Medical Student Education

Harnessing Augmented Reality and CT to Teach First-Year Medical Students Head and Neck Anatomy

Joanna K. Weeks, MD, Jina Pakpoor, MD, Brian J. Park, MD, MS, Nicole J. Robinson, BA,
Neal A. Rubinstein, MD, PhD, Stephen M. Prouty, PhD, Arun C. Nachiappan, MD

Rationale and Objectives: Three-dimensional (3D) visualization has been shown to benefit new generations of medical students and physicians-in-training in a variety of contexts. However, there is limited research directly comparing student performance after using 3D tools to those using two-dimensional (2D) screens.

Materials and Methods: A CT was performed on a donated cadaver and a 3D CT hologram was created. A total of 30 first-year medical students were randomly assigned into two groups to review head and neck anatomy in a teaching session that incorporated CT. The first group used an augmented reality headset, while the second group used a laptop screen. The students were administered a five-question anatomy test before and after the session. Two-tailed t-tests were used for statistical comparison of pretest and posttest performance within and between groups. A feedback survey was distributed for qualitative data.

Results: Pretest vs. posttest comparison of average percentage of questions answered correctly demonstrated both groups showing significant in-group improvement ($p < 0.05$), from 59% to 95% in the augmented reality group, and from 57% to 80% in the screen group. Between-group analysis indicated that posttest performance was significantly better in the augmented reality group ($p = 0.022$, effect size = 0.73).

Conclusion: Immersive 3D visualization has the potential to improve short-term anatomic recall in the head and neck compared to traditional 2D screen-based review, as well as engage millennial learners to learn better in anatomy laboratory. Our findings may reflect additional benefit gained from the stereoscopic depth cues present in augmented reality-based visualization.

Key Words: 3D visualization; Augmented reality; Medical education; Mixed reality; Near-peer; Technology in education.

Abbreviations: 2D two-dimensional, **3D** three-dimensional, **AR** augmented reality, **CT** computed tomography, **NBME** National Board of Medical Examiners, **SD** standard deviation

INTRODUCTION

Compared to previous generations, millennial learners have created new challenges for creators of preclinical curricula by placing priority on hands-on learning experiences, individualized feedback, and the use of technology (1). Medical educators have risen to this challenge by designing and implementing innovative mixed–media learning tools in fields which traditionally have depended heavily on didactic lectures supplemented by texts. Anatomy teaching in particular has undergone a rigorous reinvention during its transition from a 2-year dissection-based course to a 12-week integrated curriculum across most medical schools, and has embraced radiologic correlation and anatomic visualization technology in the process (2,3). Examples include computer-generated models available on smartphones, tablets, and laptops, three-dimensional (3D) printing from CT scans, and recently, experiments with augmented and virtual reality (2,4−6).

In augmented reality (AR), a computer-generated 3D world is seamlessly merged with the physical world. This allows the user to see and interact normally with the physical world while still engaging with displayed digital objects. Thus far, the technology has been successfully used for clinical training in laparoscopic surgery, neurosurgical procedures, and echocardiography, as well as for resident and medical student training in ear, nose, and throat surgery (7,8). Early results indicate that AR surgical training is as effective as cadaver-based training, and may save time (8−10).

In this study, we utilized a randomized-control method to investigate the effectiveness of a CT-based head and neck anatomy module for first-year medical students using a 3D AR as compared to an identical session using a navigable two-dimensional (2D) screen instead. We hypothesized that medical students using the AR tool with its benefit of stereoscopic depth cues would perform better than students using traditional 2D screens.

MATERIALS AND METHODS

This study was exempted by our institutional review board. The study was carried out in accordance with the Declaration of Helsinki, including, but not limited to, the anonymity of participants being guaranteed and the informed consent of participants being obtained.

Study Population

All first-year medical students enrolled in a large U.S. medical school were given the opportunity to participate in this study through an announcement at the end of a mandatory lecture, followed by an email detailing participation requirements. Approximately 80 interested students emailed the investigator. Using first-come-first-serve priority, 30 students were enrolled into the study and assigned a random 6-digit identifier. Using a random sorter, half of the participants were selected for a 3D AR viewing group, and the other half were selected for a 2D screen viewing group.

