Journal of Medical Robotics Research, Vol. 0, No. 0 (2015) 1–12

Augmented reality training platform for neurosurgical burr hole localization

Zachary M. C. Baum*, Andras Lasso*, Sarah Ryan*, Tamas Ungi*,†,
Emily Rae*, Boris Zevin†, Ron Levy†,‡ , Gabor Fichtinger*,†

*Laboratory for Percutaneous Surgery, School of Computing, Queen's University, Kingston, Ontario, Canada
E-mail: zac.baum@queensu.ca

†Department of Surgery, School of Medicine, Queen's University, Kingston, Ontario, Canada

‡Center for Neuroscience Studies, Queen's University, Kingston, Ontario, Canada

Augmented reality (AR) is used in neurosurgery to visualize lesions and plan procedures pre-operatively and intra-operatively, though its use has not been widely adopted in simulation-based neurosurgical training for the same tasks. This work defines metrics to determine performance in drill position and angle identification for neurosurgical training. The metrics were validated intra-operatively and in a simulated training environment, demonstrating that trainees identify drill position and angle faster and more accurately with AR compared to standard techniques. Training using AR and the proposed metrics stands to add value to neurosurgical curricula development.

Keywords: Augmented reality, competency-based medical education, holography, neurosurgery, surgical planning, visualization.

1. Introduction

Neurosurgery encompasses a variety of different procedures, all of which vary in their use of intra-operative imaging in current clinical practice. Brain tumor biopsies, subdural hemorrhage evacuations, ventriculostomies, and many other procedures attain intracranial access through what is called a burr hole. In procedures where the target is easily distinguishable in the patient's imaging, neurosurgeons begin by planning and identifying the optimal trajectory. The created plan for creating a neurosurgical burr hole encompasses two aspects; the entry point and the entry trajectory. The entry point is where the burr hole will be placed on the patient's skull. The entry trajectory is the angle at which the burr hole will be drilled – this is effectively the direction in which they will enter the skull from the planned entry point. It is critical that the entry point and entry trajectory are safe and effective for the given procedure – all the while avoiding critical structures in the patient's brain to reduce the risk of any post-operative complications. Ensuring that a safe and effective plan is determined prior to surgery is paramount to patient survival and surgical outcomes as each procedure is different and requires a specific approach to ensure satisfactory treatment.[1] Furthermore, this plan should be chosen in such a way that it also minimizes the size of the skull opening – ideally through the use of a burr hole craniostomy. Burr hole craniostomies have the best cure to complication ratio, are considered the safest and most effective when compared to twist drill craniostomy or craniotomy.[1] Additionally, minimizing the size of the skull opening can aid in providing a cosmetically good outcome for the patient. The ability to effectively and correctly plan and identify the optimal drill location and drill angle of a burr hole for a given procedure is a fundamental element of a neurosurgeon's skillset and a core piece of neurosurgical training curricula.

If the target anatomy is clearly enhanced in the patient's imaging, the principal challenge for a neurosurgeon is to mentally transfer their planned drill path from what they can see in the patient's images to the physical patient. As such, some procedures have become increasingly reliant on the use of medical imaging as a tool for determining pre-operative plans and intra-operative guidance. In some cases, neuronavigation systems can help to determine drill location and drill angle.[2] However, these methods still rely on a surgeon's ability to interpret, reconstruct and visualize two-dimensional (2D) medical images into three dimensions (3D).[3–6] These tasks rely heavily on surgical knowledge, experience, and spatial reasoning skills.

Currently, neurosurgical trainees learn these planning and spatial reasoning skills through apprenticeship inside and outside of the operating room for approximately 6-8 years after completing medical school. Much of this training is complex, hands-on, and leaves trainees able to acquire fundamental skills only when specific procedures occur in a clinical context.

Of late, medical education has begun moving towards the model of competency-based medical education (CBME). CBME allows trainees to progress through given curricula at their own pace, and to proceed past the current curriculum once they have demonstrated competency in specific, objective benchmarks and metrics.[7] The CBME model ensures that trainees who have not reached competency cannot provide care without supervision in a clinical setting and during their interactions with patients.

One of the challenges associated with CBME is the need for ongoing tracking of an individual's learning curve through objective measures as they progress towards competency.[8] Recently, to ensure compliance with CBME, there has been a shift towards methods for quantitative assessment of skills, often using external position tracking, as this does not require direct expert supervision.[9] In the specific case of neurosurgical applications, in addition to the difficulties in tracking progress, there is a lack of demonstrated neurosurgical performance metrics for planning which translate to successful patient and surgical outcomes and are usable for providing trainees with meaningful feedback.[9-10] This leaves trainees unable to independently train or develop core skills such as planning and identifying optimal drill locations and drill angles on simulated or real patients. As such, a neurosurgical curriculum which follows the CBME model may prove important for the training and skill development of future neurosurgical trainees. Practice with simulation-based training platforms in other surgical specialties has been heralded as an effective learning strategy and has been thought to be the next step for neurosurgical training curricula.[10-11]

Augmented reality (AR) is a combination of visualization and imaging technologies which superimpose data and images into the real world, eliminating the need to mentally reconstruct or project images onto patients. AR images can be coupled with other real-time data sources such as cameras,[5] endoscopes,[5] fluoroscopes,[5] operating microscopes,[12] or even displayed over a movable tablet computer in various neurosurgical contexts.[13] AR is already being used in neurosurgery to aid in the visualization of intra-cortical lesions,[12-14] hemorrhages,[15] and hydrocephalus.[5,16] Additionally, when manifested in the form of head-mounted displays (HMD), AR has shown to be beneficial for surgical planning and visualizations.[5,17–20] AR HMD are well positioned to address these problems given their ability to display three-dimensional (3D) anatomical models, imaging, and surgical data aligned with the patient and in the user's view.

The *Microsoft HoloLens* (Microsoft Corp., Redmond, Washington, USA) is a cost-effective, light-weight mixed reality platform that is considered the highest performing commercially available AR HMD platform based on its capabilities for contrast perception and frame rate as well as ergonomics,[25-26] among multiple other factors.[6,26] The *HoloLens* is a fully untethered holographic computer which combines various sensors such as accelerometers, infrared lasers, microphones and cameras into a wearable headset capable of generating 3D visualizations through a reflection on to the user's retinas, all without impeding their view of the surrounding environment. As such, the *HoloLens* has been used to provide hands-free holographic visualizations in neurosurgical applications, [21–24] though it has not yet been fully utilized for demonstrating neurosurgical performance metrics which may translate to successful patient or surgical outcomes.

