



Title	Deep learning-based video-analysis of instrument motion in microvascular anastomosis training
Author(s)	Sugiyama, Taku; Sugimori, Hiroyuki; Tang, Minghui et al.
Citation	Acta neurochirurgica, 166(1), article number 6 https://doi.org/10.1007/s00701-024-05896-4
Issue Date	2024-01-12
Doc URL	https://hdl.handle.net/2115/94207
Rights	This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature' s AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.1007/s00701-024-05896-4
Type	journal article
File Information	Acta Neurochir 166(1) 6.pdf



Title: Deep learning-based video-analysis of instrument motion in microvascular anastomosis training

Taku Sugiyama, MD, PhD^{1*}, Hiroyuki Sugimori, PhD^{2*}, Minghui Tang, PhD³, Yasuhiro Ito, MD, PhD¹, Masayuki Gekka, MD, PhD¹, Haruto Uchino, MD, PhD¹, Masaki Ito, MD, PhD¹, Katsuhiko Ogasawara, PhD², and Miki Fujimura, MD, PhD¹

* These two authors contributed equally to this work.

Author affiliations:

¹ Department of Neurosurgery, Hokkaido University Graduate School of Medicine, North 15 West 7, Kita-ku, Sapporo 060-8638, Japan

² Faculty of Health Sciences, Hokkaido University, Sapporo 060-0812, Japan

³ Department of Diagnostic Imaging, Hokkaido University Graduate School of Medicine, North 15 West 7, Kita-ku, Sapporo 060-8638, Japan

Correspondence to: Taku Sugiyama, M.D., Ph.D.

Department of Neurosurgery, Hokkaido University Graduate School of Medicine, North 15 West 7, Kita-ku, Sapporo 060-8638, Japan

Tel: +81-11-706-5987, Fax: +81-11-708-7737

E-mail: takus1113@med.hokudai.ac.jp

Running Title: Tool motion analysis using an AI model

Keywords: artificial intelligence, EC-IC bypass, jerk index, kinematic analysis, objective assessment

Abstract

Purpose: Attaining sufficient microsurgical skills is paramount for neurosurgical trainees. Kinematic analysis of surgical instruments using video offers the potential for an objective assessment of microsurgical proficiency, thereby enhancing surgical training and patient safety. The purposes of this study were to develop a deep-learning-based automated instrument tip-detection algorithm, and to validate its performance in microvascular anastomosis training.

Methods: An automated instrument tip-tracking algorithm was developed and trained using YOLOv2, based on clinical microsurgical videos and microvascular anastomosis practice videos. With this model, we measured motion economy (procedural time and path distance) and motion smoothness (normalized jerk index) during the task of suturing artificial blood vessels for end-to-side anastomosis. These parameters were validated using traditional criteria-based rating scales and were compared across surgeons with varying microsurgical experience (novice, intermediate, and expert). The suturing task was deconstructed into four distinct phases and parameters within each phase were compared between novice and expert surgeons.

Results: The high accuracy of the developed model was indicated by a mean Dice similarity coefficient of 0.87. Deep learning-based parameters (procedural time, path distance, and normalized jerk index) exhibited correlations with traditional criteria-based rating scales and surgeons' years of experience. Experts completed the suturing task faster than novices. The total path distance for the right (dominant) side instrument movement was shorter for experts compared to novices. However, for the left (non-dominant) side, differences between the two groups were observed only in specific phases. The normalized jerk index for both the right and left sides was significantly lower in the expert than in the novice groups, and receiver operating characteristic analysis showed strong discriminative ability.

Conclusion: The deep learning-based kinematic analytic approach for surgical instruments proves beneficial in assessing performance in microvascular anastomosis. Moreover, this methodology can

be adapted for use in clinical settings.

Introduction

Microvascular anastomosis skills are important techniques for neurosurgeons, as they are beneficial for repairing unforeseen vascular injuries during various intracranial surgeries. These skills also contribute to extracranial-intracranial (EC-IC) bypass surgery in specific cerebral pathologies such as ischemic disease [10, 31], moyamoya disease [7, 19], complex cerebral aneurysms [13, 27], and skull base tumors [25, 35]. Due to the delicate and fragile nature of the intracranial arteries involved in EC-IC bypass surgery [22, 34], microvascular anastomosis skills require exceptional finesse. This includes precise and coordinated instrument handling with smooth motions to ensure complete approximation of the intimal layer of both donor and recipient arteries while avoiding vascular wall injuries. Furthermore, as the prevalence of modern endovascular procedures and work-hour restriction policies have reduced residents' exposure to microvascular surgery, the role of laboratory training outside the operating theater has become increasingly important.

A variety of off-the-job training methods have been developed. These methods include “dry lab” exercises using gauze, and artificial vessels, as well as “wet lab” exercises utilizing chicken wings, turkey wings, living rats, and placentas [2, 18]. The objective assessment of the microsurgical skill has been introduced not only to provide feedback information to trainees about training progress, goals and challenges, but also to accredit and standardize surgical competence [20]. Criteria-based rating scales, such as the Stanford Microsurgery and Resident Training scale [17, 24], Objective Structured Assessment of Technical Skills [15], and the Northwestern Objective Microanastomosis Assessment Tool [1, 21], serve as good examples of such scales. However, they still face the problem of inter-rater bias and require significant human and time resources [3, 12, 32]. In order to achieve more objective and quantifiable methods, various approaches have been introduced, including hand motion

measurement (e.g., the Imperial College Surgical Assessment Device [5, 14]), instrument motion tracking [8, 16], and gripping force measurement [9, 36]. Nevertheless, these methodologies generally require specialized equipment, including sensors or virtual reality simulators for data acquisition and processing, hindering their clinical application.

With recent advancements in computer vision and machine learning technologies, attention has shifted towards automated video-based motion analysis, which does not require special equipment [6, 11, 17, 30, 33]. In line with this trend, we have previously developed a deep learning-based object detection algorithm which can automatically detect the tips of surgical instruments and track them [26]. To enhance training efficacy, we performed reinforcement learning of this algorithm and applied it to microvascular training videos to evaluate the kinematic characteristics of the surgical instruments. The aim of this study is to validate the feasibility of these methods in evaluating microvascular anastomosis skills based on the obtained temporal coordinate information.

Materials and Methods

Creation of deep learning-based algorithm for surgical instrument detection

To analyze the motion quality and smoothness of surgical instrument during the microvascular anastomosis procedure, a deep learning-based algorithm, which could separately detect right and left tips of the surgical instruments, was developed by transfer learning of a previously described surgical instrument detection convolutional neural network (CNN) algorithm [26].

Training data for the instrument detection algorithm were obtained from microsurgical videos of previous microvascular training sessions. For consideration of the future clinical application, images with a variety of colored surgical backgrounds were used in this study. Each microsurgical video was an AVI file at 29.97 fps, converted to 512×512 JPEG data. Then, a 32×32 pixel square region of interest (ROI) centered on the tips of surgical instruments appearing on the left and right sides was manually set to create supervised image data for the algorithm, using self-made annotation software in

MATLAB (Figure 1).

For accuracy evaluation of the instrument detection algorithm, a total of 60 surgical videos were used and randomly divided into five groups for five-fold cross validation (12 videos in each group). The algorithm was trained using training image data from all 60 videos, and the image data from the 12 videos were used as test data. The accuracy of the algorithm was assessed by the mean Average Precision (mAP) for both left and right instruments as previously described [26]. The mean Dice similarity coefficient (DSC) was also calculated for the evaluation of area coverage of surgical instruments.

In the subsequent experimental study, the algorithm was defined to limit the number of detected bounding boxes to two in the order of the highest DSC for each of the left and right surgical instruments. The coordinates of the center of the detected bounding box were used as the tip of the surgical instrument in all subsequent analyses, and the average coordinates were used for images with two detections on each side.

Study participants for experimental surgical performance analysis

This study was performed with institutional approval from Hokkaido University Hospital (No. 018-0291). Since our facility held periodic off-the-job microvascular training sessions for educational purposes, the surgeon participants (both instructor and trainee) were recruited from the sessions. Fourteen surgeons participated and were classified into three groups based on their experience with microvascular anastomosis and microsurgery at the time of participation. The *novice group* mainly consisted of junior neurosurgical residents (postgraduate year [PGY] 1–3, n=7) who had never performed EC-IC bypass surgery. The *intermediate group* included clinical fellows who had performed fewer than 10 EC-IC bypass surgeries (PGY 6–8, n=2). The *expert group* comprised board-certified neurosurgeons who had performed more than 20 EC-IC bypass surgeries independently (PGY 12–28, n=5). All surgeons were right-handed. Informed consent for the participation of this study was obtained

from all surgeons. Surgeon characteristics are summarized in **Table 1**

Surgical task and trial

In the experiment, the assigned surgical task was a superficial (zero depth) end-to-side anastomosis with the use of 2.0-mm artificial blood vessels and 10-0 surgical sutures. A camera attached to the operating microscope recorded the procedure. Two artificial vessels were prepared as donor and recipient, and the recipient vessel was fixed by microvascular clips or pins. The end of the donor vessel was cut to make a *fish-mouth* opening, and the recipient vessel was cut in a linear fashion. Two stay sutures were placed at both toe and heel ends. Then, interrupted sutures on one then the other side were performed to complete the end-to-side anastomosis. Each stitch was secured with three knots.

