Comparative effectiveness and safety of image guidance systems in neurosurgery: a preclinical randomized study

Hani J. Marcus, MRCS,1,2 Philip Pratt, PhD,1 Archie Hughes-Hallett, MRCS,1 Thomas P. Cundy, MBBS,1 Adam P. Marcus,3 Guang-Zhong Yang, PhD,1 Ara Darzi, MD,1 and Dipankar Nandi, DPhil2

1The Hamlyn Centre for Robotic Surgery, Institute of Global Health Innovation, and 2Faculty of Medicine, Imperial College London, United Kingdom; and 3Department of Neurosurgery, Imperial College Healthcare NHS Trust, London, United Kingdom

OBJECT Over the last decade, image guidance systems have been widely adopted in neurosurgery. Nonetheless, the evidence supporting the use of these systems in surgery remains limited. The aim of this study was to compare simultaneously the effectiveness and safety of various image guidance systems against that of standard surgery.

METHODS In this preclinical, randomized study, 50 novice surgeons were allocated to one of the following groups: 1) no image guidance, 2) triplanar display, 3) always-on solid overlay, 4) always-on wire mesh overlay, and 5) on-demand inverse realism overlay. Each participant was asked to identify a basilar tip aneurysm in a validated model head. The primary outcomes were time to task completion (in seconds) and tool path length (in mm). The secondary outcomes were recognition of an unexpected finding (i.e., a surgical clip) and subjective depth perception using a Likert scale.

RESULTS The time to task completion and tool path length were significantly lower when using any form of image guidance compared with no image guidance (p < 0.001 and p = 0.003, respectively). The tool path distance was also lower in groups using augmented reality compared with triplanar display (p = 0.010). Always-on solid overlay resulted in the greatest inattentional blindness (20% recognition of unexpected finding). Wire mesh and on-demand overlays mitigated, but did not negate, inattentional blindness and were comparable to triplanar display (40% recognition of unexpected finding in all groups). Wire mesh and inverse realism overlays also resulted in better subjective depth perception than always-on solid overlay (p = 0.031 and p = 0.008, respectively).

CONCLUSIONS New augmented reality platforms may improve performance in less-experienced surgeons. However, all image display modalities, including existing triplanar displays, carry a risk of inattentional blindness.

http://thejns.org/doi/abs/10.3171/2014.10.JNS141662

KEY WORDS neurosurgery; minimally invasive surgery; image guidance; augmented reality; diagnostic and operative techniques

Image guidance systems are increasingly important tools in surgery and have been widely adopted in neurosurgery over the last decade. Nonetheless, the evidence for the effectiveness and safety of these image guidance systems remains limited. Case-control studies report that the use of image guidance is associated with improved patient outcomes when compared with standard surgery,14 but commentators have suggested that clinical randomized studies are now neither practical nor ethical.17

The final component of all image guidance systems is the display of images in such a way that the surgeon is able to localize a point in the operative field unambiguously. Multiplanar reformatted mode has been used since the advent of CT imaging, and triplanar displays presenting axial, sagittal, and coronal views remain the most widely used technique. A drawback of triplanar displays is that surgeons must use the information from image slices to construct, in their own minds, a potentially complex 3D representation of anatomical and pathological structures. Moreover, triplanar displays require that surgeons stop operating momentarily, apply a probe to the region of interest (potentially near critical neurovascular structures), and then take their eyes off the surgical field to view the image guidance monitors. The fusion of virtual 3D models and the actual operating field to provide an augmented reality may enhance the operating room workflow.
Augmented reality systems have been described, albeit to a limited extent, in combination with endoscopy to assist during minimally invasive surgical procedures.\textsuperscript{1,2,7,19} Results of case series using augmented reality to aid surgical localization have generally been encouraging, suggesting that such technology may improve the operating room workflow.\textsuperscript{2,7,10,13} But thus far, no randomized studies have confirmed the effectiveness of such systems beyond standard triplanar image display. In addition, recent studies have highlighted several concerns with the use of augmented reality.\textsuperscript{3,6,13,18}

Arguably the greatest issue with augmented reality systems is that overlays may alter the attention of surgeons such that the surgeons fail to recognize critical events within the surgical field, such as unexpected complications. Preclinical studies have confirmed that augmented reality displays using always-on solid overlays may exacerbate inattentional blindness, raising major concerns over patient safety that must be addressed before widespread adoption of the technology into mainstream surgical practice can occur.\textsuperscript{3} The use of wire mesh rather than solid overlays, or on-demand rather than always-on augmented reality displays, may mitigate or negate inattentional blindness.\textsuperscript{3}

