**University of
Zurich^{UZH}**

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2021

Current Accuracy of Augmented Reality Neuronavigation Systems: Systematic Review and Meta-Analysis

Fick, T ; van Doormaal, J A M ; Hoving, E W ; Willems, P W A ; van Doormaal, T P C

Abstract: BACKGROUND Augmented reality neuronavigation (ARN) systems can overlay three-dimensional anatomy and pathology without the need for a two-dimensional external monitor. Accuracy is crucial for their clinical applicability. We performed a systematic review regarding the reported accuracy of ARN systems and compared them with the accuracy of conventional infrared neuronavigation (CIN). **OBJECTIVE** Explore the current navigation accuracy of ARN systems and compare them with CIN. **METHODS** Pubmed and Embase were searched for ARN and CIN systems. For ARN: type of system, method of patient-to-image registration, accuracy method and accuracy of the system was noted. For CIN: navigation accuracy, expressed as target registration error (TRE), was noted. A meta-analysis was performed comparing the TRE of ARN and CIN systems. **RESULTS** 35 studies were included, 12 for ARN and 23 for CIN. ARN systems were divided into head-mounted display and heads-up display. In ARN, four methods were encountered for patient-to-image registration, of which point-pair matching was the one most frequently used. Five methods for assessing accuracy were described. 94 TRE measurements of ARN systems were compared with 9058 TRE measurements of CIN systems. Mean TRE was 2.5 mm (CI 95% 0.7 - 4.4) for ARN systems and 2.6 mm (CI 95% 2.1 - 3.1) for CIN systems. **CONCLUSIONS** In ARN, there seems to be lack of agreement regarding the best method to assess accuracy. Nevertheless, ARN systems seem able to achieve an accuracy comparable with CIN systems. Future studies should be prospective and compare TREs which should be measured in a standardized fashion.

DOI: <https://doi.org/10.1016/j.wneu.2020.11.029>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-192552>

Journal Article

Accepted Version

Originally published at:

Fick, T; van Doormaal, J A M; Hoving, E W; Willems, P W A; van Doormaal, T P C (2021). Current Accuracy of Augmented Reality Neuronavigation Systems: Systematic Review and Meta-Analysis. *World Neurosurgery*, 146:179-188.

DOI: <https://doi.org/10.1016/j.wneu.2020.11.029>

Journal Pre-proof

Current Accuracy of Augmented Reality Neuronavigation Systems: Systematic Review and Meta-Analysis

T. Fick, MD, J.A.M. van Doormaal, BSc, E.W. Hoving, MD PhD, P.W.A. Willems, MD PhD, T.P.C. van Doormaal, MD PhD

PII: S1878-8750(20)32404-9

DOI: <https://doi.org/10.1016/j.wneu.2020.11.029>

Reference: WNEU 16300

To appear in: *World Neurosurgery*

Received Date: 19 August 2020

Revised Date: 4 November 2020

Accepted Date: 5 November 2020

Please cite this article as: Fick T, van Doormaal JAM, Hoving EW, Willems PWA, van Doormaal TPC, Current Accuracy of Augmented Reality Neuronavigation Systems: Systematic Review and Meta-Analysis, *World Neurosurgery* (2020), doi: <https://doi.org/10.1016/j.wneu.2020.11.029>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Elsevier Inc. All rights reserved.

Current Accuracy of Augmented Reality Neuronavigation Systems: Systematic Review and Meta-Analysis

T. Fick MD¹, J.A.M. van Doormaal BSc², E.W. Hoving MD PhD¹, P.W.A. Willems MD PhD³, T.P.C. van Doormaal MD PhD^{3, 4}

¹ Department of Neuro-oncology, Princess Máxima Center for Pediatric Oncology, Heidelberglaan 25, 3584 CS Utrecht, The Netherlands

²Department of Oral and Maxillo-facial surgery, University Medical Centre Utrecht, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands

³Department of Neurosurgery, University Medical Centre Utrecht, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands

⁴ Department of Neurosurgery, University Hospital of Zürich, Rämistrasse 100, 8091 Zürich, Switzerland

Corresponding author: T. Fick. E-mail address: t.fick-4@prinsesmaximacentrum.nl

Short title: Accuracy of Augmented Reality in Neurosurgery

Key words: Accuracy, Augmented Reality, Neuronavigation

Declaration of interest: T.P.C. van Doormaal is founder and CMO of AugmedIT, a startup company that develops Augmented Reality tools for surgeons.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Abstract

Background: Augmented reality neuronavigation (ARN) systems can overlay three-dimensional anatomy and pathology without the need for a two-dimensional external monitor. Accuracy is crucial for their clinical applicability. We performed a systematic review regarding the reported accuracy of ARN systems and compared them with the accuracy of conventional infrared neuronavigation (CIN).

Objective: Explore the current navigation accuracy of ARN systems and compare them with CIN.

Methods: Pubmed and Embase were searched for ARN and CIN systems. For ARN: type of system, method of patient-to-image registration, accuracy method and accuracy of the system was noted. For CIN: navigation accuracy, expressed as target registration error (TRE), was noted. A meta-analysis was performed comparing the TRE of ARN and CIN systems.

Results: 35 studies were included, 12 for ARN and 23 for CIN. ARN systems were divided into head-mounted display and heads-up display. In ARN, four methods were encountered for patient-to-image registration, of which point-pair matching was the one most frequently used. Five methods for assessing accuracy were described. 94 TRE measurements of ARN systems were compared with 9058 TRE measurements of CIN systems. Mean TRE was 2.5 mm (CI 95% 0.7 – 4.4) for ARN systems and 2.6 mm (CI 95% 2.1 – 3.1) for CIN systems.

Conclusions: In ARN, there seems to be lack of agreement regarding the best method to assess accuracy. Nevertheless, ARN systems seem able to achieve an accuracy comparable with CIN systems. Future studies should be prospective and compare TREs which should be measured in a standardized fashion.

