UNIVERSITYOF BIRMINGHAM

University of Birmingham

Sustained clinical knowledge improvements from simulation experiences with Simulation via Instant Messaging—Birmingham Advance

SIMBA and CoMICs team

Document Version
Peer reviewed version

[Link to publication on Research at Birmingham portal](#)

Sustained clinical knowledge improvements from simulation experiences with SIMBA

Authors and Affiliations

Dengyi Zhou[a*], Meri Davitadze[b*], Emma Ooi[c], Cai Ying Ng[c], Isabel Allison[a], Lucretia Thomas[a], Thia Hanania[a], Parisha Blaggan[a], Nia Evans[d], Wentin Chen[a], Eka Melson[e,f], Kristien Boelaert[g,h], Niki Karavitaki[f,g,i], Punith Kempegowda[f,g], SIMBA and CoMICs team[f]

[*]Joint first authors

[a]College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK;
[b]Georgian-American Family Medicine Clinic "Medical House", Tbilisi, Georgia;
[c]RCSI & UCD Malaysia Campus, Malaysia;
[d]Royal Glamorgan Hospital, Cwm Taf Morgannwg University Health Board, Rhondda Cynon Taf, UK;
[e]Ninewells Hospital, NHS Tayside, Dundee, UK;
[f]Institute of Metabolism and Systems Research, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK;
[g]Queen Elizabeth Hospital, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK;
[h]Institute of Applied Health Research, University of Birmingham, Birmingham, UK;
[i]Centre for Endocrinology, Diabetes and Metabolism, Birmingham Health Partners, Birmingham, UK

Corresponding author details
Dr. Punith Kempegowda
Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK
Telephone: +44 7721 930 777
E-mail: P.Kempegowda@bham.ac.uk
Twitter: @drpunith

Word count: 2805

What is already known on the subject?

- Simulation via Instant Messaging - Birmingham Advance (SIMBA) is a real-time simulation-based medical training programme using WhatsApp and Zoom.
- Assessment encourages learning and promotes motivation by emphasising progress and achievement
- Multiple choice questions have been successfully implemented in both under- and post-graduate medical education as a mode of assessing knowledge

The study's main messages

- The assessment activities for the novel simulation model were successfully adapted using Moore's 7 Levels of continuing medical education outcomes measurements
- SIMBA has shown to increase participants' clinical knowledge on simulated endocrine clinical scenarios assessed using multiple choice questions
- The improvement in knowledge was retained up to 6-12 weeks after the session
- SIMBA is an effective teaching model and has a sustained impact on participants' knowledge on simulated topics

Abstract

Background

Simulation via Instant Messaging - Birmingham Advance (SIMBA) delivers simulation-based learning through WhatsApp and Zoom helping to sustain continued medical education (CME) for postgraduate healthcare professional otherwise disrupted by the COVID-19 pandemic. This study aimed to assess whether SIMBA helped to improve clinical knowledge and if this improvement in knowledge was sustained over time.

Methods

Two SIMBA sessions—Thyroid and Pituitary—were conducted in July-August 2020. Each session included simulation of various real-life cases and interactive discussion. Participants' self-reported confidence, acceptance, and knowledge were measured using surveys and multiple-choice questions in pre- and post-simulation and 6-12-weeks follow-up period. The evaluation surveys were designed using Moore's 7 Levels of CME outcomes measurements.

Results

A total of 116 participants were included in the analysis. Significant improvement was observed in participants' self-reported confidence in approach to simulated cases [Thyroid (n=37) (p<0.0001), Pituitary (n=79) (p<0.0001)]. Significant improvement in clinical knowledge was observed following simulation [Thyroid (n=37) (p<0.0001), Pituitary (n=79) (p<0.0001)]. For both sessions, retention of confidence and knowledge was seen at 6-12 weeks' follow-up.

Discussion

SIMBA increased participants' clinical knowledge on the simulated cases and this improvement was retained up to 6-12 weeks after the session. Further studies are required to explore long-term retention and whether it translated to improved real-world clinical practice.