In this study, a ‘surgical trial’ was defined as completing one suturing which included getting hold of the needle by forceps, insertion and extraction of the needle, tying knots, and cutting threads. The surgical process of interrupted suture after the two stay sutures was analyzed. For detailed surgical performance analysis, four phases of one trial were also analyzed (**Figure 2**), phase A: grasping and inserting the needle; phase B: pushing and extracting the needle; phase C: pulling threads to first knot; phase D: tying knots and cutting threads.[9]

Deep learning-based analysis of microsurgical performance

A total of 66 trials, comprising 30, 6, and 30 trials from 5 expert, 2 intermediate, and 7 novice surgeons respectively, were evaluated using the developed surgical instrument detection algorithm. The algorithm was applied to the JPEG-converted video data of each surgeon to obtain the temporal coordinate information of the left and right instrument tip data. Calibration of the size per pixel was performed using a scale located within the microscopic field. The total number of frames for each trial was calculated, and the start and end frames were obtained for each process from phase A to phase D.

To assess the motion economy, procedural time and total path distance were used. To assess

the motion quality and smoothness (i.e. dexterity), the jerk index was employed. In this study, to mitigate the effects of task duration and movement amplitude, we calculated the normalized jerk index using the following equation based on the change in the left and right coordinates of the surgical instrument over time, the distance traveled, and the time required for the process (**Figure 3**) [8].

$$\text{Normalized jerk index} = \frac{1}{2} * \frac{t^5}{D^2} * \int \left(\frac{da}{dt}\right)^2$$

a: acceleration, t: movement time, D: movement distance

Considering that the movement of instruments within the microscopic view may also be an important factor in technical assessment, we calculated the duration during which the surgical instrument tips were outside the microscopic view. The time when the surgical instruments were out of view was represented as a percentage of the total procedural time.

Criteria-based analysis of microsurgical performance

To validate the efficacy of deep learning-based video analysis in assessing surgical performance, all trials were additionally evaluated using a previously validated criteria-based rating scale (the Stanford Microsurgery and Resident Training scale) [17, 24]. Three expert surgeons independently rated all trials, with their knowledge of the surgeon's identity blinded. As previously described [17], they rated each trial on a scale from 9 to 45 across nine categories (1–5 points each), encompassing instrument handling, respect for tissue, efficiency, suture handling, suturing technique, quality of the knot, final product, operational flow, and overall performance.

Statistical analysis

The results are presented as mean $\pm$ standard deviation. Interrater reliability for the criteria-based rating scale was assessed using Cronbach's α coefficient among three raters. Correlations between deep learning-based parameters (procedural time, path distance, normalized jerk index, and the percentage of time surgical instruments were out of the microscopic view) and the criteria-based rating scale were

analyzed using Pearson's correlation coefficient (r). Correlations between microsurgical performance parameters and years of training (PGY) were analyzed using Spearman's rank correlation coefficient (ρ).

For a detailed analysis of surgical performance, deep learning-based parameters in the four subcategorized phases (phases A-D) were compared between expert and novice groups. The Shapiro–Wilk test was used to assess normality of parameters. The t-test was applied when normality was confirmed, and the Kruskal–Wallis test was employed in its absence. Receiver operating characteristic (ROC) curves were also calculated for parameters relevant to expert and novice groups. The area under the curve (AUC) and a 95% confidence interval (CI) were computed to evaluate the discriminative accuracy of kinematic parameters in predicting the surgeon's skill level. Statistical significance was set at a p-value $<.05$.

Results

Accuracy of the created instrument detection algorithm

Table 2 shows the mAP of the detector at the tip of the surgical instrument. The average mAPs were 0.65 and 0.63 for right and left instruments, respectively. The high accuracy of the developed model was reflected by the average mean DSC for the detection of all surgical instruments of 0.87.

Correlation between deep learning-based and criteria-based performance analysis

The microsurgical performance profiles of each surgeon, including their mean scores on the criteria-based rating scale, total procedural time, total path distances of both the right and left instruments, normalized jerk indexes for both instruments, and the percentages of time when both instruments were out of the microscopic view, are detailed in **Supplementary material 1**.

The high interrater reliability for the criteria-based rating scale was confirmed with a Cronbach's α coefficient of 0.97. **Table 3** presents Pearson correlation coefficients (r), 95% CI, and p-

values for the association between deep learning-based parameters and the criteria-based rating scale. Procedural time, total path distances of both right and left instruments, and normalized jerk indexes for both instruments showed a significant correlation with the criteria-based rating scale, as depicted in **Figure 4**. However, the percentage of time the surgical instruments were out of the microscopic view did not demonstrate a significant correlation.

Correlation between microsurgical performance parameters and years of training

Table 4 displays Spearman's rank correlation coefficients (ρ) and p-values for the association between microsurgical performance parameters and surgeons' PGY. The criteria-based rating scales, procedural time, total path distances of both right and left instruments, and normalized jerk indexes for both instruments demonstrated a significant correlation with the surgeons' PGY, as shown in **Figure 5**. However, the percentage of time the surgical instruments were out of the microscopic view did not show a significant correlation.

Detailed technical analysis of surgical phases between experts and novices

The procedural time in each procedural phase is shown in **Figure 6A**, and detailed data is also provided in **Supplementary material 2**. The total path distance of the instruments is shown in **Figure 6B and C**, and detailed data are also provided in **Supplementary material 3**. Experts completed a trial in shorter procedural time than novices. With regard to right (dominant) side instrument use, total path distances in the total procedure and in phases B-D were shorter in the expert than the novice group. However, for the left (non-dominant) side instrument, the path distance only in phase C was shorter in the expert than the novice group.

The normalized jerk index is shown in **Figure 6D and E**, and detailed data are also provided in **Supplementary material 4**. In terms of both right and left hands, the normalized jerk indices for the total procedure and in phases B–D were significantly lower in the expert than in the novice group.

The ROC curves for each parameter in each phase in the discrimination between expert and novice groups are shown in **Figure 7**, and the AUC and 95% CI are shown in **Table 5**. The strong associations were reflected by the $AUC \geq 0.80$ for multiple parameters, including procedural time in phases B–D, the path distance of dominant side in phase C and D, and the normalized jerk index of both right and left sides in phases B–D.

Discussion

In this study, we developed and utilized a deep learning-based algorithm that can automatically detect and track both the right and left instrument tips separately, to analyze kinematic data of instrument motion in microvascular training. We quantified motion economy and smoothness, and compared these parameters across surgeons with different skill levels. The results showed that deep learning parameters (procedural time, path distance, and normalized jerk index) correlated well with traditional criteria-based rating scales. Additionally, these parameters exhibited correlations with the surgeons' years of experience. Therefore, this approach offers potential for the objective assessment of surgical skill. Our findings align with previous studies that used criteria-based rating scales, motion analysis software, and motion sensors for microsurgical skill assessment [9, 17]. To the best of our knowledge, this is the first study to utilize a deep learning-based video-analysis algorithm for assessing microsurgical performance in microvascular anastomosis training.

Traditionally, microsurgical performance has been represented by clinical outcomes such as bypass patency rate, clamping time, and postoperative stroke rates [23]. However, these metrics are not holistic indicators of overall surgical skill. As the traditional master-apprentice surgical education model transitions towards a competence-based educational model, various objective methods have been introduced to assess surgical skill [29, 32]. Among them, the video-based kinematic analysis approach for surgical instruments is particularly appealing, as it does not require special sensors or equipment, therefore providing high clinical adaptability. The artificial intelligence-based model is

also compelling due to its potential to analyze surgical videos in real-time with fewer human resources.

Motion quality and technical considerations

Proficiency in microsurgical skills is crucial for patient safety. While surgical skills encompass various aspects, such as procedural and anatomical knowledge, instrumentation, flow of operation, and autonomy, the dexterity in manipulating surgical instruments is particularly vital in microvascular surgery [28]. In microvascular anastomosis, where the surgical object (e.g., the recipient of the intracranial artery) has a diameter of approximately 1 mm—and sometimes is as small as 0.5 mm in pediatric moyamoya disease cases [22, 34]—precise and coordinated needle manipulation is essential for accurate needle insertion and extraction. Hesitant movements, tremors, or jerky motions often complicate this procedure.

In this study, we utilized the jerk index to quantify dexterity and motion smoothness. While motion smoothness is typically measured based on factors like the rate of change in acceleration, the curvature of the path traveled, or the cumulative count of abrupt accelerations and decelerations, these metrics can be influenced by the duration of the task and the amplitude of the movement. To negate the effects of time and amplitude, we used the normalized jerk index, as previous studies have shown that the normalized jerk index is effective in quantifying motion smoothness and coordination [4, 8]. This research marks the first time the normalized jerk index has been applied to evaluate surgical performance in microvascular anastomosis.

