Eur J Vasc Endovasc Surg (2017) 54, 34–41

Assessment of Competence in EVAR Procedures: A Novel Rating Scale Developed by the Delphi Technique

M. Strøm [a,b,c,*], L. Lönn [c,d], B. Bech [d], T.V. Schroeder [a,c], L. Konge [a,c], on behalf of the "EVARATE Delphi Panel"

[a] Copenhagen Academy for Medical Education and Simulation, The Capital Region of Denmark, Rigshospitalet, Copenhagen, Denmark
[b] Department of Vascular Surgery, Rigshospitalet, Copenhagen, Denmark
[c] University of Copenhagen, Copenhagen, Denmark
[d] Department of Radiology, Rigshospitalet, Copenhagen, Denmark

WHAT THIS PAPER ADDS

This paper presents the first procedure specific tool for assessing competence in endovascular aortic repair. A Delphi approach was used to create the tool based on consensus among 32 international experts from ten countries. This tool can be used to provide structured feedback to trainees.

Objective/Background: To develop a procedure specific global rating scale for assessment of operator competence in endovascular aortic repair (EVAR).

Methods: A Delphi approach was used to achieve expert consensus. A panel of 32 international experts (median 300 EVAR procedures, range 200–3000) from vascular surgery ($n = 21$) and radiology ($n = 11$) was established. The first Delphi round was based on a review of endovascular skills assessment papers, stent graft instructions for use, and structured interviews. It led to a primary pool of 83 items that were formulated as global rating scale items with tentative anchors. Iterative Delphi rounds were executed. The panellists rated the importance of each item on a 5 point Likert scale. Consensus was defined as 80% of the panel rating an item 4 or 5 in the primary round and 90% in subsequent rounds. Consensus on the final assessment tool was defined as Cronbach's alpha $>$ .8 after a minimum of three rounds.

Results: Thirty-two of 35 invited experts participated. Three rounds of surveys were completed with a completion rate of 100% in the first two rounds and 91% in round three. The 83 primary assessment items were supplemented with five items suggested by the panel and reduced to seven pivotal assessment items that reached consensus, Cronbach's alpha = 0.82. The seven item rating scale covers key elements of competence in EVAR stent placement and deployment. Each item has well defined grades with explicit anchors at unacceptable, acceptable, and superior performance on a 5 point Likert scale.

Conclusion: The Delphi methodology allowed for international consensus on a new procedure specific global rating scale for assessment of competence in EVAR. The resulting scale, EndoVascular Aortic Repair Assessment of Technical Expertise (EVARATE), represents key elements in the procedure. EVARATE constitutes an assessment tool for providing structured feedback to endovascular operators in training.

Article history: Received 25 November 2016, Accepted 5 April 2017, Available online 23 May 2017

Keywords: Abdominal aortic aneurysm, Clinical competence, Delphi technique, Educational assessment, Endovascular procedures, EVAR

INTRODUCTION

Endovascular aortic repair (EVAR) aims to exclude an abdominal aortic aneurysm (AAA) by implanting a modular stent graft system in the aorta. This involves several more or less complex procedural steps reflected in a shallow learning curve and outcome depending on operator experience.[1–3]

Competency based education aims to ensure that the required skills in relation to an objectified standard are met.[4,5] The emphasis is to demonstrate skilful application of knowledge.[6,7] Trainees learn at different paces and a certain number of procedures do not ensure competency.[8] Evaluation of skills should be performed with reliable assessment tools and structured feedback to increase learning and performance of trainees.[9,10]

Several assessment tools have been developed for task specific endovascular procedures.[9,11–29] However, these must be domain specific with regard to anatomy and

* Corresponding author. Copenhagen Academy for Medical Education and Simulation, Rigshospitalet 5404, Blegdamsvej 9, 2100 Copenhagen, Denmark.

E-mail address: michael.stroem@regionh.dk (M. Strøm).

1078-5884/

http://dx.doi.org/10.1016/j.ejvs.2017.04.001

endovascular instruments.[30] No validated assessment tool of specific competence in EVAR exists and available rating scales do not encompass the procedure.[31]

The aim of this study was to develop an assessment tool of EVAR competence, that is, the EndoVascular Aortic Repair Assessment of Technical Expertise (EVARATE).

METHODS

A modified Delphi approach to obtain expert consensus on the content of an assessment of competence in infrarenal EVAR was employed. The Delphi technique uses a blinded, structured, and iterative process to generate expert opinion where experts evaluate and re-evaluate items of interest until consensus is met.[32,33] In its original form even the initial questions are developed by the panel. A modified Delphi approach was applied developing the items for the first round based on literature review and expert interviews.

The study was submitted to the Ethical Committee of the Capital Region of Denmark and ethical approval was waived (FSP-15003010).