Introduction

Assessment is one of the most potent tools to encourage learners (1). To promote learning, the assessment should be educational and informative thus challenging the educators to provide efficient methods. Miller proposed a framework for the clinical assessment by creating a hierarchy of skills to be demonstrated by the healthcare professional and how these skills can be evaluated (Figure 1) (2). Learner's knowledge is assessed to demonstrate readiness to accomplish the professional requirements. However, learners must also *know how* to use the knowledge they possess and this can be assessed by oral examinations, by using multiple choice questions (MCQs) or essays. Although the testing procedures are valid and reliable, they may fail to evaluate how medical professionals perform when they face the patients in the clinical setting.

Since the 1950s, MCQs have been used as a mode of assessing knowledge in medical education. Demonstrating superior reliability compared to the traditional essay questions, MCQs have been successfully implemented in both under- and post-graduate medical education (3). Traditionally, MCQs were able to assess factual knowledge whereas application and interpretation were evaluated by essays or oral examinations. However, MCQs have evolved overtime to incorporate application and interpretation. The most common approach to MCQ writing involves structuring a single best answer (SBA) using a stem and a lead-in question followed by the list of options—so called distractors (4). Ideally, the question should present a clear problem for the learner to propose the correct answer without even looking at the options. After completing the MCQs, the learner should be provided with the rationale explaining the correct and incorrect answers. Applying stems in the MCQs potentiated evaluating the learner's higher cognitive skills, such as problem-solving, and competence (5). MCQs as an assessment method provides easily scored, objective, and unbiased data, which can be used to identify the strengths and weaknesses of any teaching technique (6).

The coronavirus (COVID-19) pandemic has disrupted medical education worldwide. This necessitated rapid transitioning from in-person lessons to video-conferencing services to maintain the same level of teaching standards (7). Simulation via Instant Messaging - Birmingham Advance (SIMBA) is one such initiative delivering high calibre training through virtual sessions in various fields of medicine and contribute to sustained medical training during the pandemic. It is based upon Kolb's experiential learning theory (8) and the change in confidence is assessed using pre- and post-simulation surveys based on the 7-point Likert scale. The sessions involve participants interacting with moderators via WhatsApp to solve complex real-life clinical scenarios and, following the simulation, engaging in discussion session over Zoom chaired by an expert. These sessions significantly improved healthcare professionals' confidence in managing various conditions (9). However, participants' self-reported confidence levels can be subjective. It is also important for participants to be able to understand how to effectively apply knowledge gained and self-reported confidence levels alone are insufficient to assess this. The long-term effectiveness of SIMBA as a teaching method has also yet to be determined.

This study aimed to assess whether SIMBA helped to improve clinical knowledge and if this improvement in knowledge was sustained over time.

Materials and methods

The study was conducted in two sessions in July 2020 (Thyroid) and August 2020 (Pituitary) by the SIMBA team with the support of the Institute of Metabolism and Systems Research and the Institute of Applied Health Research, both at the University of Birmingham. The training sessions were aimed at specialistty registrars in endocrinology.

The SIMBA sessions

The detailed description of the steps building up to the SIMBA session and the simulation process itself is published elsewhere, with the key steps summarised below (9) (Figure 2).

SIMBA was based on interactive simulation-based learning through WhatsApp. For each session, real-life case scenarios were identified to prepare standardised transcripts, which were validated and approved by consultant endocrinologist with specialist expertise experts in the relevant specialities. Any images used were validated and approved by consultant radiologists. The transcripts included sufficient medical history, clinical examinations, diagnostic test results and imaging findings that would enable participants to reach diagnosis and propose management and follow-up plan. No patient identifiable data was included in the transcripts. Transcripts of five thyroid clinical scenarios (autoimmune hypothyroidism, Graves' disease during pregnancy, amiodarone-induced thyrotoxicosis, thyroid cancer, and TSH-secreting pituitary adenoma [TSHoma]) and five pituitary clinical scenarios (non-functioning pituitary adenoma [NFPA], craniopharyngioma, acromegaly, macroprolactinoma, and Cushing's disease) were used for relevant sessions.

