RESEARCH

Evaluation of the quality of transesophageal echocardiography images and verification of proficiency

Robina Matyal MD[1], Faraz Mahmood MD[2], Ziyad Omar Knio BS[2], Stephanie B Jones MD[1], Lu Yeh MD[3], Rabia Amir MD[1], Ruma Bose MD[1] and John D Mitchell MD[1]

[1]Department of Anesthesia, Critical Care and Pain Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts, USA
[2]Department of Surgery, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts, USA
[3]Department of Anesthesia and Critical Care Medicine, University Medical Center Groningen, Groningen, the Netherlands

Correspondence should be addressed to F Mahmood: fmahmoo1@bidmc.harvard.edu

Abstract

Various metrics have been used in curriculum-based transesophageal echocardiography (TEE) training programs to evaluate acquisition of proficiency. However, the quality of task completion, that is the final image quality, was subjectively evaluated in these studies. Ideally, the endpoint metric should be an objective comparison of the trainee-acquired image with a reference ideal image. Therefore, we developed a simulator-based methodology of preclinical verification of proficiency (VOP) in trainees by tracking objective evaluation of the final acquired images. We utilized geometric data from the simulator probes to compare image acquisition of anesthesia residents who participated in our structured longitudinal simulator-based TEE educational program vs ideal image planes determined from a panel of experts. Thirty-three participants completed the study (15 experts, 7 postgraduate year (PGY)-1 and 11 PGY-4). The results of our study demonstrated a significant difference in image capture success rates between learners and experts ($\chi^2 = 14.716$, $df = 2$, $P < 0.001$) with the difference between learners (PGY-1 and PGY-4) not being statistically significant ($\chi^2 = 0$, $df = 1$, $P = 1.000$). Therefore, our results suggest that novices (i.e. PGY-1 residents) are capable of attaining a level of proficiency comparable to those with modest training (i.e. PGY-4 residents) after completion of a simulation-based training curriculum. However, professionals with years of clinical training (i.e. attending physicians) exhibit a superior mastery of such skills. It is hence feasible to develop a simulator-based VOP program in performance of TEE for junior anesthesia residents.

Introduction

The ability to accurately and reliably evaluate a trainee's actions is the highest level of clinical performance assessment tool (1). It is an important component of trainee feedback and critical to identification of opportunities for curricular interventions (2). Mixed simulators that are able to accurately simulate an operative or clinical imaging environment have been increasingly incorporated into clinician training (3, 4). With widespread implementation of these programs, preclinical verification of proficiency (VOP) is now considered an integral component of surgical training (5). Task performance in these programs is evaluated

with observer- and endpoint-based metrics and motion analysis (1, 2, 5, 6, 7).

Similarly in transesophageal echocardiography (TEE) teaching, the role of simulators in enhancing training has been established (8, 9, 10). While various metrics have been used in curriculum-based TEE training programs to evaluate acquisition of proficiency, the quality of task completion, that is the final image quality, has usually been subjectively evaluated in these studies (10, 11, 12, 13, 14). Ideally, the endpoint metric should be an objective comparison of the trainee-acquired image with a reference ideal image. Motion tracking software can allow positional data capture with a narrow range of error. It is therefore possible that the range of geometric-scan plane position for an ideal image can also be established (10, 14). Comparison of the geometric position of the reference/ideal and trainee's scan plane could then be used as an objective endpoint metric of image quality. Our primary hypothesis is that the novel VOP metric will allow for differentiation of trainee performance from expert performance in image acquisition. Secondly, we hypothesize that this VOP metric can objectively compare performance between differently trained cohorts of learners; consistent with comparisons made previously with other metrics, we expect that the PGY-1 residents who have completed an intensive fundamentals of ultrasound (FUS) course will have similar VOP results to longitudinally trained PGY-4 residents (15).

Materials and methods

Study design

This study was conducted between October 2015 and March 2016 with institutional review board approval with waiver of informed consent. Anesthesia faculty members who were Diplomats of the National Board of Echocardiography (NBE) in Advanced Perioperative TEE (PTEeXAM), PGY-1 and PGY-4 anesthesia residents were invited to participate in this study as part of a departmental education initiative. The goal was to compare the image acquisition of the PGY-1 and PGY-2 trainees throughout the training program with that of the trained anesthesia staff.

