CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Validation studies of virtual reality simulation performance metrics for mechanical thrombectomy in ischemic stroke
Authors	Crossley, Robert;Liebig, Thomas;Holtmannspoetter, Markus;Lindkvist, Johan;Henn, Pat;Lonn, Lars;Gallagher, Anthony G.
Publication date	2018
Original Citation	Crossley, R., Liebig, T., Holtmannspoetter, M., Lindkvist, J., Henn, P., Lonn, L. and Gallagher, A.G., 2019. Validation studies of virtual reality simulation performance metrics for mechanical thrombectomy in ischemic stroke. Journal of neurointerventional surgery, 014510 (7pp.). DOI: 10.1136/neurintsurg-2018-014510
Type of publication	Article (peer-reviewed)
Link to publisher's version	https://jnis.bmj.com/content/neurintsurg/early/2019/01/17/neurintsurg-2018-014510.full.pdf - 10.1136/neurintsurg-2018-014510
Rights	© Author(s) (or their employer(s)) 2018. This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/. - https://creativecommons.org/licenses/by/4.0/
Download date	2026-11-06 20:29:09
Item downloaded from	https://hdl.handle.net/10468/7602



University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

OPEN ACCESS

ORIGINAL RESEARCH

Validation studies of virtual reality simulation performance metrics for mechanical thrombectomy in ischemic stroke

Robert Crossley,[1] Thomas Liebig,[2] Markus Holtmannspoetter,[3] Johan Lindkvist,[4] Pat Henn,[5] Lars Lonn,[6] Anthony Gerald Gallagher[7]

[1]Neuroradiology, North Bristol NHS Trust, Southmead Hospital, Bristol, UK
[2]Institute of Neuroradiology, Ludwig Maximilians University of Munich, Munich, Germany
[3]Department of Neuroradiology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark
[4]Mentice AB, Gothenburg, Sweden
[5]School of Medicine, University College Cork, Cork, Ireland
[6]Department of Cardiovascular Radiology, National Hospital, Copenhagen University, Copenhagen, Denmark
[7]Faculty of Life and Health Sciences, Ulster University, Londonderry, UK

Correspondence to
Professor Anthony Gerald Gallagher, Faculty of Life and Health Sciences, Ulster University, Londonderry BT48 7JL, UK; anthonyg.gallagher@btinternet.com

Received 17 October 2018
Revised 10 December 2018
Accepted 18 December 2018

Check for updates

© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Crossley R, Liebig T, Holtmannspoetter M, *et al*. *J NeuroIntervent Surg* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/neurintsurg-2018-014510

ABSTRACT

Introduction Mechanical thrombectomy (MT) has transformed the treatment of ischemic stroke. However, patient access to MT may be limited due to a shortage of doctors specifically trained to perform MT. The studies reported here were done to (1) develop, operationally define, and seek consensus from procedure experts on the metrics which best characterize a reference procedure for the performance of an MT for ischemic stroke and (2) evaluate their construct validity when implemented in a virtual reality (VR) simulation.

Methods In study 1, the metrics for a reference approach to an MT procedure for ischemic stroke of 10 phases, 46 steps, and 56 errors and critical errors, were presented to an international Delphi panel of 21 consultant level interventional neuroradiologists (INRs). In study 2, the metrics were used to assess 8 expert and 10 novice INRs performing a VR simulated routine MT procedure.

Results In study 1, the Delphi panel reached consensus on the appropriateness of the procedure metrics for a reference approach to MT in ischemic stroke. Group differences in median scores in study 2 demonstrated that experienced INRs performed the case 19% faster (P=0.029), completed 40% more procedure phases (P=0.009), 20% more steps (P=0.012), and made 42% fewer errors (P=0.016) than the novice group.

Conclusions The international Delphi panel agreed metrics implemented in a VR simulation of MT distinguished between the computer scored procedure performance of INR experts and novices. The studies reported here support the demonstration of face, content, and construct validity of the MT metrics.

INTRODUCTION

Mechanical thrombectomy (MT) has transformed the treatment for large vessel occlusive stroke. Functional independence outcomes are now achieved in up to 70% compared with 10–20% of patients treated with MT versus traditional medical therapy.[1] Multiple randomized controlled trials have demonstrated considerable treatment effects, with numbers needed to treat as little as 2.8. To place this figure into perspective, the numbers needed to treat for cardiac coronary interventions and long term mortality is 25.[2] The HERMES collaborative meta-analysis provides an excellent summary of these MT data as a comprehensive description of the benefits is outside the scope of this article.[3]

Despite the proven effectiveness of MT, access is limited in regions and countries, primarily due to the shortage of interventional neuroradiologists (INR) specifically trained to perform MT. Traditionally, doctors acquire their skills to perform new procedures on patients. However, image guided procedures impose unique human factor challenges, which expose patients to potential risk during the doctor's advancement along the learning curve.[4] This traditional approach to training can be protracted for doctors to develop skills to achieve proficiency in performance. Currently, there is no objective, transparent, and reliable way to verify proficiency of performance.

