Supplementary Material

Clinician-AI Collaboration in Endovascular Neurointerventions: *In-Vitro* User Study of Copilot Assistance Modes for Robotic Navigation Across the Autonomy Spectrum

Nikola Fischer, Harry Robertshaw, S. M. Hadi Sadati, Xingyu Chen, Weiyuan Deng, Benjamin Jackson, Michael Adofo Kwakye, Sara Sciacca, Naga Kandasamy, Parthiban Balasundaram, Sahan Guruge, Charles Cummins, Joe Leyon, Nasr Abdelsalam, Matthew Benger, Andrew Falzon, Alejandro Granados, Christos Bergeles, and Thomas Booth

Corresponding author: thomas.booth@kcl.ac.uk

Contents

A. Statistical Power Analysis....3
B. Experimental Procedure – Participant Briefing....3
C. Succesrate for Fully Autonomous Navigation....4
D. Survey....5

A. Statistical Power Analysis

Experiment Design

The study was designed as a within-subject experiment involving N=10 participants, each performing n=6 repetitions per Copilot Assistance Mode (CAM). This design was expected to provide sufficient statistical power to detect differences in navigation time and error count between the higher-autonomy mode Full Copilot Assistance (4) and the lower-autonomy modes Manual (0) and Direct Robot Control (1). Based on an assumed baseline completion time of 150 seconds, with a between-subject standard deviation of 30 seconds and a within-subject standard deviation of 22 seconds, the estimated statistical power was approximately 85% for an 8% improvement, 96% for a 10% improvement, and 99% for a 12% improvement in navigation time. For error rate, assuming a baseline of approximately 0.40 errors per minute, the estimated power was lower. Statistical power was approximately 40-45% for a 30% reduction in errors, 60-65% for a 40% reduction, and 80-85% for a 50% reduction. Overall, these estimates suggest that the study was adequately powered to detect navigation-time improvements of 10% or greater, whereas substantially larger reductions in error rate, on the order of 40% or more, were required to achieve a comparable level of statistical power.

Subgroups

The relatively small sample size of N=13 exceeded the minimum planned sample size but at the same time limited the ability to investigate differences between participant subgroups. Participants were initially categorized into four subgroups according to experience. Although the omnibus subgroup analysis identified a significant effect for time-to-target (T2T) with p=0.0399, no Holm-adjusted pairwise subgroup comparisons reached statistical significance across the evaluated performance metrics. Consequently, all participants were pooled for subsequent analyses to maximize statistical power. A group-adjusted sensitivity analysis yielded results consistent with the pooled analysis, indicating that subgroup composition did not affect the observed effects or the overall study conclusions. No meaningful differences were observed between experience levels. Clinical experience was therefore considered a potential covariate; however, no evidence was found to suggest that it influenced performance outcomes.

B. Experimental Procedure – Participant Briefing

Before commencing the experimental sessions, participants received an introduction to the study, including an inspection of the mock operating room (OR) environment, the phantom model, and the robotic hardware. The navigation task was then explained in detail. Participants were informed that the objective of the task was to navigate the guidewire from its starting point in the aortic arch into the target anatomy of the right internal carotid artery (RICA) or right external carotid artery (RECA) under different CAMs.

Then, a demonstration video was shown illustrating the AI Agent autonomously navigating the instrument into both RICA and RECA.

Participants were then offered a familiarization phase prior to the experimental trials. This consisted of 120 s of direct robotic control using the input device to manipulate the robotic system, followed by an additional 120 s using Full Copilot Assistance (4), during which participants could activate the AI agent and observe or interact with the autonomous navigation behavior of the instrument.

Then, a demonstration video was shown illustrating the AI agent autonomously navigating the instrument into both RICA and RECA. Participants were explicitly informed that the demonstrations were not pre-programmed or choreographed; instead, the selected actions were computed in real time based on the conditions and events occurring during each experimental run.

C. Succes Rate for Fully Autonomous Navigation

An in-depth analysis for Fully Autonomous Navigation (5) is shown in Figure 1. It reveals that performance *in silico* substantially exceeded performance in *vitro*. Overall success rates reached 81% *in silico* but only 23% *in vitro*. As shown in stage-wise success funnels, the divergence between simulated and *in-vitro* outcomes widened across task stages and became particularly apparent in the final navigation segment from the RCCA to the target vessels. Earlier branching steps were comparatively robust. In particular, success in reaching RCCA while bypassing the right subclavian artery (RSA) is high in simulation (99%) and moderate *in vitro* (≈50%), but this does not translate into successful completion of the full task. This effect was especially pronounced when targeting RECA, which remained a major failure mode in the *in-vitro* setting.