Medical students and junior doctors were recruited based on their interest in endocrinology and their motivation to participate as moderators. 42 and 33 moderators participated in the Thyroid and Pituitary sessions, respectively. Prior to the sessions, these moderators were trained by experienced SIMBA team members using the finalised transcripts to ensure their proficiency. Experienced SIMBA team members had all previously moderated at least twice, had their moderating peer-reviewed with feedback given and received training on how to lead a moderator training session.

The sessions were advertised publicly on social media (Twitter, Facebook) and mailing lists of endorsing societies to invite interested participants to register by completing a Google form. Health Education West Midlands Diabetes and Endocrinology specialist training committee also helped recruit trainees as part of their specialist training improvement initiative. Participants received instructions, a unique ID number, and their moderator's WhatsApp number via email few days before the session.

On the day of simulation, participants interacted with moderators using WhatsApp, were provided with the presenting complaint of the patientpatient, and asked to approach and approached to solve the simulated cases as they would do in their daily clinical practice. At the start of the simulation, the moderator took up the role of a patient from whom clinicians requested a full history from history of presenting complaint to past medical history and social history. The moderator would also provide the clinician with any relevant physical, biochemicalbiochemical, and radiological results they requested. Lastly, the moderator prompted the trainee to combine all relevant information to arrive at the diagnosis, management and follow-up plans to present to the MDT. The first case for each session was run as a trial to allow participants familiarise themselves with the SIMBA model.

After completion of the remaining four cases, participants were invited to an interactive debrief via Zoom chaired by expert endocrinologists who focused on an appropriate approach to the cases with reference to national/international guidelines.

Evaluation of SIMBA

Participants' acceptance rate and improvement in their self-reported confidence levels to manage simulated cases pre- and post-SIMBA were assessed using 7-point Likert scale ranging from "strongly agree" and "strongly disagree". The evaluation survey was based on Kirkpatrick's training evaluation model (10): Level 1 (reaction) included questions regarding engagement of the session, Level 2 (learning) involved self-reported improvement in core competencies, confidence levels in approaching various endocrine cases and MCQs, and Level 3 (behaviour) was assessed by open-ended questions regarding changes they intend to make in the patient care following the session. Data from self-reported confidence levels were categorised into three groups: (i) confident: for those who responded with "strongly agree" and "agree", (ii) not confident: for those who responded with "disagree" and "strongly disagree", and (iii) unsure: for those who responded with "agree somewhat", "disagree somewhat", and "undecided". In addition, participants were asked to comment on their overall impression of the session, the consultant's contribution during the discussion, and interaction with the moderators.

Multiple choice questions (MCQs)

Ten SBA MCQs were created for each session with two MCQs dedicated to each case. The questions included one correct answer and four distractors, all labelled A-E. These MCQs were critically reviewed and approved for accuracy and validity by expert endocrinologists for the relevant sub-speciality. Participants were invited to complete the MCQs just before and after the simulations as part of pre- and post-SIMBA surveys, and again at 6-12 weeks as a follow-up.

Statistical analyses

Participants who completed both pre- or post-SIMBA evaluation forms were included in the analysis. The confidence levels in managing simulated cases pre- and post-SIMBA are reported using frequencies and proportions. MCQ scores pre- and post-SIMBA from both sessions are reported in medians and quartiles. Changes in confidence levels and MCQ scores were measured using Wilcoxon signed-rank test, comparing the paired samples. A similar analysis was conducted 6-12 weeks after the session, including participants who completed all pre-SIMBA, post-SIMBA, and follow-up evaluation forms. Additionally, data on feedback and key takeaway from both sessions were collected post-SIMBA. Findings from responses to open-ended questions were reviewed in an inductive thematic analysis and presented in tables with examples.