Concerns about training in medicine and healthcare have forced a radical rethink of how to optimally train doctors.[5 6] Simulation is proposed as a potential significant contributing element to this training dilemma.[7 8] Clinical trial data indicate that simulation based training produces a superior skill set.[9] Furthermore, simulation should augment the training process and not be just an educational experience,[10] rather simulation should be a tool to achieve a quality assured performance benchmark by the completion of training.[11]

Optimal simulation training platforms and training programs are derived from a comprehensive procedure characterization, and the identification and definition of procedure steps and deviations from optimal performance (or errors)[12] which once validated are used to give trainees formative feedback on their performance.

These metrics are then used to build a curriculum, a (simulation) training platform with the metrics integrated and the quantitate definition of performance benchmarks, which trainees must demonstrate before progression to in vivo clinical practice.[12–14]

The ultimate procedural goal in MT is the safe and fast recanalization of the occluded target artery to restore blood supply to ischemic brain tissue. This must be performed without damaging the navigated proximal arteries or the introduction of air or thrombotic emboli into new territories which were previously unaffected (embolization into new territory).

The studies reported here undertook a process to characterize in detail a safe reference approach to MT for a neurovascular trainee who has demonstrated independent competency in diagnostic

J NeuroIntervent Surg: first published as 10.1136/neurintsurg-2018-014510 on 17 January 2019. Downloaded from http://jnis.bmj.com/ on 4 March 2019 by guest. Protected by copyright.

Table 1 Phases of the mechanical thrombectomy procedure, with the beginning and end of each phase clearly defined

Procedure phase	Title	Phase begins and ends	
I	Femoral access to descending aorta	Begins	Femoral arterial access (as per institutional protocol)—wire enters the groin
		Ends	Wire at mid thoracic aorta T5–8, above celiac trunk below origin of left subclavian
IIa	Manipulation in the aortic arch using a multipurpose catheter (MPA, MPB, MPC, VER, BER,…)	Begins	Wire advances beyond mid thoracic aorta
		Ends	Tip of the select catheter is immediately proximal to brachiocephalic trunk
IIb	Manipulation in the aortic arch using a Simmons catheter (SIM1, SIM2, SIM3, VTK,…)	Begins	Wire advances beyond mid thoracic aorta
		Ends	Tip of the selected catheter is immediately proximal to brachiocephalic trunk
III	Access from aortic arch to common carotid artery	Begins	Removal of guidewire and acquisition of roadmap image
		Ends	Re-evaluation by contrast injection through select catheter positioned in the ICA
IV	Advancement and stable positioning of sheath/BGC	Begins	Advancement of BGC into ICA over select catheter.
		Ends	Acquisition of iso-centered biplane angiogram of intracranial vessels
V	Tri-axial access to and crossing of thrombus	Begins	Shaping distal microwire tip to facilitate atraumatic navigation
		Ends	Retrieving microwire from correctly positioned microcatheter
VI	Deployment of stentriever	Begins	Stentriever introduced to microcatheter hub
		Ends	Stentriever fully unsheathed
VII	Retrieval of thrombus	Begins	Reducing tension on DAC until tip starts to retract
		Ends	Deflation of BGC
VIII	Angiographic control of success of retrieval	Begins	Retraction of DAC to proximal ICA (below petrous segment, above CCA bifurcation)
		Ends	Recanalization confirmed on DSA
IX	Full head control angiogram	Begins	Remove DAC
		Ends	Control angio of cervical carotid through BGC in proximal CCA
X	Safe removal of devices (as unit), access vessel closure	Begins	Retraction and removal of devices
		Ends	Arterial access closure (according to institutional protocol)

Ber, Berenstein catheter; BGC, balloon guide catheter; CCA, common carotid artery; DAC, distal access/aspiration catheter; ICA, internal carotid artery; MPA, multipurpose A; MPB, multipurpose B; MPC, multipurpose C.

cerebral angiography. We sought to characterize the endovascular procedure per se and not cover the issues of diagnostic imaging, patient selection, or anesthesia. It was hypothesized that we could identify procedure, phases, steps, and procedure errors (deviations from optimal performance). It was also hypothesized that experienced INRs could reach consensus on these metrics, and when integrated into a physics based virtual reality simulation that they could distinguish between the performance of experienced and novice INRs performing the procedure on the simulator.

