

We extracted data extraction from validity studies according to the Cochrane Handbook for Systematic Reviews [10]. In this review, the main concerned information covered year and region of the publication, details of learners, interventions, and duration of the study. Both of authors (J.J.Z and Y.D.) assessed the risk of bias for randomized controlled trials by the Cochrane risk of bias tool [12].

Data synthesis and heterogeneity assessment

All analyses were conducted by Stata 15 (StataCorp, College Station, TX, USA). Comparators included traditional education, other forms of digital education, and other types of VR. For continuous data, we summarized the standardized mean differences (SMDs) and associated 95% CIs across studies. We were unable to identify a clinically meaningful interpretation of SMDs for different kinds of VR education interventions. Therefore, the effect size was determined by the value of SMDs based on the Cohen rules: <0.2 (none), 0.2 to 0.5 (small), 0.5 to 0.8 (moderate), and > 0.80 (large). The random-effects model was chosen in our meta-analysis based on preliminary analysis. We applied I^2 statistic to determine heterogeneity. $I^2 < 25\%$ (low), 25–75% (medium), and > 75% (high) indicate different levels of heterogeneity.

Subgroup analysis was conducted when feasible. Seven attributes of each random were coded as possible moderators: year, region, learners, course, intervention, comparator, and duration. Sensitivity analyses was conducted to determine if the individual study significantly altered the [13]. Publication bias was determined by a funnel plot and Begg's test. P value < 0.05 was defined as significant.

Results

Search results

In overall, 15 studies were determined satisfied the inclusion requirements (Fig. 1 and Table 1). There were 15 randomized controlled studies with an overall of 816 learners: 745 were medical students and 71 were residents. Seven studies were performed in the USA. A series of VR educational methods were evaluated, including interactive 3D models, VR or and VR surgical stimulations. Interventions in the control group ranged from traditional learning (lecture, dissection and/or textbooks) to other digital education interventions. The duration of the intervention varied between 10 min to 2 weeks. For all research studies, primary results were determined by evaluation or survey studies at the end. And 5 out of 15 studies assessing satisfaction levels as the secondary outcome [14–18]. Table 1 shows the study characteristics of involved studies.

First author	Participants/Country	N (VR/control)	Course	Intervention	Comparator	Duration
Anthony, 2011	medical students/UK	12/14	anatomy of the forearm	VR	dissection and textbooks	50 min
Battulga, 2012	medical students/Japan	50/50	shoulder	3D interactive models	2D images	60 min
de Faria, 2016	medical students/Brazil	28/28	neuroanatomy	3D interactive models	2D images	60 min
Ellington, 2018	residents/UK	16/15	female pelvic anatomy	VR	power point	2 weeks
Hampton, 2010	medical students 3, 4 year /USA	21/22	female pelvic anatomy	3D interactive models	dissection and textbooks	60 min
Keedy, 2011	medical students 1, 4 year/USA	23/23	anatomy of the liver	3D interactive models	2D images	1 day
Khot, 2013	medical students/Canada	20/20	pelvic anatomy	VR	power point	10 min
Kockro, 2015	medical students/Germany	89/80	spatial neuroanatomy	3D interactive models	power point	20 min
Moro, 2017	medical students/Australia	20/22	skull anatomy	VR	3D models	10 min
Nicholson, 2004	medical students 1 year /USA	29/28	ear anatomy	3D interactive models	text books	2 day
Seixas, 2010	surgical trainees/USA	5/5	human anatomy	VR	2D images	1 day
Solyar,	medical student/USA	7/8	paranasal	VR	textbooks	60 min

2008			sinuses			
Stepan, 2017	medical students 1,2 year /USA	33/33	neuroanatomy	VR	text books	1 day
Tan, 2012	residents/ Canada	21/19	laryngeal anatomy	3D interactive models	text books	45 min
Zachary, 2015	medical students/USA	41/32	neuroanatomy	3D interactive models	2D images and 3D models	65 min

Table 1
Characteristics of included studies

Risk of bias assessment

The risk of bias of majority studies involved was unclear or high risk as shown in the bias summary (Fig. 2). Most studies did not have information about allocation concealment and baseline of learners' characteristics. Due to the nature of the intervention, it is not practical for blinding of students and teachers during the study. For risk of completeness of data, and selective reporting, most studies were determined low. We assessed whether the research study was devoid of selective outcome reporting, which was checked whether outcomes mentioned adequately in manuscripts. And 5 studies were judged to be of high risk on completeness of data because of incomplete or accurate data on outcome standard deviation [15, 19–21].