Introduction

Augmented reality is a technology that superimposes a virtual image into the user's view of the real world. It is a technique that has been used in neuronavigation systems since 1986,¹ but still never found its way into routine clinical practice, despite its possible advantages. In CIN, the surgeon needs to mentally transform two-dimensional (2D) images from an external screen to the three-dimensional (3D) anatomy of the surgical field. Furthermore, a separate probe needs to be manipulated to make the neuronavigation information visible. In general, distractions and attention shifts can have a significant negative effect on task performance, especially when the task becomes more complex.^{2,3} This has also been shown in performing specific surgical tasks.⁴ Some ARN systems can superimpose the images directly onto the surgical field, without interference of an external screen. This diminishes attention shifts, which makes surgical tasks more efficient.⁵ Moreover, ARN obviates the need for the surgeon to translate 2D information into 3D anatomy, potentially reducing the chance of interpretation error and leaving more room mentally to concentrate on other surgical tasks.

In recent years, there have been major developments in this field and interest has grown, especially in neurosurgery.^{6,7} Neuronavigation systems are widely used in neurosurgery, taking advantage from the fact that the patient is often attached to the operating table with a head clamp. This turns the head of the patient and operating table into one single rigid body. To date, different ARN systems have been developed for spine surgery^{8,9} neuro-oncology^{10,11} and complex cerebral vascular surgeries.¹²⁻¹⁴ Since the beginning of neuronavigation systems, high accuracy after patient-to-image registration has remained one of the biggest challenges. With ARN, new and challenging contributors to inaccuracy arise.¹⁵ Novel ways of measuring the accuracy are therefore developed. ARN overlay accuracy has been described in various ways such as pixel reprojection error¹² or pixel misalignment error.¹⁶ In this review, we give an overview of the various ARN systems that are applicable to intracranial procedures and their accuracy after patient-to-image registration. Furthermore, we performed a meta-analysis to compare the current ARN accuracy to the accuracy of CIN systems.

Methods

A systematic review was performed using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. A first search was performed for ARN systems. In the meta-analysis we aimed to compare the navigation accuracy, expressed as

TRE, of ARN systems to CIN systems. Since there were no studies, with TRE as outcome measure, which compared both systems, a second search was performed for CIN systems.

Search strategy

We performed a literature search on Pubmed and Embase and used snowballing. The Pubmed search for ARN comprised the following terms: (“Neurosurgery”[Mesh] OR “Neurosurgical Procedures”[Mesh] OR Neurologic Surgical Procedure*[Title/Abstract] OR neurosurg*[Title/Abstract] OR Anterior Temporal Lobectomy[Title/Abstract] OR Brain Tissue Transplantation*[Title/Abstract] OR Cerebral Decortication[Title/Abstract] OR Hemispherectomy[Title/Abstract] OR Cerebrospinal Fluid Shunt*[Title/Abstract] OR Ventriculoperitoneal Shunt*[Title/Abstract] OR Ventriculostomy[Title/Abstract] OR Craniotomy[Title/Abstract] OR Denervation[Title/Abstract] OR Hypophysectomy[Title/Abstract] OR Microvascular Decompression Surgery[Title/Abstract] OR Neuroendoscopy[Title/Abstract] OR Split-Brain Procedure*[Title/Abstract] OR Stereotaxic Technique*[Title/Abstract] OR Neuronavigation[Title/Abstract]) AND (augmented reality[Title/Abstract] OR hybrid reality[Title/Abstract] OR mixed reality[Title/Abstract] OR AR[Title/Abstract] OR head mounted display[Title/Abstract] OR heads up display*[Title/Abstract]). This search was performed on October 2, 2019. We performed a similar search in Pubmed and Embase for CIN, the term ‘accuracy’ was added to optimize the search range. This search was performed on 28-01-2020.

Selection criteria

For the two separate searches, articles were included if the following selection criteria were met:

For ARN: (1) ARN accuracy after patient/phantom-to-image registration was measured and the method of measurement was described (2) The ARN system is used or applicable to open, cerebral procedures (3) 3D data is used as source for the virtual scene (4) virtual scene can be visualized in a 3D manner (5) the AR system gives a stereoscopic view and (6) ARN view is superimposed onto the surgical field without use of an external monitor. These inclusion criteria were chosen to only include ARN systems that superimpose the virtual images directly on the surgical field in a 3D manner. Furthermore, spinal surgery was not included in this study since the technical challenges in navigation, accuracy measurements, anatomy and surgical targets and goals are in big contrast compared to intracranial surgery.

We excluded studies where ARN accuracy measurement methods were not clearly defined or when there was no full-text available. Congress abstracts were included only when all the information could be extracted.

For CIN: (1) CIN accuracy after patient/phantom-to-image registration was used as primary outcome (2) navigation accuracy was measured as the TRE, defined as the Euclidean distance between a target point in physical space and the corresponding point in image space (3) systems were designed for open, cerebral procedures (4) commercially available, infrared, neuronavigation systems were used and (5) only standard clinical practice systems for imaging and patient-to-image registration were used. Intraoperative CT and/or MRI were not considered as standard clinical practice. We excluded studies that used self-designed or novel techniques, when the measurement method was not clearly defined, when there was no full-text available or when there was only a congress abstract reported. Self-designed or novel techniques were excluded to make the best possible comparison for accuracy of current standard clinical practice infrared neuronavigation systems.

Records were screened and included separately by 2 reviewers (TF and JvD). Disagreements between reviewers were assessed and resolved, if necessary through consultation of a third reviewer (TvD).

Data extraction

The following data was extracted from the ARN studies: (1) date of publication (2) display type (3) patients and/or phantom measurements (4) patient-to-image registration method (5) accuracy unit (6) accuracy, including number of measurements and standard deviation and (7) measurement method.

