

Access this article online
Quick Response Code:

Website: https://journals.lww.com/jehp/
DOI: 10.4103/jehp.jehp_2068_24

Radiology residents' learning curve for interpretation of female pelvic magnetic resonance imaging (MRI): A longitudinal study in Isfahan university of medical sciences

Atoosa Adibi, Tahereh Changiz¹, Elham Maghami²

Abstract:

BACKGROUND: Accurate interpretation of female pelvic magnetic resonance imaging (MRI) is a critical yet challenging skill for radiology residents. This study aimed to evaluate the learning curve of pelvic MRI interpretation following a structured educational intervention.

MATERIALS AND METHODS: In this longitudinal semi-experimental study, 20 third- and fourth-year radiology residents participated between 2023 and 2024. Participants underwent seven consecutive pelvic MRI interpretation assessments, including one baseline test and six post-training evaluations. The educational intervention consisted of three structured pelvic MRI training sessions followed by iterative feedback-based learning. Performance was assessed using standardized scoring (maximum score = 100). Changes in scores over time were analyzed using repeated-measures analysis of variance (ANOVA) and analysis of covariance (ANCOVA), adjusting for residency year and baseline performance. Statistical significance was defined as $P < 0.05$.

RESULTS: Mean examination scores increased significantly from 17.8 ± 13.7 at baseline to 77.1 ± 14.4 at the final assessment ($P < 0.001$). Repeated-measures ANOVA demonstrated a significant time effect (Greenhouse–Geisser correction, $P < 0.001$). Variability in scores decreased progressively over time, indicating convergence toward stable competency.

CONCLUSION: A structured, feedback-oriented pelvic MRI training program resulted in significant and sustained improvement in residents' interpretative performance, supporting its application in competency-based radiology education.

Keywords:

Clinical competence, learning curve, magnetic resonance imaging, pelvis, radiology education

Introduction

Medical techniques usually require many hours of training, practice, and purposeful repetition. Learning curves are graphical representations of improving a specific task or art over time.^[1] Medical technical performance can be evaluated using operation time and complication rates. In contrast, experience can be assessed

using other variables, such as the number of repetitions and training hours.^[2] Learning curves are often used to evaluate medical education. They provide information about the variables affecting the learning process and may help plan lessons. They have been widely used in less invasive surgical techniques.^[3]

Today, magnetic resonance imaging (MRI) has become one of the primary modalities

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License (CC BY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

How to cite this article: Adibi A, Changiz T, Maghami E. Radiology residents' learning curve for interpretation of female pelvic magnetic resonance imaging (MRI): A longitudinal study in Isfahan University of Medical Sciences. *J Edu Health Promot* 2026;15:217.

Department of Radiology,
School of Medicine,
Al-Zahra Hospital, Isfahan
University of Medical
Sciences, Isfahan, Iran,
¹Department of Medical
Education, Educational
Development Center,
Medical Education
Research Center, Isfahan
University of Medical
Sciences, Isfahan, Iran,
²Department of Radiology,
School of Medicine,
Isfahan University of
Medical Sciences, Isfahan,
Iran

Address for correspondence:

Dr. Elham Maghami,
Department of Radiology,
School of Medicine,
Isfahan University of
Medical Sciences, Isfahan,
Iran.
E-mail: elham.maghami@gmail.com