Data synthesis

The meta-analysis plots of primary and secondary outcomes are shown in Fig. 3A and B. The effectiveness of intervention on examination scores was reported in all studies. The studies assessed test scores as a primary outcome with multiple-choice questionnaires. We found that VR significantly increased learners' examination scores comparing with traditional learning in the random-effects model (SMD = 0.53; 95% CI 0.09−0.97; I^2 = 87.8%) (Fig. 3A). Nine of the studies (60%) showed that VR significantly increased students' examination scores when compared with traditional learning (lecture, dissection and/or textbooks) to other digital 2D methods; and five (15%) failed to reveal statistically significant effects between the VR and the control groups. Outcomes showed that the studies were heterogeneous ($p < 0.001$) and the true effects were not consistent among studies.

A total of 5 studies assessed satisfaction levels as secondary outcome [14–18]. The pooled results based on the fixed effects model showed that most students have a greater interest in learning via VR methods, rather than conventional or 2D teaching methods (SMD = 0.77; 95% CI 0.47–1.07; I^2 = 20.5%). (Fig. 3B), all favoring VR over control groups (large effect size). However, only one study mentioned the adverse effects that some participants using VR displayed headaches, dizziness, or blurred vision [22].

Publication bias

For the primary analyses, funnel-plots were made to check for risk of publication bias (Fig. 4). The shape of the funnel plot is symmetrical. Meanwhile, the result of Begg's test show a non-significant asymmetry (P = 0.54) [23]. Thus, there was no significant publication bias indicated in this review.

Subgroup analyses

A random-effects model was used for the subgroup analysis due to each subgroup being heterogeneous according to the results of tests (Table 2) [24]. As indicated in Tables 2, the categorical variables were as follows: region (USA or others), learners (medical students or residents), course (skeletal anatomy or neuroanatomy or others), intervention (3D interactive models or VR simulations), comparator (traditional methods or other digital methods) and duration (< 1 day or ≥ 1 day). Other potential moderators could not be analyzed because they were reported inadequate to do a subgroup analysis. According to subgroup analysis, we found that did not make a significant difference in the differences in the subgroups for all Q statistics are non-significant ($I^2 > 75\%$).

Subgroup	n	SMD	95% CI	p value	I^2
region					
USA	7	1.14	0.56, 1.72	0.00	79.8%
others	8	0.03	-0.57, 0.63	0.92	89.6%
learners					
medical students	12	0.51	0.02, 1.01	0.04	89.6%
residents	3	0.67	-0.45, 1.79	0.24	77.8%
course					
skeletal anatomy	6	-0.07	-0.95, 0.81	0.88	91.4%
neuroanatomy	4	0.52	-0.04, 1.10	0.07	84.9%
others	5	1.34	0.52, 2.14	0.00	87.8%
intervention					
3D interactive models	8	0.64	0.47, 0.81	0.00	82.5%
VR	7	-0.09	-0.37, 0.18	0.50	89.2%
comparator					
traditional methords	5	0.81	0.15, 1.47	0.02	82.6%
other digital methods	10	0.35	-0.25, 0.95	0.25	90.2%
duration					
< 1day	10	0.35	0.18, 0.52	0.00	89.4%
≥1 day	5	0.71	0.42, 1.10	0.00	84.4%

Table 2
Summary statistics for moderators related to examination scores

Interestingly, the moderator analysis revealed significant benefits of VR in the subgroup of medical students (SMD = 0.51; 95% CI 0.02–1.01, p = 0.04), whereas VR have no significant influence on residents (SMD = 0.67; 95% CI -0.45–1.01, p = 0.24). Also, moderator analysis of control type showed that test scores of the VR group was not significantly better than using other 2D digital methods (SMD = 0.35; 95% CI -0.25–0.95, p = 0.25), while there was a significant improvement when compared with the traditional intervention group (SMD = 0.81; 95% CI 0.15–1.47, p = 0.02). For the duration analysis, VR interventions for at least one day had moderately-to-large effects on scores (SMD = 0.71; 95% CI 0.42–1.10, $p < 0.001$), whereas those which were < 1 day had only a small effect (SMD = 0.35; 95% CI 0.18–0.52, $p < 0.001$).

Meta-regression analyses

To determine whether there were any moderation effects on primary outcomes, meta-regression analyses were conducted. We regressed effect sizes on 7 potential moderators: year, country, learners, course, intervention, comparator, and duration. As shown in Table 3, none of them turned out to be significant at a level of $p < 0.05$.

Factors	Coefficient	Standard error	95% CI	p value
year	-0.12	0.20	-3.06, 0.67	0.21
country	-1.19	0.95	-2.99, 0.54	0.17
learners	1.08	1.24	-1.35, 3.52	0.38
course	-0.26	0.89	-2.01, 1.49	0.77
intervention	-0.33	0.79	-0.53, 0.27	0.67
comparator	0.29	0.86	-1.40, 1.99	0.73
duration	0.09	0.95	-1.77, 1.97	0.91

Table 3
Meta-regression analysis for exploration of the sources of heterogeneity factors

Sensitivity analyses

Due to the significant heterogeneity (> 75%), a sensitivity analysis was used to verify the reliability of the result. When any research was removed from the model, the significant results of the VR effect on examination scores were unchanged in the models (SMD = 0.53, 95% CI: 0.01–1.07) (Fig. 5). Thus, the results indicated that the findings for examination scores were robust.