For CIN the mean navigation accuracy, including standard deviation and number of measurements was extracted for meta-analysis. This was also independently performed by the two reviewers. We collected all the data into an excel-file.

When there was missing data or the full-text was not available, we contacted the first authors by email.

Statistical analysis

For the meta-analysis, the TRE of ARN systems and CIN systems were compared. The mean TRE and their corresponding variances (i.e. squared standard errors, calculated from results as reported by individual studies) were analyzed with a linear mixed model. As we observed

large differences between variances of individual studies for ARN and CIN, we specified separate heterogeneity measures (i.e. separate between study variances) for ARN and CIN in the mixed models. We report estimated pooled means for both ARN and CIN methods with 95% CIs and the difference between pooled means with 95% and a p-value. Additionally, we estimated I^2 values from the results of the mixed model (ref. Higgins 2002) to quantify heterogeneity for ARN and CIN. A p-value of <0.05 was deemed statistically significant. Statistics were performed in Statistical Analysis Software (SAS v9.4).

Results

The results of our searches are shown in figure 1. A total of 3586 records were screened, 803 for ARN and 2775 for CIN. For ARN, 43 full-text articles were assessed for eligibility and 12 studies were included in this review.^{17–28} Of those 12 studies, 5 studies, which had the TRE as accuracy measurement, were included for meta-analysis.^{17,21–24} For CIN, 54 full-text articles were assessed for eligibility and 23 studies were included, all for meta-analysis.^{29–51}

Augmented reality systems

A summary of the extracted data is shown in table 1. The systems can be categorised into two main groups: the head-mounted displays (HMD) and heads-up displays (HUD). There is one self-designed, experimental system included that uses integral videography. Five different terms were used to define navigation accuracy. A description of the different terms is shown in table 2 and a graphic illustration of each term is shown in figure 2. In five studies, a TRE was calculated.^{17,21–24} This is calculated as the Euclidean distance between the real target point and virtual target point. The fiducial registration error (FRE) was used in two studies.^{17,18} The FRE is a measure for registration accuracy and is defined as the root-mean-square distance between recognized fiducial positions and their homologous virtual fiducial positions after registration(*figure 2A*). The targeting error was used in four studies^{19,20,27,28} and is similar to the TRE. However, contrary to TRE, targeting error is measured in 2D(*figure 2B*). Two studies used a more indirect measurement to calculate the accuracy. In *Li et al. 2018* the target point was set as where the tip of the external ventricle drain after insertion needed to be. The distance between the target point and the actual drain tip on postoperative CT-imaging was used as accuracy measurement(*figure 2C*) and was called the mean target deviation. *Incekara et al 2018* drew the circumference of the tumour on the patients scalp based on the ARN system and of the CIN system. The centre of the tumour was determined

for each system and the difference between the two was used as accuracy measurement and was called the deviation compared to CIN (*figure 2D*).

We calculated the means and confidence interval (CI) of 95% per accuracy measurement method for ARN systems. The mean TRE was 2.5mm (CI 95% 0.7 – 4.4), the mean target deviation 4.3mm (CI 95% 3.4 – 5.2) and the deviation compared to CIN 4.4mm (CI 95% 2.4 – 6.4). Because of the small number of studies for FRE and targeting error, their mean and CI 95% are mentioned per individual study. For FRE 1.2mm (CI 95% 1.0 – 1.4) and 6.6mm (CI 95% 5.2 – 7.9) and for targeting error 1.9mm (CI 95% 1.7 – 2.1) and 2.8mm (CI 95% 2.4 – 3.1). For two studies, which used the targeting error as accuracy measurement, the CI 95% could not be calculated since there was not enough data available in the studies. The reported means are 0.9mm¹⁹ and 1.1mm.²⁰

As registration method, point-pair matching is most frequently used.^{17–20,24,27,28} In mixed reality HMDs it is possible to manoeuvre the hologram with your hands. This way it is also possible to perform the registration by placing the hologram over the patient manually.^{21,25,26} Other techniques used were a self-designed automatic registration using intraoperative CT-imaging²⁴ and a reference frame with a fixed ratio to the phantom with live tracking.²²

Augmented reality neuronavigation accuracy compared to conventional infrared neuronavigation

The different pathologies in all the ARN studies and CIN studies are shown in table 3. Oncology is the largest group comprising 76.5% in ARN studies and 67.8% in CIN studies. Other groups are vascular (8.8% in ARN and 5.3% in CIN) and other (14.7% in ARN and 2.0% in CIN).

We performed a meta-analysis and compared the target registration error of ARN and CIN systems. Five articles were included for ARN^{17,21–24} with a total of 94 measurements and 23 studies were included for CIN^{29–51} with a total of 9058 measurements. Both measurements on patients and phantoms were included in the data. In ARN, 49 (52.1%) measurements were performed on patients with a mean of 2.22mm (SD 1.34) and 45 (47.9%) measurements were performed on phantoms with a mean of 3.76mm (SD 2.07). In CIN, 1431 (15.8%) measurements were performed on patients with a mean of 3.96mm (SD 1.98) and 7627 (84.2%) measurements were performed on phantoms with a mean of 2.49mm (SD 1.62). For both ARN and CIN we observed a high degree of heterogeneity with I² values of 0.957 for CIN and 0.831 for ARN. The overall mean TRE of ARN was 2.5mm (CI 95% 0.7 – 4.4) and

the overall mean of CIN was 2.6 (CI 95% 2.1 – 3.1) (*figure 3*). Although this difference is not statistically significant ($p=0.92$), it is noteworthy that the ARN measurements demonstrated a much larger confidence interval.

Discussion

In this paper, we offered an overview of current ARN systems and their accuracy in comparison to CIN systems.

