Comparative Effectiveness of 3-Dimensional vs 2-Dimensional and High-Definition vs Standard-Definition Neuroendoscopy: A Preclinical Randomized Crossover Study

BACKGROUND: Although the potential benefits of 3-dimensional (3-D) vs 2-dimensional (2-D) and high-definition (HD) vs standard-definition (SD) endoscopic visualization have long been recognized in other surgical fields, such endoscopes are generally considered too large and bulky for use within the brain. The recent development of 3-D and HD neuroendoscopes may therefore herald improved depth perception, better appreciation of anatomic details, and improved overall surgical performance.

OBJECTIVE: To compare simultaneously the effectiveness of 3-D vs 2-D and HD vs SD neuroendoscopy.

METHODS: Ten novice neuroendoscopic surgeons were recruited from a university hospital. A preclinical randomized crossover study design was adopted to compare 3-D vs 2-D and HD vs SD neuroendoscopy. The primary outcomes were time to task completion and accuracy. The secondary outcomes were perceived task workload using the NASA (National Aeronautics and Space Administration) Task Load Index and subjective impressions of the endoscopes using a 5-point Likert scale.

RESULTS: Time to task completion was significantly shorter when using the 3-D vs the 2-D neuroendoscopy \((P = .001)\), and accuracy of probe placement was significantly greater when using the HD vs the SD neuroendoscopy \((P = .009)\). We found that 3-D endoscopy significantly improved perceived depth perception \((P < .001)\), HD endoscopy significantly improved perceived image quality \((P < .001)\), and both improved participants’ overall impression \((P < .001)\).

CONCLUSION: Three-dimensional neuroendoscopy and HD neuroendoscopy have differing but complementary effects on surgical performance, suggesting that neither alone can completely compensate for the lack of the other. There is therefore strong preclinical evidence to justify 3-D HD neuroendoscopy.

KEY WORDS: Endoscopy, HD, High definition, Minimally invasive neurosurgery, 3D, 3-dimensional

ENDOSCOPES & ENDOSCOPE-ASSISTED APPROACHES HAVE BEEN USED WITHIN THE FIELD OF NEUROSURGERY SINCE THE WORK OF PIONEERS SUCH AS WALTER DANDY ALMOST A CENTURY AGO. \(^1\) ALTHOUGH THE POTENTIAL BENEFITS OF 3-DIMENSIONAL (3-D) VS 2-DIMENSIONAL (2-D) AND HIGH-DEFINITION (HD) VS STANDARD-DEFINITION (SD) ENDOSCOPIC VISUALIZATION HAVE LONG BEEN RECOGNIZED IN OTHER SURGICAL FIELDS, SUCH ENDOSCOPES ARE GENERALLY CONSIDERED TOO LARGE AND BULKY FOR USE IN THE BRAIN. \(^2-6\) THE RECENT DEVELOPMENT OF 3-D AND HD NEUROENDOSCOPIES MAY THEREFORE HERALD IMPROVED DEPTH PERCEPTION, BETTER APPRECIATION OF ANATOMIC DETAILS, AND IMPROVED SURGICAL PERFORMANCE COMPARED WITH CONVENTIONAL NEUROENDOSCOPY. \(^7-10\) THESE BENEFITS MUST BE BALANCED, HOWEVER, AGAINST THE HIGHER COST, LARGER SIZE, AND GREATER WEIGHT OF HD ENDOSCOPES. MOREOVER, THE HUMAN VISUAL SYSTEM IS EXQUISITELY SENSITIVE TO STEREOSCOPIC CUES, AND ALTHOUGH STEREO FUSION CAN TOLERATE QUITE MAJOR...
artifacts, symptoms such as diplopia and nausea have been described with prolonged use of 3-D endoscopy systems. Previous studies have sought to assess the impact of either 3-D or HD neuroendoscopy on surgical performance, with mixed findings. A major limitation of all these studies is that they do not allow comparison of the effectiveness of 3-D and HD neuroendoscopy. It remains unclear whether, for example, HD neuroendoscopy provides sufficient monocular cues to obviate the need for true 3-D endoscopy. The aim of this study was therefore to compare simultaneously the effectiveness of 3-D vs 2-D and HD vs SD neuroendoscopy.