Most AR systems, especially AR HMD systems such as the *HoloLens*, are in early stages of development and use compared to virtual reality (VR) systems – many of which have been shown to be capable of differentiating different levels of competence and skill in neurosurgery through proven curriculums.[27-30] VR simulation-based training systems typically comprise of a simulated intra-operative environment which contain all the components a surgeon would encounter in the procedure with additional 3D and virtual visualizations that are specific to the procedure.[29-30] These systems also include objective performance and outcome metrics,[27-29] or black-box algorithms which produce a final score or grade for the user.[30] These metrics involve quantifying performance throughout the procedure, such as tissue removed, tool path length or duration of excessive force, and have less emphasis on measuring performance during the planning phase of the procedure.[27-30]

At present, it seems clear that VR allows for simulated rehearsal of surgical procedures and to assess performance in doing so. Contrarily, AR is used for the visualization or projection of additional data to ensure a surgeon's eyes are kept looking into the surgical field and typically does not allow for the reporting of performance metrics. However, AR HMD still has the potential to benefit and improve current surgical and neurosurgical practice. As prospective and clinical feasibility and application studies have been limited,[10] AR has yet to see routine or wide-spread success as ideal applications have not yet been widely demonstrated in clinical use as much of the focus has instead been on the evaluation of core AR technologies and on phantom studies using AR.[13-14,16-17,20-21,23,31] Therefore, whether its use may provide improved outcomes remains unclear,[32] as studies to prove the effectiveness of AR in training and clinical scenarios are still needed.[31]

Our work sought to determine relevant, valid, objective, and transparent performance metrics which are CBME compliant and are usable for differentiating between novices and experts in the planning process for neurosurgical procedures. Furthermore, we seek to use these metrics to determine whether the use of HMD AR adds practical value to teaching and planning of neurosurgical procedures. By comparing the use of AR HMD to standard practice in three common neurosurgical procedures that vary in their use of neuronavigation, the utility of AR HMD was assessed in trainees and surgeons to determine the role of expertise in this technology. Additionally, as many previous works have focused solely on implementation of an AR HMD system and not on determining the effectiveness of AR for training - we further assessed the effectiveness of AR as a tool for delivering training programs. This assessment serves as an aid for determining what components of AR make it a suitable platform for a neurosurgical training curriculum based on the experience that AR HMD provide users.

This paper shows that AR HMD allows trainees to identify drill locations and drill angles more rapidly and with improved accuracy compared to standard techniques. Consequently, our aim was to examine the usefulness of AR HMD for visualizing neurosurgical lesions in clinical and simulated environments.

2. Methodology

In this work, we assess the effectiveness of AR HMD, namely the *HoloLens*, for neurosurgical planning and training. We present an intra-operative performance study to compare the benefit of AR for surgical planning to conventional methods for trainees and attending neurosurgeons.

Through this study, we aimed to determine the validity of our metrics for assessing trainee performance and attending neurosurgeon performance in localizing optimal drill location and drill angle using only 2D medical images or AR HMD technology capable of displayed 3D models. We experimentally evaluated if our metrics can differentiate between trainees and attending neurosurgeons in that same task. Additionally, we hoped to determine what type of participant stood to most benefit from AR HMD technologies for surgical planning.

As trainees most benefitted from our intra-operative study, we next conducted a simulated training study. This study was completed to aid in guiding the development of a neurosurgical training curricula by not only assessing performance, but also user experience for this type of task in a simulated environment. In addition to 2D medical images and AR HMD, in this study we provided 3D visualizations to participants to closely emulate the current clinical standard of neuronavigation systems. Here, trainees demonstrated their ability to identify optimal drill location and drill angle in space using our AR application, as well as two other conventional visualization methods. In a simulated training environment, we prove that displaying optimal surgical plans significantly aids users in identifying optimal drill location and drill angle.

2.1. *System Design and Implementation*

We designed *HoloQuickNav*, an application usable for intra-operative AR planning in neurosurgery procedures using the *HoloLens*. *HoloQuickNav* was developed using the cross-platform Unity engine (version 2018.1.0f2) for augmented reality and virtual reality software. Multiple components from Microsoft's open-source and cross-platform Mixed Reality Toolkit were incorporated in our application in order to accelerate development. All other remaining components and code were created specifically for *HoloQuickNav*.

Rae *et al*. previously assessed and investigated the use of holographic models displayed on the *HoloLens* for localization of burr holes in craniostomy procedures.[33] In these procedures, clinicians typically use a drill bit which is 4 mm in diameter, with the aim of successfully identifying the location of the drill entry point within a clinically acceptable range of 10 mm of the pre-planned target.[33] *HoloQuickNav* was validated using a similar protocol to [33], wherein multiple novice ($n_{\text{novice}} = 10$) and expert ($n_{\text{expert}} = 3$) users completed multiple registrations ($n_{\text{registration}} = 12$ per participant) on male and female phantoms. Six registrations occurred on the male phantom, three of which were performed with hair and three of which were performed without. The registration protocol for the female phantom was identical. Novice users could register the models within the acceptable range 94% of the time, while experts were able to do so 97% of the time. Overall, registrations were accurate < 5 mm 65% of the time when performed by an expert user.

To navigate menus and register holographic images to the patient, an *Xbox One Wireless Controller* (Microsoft Corp., Redmond, Washington, USA) was used to aid in navigating through the application. A handheld controller was chosen for our system as previous work with voice commands and *HoloLens* 'AirTap' gestures did not provide enough usability in our application.[33]

Anatomical models of a patient's skin surface, brain, and intra-cortical lesion can be generated from magnetic resonance (MR) or computerized tomography (CT) imaging using *3D Slicer*'s *Segmentation* module.[34] Users are able to translate or rotate the holographic model interchangeably in each direction or about a chosen axis to register holographic models to the patient using surface features through a manual alignment process (Fig. 1).

2.2. *Intra-operative Target Localization Study*

The intra-operative planning and target localization study was conducted as a prospective cohort study. The study was approved by the Queen's University Health Sciences and Affiliated Teaching Hospitals Research Ethics Board. Voluntary enrollment and signed consent were obtained prior to each procedure from patient subjects as well as from each attending neurosurgeon and trainee participant. Clinical feasibility of the *HoloQuickNav* system and software was tested in an operating room environment with human patients. This

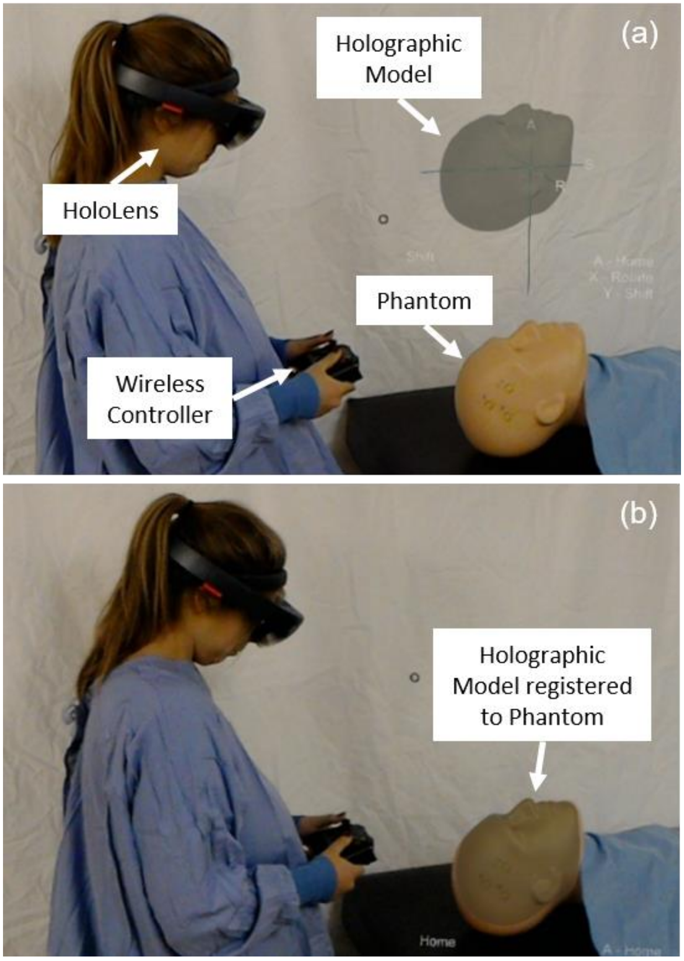

Fig. 1. The registration for the AR method with (a) a user translating models towards a phantom and (b) models aligned with the phantom.

work sought to assess whether the use of *HoloQuickNav* better informs surgeons and trainees when localizing optimal drill location and drill angle in neurosurgical procedures. Additionally, it sought to validate if the developed metrics were usable for differentiating between trainee and attending neurosurgeon performance in the localization of optimal drill location and drill angle.