Received: 25-11-2025
Accepted: 21-02-2026
Published: 29-05-2026

to diagnose and guide treatment in patients with gynecologic diseases, including staging gynecological cancers, examining the response to treatment and monitoring malignancies, determining surgical plans, and evaluating acute pelvic pain in emergencies.^[4] MRI is a selective modality for evaluating suspected developmental disorders of the vagina, which may be isolated or accompanied by uterine anomalies. MRI is used for soft tissue examination with the benefit of no ionizing radiation, to diagnose and determine treatment plans.^[5] MRI, for example, is functional for finding the location and determining the thickness of the vaginal septum, as well as for determining the treatment plan and identifying other potential causes of hematocolpus such as hymenal atresia.^[6] MRI is also helpful in investigating pathological conditions. It is used for benign and malignant vulva, urethra, uterus, ovaries, and fallopian tubes; for example, in cervical cancer, MRI without contrast injection has a high resolution in examining soft tissue and can clearly show the location of the primary tumor extension and the extent of pelvic metastases.^[7] Therefore, the ability to accurately diagnose tumor size, parametrial and pelvic sidewall invasion, bladder or rectal invasion, and lymph node metastases is helpful in disease management and treatment decisions. MRI has a high accuracy for diagnosis in the evaluation of parametrial invasion, with a sensitivity of 88%–97% and a specificity of 93%. Parametrial invasion can also be excluded entirely by MRI with a negative predictive value of 94%–100%, and MRI can determine the size of the primary tumor. Determination with 93% accuracy and primary invasion into the bladder and rectum can be excluded entirely in MRI with a predictive value of 100%.^[8]

In evaluating pelvic masses, MRI is more specific and accurate than ultrasound for characterization and is the modality of choice in doubtful female pelvic masses in ultrasound.^[9] So, it is important to get the required skills for interpretation during the radiology residency curriculum. Unfortunately, in our department, there was not any modified protocol for education on this part of imaging, and the radiology residents showed

flaws in pelvic MRI imaging interpretation. However, there was a limited number of image interpretations at our hospital. There were also several requests from levels 3 to 4 residents in our department for teaching this field of imaging. Despite all the benefits and applications, pelvic MRI is one of the challenging topics for radiology assistants due to its difficulty.^[10] This study provides a longitudinal evaluation of the learning curve for female pelvic MRI interpretation among senior radiology residents using repeated assessments with structured feedback. Unlike prior cross-sectional educational studies, it demonstrates not only progressive improvement in performance but also a reduction in score variance over time, indicating skill convergence and attainment of a stable interpretive competence. The proposed multi-step training protocol offers a practical and reproducible model for advanced radiology education.

Based on learning curve and mastery learning theories, we hypothesized that structured training combined with repeated practice and expert feedback would result in progressive improvement and convergence of pelvic MRI interpretation skills among radiology residents. Accordingly, a conceptual framework was developed to illustrate the relationship between baseline skill level, educational intervention, repeated assessments, feedback, and learning outcomes [Figure 1].

Materials and Methods

Study design and setting

This longitudinal semi-experimental study was conducted between 2023 and 2024 at Al-Zahra Hospital, the main teaching hospital affiliated with Isfahan University of Medical Sciences, Isfahan, Iran. The study was designed to evaluate the learning curve of pelvic MRI interpretation among radiology residents following a structured educational intervention.

Study participants and sampling

A convenience sample of 20 third- and fourth-year radiology residents was enrolled in the study after

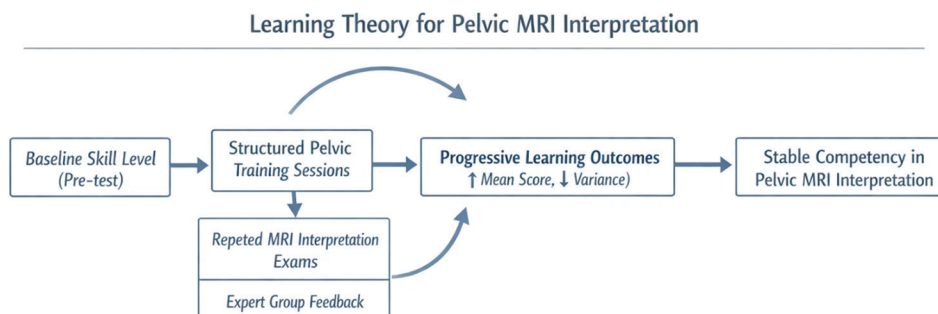


Figure 1: Conceptual framework of the learning process for female pelvic MRI interpretation among radiology residents. MRI = Magnetic resonance imaging

obtaining written informed consent. All eligible residents who agreed to participate were included. There was no exclusion criteria related to prior pelvic MRI experience, and participants represented different baseline skill levels.