Training programs

PGY-1 anesthesia categorical interns undergo an intensive, 13-day multimodal basic US course. This course has been integrated in the educational program of our hospital

(14). This course has different teaching modalities such as live lectures, online components, electronic books, hands-on sessions on phantom models, simulation and case discussions. During the course, three full days are dedicated to TEE, with approximately 14 h of TEE hands-on simulator training; the course schedule has been previously mentioned in earlier publications.

PGY-4 residents underwent an eight-session online and simulator-based TEE course during their PGY-2 year. This training program was organized into eight modules imparted over 4 weeks. Each training session started with 15 min of discussion or lectures about the topic and followed by a 75-min hands-on training, with approximately 10-h hands-on training over 4 weeks. Other TEE exposure during the residency program included 12 annual echo simulator sessions during their PGY-3 and 4 years, a cardiac-thoracic rotation of 1 month and optional TEE rotation of 1 month.

Traditional transesophageal echo training

Anesthesia residents participated in a structured longitudinal simulator-based TEE educational program during the PGY-3 and 4 years. It consists of an ultrasound training session (60 min) of formal didactics and hands-on instruction with the TEE simulator every 3–4 weeks (five total sessions). During this program, each trainee had approximately 20 h of hands-on TEE simulator training as well as access to our web-based training modules. Additionally, residents are exposed to clinical TEE training during a 1-month cardio-thoracic rotation in the PGY-3 year, and a dedicated 1-month clinical TEE training curriculum for each PGY-4 year resident.

FUS program

As an educational initiative, our department established an intensive FUS program for 'categorical anesthesia interns' during their PGY-1 period prior to start of anesthesia training (Supplementary Table 1, see section on supplementary data given at the end of this article). Components of the curriculum of the FUS educational program have been published previously (Supplementary Table 2). During the FUS program, the PGY-1 trainees were exposed to a total of 14 h of dedicated and supervised hands-on TEE simulator training at the departmental ultrasound simulation laboratory as well as live and online didactics. Image accuracy assessment was performed on the final day of the course.

PGY-1 group: This group consisted of categorical anesthesia interns who participated in the FUS educational program during their internship year.

PGY-4 group: This group comprised graduating anesthesia residents who received the traditional TEE training (as described previously) during their residency along with exposure to ultrasound training during sub-specialty rotations but not the FUS program. Image accuracy testing was performed individually in the last month of their anesthesia training prior to their graduation from the residency program.

Expert group: This group consisted of anesthesia faculty certified by NBE perioperative TEE having successfully completed the PTEeXAM and performed and reviewed the relevant required echocardiograms. Images were determined using guidelines and common imaging planes used for assessment.

Study protocol

The study consisted of the following steps:

1. An expert acquired standardized TEE images on a simulator as the reference images.
2. The final image was observed and agreed upon by consensus of experts as the ideal standard image during acquisition.
3. The motion metric and the geometric positional data of the scan plane in the 3D space for each reference image was captured and stored.
4. Experts were invited to acquire the same images on the TEE simulator with simultaneous acquisition of the motion metric and final positional data of the scan plane.
5. The final positional data of experts was analyzed to create a range of acceptable scan plane final positions in order to define an acceptable range of accuracy.
6. PGY-1 and PGY-4 trainees were invited to acquire the selected images on the TEE simulator with capture of the motion metric and final positional data.
7. The motion metric and final positional data were compared between the groups.

Acquisition of positional and orientation data

A Vimedix TEE Simulator (CAE Healthcare, Montreal, Canada) capable of capturing motion metrics during TEE probe manipulation was used for training and evaluation. The methodology used for to capture the motion data has been described in previous studies (10). Briefly, an experienced echocardiographer acquired 12 TEE images (Supplementary Table 3) that is target cut planes (TCPs) that consisted of multiple upper, mid-esophageal and transgastric windows corresponding to a standard TEE images. These TCPs were observed and approved by consensus by five experts as images representative of the standard basic exam plus an additional two TCP's, our group finds to be useful for ventricular function and hemodynamic assessment. The 4-chamber (4C) view was omitted, as it is the starting point for motion tracking in each exam.

The TCPs were (1) mid-esophageal (ME) two chamber, (2) ME long axis, (3) ME ascending aorta long axis, (4) ME ascending aorta short axis, (5) ME aortic valve short axis, (6) ME right ventricular inflow–outflow, (7) ME bicaval, (8) transgastric mid-short axis, (9) descending aorta short axis, (10) descending aorta long axis, (11) deep transgastric long axis and (12) transgastric long axis.