Medical students, general surgery residents and attending neurosurgeons were recruited from the Queen's School of Medicine, Queen's University General Surgery residency program and Kingston General Hospital Department of Surgery, respectively. Medical students and residents were considered trainees for this study. The medical students and residents (all were in postgraduate year 1 or 2) had limited previous exposure to surgical planning and limited previous experience in the task. Trainee and attending neurosurgeon outcomes were considered separately in the study as knowledge expectation in their neurosurgical planning experience differs.

Fifteen cases which underwent intracranial surgical intervention for drainage of chronic subdural hemorrhage, brain tumor resection or insertion of external ventricular drain at Kingston General Hospital in 2018 were included in our study.

Preoperative patient CT or MR images were used to provide the visualizations for surgical planning. All images and DICOM data were imported into a workstation with all identifying patient data previously removed. Patient images were reviewed by experienced technicians and attending neurosurgeons to provide expert opinion on the resulting image segmentations and holographic models which were displayed to the participants during the study.

Each case had both the attending neurosurgeon and the trainee or trainees plan their procedure using a '2D method' and an 'AR method' while in the operating room. The attending neurosurgeon would later complete the surgery using the available neuronavigation system or other standard protocols if guidance using the neuronavigation system was not possible. The 2D method has the participant denote the location of the lesion and the trajectory using a custom 3D-printed pointer tool, with only a series of 2D patient images available for reference. The AR method has the participant denote the location of the lesion and the trajectory, using the same pointer tool as in the 2D method, while wearing a *HoloLens* running *HoloQuickNav*. The 2D method requires participants to use the preoperative patient CT or MR images, which they can browse through on the workstation available in the operating room. The AR method shows participants holographic models of the surface anatomy and intra-cortical lesion virtually floating over the patient.

Prior to using the AR method for the first time in the operating room, participants were shown how to use the technology with a synthetic sample case using a mannequin head as a phantom. This case allowed them to learn how to interpret, manipulate and register the holographic models. During this sample case, it was explained to the trainees and attending neurosurgeons that they should report drill locations and drill angles which were the intended path of the catheter in subdural hemorrhages and hydrocephalus, or in the direction of the desired center of the craniotomy for a brain tumor.

Once the patient was brought into the operating room, the patient was anesthetized and positioned for surgery. The trainee reviewed the patient's CT or MR images and then used the 2D method to identify the lesion and trajectory. The reported trajectory was quantified by placing the pointer at the entry site for the subdural hemorrhage and hydrocephalus, or at the desired center of the craniotomy for brain tumors. The pointer was oriented along this desired trajectory and several coloured 3D point clouds of the scene were acquired using the *Intel RealSense D415 Depth Camera* (Intel Corp., Santa Clara, California, USA) to determine the drill location and drill angle which were identified by the trainee. Next, the attending neurosurgeon used the 2D method to identify the lesion and trajectory. Point clouds of the scene were acquired to quantify the attending neurosurgeon's planned drill location and drill angle using the 2D method.

Holographic models were registered to the patient by an expert user (Fig. 2). The trainee then used the AR method to identify the lesion and trajectory. Point clouds were acquired of the trainee's planned drill location and drill angle using the AR method (Fig. 3). Next, the attending neurosurgeon used the AR method to identify the lesion and trajectory. Point clouds were acquired of the attending neurosurgeon's planned drill location and drill angle using the AR method.

Lastly, the neuronavigation system was registered to the patient using the protocol required by the specific system used, which is typically a point-based registration, followed by a surface-based registration. The attending neurosurgeon used a frameless neuronavigation system, such as the actively tracked *NAV3i* (Stryker Corp., Kalamazoo, Michigan, USA) or the passively tracked *Brain Lab VectorVision* (Brainlab, Munich, Germany) – both of which are capable of accuracy < 1.5 mm in practice.[35] The neuronavigation systems were then used to identify an optimal drill location and drill angle. Point clouds are acquired of the attending neurosurgeon's planned drill

Fig. 2. Neurosurgeon using the *Xbox One Wireless Controller* while wearing the *HoloLens* to register the models to a patient.

Fig. 3. Trainee (right) using the *HoloLens* to indicate the drill location and drill angle as a 3D point cloud of the scene is acquired using the *Intel RealSense D415 Depth Camera* (left).

Fig. 4. Aligned 3D point clouds shown in *MeshLab*. The patient's face and identifying features, though relevant for alignment process are blurred in the image.

location and drill angle using the standard neuronavigation system. This drill location and angle was used as the clinical gold-standard. In two procedures, the neuronavigation system was not used for the surgery. In one procedure, this was due to poor quality of preoperative imaging. In the other procedure, the neuronavigation system was deemed not necessary given the limited complexity of the procedure. In these instances, the attending neurosurgeon's drill location and drill angle planned using the 2D method was defined as the gold-standard as this trajectory would be used later, during the procedure.

The point clouds containing each denoted drill location and drill angle from each trainee and attending neurosurgeon, as well as the clinical gold-standard were processed using *MeshLab* (ISTI-CNR Research Center, Pisa, Italy), a 3D mesh processing software system for processing, editing and cleaning large unstructured meshes. Each of the point clouds had any non-essential points removed (i.e. additional people, walls, equipment which were not required for determining the pointer location on the patient) and were subsequently converted to meshes to allow for simpler registration within *MeshLab*. Once all captured 3D point clouds had been converted into a mesh and had been filtered appropriately, each mesh was registered to the mesh containing the surgical gold-standard using *MeshLab*'s Align tool (Fig 4). This alignment was performed using only the facial features in the mesh; any surgical equipment, the pointer tool, or participant hands were excluded. All roughly aligned meshes were then further aligned using *Meshlab*'s automatic alignment process. The average error reported by *Meshlab* was under 0.1 mm for all sets of meshes. The patient's skin surface model was then aligned to the meshes so that the pointer-tip and the end of the pointer shaft could be annotated relative to the skin surface. A single vertex of the mesh which was closest to the pointer-tip – corresponding to a point on the patient's skin surface model – and the end of the pointer shaft were manually selected by inspection of each individual mesh. These models and annotations were exported from *MeshLab* into *3D Slicer*, where all results were computed.

Each denoted pointer-tip location and the angle resulting from the vector formed from the pointer-tip and the end of the pointer shaft were then compared to the pointer-tip location and the angle resulting from the vector formed by the pointer-tip and the end of the pointer shaft as in the clinical gold-standard. The comparison was based on four metrics, where two of the metrics measured distances and two measured angles (Fig. 5). Our four metrics - i) drill-tip distance, ii) distance to lesion, iii) drill angle error, and iv) angle to lesion - were computed based on the geometric properties of the trajectory from each trainee and attending neurosurgeon, and not based on their clinical feasibility or likelihood for surgical success. This was done as

