Improving patient safety during introduction of novel medical devices through cumulative summation analysis

Vejay N. Vakharia, MRCS,1 Roman Rodionov, PhD,1 Andrew W. McEvoy, FRCS(SN),1 Anna Miserocchi, MD,1 Rachel Sparks, PhD,2 Aidan G. O’Keeffe, PhD,3 Sebastien Ourselin, PhD,1,2 and John S. Duncan, FRCP1

1Department of Clinical and Experimental Epilepsy, Institute of Neurology, National Hospital for Neurology and Neurosurgery, 2Transitional Imaging Group, Centre for Medical Image Computing, and 3Department of Statistical Science, University College London, United Kingdom

OBJECTIVE The aim of this study was to implement cumulative summation (CUSUM) analysis as an early-warning detection and quality assurance system for preclinical testing of the iSYS1 novel robotic trajectory guidance system.

METHODS Anatomically accurate 3D-printed skull phantoms were created for 3 patients who underwent implantation of 21 stereoelectroencephalography electrodes by surgeons using the current standard of care (frameless technique). Implantation schema were recreated using the iSYS1 system, and paired accuracy measures were compared with the previous frameless implantations. Entry point, target point, and implantation angle accuracy were measured on post-implantation CT scans. CUSUM analysis was undertaken prospectively.

RESULTS The iSYS1 trajectory guidance system significantly improved electrode entry point accuracies from 1.90 ± 0.96 mm (mean ± SD) to 0.76 ± 0.57 mm (mean ± SD) without increasing implantation risk. CUSUM analysis was successful as a continuous measure of surgical performance and acted as an early-warning detection system. The surgical learning curve, although minimal, showed improvement after insertion of the eighth electrode.

CONCLUSIONS The iSYS1 trajectory guidance system did not show any increased risk during phantom preclinical testing when used by neurosurgeons who had no experience with its use. CUSUM analysis is a simple technique that can be applied to all stages of the IDEAL (idea, development, exploration, assessment) framework as an extra patient safety mechanism. Further clinical trials are required to prove the efficacy of the device.

https://thejns.org/doi/abs/10.3171/2017.8.JNS17936

KEY WORDS stereoelectroencephalography; robotic surgery; CUSUM; cumulative summation; medical device; iSYS1; surgical technique

New surgical procedures and devices have the potential to improve patient outcomes and safety through the translation of technological advancements to health care. The introduction of robotic devices to a number of surgical specialties over the last decade has resulted in their application to numerous procedures. Novel medical devices or procedures are associated with a learning curve, and patients are exposed to potential risks until efficacy can be proven and long-term outcome data are acquired. Surgical device trials have the added complication of user-dependent outcomes and the learning curve/surgeon contribution to outcomes is difficult to predict, especially when trying to assess the external validity of a particular device or procedure.

Regulation with regard to the approval of devices varies among different countries. A recent cross-sectional study showed that fewer than half of devices for which clinical studies are undertaken achieve regulatory (FDA) approval. Of devices that receive regulatory approval, 43% were cleared without the publication of a single clinical study, but rather under the 510(k) clearance, where only substantial equivalence to another approved device is necessary.
Alarminglly, a study of medical devices for orthopedic surgery revealed that those approved using the 510(k) clearance were 11.5 times more likely to be recalled than devices that underwent clinical studies prior to premarket approval.

The IDEAL (idea, development, exploration, assessment, long-term follow-up) framework is a collaborative approach between surgeons and trial methodologists to provide structured guidance (similar to that used for drug trials) for the transition of medical devices from ideas (stage 1) to long-term outcome studies (stage 4). The IDEAL framework, however, does not provide a continuous method for surgical vigilance toward the early detection of harm or the potentially negative effect of learning curves.