Data collection tool and technique

Participants completed seven consecutive pelvic MRI interpretation assessments, including one baseline test and six post-intervention tests. The examinations were designed by the primary investigator and covered essential pelvic MRI interpretation topics, including the uterus, cervix, and ovaries. Each assessment consisted of previously unseen cases, and the maximum attainable score was 100.

The educational intervention comprised three structured pelvic MRI training sessions, followed by iterative group feedback sessions after subsequent examinations. Test intervals ranged from 1 to 4 weeks depending on residents' schedules and national holidays. Image interpretation was performed using sent images or the PACS (Picture archiving and communication system) system, depending on the test phase.

All answer sheets were anonymized and coded to ensure assessor blinding. A single experienced rater evaluated all examinations using standardized scoring criteria to maintain consistency.

To construct learning curves, the number of examinations represented the horizontal axis, and exam scores represented the vertical axis. Individual learning trajectories (Panel A), group mean learning curves (Panel B), mean curves with 95% confidence intervals (Panel C), and score variance trends (Panel D) were generated.

Ethical consideration

The study (project ID: 3401553) protocol was approved by the Institutional Review Board of Isfahan University of Medical Sciences (Ethics code: IR.MUI.MED.REC.1401.415). All procedures were conducted in accordance with the Declaration of Helsinki, and informed consent was obtained from all participants.

Statistical analysis method

Descriptive statistics were reported as mean $\pm$ standard deviation (SD) for continuous variables and frequency and percentage for categorical variables. Changes in examination scores over time were analyzed using repeated-measures analysis of variance (ANOVA). When the assumption of sphericity was violated, the Greenhouse–Geisser correction was applied. Additionally, repeated-measures analysis of covariance (ANCOVA) was performed to adjust for residency year and baseline performance as potential confounders. Pairwise comparisons between consecutive examinations were conducted using paired sample *t*-tests. Statistical significance was defined as a two-sided *P*-value < 0.05 , and exact *P*-values were reported. All analyses were performed using statistical package for the social sciences (SPSS) software (version 25.0; SPSS Inc., Chicago, IL, USA).

Results

In total, 20 third- and fourth-year radiology residents participated in this study. Table 1 shows the Specifications of the tests and interventions before each test and the average and standard deviation of the scores of each test. Table 2 shows the scores of each resident on each test and the average and standard deviation of each test.

As summarized in Table 1, mean examination scores increased progressively over time, accompanied by a gradual reduction in score variability. Repeated-measures ANOVA confirmed a significant time effect on performance (Greenhouse–Geisser corrected, $P < 0.001$). Pairwise comparisons [Table 2] showed significant improvements following structured training and feedback phases, whereas intervals without new educational input demonstrated relative score stabilization.

We performed Mauchly's Test of Sphericity, which was significant (P value = 0.000), followed by a significant Greenhouse–Geisser test (P value = 0.000). The abovementioned tests showed an existing time process regarding different grades in different exams. The results of paired sample *T*-tests are presented in Table 3. This test was performed to obtain the cut point

Table 1: The Specifications of the tests and interventions before each test, and the average and standard deviation of the scores of each test

	Date	Number of Cases	Test execution Method	Training Before the Test	Mean	Standard Deviation
Test 1	23/2/2024	20	Sent pictures	No training	17.80	13.687
Test 2	9/3/2024	20	Sent pictures	Three offline training classes	28.30	12.690
Test 3	15/3/2024	3	Sent pictures	No new training	57.20	20.529
Test 4	28/4/2024	3	Sent pictures	No new training	64.10	19.309
Test 5	31/5/2024	6	Sent pictures	Feedback and explanation of the previous test	58.90	16.973
Test 6	25/6/2024	6	Sent pictures	Feedback and explanation of the previous test	74.50	17.572
Test 7	13/8/2024	3	View on PACS	Feedback and explanation of the previous test	77.10	14.385

PACS: Picture archiving and communication system

Table 2: The scores of each resident in each test and the average and standard deviation of each test. The maximum attainable score was 100