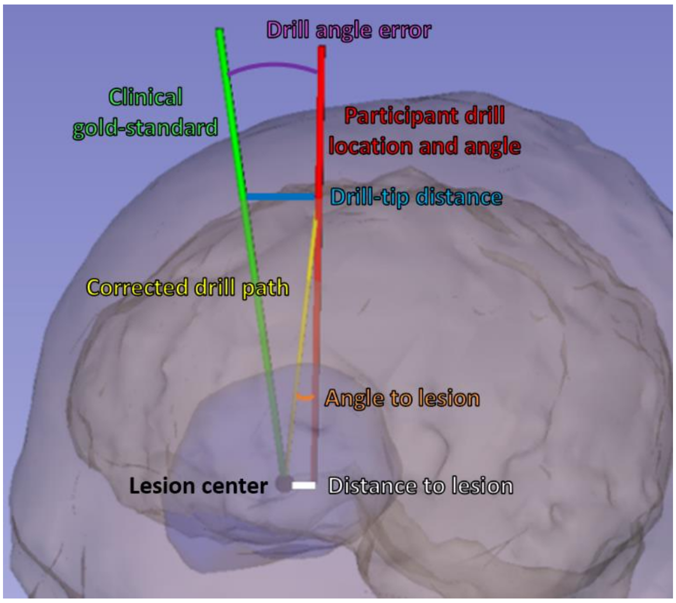

Fig. 5. 3D models shown in *3D Slicer* of surface anatomy, brain, intra-cortical lesion, and user defined trajectories from one localization task in the simulated study. The *black point* shows the lesion center of mass; the *green line* shows the gold-standard access point and trajectory; the *red line* shows the participant's access point and trajectory; the *yellow line* shows the trajectory from the participant's access point to lesion's center; the *blue line* shows the drill-tip distance; the *white line* shows the distance to lesion; the *purple arc* shows the drill angle error*;* the *orange arc* shows the angle to lesion.

we sought to assess how trainees performed and selected trajectories compared to the clinical gold-standard at the time of the procedure.

The drill-tip distance, shown in Eq. (1), measures the distance between participant drill-tip location and clinical gold-standard drill-tip location. Drill-tip distance was computed as the Euclidean distance between point *A*, the participant's pointer-tip location, and point *B*, the clinical gold-standard drill-tip location.

$$\sqrt{(A_x - B_x)^2 + \left(A_y - B_y\right)^2 + (A_z - B_z)^2} \tag{1}$$

The distance to lesion, shown in Eq. (2), measures the distance that the participant's trajectory was from intersecting the center of the lesion. Distance to lesion was computed as the closest distance between a line and a point. The line used was defined by point *X*, the participant's pointer-tip location, and point *Y*, the end of the pointer shaft. Point *Z* was defined as the center of the lesion determined by the clinical gold-standard.

$$\frac{|(Y - X) \times (X - Z)|}{|Y - X|} \tag{2}$$

The drill angle error, shown in Eq. (3), is measured as the angular error in the participant's trajectory relative to the clinical gold-standard. Drill angle error was computed as the angle between the vector *P*, defined by the participant's pointer-tip location and the end of the pointer shaft, and the vector *Q*, defined by the clinical gold-standard drill-tip location and the center of the lesion given by the clinical gold-standard.

$$\cos^{-1}\left(\frac{\vec{P} \cdot \vec{Q}}{\|\vec{P}\|\|\vec{Q}\|}\right) \tag{3}$$

The angle to lesion is measured as the angular error if the participant's pointer-tip was assumed to be optimal, in the participant's trajectory relative to the center of the lesion given by the clinical gold-standard. Angle to lesion was computed through the same method as drill angle error, using Eq. (3), where the resultant angle was between the vector *A*, defined by the participant's pointer-tip location and the end of the pointer shaft, and the vector *B*, defined by the participant's pointer-tip location and the center of the lesion given by the clinical gold-standard.

Completion time was not measured in this study. Though efficiency will be relevant for a fully-fledged training curriculum, it was not relevant in this study wherein we sought to determine if AR was suitable for creating a teaching platform and if the drill location performance metrics could be differentiated between trainees and attending neurosurgeons.

2.3. *Simulated Target Localization Study*

To assess the effectiveness of *HoloQuickNav* for training how to plan an optimal drill location and drill angle identification, we conducted a simulated patient study to assess the performance and effectiveness of our software compared to other conventional medical image visualization methods for delivering a suitable platform for neurosurgical training.

Seven medical student trainees were recruited to localize drill locations and drill angles on a phantom using three different visualization methods. All medical students had little or no prior simulated or clinical surgical experience.

The three visualization methods consisted of a computer display with 2D CT or MR images (2D method) (Fig. 6a), a computer display with 2D CT or MR images and a 3D visualization of the simulated patient's skin surface, brain, and intra-cortical lesion (3D method) (Fig. 6b), and a 3D holographic visualization of the simulated patient's skin surface, brain, and intra-cortical lesion shown on a *HoloLens* running *HoloQuickNav* (AR method) (Fig. 6c). The 3D method was additionally introduced in this study as it most closely resembles the visualizations given on neuronavigation systems that are used in clinical practice. Each visualization displayed an annotated surgical path, defined by an attending neurosurgeon, to be localized by the participants during the study. In the 2D images, the annotated surgical path was indicated by a red line shown on the images. In the 3D view and holographic visualizations, the surgical path was indicated by a model of a red cylinder.

An optical tool tracking setup was implemented using a series of four *OptiTrack Prime 17W* cameras (NaturalPoint Inc.,

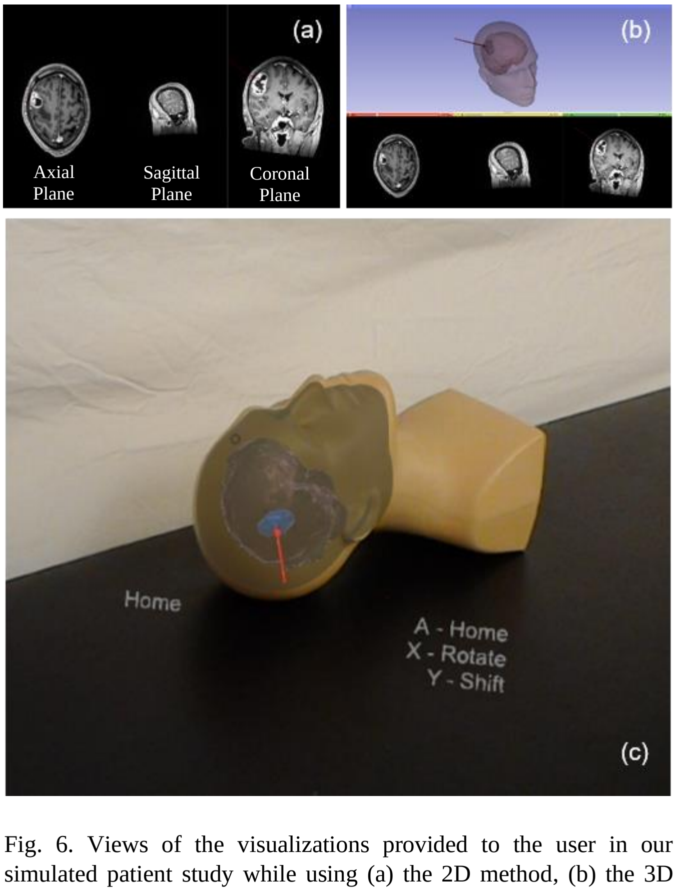

Fig. 6. Views of the visualizations provided to the user in our simulated patient study while using (a) the 2D method, (b) the 3D method, and (c) the AR method in one of the image series used in the study. Slice orientations for the visualizations are given in (a); images in (b) are presented in the same orientations.

Corvallis, Oregon, USA) to record the drill locations and drill angles during each trial. The *OptiTrack* system was used here as it simplified the experimental workflow and eliminated the need for any mesh alignments. Additionally, there were no restrictions on equipment which could be used in the simulated environment – unlike the previous study which occurred in an intra-operative environment. A series of optical tracking markers were affixed to the simulated patient and pointer tool to capture participant localizations. The PLUS Server Application software interface for hardware components was used to acquire and send tracking data to *3D Slicer* via the OpenIGTLink network protocol.[36-37] Within *3D Slicer*, all pose information was recorded using the *Sequence Browser* module within the *Sequences* extension.[38]

Participants were assigned a set of 14 image series; each of which represented a specific and different optimal trajectory are each of which were used with only one of the three visualization methods. All trials were completed in a random order. The images and associated models from each image series were deformably registered by the author conducting the study to a CT image series of the mannequin head prior to the study. This mannequin head was used as a phantom for the study to ensure that the patient images properly aligned with the phantom when participants were viewing the images on the computer screen or on the *HoloLens*. Participants were able to browse through the images and models displayed in the various visualization methods. Additionally, participants were timed and given a target time to complete each target localization of two minutes. They were not stopped if they had not completed the task within two minutes. In each trial, participants localized the drill location and drill angle to the best of their ability based on the information shown on the computer display or on the *HoloLens* using the optically tracked pointer tool.