In this study, multiple parameters were different between experts and novices in a majority of the surgical phases, suggesting that experts employ higher-quality motion when manipulating needles, threads, and the vascular wall, while also completing surgical tasks more swiftly. However, while the total path distance of the right (dominant) side instrument was shorter for experts than for novices, that of the left (non-dominant) side instrument was shorter only in a limited phase (phase C). This discrepancy might be influenced by the fact that expert surgeons more frequently use their non-

dominant hand efficiently to avoid tearing the delicate vascular wall during anastomosis. Notably, the normalized jerk index distinctly differentiated between the two groups with high precision, in terms of instrument motion for both dominant and non-dominant sides. This suggests that the normalized jerk index might be more useful than path distance for assessing microsurgical skill.

Limitations and future directions

A primary technical limitation of our methods was the absence of three-dimensional kinematic data, specifically the lack of depth information. However, given that most microsurgical procedures are conducted within the same plane under the microscope's focus, this limitation's impact was considered minimal. Another constraint was that when the surgical tool was outside the microscope's visual field, kinematic data could not be captured.

Our study had a relatively small sample size concerning the number of participating surgeons. Moreover, we did not evaluate data from repeated training sessions of trainees to estimate the learning curve and determine if feedback could enhance training efficacy. Further studies involving a larger participant pool with diverse microsurgical experiences are essential to better differentiate surgeon skill levels and assess whether such training approaches can improve training efficiency.

Owing to the versatility of this analytic method, analyzing videos in clinical settings and real-time monitoring present intriguing research avenues for the future, especially since laboratory training does not replicate the stressors surgical trainees face in operating theatre environments.

Conclusions

We developed a deep learning-based algorithm capable of measuring kinematic parameters of surgical instruments, encompassing motion economy and smoothness. The measured parameters using this deep learning approach correlated well with the skill levels of surgeons. Such methodologies promise to enrich microvascular anastomosis training, with anticipated clinical

applications.

Acknowledgements: We would like to thank Editage (www.editage.com) for English language editing. We thank Masaki Mizuhara (Integra Japan Co., Ltd.), Kazuhiro Goto (Zeiss Japan Co., Ltd.), and Shusaku Tsutsui (Johnson & Johnson Inc.) for their support in organizing microvascular hands-on training courses.

Author contribution: Taku Sugiyama, Hiroyuki Sugimori, and Minghui Tang contributed to the study conception and design. Data collection was performed by Taku Sugiyama, Yasuhiro Ito, Masayuki Gekka, Haruto Uchino, and Masaki Ito. Data analysis was performed by Hiroyuki Sugimori and Minghui Tang. Katsuhiko Ogasawara and Miki Fujimura contributed to supervision of the study. The first draft of the manuscript was written by Hiroyuki Sugimori and Taku Sugiyama. Taku Sugiyama critically revised the manuscript, and all authors approved the final version of the manuscript.

Funding: This work was supported by the Japanese Society for the Promotion of Science KAKENHI Grant Number 21K09091

Data availability: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Statements and Declarations

Conflict of Interest: The authors declare that they have no conflict of interest.

Ethical approval: All procedures performed in studies involving human participants were in

accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This study was approved by the institutional review board of Hokkaido University Hospital, Sapporo, Japan (No. 018-0291).

Informed consent: Informed consent was obtained from all individual participants included in this study.

References

1. Aoun SG, El Ahmadieh TY, El Tecle NE, Daou MR, Adel JG, Park CS, Batjer HH, Bendok BR (2015) A pilot study to assess the construct and face validity of the Northwestern Objective Microanastomosis Assessment Tool. *J Neurosurg* 123(1):103–109
2. Belykh E, Abramov I, Bardanova L, et al (2022) Seven bypasses simulation set: description and validity assessment of novel models for microneurosurgical training. *J Neurosurg* 138(3):732–739
3. Castillo-Segura P, Fernández-Panadero C, Alario-Hoyos C, Muñoz-Merino PJ, Delgado Kloos C (2021) Objective and automated assessment of surgical technical skills with IoT systems: A systematic literature review. *Artif Intell Med* 112:102007
4. Chmarra MK, Kolkman W, Jansen FW, Grimbergen CA, Dankelman J (2007) The influence of experience and camera holding on laparoscopic instrument movements measured with the TrEndo tracking system. *Surg. Endosc. Other Interv. Tech. Surg Endosc*, pp 2069–2075
5. Datta V, Mackay S, Mandalia M, Darzi A (2001) The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model. *J Am Coll Surg* 193(5):479–485
6. Davids J, Makariou SG, Ashrafian H, Darzi A, Marcus HJ, Giannarou S (2021) Automated

Vision-Based Microsurgical Skill Analysis in Neurosurgery Using Deep Learning: Development and Preclinical Validation. *World Neurosurg* 149:e669–e686

7. Fujimura M, Tominaga T, Kuroda S, Takahashi JC, Endo H, Ogasawara K, Miyamoto S (2022) 2021 Japanese Guidelines for the Management of Moyamoya Disease: Guidelines from the Research Committee on Moyamoya Disease and Japan Stroke Society. *Neurol Med Chir (Tokyo)* 62(4):165–170
8. Ghasemloonia A, Maddahi Y, Zareinia K, Lama S, Dort JC, Sutherland GR (2017) Surgical Skill Assessment Using Motion Quality and Smoothness. *J Surg Educ* 74(2):295–305
9. Harada K, Morita A, Minakawa Y, Baek YM, Sora S, Sugita N, Kimura T, Tanikawa R, Ishikawa T, Mitsuishi M (2015) Assessing Microneurosurgical Skill with Medico-Engineering Technology. *World Neurosurg* 84(4):964–971
10. Housley SB, Vakharia K, Gong AD, Waqas M, Rho K, Levy EI, Davies JM, Siddiqui AH (2022) Extracranial-to-Intracranial Bypass for Distal Internal Carotid Artery and/or Proximal Middle Cerebral Artery Steno-Occlusive Disease: A Case Series of Clinical Outcomes at a Single, High-Volume Cerebrovascular Center. *Oper Neurosurg* 23(3):177–181
11. Khalid S, Goldenberg M, Grantcharov T, Taati B, Rudzicz F (2020) Evaluation of Deep Learning Models for Identifying Surgical Actions and Measuring Performance. *JAMA Netw open* 3(3):e201664
12. Lam K, Chen J, Wang Z, Iqbal FM, Darzi A, Lo B, Purkayastha S, Kinross JM (2022) Machine learning for technical skill assessment in surgery: a systematic review. *npj Digit Med* 5(1):24
13. Lawton MT, Lang MJ (2019) The future of open vascular neurosurgery: Perspectives on cavernous malformations, AVMs, and bypasses for complex aneurysms. *J Neurosurg* 130(5):1409–1425
14. Mackay S, Datta V, Mandalia M, Bassett P, Darzi A (2002) Electromagnetic motion analysis

in the assessment of surgical skill: Relationship between time and movement. *ANZ J Surg* 72(9):632–634

15. Martin JA, Regehr G, Reznick R, Macrae H, Murnaghan J, Hutchison C, Brown M (1997) Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 84(2):273–278
16. McBeth PB, Louw DF, Yang F, Sutherland GR (2005) Quantitative measures of performance in microvascular anastomoses. *Comput Aided Surg* 10(3):173–180
17. McGoldrick RB, Davis CR, Paro J, Hui K, Nguyen D, Lee GK (2015) Motion analysis for microsurgical training: Objective measures of dexterity, economy of movement, and ability. *Plast Reconstr Surg* 136(2):231e–240e
18. Mikami T, Suzuki H, Ukai R, Komatsu K, Kimura Y, Akiyama Y, Wanibuchi M, Mikuni N (2018) Surgical Anatomy of Rats for the Training of Microvascular Anastomosis. *World Neurosurg* 120:e1310–e1318
19. Miyamoto S, Yoshimoto T, Hashimoto N, et al (2014) Effects of extracranial-intracranial bypass for patients with hemorrhagic moyamoya disease: Results of the Japan adult moyamoya trial. *Stroke* 45(5):1415–1421
20. Niveditha M, Sharma R, Suri A (2022) Microsurgical suturing assessment scores: a systematic review. *Neurosurg Rev* 45(1):119–124
21. Pines AR, Alghoul MS, Hamade YJ, et al (2017) Assessment of the interrater reliability of the congress of neurological surgeons microanastomosis assessment scale. *Oper Neurosurg* 13(1):108–112
22. Rashad S, Fujimura M, Niizuma K, Endo H, Tominaga T (2016) Long-term follow-up of pediatric moyamoya disease treated by combined direct–indirect revascularization surgery: single institute experience with surgical and perioperative management. *Neurosurg Rev* 39(4):615–623