METHODS

Procedure characterization

This study received expedited institutional review board approval from the Cork University Hospital ethics committee (ECM 4 (d) 04/07/17; June 22, 2017). The research was supported by a grant from the Swedish government agency for innovation (Vinnova) to Mentice AB (Gothenburg, Sweden). Three INRs (MH/TL and RC), a behavioral scientist (AGG), an interventional radiologist (LL), and a senior project engineer (JL) formed the procedure characterization group. The INRs had >8 years of practice experience and mastered the full range of intracranial and spinal neurovascular interventions.

The project focused on the anterior circulation, while noting there is definite overlap with performance of thrombectomy in the posterior circulation. Procedure characterization was performed over a 7 month period in five face to face meetings. Additional time was spent at the simulator development facility to optimize technical features regarding the fidelity of the simulation platform. The procedure was deconstructed into 10 phases (table 1), consisting of 46 steps; 57 possible errors were defined. Error severity was defined as either critical or non-critical.

Phases I and X are generic skills in endovascular therapies. Safe retrograde femoral access and closure were defined as being in accordance with institutional protocols.

There are certain tenets in angiography that are constant, irrespective of which organ is being studied. In general, a catheter should only be advanced over a suitable guidewire. If resistance to catheter or wire advancement is felt, pressure should be released in order to ensure that the tip of the wire or catheter has not damaged the vessel lining. When a catheter is being manipulated in the absence of a wire movement it should be backwards rather than pushing forwards to avoid 'vessel scraping'. Deviation from these basic principles were deemed to constitute an error.

Phase II was divided into subsections a and b. Selection of the innominate and right internal carotid arteries in a type I arch is generally straightforward and can be performed by a forward curve catheter. These catheters can be used directly once the leading guidewire has selected the target cervical vessel. Selection of the left common and internal carotid arteries can be more challenging. These often require the use of a reverse curve Simmonds (SIM 1, 2, 3 etc) catheter. This catheter requires manipulation in situ to form the working shape. This is technically demanding and requires additional steps. There are several recognized ways to form the working shape of the SIM catheters, in the ascending aorta over the aortic valve, in the left subclavian artery over the abdominal aortic bifurcation. The group all form the catheter in the left subclavian artery by preference. This was felt to be the safest and most reproducible technique by the group (unanimously agreed by the subsequent Delphi panel).

Phases II and IV relate to the attainment of a stable balloon guide catheter position in the internal carotid artery. In the original simulated run, steps within phases III and IV had to

J NeuroIntervent Surg: first published as 10.1136/neurintsurg-2018-014510 on 17 January 2019. Downloaded from http://jnis.bmj.com/ on 4 March 2019 by guest. Protected by copyright.

be performed sequentially such that if a step in phase IV was performed prior to the completion of any of the phase III steps an error was recorded. Experiences within the group using the simulator and Delphi discussions demonstrated that these two phases formed a continuum and that there are many occasions where performing a step from phase IV before the end of phase III would be reasonable, safe, and would not constitute an error. Variation in the order of steps in these phases was not penalized. Attempts to perform any steps from phase V prior to the completion of phase IV was judged to represent unacceptable deviation from the protocol.

Phase III, step 13, was further subdivided depending on the capabilities of the angiographic equipment. The majority of intracranial procedures are performed on biplane equipment but the use of single plane systems in MT is not uncommon.

Phase III, step 13a, pertained to units with biplane equipment. The frontal tube is centered on the aortic arch and the lateral tube is centered more cranially to visualize the common carotid artery bifurcation. The roadmap imaged acquired here allows the operator to assess the movements of the proximal catheter construct in the arch while visualizing the carotid bifurcation and distal wire position.

Step 13b would be employed if the procedure was being used on single plane equipment. The degree of x-ray field magnification was described in non-specific terms as large/small field of view (FOV). This was in response to the fact that different manufacturers of angiography equipment provide different FOVs. In addition, different operators have different preferences for different FOVs. As long as the choice of FOV did not compromise a safe technique, for instance selecting a FOV that precluded visualization of the distal wire tip position, no comment was made on FOV.

Study 1: Delphi consensus meeting

Participants

A panel of 21 experienced INRs (mean age 48 (SD 7) years; 18 men; all consultant level, including 10 professors), who performed a mean of 45 (SD 10) MT stroke cases per annum, from seven European and Scandinavian countries (Belgium (n=1), Denmark (n=4), Germany (n=12), Norway (n=1), Netherlands (n=1), Sweden (n=1) and UK (n=1)) convened in Aachen, Germany in May 2017.

Results study 1

At the start of the Delphi meeting, the project and concepts of 'proficiency based progression' were outlined and the procedure metrics for a reference approach to MT presented.