Review of literature for the initial framework

A scoping review in PubMed was conducted in May 2015 with the following string including asterisks and Boolean operators: "("endovascular surg*") AND (rating* OR scale* OR Assessment* OR Skill*) AND (simulat* OR training OR education OR technical OR psychomotor OR Competence)". The search design encompassed publications within endovascular assessment and training. Inclusion criteria were procedural steps in EVAR and ratings scales in endovascular surgery. Exclusion criteria were publications without disclosure of an assessment tool for endovascular competencies or no disclosure of the key elements in endovascular procedural steps. Furthermore, instructions for use (IFU) from infrarenal stent grafts were reviewed regarding procedural key points until data saturation was achieved. Therefore, IFU from four companies were included.[34–37]

The literature search yielded 852 publications. Thirty papers were identified from abstract review and 20 full length papers fulfilled the inclusion criteria.[9,11–29] In all, 163 previously validated assessment items (126 global rating scale [GRS] items, 37 checklist items) were extracted from the literature review, while 77 were omitted owing to definition as procedure specific, for example use of carotid artery embolism protection device or additional renal artery intervention.

Expert opinions for the initial framework

Structured interviews with experts ($n = 3$, number of EVAR procedures 350–1000 as primary operator) were conducted on procedural steps in EVAR. Experts were identified as current practising endovascular operators with long-standing experience in general endovascular procedures and EVAR with a publication record in this context (31–160 published papers, 4–30 concerning EVAR). The interviews were recorded, transcribed, and progressively added to a preliminary item pool. Interviews continued until the pool was saturated, that is, no new items added additional information. Following each interview the items were grouped and redundant items removed. Saturation was met after two rounds, that is, no new items were added after the third interview. The structured interviews yielded 67 specific checklist items and 76 GRS items.

Development of the initial assessment tool

The structured interview items were grouped in overall themes. The pool was saturated with additional information from the scoping review and the IFUs and finally phrased as GRS items.[38] GRSs typically ask raters to assess participants' overall performance on subtasks, whereas checklists prompt raters to binary attest to the performance or neglect of directly observable actions.[38]

Domains outside endovascular aortic interventions were omitted. Further, three groups of items were omitted at the discretion of the research team: (i) pre-operative "planning" at the workstation, (ii) "closing of puncture sites"; and (iii) "complications" that were not intricate results of the stent graft implantations, such as, for example, type-one leak or inadvertently over-stenting. Handling of complications and closure of puncture sites were not included in the assessment tool to ensure that all items were applicable to all procedures (i.e., a perfectly performed procedure without any complications can achieve maximum points).

The initial item bank for the first Delphi round resulted in 83 items with assigned tentative anchors at rating 1, 3, and 5 on a 5 point Likert scale.

Delphi panel recruitment

A panel of international voluntary experts was established by email invitation and informed consent obtained. Purposive sampling was used to ensure wide international representation. Expert level was defined as currently active EVAR operator having planned, sized, and performed a minimum of 200 EVAR procedures as primary operator, working in a teaching hospital or university, and having an endovascular publication record (range 2–249, median 53). Although consensus on the sample size of a Delphi panel is absent,[39] there is a general recommendation of 15–30 participants for a homogeneous (e.g., few overlapping specialties) survey.[32,33,40] Aiming for a response rate of >80%, 35 physicians from 10 countries were invited to participate. Membership of the Delphi panel was kept confidential during the study.

Executing the Delphi study

The Delphi technique was designed to obtain expert consensus by iterative rounds of progressive blinded feedback.[41] The study was executed using an online survey format (Surveyexact.com, version 7.1; Rambøll Management Consulting, Aarhus, Denmark) from November 2015 to June 2016. Each Delphi round started with a standardised email, including a unique link to the survey with a mandatory instruction for raters. Prospective validation of

the online survey ensured no missing data. Up to four reminders were sent to non-responders.

The panellists were instructed to rate each item independently on relevance for a rating scale for EVAR competence on a Likert scale, from 1 ("should *definitely not* be part of the final assessment tool") to 3 ("neutral") to 5 ("should *definitely* be part of the final assessment tool").

The panellists were allowed to comment and suggest rephrasing of all items and tentative anchors. Furthermore, the Delphi panel was encouraged to formulate new items of important aspects. After each round, the research team evaluated the agreed items. Each panellist's evaluation was given same weight.

Consensus was a priori defined as 80% of participants scoring 4 or 5 on an item in round one and 90% in sequent rounds. Items where a priori consensus was not met were removed from the rating scale. Redundancies were resolved when more than two panellists suggested merging of items. New items proposed by the panellists were grouped and rephrased for clarity and uniformity.

The revised rating scale was reissued and the process repeated until consensus was reached.

A priori consensus on items was defined as Cronbach's alpha $>$ .8 after a minimum of three Delphi rounds.