23. Reynolds MR, Grubb RL, Clarke WR, Powers WJ, Zipfel GJ, Adams HP, Derdeyn CP (2013) Investigating the mechanisms of perioperative ischemic stroke in the Carotid Occlusion Surgery Study. *J Neurosurg* 119(4):988–995
24. Satterwhite T, Son J, Carey J, Echo A, Spurling T, Paro J, Gurtner G, Chang J, Lee GK (2014) The stanford microsurgery and resident training (SMaRT) scale: Validation of an on-line global rating scale for technical assessment. *Ann Plast Surg* 72(SUPPL. 1):S84–S88
25. Sekhar LN, Kalavakonda C (2002) Cerebral revascularization for aneurysms and tumors. *Neurosurgery* 50(2):321–331
26. Sugimori H, Sugiyama T, Nakayama N, Yamashita A, Ogasawara K (2020) Development of a Deep Learning-Based Algorithm to Detect the Distal End of a Surgical Instrument. *Appl Sci* 10(12):4245
27. Sugiyama T (2017) *Mastering Intracranial Microvascular Anastomoses -Basic Techniques and Surgical Pearls*. MEDICUS SHUPPAN, Publishers Co., Ltd
28. Sugiyama T, Clapp T, Nelson J, et al (2021) Immersive 3-Dimensional Virtual Reality Modeling for Case-Specific Presurgical Discussions in Cerebrovascular Neurosurgery. *Oper Neurosurg* 20(3):289–299
29. Sugiyama T, Gan LS, Zareinia K, Lama S, Sutherland GR (2017) Tool-Tissue Interaction Forces in Brain Arteriovenous Malformation Surgery. *World Neurosurg* 102:221–228
30. Sugiyama T, Ito M, Sugimori H, et al (2023) Tissue Acceleration as a Novel Metric for Surgical Performance During Carotid Endarterectomy. *Oper Neurosurg* 25(4):343–352
31. Sugiyama T, Kazumata K, Asaoka K, Osanai T, Shimbo D, Uchida K, Yokoyama Y, Nakayama N, Itamoto K, Houkin K (2015) Reappraisal of microsurgical revascularization for anterior circulation ischemia in patients with progressive stroke. *World Neurosurg* 84(6):1579–1588
32. Sugiyama T, Lama S, Gan LS, Maddahi Y, Zareinia K, Sutherland GR (2018) Forces of tool-

tissue interaction to assess surgical skill level. *JAMA Surg* 153(3):234–242

33. Sugiyama T, Nakamura T, Ito Y, Tokairin K, Kazumata K, Nakayama N, Houkin K (2019) A Pilot Study on Measuring Tissue Motion During Carotid Surgery Using Video-Based Analyses for the Objective Assessment of Surgical Performance. *World J Surg* 43(9):2309–2319
34. Takagi Y, Kikuta KI, Nozaki K, Hashimoto N (2007) Histological features of middle cerebral arteries from patients treated for moyamoya disease. *Neurol Med Chir (Tokyo)* 47(1):1–4
35. Wolfswinkel EM, Landau MJ, Ravina K, Kokot NC, Russin JJ, Carey JN (2018) EC-IC bypass for cerebral revascularization following skull base tumor resection: Current practices and innovations. *J Surg Oncol* 118(5):815–825
36. Zareinia K, Maddahi Y, Gan LS, Ghasemloonia A, Lama S, Sugiyama T, Yang FW, Sutherland GR (2016) A Force-Sensing Bipolar Forceps to Quantify Tool-Tissue Interaction Forces in Microsurgery. *IEEE/ASME Trans Mechatronics* 21(5):2365–2377

Figure legends

Figure 1

Training data for instrument detection YOLOv2 algorithm and reprocessing flowchart

Figure 2

Deconstructed smaller surgical phases .

Figure 3

Surgical instrument tip trajectory analysis software.

Figure 4

Correlation between deep learning-based parameters and the criteria-based rating scale.

A: Procedural time

B: Path distance for right (dominant) side instrument tip

C: Path distance for left (non-dominant) side instrument tip

D: Normalized jerk index for right (dominant) side instrument tip

E: Normalized jerk index for left (non-dominant) side instrument tip

Figure 5

Correlation between microsurgical performance parameters and surgeons' years of training.

Smoothing splines are provided for reference of learning curves.

PGY, postgraduate year

A: Rating scale

B: Procedural time

C: Path distance for right (dominant) side instrument tip

D: Path distance for left (non-dominant) side instrument tip

E: Normalized jerk index for right (dominant) side instrument tip

F: Normalized jerk index for left (non-dominant) side instrument tip

Figure 6

Box-and-whisker diagrams for each parameter in each phase.

A: Procedural time

B: Path distance for right (dominant) side instrument tip

C: Path distance for left (non-dominant) side instrument tip

D: Normalized jerk index for right (dominant) side instrument tip

E: Normalized jerk index for left (non-dominant) side instrument tip

Figure 7

Receiver operating curves for each parameter in each phase

A: Phase A, **B:** Phase B, **C:** Phase C, **D:** Phase D.

Surgeon No.	Surgeon Group	PGY	Experience (N)		
			Laboratory training	Microsurgery	EC-IC bypass
1	Novice	1	<10	<5	0
2		1	<2	<5	0
3		2	<5	<10	0
4		2	<5	<20	0
5		2	<5	<30	0
6		3	<15	<40	0
7		3	<20	<30	0
8	Intermediate	6	>20	<100	<5
9		8	>30	<200	<10
10	Expert	12	>30	<700	<100
11		13	>30	<1000	<100
12		14	>30	<300	<30
13		20	>50	>1500	>150
14		28	>50	>2000	>300

Table 1 Surgeons' characteristics

PGY, postgraduate year; N, number; EC-IC, extracranial-intracranial

	mAP			Mean DSC
	Total	R	L	
Fold 1	0.7	0.72	0.7	0.89
Fold 2	0.65	0.67	0.65	0.88
Fold 3	0.65	0.67	0.65	0.88
Fold 4	0.59	0.63	0.59	0.85
Fold 5	0.53	0.56	0.53	0.83
Mean	0.63	0.65	0.63	0.87

Table 2 Mean average precision and dice similarity coefficient in each fold.
DSC, Dice similarity coefficient; L, left; mAP, mean average precision; R, right

Parameter	Side	r	95% CI	p-value
Time		-0.77	-0.92 – -0.40	<0.01**
Total path distance	R	-0.85	-0.95 – -0.58	<0.01**
	L	-0.58	-0.85 – -0.07	0.03*
Normalized jerk index	R	-0.60	-0.86 – -0.10	0.02*
	L	-0.57	-0.84 – -0.06	0.03*
% out of operative field	R	-0.19	-0.65 – -0.38	0.52
	L	-0.11	-0.61 – -0.45	0.71

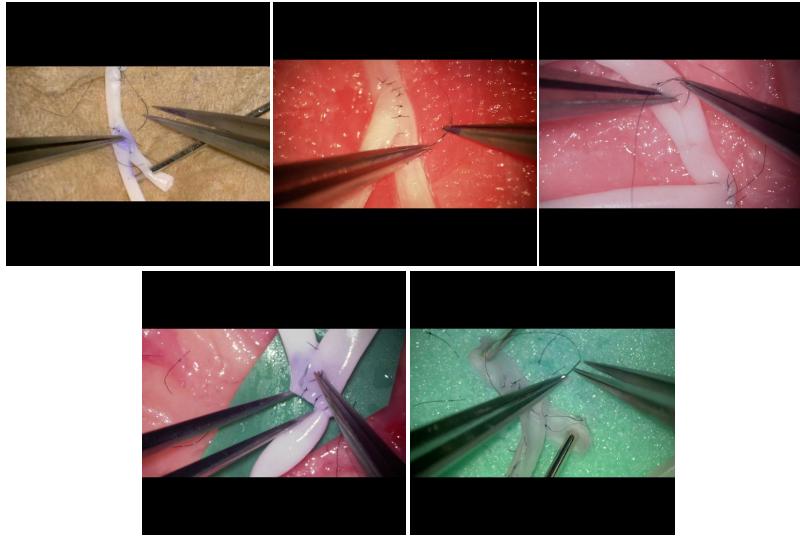
Table 3 The association between deep learning-based parameters and the criteria-based rating scale
Pearson correlation coefficients (r), 95% confidence interval (CI), and p-values are provided.

Parameter	Side	ρ	p-value
Rating scale	-	0.82	<0.001***
Time	-	-0.77	<0.01**
Total path distance	R	-0.80	<0.001***
	L	-0.57	0.03*
Normalized jerk index	R	-0.67	<0.01**
	L	-0.76	<0.01**
% out of operative field	R	-0.06	0.84
	L	-0.20	0.49

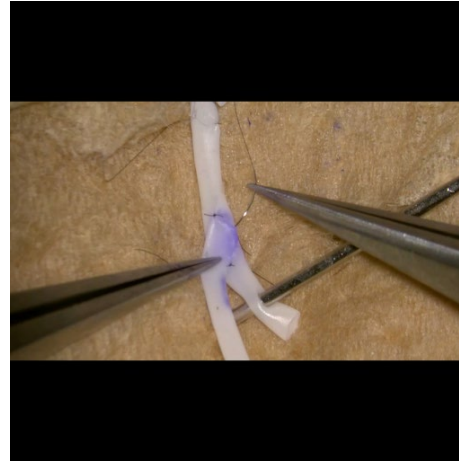
Table 4 The association between microsurgical performance parameters and years of training (postgraduate year)
Spearman's rank correlation coefficients (ρ) and p-values are provided.