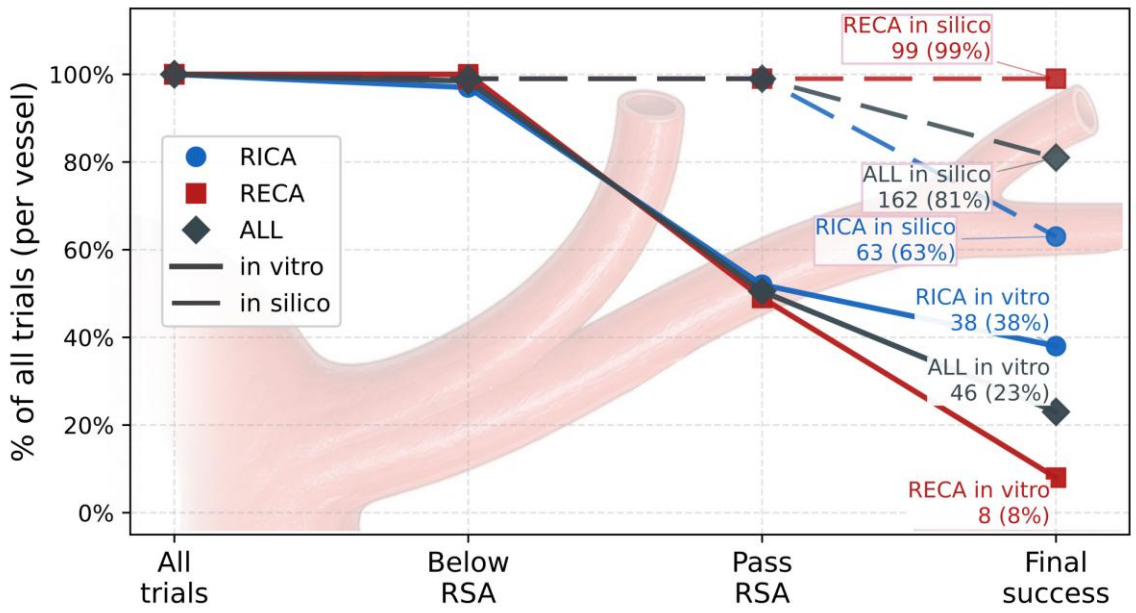

Figure 1. Stage-wise success rate funnel for Fully Autonomous Navigation (5) trials, comparing *in-vitro* and *in-silico* performance across the targets of right internal carotid artery (RICA), right external carotid artery (RECA), and all both combined. Results are based on n = 200 *in-vitro* runs and n = 200 *in-silico* runs.

D. Survey

Modification of NASA Task Load Index (NASA-TLX)

Following the experimental trials, participants completed a survey designed to capture subjective evaluations of the different CAMs. The survey assessed perceived workload for each to complement the objective performance measures collected during the experiments. Subjective workload was assessed using a modified version of the NASA Task Load Index (NASA-TLX) where all workload dimensions were rated on a 5-point Likert scale (1 = low workload, 5 = high workload).

To reduce respondent burden and facilitate repeated assessments within the surgical simulation workflow, all workload dimensions were rated on a 5-point Likert scale (1 = low workload, 5 = high workload) rather than the original 21-point scale. The temporal demand dimension was omitted because the study was designed to evaluate navigation performance and user interaction rather than workload associated with time pressure. The imposed task time limit served solely as a procedural upper bound and was not intended to create temporal constraints on participant performance.

Survey Results

The material presents the original survey responses from all 11 clinician participants. Participants are labelled P01-P11 in questionnaire response order. The name of participants is omitted to remove direct identifiers.

Raw-item abbreviations used in the tables: MD = Mental Demand; PD = Physical Demand; PERF = Performance; EFF = Effort; FRUS = Frustration. SUS1-SUS10 correspond to the ten System Usability Scale item statements in their original questionnaire order. These are raw item responses, not calculated.

Questionnaire item key

NASA-TLX:
Usefulness: How do you rate the usefulness of controlling the endovascular instruments with the Robot User Controller?

MD: Mental Demand: How mentally demanding was the task?

PD: Physical Demand: How physically demanding was the task?

PERF: Performance: How successful do you think you were in accomplishing the task?

EFF: Effort: How hard did you have to work to accomplish your level of performance?

FRUS: Frustration: How insecure, discouraged, irritated, stressed, and annoyed were you?

SUS

SUS 1: I think that I would like to use this system frequently.

SUS2: I found the system unnecessarily complex.

SUS3: I thought the system was easy to use.

SUS4: I think that I would need the support of a technical person to be able to use this system.

SUS5: I found the various functions in this system were well integrated.

SUS6: I thought there was too much inconsistency in this system.

SUS7: I would imagine that most people would learn to use this system very quickly.

SUS8: I found the system very cumbersome to use.

SUS9: I felt very confident using the system.

SUS10: I needed to learn a lot of things before I could get going with this system.