Resident's Code	Grade	Ex 1	Ex 2	Ex 3	Ex 4	Ex 5	Ex 6	Ex 7
a	4	31	38	77	82	80	98	97
b	4	28	39	78	62	84	95	97
c	4	30	40	76	77	76	90	93
d	4	30	38	75	73	63	83	83
e	4	39	45	77	90	81	95	90
f	4	27	38	70	77	63	78	77
g	4	28	38	75	73	73	88	90
h	4	25	34	62	73	63	75	63
i	4	32	40	78	80	71	80	87
j	4	30	39	73	80	69	88	90
k	3	0	12	25	22	48	47	56
l	3	20	35	54	87	62	88	83
m	3	8	24	48	37	60	73	73
n	3	10	23	19	63	55	84	86
o	3	2	10	47	63	51	70	60
p	3	2	9.5	57	55	37	50	57
q	3	0	9.5	31	60	49	52	77
r	3	0	14	38	27	31	47	60
s	3	14	28	61	62	31	59	63
t	3	0	12	23	28	31	50	60
Mean		17.80	28.30	57.20	64.10	58.90	74.50	77.10
SD		13.687	12.690	20.529	19.309	16.973	17.572	14.385

Table 3: Comparison of consecutive exam scores using paired sample *t*-tests

Executive exams	Mean	SD	95% Confidence Interval of the Difference		<i>P</i>
			Lower	Upper	
Pair 1 Ex 1–Ex 2	–10.500	2.833	–11.826	–9.174	0.000
Pair 2 Ex 2–Ex 3	–28.900	12.138	–34.581	–23.219	0.000
Pair 3 Ex 3–Ex 4	–6.900	14.846	–13.848	0.048	0.051
Pair 4 Ex 4–Ex 5	5.200	13.851	–1.283	11.683	0.110
Pair 5 Ex 5–Ex 6	–15.600	7.373	–19.051	–12.149	0.000
Pair 6 Ex 6–Ex 7	–2.600	8.191	–6.434	1.234	0.172

or the pattern of progress, but the cut point was not obtained [Table 3].

As shown in Table 3, the difference between the mean of consecutive exam scores is statistically significant only for exams 1 and 2, exams 3, and exams 5 and 6, as there is a drop in the mean of exam scores in exam 5 compared to exam 4.

In panel A, individual educational charts for each resident participating in the project are shown [Figure 2]. In panel B, the average learning curve of the group is drawn [Figure 3]. Panel C shows the average curve of the group (the mean curve) with 95% confidence [Figure 3]. In panel D, changes in the variance of exam scores are depicted [Figure 4].

Discussion

The primary aim of this study was to evaluate the learning curve of female pelvic MRI interpretation

among senior radiology residents following a structured educational intervention. We hypothesized that repeated practice combined with targeted training and expert feedback would lead to both progressive improvement in performance and convergence of interpretive skills over time. Our results support this hypothesis. Repeated-measures ANOVA demonstrated a significant time effect, confirming a learning process across consecutive examinations. The significant increases in mean scores following structured training and feedback phases indicate progressive skill acquisition, while the eventual reduction in score variance reflects convergence toward a stable and acceptable level of competence, consistent with mastery learning theory. The transient decrease in mean score observed in the fifth examination likely reflects increased case complexity rather than regression in learning. This finding underscores the non-linear nature of learning curves and highlights the importance of repeated assessments over time rather than reliance on single performance metrics.

Due to the relatively small sample size, we could not stratify and analyze the scores based on the participants' residency year through ANCOVA analysis. Our study found that all the third- and fourth-year residents reached scores greater than 50 out of 100 after six tryouts. It should be noted that in our academy, third-year radiology residents usually have no training or skills to interpret pelvic MRIs.