Content Selection

The spatially complex and visually challenging anatomy of the head and neck was selected as the topic for review by the research team. Particular focus was given to the interconnected sinuses, the nasopharynx and larynx, and the vasculature of the head and neck. Relevant test questions were selected by a resident in diagnostic radiology with prior tutoring experience and derived from publicly available National Board of Medical Examiners (NBME)-style head and neck anatomy questions. Each question was vetted for accuracy before inclusion. We coordinated the presentation of material with the students' anatomy lecture schedules to ensure that the students had not previously been taught the material to be covered. To discourage memorization, students were not informed that the same questions would appear on the pre- and posttests.

Cadaver CT Scan

We obtained 0.5 mm slice thickness noncontrast CT scans of the cadaver under study through collaboration with anatomy laboratory staff and radiology department personnel. The digital imaging and communications in medicine (DICOM) standard images were exported from our picture archiving and communication system (PACS) and uploaded to the AR headset device and a laptop for viewing in the anatomy laboratory. The AR CT hologram was "hung" in space parallel to the actual cadaver so that students in the AR group could see the cadaver's CT hologram in direct physical relationship to the body itself (see Fig 1). Display settings were adjusted to highlight structures such as bones (see Fig 2). CT in the screen group was displayed in standard 2D slices in the axial, coronal, and sagittal orientations.

Session Design

A resident in diagnostic radiology with training on the use of both the augmented reality headset and the laptop software taught both groups of students for purposes of consistency and standardization. Each group took part in a 60-minute session consisting of the following: (1) introduction to the project, (2) five-question NBME-style pretest focused on spatial relationships (see Supplement A), (3) 15-minute introduction
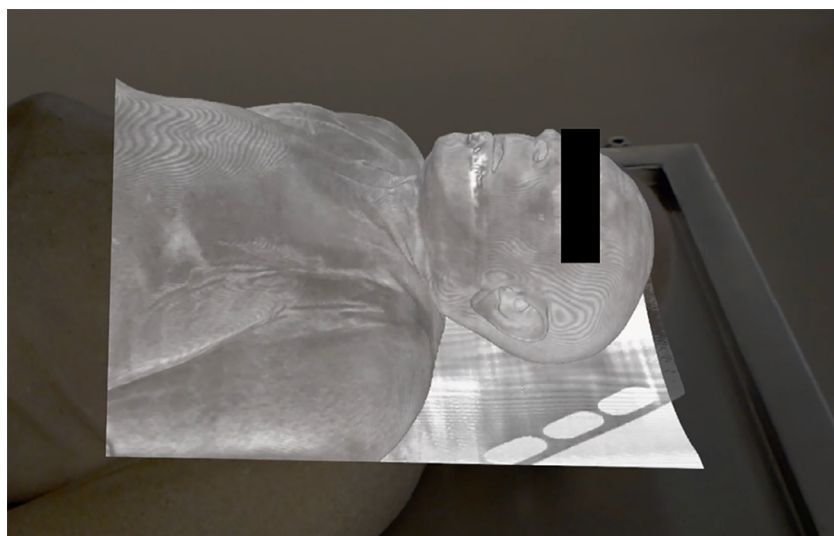

Figure 1. *Foreground*: Screen capture taken from the augmented reality (AR) headset looking at the 3D hologram of the noncontrast CT of the cadaver's head and neck as seen "hung" in space parallel to the actual physical cadaver resting on the surgical table. The black box covering the eyes is superimposed on the image for anonymity. *Background:* Donated cadaver, utilized to obtain CT images and also for gross dissection, resting on surgical table.
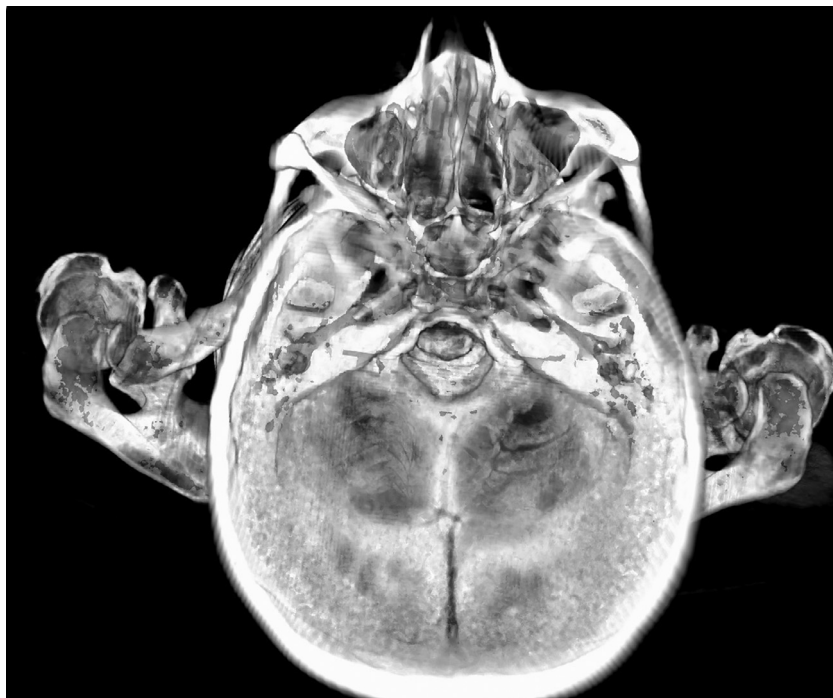
Figure 2. Example view of the cadaver CT head and neck used in the teaching sessions showing the sinuses. This view was seen in 3D by students in the augmented reality (AR) group using an AR-capable headset and Surgical AR (Medivis, Inc., Brooklyn, New York) software.