Statistics

Cronbach's alpha was used to determine the internal consistency of the assessment tool after each round. The Mann–Whitney *U*-test was used to compare differences in scoring of items between groups.

A p value $<$ .05 was considered statistical significant. Pairwise p values were adjusted for multiple comparisons using simple Bonferroni correction. SPSS 23.0 (IBM, Armonk, NY, USA) was used for statistical analysis.

RESULTS

Thirty-five experts were contacted of which 32 accepted and completed the first online survey. The panel comprised of vascular surgeons ($n = 21$) and interventional radiologists ($n = 11$) representing two continents, 10 countries, and 22 institutions (Table 1). In all, 32 (100%) completed the first two rounds and 29 (91%) completed round three.

The first Delphi round resulted in 39 items (Table 2). As suggested by the panel 18 items were merged into three to avoid redundancy. Furthermore, five new items suggested by the panel were added. The second round rated 29 items and consensus was reached on 14 items with a Cronbach's alpha of .84. In the third round consensus was reached on seven pivotal items (Cronbach's alpha = 0.82). The anchors were prospectively refined as suggested by the panel and finally determined by the research group. There were no missing data. No significant differences in scoring of individual items were found between radiologists and vascular surgeons in any round using the Mann–Whitney *U*-test. The Delphi process resulted in the EVARATE rating scale (Figs. 1 and 2).

Table 1. Participant demographics.

Mean (range) age (y)	50 (40–69)
Mean (range) no. years as specialist	16 (5–30)
Male, *n*	29 (91%)
Specialty, *n*	
Radiology	11 (34%)
Vascular surgery	21 (66%)
EVAR procedures (*n*)	
200–399	18
400–999	10
1000–3000	4
Country	
Belgium	2
Denmark	4
Finland	2
Germany	2
The Netherlands	1
Norway	1
Sweden	13
Switzerland	1
UK	3
USA	3

DISCUSSION

Development of the novel rating scale, EVARATE

This study investigated fundamentals of competence within the EVAR context and a final rating scale was developed. The EVARATE rating scale targets the access vessels, the deployment of the modular stent graft, and quality control of the implantation, and comprises pivotal elements describing key points in the procedure. Several assessment scales have been developed for specific endovascular procedures.[9,11–29] The resulting final seven- item rating scale is in line with previously published endovascular assessment tools.[11,12,20,23–27] The low number of items is in contrast to other endovascular rating scales, which have encompassed pre- and post-planning and a more vigorous part task approach.[13,18,22,28,42] However, from the literature review the previously proposed assessment items were incorporated for the first Delphi round and successively eliminated, thus warranting the resulting number. The rating scale aimed at EVAR specific actions. However, the individual items encompassed several implied procedural steps and lead to an overall assessment of part task actions, emphasized in the anchors.

Structure of the assessment items

Assessments of technical skills by external raters must be based on an objectified standard.[43] Checklists and GRSs are often used to objectify the assessment. Checklists prompt raters to assess the performance directly on observable actions using binary items, whereas GRSs typically ask raters to assess participants' overall performance on part tasks. Compared with checklists, GRS have been shown to have higher reliability and may better capture intrinsic elements of expertise.[38] While checklists can be used for structured feedback, they offer the risk of the trainee performing the task poorly; however, scoring points for completing the

Table 2. The result of the first Delphi round.