A further problem with augmented reality systems is that, at present, endoscopic live images are generally 2D, and virtual overlays, therefore, are presented in kind, significantly impairing depth perception. The increasing availability of 3D endoscopy permits stereoscopic augmented reality systems, but many users continue to have difficulty appreciating the depth of solid overlays despite binocular and kinetic cues.\textsuperscript{6,13,18} It has been suggested that wire mesh rather than solid overlays might improve depth perception.\textsuperscript{13} Another potential solution, termed “inverse realism,” provides “see-through vision” of the embedded virtual object while maintaining the salient anatomical structures of the exposed surface, which partially occlude the object.\textsuperscript{9}

The aim of this study was to compare simultaneously the effectiveness and safety of various image guidance systems using triplanar and augmented reality displays, with standard surgery, using a keyhole neurosurgical approach as an exemplar.

**Methods**

The Imperial College Joint Research Compliance Office approved the study protocol. The Consolidated Standards of Reporting Trials statement was used in the preparation of this manuscript.\textsuperscript{16}

**Participants and Study Settings**

Fifty novices were recruited from one university hospital. Participants were deemed suitable for inclusion if they had no prior experience of endoscopic or endoscope-assisted surgery (performed zero). Written informed consent was obtained from all participants.

**Trial Design**

A preclinical, randomized study design was adopted, comparing the following: 1) no image guidance; 2) triplanar display of axial, sagittal, and coronal images; 3) always-on solid overlay augmented reality; 4) always-on wire mesh overlay augmented reality; and 5) on-demand inverse realism augmented reality.

The Modelled Anatomical Replica for Training Young Neurosurgeons (MARTYN) head (Royal College of Surgeons of England), with an accompanying circle of Willis including a basilar tip aneurysm, was used.\textsuperscript{11} The model consists of a gelatin-based brain encased within a latex dura and a polyurethane skull. A previous study has confirmed that the MARTYN head is realistic (face validity), useful (content validity), and able to discriminate between surgeons of different experience (construct validity), with respect to the supraorbital subfrontal approach.\textsuperscript{12} A 25 × 15–mm left supraorbital craniotomy was fashioned using a high-speed drill (B. Braun, Melsungen AG). The high-speed drill was then used to remove the inner edge of the bone above the orbital rim, and the jugae cerebrali. A simple “C” durotomy was performed and the flap retracted basally.

A VisionSense III neuroendoscopy system (VisionSense Corp.) was used for visualization. The high-definition, zero degree, rigid endoscope is 4 mm in diameter and 18 cm long, providing a resolution of 1920 × 1080 pixels. Images were displayed using a 42-in (about 107 cm) stereoscopic screen.

A CT scan of the MARTYN head was performed and the vascular tree and aneurysm manually segmented using itk-SNAP version 2.4.0 software (www.itksnap.org),\textsuperscript{20} and smoothed and decimated using MeshLab version 1.3.2 (www.meshlab.sourceforge.net). The model head was fixed in place with a Mayfield clamp and a Budde-halo retractor system attached (Integra LifeSciences Corp.). An NDI Polaris Optical Tracking System (Northern Digital Inc.) was used to track the endoscope and image guidance probe with respect to a reference frame. Rigid registration of surface fiducials was used to relate the head and CT scan coordinate frames. Manual alignment of a reference object was then used to determine the hand-eye transformation from the camera tracking frame to the camera frame defining the projection to display coordinates. Concatenated together, the results of these calibrations were used to map renderings of the CT segmentations onto captured images of the model head. Custom software was used to generate the different image display modalities and to record the probe path length (Figs. 1 and 2).

By a computer-generated sequence participants were randomly allocated into 5 groups to determine which image display modality was used. Blocked randomization was used to ensure that 10 participants were evenly allocated into each group. Each participant was shown a short video demonstrating the endoscopic supraorbital approach in the MARTYN head and given several minutes to familiarize themselves with the image guidance system. They were then asked to identify the basilar tip aneurysm using a probe, with instructions to minimize their exposure and manipulation of brain tissue with use of cottonoid patties (Codman and Shurtleff Inc.), as they deemed appropriate. The task was considered complete when users applied the probe to the aneurysm.

A surgical clip was placed over the left posterior cere-
bral artery, approximately 5 mm from the aneurysm, and within the surgical trajectory, providing an unexpected finding to assess inattentional blindness (Fig. 3).