Participant background, scenario ranking, and phantom

Experience level: Experience level as a Neurointerventionalist?

Years: Number of years doing Neurointerventions (rounding to the closest full year)

Experiment date: Please set the date when this experiment was run.

Rank 1-Rank 6: Original six ranking fields from the survey, from first-ranked to sixth-ranked scenario.

Phantom rating: How realistic do you rate the phantom in general?

Table S1a. Participant background and phantom realism rating (raw responses)

ID	Experience level	Years	Date	Phantom realism
P01	Consultant	13	2026-02-18	4
P02	Consultant	14	2026-03-04	5
P03	Final year fellow	3	2026-03-04	3
P04	Consultant	18	2026-03-05	4
P05	Consultant	8	2026-03-06	5
P06	Consultant	6	2026-03-06	3
P07	Fellow	5	2026-03-09	4
P08	Consultant	10	2026-03-10	4
P09	None - Body IR Resident	0	2026-03-12	4
P10	Fellow	2	2026-03-12	5
P11	Fellow	2	2026-03-13	4

Note: UK consultant grade is US attending equivalent.

Table S1b. Scenario ranking (raw responses)

ID	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
P01	0 Manual	4a/b Copilot	3b AI Rotation	1 Direct Robot Control	2 AI Advice	3a AI Translation
P02	4a/b Copilot	1 Direct Robot Control	0 Manual	3b AI Rotation	3a AI Translation	2 AI Advice
P03	0 Manual	4a/b Copilot	3a AI Translation	3b AI Rotation	1 Direct Robot Control	2 AI Advice

ID	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
P04	2 AI Advice	0 Manual	1 Direct Robot Control	3a AI Translation	3b AI Rotation	4a/b Copilot
P05	0 Manual	1 Direct Robot Control	2 AI Advice	3b AI Rotation	4a/b Copilot	3a AI Translation
P06	4a/b Copilot	1 Direct Robot Control	2 AI Advice	3b AI Rotation	0 Manual	3a AI Translation
P07	0 Manual	3a AI Translation	4a/b Copilot	3b AI Rotation	1 Direct Robot Control	2 AI Advice
P08	0 Manual	2 AI Advice	1 Direct Robot Control	4a/b Copilot	3b AI Rotation	3a AI Translation
P09	0 Manual	1 Direct Robot Control	4a/b Copilot	3b AI Rotation	3a AI Translation	2 AI Advice
P10	0 Manual	1 Direct Robot Control	3b AI Rotation	2 AI Advice	4a/b Copilot	3a AI Translation
P11	0 Manual	4a/b Copilot	1 Direct Robot Control	2 AI Advice	3b AI Rotation	3a AI Translation

Table S2. Phantom realism explanation (verbatim responses)

ID	Why do you rate the phantom this way?
P01	Anatomically seemed correct, interface between vessel parts visible, missing pulsatility, fluid, etc.
P02	Realistic tactile and anatomical experience.
P03	Generally accurate but unable to differentiate subclavian and CCA origins.
P04	Anatomical localization and the origin of branches (mainly the subclavian on the right).
P05	Shape like real life, more stiff than real life, good model

ID	Why do you rate the phantom this way?
P06	Anatomy is not as detailed as in vivo; for example, when navigating the arch/subclavian, other vessels such as the internal mammary, thyrocervical trunk, and costo cervical trunk are occasionally encountered It is difficult to assess the haptics of the phantom's material (PLA) during the manual phase as there was no/little water/lubricant. There was an increased friction/coefficient of friction, which was significantly increased compared to in vivo. Likely a combination of friction in the catheters and the interaction with the phantom.
P07	Anatomically accurate, but no flow dynamics
P08	Similar to what we see, right size, the only thing is the vessel is stiff or rigid
P09	Realistic phantom with similar anatomy to average patient
P10	Similar to real-life situations where the wire keeps going into the right subclavian
P11	Anatomically correct - movement of wire as expected in a normal patient, although no replication of flow

Direct Robot Control

Experience: Experience with the Robot User Controller as an input device used in this study?

Table S3. Direct Robot Control: usefulness and workload item responses

ID	Experience	Usefulness	MD	PD	PERF	EFF	FRUS
P01	2nd times	4	1	1	4	2	2
P02	1st time	4	4	2	4	3	1
P03	2nd times	2	2	1	3	3	3
P04	2nd times	3	3	2	4	3	2

ID	Experience	Useful-ness	MD	PD	PERF	EFF	FRUS
P05	3rd times	4	4	1	4	4	2
P06	1st time	4	3	1	5	3	2
P07	2nd times	2	4	2	2	5	4
P08	2nd times	5	3	4	5	2	1
P09	1st time	3	2	1	2	3	2
P10	3rd times	4	2	1	4	3	1
P11	1st time	2	3	3	2	4	5

Note: Our group has performed other different robotic experiments - for example validating robotic controller interfaces - but none testing AI Copilot Assistance Modes.