Each phase and step were discussed, and the proposed metrics were edited in real time such that a vote was taken on an agreed consensus statement. The most significant alteration to the proposed metrics was that the use of balloon occlusion guide catheters (BGCs) was mandated. Additional changes and edits were made in real time, and mainly concentrated on the precision of the language and operational definitions of procedure steps and errors. A summary of the changes that were made to the procedure steps and errors are reported in table 2.

At the end of each phase discussion a vote was taken to ensure majority consensus (table 2). All phases were passed with large majorities. The smallest consensus vote was for phase V, with 85% in favor of the characterization. There were zero votes against the characterization in phase V but there were three abstentions. No votes (ie, against the phases characterized) were only recorded for phases I and IV, and these were by a single individual.

Study 2: Construct validity

Participants

Eight consultant/professor level and 10 trainee INRs participated in this study. Mean age of the consultant/professor level INRs was 48 years (range 40–60 years) and for trainees, 36 years (range 32–40 years). Consultants were from Germany (n=5), the UK (n=2), and Denmark (n=1). Trainees were from the UK and Ireland (n=5), Germany (n=4), and The Netherlands (n=1). The consultant/professor level INRs completed on average 42 thrombectomy procedures per year (range 40–50). The trainees had on average been practicing for 15.8 (0–60) months, supervised cerebral angiograms 70 (0–200), independent cerebral angiograms 123 (0–300), assisted interventions 95 (5–300), first operator interventions 19 (0–80), assisted in thrombectomy cases 12 (0–40), and first operator thrombectomy 3 (0–10).

Apparatus

The vascular interventional simulation trainer (VIST, figure 1), described elsewhere, was used for the study.[15 16] The VIST virtual reality simulator utilizes a physics based high fidelity

Table 2 Summary of the changes agreed and voted on by the Delphi panel to the procedure steps and procedure errors of the reference approach to mechanical thrombectomy

	Steps		Errors		Voting			
Procedure phase	Before Delphi	After Delphi	Before Delphi	After Delphi	Yes	No	Abstained	% Majority
I	2	2	11	12	20	1	0	95
IIa	1	1	10	10	21	0	0	100
IIb	6	6	9	9				
III	6	6	12	12	21	0	0	100
IV	5	3	9	9	20	1	0	95
V	7	7	14	13	18	0	3	85
VI	2	2	12	9	18	0	2	90
VII	11	11	12	11	19	0	1	95
VIII	3	3	3	3	20	0	0	100
IX	4	4	1	1	19	0	1	95
X	2	2	1	1	20	0	0	100

J NeuroIntervent Surg: first published as 10.1136/neurintsurg-2018-014510 on 17 January 2019. Downloaded from http://jnis.bmj.com/ on 4 March 2019 by guest. Protected by copyright.



Figure 1 The vascular interventional simulation trainer (VIST) virtual reality simulator.

endovascular simulator which enables hands on procedural training for clinicians using real patient cases. This technology allows the use of the medical devices for simulated endovascular procedures. In replicated real world scenarios, it provides step by step guidance and metric based feedback throughout the procedure.[15 16] The haptic feedback to the users' actions are calculated in real time and depend on the interaction between the virtual devices and the virtual anatomy (vessel geometry and vessel properties), and give realistic tactile feedback to the operator during the training procedure.[15 16]

The metric based performance characterization and operational definitions of MT, as described in study 1, were used to establish the simulation and assessments.

Procedure

Neuroradiology trainees received a didactic explanation of acute stroke interventions. They then observed a simulated thrombectomy being performed by a faculty expert INR. The demonstration was performed with full audiovisual support, and the phases and steps of the procedure were verbally described in real time. Trainees then performed a reference thrombectomy case on the VIST physics based virtual reality simulation platform and enabled performance metrics agreed at the Delphi meeting (US Provisional Patent Application No 62667500). The case was performed in the presence of one of the INR faculty members. The faculty members assisted the delegates with the use of the simulator and preparation of the modified procedure apparatus required. No guidance regarding the performance of the thrombectomy procedure (order of steps, phases/catheter or wire manipulation/x-ray screening/table positioning, etc) was offered by the supervising faculty.

Statistical analysis

Performance differences were compared for statistical significance with Mann–Whitney U tests using SPSS statistical package (V.24).[17] Statistical power calculations were extrapolations based on previous research using the VIST simulator and performance metrics for carotid angiography.[16] The mean number of errors for the attending clinicians in the control arm of the study was 15.17 (SD 3.1). We estimated a 26–42% difference between the experienced and novice INR groups based on previous studies.[18] The statistical power of a 26% difference between the groups (ie, consultants INRs=15.17 vs novice MT INRs=19.11 was calculated for n=8 in the consultant group and n=10 in the novice group with an α of 5% and a β of 20%) was found to be 0.815 for a two tailed test. A difference of 42% between the groups (ie, consultants INRs=15.17 vs novice INRs=21.54) and using the same statistical power calculation methodology gave a statistical power of 0.996 for the same sample sizes.