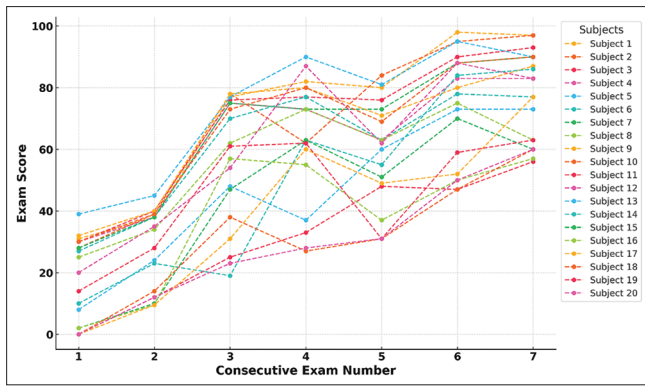


Figure 2: The trend of exam scores over time for individual participants (Panel A)

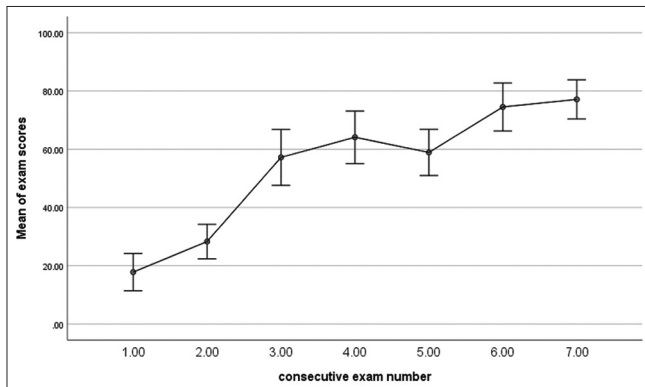


Figure 3: Changes in the mean of group scores over time (Panel B) with 95% confidence intervals (Panel C)

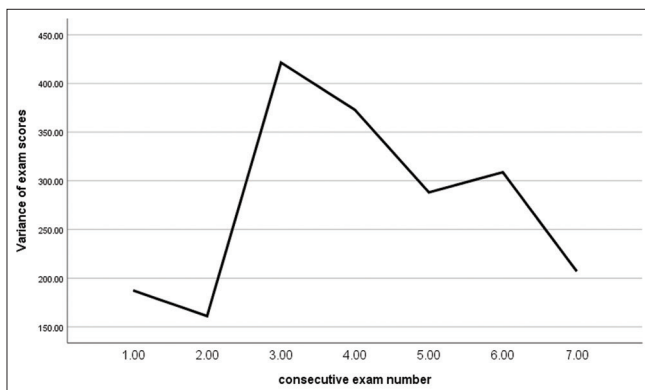


Figure 4: Changes in the variance of exam results over time (Panel D)

Previous studies have demonstrated the utility of learning curve analysis in medical imaging education. Like our findings, studies in radiology and ultrasound training have shown that structured educational interventions combined with repeated case interpretation significantly improve diagnostic performance over time. For example, prior research in radiology residents interpreting CT and MRI examinations has reported progressive improvement in accuracy following deliberate practice and feedback-based learning models. Moreover, competency-based educational frameworks

emphasize not only improvement in mean performance, but also reduction in inter-individual variability as indicators of skill mastery. Our observation of decreasing score variance in later examinations aligns with these models and suggests that the applied training protocol facilitates convergence toward uniform interpretive competence. Compared with prior studies that relied on cross-sectional assessments, the longitudinal design of the present study provides a more robust depiction of the learning process in pelvic MRI interpretation.

Another study by Indrielle-Kelly *et al.*^[11] compared the learning curves of an ultrasound trainee (obstetrics and gynecology resident) and a radiology trainee when assessing pelvic endometriosis. Consecutive patients with suspected endometriosis were prospectively enrolled and underwent an ultrasound and magnetic resonance imaging, which was reported according to the international deep endometriosis analysis (IDEA) group consensus. Trainees reported deep endometriosis (DE), endometriomas, frozen pelvis, and adenomyosis. Using the Kappa agreement, their findings were compared against laparoscopy/histology and expert findings. Their study showed a positive learning curve in some areas of pelvic endometriosis mapping after as few as 35 cases. Still, a more extensive caseload was required to demonstrate the curve fully. In our study, the number of cases in different exam trials was different, while the increasing scores over exam trials, even by a few cases [Table 2] implies that time interval and repetition of trials may be as important as the number of encountered cases in the advancement of learning curves.