PATIENTS AND METHODS

The trial protocol was approved by the Imperial College Joint Research Compliance Office. The Consolidated Standards of Reporting Trials statement was used in the preparation of this manuscript.

Participants and Study Settings

Ten novice neuroendoscopic surgeons were recruited from a university hospital. Participants were deemed suitable for inclusion if they had no earlier experience with endoscopic or endoscope-assisted neurosurgery (performed zero). Informed consent was obtained from all participants.

Trial Design

A preclinical randomized crossover study design was adopted comparing 3-D against 2-D and HD against SD neuroendoscopy. A Sawbones skull and brain (Pacific Research Laboratories, Inc, Vashon Island, Washington) with an accompanying circle of Willis including an anterior communicating artery aneurysm was used. A 25 × 15-mm left supraorbital craniotomy was fashioned, and 5 colored targets were placed around the surgical field approximately 30 mm in diameter and 30 to 60 mm in depth.

A VisionSense III neuroendoscopy system (VisionSense, Petach Tikva, Israel) was used for visualization. The SD 0° rigid VisionSense endoscope was 4.9 mm in diameter and 20 cm in length, providing a resolution of 640 × 480 pixels. The HD 0° rigid VisionSense endoscope was 4 mm in diameter and 18 cm in length, providing a resolution of 1920 × 1200 pixels. Images were displayed using a 24-in stereoscopic flat-screen system and switched between 3-D and 2-D (left eye) using the system’s toggle.

Participants were randomly allocated using a computer-generated sequence into groups to determine the order in which 2-D SD, 3-D SD, 2-D HD, or 3-D HD neuroendoscopy was used. Blocked randomization was used to ensure that an equal number of participants began with 3-D vs 2-D and HD vs SD neuroendoscopy.

Each participant was asked to place a probe on targets in a predetermined random sequence of 10 colors (see Figure 1). Participants were instructed to be both quick and accurate, placing the probe as close to the center of the colored targets as possible. This process was repeated on 3 occasions with each endoscope configuration.

Outcomes

The primary outcomes were time to task completion (seconds) and accuracy with which probes were placed on targets (score). To determine accuracy, all recorded videos were reviewed and scored independently by 2 observers (H.J.M. and A.H.H.) on how closely each probe was to the center of the colored targets; participants were scored 3 points if they directly made contact with the 1-mm colored target, 2 points if they made contact with the surrounding 1-mm black line, 1 point if they made contact with the surrounding 2-mm white line, and zero points if they missed the target entirely (see Figure 2). Whereas participants were aware of the endoscope they were using, the data analysts were partially blinded to the allocation (to 3-D vs 2-D but not HD vs SD).

The secondary outcomes were perceived task workload using the NASA (National Aeronautics and Space Administration) Task Load Index and subjective impressions of the endoscopes using a 5-point Likert scale. Participants were asked after using each endoscope to what extent they considered the system provided high-quality images, if it allowed high-fidelity depth perception, and if they would like to use the visualization modality again.

Statistical Analysis

The sample size was calculated on the basis of recently published work. We estimated that to detect a reduction in time to task...
accuracy of probe placement with different neuroendoscopes were significantly different \( (P = .005 \text{ and } P = .021, \text{ respectively}) \); they are illustrated in Figures 3 and 4.

The medians and interquartile ranges of the primary outcome data stratified according to 3-D vs 2-D and HD vs SD neuroendoscopy are summarized in Tables 3 and 4, respectively. Post hoc statistical analysis suggested that time to completion was significantly shorter when using 3-D vs 2-D neuroendoscopy \( (P = .001) \) and that the accuracy of probe placement was significantly greater when using HD vs SD neuroendoscopy \( (P = .009) \).