	Item title	Average (SD)	Consensus (%)
1	Demonstrates an ability to project renals	4.4 (0.6)	96.9
2	Demonstrates an ability to adjust and use C-arm and table appropriately	4.2 (0.6)	90.6
3	Contrast use	3.9 (0.5)	84.4
4	Demonstrates optimal image when releasing main graft	4.4 (0.6)	96.9
5	Demonstrates correct use of fluoroscopy and good overview of manipulated instruments	3.9 (0.6)	81.3
6	Can select appropriate instruments	4.1 (0.5)	93.8
7	Demonstrates correct preparation and safe handling of guidewires, catheters, and sheaths	4.2 (0.9)	81.3
8	Demonstrates maintenance of wire stability	4.3 (0.7)	87.5
9	Demonstrates awareness of tips of guidewires	4.2 (0.7)	84.4
10	Demonstrates knowledge of instruments	4.2 (0.7)	84.4
11	Able to perform angiogram	4.1 (0.9)	81.3
12	Demonstrates respect for tissue	4.2 (0.7)	81.3
13	Demonstrates respect for the puncture sites	4.3 (0.7)	87.5
14	Demonstrates correct sequencing of the procedure	4.3 (0.8)	81.3
15	Awareness of haemodynamics	4.1 (0.7)	81.3
16	Aortic angiogram	3.9 (0.5)	84.4
17	Overall endovascular technique	4.4 (0.6)	93.8
18	Can name relevant anatomical structures during the procedure	4.2 (0.9)	81.3
19	Can evaluate the angiographic images properly	4.6 (0.5)	100
20	Able to advance and secure a stiff guidewire over catheter in aorta before advancement of main body	4.3 (0.7)	84.4
21	Ensures that graft has not turned and realign if needed	4.3 (0.6)	90.6
22	Ensures that contralateral leg is over bifurcation	4.4 (0.9)	90.6
23	Releases main graft securely and in correct place	4.8 (0.4)	100
24	Demonstrates knowledge of deploying and docking top cap	4.4 (0.8)	84.4
25	Uses whole neck length for proximal sealing	4.3 (0.9)	87.5
26	Ensures that guidewire is in right lumen at contralateral gate and proximal neck	4.8 (0.4)	100
27	Securely deploys extension limbs without covering the internal iliacs	4.3 (0.6)	90.6
28	Evaluates leaks	4.3 (0.6)	93.8
29	Ensures renal arteries are patent	4.8 (0.4)	100
30	Ensures that the extension legs do not kink	4.2 (0.5)	93.8
31	Able to perform complete post insertion angiogram	4.4 (0.7)	90.6
32	Removes the instruments safely	4.1 (0.9)	84.4
33	Quality of final product	4.3 (0.8)	96.9
34	Overall EVAR deployment ability	4.4 (0.6)	93.8
35	Overall communications skill	4.1 (0.7)	81.3
36	Ensures sterility	4.2 (0.9)	84.4
37	Demonstrates how the graft system/handles work	4.4 (0.7)	87.5
38	Can complete the case	4.4 (0.6)	93.8
39	Overall procedure assessment	4.3 (0.6)	93.8

individual procedural steps. GRSs are considered more sensitive to level of expertise[10]; hence, the use of detailed checklists may not provide better assessment of skills than a more general scale that evaluates overall skills intrinsic to the procedure.[38,40]

Some substeps conceived important for the procedure such as “Negotiating tortuous vessels”, “post-dilatation with moulding balloon in proximal and distal sealing zones”, or “demonstrate skill to ensure radiation hygiene” were voted out after the first iteration. Those skills are so intrinsic to other skills and hence were found redundant by the panel. It is also noteworthy that “classical” GRS such as “overall assessment” (deleted in round two) and “flow of operation” (deleted in round three) were omitted from the final scale. Of the five items suggested by the panel only one regarding stability of the main graft was carried into the final scale (item 3). The Delphi panellists voted out items that should be considered fundamental and prerequisite before being allowed to perform EVAR. In a recent study by Maertens et al.,[42] 90% of proposed endovascular skills were considered fundamental by a group of experts. This result is in agreement with the development of EVARATE showing the distinction between fundamental skills and advanced procedure specific skills.

The Delphi approach

The Delphi approach is accepted as an effective method of obtaining consensus drawing on the combined intelligence of

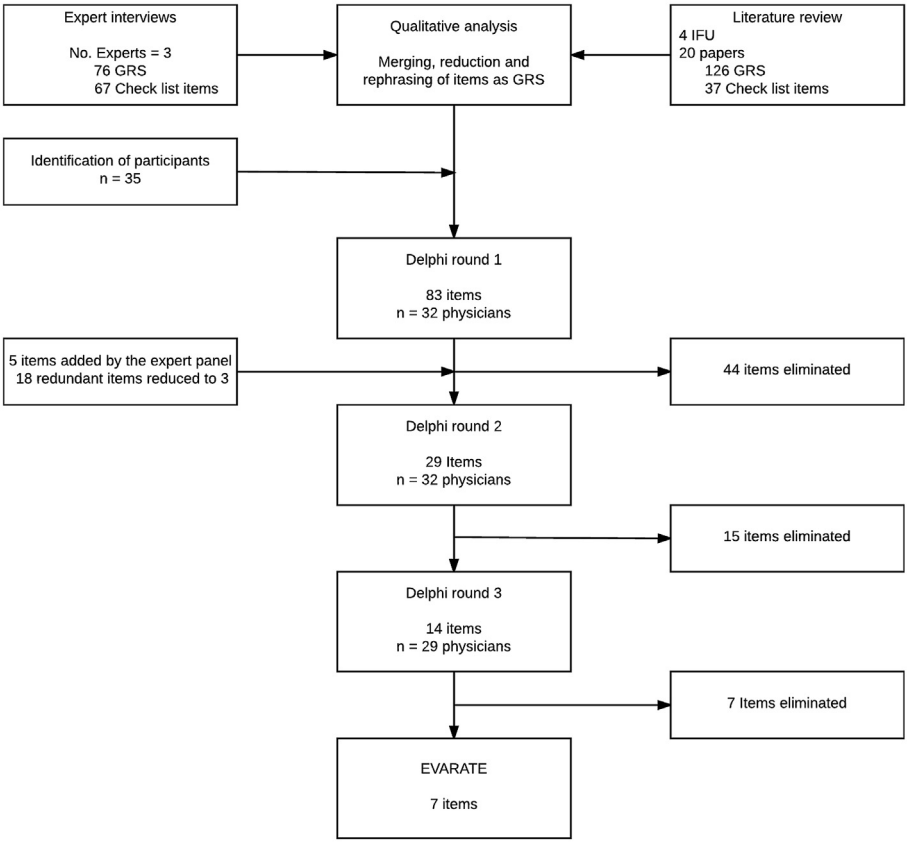

Figure 1. Flowchart describing the development of "EndoVascular Aortic Repair Assessment of Technical Expertise" (EVARATE). *Note.* GRS = global rating scale; IFU = instructions for use.