Results

A total of 37 and 79 participants from the SIMBA Thyroid and Pituitary sessions, respectively, completed both pre- and post-SIMBA evaluations and were included in the analysis. SIMBA Thyroid participants include 20 (54.1%) from the UK, and 17 (45.9%) participants internationally, across three

continents (Africa, Asia, Europe), mainly comprised of specialty training registrars (n=22/37, 59.5%). Of the 79 SIMBA Pituitary participants, 45 (57.0%) were from the UK, and 34 (43.0%) from the rest of the world, across three continents (Africa, Asia, Europe). Similar to the thyroid session, majority of these participants are specialty training registrars (n=50/79, 63.3%).

Significant improvement was observed in participants' self-reported confidence in their approach to simulated cases post-SIMBA Thyroid session (Graves' disease during pregnancy (p=0.0032), amiodarone-induced thyrotoxicosis (p<0.0001), and thyroid cancer (p=0.0005), and TSHoma (p<0.0001)) and all the simulated scenarios in Pituitary session (p<0.0001) (Table 1). From SIMBA thyroid, TSHoma saw the greatest increase in percentage of participants who became confident in their approach post-session (+56.8%). From SIMBA pituitary, craniopharyngioma saw the greatest increase in percentage of participants who became confident in their approach post-session (+50.6%). There was a trend towards improved confidence level for autoimmune hypothyroidism (p=0.0625). There was no decrease in confidence seen in any of the simulated cases. Overall, there was a greater increase in participants who reported they were confident following the pituitary session compared to thyroid (+43.0% vs. +35.1%).

There was a significant improvement in MCQs seen for both thyroid and pituitary sessions (p<0.0001). A larger improvement in MCQ scores was seen for thyroid session [pre-SIMBA: median (IQR)- 60% (40%-70%) vs 90% (80%-100%) post-SIMBA; (p<0.0001)] compared to pituitary session post-simulation [Pre-SIMBA: 80% (60%-90%) vs 80% (70%-90%) post-SIMBA; (p<0.0001)].

During follow-up, 36 participants (Thyroid n=18, Pituitary n=18) completed the MCQs. Overall, there was a retention of confidence levels in participants' approach to simulated thyroid (p=1.0000) and pituitary (p=0.1696) cases in comparison to post-SIMBA results (Table 2). Minimal change in MCQ scores were observed for both sessions at follow-up [Thyroid (post-SIMBA: median (IQR)- 90% (70%-100%) vs follow-up median (IQR)- 90% (80%-100%) (p=0.3829)), Pituitary (post-SIMBA: median (IQR)-40% (20%-40%) vs follow-up median (IQR)- 45% (10%-68%) (p=0.7825)] (Figure 3).

Discussion

SIMBA proved to be an effective learning model to increase self-reported confidence level and clinical knowledge in managing the simulated thyroid and pituitary cases. Retention of confidence and knowledge was seen at 6-12 weeks' follow-up. This study builds on to the established SIMBA model (9) and further evaluated the impact of simulation on change in clinical knowledge and its retention. SIMBA is based on Kolb's experiential learning theory, which is effective in guiding simulation-based medical education, allowing participants to gain knowledge during each phase of the learning cycle (8,11). The cycle begins with a concrete experience, which is represented by the simulated virtual cases conducted through WhatsApp. Reflective observation is facilitated by discussion with the expert via Zoom, who discusses each case with evidence-based rationale allowing the participants to compare what was done differently and reflect on their personal performance. Abstract conceptualisation was enabled by using post-SIMBA MCQs to ensure participants can incorporate the knowledge they learnt. The final phase of active experimenting was shown by retained knowledge during the follow-up period. These findings prove SIMBA helped improve participants' knowledge, competencecompetence, and performance. We saw a larger improvement in clinical knowledge in thyroid session compared to pituitary session post-simulation,

which may be due to higher median score at baseline for the latter and hence less scope for further improvement.

The assessment activities were designed using Moore's 7 Levels of CME Outcomes Framework (12). Level 1 (participation) was assessed using pre-SIMBA questionnaires which included participants' demographic information, whereas level 2 (satisfaction) was assessed using post-SIMBA questionnaires revealing that SIMBA had a high acceptance rate among participants and was an effective learning model. MCQs tested participants' declarative knowledge (Level 3a), procedural knowledge (Level 3b), and competence (Level 4), which increased significantly after both SIMBA sessions. The participants' performance (level 5) was assessed via self-reported performance in the 6-12-week follow-up survey, which showed sustained improvement in knowledge and competence.