Motion metric data collection and position tracking

A participant was first asked to obtain a ME-4C view so that the starting point for all TCPs, across all participants, was similar. Metrics tracking was started by the instructor, and the participant was asked to capture the first TCP. Metrics tracking was stopped when the participant was satisfied with the quality of his/her TCP. Participants were asked to return to the ME-4C view, and the procedure was repeated for each of the remaining 11 TCPs.

The simulator recorded the probe's position and orientation in the form of *x*, *y* and *z* coordinates and roll, pitch and yaw degrees (Supplementary Fig. 1). The data were exported from the simulator as a comma-separated values (.csv) file, and the final positions and orientations of each image capture were imported to the geometric analytical software 'R' (R Core Group, Vienna, Austria) for further analyses.

Data collected

In addition to tracking the probe position as Cartesian coordinates (*x*, *y*, *z*), and the probe orientation in aircraft principal axes (roll, pitch and yaw), the following motion metrics were also recorded during image acquisition:

Capture time: The total time, in seconds, from the start of metrics tracking to the time of image acquisition.

Lag time: The time, in seconds, from the start of metrics tracking to the time of the first probe acceleration peak.

Path length: The total distance, in cm, traveled by the probe tip during metrics tracking.

Probe accelerations: The number of times the probe accelerated $>0.5\,cm/s^2$ during metrics tracking.

The metrics data were exported from the simulator as .csv file and imported to R (R Core Group) for further analyses.

Assessment of image accuracy

Each participant's probe manipulation at the time of image capture was defined as a unit vector with the following endpoints:

Position:

$$(x, y, z)$$

Orientation:

$$(x + \cos(yaw)\cos(pitch), y - \cos(pitch)\sin(yaw), z + \sin(pitch))$$

where x, y, z, pitch and yaw were obtained from the .csv output.

For each TCP, the set of probe manipulations of the expert group were plotted in 3D space in R (R Core Group) (Fig. 1A), and ranges of acceptable probe position (Fig. 1B) and orientation (Fig. 1C) were defined.

These ranges were developed following the concept of a 95% CI; the sample of expert positions yielded an acceptable range sphere centered at their geometric center with a radius equal to the product of their s.d. and the

95% critical z score (z score ≈ 1.96). The range of acceptable probe orientations was constructed similarly.

The range of acceptable probe positions was a sphere centered at point p with radius r_1. The range of acceptable probe orientations was a sphere centered at point o with radius r_2:

p: The geometric centroid of expert *positions*.

r_1: $\phi^{-1}(0.975) * SD(A) + \bar{A}$, where A was the set of distances between p and expert positions.

o: The geometric centroid of expert *orientations*.

r_2: $\phi^{-1}(0.975) * SD(B) + \bar{B}$, where B was the set of distances between o and expert orientations.

$\phi(z)$ is the standard normal density (Gaussian) function $\phi(z) = 1/\sqrt{2\pi}\ e^{-\frac{1}{2}z^2}$.

The probe manipulations of the PGY-1 and the PGY-4 groups were then overlaid on the same plot (Fig. 1D), and each participant's probe manipulation was deemed successful if, and only if, their position and orientation fell within the respective acceptable ranges.

Statistical methods

A Pearson's chi-square test of independence was used to assess whether there was a significant difference in the success rate between the PGY-1 group, the PGY-4 group and the expert group. A Pearson's chi-square test of independence was then applied to the PGY-1 group and the PGY-4 group alone in order to assess whether the PGY-1 group had attained a level of proficiency comparable to the PGY-4 group.
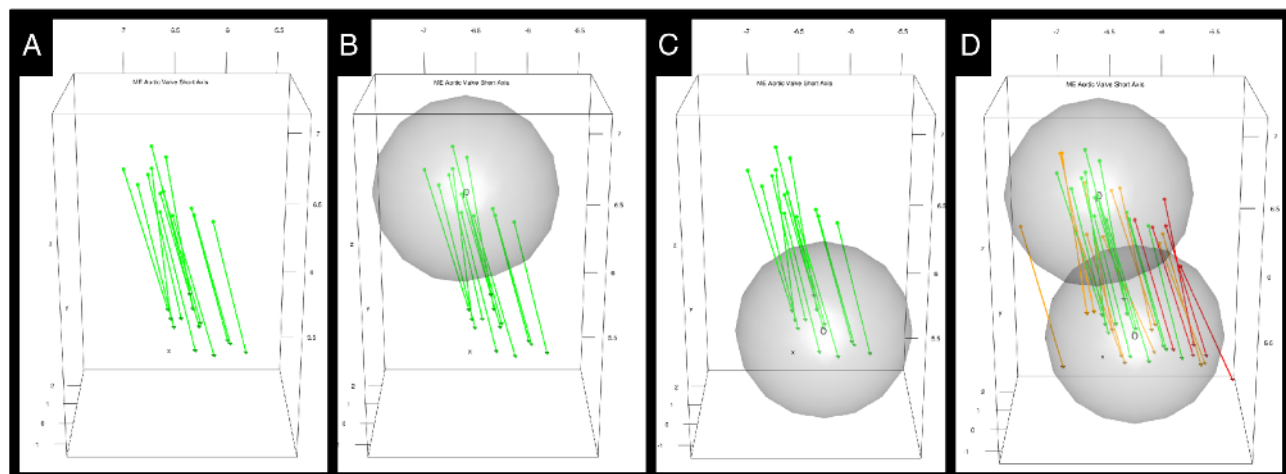