to the relevant anatomy using an Atlas of Human Anatomy (Netter, 2014) book, (4) 15- to 30-minute guided review of the head and neck portion of the cadaver CT scan, either using the AR-viewer headset or the laptop screen, with time duration depending on the student's pace, (5) correlation of the reviewed anatomy on the bisected cadaver from which the CT was obtained, and (6) an identical five-question NBME-style posttest (see Supplement A). After the session, students were emailed a qualitative feedback form and asked to complete and return it to the investigators within 2 weeks. In free-response questions, they were asked what they liked and what they would change about the technology and the way it was used to teach (see Supplement B).

Statistical Analysis

Within-group pre- vs. posttest mean comparison was calculated using two-tailed paired t tests. Between-group posttest mean comparison was calculated using two-tailed unpaired t tests. Effect size was calculated as Glass's *delta*. A post hoc power analysis was performed for the between-group posttest mean comparison to compute achieved power. Fisher's exact test was used to analyze their pre- and postsession perceptions reported on the completed qualitative feedback surveys. Difference scores were calculated for student's perceived understanding pre- and postsession, and these were compared between the two groups with a two-tailed unpaired t test. Their qualitative comments were read and coded thematically by one of the investigators. Significance was set at alpha level 0.05 for all statistical analyses. Significance and effect size were performed using Microsoft Excel (Microsoft Corp.,

Redmond, WA). Power analysis was performed using G★Power (11).

Hardware

Augmented reality headset: Microsoft HoloLens (Microsoft Corp., Redmond, WA)

Screen-based viewer: Dell Precision 7720 Mobile Workstation with Intel Core i7 CPU, 17.3 inch screen, 32GB RAM, Nvidia Quadro P5000 GPU (Dell, Round Rock, TX).

Software

Augmented reality 3D visualization and laptop screen manipulation software: SurgicalAR (Medivis Inc., Brooklyn, NY), viewed through HoloLens for the AR group and on a Dell Laptop screen for the screen group.

SurgicalAR is a software used for display of medical images and can be used for image manipulation, basic measurements and 3D visualization (multiplanar reconstructions and 3D volume rendering).

RESULTS

Objective Measures: Pre- and Posttest Results

Two groups of 15 participants each, for a total of 30 students, completed the study (15 women, 15 men, average age of the participants of 24 years, ranging 21−31 years). The pretest confirmed an overall lack of head and neck anatomy knowledge among the medical students in both groups; AR group mean 2.93 out of 5 (59% correct), SD 0.96 vs. screen group mean
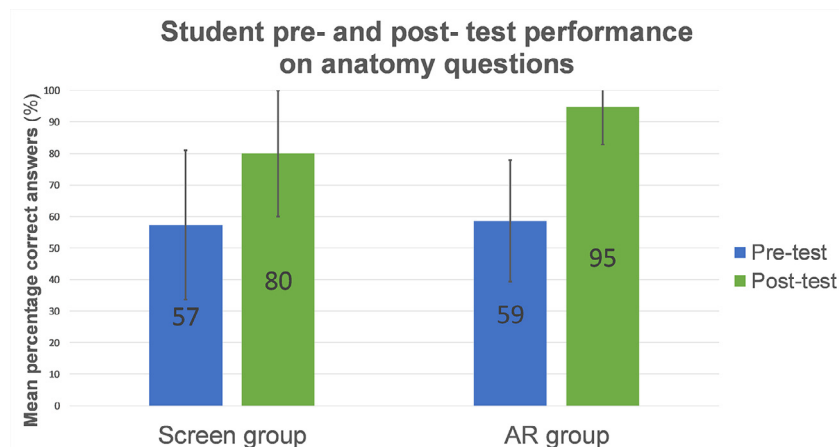
Figure 3. Student performance on the pre- (blue) and post- (green) test consisting of NBME-style anatomy questions, showing mean percentage correct for the screen group and for the AR group. (Color version of figure is available online.)