Parameter	Side	Phase A	Phase B	Phase C	Phase D
Time	-	0.59 (0.41-0.77)	0.93 (0.84-1.00)	0.95 (0.86-1.00)	0.91 (0.79-1.00)
Path distance	R	0.36 (0.17-0.54)	0.76 (0.61-0.91)	0.84 (0.72-0.96)	0.91 (0.82-1.00)
	L	0.55 (0.37-0.73)	0.59 (0.41-0.77)	0.70 (0.53-0.87)	0.54 (0.35-0.72)
Normalized jerk index	R	0.56 (0.37-0.74)	0.94 (0.87-1.00)	0.90 (0.8-0.99)	0.86 (0.74-0.98)
	L	0.48 (0.28-0.67)	0.87 (0.75-0.98)	0.92 (0.84-1.00)	0.85 (0.72-0.98)

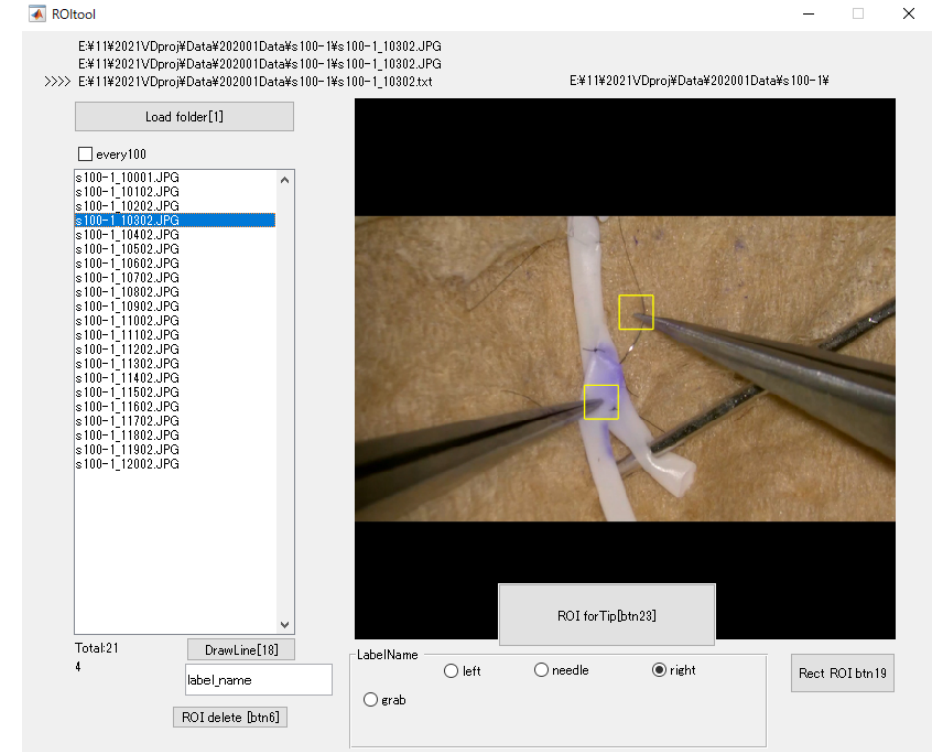
Table 5 Area under the curve and 95% CI based on receiver operating characteristic analysis for each parameter in each phase



AVI file

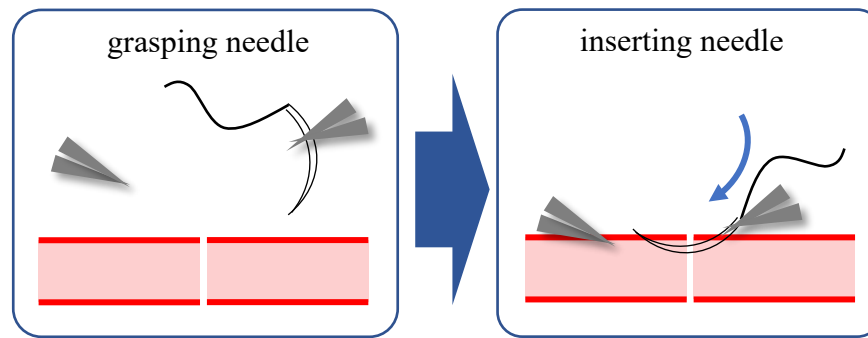


512 × 512 JPEG files
with zero padding

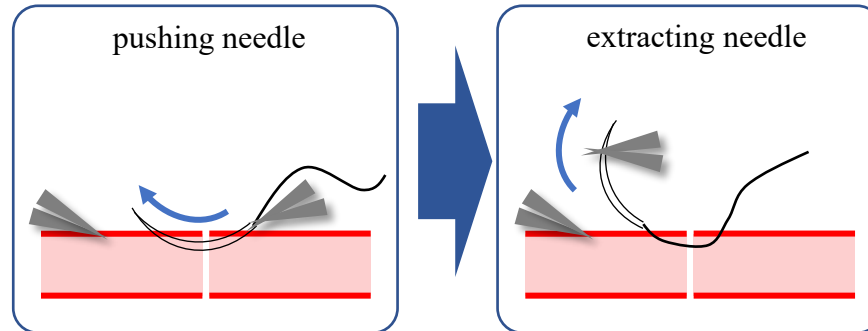


Annotation
(supervised data creation)