Figure 1

3D rendering of the unit direction vectors of the ultrasound beam from 15 experts capturing the mid-esophageal aortic valve short axis transesophageal echocardiography target cut plane (A). The derived range of acceptable probe positions (B) and orientations (C) are displayed as spheres with centers and radii generated from the distribution of expert unit vectors as described. The probe manipulations of the PGY-1 group and the PGY-4 groups were then overlaid on the same plot (D), with red vectors corresponding to the PGY-1 group and orange vectors corresponding to the PGY-4 group. PGY, postgraduate year.

One-way ANOVA was used to evaluate whether there was a difference in the path length, capture time, number of accelerations from rest and lag time between the three groups. A *post hoc* Tukey HSD test quantified the significance of pairwise differences between the PGY-1 group and the PGY-4 group in order to assess whether the PGY-1 group had attained a level of proficiency comparable to the PGY-4 group.

Results

Thirty-three participants completed the study (15 experts, 7 PGY-1 and 11 PGY-4). Fourteen image acquisitions were inaccessible due to technical errors, resulting in 382/396 images (96.5%) available for analysis.

Successful image captures

The proportion of successful image captures, categorized by training level is summarized in Table 1.

Within the PGY-1 group, 68 out of 82 captured images were acceptable, yielding an aggregate success rate of 82.9%. Within the PGY-4 group, 103 out of 124 captured images were acceptable, yielding an aggregate success rate of 83.1%. Within the expert group, 168 out of 176 captured images were acceptable under the defined spheres for each respective target cut plane, yielding an aggregate success rate of 95.5%.

A Pearson's chi-square test of independence suggests that there is a significant difference in success rate between the PGY-1 group, the PGY-4 group and the expert group ($\chi^2 = 14.716$, $df = 2$, $P < 0.001$). However, the difference between the PGY-1 and PGY-4 group alone was not significant ($\chi^2 = 0$, $df = 1$, $P = 1.000$).

The distribution of individual subjects' proportion of successful image captures within each of the three groups is summarized in Table 2.

Table 1 Counts and proportions of successful image captures across three training groups.

Training group	Successful captures	*n*	%	*P*	
PGY-1	213	221.33333	99.7	1.000	
PGY-4	263	268.33333	105.9		0.001
Expert	313	315.33333	112.2		

A Pearson's chi-square test was used to assess the difference in proportions between all three groups, as well as between the PGY-1 group and PGY-4 group alone.
PGY, postgraduate year.

Table 2 Minimum, median and maximum proportion of successful image captures for subjects within the three starting groups.

Training group	Min (%)	Median (%)	Max (%)
PGY-1	88.8	105.6	105.5
PGY-4	97.1	113.9	109.7
Expert	105.4	122.3	113.8

PGY, postgraduate year.

Motion metrics of proficiency (independent of image accuracy; i.e. capture time, lag time, number of accelerations from rest and path length) are summarized in Table 3.

When comparing all three groups, capture time ($P < 0.001$), lag time ($P = 0.002$), number of accelerations from rest ($P < 0.001$) and path length ($P < 0.001$) differed significantly with training level.

When comparing the PGY-1 group to the PGY-4 group alone, capture time ($P < 0.001$) and lag time ($P = 0.006$) differed significantly with training level. The differences in average accelerations from rest ($P = 0.721$) and path length ($P = 0.728$) were not significant.