Cumulative summation (CUSUM) analysis is a simple early-warning system that compares outcomes of a new intervention or procedure against an established risk or failure rate that can be used longitudinally to monitor outcome and surgical performance. CUSUM analysis has been used in prospective robot-assisted randomized control trials and case control studies in which historical outcome data are used to provide a baseline for comparison. The latter use has been successfully applied to assess learning curves of particular surgical and nonsurgical interventions. CUSUM analysis acts as an early-warning mechanism to inform investigators if an intervention is exposing patients to a higher-than-expected risk of adverse events, but does not replace conventional statistical methods.

Stereoelectroencephalography (SEEG) is a neurosurgical procedure in which multiple electrodes (usually 8–14) are placed within the brain to identify the seizure-onset zone in patients with drug-refractory focal epilepsy, to determine if a resection would be feasible. To date, 4 techniques have been described for the implantation of multiple intracerebral electrodes, including stereotactic frame-based, frameless, robotic, and custom 3D-printed fixture methods. Electrode trajectories are preplanned to ensure that electrodes are a safe distance from intracranial arteries and veins because damage to these vessels could result in a life-threatening hemorrhage that could cause mortality or significant morbidity. The accuracy with which the electrode conforms to the preplanned trajectory is therefore dependent on the implantation method.

We previously performed a meta-analysis of accuracy related to the surgical implantation methodology and found a paucity of evidence within the literature comparing implantation techniques. Studies to date have been of poor quality, amounting to level 3 evidence. There have been no prospective comparisons of different implantation techniques. Herein, we provide an example in which CUSUM analysis was used as an early-warning tool to compare a novel robotic device for the insertion of intracerebral SEEG electrodes with the currently used frameless technique.

**Methods**

**SEEG Technique**

The frameless implantation technique, which is used as the standard of care at the National Hospital for Neurology and Neurosurgery, has been described previously. Briefly, the technique involves the use of a mechanical arm in combination with a precision aiming device and the StealthStation S7 neuronavigation system (Medtronic, Inc.). After registration of the patient to the neuronavigation system using bone fiducial markers as registration points, preplanned trajectories on the StealthStation are used to align the mechanical arm and the precision aiming device. Using a series of reduction tubes, the trajectory is then drilled through the skull and the electrode bolt is screwed into the skull. Next, a stylet is passed through the bolt and the electrode is inserted to the target point.

In a similar fashion, the novel iSYS1 trajectory guidance system (AS Medizintechnik GmbH) is a small device that interfaces with the StealthStation S7 neuronavigation system and, through a series of iterative steps, aligns with the preplanned trajectory. Similar to the precision aiming device, the iSYS1 has a working channel through which reduction tubes are placed to allow drilling and insertion of the skull bolt followed by the electrode. Both procedures were performed by the same 2 neurosurgeons working together.

**Phantom Generation**

Three patients, who underwent a total of 21 electrode implantations by surgeons using the conventional frameless method, were selected on the basis of a power calculation, a representative range of anatomical targets, and drilling angles to the skull. Skull models for each patient were 3D printed (3D Systems, Inc.) with bone fiducials in situ using a commercially available realistic bone-like substitute (DuraForm PA) and covered with a synthetic skin substitute (Fig. 1).

The 2 neurosurgeons who performed implantations in the patients repeated the implantation procedure on these phantoms, using the iSYS1 robotic trajectory guidance system. All equipment, including the drill and electrode bolts, was consistent for both implantation techniques. The 2 neurosurgeons had seen a demonstration of the iSYS1 system and were aware of instructions for its use, but had not received any practical training.
Following implantation, the 3D-printed skulls underwent CT scanning. The planned and actual (implanted) bolt trajectories were compared using a lateral deviation method (Fig. 2) for the entry point, projected target point, and angle error. The results were compared with the actual postoperative patient implantation for the same planned trajectories using EpiNav. 16 The end of the learning process is where the positive gradient of the curve becomes most negative. CUSUM analysis curves (Fig. 4) suggest that the overall learning effect was minimal and that the maximal improvement with iSYS1 compared with the frameless technique occurred after the eighth electrode.