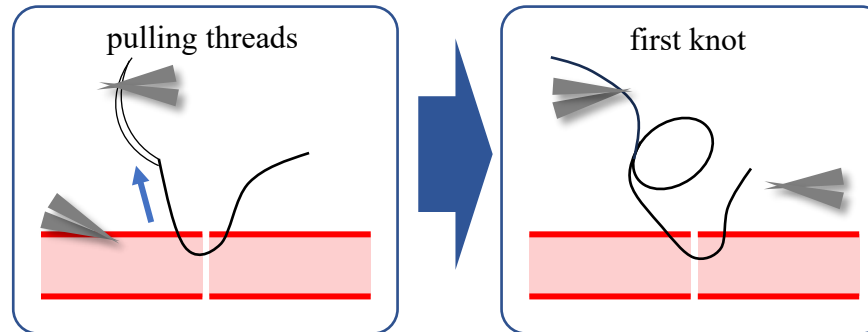
Phase A



Phase B



Phase C



Phase D

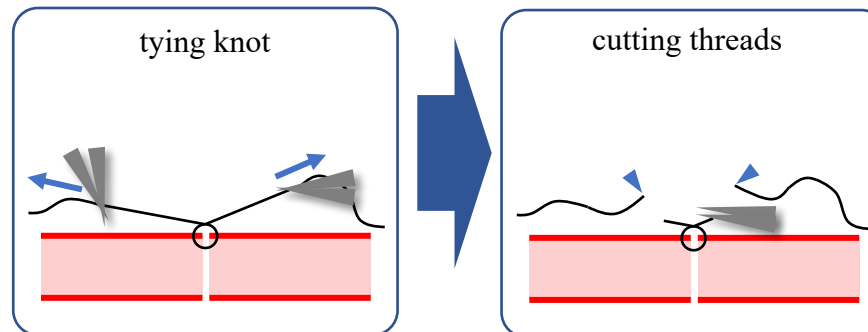


Fig. 2

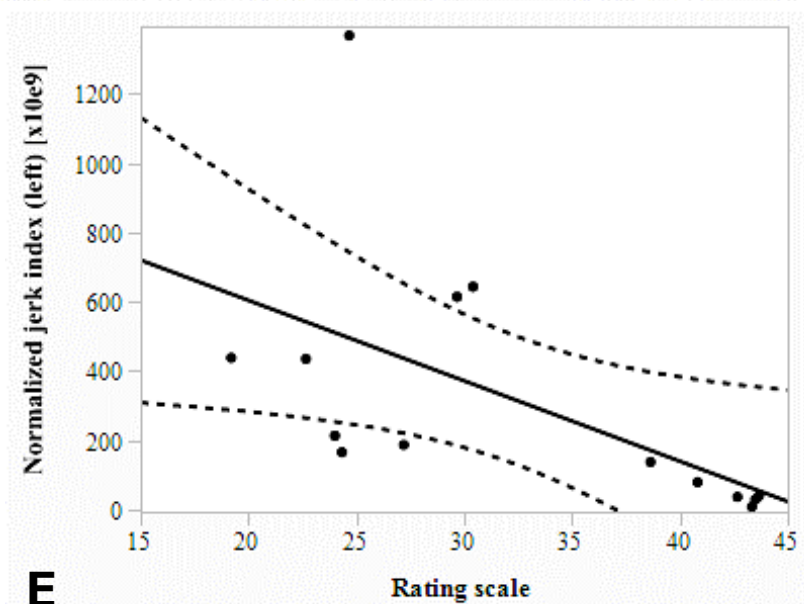
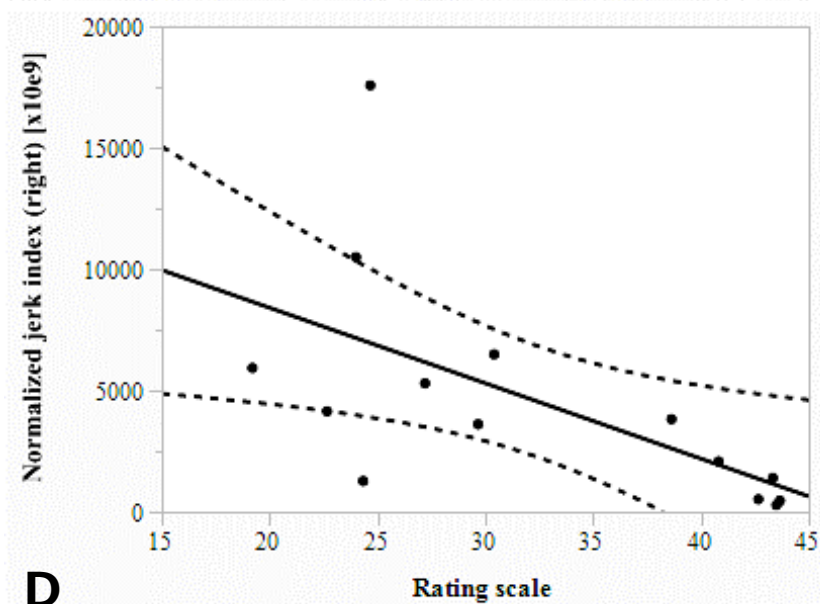
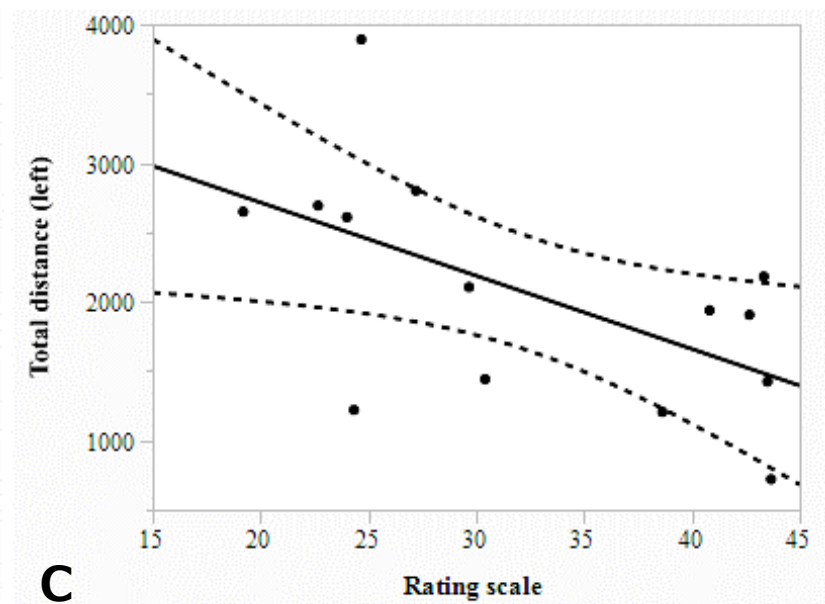
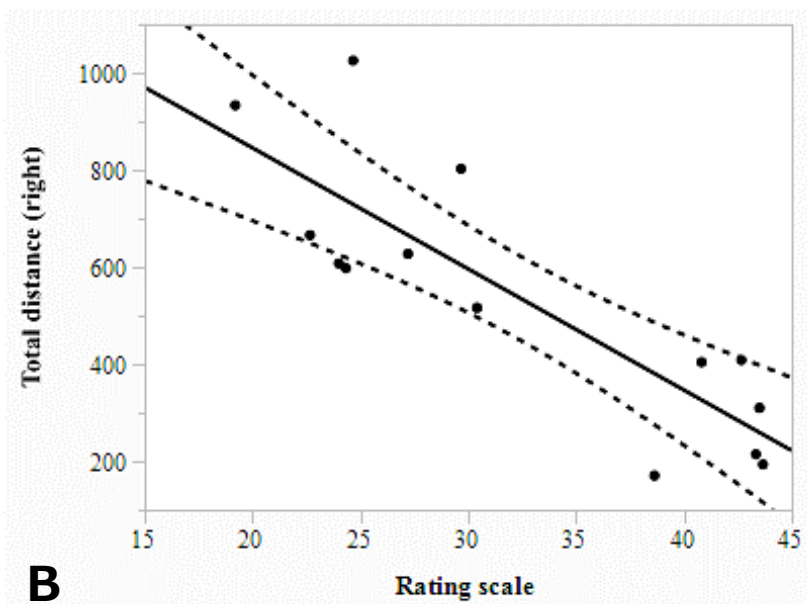
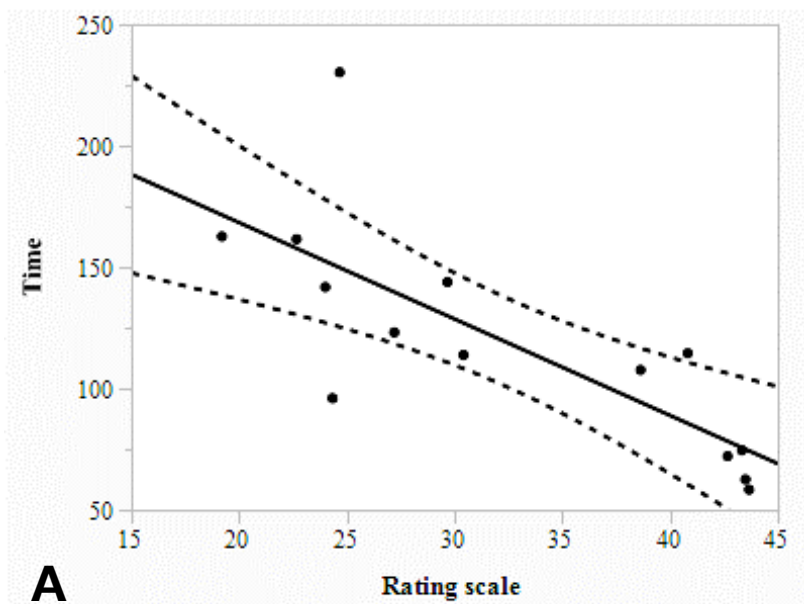