Well-constructed MCQs helps to evaluate application of knowledge, interpretation or creation (13). Case-based MCQs also lead to a higher level of learning and deeper information processing compared to traditional MCQs (7,14). In this study, the MCQs were matched to the simulated cases to test participants' factual knowledge and their ability to interpret the given case scenario using this knowledge. On the other hand, MCQs have also been used to compare the efficacy between virtual and face-to-face teaching demonstrating that online teaching formats are sufficient educational tools in terms of providing comparable degrees of satisfaction and comprehension of learners (15).

Limitations and future research

While providing further evidence proving SIMBA model's effectiveness in medical education, the study has certain limitations. Whilst we were able to design assessments using Miller's levels 1 (knows), 2 (knows how), and 3 (shows how), level 4 (does) translating to direct observation in clinical settings remains a challenge. This could be addressed by requesting participants to ask their supervisors to complete an assessment form whilst observing them during clinical practice. This assessment form would be focused on the area of study and would allow us to gauge the effects of SIMBA on the competencies of doctors from a key performance indicator's perspective.

We also could not measure level 6 (patient health) and level 7 (community health) of Moore's CME outcomes due to the inherent limitations of simulation models. Furthermore, only a proportion of participants responded to the follow-up survey, so our findings of knowledge retention is not generalisable to all participants. There may also be some selection bias, as the most enthusiastic participants were most likely to answer the questionnaire. The same set of MCQs was utilised in the pre-, post-, and follow-up surveys meaning that the participants may have simply memorised the answers and may not be an accurate reflection of the actual knowledge and competency.

Another limitation would be our relatively short follow-up period of 6-12 weeks. It would be useful to conduct a further longitudinal evaluation at 6 months and 1 year following the SIMBA session to identify the longstanding impact of SIMBA as a teaching method.

It has been demonstrated that the practical solution for the post-CODIV era appears to be the blended model of learning—providing theoretical teaching virtually while practicing practical skills face-to-face (16). Further research is required to evaluate whether SIMBA can supplement or replace aspects of traditional face-to-face teaching in standard practice. A comparative study could be conducted with matched clinicians at similar levels of their specialty training with half allocated to receive SIMBA training and the other half to standard face-to-face teaching, with pre- and post-MCQs used to evaluate if SIMBA was as effective.

Conclusions

SIMBA proved to be an effective teaching model and has an intermediate-term positive impact on participants' confidence in managing simulated cases. Further studies are required to explore whether confidence levels and increase in knowledge can be translated to real-world practice, as well as improved patient care.

Declaration

Conflict of interests

The authors declare no conflict of interests.

Funding

Ethics approval

The protocol was approved, and the study was conducted as part of specialist training improvement initiative commissioned by Health Education West Midlands Diabetes and Endocrinology specialist training committee. Participation into this session was voluntary and all participants were informed using a General Data Protection Regulation (GDPR) statement during registration. A mandatory tick-box consenting to participation was included in registration form to ensure voluntary consent. All methods were carried out in accordance with relevant guidelines and regulations.

Authors' Contribution

DZ, SIMBA Pituitary session lead, and MD, SIMBA Thyroid session lead, are the joint first authors having made all round contributions to the study. EO has analysed and interpreted the data analysis. WC helped with writing the manuscript. CYN and EO have constructed the MCQs for both sessions. LC, TH, PB, and NE were core moderators during the sessions and have created the transcripts, advertisements, and other relevant materials. KB and NK were the expert endocrinologists in the study, who have approved the finalised transcripts and have chaired the discussion and Q&A sessions for Thyroid and Pituitary, respectively. PK and EM conceptualised and supervised the

delivery of all aspects of SIMBA, and critically reviewed the manuscript. The final version has been reviewed and approved by all the named authors.

References:

1. Wass V, Van Der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. Vol. 357, Lancet. Elsevier Limited; 2001. p. 945–9.