CUSUM Analysis

CUSUM analysis is calculated using the following equation:

$$\delta_n = (\delta_{n-1} + X_{i_n}) - X_{o_n}$$

where \(\delta_n\) is the CUSUM after \(n\) attempts, \(X_{i_n}\) is the result of the intervention following the \(n\)th attempt, and \(X_{o_n}\) is the established risk or failure rate of the control with which ongoing attempts are compared. \(X_{o_n}\) can be calculated either on a case-by-case basis (\(X_{o_n}\)), as with paired control trials in this case, or as an overall frequency if this is known. When \(\delta_n\) is plotted for subsequent attempts, the gradient of the graph provides information regarding whether the intervention is performing better (negative gradient) or worse (positive gradient) than the control intervention. A change in the gradient from negative to positive following the introduction of a new intervention therefore serves as an early warning that outcomes are worse than in the control group, even though this may not have reached statistical significance. Each of the electrode bolt insertion accuracies using the iSYS1 on the phantom (\(X_{i_n}\)) were compared with the patient insertions using the frameless technique (\(X_{o_n}\)). Analysis was also done using a 3-mm-accuracy safety margin based on accuracy data provided by Cardinale et al. 3

Statistical Analysis

The power calculation assumed a significance level of \(\alpha = 0.05\), power of 1 – \(\beta = 0.95\), to detect a 0.8-mm improvement in entry point accuracy with an estimated standard deviation of 0.7 mm based on previously published data. 7 According to this calculation, paired results from 20 electrodes would be required. Following implantation, paired electrode bolt insertion accuracies for the entry point, projected target point, and angle error were tested using both the Kolmogorov-Smirnov and Shapiro-Wilk tests to confirm a Gaussian distribution. Next, Student’s paired t-test (2-tailed) was performed using SPSS version 24.

Results

Phantom Testing Accuracy

Comparison of the frameless insertion of SEEG electrodes in the patients with the iSYS1 system on the 3D phantoms resulted in a statistically significant (\(p < 0.01\)) improvement in the entry point accuracy from 1.90 ± 0.96 mm (mean ± SD) to 0.76 ± 0.57 mm (mean ± SD), respectively. Projected target point accuracy improved from 1.72 ± 0.98 mm (mean ± SD) to 1.34 ± 0.86 mm (mean ± SD), but was not statistically significant (\(p = 0.17\)). Angle error from the plan nonsignificantly improved from 0.95° ± 0.39° (mean ± SD) to 0.88° ± 0.55° (mean ± SD) (\(p = 0.59\)) (Fig. 3).

CUSUM Analysis Results

The CUSUM analyses for entry point, target point, and angle error are shown in Fig. 4. The entry point and target point plots reveal a negative trend line with high correlation (\(R^2 = 0.98\) and 0.69, respectively), indicating that the iSYS1 implantation technique was beneficial and did not increase risk. The CUSUM analysis for angle error, however, showed wide variation with poor correlation (\(R^2 = 0.15\)), suggesting that the implantation method had little or no effect on this measurement.

Learning Curve Assessment

The end of the learning process is where the positive gradient of the curve becomes negative or where the gradient becomes most negative. CUSUM analysis curves (Fig. 4) suggest that the overall learning effect was minimal and that the maximal improvement with iSYS1 compared with the frameless technique occurred after the eighth electrode.

Discussion

The transition of medical devices to the clinical setting through use by early adopters of technology has the potential to cause patient harm before the long-term risks and benefits can be determined through methodologically sound clinical trials. In contrast to drug trials, where phase I and II trials are performed in small numbers of patients to prevent harm, only 60% of devices were found to have published clinical trials prior to attaining regulatory approval. The IDEAL collaboration is an attempt to provide a framework for device trials analogous to that of a drug trial, in which small-scale studies are performed using fewer patients to determine the device’s safety and efficacy pri-
or to larger comparative studies in which long-term data can be gathered. During this period, robust early-warning mechanisms are required that will detect any potentially deleterious effects of the device and thereby prevent patient harm. Herein, we used CUSUM analysis to compare paired electrode insertions in an anatomically accurate phantom using the iSYS1 robotic trajectory guidance system with a frameless implantation performed in patients.