**Outcomes**

The primary outcomes were time to task completion (in seconds), and tool path length (in mm). The secondary outcomes were recognition of the unexpected finding (prompted), and subjective depth perception (on a 5-point Likert scale). Whereas participants were aware of the image display modality they were using, the data analysts were blinded to their allocation.

**Statistical Analysis**

The sample size was calculated on the basis of recently published work, preliminary data, and anticipated ease of participant recruitment. It was estimated that to detect a reduction in time to task completion from 152 to 92 sec (SD 48 sec), with a 2-sided 5% significance level and a power of 80%, a sample size of at least 10 participants was necessary in each group.

Data were analyzed with SPSS version 20.0 (IBM Corp.). The median and interquartile ranges were calculated for all outcome measures, and nonparametric tests were performed, with a value of \( p < 0.05 \) considered statistically significant. We compared the time to task completion, tool path length, and subjective depth perception using the Kruskal-Wallis 1-way ANOVA. We compared the proportion of users recognizing the unexpected finding using the chi-square or Fisher exact test (if less than 80% of the cells had an expected frequency of 5 or greater). If a significant difference was identified in any of the outcomes, we then used the Bonferroni correction to compare the following groups directly (\( n = 4; p < 0.0125 \)): no image guidance versus any image guidance, triplanar image display versus any augmented reality display, always-on solid overlay versus on-demand inverse realism, and always-on solid overlay versus always-on mesh overlay.
Results

Baseline Demographic Data

The demographics of the participants are summarized in Table 1. Of the 50 participants recruited, 35 were medical students and 15 were junior doctors; none had yet embarked on formal surgical training. There was no significant difference in demographics or experience between the groups. All participants enrolled completed the study, and no losses occurred after randomization.

Primary Outcomes

The median and interquartile ranges of the primary outcomes are summarized in Table 2. The time to task completion and tool path length differed significantly among the study groups (p = 0.007 and p = 0.002, respectively) (Figs. 4 and 5). In subsequent analysis, the time to task completion and tool path length were significantly lower when using any form of image guidance (triplanar display, always-on solid overlay, always-on wire mesh overlay, and on-demand inverse realism overlay) compared with no image guidance (p < 0.001 and p = 0.003, respectively). The tool path distance was also lower in groups using augmented reality (always-on solid overlay, always-on wire mesh overlay, and on-demand inverse realism) compared with triplanar image display (p = 0.010). There was no significant difference in the primary outcomes between the different augmented reality overlays.

Secondary Outcomes

The secondary outcomes are summarized in Table 2. All videos were reviewed and the surgical clip was visible in all cases. The proportion of participants who recognized the unexpected finding differed significantly among the various study groups (p = 0.025). All but 1 participant in the control arm recognized the aneurysm clip (90% recognition of unexpected finding), while less than 50% recognized it when using any form of image guidance, including conventional triplanar display (p = 0.003). Always-on solid overlay resulted in the greatest inattentional blindness (20% recognition of unexpected finding). Always-on wire mesh and on-demand overlays mitigated, but did not negate, inattentional blindness and were comparable to triplanar displays (40% recognition of unexpected finding in all groups). The subjective depth perception also varied significantly among the augmented reality groups (p = 0.015). Solid overlay resulted in significantly worse subjective depth perception than wire mesh and inverse realism (p = 0.031 and p = 0.008, respectively).

Discussion

Image guidance technology is among the most rapidly emerging innovations in surgery. Although the evidence for the use of image guidance in surgery remains limited, systems using triplanar display have already achieved substantial clinical penetration in neurosurgery. The recent development and refinement of augmented reality displays offer the possibility of improved operating room workflow. In the coming years, technological advances are likely to promote further the dissemination of augmented reality technology in surgery. This randomized study is the first to simultaneously compare the effectiveness and safety of image guidance systems using triplanar and augmented reality displays, with that of standard surgery, in a validated preclinical model. To this end, this study’s findings confirm the utility of image guidance systems in neurosurgery and suggest that new augmented reality platforms with always-on wire mesh or on-demand inverse realism overlays may result in improved surgical performance. However, it must be acknowledged that all image display modalities, including existing triplanar display, carry a risk of inattentional blindness.