Table S4. Direct Robot Control: open-text responses

ID	Why rated this way	Suggested improvement
P01	Intuitive interface, responsive; can't do speedy spinning of the GW	speedy spinning
P02	No elements that were not controllable, same motions from clinic, preferred over joystick-based interfaces, liked continuous control rather than click/iteratively,	none
P03	It is very good but does not advance me from manual control	lag reduction
P04	Still prefer the tactile feedback when being fully manual, but AI advice was helpful.	Prefer this to joystick
P05	Does correspond to what I want to do; right learning curve needed to work well	1. Controller speed too sensitive, move slightly faster it goes super fast, subtle movement like rotation, is not doing anything; 2 latency between imaging and the action, for fine movement it is a problem, hard to see little movement of the wire, finally see it it was too much; 3. seems like the latency of machine, also observed in OR, maybe framerate of fluoroscopy
P06	The controller is intuitive and easy to learn. The translational element seems much more responsive than the rotational. The rotational element seems to be more effective when performed in tandem/simultaneously with translation. I like the fact that the controller can be used with a single hand, allowing for rotation and translation in single coordinated movements. This freed up the left hand to control the AI/controller reset button (green). Allows manipulation over a short distance.	Overall latency (could be controller and/or latency due to other reasons); and 1:1 haptics, particularly for rotational elements. For the rotational element, infinite rotation in any direction may lead to problems with the wire (particularly if 014 or smaller wires).

ID	Why rated this way	Suggested improvement
P07	Time lag between the manual input and the robot output makes movements less smooth	To reduce the time lag mentioned above as well as the smoothness of movements (especially rotations)
P08	More like how we use the device (button not similar)	Rotation is not so responsive, difficult to rotate, difficult to fine movement
P09	Reasonable. Translation works well. Rotation less so. Difficult to judge degree of rotation of wire when rotating controller. Slight lag between action of user and movement of robot	Consider making rotation mechanism thinner so it feels closer to wire manipulation. Maybe more ergonomic/comfortable structure. Could become uncomfortable with prolonged use
P10	Easy to learn - fair response time	The speed to reduce lag
P11	Friction within the slide mechanism limited the fine translation control. Replication of wire torque using rotation control to actual wire movement (one to one) was not provided, resulting in difficulty in fine-tuning the wire angle. Ergonomically challenging to find an appropriate grip that would allow the use of the robot controller for extended periods of time	Lower friction in translation/slider mechanism. Smaller rotation control knobs. Method to remove noise/over input, particularly in the rotation controller given the delay between action and apparent robot/wire movement

Table S5. Direct Robot Control: raw SUS item responses (SUS1-SUS10)

ID	SUS1	SUS2	SUS3	SUS4	SUS5	SUS 6	SUS 7	SUS 8	SUS 9	SUS 10
P01	5	2	4	1	5	1	5	2	5	2
P02	5	1	5	2	4	2	5	1	5	2
P03	2	1	5	5	4	2	5	2	4	1
P04	4	2	4	4	4	1	4	1	4	2
P05	4	2	4	1	4	1	4	2	4	4
P06	4	2	5	1	3	1	4	1	4	1
P07	2	1	3	2	4	1	3	3	2	5
P08	4	1	1	1	5	1	5	1	5	1
P09	1	2	3	1	3	3	5	4	3	2
P10	4	1	5	1	4	1	4	1	4	1
P11	1	2	2	1	4	5	4	4	1	2

AI Advice

Table S6. AI Advice: usefulness and workload item responses

ID	Usefulness	MD	PD	PERF	EFF	FRUS
P01	2	4	2	2	4	4
P02	2	4	4	2	4	4
P03	1	1	1	3	3	3
P04	5	2	2	5	1	1

ID	Usefulness	MD	PD	PERF	EFF	FRUS
P05	1	2	1	3	4	2
P06	3	1	1	5	3	1
P07	1	5	1	2	5	4
P08	4	2	1	5	1	1
P09	1	4	1	1	4	4
P10	3	4	1	4	2	1
P11	1	2	3	2	3	4