Comparison With Other Studies

CUSUM analysis has been used in a number of surgical and nonsurgical fields to assess the learning curve of operators and as a continuous quality assurance indicator. Through the collection of prospective outcomes, real-time comparisons can be made among prospective control groups, retrospective cohorts, or previously established risk/failure rates. Currently, the main use of CUSUM analysis is to assess surgeon learning curves and training for new techniques. However, CUSUM analysis also has the potential to be applied as a quality assurance indicator, to surgeon revalidation, and as an early-warning detector in clinical trials.

To date, there have been no prospective control trials comparing SEEG insertion techniques. This is probably due to the requirement for a single unit to have a surgeon or group of surgeons who are capable of performing more than one implantation technique. There is some evidence from studies in which one technique has replaced another. Cardinale et al.² compared historical SEEG electrode implantation accuracy using the Talairach frame with that of the neuromate robot (Renishaw). No prospective controlled trial data are available to suggest the superiority of one over the other. Given that the neuromate robot is now the standard of care in Cardinale and associates’ unit, it would be ethically challenging to perform a prospective trial comparing it with the previous technique.

The scenario of a single neurosurgical unit not having the surgical expertise to use more than one technique could be overcome through multicenter trials, but individual surgeon-specific performance is difficult to account for methodologically. Furthermore, the comparison of techniques between different units may introduce a systematic bias. Another important consideration is how the learning effect will be overcome when comparing a new technique that has less familiarity among surgeons with one that has been established. CUSUM analysis may overcome this.

Using paired electrode data, we showed that the performance of a novel device can be continuously monitored, and any change in safety performance over time can be detected. By recreating an anatomically accurate phantom replica of a patient’s skull using 3D printing technology, the same electrode trajectories could be implemented by the same surgeons to control for any systemic bias. During clinical trials, data monitoring committees are established to preside over serious adverse events and have the power to close trials prematurely when one arm of a trial shows significant benefit over another.

For an adverse event to reach statistical significance, a significant number of patients are exposed to risk. Our prospective power calculation revealed that 20 electrodes would be needed to statistically detect a 0.8-mm improvement in entry point accuracy. CUSUM analysis cannot replace statistical tests. However, it does allow trends in beneficial or adverse events to be monitored closely, and it could potentially alert investigators to deleterious outcomes before they become statistically significant. We found that by using the iSYS1 trajectory guidance system, implantation entry point accuracies were significantly improved (p < 0.01) from 1.90 ± 0.96 mm (mean ± SD) with the frameless technique to 0.76 ± 0.57 mm (mean ± SD).

These results are consistent with a study by Dorfer et
al., in which preclinical testing of the iSYS1 device improved entry point accuracy to 0.6 ± 0.4 mm (mean ± SD) from 1.4 ± 0.5 mm (mean ± SD) with the frameless technique. The same group also reported entry point accuracy of 1.18 ± 0.5 mm (mean ± SD) for 93 electrodes in 16 patients, after slight modification of the technique. Cardinale et al. found a mean entry point accuracy of 0.78 mm using the neuromate robot for SEEG, and Mullin et al. found a mean entry point accuracy of 1.2 mm using the ROSA system.

Based on these accuracy data, Cardinale et al. recommended the institution of a safety margin of 3 mm according to the following formula:

\[
\text{Safety margin} = \text{bolt diameter} + \text{mean error} + 3 \text{ SD}
\]

One percent of electrodes would deviate outside of a 3-SD safety margin, which was deemed acceptable. When the CUSUM analysis was repeated on the basis of a 3-mm entry point safety margin and an accepted 1% violation rate, the iSYS1 device performed within this threshold for all implanted electrodes.

**Study Limitations**

Limitations of our study include the small sample (21 electrodes) and differences in the use of a 3D-printed phantom compared with a patient. We tried to use a material that had similar properties to real bone, but the bone substitute was slightly harder.