2.86 out of 5 (57% correct), SD 1.19. Mean pretest scores were not significantly different between groups ($p = 0.87$).

The AR group's posttest score mean was 4.73 out of 5 (95% correct) compared to 4 out of 5 (80% correct) for the screen group (Fig 3). In pretest vs. posttest mean comparison, the AR group showed significant in–group improvement ($p < 0.001$, effect size = 1.87), noting a larger effect size than the screen group, which showed significant in–group improvement ($p = 0.021$, effect size = 0.95) as well. Between–group analysis comparing the AR group and the screen group indicated posttest performance was significantly better in the AR group ($p = 0.022$, effect size = 0.73). Post hoc achieved power was 0.65.

Subjective Measures: Feedback Survey

All 15 of 15 (100%) students in the AR group returned the survey after a maximum of two reminder emails compared to 14 of 15 (93%) students in the screen group. For both groups, students reported that their perceived understanding of the anatomic relationships explored, had significantly improved after the session compared to before the session (AR $p < 0.001$, screen $p < 0.001$) (Table 1). There was no significant difference between the AR and the screen groups' difference scores, measuring the level of increase in student's perceived understanding ($p = 0.62$). As many as 11 of 15 (73%) students in the AR group and 12 of 14 (86%) students in the screen group described the technology as efficient/"high-yield." Also, 13 of 15 (87%) students in the AR group and 13 of 14 (93%) students in the screen group wanted this type of session integrated into their formal curriculum. Generally, the AR group provided lengthier and more varied responses and used more emphatic language than the screen group. Themes have been arranged by positive or negative emotional valence (Table 2).

DISCUSSION

This study demonstrates significantly superior performance on NBME–style anatomy questions in students who use 3D augmented reality for learning compared to students who use a 2D computer screen. Our findings also provide evidence to support the successful implementation of a mixed-media module to teach anatomy-radiology correlation and interpretation. Both the AR– and screen–based educational sessions were highly popular with students. In written feedback, the AR group provided lengthier and more varied responses and used more emphatic language. These findings add to results from similar studies which have shown increased benefit from the use of holograms or stereoscopic viewers when compared to 2D screens, and improved performance after the integration of enhanced visualization techniques (12−18). Our findings extend that concept to include augmented reality software and headsets which can allow for quick

---

TABLE 1. Student's Perceived Understanding of Neuroanatomical Structures

	Presession Mean Score	Postsession Mean Score	In-Group Comparison (Pre- Vs Postsession Mean Score)
AR group	1.07, SD 0.70	3.07, SD 0.80	$p < 0.001$*
Screen-based group	0.79, SD 0.70	3.21, SD 0.80	$p < 0.001$*
Between-group comparison (AR vs Screen-based Difference Score)		$p = 0.62$	

* Significant at < 0.001. SD = standard deviation.
Student's perceived understanding of neuroanatomical structures before (presession) and after (postsession) exploring them in the teaching session, as reported on a scale from 0 to 5 (0 = not at all, 1 = vaguely, 2 = mostly understood with some misconceptions, 3 = well, 4 = very well, 5 = ready for exam).

TABLE 2. Representative Examples of Positive and Negative Feedback From Both the AR and Screen Groups

	AR Group	Screen Group
Positive Valence	*"AR is immersive and fun. I'm able to literally travel down the nasopharynx and oropharynx into the trachea along the axial plane…coronal…and sagittal plane... I've learned a lot more with these images."*	*"As physicians we will be viewing or interpreting many of the structures we learn about in anatomy via radiology image…so it is important to be comfortable navigating the human body via radiology rather than just cadavers or idealized illustrations alone."*
	"It's much easier to get a sense of orientation of structures relative to each other from the AR tech than from an anatomy textbook … it's hard to have a sense of things like relative size of structures and relative distances when you're viewing anatomy one page at a time."	*"It is extremely helpful to be able to explore complex spatial relationships and practice spatial reasoning skills using radiology images."*
Negative Valence	*"It was hard to keep my head perfectly still long enough to view the structure of interest."*	*"Since it's identification of structures, I believe that's something that can be done individually."*
	"It took longer for the instructor and student to identify the structures they were referring to, rather than just physically pointing them out on the cadaver."	

translation of DICOM data into educational 3D models. Previous research has suggested that advanced visualization is particularly helpful for the bottom quartile of medical students; as medical educators adapt to entering first-year student classes with a range of scientific and nonscientific backgrounds, this tool may become particularly useful (4,6).