Several pathologies are assessed where in both groups the oncology group is the largest. A noticeable difference between these groups is in the ‘other’ category with 14.7% in the ARN group and 2.0% in the CIN group. This is due to the larger amount of intracranial hypertension cases in the ARN group in which HMDs can have a big impact due to the macroscopic nature of the procedure as supposed to CIN which is more often used in microscopic procedures.

In ARN, multiple terms have been used for expressing navigation accuracy, each with its own limitations. TRE is the measurement to express the navigation accuracy in the region of interest to the neurosurgeon in 3D space. However, TRE measurements in ARN remains a challenge, partly due to the poor depth perception in holograms. The targeting error can therefore be used to overcome this issue, though this is a 2D measurement which as a consequence will always be an underestimation of the true navigation accuracy. FRE is a frequent used term to express the navigation accuracy. With point-based matching the FRE can automatically be calculated when using an ICP-algorithm. Most CIN systems show the FRE after patient-to-image registration automatically and therefore the neurosurgeon is used to this term. For each individual system FRE can only be used as check to inform the neurosurgeon whether the registration process has been performed successfully since a specific FRE does not correlate with a specific navigation accuracy in the region of interest and is thus not ideal to use as expression of navigation accuracy.⁵² However, for evaluation of a new designed system it can be used to assess the accuracy of the overall registration process. The preferred term to assess accuracy can also depend on the surgical goal. When inserting depth electrodes the goal is to accurately place the electrodes at a predefined target.⁵³ In comparison, when inserting an external ventricle drain, the accuracy outcome can also be the deviation to a predefined target. The initial registration accuracy has influence on the outcome, but is not the only factor contributing to the outcome and is therefore not a good

measure to express navigation accuracy. Thus TRE remains the term of most interest to the neurosurgeon and we believe this should be set as golden standard to express the application accuracy in ARN systems.

We found an overall mean TRE for ARN of 2.5mm (CI 95% 0.7 – 4.4) and for CIN of 2.6 (CI 95% 2.1 – 3.1). These values are not significantly different ($p = 0.92$). However, as seen in figure 3, the variation in data in the ARN group is quite large, including the overall lowest and highest mean accuracy while only comprising 5 studies. Several factors may contribute to this diversity. In ARN, TRE measurements are performed in various, self-designed ways. This in contrast to CIN where these measurements are much more standardized. Calculating the TRE in ARN systems remains a challenge and especially in the HMD group new challenges arise. Simply measuring the distance between the actual and virtual target is unreliable; the virtual target can be underneath the surface of the patient and therefore be unreachable. Furthermore, depth perception can be difficult which makes the placement of the measuring device on the virtual target imprecise.⁵⁴ Measuring the X-, Y- and Z-axis in 2D from different views can overcome the difficulties in depth perception. However, this can cause random holographic drift, where the hologram shifts from its original position when moving around, which makes the results difficult to interpret.⁵⁵ Another factor which may influence the accuracy is how the device is worn. To the best of our knowledge there are now studies which have evaluated this variable. In CIN, a probe is tracked at a submillimeter level and is used to acquire 3D coordinates of the targets. HUDs in microscopes that are tracked by infrared cameras use the same technique and add the microscope calibration error to calculate the total augmented reality overlay accuracy. HMDs can be programmed to track probes as well, however with a relatively much larger inaccuracy making it unusable for accuracy measurements. HMDs are therefore sometimes incorporated in an infrared workflow to allow coordinates to be obtained with an infrared tracked probe. For this the HMD needs to be adapted and additional equipment is necessary in the operating room making the HMD no longer a stand-alone, low-cost, neuronavigation system. Effort needs to be put into developing an easy-to-use accuracy measurement method for TRE in order for the ARN systems to be evaluated. Moreover, in the ARN group, different systems were used to overlay the virtual images. HUDs, HMDs and a self-designed, experimental system that uses integral videography. These systems have their own properties in how they perform and in which way the TRE measurements can be performed. All these different features may contribute in the heterogeneity of the data.

Limitations

This review has several limitations. In the meta-analysis different factors influencing accuracy are not accounted for separately. In CIN it is shown that the type of patient-to-image registration method can influence the final navigation accuracy.^{32,46,51} In point-pair matching alone, the type of fiducials (anatomical landmarks, stickers or bone implants), the number of fiducials and the configuration all have effect on the overall navigation accuracy.⁵⁶⁻⁵⁸ The data does not allow to account for these different factors separately since the groups would be too small. Moreover, not all registration methods are applicable to both ARN and CIN and therefore a direct comparison is not possible. The accepted accuracy depends mainly on the goal and type of neurosurgery. Deep brain stimulation, neuronavigated biopsies or open neuro-oncological tumour resection all require different levels of accuracy and the type of patient-to-image registration method is chosen accordingly. For future research it would be of benefit to compare the impact of different registration strategies in ARN systems.

Furthermore, phantoms as well as patients are used. Fiducials on phantoms are rigid, bypassing the error of skin movement and phantom measurements are done in a controlled, laboratory setting avoiding any other variables that may cause inaccuracy. In our meta-analysis 47.9% of measurements in the ARN group were performed on phantoms in comparison to 84.2% in the CIN group which can be an advantage for a better outcome in the CIN group. In contrast, in the ARN group the TRE in phantom measurements was higher compared to the TRE in patient measurements. This means other factors (method of registration, type of system) besides patient or phantom measurements are of bigger influence on the overall navigation accuracy.

The meta-analysis was performed with the assumption that the data follows a normal distribution. This was not applicable to every study, where it varied in normal distributed data and a slight positive skew. However, since the means were not statistically significant by a large margin, we do not expect for our conclusion to be different with other analyses.

ARN has seen a rapid development with various novel systems and self-designed techniques which may lead to publication bias. Therefore, the reported 2.5mm accuracy might be optimistic. Nonetheless, it indicates that this accuracy is technically possible and can approach the CIN accuracy. This number is clinically acceptable since systems with this accuracy are currently used in clinical practice. Therefore, this number could be used as benchmark for future neuronavigation systems.