The control group was derived from in vivo frameless implantations. We accept that it may have been methodologically better if the frameless implantation method was also used on the phantoms. We compared the same surgeons to prevent intersurgeon variability, but they had more experience with the frameless technique than with the iSYS1 robotic device. We found that there was a minimal learning curve associated with first-time use of the iSYS1 device. At all times, entry point accuracy was found to be higher with the iSYS1 device than with the frameless technique.

We do not believe that experience gained through implantation of electrodes in patients using the frameless method would have resulted in an increase in the accuracy of the iSYS1 phantom implantations. This is because the robotic implantations were performed many months after the implantations in patients. Furthermore, given that the iSYS1 trajectory guidance system performs the alignment automatically, we cannot envisage how prior experience with the patients’ skulls could improve the accuracy of the implantation.

**FIG. 4.** CUSUM analyses for entry point error (A), target point error (B), and angle error (C). A negative gradient implies that the intervention (iSYS1) is more beneficial than the control (frameless) implantation. Any change to a positive gradient should alert investigators to an increase in potential risk of the implantation (as can be seen between electrodes 6 and 8 in panel B). The end of the learning curve is taken as the point where the positive gradient becomes negative or where the gradient becomes most negative. In panel A, the learning curve for entry point error becomes most negative after the implantation of electrode 8. In panel B, the target point error becomes negative after the implantation of electrode 9. In panel C, the angle error learning curve becomes negative after the implantation of electrode 12, but there is poor correlation between the intervention and angle error. Figure is available in color online only.
We focused more on entry point accuracy than target point accuracy. The reason is that entry point accuracy and angle are the main factors that can be controlled by the surgeon. Given that the electrodes are flexible and inserted in a blind fashion, there is a potential for them to deviate within the brain. Our phantoms did not have brain material within them and, as such, we could not accurately compare actual target points. To account for this, we calculated projected target points based on the bolt tip and extrapolated these data to a uniform distance for both manual and iSYS1 implantations. The majority of intracranial hemorrhages following SEEG implantation are extraxial, most likely as a result of damage to cortical veins. Although there does remain a risk of hemorrhage along the entire intracranial length of the electrode, small inaccuracies at the target are unlikely to prevent measurement of the interictal and ictal electrophysiology from the target structure of interest.

Conclusions

The introduction of novel medical devices to clinical practice has inherent risk. A large proportion of devices are approved without rigorous clinical trial data or long-term follow-up. In this study, we have shown that CUSUM analysis is an effective tool for the assessment of a novel robotic device for SEEG electrode insertion. As part of preclinical testing, we recreated implantation schemes and implemented these using an anatomically accurate skull phantom. Entry point accuracy was statistically improved using the iSYS1 robotic trajectory guidance system. CUSUM analysis can be used as an early-warning tool in conjunction with all stages of the IDEAL framework to enhance patient safety. A thorough independent appraisal of clinical and economic factors is required before medical devices can be widely adopted. Even with the use of methodologically sound clinical trials, patients are exposed to potential risks. It is an ethical obligation incumbent on all trial investigators to mitigate this risk to the extent possible through the early detection of complications.

Acknowledgments

We thank Katie Konyn for performing the photography during phantom testing. This work was supported by the Wellcome Trust (Innovation Grant no. 106882) and the National Institute for Health Research University College London Hospitals Biomedical Research Centre.

References


**Disclosures**

The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

**Author Contributions**

Conception and design: Vakharia. Acquisition of data: Vakharia, Rodionov, McEvoy, Miserocchi, Sparks. Analysis and interpretation of data: Vakharia. Drafting the article: Vakharia. Critically revising the article: Vakharia, Rodionov, McEvoy, Miserocchi, Sparks, Ourselin, Duncan. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Vakharia. Statistical analysis: Vakharia, O’Keeffe. Study supervision: Ourselin, Duncan.

**Correspondence**

Vejay N. Vakharia: University College London, London, United Kingdom. vvakharia@ucl.ac.uk.