Fig. 4

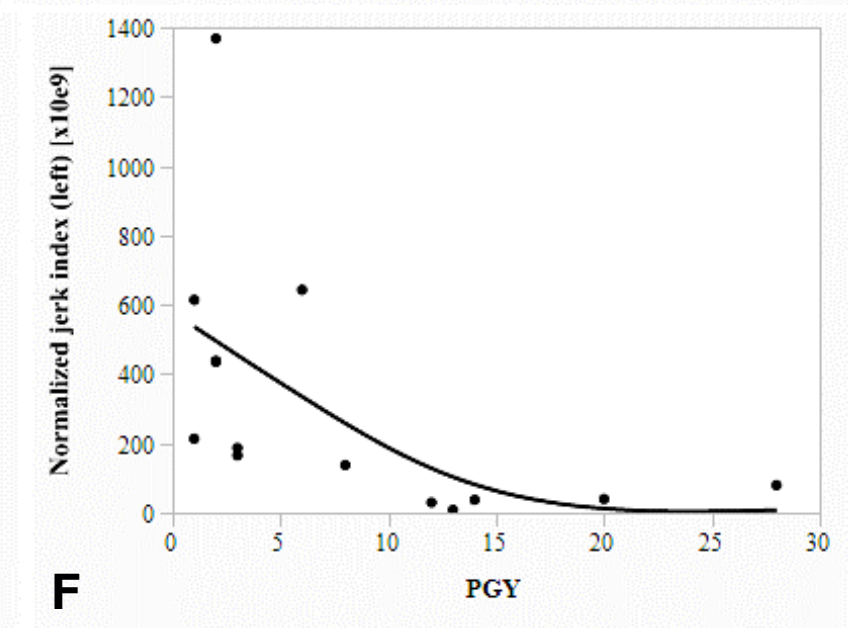
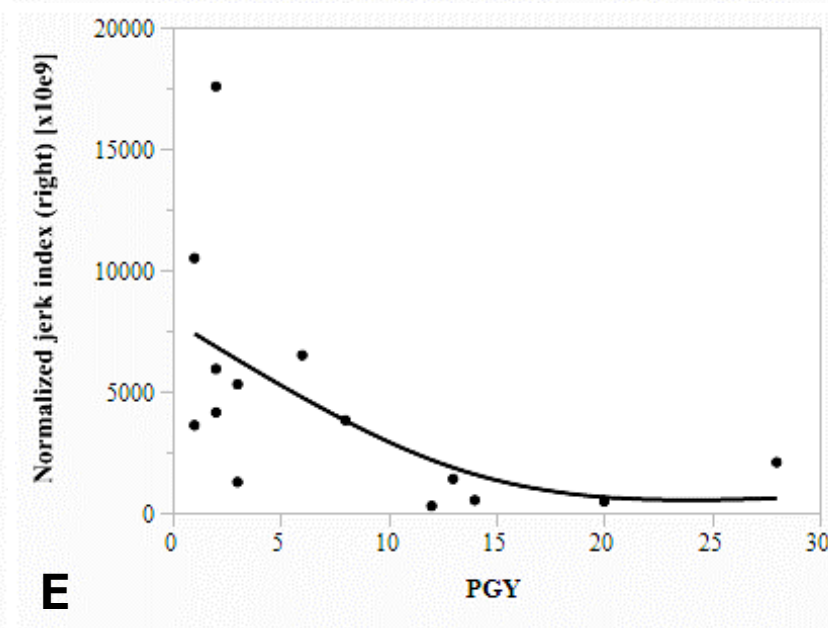
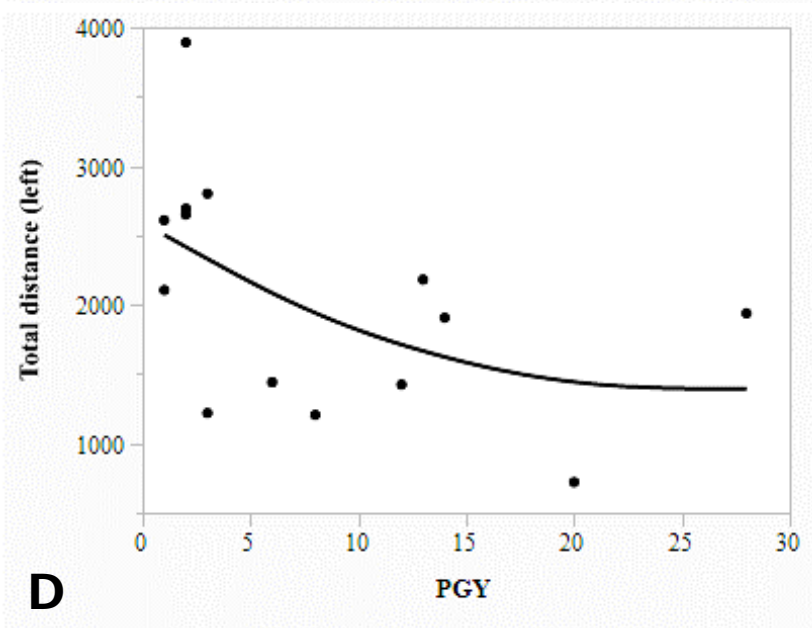
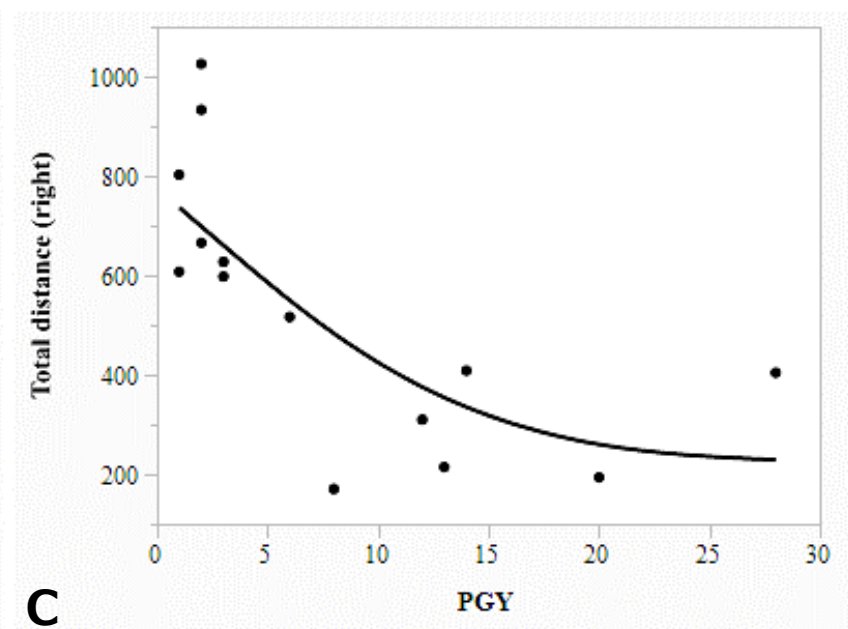
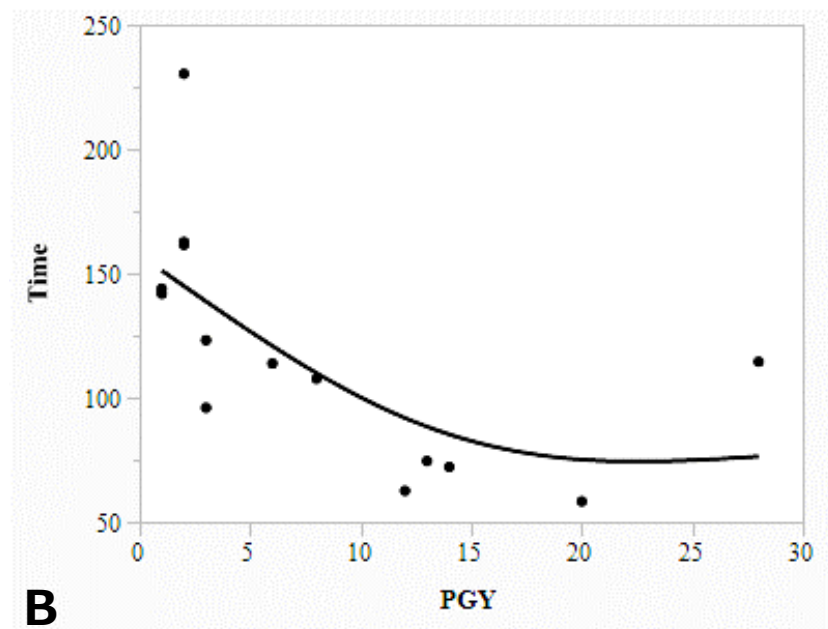
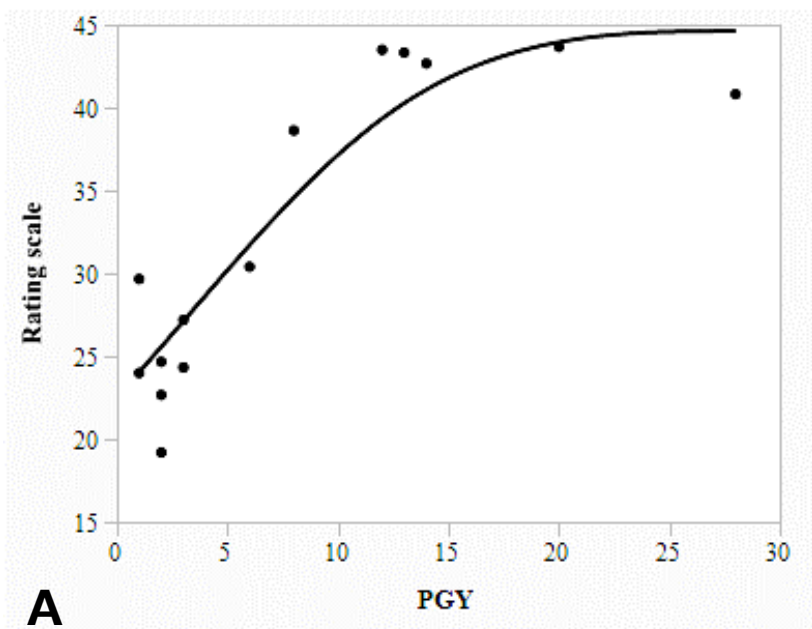