Conclusions

In ARN, a universal method to report the navigation accuracy is lacking. A TRE of 2.5mm was demonstrated to be currently achievable with ARN and did not show a significant difference compared to CIN systems. Future studies should be prospective and compare accuracy between ARN and CIN using TRE as golden standard. Furthermore, effort needs to be put into developing a standardized method for measuring the TRE.

Acknowledgements: We thank N.P.A. Zuithoff of Julius Support at the University of Utrecht for performing the meta-analysis.

References:

1. Roberts DW, Strohbehn JW, Hatch JF, Murray W, Kettenberger H. A frameless stereotaxic integration of computerized tomographic imaging and the operating microscope. *J Neurosurg.* 1986;65(4):545-549. doi:10.3171/jns.1986.65.4.0545
2. Weerdesteyn V, Schillings AM, Duysens J, Van Galen GP. Distraction Affects the Performance of Obstacle Avoidance During Walking. *J Mot Behav.* 2003;35(1):53-63. doi:10.1080/00222890309602121
3. Graydon J. Distraction and Cognitive Performance. *Eur J Cogn Psychol.* 1989;1(2):161-179. doi:10.1080/09541448908403078
4. Goodell KH, Cao CGL, Schwaitzberg SD. Effects of cognitive distraction on performance of laparoscopic surgical tasks. *J Laparoendosc Adv Surg Tech.* 2006;16(2):94-98. doi:10.1089/lap.2006.16.94
5. Léger É, Drouin S, Collins DL, Popa T, Kersten-Oertel M. Quantifying attention shifts in augmented reality image-guided neurosurgery. *Healthc Technol Lett.* 2017;4(5):188-192. doi:10.1049/htl.2017.0062
6. Meola A, Cutolo F, Carbone M, et al. Augmented reality in neurosurgery: a systematic review. *Neurosurg Rev.* 2018;40(4):537-548. doi:10.1007/s10143-016-0732-9. Augmented
7. Contreras López WO, Navarro PA, Crispin S. Intraoperative clinical application of

augmented reality in neurosurgery: A systematic review. *Clin Neurol Neurosurg.* 2019;177:6-11. doi:10.1016/j.clineuro.2018.11.018

8. Molina CA, Theodore N, Ahmed AK, et al. Augmented reality–assisted pedicle screw insertion: a cadaveric proof-of-concept study. *J Neurosurg Spine.* 2019;31(July):1-8. doi:10.3171/2018.12.spine181142

9. Elmi-Terander A, Burström G, Nachabe R, et al. Pedicle Screw Placement Using Augmented Reality Surgical Navigation With Intraoperative 3D Imaging: A First In-Human Prospective Cohort Study. *Spine (Phila Pa 1976).* 2019;44(7):517-525. doi:10.1097/BRS.0000000000002876

10. Kawamata T, Iseki H, Shibasaki T, Hori T, Heilman CB, Barnett GH. Endoscopic augmented reality navigation system for endonasal transsphenoidal surgery to treat pituitary tumors: Technical note. *Neurosurgery.* 2002;50(6):1393-1397. doi:http://dx.doi.org/10.1097/00006123-200206000-00038

11. Mascitelli JR, Schlachter L, Chartrain AG, et al. Navigation-linked heads-up display in intracranial surgery: Early experience. *Oper Neurosurg.* 2018;15(2):184-193. doi:10.1093/ons/oxp205

12. Cabrilo I, Bijlenga P, Schaller K. Augmented reality in the surgery of cerebral arteriovenous malformations: Technique assessment and considerations. *Acta Neurochir (Wien).* 2014;156(9):1769-1774. doi:10.1007/s00701-014-2183-9

13. Rychen J, Goldberg J, Raabe A, Bervini D. Augmented Reality in Superficial Temporal Artery to Middle Cerebral Artery Bypass Surgery: Technical Note. *Oper Neurosurg.* 2019;0(0):1-7. doi:10.1093/ons/opz176

14. Kersten-Oertel M, Gerard I, Drouin S, et al. Augmented reality in neurovascular surgery: feasibility and first uses in the operating room. *Int J Comput Assist Radiol Surg.* 2015;10(11):1823-1836. doi:10.1007/s11548-015-1163-8

15. Holloway R. Registration Error Analysis for augmented reality. *PRESENCE Virtual Augment Real.* 1997;6(4):413-432.
https://books.google.co.uk/books/about/Augmented_Reality_for_Developers.html?id=8xhKDwAAQBAJ&printsec=frontcover&source=kp_read_button&redir_esc=y#v=onepage&q&f=false%0Ahttps://mediatum.ub.tum.de/1286695