Denoted drill locations and drill angles were compared to that which was defined by an attending neurosurgeon prior to the study. The comparison was based on the same metrics as in the previous study, with the addition of completion time to provide a sense of time pressure between the different visualization tasks for the participants. Following the study, participants answered a subjective multidimensional questionnaire to assess the workload and effectiveness of the 2D, 3D, and AR methods. This questionnaire was based on the NASA Task Load Index, a subjective and multidimensional assessment tool that rates the perceived workload for assessing a task's effectiveness.[39]

3. Results and Discussion

All statistical testing was calculated, and subsequent results were obtained using the *MATLAB Statistics and Machine Learning Toolbox* (The MathWorks Inc., Natick, Massachusetts, USA).

3.1. *Intra-operative Target Localization Study*

Differences between metrics from 2D and AR methods for trainees and attending neurosurgeons, as well as differences between trainees and attending neurosurgeon performance when using 2D and AR methods were tested using a two-tailed Mann-Whitney U test, where ($\alpha = 0.05$). Results for each metric are presented as median [minimum–maximum]. *P*-values are only given for significant results. Trainee (n = 15) and attending neurosurgeon (n = 15) performance between methods are summarized in Table 1 and performance between participant experience level are summarized in Table 2.

Table 1. 2D and AR method performance comparison.

	Metric	2D Method	AR Method	*p*
Trainee	Drill-tip distance [mm]	27 [13–68]	20 [2–44]	
	Distance to lesion [mm]	19 [5–49]	9 [1–43]	
	Drill angle error [°]	29 [9–84]	31 [11–57]	
	Angle to lesion [°] *	24 [4–67]	13 [1–53]	0.03
Attending Neurosurge	Drill-tip distance [mm]	12 [5–28]	9 [4–21]	
	Distance to lesion [mm]	11 [4–20]	8 [3–20]	
	Drill angle error [°]	20 [3–28]	16 [2–35]	
	Angle to lesion [°]	11 [5–27]	11 [3–36]	

* Indicates significance between 2D and AR metrics in a participant group.

Table 2. Trainee and attending neurosurgeon performance comparison.

	Metric	Trainee	Attending Neurosurgeon	*p*
2D Method	Drill-tip distance [mm] *	27 [13–68]	12 [5–28]	0.001
	Distance to lesion [mm]	19 [5–49]	11 [4–20]	
	Drill angle error [°]	29 [9–84]	20 [3–28]	
	Angle to lesion [°] *	24 [4–67]	11 [5–27]	0.009
AR Method	Drill-tip distance [mm] *	20 [2–44]	9 [4–21]	0.011
	Distance to lesion [mm]	9 [1–43]	8 [3–20]	
	Drill angle error [°] *	31 [11–57]	16 [2–35]	0.032
	Angle to lesion [°]	13 [1–53]	11 [3–36]	

* Indicates significance between trainee and attending neurosurgeon metrics for a given planning method.

The intra-operative study provides a proof of concept that the *HoloLens* and *HoloQuickNav* may have the potential to be used for identifying optimal drill locations and drill angles in neurosurgical procedures. While changes in performance were observed in the trainee group, only one of four of these differences were statistically significant. There were measurable changes between the neurosurgeon group as well, however, these changes were smaller in magnitude and not statistically significant. As such, while the *HoloLens* may provide neurosurgeons with enhanced visualizations of patient anatomy and potentially reduce the level of difficulty in determining the optimal drill location and drill angle to place a burr hole for a given procedure, our results show that it does not significantly affect their ability to form a surgical plan when compared to conventional methods.

Significant changes in performance are observed between trainees and attending neurosurgeons when using the same planning method, with two of four differences in metrics being statistically significant for both the 2D and AR methods. However, the differences between metrics which describe a participant's ability to be closer to the internal target (distance to lesion and angle to lesion) are not statistically significant. Upon further inspection of the performance of the attending neurosurgeons compared to the trainee performance, it seems that the trainees can perform similarly well, on average, when using the AR method. These results indicate that the HMD AR system allows the trainees to select a drill angle which gets them sufficiently near to the lesion center, but that their chosen drill location is just not optimal. As such, the trainees appear to be able to perform at a higher level intra-operatively using this technology but will need ongoing practice – likely in a proficiency-based curriculum – to achieve the level of expertise of an attending surgeon.

CBME compliant metrics must be relevant, valid, objective and transparent. From our study, it also is clear that our metrics are relevant. Creating a burr hole in any our target procedures is a key task wherein the learning objectives involve selecting the proper location of the drill site on the skull, and the key performance metrics require a burr hole placed in the acceptable region and an appropriate perforation angle at the surface.[40] The proposed metrics are valid, as they can differentiate between trainees and attending neurosurgeon skill levels, as seen in Table 2. The proposed metrics are objective, as they are computed directly from the geometric properties of the user's drill location and drill angle localization with respect to the optimal. As such, the reported metric is not influenced by the preceptor and relies solely on the user's ability Additionally, the metrics are transparent as they are simple to interpret geometrically and the concept surrounding the choice of metric and its calculation is clear. However, our defined metrics cannot measure surgical outcomes and have no defined threshold for determining success. Though they appear clinically relevant, it is not clear if the improvements in performance observed while using the *HoloLens* has any effect on the success of the procedure or to patient outcome.

Through the period in which the study was conducted, 19 cases met our study's inclusion criteria and were available for patient enrolment. Fifteen cases were attended, in the remaining four cases; one patient did not consent to participate, one case could not be completed due to technical problems with the *HoloLens*, and two cases could not be completed due to the absence of research personnel. As such, cases were not selected based on perceived difficulty or potential results. However, additional cases are required to determine whether experts can benefit from this technology by adapting it to their routine, or if it can only help less experienced operators in achieving better accuracy at an early phase of their learning curve.

3.2. *Simulated Target Localization Study*

A summary of participant (n = 7) metrics from target localizations (n = 14 per participant) using the 2D, 3D and AR methods are presented in Table 3. Results are presented as median [minimum–maximum].

Table 3. 2D, 3D, and AR method performance summary.

Metric	2D Method	3D Method	AR Method
Drill-tip distance [mm]	40 [8–156]	19 [3–171]	13 [2–26]
Distance to lesion [mm]	23 [5–80]	15 [1–58]	13 [1–23]
Drill angle error [°]	32 [2–173]	19 [1–151]	8 [1–19]
Angle to lesion [°]	25 [8–75]	23 [1–106]	15 [2–55]
Completion time [s]	134 [41–225]	76 [17–171]	43 [15–126]

Differences between metrics from different visualization methods were tested using the Mann-Whitney U test, using Bonferroni correction for multiple tests ($\alpha = 0.01$). Resulting *p*-values from each statistical test are presented in Table 4.

Table 4. Pairwise performance comparison of 2D, 3D, and AR method target localization metrics.