2. Miller GE. The assessment of clinical skills/competence/performance. Vol. 65, Academic Medicine. Acad Med; 1990. p. S63–7.

3. Moss E. Multiple choice questions: Their value as an assessment tool. Vol. 14, Current Opinion in Anaesthesiology. Curr Opin Anaesthesiol; 2001. p. 661–6.

4. Coughlin PA, Featherstone CR. How to Write a High Quality Multiple Choice Question (MCQ): A Guide for Clinicians. Vol. 54, European Journal of Vascular and Endovascular Surgery. W.B. Saunders Ltd; 2017. p. 654–8.

5. Collins J. Education techniques for lifelong learning: Writing multiple-choice questions for continuing medical education activities and self-assessment modules. Vol. 26, Radiographics. 2006. p. 543–51.

6. Farooqui F, Saeed N, Aaraj S, Sami MA, Amir M. A Comparison Between Written Assessment Methods: Multiple-choice and Short Answer Questions in End-of-clerkship Examinations for Final Year Medical Students. Cureus. 2018 Dec;10(12):e3773–e3773.

7. Donnelly C. The use of case based multiple choice questions for assessing large group teaching: implications on student's learning. Irish J Acad Pract. 2014;3(1):12.

8. Kolb DA. Experiential Learning: Experience as the Source of Learning and Development. Englewood Cliffs, NJ Prentice Hall. 1984;

9. Melson E, Davitadze M, Aftab M, Ng CY, Ooi E, Blaggan P, et al. Simulation via instant messaging-Birmingham advance (SIMBA) model helped improve clinicians' confidence to manage cases in diabetes and endocrinology. BMC Med Educ [Internet]. 2020 Dec 18 [cited 2020 Aug 25];20(1):274.

10. Kirkpatrick DL, Kirkpatrick JD. Evaluating training programs : the four levels. Berrett-Koehler; 2006. 379 p.

11. Poore JA, Cullen DL, Schaar GL. Simulation-based interprofessional education guided by Kolb's experiential learning theory. Vol. 10, Clinical Simulation in Nursing. Elsevier Inc.; 2014. p. e241–7.

12. Moore DEJ, Green JS, Gallis HA. Achieving desired results and improved outcomes: integrating planning and assessment throughout learning activities. J Contin Educ Health Prof. 2009;29(1):1–15.

13. Case S, Swanson D. Constructing Written Test Questions For the Basic and Clinical Sciences. Natl Board Exam. 2002 Jan;

14. Chéron M, Ademi M, Kraft F, Löffler-Stastka H. Case-based learning and multiple choice

questioning methods favored by students. BMC Med Educ. 2016;16(1):41.

15. Kenzaka T, Goda K, Kumabe A. A Comparison of the Learning Effects of Face-to-Face Versus Online Formats of a Clinical Reasoning Lecture. 2022 Jun 20;14(6).

16. Atwa H, Shehata MH, Al-Ansari A, Kumar A, Jaradat A, Ahmed J, et al. Online, Face-to-Face, or Blended Learning? Faculty and Medical Students' Perceptions During the COVID-19 Pandemic: A Mixed-Method Study. Front Med. 2022 Feb 3;9:791352.

Figure 1. Miller's pyramid of framework for clinical assessment (Georgie E. Miller, 1990).

Figure 2. Working model of SIMBA is divided into two main phases: preparation and on-the-day. In the preparation phase, suitable cases are identified from outpatient clinics. Anonymised transcripts containing relevant history, examination, and investigations are prepared. These are then approved by an expert in the field. The approved transcripts are used to train the moderators. Prior to the actual day of the session, the sessions are advertised and registered healthcare professionals are provided with detailed instructions on how to join the session. On the day, participants are assigned to moderators usually in a 3:1 ratio. The participant contacts the moderator who then initiates the mock case with a focus to let the participant familiarise with the SIMBA model. Usually, this lasts about 20 min following which the coordinator for the day liaises with the participant to clarify any technical issues. This is followed by SIMBA session with 4-8 cases with a 15-minute comfort break in the middle of the session. Upon completion of the simulation, an expert discusses the simulated cases based on the transcripts and current relevant guidelines via Zoom. The participants have ample time to engage with the expert to clarify any doubts.