an expert panel and has recently been used to determine international consensus on fundamental endovascular skills.[42] The method has been criticized because the steering committee to some extent controls the response process.[41] Previous studies have countered this either by letting the expert panel develop the items for the first iteration,[44] or by developing very elaborate lists of substeps for evaluation.[39,42] Cumbersome rounds with several items to be rated have led to low response rates.[39] The method of letting the panel develop the items for the first round could be conceived problematic for a complex procedure with the risk of the participants getting weary early in the process leading to an early high dropout rate. The item pool for the first round of this study was very elaborate and substeps were added until saturation was met. However, items were presented in several overall themes for clarity and all communication was personalised resulting in a very high response rate.

An analogue approach to the one used in this study has previously been used in colonoscopy also producing a 7 point global rating scale.[10,45] While the previously mentioned studies had more participants (n = 55 and n = 44, respectively), they also had higher dropout rates.

Panel size

Previous studies have used 5—10 participants per professional group and 15—30 participants has been recommended for a homogeneous (e.g., few overlapping specialties) survey.[32,33,40] Sample sizes of more than 30 have been shown to add little to the results, while being difficult to maintain following low response rates.[32] Thus, the Delphi panel is a representative group in terms of group size and subgroup composition, and in alignment with current recommendations.

The response rate in this study was acceptable compared with similar studies.[10,39—42,44,46] The high response rate, with no loss of participants in round two could be attributed to the number of items to be rated in each round being substantially reduced.

LIMITATIONS

Although the sample size is in line with current recommendations for a Delphi study,[32,33,40] and the response rate was very good, the dropout rate could influence the results. Cronbach's alpha is sensitive to the number of participants, and the loss of three participants in the final round is reflected in the decrease in Cronbach's alpha from .84 to .82. Furthermore, there is no agreed standard in the literature to establish consensus. The predetermined method commonly used in comparative reports was chosen and other modes were not incorporated post hoc.

Participants included relevant specialties and a broad representation from Europe and North America. However, a possible limitation was to not include other continents. Many participants were Swedish. However, this had no

EndoVascular Aortic Repair Assessment of Technical Expertise – EVARATE

Study: **ID number:** **Date:** **Assessor:**

	1 Unacceptable	2	3 Acceptable	4	5 Superior
1. **Respect for the puncture sites and the access vessel**	Introduces and advances delivery devices with unacceptable force. Does not secure the sheaths when exchanging instruments or the sheaths travels.		Introduces and advances delivery devices with acceptable care and force. Secures the sheaths when exchanging instruments.		Introduces and advances delivery devices with optimal care. Keeps sheaths in place at all times.
2. **Demonstrates ability to accurately and safely deploy the top-stent of the stent graft system**	Deploys top-stent at clearly wrong place, wrong orientation or unsafe.		Deploys the top-stent safely in acceptable position and orientation. Some time consuming.		Deploys the top-stent safely and accurately in optimal position and orientation.
3. **Demonstrates ability to release the main graft securely and accurately**	Displaces or rotates the graft while releasing.		Releases the main graft stable, and securely in acceptable orientation using some effort.		Confidently releases the main graft stable and securely in optimal position and orientation.
4. **Demonstrates optimal fluoroscopy view when releasing the main graft**	Poor visualization of the proximal sealing zone. Does not correct for parallax. Does not have renals, bifurcation, or contra-leg tick-marks in the image.		Acceptable visualization of the proximal sealing zone. Projects renals and corrects for parallax. Has contralateral-leg tick-marks and the aortic bifurcation in the image		Superior and fluently visualization of the proximal sealing zone with optimal projection of lowest renal and optimal correction for parallax. Image centred and magnified on renals and with tick-marks on the contralateral-leg and the aortic bifurcation in the image.
5. **Catheterization of the contralateral limb of the main graft**	Cannot cannulate and/or does not ensure the right lumen in the main body.		Ensures the entry-side/gate in two projections upon catheterization and ensures the right lumen in the main body with pigtail. Secures safe placement of a stiff wire. Some time consuming.		Cannulates confidently. Ensures the entry-side/gate in two projections upon catheterization and ensures the right lumen in the main body with pigtail. Secures safe placement of a stiff wire with superior technique.
6. **Demonstrates ability to deploy the extension limbs**	Covers one or both internal iliac arteries and/or extension limbs kinks without correction. Poor visualization. Awkward movements and/or time consuming.		Deploys extension limbs correctly under acceptable visualization. Corrects kinks of extension limbs. Some time consuming.		Secure deployment of the extension limbs at optimal distal landing zones with optimal visualization. Corrects kinks of extension limbs.
7. **Able to perform and analyse a completion angiogram**	Insufficient visualization of renal and visceral arteries, proximal and distal sealing zones, overlapping stent grafts, and insufficient check for endo-leaks.		Acceptable visualization of renal and visceral arteries, proximal and distal sealing zones, overlapping stent grafts, and acceptable check for endo-leaks.		Superior visualization of renal and visceral arteries, proximal and distal sealing zones, overlapping stent grafts, and superior check for endo-leaks.