Our study compared 3D augmented reality with 2D screen-based visualization of the same DICOM data set and still detected a difference in student performance following an otherwise identical mixed-media educational module. The factors underlying the observed findings are likely multifactorial. The survey feedback from the students in our study, in conjunction with prior literature on the topic, suggests that potential mechanisms behind the improved effect from AR visualization may include: (1) increased student control and personal motivation in the learning material given an interactive environment, and (2) improved spatial understanding from stereoscopic depth cues provided by AR, but absent in screen-based or printed visualization (19−22). Other research has shown there to be a decreased cognitive load while localizing anatomic structures in 3D as compared to 2D environments (12,15), which may additionally explain the success of these methods for lower-performing students who may feel frustrated by higher cognitive-load methods.

The usefulness of 3D physical models has been previously well-established; however, it can be costly to obtain and store them and they cannot be readily used to teach radiologic-anatomic correlation. Augmented reality platforms may allow educators to create an immersive experience similar to viewing a physical model but allowing for radiologic overlays and direct radiologic-anatomic correlation. Medical schools or radiologic education departments would have to make the financial investment in an augmented reality capable headset and the software to display the CT data. As in our study, this cost could be partly offset by the use of "near-peer" instructors such as senior medical students or residents, instead of faculty. Near-peer instructors are favored by many millennial learners and could additionally facilitate an increased instructor-to-student ratio (1,23,24).

Limitations

There are limitations in this study that should be noted. Because the pretest and posttest were separated by the short period of time that it took to complete the educational session, the demonstrated gain reflected short-term recall rather than long-term knowledge acquisition. Further studies incorporating a longer exposure to immersive visualization and correlation with class performance could address this issue better. As the same five questions were used, a practice effect may account for a portion of the score improvement, although it would be expected to affect all groups at the same level. This study was limited by resident and student time constraints; however, increasing the number of both participants and questions on each test would more accurately and reproducibly assess student mastery of material. In terms of participant selection, because students volunteered for participation, those students interested in spending time outside of class in an educational session may have had a partiality toward technology, or an above-average dependence on visualization for learning. Participants not assigned to the AR system may also have been disappointed, which could also potentially reflect their subsequent feedback. These represent possible sources of bias in this study. Finally, the participating students all attended the same large and selective U.S. medical school

and may not be representative of all medical students globally. Larger and multi–institutional trials would improve the representativeness of the sample.

CONCLUSIONS

As seen in this study, immersive 3D visualization has the potential to improve anatomic recall of the head and neck, as well as engage millennial learners to better learn radiologic correlation in the anatomy laboratory. Augmented reality viewers and the software to support translation of DICOM images into immersive viewing are rapidly advancing (25−28). This study demonstrates that there is value in fur–ther development of AR technology for radiologic-anatomic education, and in continued investigation and comparative analysis of AR-based learning modules.

CONFLICT OF INTEREST

Authors declare that they have no conflicts of interest.