Determining if MRI in diagnosing endometriosis is related to a radiologist's expertise was the subject of another study by Saba *et al.*^[12] This study was compliant with the STARD method. Thirty patients (mean age 34; range 21–45 years), who had undergone an MRI study for suspected endometriosis underwent surgery and were retrospectively evaluated. Sensitivity, specificity, and accuracy for the diagnosis of ovarian problems in the first analysis were 88.9, 87, and 88%; in the second, 92.6, 87, and 90%, whereas in the third, 92.6, 91.3, and 92%. Sensitivity, specificity, and accuracy for the USLs in the first analysis were 62.5, 76.9, and 70%; in the second, 72, 80.8, and 76%, whereas in the third, 80, 84.6, and 82%. Sensitivity, specificity, and accuracy for the diagnosis of vaginal fornix problems in the first analysis were 63.2, 64.5, and 64%; in the second, 73.7, 77.4, and 76%, whereas in the third, 73.7, 83.9, and 80%. Sensitivity, specificity, and accuracy for the rectum\sigma\douglas (RSD) in the first analysis were 39.1, 81.5, and 62%; in the second, 62.5, 85.2, and 76%, whereas in the third, 73.9, 88.9, and 82%. The McNemar test indicated a significant statistical difference in sensitivity in detecting nodules

of endometriosis in RSD between the first and third analysis ($P = 0.0215$). The authors concluded that the accuracy of MRI in diagnosing endometriosis increased with the radiologist's expertise, and the improvement was statistically significant in determining RSD involvement. Unlike Saba, we did not assess the specificity and sensitivity of participants' reports according to the other diagnostic modalities and used the examiners' interpretation as the standard for comparison and scoring. This may limit the value of the study protocol to the expertise of examiners.

Our findings align with and extend the results reported by Stanzone *et al.*, who analyzed the learning curve of radiology residents interpreting MRI for deep myometrial invasion in endometrial cancer staging. They demonstrated a progressive improvement in diagnostic accuracy and inter-observer agreement after structured, repetition-based training. Like their results, our study showed a steady rise in interpretation scores and a decrease in variance across consecutive trials, indicating skill convergence among trainees. However, while Stanzone's study focused on a specific oncologic diagnostic task within endometrial cancer, our protocol encompassed a broader range of pelvic MRI cases (uterine, ovarian, and cervical), providing a more generalizable assessment of radiologic competency development. This comparison underscores that repeated exposure, standardized feedback, and competency-based evaluation constitute effective pillars of radiology education across different imaging contexts.^[10]

This study has several strengths. First, it employed a longitudinal repeated-measures design, allowing precise evaluation of individual and group learning trajectories over time rather than relying on cross-sectional comparisons. Second, the use of standardized assessments, identical scoring criteria, and a single blinded rater enhanced internal consistency and reduced measurement bias. Third, in addition to improvement in mean scores, the study uniquely examined changes in score variance, providing insight into skill convergence and stabilization, which is rarely addressed in radiology education research. However, some limitations should be acknowledged. The study was conducted at a single academic center with a relatively small sample size, which may limit generalizability. In addition, the absence of a parallel control group restricts causal inference regarding the effectiveness of the intervention. Finally, test intervals were not uniform due to scheduling constraints, which may have influenced the pace of learning. Despite these limitations, the study provides a practical and reproducible framework for evaluating learning curves in advanced imaging education.

Conclusions

This study demonstrates that a structured, competency-based training protocol incorporating repeated assessments and group feedback enables senior radiology residents to achieve stable and acceptable proficiency in pelvic MRI interpretation, regardless of baseline performance or residency year. From a health policy perspective, such standardized educational interventions can contribute to improving the quality and consistency of diagnostic imaging services, particularly in settings with limited case volume and training opportunities. Implementing similar protocols within residency curricula may support workforce capacity building, optimize educational resources, and promote uniform diagnostic standards across training centers. Future studies are warranted to evaluate the scalability and long-term impact of this model on clinical performance and patient care outcomes.