Discussion

Our meta-analysis of randomized controlled studies examined the effectiveness of VR-based technology in anatomy teaching. In this study, the authors discovered that VR interventions have a moderate enhancement (SMD = 0.53) in test scores of learners in comparation with conventional or other 2D digital methods ($p < 0.01$). As is well-known, more interactive VR interventions could moderately improve medical learners' academic scores in anatomy. Among 15 studies, only 5 studies assessed satisfaction scores as a secondary outcome with a result that most of students more interested in using VR to learn anatomy. Naturally, the fact that no included randomized controlled researches were found in databases before 2004 suggests that VR is an emerging academic method [25], attracting increasing interest from the world of education. In general, the risk of bias for most studies was unclear for a lack of description or data. Potentially high risk of incomplete reporting bias identified in some studies. However, results of sensitivity and subgroup analyses were nonsignificant for variables (year, country, learners, course, intervention, comparator, and duration) on the outcome variables. Since the different types of learners and interventions in researches in this review, inconsistent methodological method makes it difficult to draw accurate conclusions.

In the subgroup analysis for levels of learners, the source of high heterogeneity could be diverse phases of participants' medical education among included studies. In several studies, learners are first-year medical students [17, 26], while others are forth-year medical students [14, 15]. Of course, the longer learners got more knowledge of anatomy, which leads to comparing results complex or paradoxical. As Hattie et al. had concluded, the different expertise degrees of learners are remarkable in education [27]. Therefore, medical students in school could be easier motivated and effective in front of fictitious scenarios, since of their less clinical experience versus residents. In addition, various organs or body parts learned present different levels of complexity, leading to the heterogeneity in results. For example, learning the anatomy of the brain is harder than learning skeletal parts [17, 18]. In terms of duration, the results of this review showed that a course for 1 day or longer had a larger effect size than a course for several hours (0.71 vs 0.35). Thus, the learning duration has influenced the educational efficiency of VR methods, which should be considered and adjusted in practice.

Types of comparator is another source of variation. Only 5 of 15 studies were found where this technology was compared to traditional methods such as lectures, dissection or textbooks. However, it would be more meaningful to conduct evaluations of studies that compare the different features of digital-based methods rather than those which compare digital-based to traditional methods [19]. Dissection is regarded as the standard teaching method for anatomy all the time. in this review, only 2 of 15 studies compared VR with dissection for anatomy teaching [14, 19]. In fact, VR could be used as an adjunct to dissection in class with fewer lab hours or resources. For example, a study found that better results for group with dissection for anatomy teaching compared to computer-based teaching group, while the best scores came from the group who were taught by blended methods of both [28].

For satisfaction scores, the pooled result of the comparison of VR versus others was significantly in favor of VR, which could be due in part to the novelty of the method. Most of the participants in the studies reported that the VR methods were easier and more enjoyable to use. Some researchers had revealed that there is a significant positive correlation between motivation and academic record of students [29]. However, due to the complicated anatomical configuration, in one study, some participants found the VR methods disorienting and frustrating [22]. Thus, more studies should focus on the adverse effect caused by this new technology as well as the attitude of participates.

As a fast-moving technology, the cost of VR will be a critical aspect when considering to apply it into education especially for low-income settings. In this review, only one study is from Brazil [30], which reduces the applicability of innovative educational methods to developing regions. Unfortunately, none random controlled study on cost-effectiveness of VR comparing with others.

Strengths And Limitations

VR is currently a new visualization technique, so there was no high-quality evidence on the effectiveness of VR-based technology. It is hard to offer an overall conclusion of the efficiency of these strategies. The strengths of this meta-analysis were detailed search on randomized controlled studies, and the data was drawn out by two of authors independently. Because of the variability in studies, we also assess the risk of bias, sensitivity analyses and meta-regression analyses on outcomes from articles. Results of sensitivity and subgroup analyses were nonsignificant, indicating that the findings were robust.

This review also has several limitations. Firstly, the included researches mainly reported postintervention information, so we did not compute pre-to post-intervention modification. The validity of the different assessments used in the included studies might constitute a bias. Gender information was not easy to obtain in the current meta-analysis, but it is an important factor influencing teaching effect [31]. In future studies, information on gender ratio for treatment and control group may be collected for analysis. Besides, none of the studies assessed the cost of setup and maintenance of the VR-based intervention, whereas none of the included studies assessed cost-effectiveness. Further research should evaluate the effectiveness of VR in a variety of settings and evaluate outcomes such as attitude, adverse effects, and cost-effectiveness.