RESULTS

The main performance parameters assessed and compared were (1) duration of the procedure in minutes, (2) number of procedure phases completed, (3) number of procedure steps completed, (4) number of handling errors made, (5) amount of contrast agent used during the procedure (in milliliters), and (6) amount of fluoroscopy (in minutes) used in performance of the procedure.

Figure 2A shows the median, 25th and 75th percentile ranks (PR) of the number of minutes it took each group to perform the procedure. Based on the median scores, the consultant INRs performed the procedure 19% faster than the trainees, and their performance times were also more homogeneous, as indicated by the smaller range of scores (consultant range=9.45 and trainee range=34.03). This difference was statistically significant (median=24 (25th PR=23 and 75th PR=28) vs trainee median=31 (25th PR=24 and 75th PR=50), Mann–Whitney U=17.0, Z=−2.04, P=0.043). A similar performance pattern was observed for procedure phases completed (figure 2B). The consultant INRs completed 40% more procedure phases and this difference was statistically significant (consultant median=7 (25th PR=6 and 75th PR=8) vs trainee mean=5 (25th PR=5 and 75th PR=6), Mann–Whitney U=11.5, Z=−2.61, P=0.009). Figure 2C shows that the consultant INRs also completed 20% more procedure steps and the difference was statistically significant (consultant median=34 (25th PR=32 and 75th PR=35) vs trainee median=28 (25th PR=25 and 75th PR=32), Mann–Whitney U=12.0, Z=−2.5, P=0.012). The largest difference between the two groups was observed for handling errors (figure 2D). The trainees made 42% more handling errors than the consultant INRs and this difference was statistically significant (consultant median=26 (25th PR=12 and 75th PR=29) vs trainee median=44 (25th PR=27 and 75th PR=60) Mann–Whitney U=13.5, Z=−2.36, P=0.016).

Although the trainees used 11% more contrast agent (consultant median=36 (25th PR=29 and 75th PR=39) vs trainee median=40 (25th PR=29 and 75th PR=77) this difference was not statistically significant (P=0.46). The consultant group also used 24% less fluoroscopy than the trainees but this difference was not statistically significant (consultant median=11 (25th PR=10 and 75th PR=12) vs trainee mean=15 (25th PR=11 and 75th PR=22).

Using the IQR scores of the consultant INR group, we calculated the variability range for each dependent variable score. Trainee time to complete the procedure demonstrated 3.6 times greater variability than the consultant INR group. A similar pattern was observed for phases completed (two times greater), procedure steps completed (1.7 times greater), handling errors made (2.235 times greater), contrast agent used during the procedure (seven times greater) and fluoroscopy time used during the procedure (3.7 times greater).

DISCUSSION

In the two studies reported here we developed the performance metrics for MT as proposed previously.[12–14 19] The metrics and their operational definitions were presented to the Delphi panel

J NeuroIntervent Surg: first published as 10.1136/neurintsurg-2018-014510 on 17 January 2019. Downloaded from http://jnis.bmj.com/ on 4 March 2019 by guest. Protected by copyright.