Table S7. AI Advice: open-text responses

ID	Why rated this way	Suggested improvement
P01	Mixed: counterintuitive recommendations at the start, particularly for translation; tried to follow rotation at the RECA, difficult to succeed by following it; therefor ignored at RECA instead of following it; For subclavian, advice was ignored and user got trapped in subclavian; next time, following advice would protect from falling into subclavian; Alternative: Voice could be an alternative, but careful as it could be irritating and distracting	Interesting to share these insights for fully copilot control, but delivered as extra AI-Advise not very useful
P02	Too much time-lag, more parameters to consider therefore hard to control; flow of information not smooth	Less lag, more in tune/sychronised with user
P03	Because you cannot look at advice whilst navigating easily, and also advice is slightly counter-intuitively designed	Instead of max/min, you could provide an estimate of distance to torque for rotations, and it may be more intuitive to switch direction to align with the actual physical motion of the hand
P04	The combination of translation and rotation movement suggestion was helpful, along with the degree of force required, although it would be difficult to exert the exact amount of force recommended.	Nothing in particular
P05	Noticed different from what I am doing, but mine works	It would be better for it to learn what I am doing
P06	I understood the power gauge to imply velocity; I did not understand the Max and Min	More clearly depict the AI advice; ie Max+ve and Min -ve are not intuitive
P07	The advice panel is separate from the images (which makes it difficult to concentrate on both). The advice provided wasn't very helpful as it kept rapidly alternating, making it difficult to follow.	If the advice could be overlaid on the images, that would make it easier to follow.

ID	Why rated this way	Suggested improvement
P08	Difficulty following the advice; it is useful, smooth, no need for the advice	Giving translation and rotation together, in practice, do the trans and rotate separately
P09	Not helpful at all. There was information overload with all 4 possible movements (i.e. forward, back, clockwise, anticlockwise) changing rapidly, and sometimes unintuitively. For example, advice would go from max clockwise to max anticlockwise and then back again in 1 second. This was confusing. Step-by-step instructions on how to manipulate the wire might be better. Example instruction could be "move forward x cm, then rotate clockwise by y degrees, then move forward again by z cm".	See above
P10	Pros - sometimes helpful with direction of rotation; cons - too fast	Slower instructions - more focused on direction of rotation
P11	Unclear on the rotation direction and velocity (what is max speed). I did not encounter a scenario where following the AI advice was crucial to navigation. Limiting factor in navigation was robot usability and control rather than required actions.	Clearer values for min/max with real-world references (faster/slower compared to current movement speed or more/less rotation). May be more useful in more anatomically complex scenarios where direction of motion or rotation is not obvious based on the location.

Table S8. AI Advice: raw SUS item responses (SUS1-SUS10)

ID	SUS1	SUS2	SUS3	SUS4	SUS5	SUS 6	SUS 7	SUS 8	SUS 9	SUS 10
P01	1	2	4	1	2	5	5	3	3	1
P02	1	4	2	2	4	3	4	3	2	3
P03	1	3	2	4	3	3	1	4	1	4
P04	5	1	5	2	4	1	5	1	5	2
P05	1	2	4	1	3	2	4	2	4	4
P06	3	4	3	1	3	1	3	2	3	1
P07	1	4	5	3	1	5	1	5	1	5
P08	5	1	5	1	5	3	5	1	4	1
P09	1	5	1	1	2	4	2	5	2	2
P10	3	1	3	1	4	1	4	4	3	1
P11	1	2	1	4	4	5	2	3	1	3

Single-motion AI Assist: Translation

Table S9. Single-motion AI Assist: Translation: usefulness and workload item responses

ID	Usefulness	MD	PD	PERF	EFF	FRUS
P01	1	4	3	1	4	4
P02	3	3	3	4	4	3
P03	4	2	1	4	3	2

ID	Usefulness	MD	PD	PERF	EFF	FRUS
P04	3	2	2	4	4	2
P05	1	4	1	1	5	3
P06	1	3	1	4	4	2
P07	5	2	1	4	2	2
P08	1	5	2	1	4	4
P09	2	2	1	3	4	2
P10	2	3	1	3	4	4
P11	1	1	1	1	3	5

Table S10. Single-motion AI Assist: Translation: open-text responses

ID	Why rated this way	Suggested improvement
P01	Translation of AI would trap user, e.g., in subclavian or in the wrong RECA/RICA, and not autonomous retraction useful; difficult to get in sync with Agent	Overcome trapping, more practice might be beneficial
P02	Act together, AI controls one, difficult to control another	Not sure
P03	Because I felt that it was a nice balance and gave me one less thing to think about	NA
P04	Difficult to predict what the other system is working on and apply rotation at the appropriate time when translation is done.	As last question
P05	Because find it difficult to rotate from advancement	Prefer something either not controlled by AI or controlled by AI, unless it is a system learn from my movement