Discussion

It has been previously demonstrated by multiple investigators that a foundation of basic knowledge and psychomotor skills for perioperative TEE can be acquired in the skills laboratory (15, 16, 17). The VOP metric allowed for objective differentiation between experts and trainees in terms of final images acquired. The establishment of an endpoint image provides a critical approach to the assessment of image quality, which was previously subjective. In this study, echo-naïve junior residents who underwent the intensive FUS program demonstrated a proficiency level in TEE that was comparable to graduating resident who underwent traditional training. This finding was verified both with the new VOP metric and previously described hand motion metrics (15, 16). Future studies should explore how to combine the VOP metric and existing hand motion, knowledge and workflow metrics into a comprehensive proficiency index that could be used to track the progression of learners toward proficiency over time. With the ability of simulators to incorporate pathologies, Doppler information and 3D imaging, advanced evaluations for certification could also be developed.

Expertise is a trait acquired through repetitive high-level clinical exposure. While we have previously had

Table 3 Motion metrics (capture time, lag time, accelerations from rest and path length) across the three training groups.

Motion metric	PGY-1		PGY-4		P (difference between PGY-1 and PGY-4)	Expert		P (difference among all three groups)
	Mean	s.d.	Mean	s.d.		Mean	s.d.	
Capture time (s)	31.27	44.81	32.22	41.61	<0.002	13.01	15.17	<0.002
Lag time (s)	33.58	51.37	36.70	49.01	1.207	14.07	17.65	0.002
Accelerations from rest	35.89	57.93	41.18	56.41	1.568	15.13	20.13	<0.001
Path length (cm)	38.19	64.50	45.66	63.81	1.929	16.19	22.61	<0.001

A one-way ANOVA was used to assess the difference each in mean capture time between all three groups, and a *post hoc* Tukey HSD test was used to assess the difference between the PGY-1 group and the PGY-4 alone.
PGY, postgraduate year.

some limited data on expert performance, this is the first study in which we had a large enough group of experts to do a thorough comparison of all metrics between experts and trainees; it is also the first time objective endpoint data was quantifiable with the new VOP metric. Neither group of residents achieved a level of accuracy demonstrated by the experts. This implies that simulation training is not a substitute for clinical training and our program establishes readiness to learn and perform and not expertise.

There are a few limitations to this study. First, while these simulators do an excellent job in maintaining of TEE movements and representations of the anatomy, the imaging in them is more idealized and not completely concordant with human anatomy. Secondly, while the ability to capture motion metrics and assess image quality is available, not all commercially available TEE simulators offer such features. Thirdly, the determination of image quality is based on complex geometric algorithms, which most centers will not yet be able to employ due to lack of requisite hardware or software. This process can, however, be automated and possibly incorporated into simulators themselves, rendering assessment less cumbersome. Finally, while essential, image acquisition skills are only one component of clinical proficiency; other techniques exist and must be employed to assess workflow and knowledge as well.

In conclusion, based on the positional data of the TEE probe and the scan plane, we demonstrated that the quality of the acquired image could also be objectively evaluated and mathematically compared to an acceptable range of expert images. Possession of psychomotor skills for an invasive procedure is considered an integral component of VOP, and we have successfully demonstrated acquisition of psychomotor skills for TEE utilizing our training protocols. After a simulator-based training, echo-naïve junior residents were able to demonstrate psychomotor skills and image quality for TEE image acquisition that were comparable to senior residents who underwent traditional training.

References

1 Miller GE. The assessment of clinical skills/competence/performance. *Academic Medicine* 1990 **65** (Supplement 9) S63–S67. (https://doi.org/10.1097/00001888-199009000-00045)

2 Laguna MP, de Reijke TM & de la Rosette JJ. How far will simulators be involved into training? *Current Urology Reports* 2009 **10** 97–105. (https://doi.org/10.1007/s11934-009-0019-6)

3 Fried GM. FLS assessment of competency using simulated laparoscopic tasks. *Journal of Gastrointestinal Surgery* 2007 **12** 210–212. (https://doi.org/10.1007/s11605-007-0355-0)

4 Vassiliou MC, Feldman LS, Andrew CG, Bergman S, Leffondré K, Stanbridge D & Fried GM. A global assessment tool for evaluation of intraoperative laparoscopic skills. *American Journal of Surgery* 2005 **190** 107–113. (https://doi.org/10.1016/j.amjsurg.2005.04.004)