Fig. 5

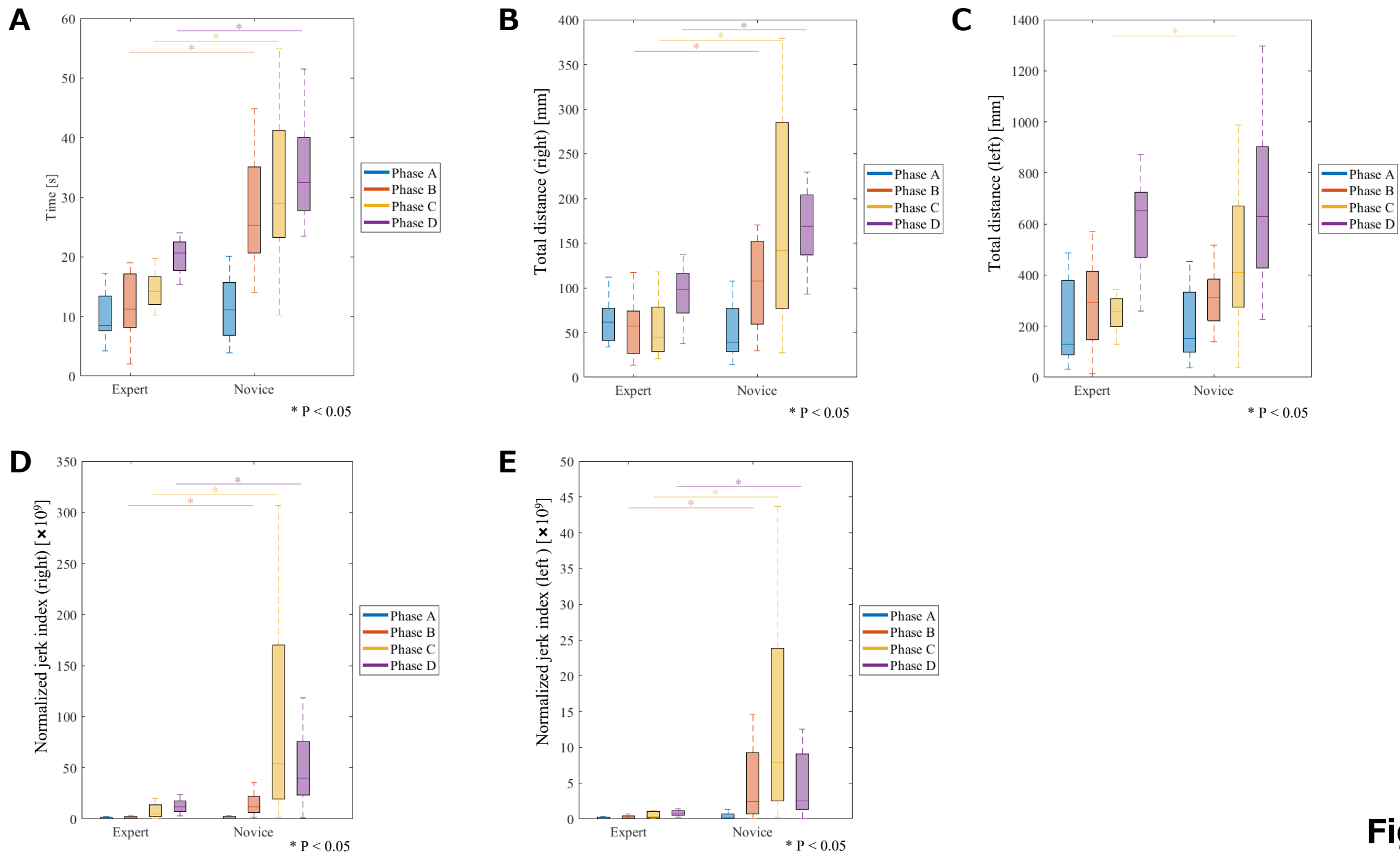


Fig. 6

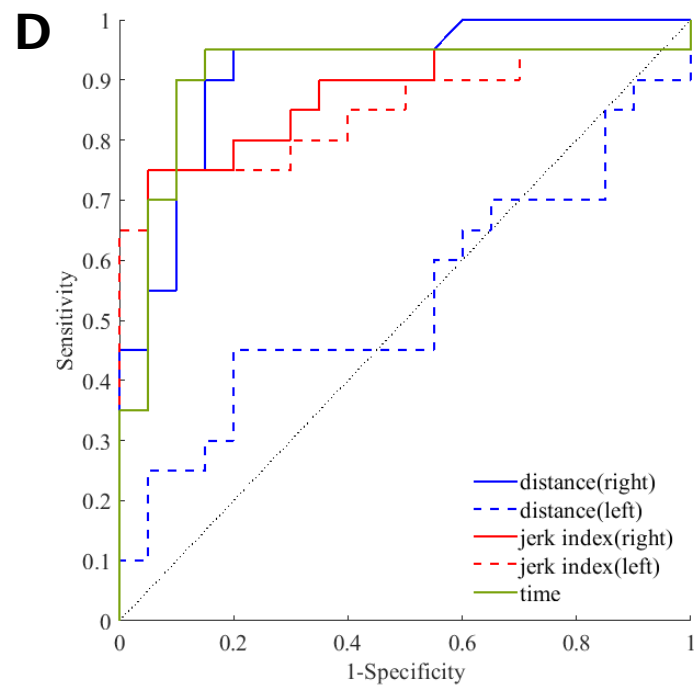
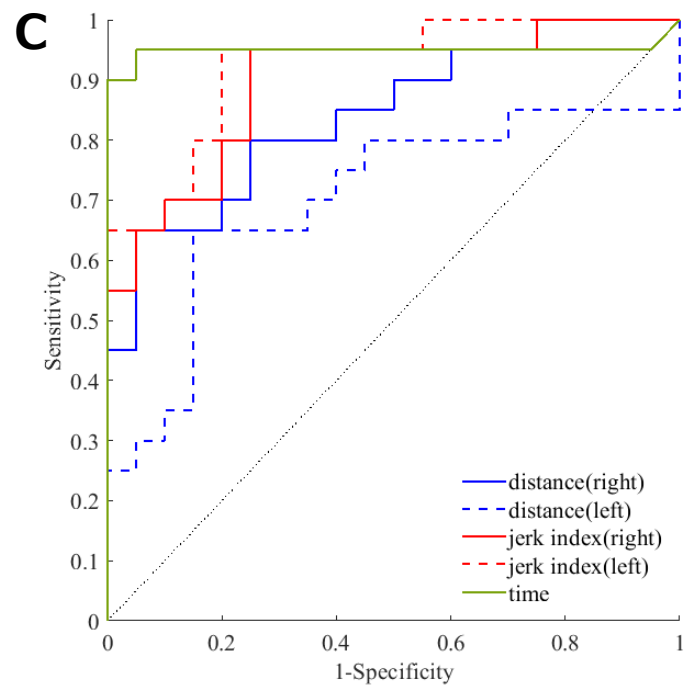
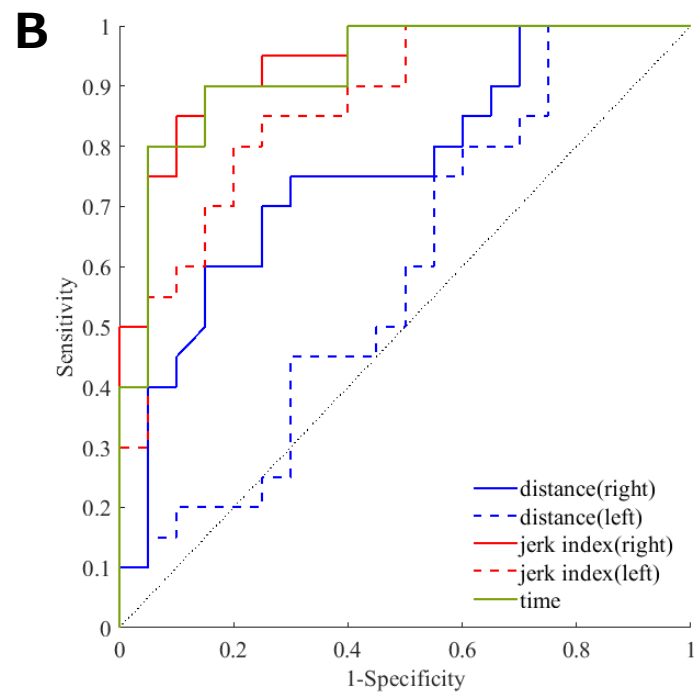
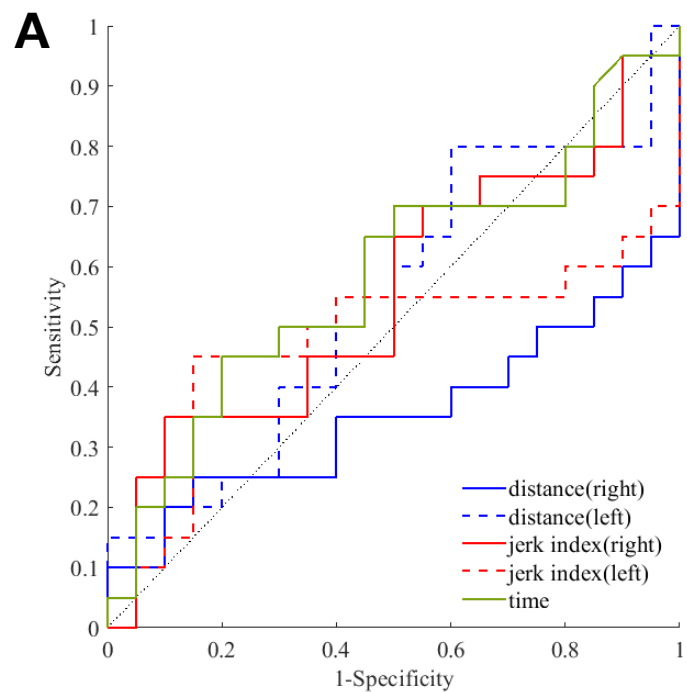


Fig. 7