16. Gerard IJ, Kersten-Oertel M, Drouin S, et al. Combining intraoperative ultrasound brain shift correction and augmented reality visualizations: a pilot study of eight cases. *J Med Imaging*. 2018;5(02):1. doi:10.1117/1.JMI.5.2.021210
17. Zhang X, Chen G, Liao H. High-quality see-through surgical guidance system using enhanced 3-D autostereoscopic augmented reality. *IEEE Trans Biomed Eng*. 2017;64(8):1815-1825. doi:10.1109/TBME.2016.2624632
18. van Doormaal T, van Doormaal J, Mensink T. Clinical Accuracy of Holographic Navigation Using Point-Based Registration on Augmented-Reality Glasses. *Oper Neurosurg*. 2019;0(0):1-6. doi:10.1093/ons/opz266
19. King AP, Edwards PJ, Maurer CR, et al. A system for microscope-assisted guided interventions. *Stereotact Funct Neurosurg*. 1999;72(2-4):107-111. doi:10.1159/000029708
20. Edwards PJ, Hawkes DJ, Hill DLG, et al. Augmentation of reality using an operating microscope for otolaryngology and neurosurgical guidance. *Comput Aided Surg*. 1995;1(3):172-178. doi:10.3109/10929089509105692
21. McJunkin JL, Jiramongkolchai P, Chung W, et al. Development of a mixed reality platform for lateral skull base anatomy. *Otol Neurotol*. 2018;39(10):e1137-e1142. doi:10.1097/MAO.0000000000001995
22. Meulstee JW, Nijsink J, Schreurs R, et al. Toward Holographic-Guided Surgery. *Surg Innov*. 2019;26(1):86-94. doi:10.1177/1553350618799552
23. Carl B, Bopp M, Voellger B, Saß B, Nimsky C. Augmented Reality in Transsphenoidal Surgery. *World Neurosurg*. 2019;125:e873-e883. doi:10.1016/j.wneu.2019.01.202
24. Carl B, Bopp M, Chehab S, Bien S, Nimsky C. Preoperative 3-Dimensional Angiography Data and Intraoperative Real-Time Vascular Data Integrated in Microscope-Based Navigation by Automatic Patient Registration Applying Intraoperative Computed Tomography. *World Neurosurg*. 2018;113:e414-e425. doi:10.1016/j.wneu.2018.02.045
25. Incekara F, Smits M, Dirven C, Vincent A. Clinical Feasibility of a Wearable Mixed-Reality Device in Neurosurgery. *World Neurosurg*. 2018;118:e422-e427. doi:10.1016/j.wneu.2018.06.208

26. Li Y, Chen X, Wang N, et al. A wearable mixed-reality holographic computer for guiding external ventricular drain insertion at the bedside. *J Neurosurg*. Published online 2018;1-8. doi:10.3171/2018.4.jns18124
27. Maruyama K, Watanabe E, Saito K, et al. Novel augmented reality-based neuronavigation using smart glasses and clinical application to brain tumor surgery. *Neuro Oncol*. 2017;19(Suppl 3):43-44.
28. Maruyama K, Watanabe E, Kin T, et al. Smart Glasses for Neurosurgical Navigation by Augmented Reality. *Oper Neurosurg*. 2018;15(5):551-556. doi:10.1093/ons/oxx279
29. Ji D, Dong Y, Wang M, Song Z. Accuracy analysis of line-based registration for image guided neurosurgery at different operating areas—a phantom study. *Comput Assist Surg*. 2017;22(0):148-156. doi:10.1080/24699322.2017.1389392
30. Schwam ZG, Kaul VZ, Cosetti MK, Wana GB. Accuracy of a Modern Intraoperative Navigation System for Temporal Bone Surgery in a Cadaveric Model. *Otolaryngol - Head Neck Surg (United States)*. 2019;161(5):842-845. doi:10.1177/0194599819864292
31. Li Q, Zamorano L, Jiang Z, et al. Effect of optical digitizer selection on the application accuracy of a surgical localization system - A quantitative comparison between the OPTOTRAK and FlashPoint tracking systems. *Comput Aided Surg*. 1999;4(6):314-321. doi:10.1002/(SICI)1097-0150(1999)4:6<314::AID-IGS3>3.0.CO;2-G
32. Pfisterer WK, Papadopoulos S, Smith K. Fiducial versus nonfiducial neuronavigation registration assessment and considerations of accuracy. *Neurosurgery*. 2008;62(Mar):201-207. doi:10.1227/01.NEU.0000297007.04975.C6
33. Raabe A, Krishnan R, Wolff R, et al. Laser surface scanning for patient registration in intracranial image-guided surgery. *Neurosurgery*. 2002;50(4):797-803. doi:10.1097/00006123-200204000-00021
34. Shamir RR, Joskowicz L, Spektor S, Shoshan Y. Localization and registration accuracy in image guided neurosurgery: A clinical study. *Int J Comput Assist Radiol Surg*. 2009;4(1):45-52. doi:10.1007/s11548-008-0268-8
35. Kral F, Riechelmann H, Freysinger W. Navigated surgery at the lateral skull base and registration and preoperative imagery: Experimental results. *Arch Otolaryngol - Head*

Neck Surg. 2011;137(2):144-150. doi:10.1001/archoto.2010.249

36. Gerard IJ, Hall JA, Mok K, Collins DL. New protocol for skin landmark registration in image-guided neurosurgery: Technical note. *Clin Neurosurg.* 2015;11(3):376-381. doi:10.1227/NEU.0000000000000868
37. Mascott CR, Sol JC, Bousquet P, Lagarrigue J, Lazorthes Y, Lauwers-Cances V. Quantification of true in vivo (application) accuracy in cranial image-guided surgery: Influence of mode of patient registration. *Neurosurgery.* 2006;59(1 SUPPL. 1):146-156. doi:10.1227/01.NEU.0000220089.39533.4E
38. Stieglitz LH, Fichtner J, Andres R, et al. The silent loss of neuronavigation accuracy: A systematic retrospective analysis of factors influencing the mismatch of frameless stereotactic systems in cranial neurosurgery. *Neurosurgery.* 2013;72(5):796-807. doi:10.1227/NEU.0b013e318287072d
39. Helm PA, Eckel TS. Accuracy of Registration Methods in Frameless Stereotaxis. *Comput Aided Surg.* 1998;3(2):51-56. doi:10.3109/10929089809148129
40. Makiese O, Pillai P, Salma A, Sammet S, Ammirati M. Accuracy validation in a cadaver model of cranial neuronavigation using a surface autoregistration mask. *Neurosurgery.* 2010;67(SUPPL. 1):8-11. doi:10.1227/01.NEU.0000383751.63835.2F
41. Brinker T, Arango G, Kaminsky J, et al. An experimental approach to image guided skull base surgery employing a microscope-based neuronavigation system. *Acta Neurochir (Wien).* 1998;140(9):883-889. doi:10.1007/s007010050189
42. Steinmeier R, Rachinger J, Kaus M, Ganslandt O, Huk W, Fahlbusch R. Factors influencing the application accuracy of neuronavigation systems. *Stereotact Funct Neurosurg.* 2000;75(4):188-202. doi:10.1159/000048404
43. Smith TR, Mithal DS, Stadler JA, Asgarian C, Muro K, Rosenow JM. Impact of fiducial arrangement and registration sequence on target accuracy using a phantom frameless stereotactic navigation model. *J Clin Neurosci.* 2014;21(11):1976-1980. doi:10.1016/j.jocn.2014.04.006
44. Li Q, Guthikonda M, Mittal S. The analysis of affecting factors on the application accuracy for frameless. *IEEE.* 2016;(June). doi:10.1109/IEMBS.1997.757748
45. Omara AI, Wang M, Fan Y, Song Z. Anatomical landmarks for point-matching