Metric	2D v. 3D	2D v. AR	3D v. AR
Drill-tip distance	0.003*	< 0.001*	0.008*
Distance to lesion	0.007*	< 0.001*	0.047
Drill angle error	0.004*	< 0.001*	< 0.001*
Angle to lesion	0.57	< 0.001*	0.034
Completion time	< 0.001*	< 0.001*	< 0.001*

* Indicates significance at the Bonferroni corrected alpha value.

Participants localized a drill location and drill angle. Participants localized these on the correct side of the phantom's head in 67% of tasks using the 2D method, 97% of tasks using the 3D method, and 100% of tasks using the AR method. Furthermore, participants localized the drill location and drill angle within the target time of two minutes in 55% of tasks using the 2D method, 88% of tasks using the 3D method, and 97% of tasks using the AR method. These values are particularly encouraging when looking not only at the skill of drill location and drill angle localization but at one which is even more fundamental; that of correctly interpreting the orientation of a series of CT or MRI images. As only 67% of the trials completed using the 2D Method saw the participant localize a trajectory on the correct side of the phantom, it stands that several of the medical students who were recruited as participants lack this fundamental skill. However, when given access to the imaging alongside 3D or AR visualizations, this increased to 97% and 100% respectively.

The simulated study demonstrated significant improvements in performance when users completed their drill location and drill angle localizations using the 3D and AR methods when compared to traditional 2D methods. This result follows the literature wherein 3D and AR visualizations better facilitate and provide benefits to a user's perception of spatial relations between images or models on a screen and real-world objects, such as patients or phantoms.[18] Additionally, our study demonstrated significant improvements over the 3D method when using the AR method, most saliently with a reduction in the mean completion time of over 40% between the two methods. This illustrates the potential impact of augmented environments and AR technology for increasing the efficiency and effectiveness of surgical training in simulated environments rather than as a benefit to those with more experience, such as attending neurosurgeons.

The holographic instability, defined as the drift or variability in the position of the holographic images and models as the user moves their position in space or changes the direction in which they are observing the scene, produced by the *HoloLens* as the user moves has been estimated to be approximately 5 mm.[41] It is of note that the median drill-tip distance to the preplanned drill location for trainees in a simulated training environment (13 mm) and clinical setting (20 mm) were higher than the estimated holographic instability, even when the variability and registration error are combined. This reveals that participants were not able to mitigate the holographic instability. While it is valuable to see that some trainees were able to localize the access points within the range of the variability and registration error, at present this error is still too high for clinical use. The holographic variability, registration processes, and training that accompany this technology must be improved before use for decision making on patients can begin. However, given the *HoloLens*'s ability to provide direct 2D and 3D visualizations in the operative field, this type of technology may prove beneficial for simulation and training of medical students and surgical residents in simulated and clinical settings. As such, the *HoloLens* and related AR HMD technologies may hold potential for training and simulation-based education or planning of neurosurgical procedures and should continue to be explored.[19,42-43] It is of need to assess whether or not trainees are more capable after being trained with AR technologies than without it – or with other training methods. Furthermore, it stands to be assessed if skills acquired by trainees using AR in simulated environments will extend to clinical settings when the additional pressures that are present in that environment become part of the overall scenario.

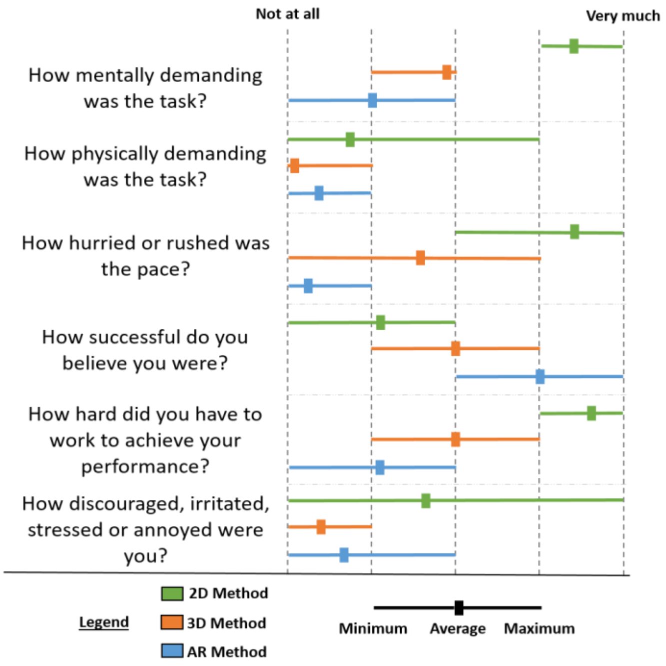

Fig. 7. Min-max-average assessment of post-study questionnaire responses.

A summary of responses to the post-study questionnaire are shown in Fig. 7. Responses showed that, on average, users felt the AR method was less mentally demanding, less hurried or rushed, more successful, and required users to work less hard than the 2D and 3D methods. The AR method was comparable to the 3D method in terms of how discouraged, irritated, stressed or annoyed it made the users. It was comparable to the 2D and 3D methods in terms of how physically demanding the task was. These results illustrate that in addition to quantitatively allowing participants to perform better, they also underwent less mental work and were under less time pressure while performing at overall higher levels of performance. Using the AR method decreases cognitive load for trainees, especially when they are in the early stages of acquiring and refining their technical and surgical skills. The positive results from our questionnaire highlight the potential for this technology to be used in surgical applications or simulation-based approaches for teaching skills and building confidence in trainees. Lastly, a limitation of both the simulated and clinical studies involved recruitment of all participants from the same institution. Trainees were recruited from the Queen's School of Medicine and Queen's University General Surgery residency program and most had little or no prior experience with HMD AR technology or with surgical planning. Furthermore, many trainees in the simulated study had little to no experience in reading and interpreting medical images. To further demonstrate our results, future studies would be required at multiple institutions with participants of all skill levels. In the clinical feasibility study, our sample size is a further limitation.

When compared to similar systems *HoloQuickNav* controller-based registration method proved to provide similar or improved registration accuracy,[23,31,42-43]. This work further provided a detailed, multidimensional and subjective analysis of the user experience and task load that AR HMD can deliver compared to other visualization methods when used for surgical planning or guidance – a detail often excluded in other studies involving the *HoloLens*, or in studies using other AR HMD devices for surgical planning and guidance.[13-14,16-17,21,23,27-31,33] Furthermore, the metrics developed in this work focus on the assessment of planning skills used prior to the procedure – something which several other systems do not quantify.[27-30]

Finally, the proposed system shows promise for use in a neurosurgical training curriculum.

This technology may be used as a training platform which would allow trainees to have their improvements monitored relative to established performance benchmarks, using our established and demonstrated metrics for assessing proficiency in surgical planning, as they progress through a curriculum to aid them in planning and targeting neurosurgical procedures. However, for practical curriculum development, there are two changes that our platform would require. First, metrics must be computed from a gold-standard that is derived from a consensus of expert attending neurosurgeons. This will ensure that instead of needing to match exactly with the optimal trajectory defined by one attending neurosurgeon, at the time of the procedure, trainees will seek to ensure their trajectories fall into a 3D cone of consensus. This cone may be defined to ensure that all expert trajectories are encapsulated, or it may be defined as the mean of all expert trajectories with a tolerance of one standard deviation. In this sense, the essence of the metric is unchanged, but the computation and tolerance for success are modified. Secondly, determining the number and frequency of simulated training sessions required on the path to proficiency will be critical to ensuring that the developed curriculum gives trainees the ability to practice until they become competent.