Conclusions

As an emerging and new technology, VR has the potency in transforming medical teaching. In this meta-analysis, results showed that when compared with traditional or 2D digital methods, VR at least as or potentially improve teaching effectiveness of anatomy. However, our results are not certain for lack of standardized measures and high heterogeneity among studies, and the appropriate mode of integrating VR into class needs to be further explored. To enhance the teaching quality, VR as an implement could be considered on the medical teaching situations by universities and hospitals.

Abbreviations

VR: Virtual reality (); SMD: Standardized mean difference; three-dimensional: 3D

Declarations

Acknowledgments

This work was funded by research fund program of xijing university (XJ180203).

Authors' contributions

HLJ and XLX involved in designing the study, collecting the data, analyzing the data, and writing the manuscript. YD and JJZ contributed to the study methodology and corrected the final version of the manuscript.

Ethics approval

Not applicable.

Conflict of interest

None.

Funding

The authors declare that they have no source of funding for the research.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. McLachlan JC, Patten D: Anatomy teaching: ghosts of the past, present and future. *Medical education* 2006, 40(3):243-253.
2. Wainman B, Wolak L, Pukas G, Zheng E, Norman GR: The superiority of three-dimensional physical models to two-dimensional computer presentations in anatomy learning. *Medical education* 2018, 52(11):1138-1146.
3. Yammine K, Violato C: The effectiveness of physical models in teaching anatomy: a meta-analysis of comparative studies. *Advances in Health Sciences Education* 2016, 21(4):883-895.
4. Sugand K, Abrahams P, Khurana A: The anatomy of anatomy: a review for its modernization. *Anatomical sciences education* 2010, 3(2):83-93.
5. Pilot A: Virtual Reality and Cardiac Anatomy: Exploring. 2018.

6. Erolin C, Reid L, McDougall S: Using virtual reality to complement and enhance anatomy education. *Journal of visual communication in medicine* 2019:1-9.
7. Rose K, Pedowitz R: Fundamental arthroscopic skill differentiation with virtual reality simulation. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2015, 31(2):299-305.
8. Maresky H, Oikonomou A, Ali I, Ditkofsky N, Pakkal M, Ballyk B: Virtual reality and cardiac anatomy: Exploring immersive three-dimensional cardiac imaging, a pilot study in undergraduate medical anatomy education. *Clinical Anatomy* 2019, 32(2):238-243.
9. Yammine K, Violato C: A meta-analysis of the educational effectiveness of three-dimensional visualization technologies in teaching anatomy. *Anatomical sciences education* 2015, 8(6):525-538.
10. Moher D, Liberati A, Tetzlaff J, Altman DG: Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine* 2009, 151(4):264-269.
11. Fernández-Palacios BJ, Morabito D, Remondino F: Access to complex reality-based 3D models using virtual reality solutions. *Journal of cultural heritage* 2017, 23:40-48.
12. Higgins JP: Cochrane handbook for systematic reviews of interventions version 5.0. 1. The Cochrane Collaboration. *http://www cochrane-handbook org* 2008.
13. Kong Y, Seo YS, Zhai L: Comparison of reading performance on screen and on paper: A meta-analysis. *Computers & Education* 2018, 123:138-149.
14. Hampton BS, Sung VW: Improving medical student knowledge of female pelvic floor dysfunction and anatomy: a randomized trial. *American journal of obstetrics and gynecology* 2010, 202(6):601. e601-601. e608.
15. Keedy AW, Durack JC, Sandhu P, Chen EM, O'Sullivan PS, Breiman RS: Comparison of traditional methods with 3D computer models in the instruction of hepatobiliary anatomy. *Anatomical sciences education* 2011, 4(2):84-91.
16. Solyar A, Cuellar H, Sadoughi B, Olson TR, Fried MP: Endoscopic Sinus Surgery Simulator as a teaching tool for anatomy education. *American journal of surgery* 2008, 196(1):120-124.
17. Stepan K, Zeiger J, Hanchuk S, Del Signore A, Shrivastava R, Govindaraj S, Iloreta A: Immersive virtual reality as a teaching tool for neuroanatomy. *International forum of allergy & rhinology* 2017, 7(10):1006-1013.
18. Drapkin ZA, Lindgren KA, Lopez MJ, Stabio ME: Development and assessment of a new 3D neuroanatomy teaching tool for MRI training. *Anatomical sciences education* 2015, 8(6):502-509.
19. Codd AM, Choudhury B: Virtual reality anatomy: Is it comparable with traditional methods in the teaching of human forearm musculoskeletal anatomy? *Anatomical sciences education* 2011, 4(3):119-125.
20. Khot Z, Quinlan K, Norman GR, Wainman B: The relative effectiveness of computer-based and traditional resources for education in anatomy. *Anatomical sciences education* 2013, 6(4):211-215.
21. Seixas-Mikelus SA, Adal A, Kesavadas T, Baheti A, Srimathveeravalli G, Hussain A, Chandrasekhar R, Wilding GE, Guru KA: Can image-based virtual reality help teach anatomy? *Journal of endourology* 2010, 24(4):629-634.
22. Moro C, Stromberga Z, Raikos A, Stirling A: The effectiveness of virtual and augmented reality in health sciences and medical anatomy. *Anatomical sciences education* 2017, 10(6):549-559.
23. Begg CB, Mazumdar M: Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994:1088-1101.
24. Harris RJ, Deeks JJ, Altman DG, Bradburn MJ, Harbord RM, Sterne JA: Metan: fixed-and random-effects meta-analysis. *The Stata Journal* 2008, 8(1):3-28.
25. Liu D, Bhagat KK, Gao Y, Chang T-W, Huang R: The potentials and trends of virtual reality in education. In: *Virtual, augmented, and mixed realities in education.* edn.: Springer; 2017: 105-130.
26. Nicholson DT, Chalk C, Funnell WRJ, Daniel SJ: Can virtual reality improve anatomy education? A randomised controlled study of a computer-generated three-dimensional anatomical ear model. *Medical education* 2006,