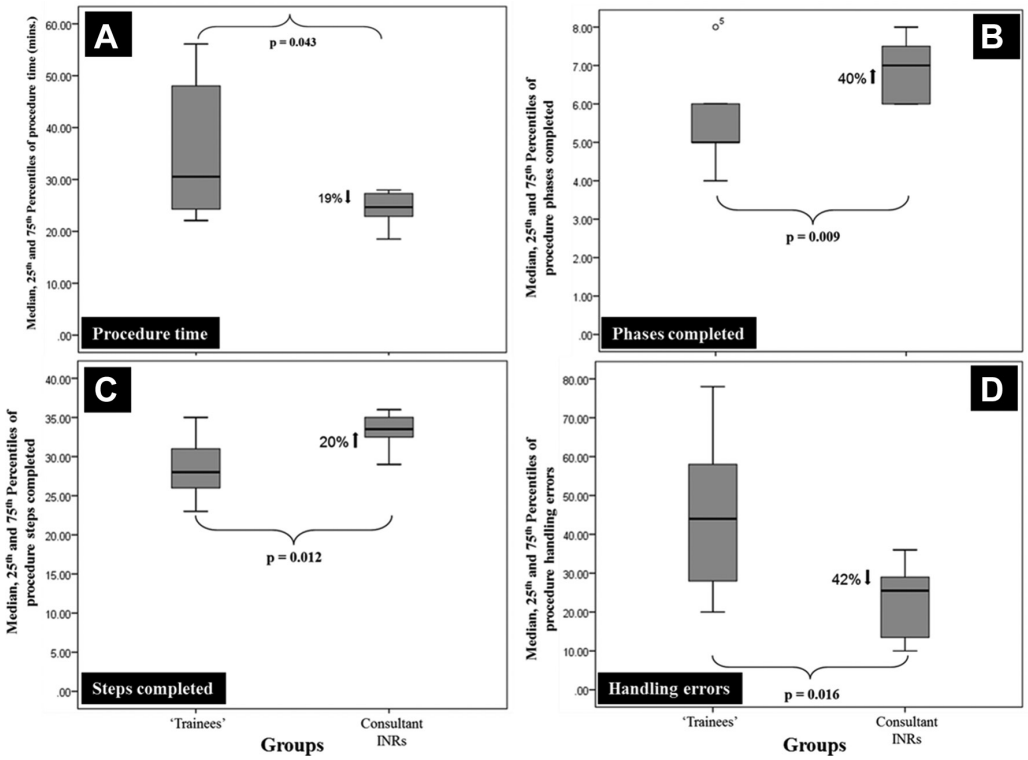


Figure 2 Median, 25th, and 75th rank scores of (A) performance time (in min), (B) procedure phases completed, (C) procedure steps completed, and (D) procedure handling errors made by trainees and consultant interventional neuroradiologists (INRs).

for their informed consideration on how well they characterized a safe and effective way of performing an MT by a trainee at the start of their learning curve. The only fundamental change to the method made at the Delphi meeting was the mandated use of proximal BGCs. Conversations centered on the fact that at the time of discussion, two of the main papers included in the HERMES collaboration[3] and the Solitaire With the Intention For Thrombectomy as Primary Endovascular Treatment (SWIFT PRIME)[20] and Extending the Time for Thrombolysis in Emergency Neurological Deficits-Intra-Arterial (EXTEND-IA)[21] mandated the use of BGCs (recommendation 11AHA guidelines 2015, class IIa; level of evidence C).[22]

In study 2, we sought to establish the construct validity[23] of the metrics agreed at the Delphi panel meeting and implemented in the VIST virtual reality simulation. In this study, we compared the computer based assessment of procedure experts and novices performing a straightforward MT procedure. The results showed that the performance metrics coupled to the VR simulation distinguished between the objectively scored performance of experienced practitioners and procedure novices. The experienced group completed significantly more phases and steps of the procedure in a shorter time frame. More importantly, they also made fewer objectively assessed procedure errors. The experienced group also used less contrast agent and fluoroscopy, but these differences were not statistically significant. This is almost certainly due to the large variability of scores in the novice group. Although these measures were not statistically significantly different, the measures were detecting performance differences in the two groups. Indeed, performance variability is a very good indicator of 'skill'. Individuals who are skilled at what they do perform better than less skilled individuals, but they also perform very homogeneously.[24]

Acute ischemic stroke is a leading cause of death and long term disability. MT is now the recommended treatment of choice for ischemic stroke due to large vessel occlusion in the anterior cerebral circulation. It has led to an improvement in outcomes in comparison with other treatments[3] with a reduction in long term disability.[25] Worldwide, there is a shortage of clinicians trained and skilled enough to perform the procedure, which is a significant impediment to the benefits that this treatment can confer on patients and healthcare systems.

In prospective randomized studies, virtual reality simulation in the surgical arena has been well validated as improving intraoperative performance of trainees.[9 12] Furthermore, it has been demonstrated that the benefits conferred with surgical simulation are optimized with metric based training to proficiency.[26] The metrics, derived from experienced and proficient practitioners, once validated are used to construct a curriculum, which uses metric based performance feedback to trainees. The metrics are also used to establish performance benchmarks (ie, proficiency levels), which trainees must unambiguously demonstrate before training progression. Additionally, trainees (no matter how senior) do not progress to performing the procedure on real patients until they have demonstrated that they 'know' how to do the procedure and can 'do it' to a quantitatively defined performance level. Furthermore, the performance level is not estimated, rather it is based on the average of the objectively assessed performance of INRs experienced in MT on the exact same virtual reality simulated procedure which trainees must perform. Prospective, randomized, and blinded clinical studies have demonstrated that metric based training to proficiency (ie, proficiency based progression or PBP) improves intraoperative performance of image guided surgical procedures[26–30] and an endovascular procedure by very experienced clinicians learning to perform a procedure that is novel to them.[16] Furthermore, there is also evidence that a PBP impacts on clinical outcomes. In a prospective, randomized, and blinded study of epidural analgesia for labor, the PBP trained anesthetists had a 54% lower

J NeuroIntervent Surg: first published as 10.1136/neurintsurg-2018-014510 on 17 January 2019. Downloaded from http://jnis.bmj.com/ on 4 March 2019 by guest. Protected by copyright.

epidural failure rate (13.3 vs 28.7) compared with the standard simulation based trained group.[31]

Virtual reality simulators for endovascular procedures are orders of magnitude superior to simulators that are used in image guided surgery.[12] They are however not much better than expensive video[26] games without procedure performance metrics. Angelo *et al*[26] demonstrated that simulation based training without metric based feedback and a requirement to demonstrate proficiency benchmarks conferred little benefit over traditional training.