Procedure time: **Total score:**

Figure 2. The final assessment tool: "EndoVascular Aortic Repair Assessment of Technical Expertise" (EVARATE).

significant impact on the items scores. Overall, the EVERATE scale was developed through a stringent set-up ensuring content validity and a controlled response process.

FUTURE PERSPECTIVES

The long-term objective of this project is to develop an assessment tool with good evidence of validity that can be used in credentialing future EVAR operators. Further evidence according to the contemporary framework for validity should be examined by applying the novel rating scale to the EVAR procedure, in simulated and clinical settings.[43] A validated assessment tool would allow training to a predefined standard in a simulated environment countering the shallow learning curve for the procedure using structured feedback. The vision is an earlier and safer access to advanced clinical procedures. An assessment tool that deconstructs the key issues of a complex procedure allows the trainee to identify the critical steps of the procedure in the correct order and thereby better comprehend each procedural substep.[39,47] Structured and formative feedback using procedure specific GRSs have been shown to improve the performance of trainees.[4,48] EVARATE can be a platform for individualised formative feedback to trainees for the experienced supervisor involved in the training and education of future operators. EVARATE can be used for both full procedure and part- task training based on essential focus points.

CONCLUSION

The Delphi methodology allowed for the determination of international two speciality expert consensus on the content of a new procedure specific GRS for assessment of competence in EVAR. The resulting scale, EVARATE, is representing key elements in the procedure and draws on international consensus between radiologists and vascular surgeons.

A specific assessment tool is essential to ensure valid assessment of physicians for this complex, high stakes procedure. EVARATE could be used for simulation based mastery learning and for providing structured feedback to trainees both in a simulation based and in a clinical context.

ACKNOWLEDGEMENTS

The authors sincerely acknowledge the participation of the members of the EVARATE Delphi panel: Pekko Aho, Martin Back, Colin Bicknell, Katarina Björses, Jan Brunkwall, Michael Dake, Nuno Dias, Eric Dorenberg, Stevo Duvnjak, Mårten Falkenberg, Johan Formgren, Jan Holst, Thorarinn Kristmundsson, Sten Langfeldt, Hans Lindgren, Hossain Mafi, Martin Malina, Kevin Mani, Bijan Modarai, Robert Morgan, Niklas Nyman, Håkan Pärsson, John Rasmussen, Timothy Resch, Murray Shames, Josua C. Van Den Berg, Isabelle Van Herzeele, Hence Verhagen, Eric Verhoeven, Maarit Venermo, Frank Vermassen, and Anders Wanhainen.

CONFLICTS OF INTEREST

L.L. is a consultant for Mentice AB.

FUNDING

None.

REFERENCES

1 Walker TG, Kalva SP, Yeddula K, Wicky S, Kundu S, Drescher P, et al. Clinical practice guidelines for endovascular abdominal aortic aneurysm repair: written by the Standards of Practice Committee for the Society of Interventional Radiology and Endorsed by the Cardiovascular and Interventional Radiological Society of Europe. *J Vasc Interv Radiol* 2010;**21**(11):1632—55.

2 Lobato AC, Rodriguez-Lopez J, Diethrich EB. Learning curve for endovascular abdominal aortic aneurysm repair: evaluation of a 277-patient single-center experience. *J Endovasc Ther* 2002;**9**(3):262—8.

3 Laheij RJF, van Marrewijk CJ, Buth J, Harris PL. The influence of team experience on outcomes of endovascular stenting of abdominal aortic aneurysms. *Eur J Vasc Endovasc Surg* 2002;**24**(2):128—33.

4 Aggarwal R, Grantcharov T, Moorthy K, Milland T, Darzi A. Toward feasible, valid, and reliable video-based assessments of technical surgical skills in the operating room. *Ann Surg* 2008;**247**(2):372—9.