40(11):1081-1087.

27. Hattie J: The applicability of Visible Learning to higher education. *Scholarship of Teaching and Learning in Psychology* 2015, 1(1):79.
28. Biasutto SN, Caussa LI, del Río LEC: Teaching anatomy: cadavers vs. computers? *Annals of Anatomy-Anatomischer Anzeiger* 2006, 188(2):187-190.
29. Kim K-J, Frick TW: Changes in student motivation during online learning. *Journal of Educational Computing Research* 2011, 44(1):1-23.
30. de Faria JWV, Teixeira MJ, Júnior LdMS, Otoch JP, Figueiredo EG: Virtual and stereoscopic anatomy: when virtual reality meets medical education. *Journal of neurosurgery* 2016, 125(5):1105-1111.
31. Bleakley A: Gender matters in medical education. In: *Patient-Centred Medicine in Transition.* edn.: Springer; 2014: 111-126.

Figures

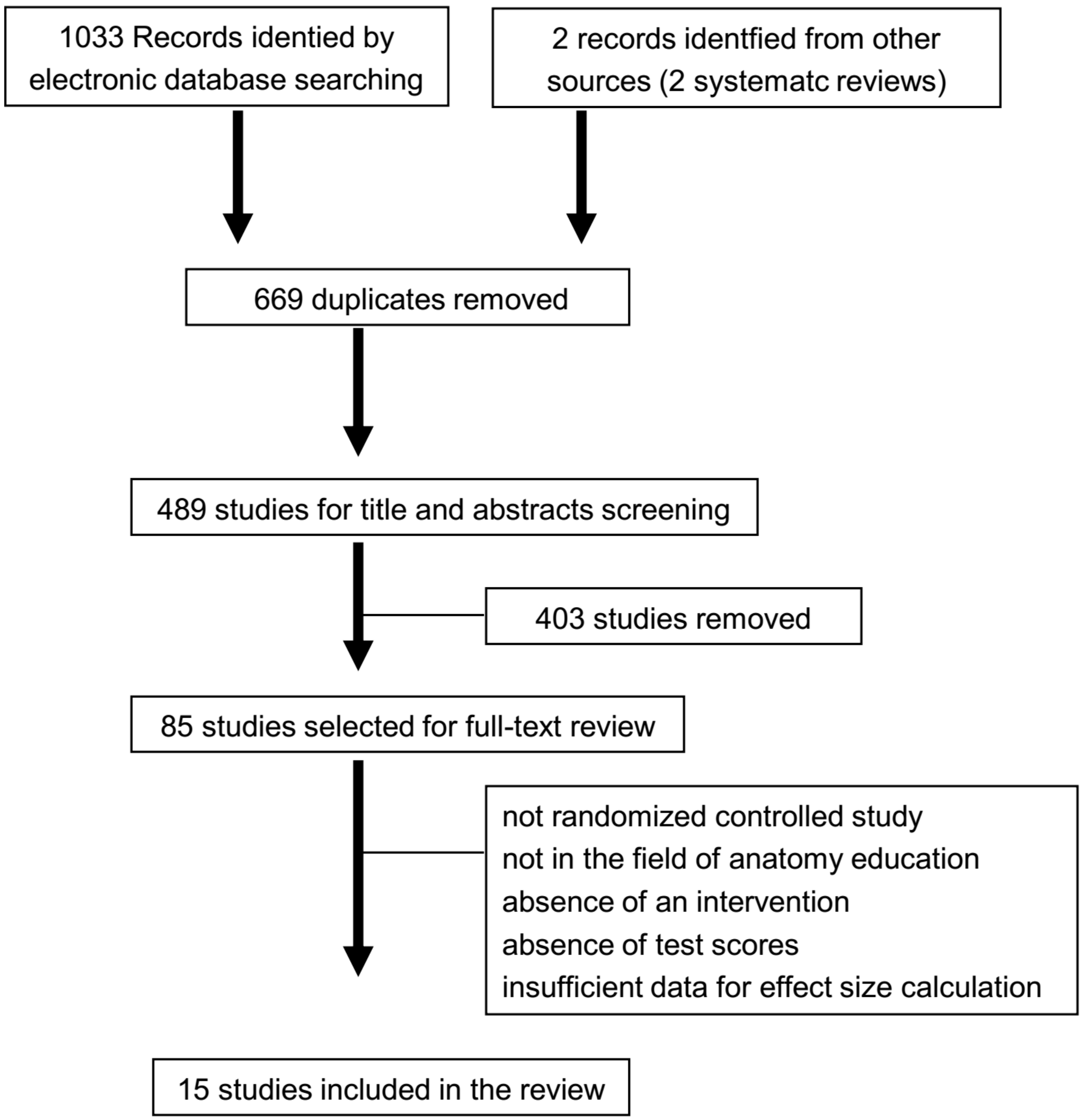


Figure 1

Flowchart of the search strategy

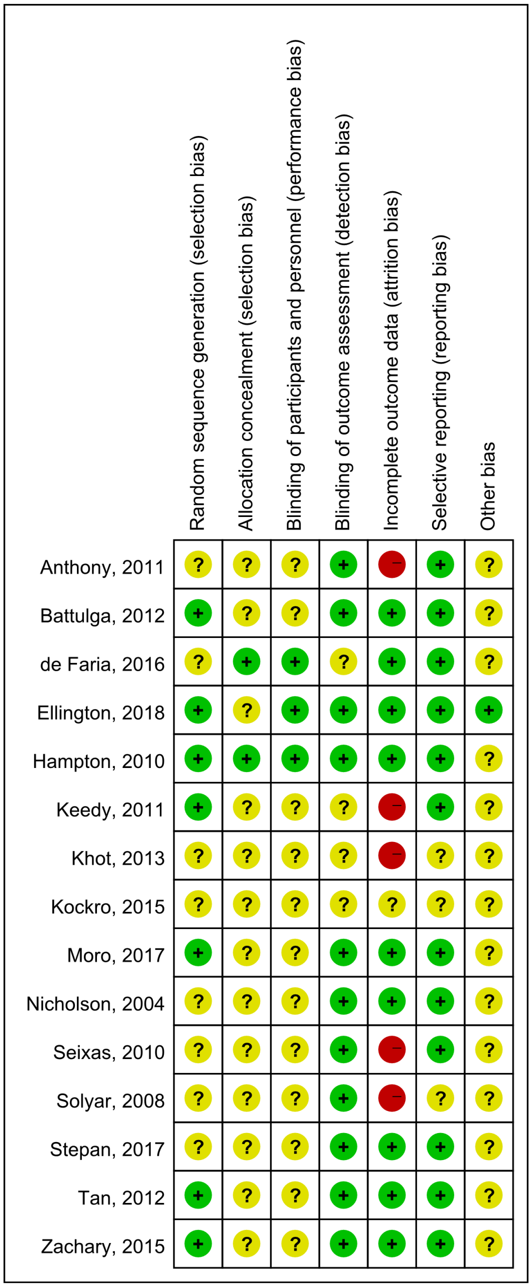
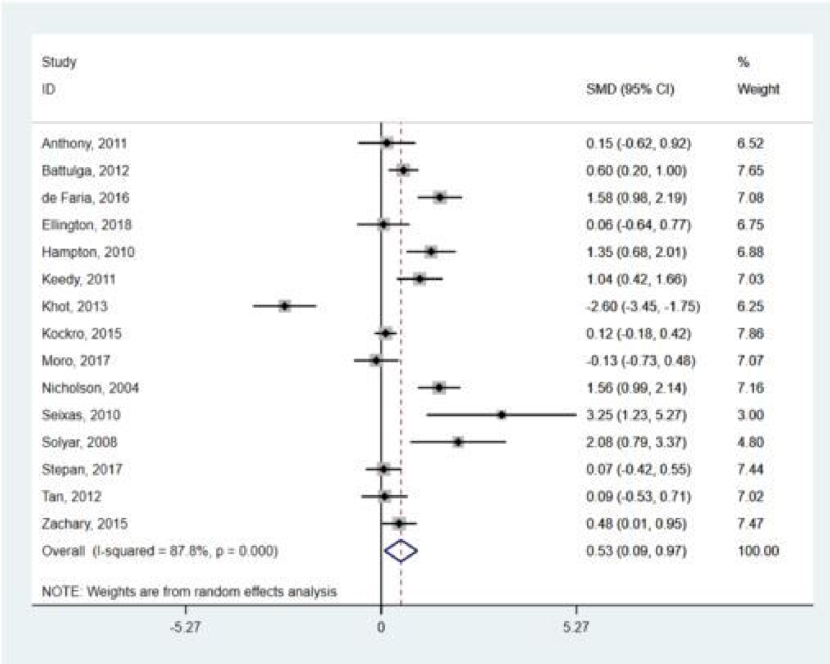