### Secondary Outcomes

The medians and interquartile ranges of the secondary outcome data are summarized in Table 2. The subjective impressions of the endoscopes using a 5-point Likert scale varied significantly \( (P < .001) \), but the perceived task workload according to the Task Load Index did not reach statistical significance \( (P = .161) \).

The medians and interquartile ranges of the secondary outcome data stratified according to 3-D vs 2-D and HD vs SD neuroendoscopy are summarized in Tables 2 and 3, respectively. Post hoc statistical analysis suggested that 3-D neuroendoscopy significantly improved perceived image quality \( (P < .001) \), and both improved the overall likelihood that participants would use the modality again \( (P < .001) \). Because the Kruskal-Wallis analysis failed to demonstrate a significant difference in the cognitive workload with different endoscopes, no further analysis was performed.

### DISCUSSION

The advantages of 3-D and HD endoscopy have been demonstrated in other surgical fields, but until recently, the large sizes of such endoscopes restricted their use in the brain.\textsuperscript{2-6} This preclinical randomized crossover study is the first to compare simultaneously the effectiveness of 3-D vs 2-D and HD vs SD neuroendoscopy. Interestingly, the effects of 3-D and HD neuroendoscopy appear to be distinct and complementary; the use of 3-D vs 2-D neuroendoscopy led to a significant reduction in completion from 90 to 66 seconds (standard deviation, 30 seconds), with a 2-sided 5% significance level and a power of 80%, a sample size of 10 participants was necessary.

Data were analyzed with SPSS version 20.0 (Chicago, Illinois). The Cronbach \( \alpha \) was used to assess the interrater reliability of accuracy scores. The medians and interquartile ranges were calculated for all outcome measures and nonparametric tests performed, with a value of \( P < .05 \) considered statistically significant. We compared 2-D SD, 3-D SD, 2-D HD, and 3-D HD neuroendoscopy using the Kruskal-Wallis 1-way analysis of variance. Subsequently, if a significant difference was identified, the Mann-Whitney \( U \) test was used to compare 3-D vs 2-D and HD vs SD neuroendoscopy with the Bonferroni correction. We also performed a post hoc analysis comparing the data from the 2 novice neuroendoscopic surgeons with extensive laparoscopic experience against the other participants to determine whether there was a significant difference in performance.

### RESULTS

#### Baseline Demographic Data

The demographics of the 10 participants are summarized in Table 1. All participants completed the study, and no losses occurred after randomization. Two of the subjects had considerable experience with laparoscopic surgery (performed \( \geq 50 \) cases) but had no experience with neuroendoscopy and thus were included. Post hoc analysis confirmed no significant difference in performance.

#### Primary Outcomes

The Cronbach \( \alpha \) demonstrated excellent interrater reliability when scoring the accuracy of probe placement \( (\alpha = 0.925) \). The medians and interquartile ranges of the primary outcomes are summarized in Table 2. The time to task completion and

<table>
<thead>
<tr>
<th>TABLE 1. Demographics of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex, male/female</td>
</tr>
<tr>
<td>Age, median (range), y</td>
</tr>
<tr>
<td>Handedness, R:L</td>
</tr>
</tbody>
</table>

#### Secondary Outcomes

<table>
<thead>
<tr>
<th>TABLE 2. Summary of Results According to Neuroendoscope Used\textsuperscript{a}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Time, s\textsuperscript{b}</td>
</tr>
<tr>
<td>Accuracy\textsuperscript{b}</td>
</tr>
<tr>
<td>NASA-TLX</td>
</tr>
<tr>
<td>Likert scale Quality\textsuperscript{b}</td>
</tr>
<tr>
<td>Depth\textsuperscript{b}</td>
</tr>
<tr>
<td>Overall\textsuperscript{b}</td>
</tr>
</tbody>
</table>

\textsuperscript{a}HD, high definition; NASA-TLX, National Aeronautics and Space Administration Task Load Index; SD, standard definition; 2-D, 2-dimensional; 3-D, 3-dimensional.