5 ten Cate O, Scheele F. Viewpoint: competency-based postgraduate training: can we bridge the gap between theory and clinical practice? *Acad Med* 2007;**82**(6):542—7.

6 McGaghie WC. Mastery learning. *Acad Med* 2015;**90**(11):1438—41.

7 Carraccio C, Englander R, Van Melle E, ten Cate O, Lockyer J, Chan M-K, et al. Advancing competency-based medical education. *Acad Med* 2016;**91**(5):645—9.

8 Gallagher AG, Ritter EM, Champion H, Higgins G, Fried MP, Moses G, et al. Virtual reality simulation for the operating room. *Ann Surg* 2005;**241**(2):364—72.

9 Boyle E, O'Keeffe DA, Naughton PA, Hill ADK, McDonnell CO, Moneley D. The importance of expert feedback during endovascular simulator training. *J Vasc Surg* 2011;**54**(1):240—248.e1.

10 Walsh CM, Ling SC, Khanna N, Cooper MA, Grover SC, May G, et al. Gastrointestinal endoscopy competency assessment tool: development of a procedure-specific assessment tool for colonoscopy. *Gastrointest Endosc* 2014;**79**(5):798—807.e5.

11 Lee JT, Qiu M, Teshome M, Raghavan SS, Tedesco MM, Dalman RL. The utility of endovascular simulation to improve technical performance and stimulate continued interest of preclinical medical students in vascular surgery. *J Surg Educ* 2009;**66**(6):367—73.

12 Rudarakanchana N, Desender L, Van Herzeele I, Cheshire NJ. Virtual reality simulation for the optimization of endovascular procedures: current perspectives. *Vasc Health Risk Manag* 2015;**11**:195.

13 Räder SBEW, Abildgaard U, Jørgensen E, Bech B, Lönn L, Ringsted CV. Association between endovascular performance in a simulated setting and in the catheterization laboratory. *Simul Healthc* 2014;**9**(4):241—8.

14 Rosenthal R, Mujagic E, Jacob AL, Seelos R, Schäfer J, Gürke L. Impact of an intensive 2-day endovascular training course on technical performance of trainees. *Ann Vasc Surg* 2013;**27**(8):1173—81.

15 Weisz G, Smilowitz NR, Parise H, Devaud J, Moussa I, Ramee S, et al. Objective simulator-based evaluation of carotid artery stenting proficiency (from Assessment of Operator