Figure 2

Risk of bias assessment of included studies

A



B

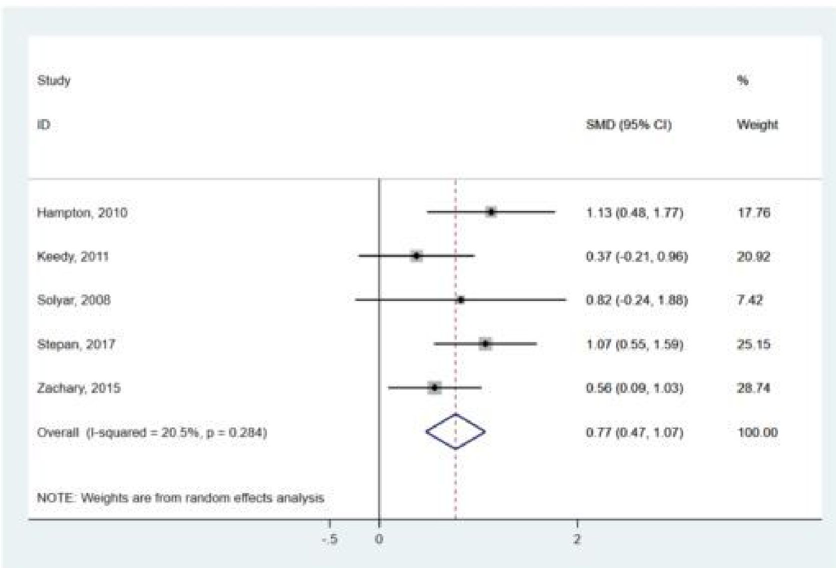


Figure 3

Forest plots for examination scores (A) and satisfaction outcomes (B)