Image guidance systems have 2 potential roles in surgery: first, to help guide the overall surgical approach to pathology; second, to facilitate unambiguous tissue dissection, particularly in the context of oncological resection. Currently, most image guidance systems are designed for the former, allowing surgeons to define precisely a narrow surgical corridor to deep-seated pathology, thus minimiz-
ing the risk of iatrogenic injury. In the present preclinical, randomized study, all image guidance systems resulted in significantly reduced time to completion and tool path length compared with standard surgery (p < 0.001 and p = 0.003, respectively). Surprisingly, no clinical randomized studies have evaluated the role of contemporary image guidance systems to help define the surgical trajectory, and some commentators have suggested that performing such studies now would be neither practical nor ethical.17 Perhaps the largest case-control study in the surgical literature reported on patients undergoing meningioma resection at the National Hospital for Neurology and Neurosurgery, and concluded that the use of image guidance was associated with shorter operating times, reduced blood loss, and fewer major complications than standard surgery.14

In the present study, image guidance with augmented reality displays led to a significantly reduced tool path length compared with image guidance with existing triplanar display (p = 0.010). Augmented reality displays may improve the surgical workflow by obviating the need for the surgeon to repeatedly stop operating, apply the probe to the region of interest, and turn away from the operating field to view the image guidance monitor. In a previous preclinical, randomized study, augmented reality improved the accuracy of surgical trainees in identifying skull-base landmarks on a cadaver, compared with no image guidance.2 The investigators noted that they did not include a triplanar display as a comparison and, therefore, could not comment on the relative merit of augmented reality over existing image guidance platforms. Augmented reality systems have also been developed and applied to minimally invasive surgery clinically.7,8 The most cited of these studies is that of Kawamata et al., who developed an augmented reality system consisting of a 2.7-mm rigid endoscope (Olympus Corp.), an optical tracking system, and a controller to overlay endoscopic live images with wire mesh models of tumors and neighboring anatomical landmarks.7 In a series of 12 patients with pituitary tumors undergoing endoscopic transnasal transsphenoidal hypophysectomy, the authors felt that the system improved operative workflow.

The use of augmented reality has been reported to be associated with inattentional blindness in surgeons, rais-

### TABLE 2. Summary of results according to image guidance used

<table>
<thead>
<tr>
<th>Study Group</th>
<th>Median Time to Task Completion in Sec (IQR)</th>
<th>Median Path Length in mm (IQR)</th>
<th>Recognition of Unexpected Finding, No. (%)</th>
<th>Depth Perception*</th>
</tr>
</thead>
<tbody>
<tr>
<td>No image guidance</td>
<td>88.5 (61.8–150.5)</td>
<td>2938.5 (1411.8–5955.8)</td>
<td>9/10 (90.0)</td>
<td>—</td>
</tr>
<tr>
<td>Triplanar display</td>
<td>30.5 (16.5–64.8)</td>
<td>1384.0 (1183.8–2873.0)</td>
<td>4/10 (40.0)</td>
<td>—</td>
</tr>
<tr>
<td>Always-on solid overlay</td>
<td>19.5 (16.3–64.3)</td>
<td>795.0 (371.0–1624.8)</td>
<td>2/10 (20.0)</td>
<td>2.0 (1.0–3.0)</td>
</tr>
<tr>
<td>Always-on wire mesh overlay</td>
<td>25.5 (17.8–52.8)</td>
<td>641.0 (406.5–1798.5)</td>
<td>4/10 (40.0)</td>
<td>3.5 (2.0–4.0)</td>
</tr>
<tr>
<td>On-demand inverse realism overlay</td>
<td>26.0 (13.0–41.5)</td>
<td>639.5 (314.5–1068.8)</td>
<td>4/10 (40.0)</td>
<td>4.0 (2.8–4.0)</td>
</tr>
</tbody>
</table>

* Data are given as median Likert scale score (Likert scale score range), where 5 represents excellent depth perception and 1 represents poor depth perception.
† Kruskal-Wallis 1-way analysis of variance.
‡ Fisher exact test.

![FIG. 4](image)

**FIG. 4.** Graph illustrating time to completion with different image guidance displays. Wireframe = wire mesh.

![FIG. 5](image)

**FIG. 5.** Graph illustrating path distance with different image guidance displays. Circle represents an outlier (greater than 1.5 times the interquartile range).
ing important safety concerns. Dixon et al. compared 32 surgeons of varying experience performing an endonasal navigation exercise on a cadaver. Surgeons were randomized into groups performing the exercise with or without augmented reality. Although surgeons in the group using augmented reality were more accurate, they were less likely to identify unexpected findings. The authors of the study speculated that augmented reality may have led to perceptual blindness in a number of ways, including attentional tunneling, increased visual clutter and jitter, and an additional camouflage effect. In the present study, all image guidance displays resulted in considerable inattentional blindness (p = 0.003), despite the surgical clip being clearly visible in all the videos reviewed. Interestingly, image guidance with existing triplanar displays also resulted in comparable inattentional blindness (40% recognition of unexpected findings), suggesting that it is the cognitive load rather than the overlay per se that is important.