Figure 3. Summary of MCQ scores of participants who completed all three pre-, post-, and follow-up evaluations. Wilcoxon-signed rank test was used to compare MCQ scores pre-SIMBA, post-SIMBA and after follow-up. Compared to pre-SIMBA, there was a significant improvement in MCQ scores post-SIMBA thyroid [pre-SIMBA: median (IQR) 60% (40%-70%) vs 90% (80%-100%) post-SIMBA; (p<0.0001)] and post-SIMBA pituitary [Pre-SIMBA: 80% (60%-90%) vs 80% (70%-90%) post-SIMBA; (p<0.0001)]. Compared to pre-SIMBA, there were no significant differences in MCQ scores for both sessions at 6-12 weeks' follow-up [Thyroid (post-SIMBA: median (IQR)- 90% (70%-100%) vs follow-up median (IQR)- 90% (80%-100%) (p=0.3829)), Pituitary (post-SIMBA: median (IQR)- 40% (20%-40%) vs follow-up median (IQR)- 45% (10%-68%) (p=0.7825)].

Table 1. Changes in participants' confidence levels in their approach to simulated cases, post-SIMBA, in all sessions individually and combined (overall). (*P<0.05)

AIT, Amiodarone-induced thyrotoxicosis; TSHoma, Thyrotropinoma; NFPA, Non-functioning pituitary adenoma.

Session	Case	Confident	Unsure	Not confident	Significance
Thyroid (n=37)	Autoimmune hypothyroidism	+16.2%	-13.5%	-2.7%	P=0.0625
	Graves' disease during pregnancy	+29.7%	-27.0%	-2.7%	P=0.0032*
	AIT	+40.5%	-32.4%	-8.1%	P<0.0001*
	Thyroid cancer	+32.4%	-29.7%	-2.7%	P=0.0005*
	TSHoma	+56.8%	-48.5%	-8.1%	P<0.0001
	Overall	+35.1%	-30.3%	-4.9%	P<0.0001*
Pituitary (n=79)	NFPA	+41.8%	-41.8%	0.0%	P<0.0001*
	Craniopharyngioma	+50.6%	-45.6%	-5.1%	P<0.0001*
	Macroprolactinoma	+31.6%	-30.4%	-1.3%	P<0.0001*
	Acromegaly	+49.4%	-49.4%	0.0%	P<0.0001*
	Cushing's Disease	+41.5%	-41.5%	0.0%	P<0.0001*
	Overall	+43.0%	-41.8%	-1.3%	P<0.0001*

Table 2. Changes in participants' confidence levels in their approach to simulated cases, at follow-up compared to the post-SIMBA.

AIT, Amiodarone-induced thyrotoxicosis; TSHoma, Thyrotropinoma; NFPA, Non-functioning pituitary adenoma.

Session	Case	Confident	Unsure	Not-Confident	Significance
Thyroid (n=18)	Autoimmune hypothyroidism	0.0%	0.0%	0.0%	P=1.0000
	Graves' disease during pregnancy	0.0%	0.0%	0.0%	P=1.0000
	AIT	+5.6%	-11.1%	+5.6%	P=1.0000
	Thyroid cancer	0.0%	0.0%	0.0%	P=1.0000
	TSHoma	-5.5%	+11.1%	-5.6%	P=1.0000
	Overall	0.0%	0.0%	0.0%	P=1.0000
Pituitary (n=18)	NFPA	+5.6%	0.0%	-5.6%	P=1.0000
	Craniopharyngioma	+11.1%	-5.6%	-5.6%	P=0.5312
	Macroprolactinoma	+16.7%	-11.1%	-5.6%	P=0.2500
	Acromegaly	-11.1%	+16.7%	-5.6%	P=1.0000
	Cushing's Disease	0.0%	+5.6%	-5.6%	P=1.0000
	Overall	+4.5%	+1.1%	-5.6%	P=0.1696