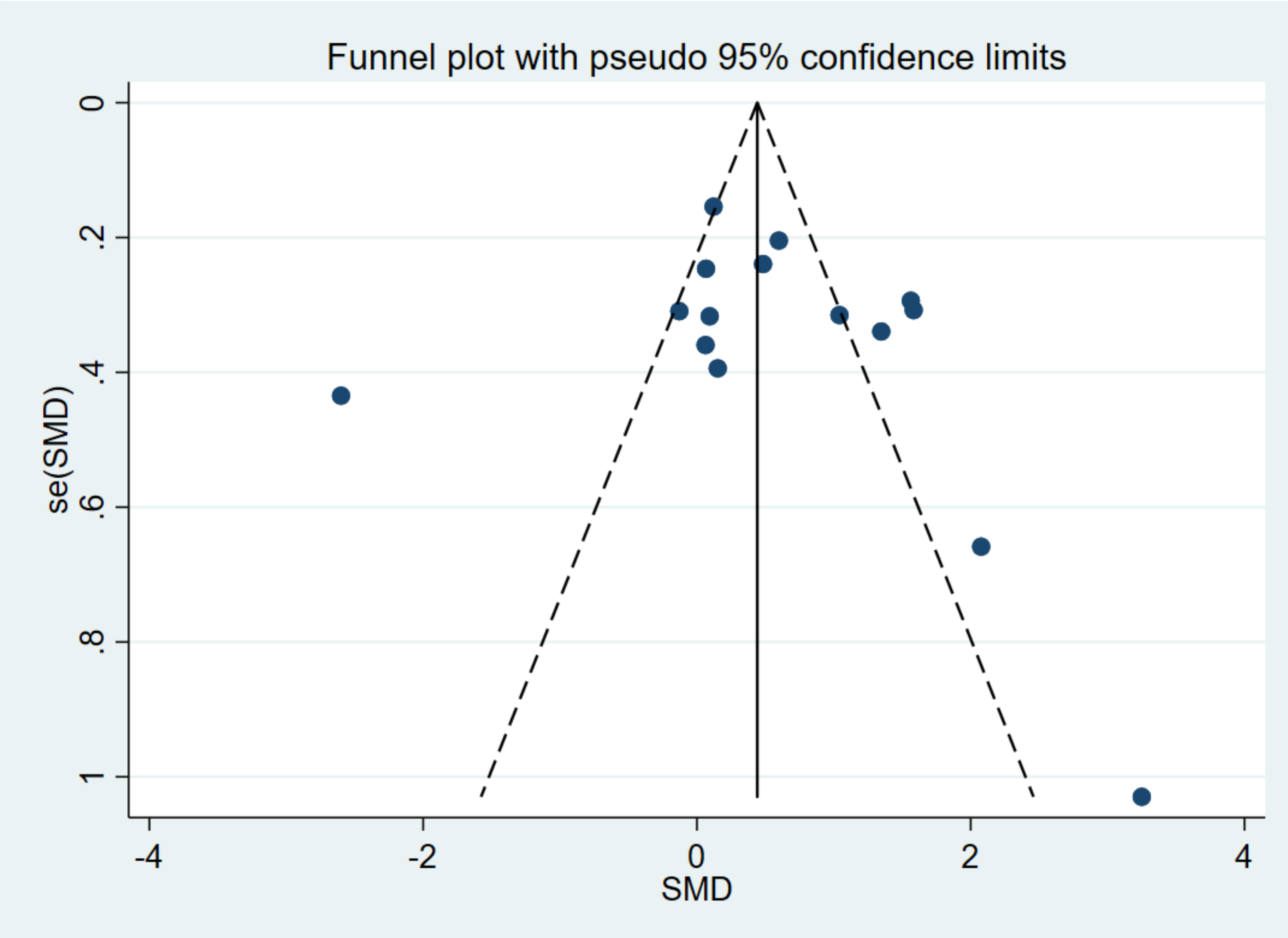


Figure 4

Funnel plot analysis for examination scores

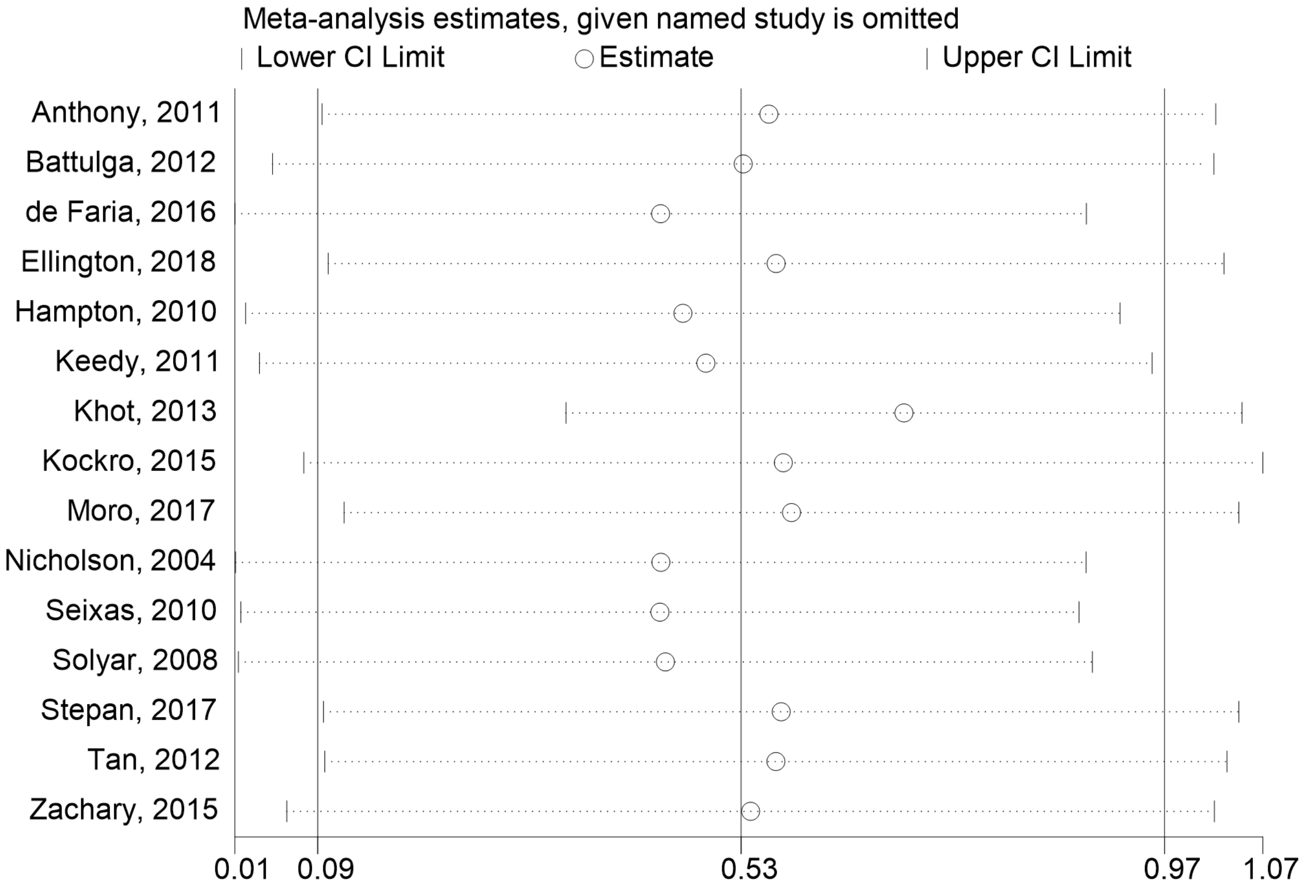


Figure 5

Sensitivity analysis assessing the influence of each study on the pooled analysis