An additional concern with augmented reality systems has been the issue of subjective depth perception. At present, the majority of endoscopes in clinical use are 2D, with augmented reality overlays presented in kind. The recent introduction of 3D and high-definition endoscopes may allow for stereoscopic augmented reality systems. In keeping with existing literature in the field, the present study found that even with such binocular cues, many users continue to struggle to appreciate the depth of solid overlays (median, 2.0 on a Likert scale ranging from 1 to 5).6,9,13,18 To this end, wire mesh and inverse realism overlays resulted in better subjective depth perception (median, 3.5 and 4.0, respectively, on a Likert scale ranging from 1 to 5).

**Limitations**

It should be noted that this study has a number of limitations. All the study participants were novices yet to embark on formal surgical training. Selection of novices, rather than intermediates or experts, allowed for a relatively homogenous sample; this was essential, as the assessment of inattentional blindness with an unexpected finding meant a crossover study design could not be adopted.

Novice participants may have had difficulty identifying the aneurysm or recognizing the surgical clip as an abnormal finding. However, all participants observed a short video introducing the MARTYN head and demonstrating the circle of Willis and basilar tip aneurysm. Moreover, the fact that 90% of participants in the control group did recognize the aneurysm clip suggests the observed inattentional blindness was the result of the use of image guidance rather than surgical inexperience.

None of the study participants had prior experience with image guidance, and all were given a few minutes to familiarize themselves with the equipment. It is likely that greater familiarity would have resulted in improved performance and less inattentional blindness.

The supraorbital subfrontal approach was selected as an exemplar keyhole approach, but it is infrequently performed even among experienced neurosurgeons. The approach was chosen as it is recognized as technically challenging, and may benefit from image guidance.15 In addition, a preclinical model with face, content, and construct validity was readily available.12

The MARTYN head is of comparatively low fidelity, particularly with respect to its internal structure. The result is that the triplanar view lacks anatomical detail, and this may have had an impact on task performance. However, the use of segmented rather than plain images likely mitigated this (Fig. 2).

**Generalizability**

The generalizability of this study likely depends on several factors, including the experience of the surgeon and the complexity of the surgical approach. With greater experience, surgeons learn to use anatomical landmarks to guide their surgical trajectory, and the benefits of image guidance almost certainly lessen. Similarly, in relatively straightforward surgical approaches, there may be little advantage to the use of image guidance.

**Conclusions**

In this study, the use of image guidance systems with augmented reality overlays significantly reduced the time to task completion and tool path distance, but also increased the risk of inattentional blindness. These findings support the need for less-experienced surgeons using image guidance systems, particularly when undertaking complex approaches, to be carefully supervised by experts who are less cognitively loaded and better able to identify potential complications.

**Acknowledgments**

We thank Y. Saleh, G. Herbert, M. Edmondson, L. Chacko, and S. Morris for their assistance in recruiting participants. We also thank M. Cooke and the Royal College of Surgeons of England for developing and providing the MARTYN model head.

**References**


Author Contributions
Conception and design: all authors. Acquisition of data: Marcus, Pratt, Marcus. Analysis and interpretation of data: Marcus, Pratt, Hughes-Hallett, Cundy, Marcus. Drafting the article: Marcus. Critically revising the article: Pratt, Hughes-Hallett, Cundy, Marcus, Yang, Darzi, Nandi. Reviewed submitted version of manuscript: Pratt, Marcus, Yang, Darzi, Nandi. Approved the final version of the manuscript on behalf of all authors: Marcus. Statistical analysis: Marcus. Administrative/technical/material support: Yang, Darzi, Nandi. Study supervision: Yang, Darzi, Nandi.

Supplemental Information
Previous Presentation
Portions of this work were presented in poster form at the Spring Meeting for Clinician Scientists in Training 2015 held at the Royal College of Physicians in London, United Kingdom, on February 26, 2015.

Correspondence
Hani J. Marcus, Hamlyn Centre, Imperial College London and Imperial College Healthcare NHS Trust, Praed St., Paterson Bldg. (Level 3), London W2 1NY, UK. email: hani.marcus10@imperial.ac.uk.