ID	Why rated this way	Suggested improvement
P06	Rotation is much more effective simultaneously with translation; as a result, it is difficult to rotate when you are reliant on AI's translation, which may be different to your intuition	Ability to prompt AI to increase or decrease translation to allow for simultaneous manoeuvres
P07	It was easy to use as the rotations translated quickly, which made wire guidance easy	Nothing
P08	For each rotation, it needs to be still, but AI is doing it with me in parallel, like two people working on the same thing; hard to predict	Should know what AI is going to do, like extra advice showing what it is going to
P09	In theory, these should be easy to use; however, in reality I found that I was out of sync with the AI movements for the majority of the time. My rotation of the wire was often slower than the translational movement by the AI. This led to the wire ending up in unintended positions. Any subsequent correction I made would, in turn, confuse the AI in its movement, and vice versa.	Could try slowing things down slightly and segmenting the movements. E.g. I rotate while wire is stationary, AI moves a certain distance until it stucks. I then reorientate wire before AI starts moving again.
P10	It's too fast to work with human rotation - very difficult to salvage if it goes to the wrong anatomy	More human control, including pausing AI control or controlling the speed of translation
P11	Translation control was the easier aspect of the robot control scheme. I rarely felt limited by my ability to translate when using robot control without AI-assist. Often led to failure to access the required vessel as there were points from which the AI translation could not recover	AI to modulate human input rather than entirely controlling the translation - e.g. prevent from losing position by moving too quickly at the incorrect angle or overshooting a location.

Table S11. Single-motion AI Assist: Translation: raw SUS item responses (SUS1-SUS10)

ID	SUS1	SUS2	SUS3	SUS4	SUS5	SUS6	SUS7	SUS8	SUS9	SUS10
P01	1	4	2	1	1	4	2	4	2	4
P02	3	3	4	3	3	2	4	3	4	2
P03	3	1	5	5	4	2	4	1	4	1
P04	3	2	2	2	4	2	4	2	4	2
P05	1	5	1	1	3	4	1	4	1	5
P06	3	2	2	1	4	1	4	3	3	1
P07	3	1	5	1	5	1	5	1	4	2
P08	1	4	3	1	4	5	3	5	2	1
P09	1	2	3	1	2	3	3	4	3	2
P10	1	3	1	1	1	4	4	5	3	1
P11	1	3	4	2	4	5	5	5	2	3

Single-motion AI Assist: Rotation

Table S12. Single-motion AI Assist: Rotation: usefulness and workload item responses

ID	Usefulness	MD	PD	PERF	EFF	FRUS
P01	4	2	2	4	2	1
P02	3	2	3	4	2	2
P03	3	4	1	2	3	4
P04	2	4	3	3	4	3

ID	Usefulness	MD	PD	PERF	EFF	FRUS
P05	3	2	1	5	1	1
P06	4	2	3	5	2	1
P07	4	2	1	4	2	2
P08	3	2	1	5	1	2
P09	3	2	1	3	2	2
P10	4	1	1	4	3	1
P11	2	2	2	3	4	4

Table S13. Single-motion AI Assist: Rotation: open-text responses

ID	Why rated this way	Suggested improvement
P01	Higher level of perceived control; possible to slow down translation before critical bifurcations; more synergistic compared to 3a	no comment
P02	cannot do trans and rotate it together	na
P03	OK, but more frustrating than translation	Just better rotation from AI
P04	Lack of co-ordination between translation and rotation.	As above
P05	Not useful association; easier to advance and let the machine rotate	Learn from operators, both training and real-time
P06	AI seems to have a good understanding of this element and can be done simultaneously with translation that is user controlled	Speed of rotation, if possible
P07	I cannot recall all the details, but it was not difficult to use	Nothing

ID	Why rated this way	Suggested improvement
P08	Same issue, not knowing what AI is doing	Like previous, additional advice will be better
P09	Slightly better when I was controlling translation compared to rotation. Felt I had more control and could stop translational movements while AI was rotating. Once AI had stopped rotating, I could then move the wire	Maybe a suggestion from AI of how far to translate the wire before further rotational calibration will be required.
P10	I found it useful as in some cases it was better than human trial and error attempts	NA
P11	AI-assisted seemed better able to predict rotation input relative to wire movement than I was, allowing for often accurate wire angulation for stage of task. This solves the limiting factor of robot control that the rotation control function was challenging to use. If robot control were easier, then the AI would not have been particularly useful.	None

Table S14. Single-motion AI Assist: Rotation: raw SUS item responses (SUS1-SUS10)

ID	SUS1	SUS2	SUS3	SUS4	SUS5	SUS6	SUS7	SUS8	SUS9	SUS10
P01	4	2	4	1	4	2	4	2	4	2
P02	3	2	4	2	3	2	4	2	4	2
P03	1	2	4	5	4	2	4	2	3	2
P04	2	2	4	3	4	2	4	2	3	2
P05	3	1	4	1	4	2	4	2	4	4
P06	4	1	5	1	4	1	4	2	4	1

ID	SUS1	SUS2	SUS3	SUS4	SUS5	SUS6	SUS7	SUS8	SUS9	SUS10
P07	3	1	5	1	1	1	5	1	4	2
P08	3	2	5	1	5	1	4	2	5	1
P09	3	2	3	1	3	2	4	3	3	1
P10	4	1	5	1	5	1	5	1	5	1
P11	1	2	3	2	3	5	4	3	2	2