Study limitations

The original simulation platform utilized air filled contrast syringes. All in the characterization group recognized that this detracts from the fidelity of the scenario. In tandem with this project, a closed pressure monitored fluid injection system has been developed which will be incorporated into any future work. This technical development will considerably improve the fidelity of the physics based VR MT simulation.

In this study, we only characterized a straightforward and reference approach to an MT procedure. We believed that if a trainee could not perform a straightforward procedure it was unlikely they would be competent to perform a more complex procedure. Furthermore, learning to perform MT using a 'standardized' or reference approach facilitates learning by giving trainees a safe and effective approach (which they can learn to perform to a quality assured performance level) from which they can hone the approach to the procedure that best suits them. There are other approaches to the performance of MT which were not characterized in this study. The methodology employed here can however be used to characterize and validate the metrics for these approaches in the same detail as reported here.

CONCLUSIONS

Study 1 characterized a referenced approach to MT and consensus was reached on the essential phases, steps, and procedure errors to be avoided. The metrics were then incorporated into a physics based VR endovascular simulation. The results in study 2 showed that the experienced INRs completed more of the case, faster, and with fewer errors. The results from these studies offer support for the face, content, and construct validity of the MT metrics. The next step in the validation of these VR enabled metrics is to use them as part of a systematic MT training program.

Contributors All of the authors contributed to the writing of the article. RC, TL, and MH are very experienced mechanical thrombectomy interventional neuroradiologists. LL and PH are very experienced clinicians and medical education specialists. JL is a senior computer engineer. AGG developed and helped validate proficiency based progression simulation training. TL, MH, RC, LL, JL, and AGG have characterised a reference approach to mechanical thrombectomy. JL developed and formatted the VR simulation from the CT angiography of the stroke patient reported in this paper. AGG and PH produced the first draft of this paper. All authors contributed to editing the paper post editorial review. Study 1, study design and data collection: RC, TL, MH, JL, LL, and AGG. Study 2, study design and data collection: RC, TL, MH, JL, PH, and AGG. Data analysis: RC, TL, MH, JL, and AGG. Paper writing: RC, TL, MH, JL, PH, and AGG. Results interpretation, paper critical revisions, and agreed on final draft: RC, TL, MH, JL, PH, LL, and AGG.

Funding The research and researchers on this paper were supported by a grant from the Swedish government agency for innovation (Vinnova) to Mentice AB (Gothenburg, Sweden) to characterize, develop, and then validate the metrics for a reference approach to performance of mechanical thrombectomy.

Competing interests MH has received honoraria from Microvention, Medtronic Neurovascular, Mentice AB, and Stryker Neurovascular for consulting and proctoring. RC has received honorarium for speaking (Stryker Neurovascular, UK) and educational sponsorship to attend meetings/conferences from Microvention, Stryker, Medtronic, Penumbra, and Johnson & Johnson. JL works as an engineer at Mentice and developed the VR model of the real patient data. LL has served as a clinical advisor and then Medical Director of Mentice.