\textsuperscript{b}Likert scale: 5 = high quality, excellent depth perception, and overall would use again; 1 = low quality, poor depth perception, and overall would not use again. Values are median (interquartile range) reported.
in the time to task completion and a subjective improvement in depth perception, whereas the use of HD vs SD neuroendoscopy led to a significant increase in the accuracy of probe placement and a subjective improvement in image quality.

Although no previous studies have addressed the influence of both 3-D and HD neuroendoscopy, similar studies have assessed their impact individually with varying results. Fraser et al. used the VisionSense II system to compare 3-D and 2-D neuroendoscopy in a model simulating the transnasal transsphenoidal approach to the pituitary. In all, 33 participants with varying levels of experience were asked to use a rongeur to remove the sellar floor and then take 4 small pituitary biopsies. Although 3-D neuroendoscopy resulted in improved cutting efficiency \((P = 0.04)\) and was subjectively the user preference, it was not associated with a significant difference in other primary outcomes such as time to completion or error rates. Thus far, only limited clinical studies have directly compared 3-D and 2-D neuroendoscopy, and most have failed to demonstrate any significant differences. In a recent retrospective cohort study, Barkhoudarian et al. analyzed 160 transnasal transsphenoidal procedures, of which 65 were performed with a 3-D neuroendoscope and 95 with a 2-D neuroendoscope. Although there was no significant difference in overall operating time, within the disease-specific comparison, pituitary adenoma resection was shorter with 3-D vs 2-D neuroendoscopy (174 vs 147 minutes; \(P = 0.03\)); there was no difference in the rate of gross total resection or complications.

### Table 3. Summary of Results Stratified Into 3- vs 2-Dimensional

<table>
<thead>
<tr>
<th></th>
<th>2-D (2-D SD and 2-D HD)</th>
<th>3-D (3-D SD and 3-D HD)</th>
<th>(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time, s</td>
<td>67.5 (52.0-81.3)</td>
<td>52.0 (42.5-68.3)</td>
<td>.001</td>
</tr>
<tr>
<td>Accuracy</td>
<td>15.5 (12.9-19.4)</td>
<td>17.0 (14.0-21.3)</td>
<td>0.085</td>
</tr>
<tr>
<td>NASA-TLX</td>
<td>39.7 (31.2-53.4)</td>
<td>28.9 (24.2-43.3)</td>
<td>NA</td>
</tr>
<tr>
<td>Likert scale</td>
<td>Quality</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.5 (2.0-4.0)</td>
<td>4.0 (3.0-5.0)</td>
<td>0.460</td>
</tr>
<tr>
<td></td>
<td>Depth</td>
<td>2.0 (2.0-3.0)</td>
<td>4.0 (4.0-4.0)</td>
</tr>
<tr>
<td></td>
<td>Overall</td>
<td>3.0 (2.8-4.0)</td>
<td>4.0 (4.0-5.0)</td>
</tr>
</tbody>
</table>

*HD, high definition; NASA-TLX, National Aeronautics and Space Administration Task Load Index; SD, standard definition; 2-D, 2-dimensional; 3-D, 3-dimensional.

*Likert scale: 5 = high quality, excellent depth perception, and overall would use again; 1 = low quality, poor depth perception, and overall would not use again. Values are median (interquartile range).