This technology may allow methods for providing objective and measurable feedback as our metrics are geometric in nature. This ensures that trainees can practice in such a way that it automatically allows trainees to practice their technical skills without expert supervision – an important component of the competency-based medical education paradigm that is rapidly evolving at medical schools around the world. By ensuring that trainees select a trajectory within the cone of consensus, we know that they have selected an appropriate trajectory. Given that the metrics we have established for drill location and drill angle localization are not neurosurgery-specific and the setup is readily replicable, it is foreseeable that this system could be reused for the analysis of planning and targeting effectiveness in other surgical specialties.

As the *HoloQuickNav* platform is now available – and has been validated for simulated training and intra-operative use – we propose that in future work, a curriculum consisting of a series of simulation-based training sessions be developed. The AR platform must be compared against existing learning tools and visualization methods for reaching proficiency in neurosurgical planning. It must be assessed whether this technology allows trainees to reach proficiency sooner or with reduced cognitive effort than with other methods.

4. Conclusion

The results obtained in this work indicate that our AR HMD technology can measurably improve surgical planning and target localization in clinical and simulated training settings for trainees. The feasibility and usefulness of the *HoloLens* were validated for identifying optimal drill location and drill angle in a clinical environment, and this work has led to the development of metrics which allow for significant differentiation between levels of competence in multiple areas. Furthermore, trainees rate this technology equally or more helpful compared to conventional visualization methods.

Acknowledgments

Z. Baum was supported by the Ontario Graduate Scholarship from the Government of Ontario Ministry of Training, Colleges and Universities. G. Fichtinger is supported as a Canada Research Chair in Computer-Integrated Surgery. Financial support was received from the SEAMO Educational Innovation and Research Fund. This work was financially supported as a Collaborative Health Research Project (CHRP #127797), a joint initiative between the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canadian Institutes of Health Research (CIHR).

References

[1] R. Weigel, P. Schmiedek and J. K. Krauss, “Outcome of contemporary surgery for chronic subdural haematoma: evidence based review,” Journal of Neurology, Neurosurgery & Psychiatry, vol. 73, pp. 937-943, 2003.

[2] C. R. Wirtz, F. K. Albert, M. Schwaderer, C. Heuer, A. Staubert, V. M. Tronnier, M. Knauth and S. Kunze, "The benefit of neuronavigation for neurosurgery analyzed by its impact on glioblastoma surgery," Neurological Research, vol. 22, no. 4, pp. 354-360, 2000.

[3] W. J. Elias, J. B. Chadduck, T. D. Alden and E. R. Laws, "Frameless stereotaxy for transsphenoidal surgery," Neurosurgery, vol. 45, no. 2, pp. 271-275, 1999.

[4] D. Orringer, D. Lau, S. Khatri, G. J. Zamora-Berridi, K. Zhang, C. Wu, N. Chaudhary and O. Sagher, "Extent of resection in patients with glioblastoma: limiting factors, perception of resectability, and effect on survival," Journal of Neurosurgery, vol. 117, no. 5, pp. 851-859, 2012.

[5] A. Meola, F. Cutolo, M. Carbone, F. Cagnazzo, M. Ferrari and V. Ferrari, "Augmented reality in neurosurgery: a systematic review," Neurosurgical Review, vol. 40, no. 4, pp. 537-548, 2017.

[6] O. M. Tepper, H. L. Rudy, A. Lefkowitz, K. A. Weimer, S. M. Marks, C. S. Stern and E. S. Garfein, "Mixed reality with HoloLens: where virtual reality meets augmented reality in the operating room," Plastic and Reconstructive Surgery, vol. 140, no. 5, pp. 1066-1070, 2017.

[7] J. R. Frank, L. S. Snell, O. T. Cate, E. S. Holmboe, C. Carraccio, S. R. Swing, P. Harris, N. J. Glasgow, C. Campbell, D. Dath, R. M. Harden, W. lobst, D. M. Long, R. Mungroo, D. L. Richardson, J. Sherbino, I. Silver, S. Taber, M. Talbot and K. A. Harris, "Competency-based medical education: theory to practice," Medical Teacher, vol. 32, no. 8, pp. 638-645, 2010.

[8] E. S. Holmboe, J. Sherbino, D. M. Long, S. R. Swing, J. R. Frank and f. t. I. C. Collaborators, "The role of assessment in competency-based medical education," Medical Teacher, vol. 32, no. 8, pp. 676-682, 2010.

[9] M. Kirkman, M. Ahmed, A. Albert, M. Wilson, D. Nandi and N. Sevdalis, "The use of simulation in neurosurgical education and training: A systematic review," *Journal of Neurosurgery,* vol. 121, no. 2, pp. 228-246, 2014.

[10] S. Issenberg, W. McGaghie, E. Petrusa, D. Lee Gordon and R. Scalese, "Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review," *Medical Teacher,* vol. 27, no. 1, pp. 10-28, 2005.

[11] C. Reiley, H. Lin, D. Yuh and G. Hager, "Review of methods for objective surgical skill evaluation," *Surgical Endoscopy,* vol. 25, no. 2, pp. 356-366, 2011.

[12] I. Cabrilo, P. Bijlenga and K. Schaller, "Augmented reality in the surgery of cerebral aneurysms: a technical report," *Neurosurgery,* vol. 2, pp. 252-260, 2014.

[13] W. Deng, F. Li, M. Wang and Z. Song, "Easy-to-use augmented reality neuronavigation using a wireless tablet PC," *Stereotactic and Functional Neurosurgery,* vol. 92, no. 1, pp. 17-24, 2014.

[14] J.G. Chen, K. W. Han, D. F. Zhang, Z. X. Li, Y. M. Li and L. J. Hou, "Presurgical planning for supratentorial lesions with free Slicer software and Sina app.," *World Neurosugery,* vol. 106, pp. 193-197, 2017.

[15] G. C. Sun, F. Wang, X. L. Chen, X. G. Yu, X. D. Ma, D. B. Zhou, R. Y. Zhu and B. N. Xu, "Impact of virtual and augmented reality based on intraoperative magnetic resonance imaging and functional neuronavigation in glioma surgery involving eloquent areas," *World Neurosurgery,* vol. 96, pp. 375-382, 2016.

[16] J. W. Yoon, R. E. Chen, K. ReFaey, R. J. Diaz, R. Reimer, R. J. Komotar, A. Quinones-Hinojosa, B. L. Brown and R. E. Wharen, "Technical feasibility and safety of image-guided parieto-occipital ventricular catheter placement with the assistance of a wearable head-up display," *International Journal of Medical Robotics and Computer Assisted Surgery,* vol. 13, no. 4, 2017.

[17] C. R. Maurer Jr., F. Sauer, B. Hu, B. Bascle, B. Geiger, F. Wenzel, F. Recchi, T. Rohlfing, C. M. Brown, R. S. Bakos, R. J. Maciunas and A. Bani-Hashemi, "Augmented reality visualization of brain structures with stereo and kinetic depth cues: system description and initial evalution with head phantom," in *SPIE Medical Imaging,* San Diego, 2001.

[18] K. Abhari, J. H. S. Baxter, E. C. S. Chen, A. R. Khan, T. M. Peters and S. de Ribaupierre, "Training for planning tumour resection: augmented reality and human factors," *IEEE Transactions on Biomedical Engineering,* vol. 62, no. 6, pp. 1466-1477, 2015.

[19] M. Kersten-Oertel, I. Gerard, S. Drouin, K. Mok, D. Sirhan, D. S. Sinclair and D. L. Collins, "Augmented reality in neurovascular surgery: feasibility and first uses in the operating room," *International Journal of Computer Assisted Radiology and Surgery,* vol. 10, no. 11, pp. 1823-1836, 2015.

[20] H. Brun, R. A. B. Bugge, L. K. R. Suther, S. Birkeland, R. Kumar, E. Pelanis and E. O. J, "Mixed reality holograms for heart surgery planning: first user experience in congenital heart disease," *European Heart Journal Cardiovascular Imaging,* 2018.