Full Copilot Assist: Preset Phantom Sections

Table S15. Full Copilot Assist: Preset Phantom Sections: usefulness and workload item responses

ID	Usefulness	MD	PD	PERF	EFF	FRUS
P01	5	2	2	4	2	1
P02	5	2	2	4	3	1
P03	4	2	1	3	3	2
P04	2	2	2	4	2	2
P05	1	3	1	1	4	2
P06	1	2	2	3	2	3
P07	5	1	1	5	1	1
P08	4	1	1	4	2	1
P09	3	2	1	3	2	2
P10	3	1	1	4	1	1

ID	Usefulness	MD	PD	PERF	EFF	FRUS
P11	2	2	3	2	2	4

Table S16. Full Copilot Assist: Preset Phantom Sections: open-text responses

ID	Why rated this way	Suggested improvement
P01	Worked each run, no intervention necessary	no comment
P02	Easy navigation, comfortable to let AI do the task, less human effort	Touchscreen to choose target rather than having an on/off button; info screen on how much the AI thinks it has completed (e.g., 60% of the navigation done) to omit miscommunication with AI
P03	Nice option	Nothing much
P04	The predefined section was not the most challenging section to navigate without AI	Ability to choose sections
P05	It never crossed the subclavian	Something more to take into account: tension on the wire, material fatigue
P06	The AI failed; unable to catheterise CCA	Ability to override;
P07	The Copilot was able to quickly navigate the preset segment, which made the whole process quick	It kept going into the subclavian artery in each attempt, which needs to be improved
P08	It was able to cross the arch, but always went into the subclavian	People find different difficulties, not complete AI. AI doing what it is good at, and operator doing what they are good at. Choice will be better

ID	Why rated this way	Suggested improvement
P09	AI failed a few times, by which point the wire was in a less-than-ideal position, requiring me to start again. When AI was successful in navigation to CCA, it did make my job easier, and in those cases was useful	Make AI more accurate at getting to the correct target
P10	There was a 50-50 chance of success	Again, I think having more human control, such as pausing the system when needed or changing the AI's decision about the direction of rotation, would be helpful
P11	2/4 times the AI managed to navigate to the endpoint of the predefined section without interaction. Given the challenges of the robot control, it facilitated passing the most difficult section. 50% of failure occasions, however, is not ideal.	More reliability.

Full Copilot Assist: Individual Phantom Sections

Table S17. Full Copilot Assist: Individual Phantom Sections: usefulness and workload item responses

ID	Usefulness	MD	PD	PERF	EFF	FRUS
P01	3	3	2	3	2	3
P02	4	2	2	4	2	2
P03	4	2	1	3	3	2
P04	4	2	1	5	1	1
P05	4	3	1	5	4	1
P06	4	4	3	5	4	3

ID	Usefulness	MD	PD	PERF	EFF	FRUS
P07	5	1	1	5	1	1
P08	5	1	1	5	1	1
P09	3	2	1	3	2	2
P10	5	1	1	4	1	1
P11	4	2	2	4	3	3

Table S18. Full Copilot Assist: Individual Phantom Sections: open-text responses

ID	Why rated this way	Suggested improvement
P01	Explored performance for RECA and realized less performance than in other anatomical compartments (in the set ones); mixed performance perceived	Agent performance on weak sections should be improved, RECA
P02	Same as 4a, please see	NA
P03	Because it would be a nice option to ask AI if you are struggling with something	Nothing
P04	Ability to define individual sections means challenging sections can be selected for AI use or assistance.	Ability to predefine sections and option to have assist or direct AI control.
P05	Can choose whatever stage	Learn from operator
P06	Able to use AI to complete parts of the task that you are certain it will accomplish, allowing me to take over only at the difficult/critical steps (i.e., catheterizing the CCA, not the Subclavian)	Repositioning the dead man switches for AI and for catheter reset in a more ergonomically friendly position controlled with one hand

ID	Why rated this way	Suggested improvement
P07	It was useful to navigate the brachiocephalic artery segment rapidly (which I found somewhat difficult with the manual control of the robot)	Nothing
P08	can choose the difficult part that I know is difficult for humans and do it myself when I know AI is not good at the subclavian	Nothing
P09	Again, when the AI worked and got me through a difficult section, it was very helpful. Sometimes, it failed to get me across a difficult section, and I had to take over. This wasted a lot of time.	Make AI more accurate in navigating difficult segments.
P10	Very useful to have flexibility of when AI can make decisions - it increases the speed of the task - easy to override if not successful	NA
P11	Most useful mode as possibly to take manual control when AI assist is not able to perform task. Challenges of navigating into innominate using robot control could be eliminated by using AI Assist, which seemed well able to clear this section of the model. If robot control were better, then AI Assist would be less useful	N/A

Note: Innominate is another term for brachiocephalic artery.