5 Sanfey H, Ketchum J, Bartlett J, Markwell S, Meier AH, Williams R & Dunnington G. Verification of proficiency in basic skills for postgraduate year 1 residents. *Surgery* 2010 **148** 759–766; discussion 766–767. (https://doi.org/10.1016/j.surg.2010.07.018)

6 Hafford ML, Van Sickle KR, Willis RE, Wilson TD, Gugliuzza K, Brown KM & Scott DJ. Ensuring competency: are fundamentals of laparoscopic surgery training and certification necessary for practicing surgeons and operating room personnel? *Surgical Endoscopy* 2012 **27** 118–126. (https://doi.org/10.1007/s00464-012-2437-7)

7 Dawe SR, Pena GN, Windsor JA, Broeders JA, Cregan PC, Hewett PJ & Maddern GJ. Systematic review of skills transfer after surgical simulation-based training. *British Journal of Surgery* 2014 **101** 1063–1076. (https://doi.org/10.1002/bjs.9482)

8 Bose R, Panzica P, Karthik S, Karthik S, Subramaniam B, Pawlowski J, Mitchell J & Mahmood F. Transesophageal echocardiography simulator: a new learning tool. *Journal of Cardiothoracic and*

Vascular Anesthesia 2009 **23** 544–548. (https://doi.org/10.1053/j.jvca.2009.01.014)

9 Bose RR, Matyal R, Warraich HJ, Summers J, Subramaniam B, Mitchell J, Panzica PJ, Shahul S & Mahmood F. Utility of a transesophageal echocardiographic simulator as a teaching tool. *Journal of Cardiothoracic and Vascular Anesthesia* 2011 **25** 212–215. (https://doi.org/10.1053/j.jvca.2010.08.014)

10 Matyal R & Mahmood F. Simulator-based transesophageal echocardiographic training with motion analysis. *Anesthesiology* 2014 **121** 389–399. (https://doi.org/10.1097/ALN.0000000000000234)

11 Ogilvie E, Vlachou A, Edsell M, Fletcher SN, Valencia O, Meineri M & Sharma V. Simulation-based teaching versus point-of-care teaching for identification of basic transoesophageal echocardiography views: a prospective randomised study. *Anaesthesia* 2015 **70** 330–335. (https://doi.org/10.1111/anae.12903)

12 Sharma V, Chamos C, Valencia O, Meineri M & Fletcher SN. The impact of internet and simulation-based training on transoesophageal echocardiography learning in anaesthetic trainees: a prospective randomised study. *Anaesthesia* 2013 **68** 621–627. (https://doi.org/10.1111/anae.12261)

13 Shakil O, Mahmood B, Matyal R, Jainandunsing JS, Mitchell J & Mahmood F. Simulation training in echocardiography: the evolution of metrics. *Journal of Cardiothoracic and Vascular Anesthesia* 2013 **27** 1034–1040. (https://doi.org/10.1053/j.jvca.2012.10.021)

14 Montealegre-Gallegos M, Mahmood F, Kim H, Bergman R, Mitchell JD, Bose R, Hawthorne KM, O'Halloran TD, Wong V, Hess PE, *et al*. Imaging skills for transthoracic echocardiography in cardiology fellows: the value of motion metrics. *Annals of Cardiac Anaesthesia* 2016 **19** 245. (https://doi.org/10.4103/0971-9784.179595)

15 Matyal R. Simulator-based transesophageal echocardiographic training with motion analysis. *Anesthesiology* 2014 **121** 389–399. (https://doi.org/10.1097/ALN.0000000000000234)

16 Mitchell JD, Montealegre-Gallegos M, Mahmood F, Owais K, Wong V, Ferla B, Chowdhury S, Nachshon A, Doshi R & Matyal R. Multimodal perioperative ultrasound course for interns allows for enhanced acquisition and retention of skills and knowledge. *A & A Case Reports* 2015 **5** 119–123. (https://doi.org/10.1213/XAA.0000000000000200)

17 Mitchell JD, Mahmood F, Wong V, Bose R, Nicolai DA, Wang A, Hess PE & Matyal R. Teaching concepts of transesophageal echocardiography via Web-based modules. *Journal of Cardiothoracic and Vascular Anesthesia* 2015 **29** 402–409. (https://doi.org/10.1053/j.jvca.2014.07.021)