registration in image-guided neurosurgery. *Int J Med Robot Comput Assist Surg*.

Published online 2013. doi:10.1002/rcs.1509

46. Woerdeman PA, Willems PWA, Noordmans HJ, Tulleken CAF, Van Der Sprenkel JWB. Application accuracy in frameless image-guided neurosurgery: A comparison study of three patient-to-image registration methods. *J Neurosurg*. 2007;106(6):1012-1016. doi:10.3171/jns.2007.106.6.1012
47. Rachinger J, Von Keller B, Ganslandt O, Fahlbusch R, Nimsky C. Application accuracy of automatic registration in frameless stereotaxy. *Stereotact Funct Neurosurg*. 2006;84(2-3):109-117. doi:10.1159/000094462
48. Pillai P, Sammet S, Ammirati M. Application accuracy of computed tomography-based, image-guided navigation of temporal bone. *Neurosurgery*. 2008;63(4 SUPPL.):326-333. doi:10.1227/01.NEU.0000316429.19314.67
49. Ballesteros-Zebadúa P, García-Garduño OA, Galván de la Cruz OO, Arellano-Reynoso A, Lárraga-Gutiérrez JM, Celis MA. Assessment of an image-guided neurosurgery system using a head phantom. *Br J Neurosurg*. 2016;30(6):606-610. doi:10.3109/02688697.2016.1173188
50. Paraskevopoulos D, Unterberg A, Metzner R, Dreyhaupt J, Eggers G, Wirtz CR. Comparative study of application accuracy of two frameless neuronavigation systems: Experimental error assessment quantifying registration methods and clinically influencing factors. *Neurosurg Rev*. 2011;34(2):217-228. doi:10.1007/s10143-010-0302-5
51. Mongen MA, Willems PWA. Current accuracy of surface matching compared to adhesive markers in patient-to-image registration. *Acta Neurochir (Wien)*. 2019;161(5):865-870. doi:10.1007/s00701-019-03867-8
52. Fitzpatrick JM. Fiducial registration error and target registration error are uncorrelated. *Med Imaging 2009 Vis Image-Guided Proced Model*. 2009;7261:726102. doi:10.1117/12.813601
53. Eleopra R, Rinaldo S, Devigili G, et al. Frameless deep brain stimulation surgery: A single-center experience and retrospective analysis of placement accuracy of 220 electrodes in a series of 110 patients. *Stereotact Funct Neurosurg*. 2020;97(5-6):337-346. doi:10.1159/000503335

54. Aulenta F, Lens P. Recent advances in augmented reality. *N Biotechnol*. 2011;29(1):1. doi:10.1016/S1871-6784(11)00246-9
55. Frantz T, Jansen B, Duerinck J, Vandemeulebroucke J. Augmenting microsoft's HoloLens with vuforia tracking for neuronavigation. *Healthc Technol Lett*. 2018;5(5):221-225. doi:10.1049/htl.2018.5079
56. Datteri RD, Dawant BM. Estimation and reduction of target registration error. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2012;7512 LNCS:139-146. doi:10.1007/978-3-642-33454-2_18
57. Shamir RR, Joskowicz L, Shoshan Y. Fiducial optimization for minimal target registration error in image-guided neurosurgery. *IEEE Trans Med Imaging*. 2012;31(3):725-737. doi:10.1109/TMI.2011.2175939
58. West JB, Fitzpatrick JM, Toms SA, Maurer CR, Maciunas RJ. Fiducial point placement and the accuracy of point-based, rigid body registration. *Neurosurgery*. 2001;48(4):810-817. doi:10.1227/00006123-200104000-00023

Figure legends

Figure 1: study flow diagram. ARN: augmented reality neuronavigation, CIN: conventional infrared neuronavigation.

Figure 2: different measurement methods. A: difference between TRE and FRE. 1. The distance between the registered fiducial position and the fiducial position after matching (=FRE). 2. The distance between a virtual target point, not used for registration, and the corresponding real target point (=TRE). B: difference between TRE and targeting error. Both the distance between a virtual target point and the corresponding real target point, but the targeting error is measured in 2D space and TRE in 3D space. C: target deviation (Li et al. 2018). Distance between the tip of the EVD and planned target. D: deviation compared to CIN (Incekara et al. 2018). Distance between the measured tumour centre based on ARN and CIN. ARN: Augmented Reality Neuronavigation, CIN: Conventional Infrared Neuronavigation, EVD: External Ventricle Drain, FRE: Fiducial Registration Error, TRE: Target Registration Error

Figure 3: forest plot of ARN compared to CIN for TRE measurements. ARN: augmented reality neuronavigation, CIN: conventional infrared neuronavigation, TRE: target registration error