Patient consent Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES

1 Nogueira RG, Jadhav AP, Haussen DC, *et al*. Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. *N Engl J Med* 2018;378:11–21.
2 Antman EM, Anbe DT, Armstrong PW, *et al*. ACC/AHA guidelines for the management of patients with ST-elevation myocardial infarction--executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (writing committee to revise the 1999 Guidelines for the Management of Patients With Acute Myocardial Infarction). *Circulation* 2004;110:588–636.
3 Goyal M, Menon BK, van Zwam WH, *et al*. Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. *Lancet* 2016;387:1723–31.
4 Gallagher AG, Smith CD. From the operating room of the present to the operating room of the future. Human-factors lessons learned from the minimally invasive surgery revolution. *Semin Laparosc Surg* 2003;10:127–39.
5 Institute of Medicine. *Graduate medical education that meets the nation's health needs*. Washington, DC: Institute of Medicine, 2014.
6 Imrie K. *Training 21st century clinical leaders*. Dublin, Ireland: Royal College of Physicians in Ireland, 2014.
7 Satava RM. Virtual reality surgical simulator. The first steps. *Surg Endosc* 1993;7:203–5.
8 Pellegrini CA. Surgical education in the United States: navigating the white waters. *Ann Surg* 2006;244:335–42.
9 Zendejas B, Brydges R, Hamstra SJ, *et al*. State of the evidence on simulation-based training for laparoscopic surgery: a systematic review. *Ann Surg* 2013;257:586–93.
10 Gallagher AG, Satava RM. Surgical simulation: seeing the bigger picture and asking the right questions. *Ann Surg* 2015;262:e50–1.
11 Institute of Medicine. *Graduate medical education that meets the nation's health needs*: The National Academies Press, Washington DC., 2014.
12 Gallagher AG, O'Sullivan GC. *Fundamentals of surgical simulation; principles & practices*. London: Springer Verlag, 2011.
13 Gallagher AG, Ritter EM, Champion H, *et al*. Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. *Ann Surg* 2005;241:364-72.
14 Gallagher AG. Metric-based simulation training to proficiency in medical education:- what it is and how to do it. *Ulster Med J* 2012;81:107–13.
15. Liebig T, Holtmannspötter M, Crossley R, *et al*. Metric-based virtual reality simulation: A paradigm shift in training for mechanical thrombectomy in acute stroke. *Stroke* 2018;49:e239–e242.
16 Cates CU, Lönn L, Gallagher AG. Prospective, randomised and blinded comparison of proficiency-based progression full-physics virtual reality simulator training versus invasive vascular experience for learning carotid artery angiography by very experienced operators. *BMJ Simulation and Technology Enhanced Learning* 2016;2:1–5.
17 *IBM SPSS Statistics [program]. 24 version*. New York: IBM SPSS, 2016.
18 Gallagher AG, Seymour NE, Jordan-Black JA, *et al*. Prospective, randomized assessment of transfer of training (ToT) and transfer effectiveness ratio (TER) of virtual reality simulation training for laparoscopic skill acquisition. *Ann Surg* 2013;257:1025–31.
19 Angelo RL, Ryu RK, Pedowitz RA, *et al*. Metric development for an arthroscopic bankart procedure: assessment of face and content validity. *Arthroscopy* 2015;31:1430–40.
20 Jovin TG, Chamorro A, Cobo E, *et al*. Thrombectomy within 8 hours after symptom onset in ischemic stroke. *N Engl J Med* 2015;372:2296–306.
21 Campbell BC, Mitchell PJ, Kleinig TJ, *et al*. Endovascular therapy for ischemic stroke with perfusion-imaging selection. *N Engl J Med* 2015;372:1009–18.
22 Powers WJ, Derdeyn CP, Biller J, *et al*. 2015 American Heart Association/American Stroke Association focused update of the 2013 guidelines for the early management of patients with acute ischemic stroke regarding endovascular treatment: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 2015;46:3020–35.
23 Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. *Surg Endosc* 2003;17:1525–9.

J NeuroIntervent Surg: first published as 10.1136/neurintsurg-2018-014510 on 17 January 2019. Downloaded from http://jnis.bmj.com/ on 4 March 2019 by guest. Protected by copyright.

24 Gallagher AG, Satava RM. Virtual reality as a metric for the assessment of laparoscopic psychomotor skills. Learning curves and reliability measures. *Surg Endosc* 2002;16:1746–52.
25 Albers GW, Marks MP, Kemp S, *et al*. Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging. *N Engl J Med* 2018;378:708–18.
26 Angelo RL, Ryu RK, Pedowitz RA, *et al*. A proficiency-based progression training curriculum coupled with a model simulator results in the acquisition of a superior arthroscopic bankart skill set. *Arthroscopy* 2015;31:1854–71.
27 Van Sickle KR, Ritter EM, Baghai M, *et al*. Prospective, randomized, double-blind trial of curriculum-based training for intracorporeal suturing and knot tying. *J Am Coll Surg* 2008;207:560–8.
28 Ahlberg G, Enochsson L, Gallagher AG, *et al*. Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *Am J Surg* 2007;193:797–804.
29 McClusky DA, Gallagher AG, Ritter EM, *et al*. Virtual reality training improves junior residents' operating room performance: Results of a prospective, randomized, double-blinded study of the complete laparoscopic cholecystectomy. *J Am Coll Surg* 2004;199:73.
30 Seymour NE, Gallagher AG, Roman SA, *et al*. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann Surg* 2002;236:458–63.
31 Kallidaikurichi Srinivasan K, Gallagher A, O'Brien N, *et al*. Proficiency-based progression training: an 'end to end' model for decreasing error applied to achievement of effective epidural analgesia during labour: a randomised control study. *BMJ Open* 2018;8:e020099.

J NeuroIntervent Surg: first published as 10.1136/neurintsurg-2018-014510 on 17 January 2019. Downloaded from http://jnis.bmj.com/ on 4 March 2019 by guest. Protected by copyright.