Table S19. Full Copilot Assist: Individual Phantom Sections: additional open-text response

ID	How did you choose your individual phantom sections for the full Copilot Assist?
P01	Seen agent work for subclavian and arch part, used it there and hoped for good performance in new sections, known to be difficult (RECA)

ID	How did you choose your individual phantom sections for the full Copilot Assist?
P02	Lower risk navigation with AI, free up mental load for clinician
P03	Slightly randomly but particularly helpful getting into brachiocephalic
P04	Initial attempts: did not use it to familiarise with the equipment, and didn't need it in the long run as got familiar with the equipment. However, it will be useful in challenging situations and particularly after multiple attempts when fatigue sets in.
P05	Beginning: using AI cross-subclavian, but it fails, so I have to take over
P06	In one scenario, I tried to activate AI in segments, which I thought would be successful based on previous runs.
P07	I chose the Copilot to navigate the brachiocephalic artery segment, which I found somewhat difficult with manual control of the robot
P08	The same as the previous; when I have difficulties, I will give it a try
P09	Sections where I thought I might struggle
P10	Speeding up the initial part of the task before reaching target anatomy - useful in more straight segments of the anatomy - after some attempts, I could understand the parts the AI did well, so I used it in those parts
P11	Segments I found challenging when using the robot control would be attempted with co-pilot assist. If that was not working after a period of time, then I would take manual control to either pass the challenging segment manually or reset the start point.

Table S20. Full Copilot Assist: Individual Phantom Sections: raw SUS item responses (SUS1-SUS10)

ID	SUS1	SUS2	SUS3	SUS4	SUS5	SUS6	SUS7	SUS8	SUS9	SUS10
P01	4	2	3	1	4	4	2	2	3	4
P02	4	2	4	2	4	2	4	2	4	2

ID	SUS1	SUS2	SUS3	SUS4	SUS5	SUS 6	SUS7	SUS8	SUS9	SUS1 0
P03	4	1	5	5	4	2	5	1	4	2
P04	5	1	4	2	4	2	5	2	4	2
P05	1	4	1	1	2	4	1	4	1	4
P06	4	1	4	1	4	1	4	2	4	1
P07	4	1	5	1	4	2	5	1	5	2
P08	5	1	5	1	5	1	5	1	5	1
P09	4	2	4	1	3	3	4	2	3	1
P10	4	1	3	1	4	2	4	1	5	1
P11	2	2	4	1	4	4	4	3	2	2

Additional Copilot questions and comments

Intermediate target points: How would you rate the possibility of picking intermediate target points instead of one final target of navigation?

Envisioned scenarios/features: Which Copilot scenarios or features (not presented in this study or survey) would you envision as helpful and why?

Further comments: Any further comments you would like to make?

Table S21. Additional Copilot questions and final comments (raw responses)

ID	Intermediate target-point rating	Envisioned Copilot scenarios/features	Further comments
P01	3	multiple instruments, multiple locations including the left carotid artery	NA

ID	Intermediate target-point rating	Envisioned Copilot scenarios/features	Further comments
P02	5	Time estimation for navigation task accomplishment from start to finish	NA
P03	4	If you cannot engage a vessel with a wire - final option to try.	Nice system
P04	4	Potential use in device deployment Use in device selection (eg diameter and length of device) based on individual device properties	1. Good experience and it was enjoyable 2 Lovely team
P05	4	Intracranial vessels anatomy	Improvement in ease of use, compared to last time; controller more sensitive, easy to control, doing more what I want than last time
P06	5	Robot control of catheters/intermediate would be interesting to explore. For this particular scenario, selecting targets as well as selecting "danger zone" where you are penalized or encouraged to avoid to assess accuracy	Very well run and organized.
P07	5	Copilot overlay on the images, which may be helpful	No further comments
P08	5	Give freehand to select; force them to use manual, difficulty so use AI, like AI help human (needed more for junior) and human help AI (more in senior level)	NA

ID	Intermediate target-point rating	Envisioned Copilot scenarios/features	Further comments
P09	4	In body IR, we treat a lot of peripheral vascular disease and often have to cross tight stenoses or occlusive disease. Would be interesting to see how Copilot manages these sorts of conditions. In addition, there can be complications of manipulating guidewires (e.g., dissection). Again, it would be useful to know how the AI agent performs at handling these complications.	Manual manipulation was by far easier and better still. Hopefully your work leads to better systems that change this in the future.
P10	5	AI decisions about rotation - using AI in straight anatomy to increase speed of tasks	NA
P11	4	2 planar integration of location/target point. To facilitate crossing small segments where it is not clear which direction should be travelled to access the required vessel.	NA