### Table 4. Summary of Results Stratified Into Standard Definition vs High Definition

<table>
<thead>
<tr>
<th></th>
<th>SD (2-D SD and 3-D SD)</th>
<th>HD (2-D HD and 3-D HD)</th>
<th>(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time, s</td>
<td>63.0 (49.0-76.5)</td>
<td>54.0 (45.8-69.5)</td>
<td>0.110</td>
</tr>
<tr>
<td>Accuracy</td>
<td>14.8 (13.0-18.1)</td>
<td>17.5 (14.5-22.6)</td>
<td>.009</td>
</tr>
<tr>
<td>NASA-TLX</td>
<td>39.7 (27.5-47.8)</td>
<td>31.5 (22.6-45.5)</td>
<td>NA</td>
</tr>
<tr>
<td>Likert scale</td>
<td>Quality</td>
<td>2.0 (2.0-3.0)</td>
<td>5.0 (4.0-5.0)</td>
</tr>
<tr>
<td></td>
<td>Depth</td>
<td>3.0 (2.0-4.0)</td>
<td>4.0 (3.0-4.3)</td>
</tr>
<tr>
<td></td>
<td>Overall</td>
<td>3.0 (2.8-4.0)</td>
<td>4.0 (4.0-5.0)</td>
</tr>
</tbody>
</table>

*HD, high definition; NASA-TLX, National Aeronautics and Space Administration Task Load Index; SD, standard definition; 2-D, 2-dimensional; 3-D, 3-dimensional.

*Likert scale: 5 = high quality, excellent depth perception, and overall would use again; 1 = low quality, poor depth perception, and overall would not use again. Values are median (interquartile range).
Studies have also assessed the impact of HD vs SD neuroendoscopy on surgical performance. Schroeder and Nehlsen reported higher image resolution and color fidelity with HD neuroendoscopy, particularly when discriminating tumor from neighboring tissue during transnasal transphenoidal approaches. Conrad et al captured a series of images during endoscopic approaches and found significantly improved recognition of anatomic landmarks by surgeons when using HD vs SD cameras (84.4% vs 63.0%; $P = .01$). The clinical significance of HD visualization has not yet been ascertained, although it is commonly assumed.

The major limitation of all the above studies is that they address the effects of either 3-D or HD neuroendoscopy, rather than both, making it difficult to tease apart their relative importance and the extent to which each can compensate for the other. It is well recognized, for example, that a number of monocular cues can contribute to depth perception such as motion parallax, the kinetic depth effect, and pictorial cues (eg, size, perspective, texture, interposition, lighting, and shadow).

It has been suggested that HD endoscopy, by allowing improved recognition of these monocular cues, might obviate the need for stereoscopy. The present study provides firm evidence that although 3-D and HD neuroendoscopy individually improve surgical performance, they do so in different ways, and neither can fully compensate for the lack of the other.

Limitations

It should be noted that this study has a number of limitations. First, although all participants were novices in neuroendoscopy, 2 had substantial experience in laparoscopic surgery. Post hoc analysis failed to demonstrate a significant difference between these 2 surgeons and the remaining participants, but it remains likely that this influenced their performance (albeit with a trend toward reduced time to task completion and improved accuracy of probe placement). Second, although the task was based on similar externally validated measures of surgical performance, it has not itself been validated. Moreover, the duration of the task may not have been long enough for participants to experience symptoms such as diplopia and nausea that have been described with prolonged use of 3-D endoscopy. Third, although the use of 3-D neuroendoscopes that were toggled to 2-D, rather than dedicated 2-D neuroendoscopes, allowed control of the video capture and display hardware, it might have led to a somewhat lower image quality. Finally, an inherent limitation of the methodology was that video was captured with the endoscope being assessed and the accuracy of probe placement was therefore only partially blinded, with researchers able to distinguish HD vs SD but not 3-D vs 2-D. Unfortunately, the small size of the keyhole supraorbital craniotomy made it difficult to place a further endoscope or camera without obstructing access.

Generalizability

Clearly, the generalizability of these findings is likely to depend on several factors. In this study, only novice neuroendoscopists were included, in part because of the difficulty in recruiting sufficient numbers of experienced neurosurgeons (particularly because the sample size would have to be substantially larger to detect a presumably smaller effect on performance). The general surgical literature suggests that the influence of 3-D vs 2-D endoscopy on surgical performance is reduced or negated with experience as surgeons learn to use monocular cues to judge distance. The relative influence of 3-D and HD neuroendoscopy is also likely to vary depending on the nature of the surgical task performed. We speculate that complex procedures such as neurovascular dissection would be far more influenced by the nature of the endoscope used than relatively straightforward procedures.