Performance by the Carotid Stenting Simulator Study [ASSESS]). *Am J Cardiol* 2013;**112**(2):299—306.
16 Hseino H, Nugent E, Lee MJ, Hill ADK, Neary P, Tierney S, et al. Skills transfer after proficiency-based simulation training in superficial femoral artery angioplasty. *Simul Healthc* 2012;**7**(5):274—81.
17 Willaert WIM, Aggarwal R, Daruwalla F, Van Herzeele I, Darzi AW, Vermassen FE, et al. Simulated procedure rehearsal is more effective than a preoperative generic warm-up for endovascular procedures. *Ann Surg* 2012;**255**(6):1184—9.
18 Hseino H, Nugent E, Cantwell C, Lee MJ, Given M, Hill ADK, et al. Impact of an assistant on the technical skills of the primary operator in superficial femoral artery angioplasty. *Vasc Endovascular Surg* 2012;**46**(8):635—9.
19 Willaert WIM, Cheshire NJ, Aggarwal R, Van Herzeele I, Stansby G, Macdonald S, et al. Improving results for carotid artery stenting by validation of the anatomic scoring system for carotid artery stenting with patient-specific simulated rehearsal. *J Vasc Surg* 2012;**56**:1763—70.
20 Aparajita R, Zayed MA, Casey K, Dayal R, Lee JT. Development and implementation of an introductory endovascular training course for medical students. *Ann Vasc Surg* 2011;**25**(8):1104—12.
21 Bech B, Lönn L, Falkenberg M, Bartholdy NJ, Räder SB, Schroeder TV, et al. Construct validity and reliability of structured assessment of endovascular expertise in a simulated setting. *Eur J Vasc Endovasc Surg* 2011;**42**(4):539—48.
22 Bech B, Lönn L, Schroeder TV, Räder SBEW, Ringsted C. Capturing the essence of developing endovascular expertise for the construction of a global assessment instrument. *Eur J Vasc Endovasc Surg* 2010;**40**(3):292—302.
23 Van Herzeele I, O'Donoghue KGL, Aggarwal R, Vermassen F, Darzi A, Cheshire NJW. Visuospatial and psychomotor aptitude predicts endovascular performance of inexperienced individuals on a virtual reality simulator. *J Vasc Surg* 2010;**51**(4):1035—42.
24 Van Herzeele I, Aggarwal R, Malik I, Gaines P, Hamady M, Darzi A, et al. Validation of video-based skill assessment in carotid artery stenting. *Eur J Vasc Endovasc Surg* 2009;**38**(1):1—9.
25 Tedesco MM, Pak JJ, Harris EJ, Krummel TM, Dalman RL, Lee JT. Simulation-based endovascular skills assessment: the future of credentialing? *J Vasc Surg* 2008;**47**(5):1008—14.
26 Berry M, Lystig T, Beard J, Klingestierna H, Reznick R, Lönn L. Porcine transfer study: virtual reality simulator training compared with porcine training in endovascular novices. *Cardiovasc Intervent Radiol* 2007;**30**(3):455—61.
27 Hislop SJ, Hsu JH, Narins CR, Gillespie BT, Jain RA, Schippert DW, et al. Simulator assessment of innate endovascular aptitude versus empirically correct performance. *J Vasc Surg* 2006;**43**(1):47—55.
28 Chaer R a, DeRubertis BG, Lin SC, Bush HL, Karwowski JK, Birk D, et al. Simulation improves resident performance in catheter-based intervention. *Trans Meet Am Surg Assoc* 2006;**124**(3):9—18.
29 Dayal R, Faries PL, Lin SC, Bernheim J, Hollenbeck S, DeRubertis B, et al. Computer simulation as a component of catheter-based training. *J Vasc Surg* 2004;**40**(6):1112—7.
30 Gould DA, Reekers JA. The role of simulation in training endovascular interventions. *Eur J Vasc Endovasc Surg* 2008;**35**(6):633—6.
31 Bech B, Lönn L, Schroeder TV, Räder S, Ringsted C. Response to comments regarding "Capturing the Essence of Developing Endovascular Expertise for the Construction of a Global Assessment Instrument". *Eur J Vasc Endovasc Surg* 2011;**41**(3):436—7.
32 de Villiers MR, de Villiers PJT, Kent AP. The Delphi technique in health sciences education research. *Med Teach* 2005;**27**(7):639—43.
33 Clayton MJ. Delphi: a technique to harness expert opinion for critical decision-making tasks in education. *Educ Psychol* 1997;**17**(4):373—86.
34 Bolton Medical. *Treovance instructions for use*. http://www.boltonmedical.com/. 11/22/2014.
35 Cook medical. *Cook zenith instructions for use*. www.cookmedical.com. 11/22/2014.
36 Medtronic. *Medtronic endurant instructions for use*. www.medtronic.com. 11/22/2014.
37 Gore Medical. *Gore excluder AAA endoprothesis Instructions for use*. www.goremedical.com. 11/22/2014.
38 Ilgen JS, Ma IWY, Hatala R, Cook DA. A systematic review of validity evidence for checklists versus global rating scales in simulation-based assessment. *Med Educ* 2015;**49**(2):161—73.
39 Palter VN, MacRae HM, Grantcharov TP. Development of an objective evaluation tool to assess technical skill in laparoscopic colorectal surgery: a Delphi methodology. *Am J Surg* 2011;**201**(2):251—9.
40 Tolsgaard MG, Todsen T, Sorensen JL, Ringsted C, Lorentzen T, Ottesen B, et al. International multispecialty consensus on how to evaluate ultrasound competence: a Delphi consensus survey. *PLoS One* 2013;**8**(2):e57687.
41 Graham B. Delphi as a method to establish consensus for diagnostic criteria. *J Clin Epidemiol* 2003;**56**(12):1150—6.
42 Maertens H, Aggarwal R, Macdonald S, Vermassen F, Van Herzeele I. Transatlantic multispecialty consensus on fundamental endovascular skills: results of a Delphi consensus study. *Eur J Vasc Endovasc Surg* 2016;**51**(1):141—9.
43 *Standards for educational & psychological testing*. American Educational Research Association; 2014.
44 Pucher PH, Brunt LM, Fanelli RD, Asbun HJ, Aggarwal R. SAGES expert Delphi consensus: critical factors for safe surgical practice in laparoscopic cholecystectomy. *Surg Endosc* 2015;**29**(11):3074—85.
45 Walsh CM, Ling SC, Walters TD, Mamula P, Lightdale JR, Carnahan H. Development of the gastrointestinal endoscopy competency assessment tool for pediatric colonoscopy (GiECATKIDS). *J Pediatr Gastroenterol Nutr* 2014;**59**(4):480—6.
46 Palter VN, Graafland M, Schijven MP, Grantcharov TP. Designing a proficiency-based, content validated virtual reality curriculum for laparoscopic colorectal surgery: a Delphi approach. *Surgery* 2012;**151**(3):391—7.
47 Grantcharov TP, Reznick RK. Teaching procedural skills. *J Gen Intern Med* 2008;**336**(7653):1129—31.
48 Ericsson KA. Deliberate practice and acquisition of expert performance: a general overview. *Acad Emerg Med* 2008;**15**(11):988—94.