ACKNOWLEDGMENTS

REFERENCES

1. Ruzycki SM, Desy JR, Lachman N, et al. Medical education for millenni-als: how anatomists are doing it right. Clin Anat 2019; 32:20–25.
2. Benninger B, Matsler N, Delamarter T. Classic versus millennial medical lab anatomy. Clin Anat 2014; 27:988–993.
3. Sugand K, Abrahams P, Khurana A. The anatomy of anatomy: a review for its modernization. Anat Sci Educ 2010; 3:83–93.
4. Miller M. Use of computer-aided holographic models improves perfor-mance in a cadaver dissection-based course in gross anatomy. Clin Anat 2016; 29:917–924.
5. McMenamin PG, Quayle MR, McHenry CR, et al. The production of ana-tomical teaching resources using three-dimensional (3D) printing tech-nology. Anat Sci Educ 2014; 7:479–486.
6. Moro C, Stromberga Z, Raikos A, et al. The effectiveness of virtual and augmented reality in health sciences and medical anatomy. Anat Sci Educ 2017; 10:549–559.
7. Barsom EZ, Graafland M, Schijven MP, et al. Systematic review on the effectiveness of augmented reality applications in medical training. Surg Endosc 2016; 30:4174–4183.
8. Piromchai P, Avery A, Laopaiboon M, et al. Virtual reality training for improving the skills needed for performing surgery of the ear, nose or throat. Cochrane Database Syst Rev 2015:CD010198.
9. Gurusamy KS, Aggarwal R, Palanivelu L, et al. Virtual reality training for surgical trainees in laparoscopic surgery. Cochrane Database Syst Rev 2009:CD006575.
10. Nagendran M, Gurusamy KS, Aggarwal R, et al. Virtual reality training for surgical trainees in laparoscopic surgery. Cochrane Database Syst Rev 2013:CD006575.
11. Faul F, Erdfelder E, Lang A-G, et al. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical scien-ces. Behavior Research Methods 2007; 39:175–191.
12. Hackett M, Proctor M. The effect of autostereoscopic holograms on ana-tomical knowledge: a randomised trial. Med Educ 2018; 52:1147–1155.
13. Wainman B, Wolak L, Pukas G, et al. The superiority of three-dimensional physical models to two-dimensional computer presentations in anatomy learning. Med Educ 2018; 52:1138–1146.
14. Yeo CT, Ungi T, U-Thainual P, et al. The effect of augmented reality train-ing on percutaneous needle placement in spinal facet joint injections. IEEE Trans Biomed Eng 2011; 58:2031–2037.
15. Foo JL, Martinez-Escobar M, Juhnke B, et al. Evaluating mental work-load of two-dimensional and three-dimensional visualization for anatomi-cal structure localization. J Laparoendosc Adv Surg Tech A 2013; 23:65–70.
16. Peterson DC, Mlynarczyk GS. Analysis of traditional versus three-dimen-sional augmented curriculum on anatomical learning outcome measures. Anat Sci Educ 2016; 9:529–536.
17. Yammine K, Violato C. A meta-analysis of the educational effectiveness of three-dimensional visualization technologies in teaching anatomy. Anat Sci Educ 2015; 8:525–538.
18. Ebner F, De Gregorio A, Schochter F, et al. Effect of an augmented reality ultrasound trainer app on the motor skills needed for a kidney ultrasound: prospective trial. JMIR Serious Games 2019; 7:e12713.
19. Kugelmann D, Stratmann L, Nuhlen N, et al. An Augmented Reality magic mirror as additive teaching device for gross anatomy. Ann Anat 2018; 215:71–77.
20. Ferrer-Torregrosa J, Jimenez-Rodriguez MA, Torralba-Estelles J, et al. Distance learning ects and flipped classroom in the anatomy learning: comparative study of the use of augmented reality, video and notes. BMC Med Educ 2016; 16:230.
21. Chytas D, Johnson EO, Piagkou M, et al. The role of augmented reality in anatomical education: an overview. Ann Anat 2020:151463.
22. Ma M, Fallavollita P, Seelbach I, et al. Personalized augmented reality for anatomy education. Clin Anat 2016; 29:446–453.
23. Johnson EO, Charchanti AV, Troupis TG. Modernization of an anatomy class: from conceptualization to implementation. A case for integrated multimodal-multidisciplinary teaching. Anat Sci Educ. 2012; 5:354–366.
24. Evans DJ, Cuffe T. Near-peer teaching in anatomy: an approach for deeper learning. Anat Sci Educ 2009; 2:227–233.
25. Bork F, Stratmann L, Enssle S, et al. The benefits of an augmented reality magic mirror system for integrated radiology teaching in gross anatomy. Anat Sci Educ 2019; 12:585–598.
26. Contreras Lopez WO, Navarro PA, Crispin S. Intraoperative clinical appli-cation of augmented reality in neurosurgery: a systematic review. Clin Neurol Neurosurg 2019; 177:6–11.
27. Meola A, Cutolo F, Carbone M, et al. Augmented reality in neurosurgery: a systematic review. Neurosurg Rev 2017; 40:537–548.
28. Sotgiu MA, Mazzarello V, Bandiera P, et al. Neuroanatomy, the Achille's heel of medical students. A systematic analysis of educa-tional strategies for the teaching of neuroanatomy. Anat Sci Educ. 2020; 13:107–116.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.acra.2020.07.008.