CONCLUSION

The results of this study may have considerable implications on endoscopic and endoscope-assisted neurosurgical approaches. Importantly, the fact that 3-D vs 2-D and HD vs SD neuroendoscopy did differing but complementary effects suggests that neither 3-D nor HD alone can completely compensate for the lack of the other. There is therefore strong preclinical evidence for the development and use of next-generation 3-D HD neuroendoscopes, particularly for inexperienced surgeons or when complex surgery is performed. Further studies are merited to confirm that no side effects or adverse reactions occur with prolonged use and that these findings are translated into improved surgical performance in a clinical setting.

Disclosures

Dr Marcus is supported by an Imperial College Wellcome Trust Clinical Fellowship. The authors have no personal, financial, or institutional interest in any of the drugs, materials, or devices described in this article.

REFERENCES


COMMENTS

The authors should be congratulated for this important study. 3-Dimensional (3-D) endoscopy will certainly become more important in endoscopy and endoscopic neurosurgery. Most promising here is indeed the VisionSense principle with “chip in the tip” using bee’s eye technology (“compound eye”) and with LED illumination (avoiding the heat of xenon light), which is now also available in full high-definition (HD) resolution (VisionSense III, “a single sensor divided into hundreds of thousands of tiny eyes”). This scope now has even a decreased outer diameter of <=4 mm (actual, 3.3 mm). It corresponds to a rigid HOPKINS scope required for effective 3-D imaging (should also not be below 3.5 mm in diameter; otherwise, the available 3-D minicameras will not provide an adequate resolution and brightness).

This study is of special interest and value because it is a randomized crossover study in which 10 novice neurosurgeons without neuroendoscopic experience performed endoscopic surgery in an anatomic model with a supraorbital craniotomy (2 participants had experience with laparoscopy). The randomization was quite complete and skillful; blocked randomization ensured that equal numbers of participants started with 3-D vs 2-dimensional (2-D) and HD vs standard-definition (SD) endoscopy.

As could be expected (but was not yet demonstrated in such a statistical setup), the accuracy of the task performed (probe placement) was significantly better with HD vs SD endoscopy. More interesting is the result that time to completion was significantly shorter with 3-D vs 2-D endoscopic imaging; without longer training in 3-D endoscopy, one would expect a steeper learning curve with a 3-D screen. This result shows the advanced performance of the VisionSense III system. The task work loading was not different, and the overall likelihood of the participants was greatest when 3-D neuroendoscopy was combined with HD resolution. In a clinical pilot study, which is not cited by the authors, otolaryngologists found a better depth perception and orientation and an improved completeness of surgery in sinonasal surgery comparing 2-D conventional endoscopy with an earlier version of the VisionSense 3D system.

The authors also provide an adequate, restrictive Discussion, pointing to the limitations of the study regarding the short duration of the task, which would not exclude the occurrence of discomfort with prolonged 3-D endoscopy like diplopia and nausea. The conclusion that neither 3-D nor HD alone can compensate for the lack of the other is convincing; therefore, we should encourage the industry to provide fully autoclavable 3-D HD neuroendoscopes with diameters not exceeding 4 mm in diameter. However, we have to remember that the human eye has a higher resolution than HD video—one advantage of using the operation microscope. Therefore, the next step—the ultra-HD-resolution—is already being considered and will be available in the next years in the clinical setup.

Michael R. Gaab
Hannover, Germany


learning curve for neuroendoscopic trainees, and ultimately decrease morbidity for neuroendoscopic procedures. The 3-D neuroendoscopes require further improvements to further reduce their size and to allow the addition of working ports for instruments and irrigation. This study provides the justification to further develop these relatively new technologies with the proviso that we continue to assess them both in similar preclinical models but also in clinical situations.

Mark Hamilton
Calgary Alberta, Canada