Informed consent

All patients provided informed consent before participation in this study.

Acknowledgments

The authors would like to thank all radiology residents who participated in this study for their time and commitment. We also appreciate the support of the Radiology Department of Al-Zahra Hospital, Isfahan University of Medical Sciences, for facilitating the educational sessions and assessments. The authors are grateful to the faculty members who contributed to the organization of the training program and provided constructive educational feedback throughout the study.

Abbreviations

MRI: Magnetic Resonance Imaging; ANOVA: Analysis of Variance; ANCOVA: Analysis of Covariance; SPSS: Statistical Package for the Social Sciences; USA: United States of America; SD: Standard Deviation; IDEA: International Deep Endometriosis Analysis; DE: Deep Endometriosis; R.S.D.: Rectum\Sigma\Douglas.

Financial support and sponsorship

This study received no grants or funding support.

Conflict of interest

The authors declare no conflicts of interest.

References

1. Howard NM, Cook DA, Hatala R, Pusic MV. Learning curves in health professions education simulation research: A systematic review. *Simul Healthc* 2021;16:128-35.
2. Blehar DJ, Barton B, Gaspari RJ. Learning curves in emergency ultrasound education. *Acad Emerg Med* 2015;22:574-82.
3. Pusic MV, Boutis K, Pecaric MR, Savenkov O, Beckstead JW,

- Jaber MY. A primer on the statistical modelling of learning curves in health professions education. *Adv Health Sci Educ Theory Pract* 2017;22:741-59.
4. Ratner ES, Staib LH, Cross SN, Raji R, Schwartz PE, McCarthy SM. The clinical impact of gynecologic MRI. *AJR Am J Roentgenol* 2015;204:674-80.
 5. Somberg Gunther M, Kanmaniraja D, Kobi M, Chernyak V. MRI of acute gynecologic conditions. *J Magn Reson Imaging* 2020;51:1291-309.
 6. Romanski PA, Aluko A, Bortoletto P, Troiano RN, Pfeifer SM. Aqueous vaginal contrast and scheduled hysteroscopy with magnetic resonance imaging to delineate complex müllerian anomalies. *Fertil Steril* 2022;117:221-3.
 7. Kitajima K, Suenaga Y, Ueno Y, Kanda T, Maeda T, Makihara N, *et al.* Value of fusion of PET and MRI in the detection of intra-pelvic recurrence of gynecological tumor: comparison with 18 F-FDG contrast-enhanced PET/CT and pelvic MRI. *Ann Nucl Med* 2014;28:25-32.
 8. Grueneisen J, Beiderwellen K, Heusch P, Gratz M, Schulze-Hagen A, Heubner M, *et al.* Simultaneous positron emission tomography/magnetic resonance imaging for whole-body staging in patients with recurrent gynecological malignancies of the pelvis: a comparison to whole-body magnetic resonance imaging alone. *Investigative radiology*. 2014;49(12):808-15.
 9. Allen BC, Hosseinzadeh K, Qasem SA, Varner A, Leyendecker JR. Practical approach to MRI of female pelvic masses. *AJR Am J Roentgenol* 2014;202:1366-75.
 10. Stanzione A, Cerrone F, Ferraro F, Menna F, Spina A, Danzi R, *et al.* Training radiology residents to evaluate deep myometrial invasion in endometrial cancer patients on MRI: A learning curve study. *Eur J Radiol* 2024;111:546.
 11. Indrielle-Kelly T, Fischerova D, Hanuš P, Frühauf F, Fanta M, Dundr P, *et al.* Early learning curve in the assessment of deep pelvic endometriosis for ultrasound and magnetic resonance imaging. *Biomed Res Int* 2020;2020:8757281.
 12. Saba L, Guerriero S, Sulis R, Pilloni M, Ajossa S, Melis G, *et al.* Learning curve in the detection of ovarian and deep endometriosis by using magnetic resonance: Comparison with surgical results. *Eur J Radiol* 2011;79:237-44.