[21] C. Morales, N. V. Navkar, D. Tsagkaris, A. Webb, T. A. Birbilis, I. Seimenis and N. V. Tsekos, "Holographic interface for three-dimensional visualization of MRI on HoloLens: a prototype platform for mri guided neurosurgeries," in *17th International Conference of Bioinformatics and Bioengineering*, Washington, 2017.

[22] W. Si, X. Liao, Q. Wang and P. Heng, "Augmented reality-based personalized virtual perative anatomy for neurosurgical guidance and training," in *IEEE Conference on Virtual Reality and 3D User Interfaces*, Reutlingen, 2018.

[23] T. Frantz, B. Jansen, J. Duernick and J. Vandemeulebroucke, "Augmenting Microsoft's HoloLens with vuforia tracking for neuronavigation," *Healthcare Technology Letters,* vol. 5, no. 5, pp. 221-225, 2018.

[24] F. Insekara, M. Smits, C. Dirven and A. Vincent. “Clinical feasibility of a wearable mixed-reality device in neursosurgery,” World Neurosurgery, vol. 118, pp. e422-e427, 2018.

[25] L. Qian, A. Barthel, A. Johnson, G. Osgood, P. Kazanzides, N. Navab and F. B, "Comparison of optical see-through head-mounted displays for surgical interventions with object-anchored 2D-display," *International Journal of Computer Assisted Radiology and Sugery,* vol. 12, no. 6, pp. 901-910, 2017.

[26] S. Moosburner, C. Remde, P. Tang, M. Queisner, N. Haep, J. Pratschke and I. M. Sauer, "Real world usability analysis of two augmented reality headsets in visceral surgery.," *Artificial Organs,* vol. 00, pp. 1-5, 2018.

[27] G. AlZhrani, F. Alotaibi, H. Azarnoush, A. Winkler-Schwartz, A. Sabbagh, K. Bajunaid, S. P. Lajoie and R. F. Del Maestro, “Proficiency performance benchmarks for removal of simulated brain tumours using a virtual reality simulator NeuroTouch,” *Journal of Surgical Education,* vol. 72, no. 4, pp. 685-696, 2015.

[28] T. Holloway, Z. S. Lorsch, M. A. Chary, S. Sobotka, M. M. Moore, A. B. Costa, R. F. Del Maestro and J. Bederson, “Operator experience determines performance in a simulated computer-based brain tumour resection task,” *International Journal of Computer Assisted Radiology and Surgery,* vol. 10, no. 11, pp. 1853-1862, 2015.

[29] N. Gelinas-Phaneuf, N. Choudhury, A. R. Al-Habib, A. Cabral, E. Nadeau, V. Mora, V. Pazos, P. Debergue, R. DiRaddo and R. F. Del Maestro, “Assessing performance in brain tumour resections using a novel virtual reality simulator,” *International Journal of Computer Assisted Radiology and Surgery,* vol. 9, no. 1, pp. 1-9, 2015.

[30] K. G. Hooten, J. R. Lister, G. Lombard, D. E. Lizdas, S. Lampotang, D. A. Rajon, F. Bova and G. J. A. Murad, “Mixed reality ventriculostomy simulation: experience in neurosurgical residency,” *Operative Neurosurgery,* vol. 10, no. 4, pp. 565-576, 2014.

[31] F. Cutolo, A. Meola, M. Carbone, S. Sinceri, F. Cagnazzo, E. Denaro, N. Esposito, M. Ferrari and V. Ferrari, "A new head-mounted display-based augmented reality system in neurosurgical oncology: a study on phantom," *Computer Assisted Surgery,* vol. 22, no. 1, pp. 39-53, 2017.

[32] P. E. Pelargos, D. T. Nagasawa, C. Lagman, S. Tenn, J. V. Demos, S. J. Lee, T. T. Bui, N. E. Barnette, N. S. Bhatt, N. Ung, A. Bari, N. A. Martin and I. Yang, "Utilizing virtual and augmented reality for educational and clinical enhancements in neurosurgery," *Journal of Clinical Neuroscience,* vol. 35, pp. 1-4, 2017.

[33] E. Rae, A. Lasso, M. Holden, E. Morin, R. Levy and G. Fichtinger, "Neurosurgical burr hole placement using the Microsoft HoloLens," in *SPIE Medical Imaging,* Houston, 2018.

[34] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J. C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J.

Buatti, S. Aylward, J. V. Miller, S. Pieper and R. Kikinis, "3D Slicer as an image computing plaform for the quantitative imaging network," *Magnetic Resonance Imaging,* vol. 30, pp. 1323-1341, 2012.

[35] D. Paraskevopoulous, A. Unterberg, R. Metzner, J. Dreyhaupt, G. Eggers and C. R. Wirtz, "Comparative study of application accuracy of two frameless neuronavigation systems: experimental error assessment quantifying registration methods and clinically influencing factors," *Neurosurgical Review*, vol. 34, no. 2, pp. 217-228, 2010.

[36] A. Lasso, T. Heffter, A. Rankin, C. Pinter, T. Ungi and G. Fichtinger, "PLUS: open-source toolkit for ultrasound-guided intervention systems," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 10, pp. 2527-2537, 2014.

[37] J. Tokuda, G. S. Fisher, X. Papademetris, Z. Yaniv, L. Ibanez, P. Cheng, H. Liu, J. Blevins, J. Arata, A. J. Golby, T. Kapur, S. Pieper, E. C. Burdette, G. Fichtinger, C. M. Tempany and N. Hata, "OpenIGTLink: an open network protocol for image-guided therapy environment," *International Journal of Merical Robotics and Computer Assisted Surgery,* vol. 5, no. 4, pp. 423-434, 2009.

[38] T. Kapur, S. Pieper, A. Fedorov, J. C. Fillion-Robin, M. Halle, L. O'Donnell, A. Lasso, T. Ungi, C. Pinter, J. Finet, S. Pujol, J. Jagadeesan, J. Tokuda, I. Norton, R. S. J. Estepar, D. Gering, H. J. W. L. Aerts, M. Jakab, N. Hata, L. Ibanez, D. Blezek, J. Miller, S. Aylward, W. E. L. Grimson, G. Fichtinger, W. M. Wells, W. E. Lorensen, W. Shroeder and R. Kikinis, "Increasing the impact of medical image computing using community-based open-access hackathons: the NA-MIC and 3D Slicer experience," Medical Image Analysis, vol. 33, pp. 176-180, 2016.

[39] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," *Advances in Psychology,* vol. 52, pp. 139-183, 1988.

[40] N. Choudhury, N. Gelinas-Phaneuf, S. Delorme and R. Del Maestro, "Fundamentals of neurosurgery: virtual reality tasks for training and evaluation of technical skills," *World Neurosurgery,* vol. 80, no. 5, pp. e9-e19, 2013.

[41] R. Vassallo, A. Rankin, E. Chen and T. Peters, "Hologram stability evaluation for Microsoft Hololens," in *SPIE Medical Imaging*, Orlando, 2017.

[42] P. Pratt, M. Ives, G. Lawton, J. Simmons, N. Radev, L. Spyropolou and D. Amiras, "Through the HoloLens looking glass: augmented reality for extremity reconstruction surgery using 3D vascular models with perforating vessels," *European Radiology Experimental,* vol. 2, no. 1, p. 2, 2018.

[43] M. Pfandler, M. Lazarovici, P. Stefan, P. Wucherer and M. Weigi, "Virtual reality-based simulators for spine surgery: a systematic review," *The Spine Journal,* vol. 17, no. 9, pp. 1352-1363, 2017.