Study	Display type	Measurement object	Registration method	Accuracy unit	Accuracy in mm (SD)	Measurement method
Zhang et al. (2017) ¹⁷	Integral videography	Phantom	Point-pair matching; corners of cube	FRE and TRE	FRE: 1.2 (0.11) N=4 TRE: 0.6 (0.34) N=3	Difference between real target and virtual target coordinates, obtained by an infrared tracked probe
King et al. (1999) ¹⁹	HUD	Phantom and patient	Point-pair matching; bone implants	Targeting error	0.9 (not defined)	Predicting accuracy by mathematical simulation
Edwards et al. (1995) ²⁰	HUD	Phantom	Point-pair matching; fiducials or surface matching	Targeting error	1.1 (not defined) N=18	Distance between the real target and the corresponding virtual target was measured with a caliper
Carl et al. (2019) ²³	HUD	Patient	Point-pair matching; stickers	TRE	2.9 (1.30) N=28	Difference between real target and virtual target coordinates, obtained by an infrared tracked probe
Carl et al. (2018) ²⁴	HUD	Patient	Self-designed automatic registration using intraoperative CT	TRE	1.4 (0.79) N=21	Difference between real target and virtual target coordinates, obtained by an infrared tracked probe
Li et al. (2018) ²⁶	HMD	Patient	Manual	Target deviation	AR 4.3 (1.63) N=15	Euclidean distance between planned target and actual catheter tip on postoperative CT
Van Doormaal et al. (2019) ¹⁸	HMD	Phantom and patient	Point-pair matching; stickers	FRE	Phantom: 7.2 (1.80) N=10 Patient: 4.4 (2.50) N=3	Automatically calculated by HMD application
Incekara et al. (2018) ²⁵	HMD	Patient	Manual	Deviation compared to CIN	4.4 (4.8) N=25	Difference between tumour centre on skin shown by hologram compared to conventional neuronavigation
McJunkin et al. (2019) ²¹	HMD	Phantom	Manual	TRE	5.8 (0.54) N=21	Difference between real target and virtual target coordinates with a tracked probe linked to the HMD
Maruyama et al. (2017) ²⁷	HMD	Phantom	Point-pair matching; anatomical landmarks	Targeting error	1.9 (1.1) N=100	Measuring difference with ruler between virtual and real target in X and Y axis
Maruyama et al. (2018) ²⁸	HMD	Phantom and patient	Point-pair matching; points on phantom	Targeting error	Phantom: 3.1 (1.9) N=75 Patients: 2.1 (1.1) N=40	Photographs were taken from the phantom from 5 different angles and distance were measured between real and visual target
Meulstee et al. (2018) ²²	HMD	Phantom	Reference frame with preoperative planning	TRE	2.3 (0.5) N=21	Measuring the overlay error in X, Y and Z-axis and calculating the TRE

Table 1: summary of extracted data for Augmented Reality Neuronavigation studies. SD: Standard Deviation, FRE: Fiducial Registration Error, TRE: Target Registration Error; HUD: Heads-Up Display, HMD: Head-Mounted Display; CIN: Conventional Infrared Neuronavigation

Measurement method	Description
Fiducial registration error	<p>The root-mean-squared distance between recognized fiducial positions and their homologous virtual fiducial positions after registration. This can be calculated by:</p> $FRE \equiv \sqrt{\frac{1}{n}((F_1 - F'_1)^2 + \dots + (F_n - F'_n)^2)}$ <p>Where F = recognized real fiducial position, F' = virtual fiducial position, n = total number of fiducials</p>
Target registration error	<p>The Euclidean distance between the real target point and virtual target point. This can be calculated by:</p> $TRE = T - T'$ <p>Where T is the real position of the defined target and T' is the virtual position of the defined target.</p>
Targeting error	The distance between the real target point and virtual target point, measured in 2D.
Target deviation	The distance between the preoperative determined target point for the drain tip and the actual drain tip on postoperative CT-imaging measured with the ruler function in the imaging program.
Deviation compared to CIN	The distance between the centre of the tumor circumference drawn on the patient's head using ARN and the centre of the tumor circumference drawn using CIN.

Table 2: The different measurement methods and the corresponding description

Pathologies	No. of patients	
	ARN (102)	CIN (245)
Oncology	78 (76.5%)	166 (67.8%)
Adenoid cystic carcinoma	1	
Adenoma	29	1
Chordoma	1	
Craniopharyngioma	2	1
Cysts	1	
Ependymoma	1	1
Epidermoid cyst	1	
Gliomas	14	59
Hemangioblastoma	1	1
histiocytosis		1
Lipoma	1	
Lymphoma	1	1
Meningioma	9	34
Metastasis	5	15
Neuroblastoma		1
PNET		1
Prolactinoma	1	
Rathke cyst	7	
Sarcoma	1	
Schwannoma	1	
Undefined	1	50
Vascular	9 (8.8%)	13 (5.3%)
Aneurysm	5	1
Arteriovenous malformations	4	2
Cavernous hemangioma		8
Intracerebral hemorrhage		2
Other	15 (14.7%)	5 (2.0%)
Abscess		1
Intracranial hypertension	15	4
Undefined	-	61

Table 3: An overview of the different pathologies of the patients in all included studies. ARN: Augmented Reality Neuronavigation, CIN: Conventional Infrared Neuronavigation, PNET: Primitive Neuroectodermal Tumour

Journal Pre-proof

Journal Pre-proof

Journal Pre-proof

Journal Pre-proof

Abbreviations list: 2D: 2-Dimensional, 3D: 3-Dimensional, ARN: Augmented Reality Neuronavigation, CIN: Conventional Infrared Neuronavigation, EVD: External Ventricle Drain, FLE: Fiducial Localization Error, FRE: Fiducial Registration Error, HMD: Head-Mounted Display, HUD: Heads-Up Display, RMS: Root Mean Square, TRE: Target Registration Error

Declaration of interests

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

T.P.C. van Doormaal is founder and CMO of AugmedIT, a startup company that develops Augmented Reality tools for surgeons.