Journal of Medical Education Development
e-ISSN: 2980-7670
https://edujournal.zums.ac.ir
Volume 19, Issue 3
July 2026
Pages: 126–131

Brief Report

A Pilot Educational Evaluation of a Female-Oriented Medical Emergencies Training Program

Alireza Irandegani, Amin Habibi, Masomeh Kalantarion *

Department of Medical Education, School of Medical Education and Learning Technologies, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Article info

Article history:
Received 30 Nov. 2025
Revised 1 Jan. 2026
Accepted 28 Apr. 2026
Published 1 Jul. 2026

***Corresponding author:**
Masomeh Kalantarion, Department of Medical Education, School of Medical Education and Learning Technologies, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
Email: kalantarion65@gmail.com

How to cite this article:
Irandegani A, Habibi A, Kalantarion M. A Pilot Educational Evaluation of a Female-Oriented Medical Emergencies Training Program. *J Med Edu Dev*. 2026;19(3):126-131.

Abstract

Background & Objective: Educational evaluation is essential for ensuring the effectiveness of skill-based training programs in medical education. Innovative emergency medical services (EMS) courses require structured evaluation to determine whether educational objectives are achieved. This pilot study reports a preliminary educational evaluation of the first female-oriented medical emergencies training program using Kirkpatrick's model with a focus on feasibility and initial acceptability rather than definitive effectiveness.

Materials & Methods: This descriptive cross-sectional study evaluated the program at Kirkpatrick Levels 1 (Reaction) and 2 (Learning). All six female trainees enrolled in the program participated. Reaction was assessed through semi-structured interviews analyzed using qualitative content analysis. Learning outcomes were evaluated using a validated 100-item written examination and a five-station mini-OSCE. Descriptive statistics were used for quantitative data analysis.

Results: At the Reaction level, qualitative analysis identified three main categories related to educational quality: perceived relevance of course content, effectiveness of teaching methods, and adequacy of the learning environment. Participants reported high satisfaction with practical, simulation-based instruction and instructor support. At the Learning level, all trainees passed the written examination (mean score: 74.8/100) and the mini-OSCE (mean score: 82.3/100), indicating acceptable post-program knowledge and skill performance.

Conclusion: This pilot evaluation provides preliminary evidence that the evaluated training program appears feasible and was well-received at the Reaction and Learning levels of Kirkpatrick's model among this small cohort. The findings show preliminary promise and support the use of structured educational evaluation frameworks for assessing innovative, skill-based EMS training programs. However, due to the lack of a pretest and the small sample size, causal conclusions about learning cannot be drawn; these results should be interpreted as descriptive findings of post-program performance.

Keywords: educational evaluation; Kirkpatrick model; emergency medical services education; pilot study

Introduction

Medical emergency services constitute a fundamental pillar of healthcare systems, providing rapid and life-saving responses to accidents, acute illnesses, natural disasters, and unforeseen crises.

The effectiveness of pre-hospital emergency care is highly dependent on the competency and preparedness of emergency medical technicians (EMTs), whose timely interventions often determine patient outcomes. Consequently, the quality of training provided to emergency personnel is directly associated with reduced mortality and morbidity rates, as well as improved public trust and satisfaction with healthcare services. For this reason, many developed countries have placed strategic emphasis on investing in structured and competency-based training programs for EMTs within their national health education systems [1].

Equitable access to prompt and high-quality pre-hospital emergency services is a fundamental requirement of any healthcare system. However, in certain regions of Iran, sociocultural, religious, and geographical factors have created challenges in delivering emergency care to female patients. In some situations, the exclusive availability of male EMTs may lead to delays in service delivery or reduced utilization of emergency services, potentially resulting in adverse health outcomes. These contextual challenges highlight the need for alternative workforce strategies to ensure timely and culturally appropriate emergency care for all population groups [2]. In response to these needs, Iranshahr University of Medical Sciences designed and implemented the first formal medical emergencies training program specifically for female recruits. The implementation of such a novel program necessitates an initial, in-depth pilot evaluation. While the establishment of this training course represents an important educational innovation, the mere launch does not ensure its effectiveness. Systematic educational evaluation, particularly in this pilot phase, is required to determine whether the program has achieved its intended preliminary objectives in terms of learner satisfaction, knowledge acquisition, and skill development, and to inform its refinement. Evaluation of training programs is widely recognized as a cornerstone of continuous educational improvement. Without structured assessment of learning processes and outcomes, it is not possible to make informed judgments about program effectiveness or identify areas requiring refinement. Educational evaluation enables educators to review achievements, recognize limitations, optimize instructional strategies, and support evidence-informed decision-making in curriculum development [3]. The application of standardized evaluation frameworks, such as the Kirkpatrick model, allows for a more comprehensive and structured analysis of training effectiveness. This model facilitates evaluation across multiple levels, including learners' reactions, learning outcomes, behavioral changes, and broader results [4]. In a pilot evaluation of a newly implemented educational program, assessment at the Reaction and Learning levels is particularly appropriate, as it provides foundational evidence of feasibility and early effectiveness. Several national studies have examined aspects of emergency and clinical training quality. For example, Kheirkhah et al. reported moderate levels of satisfaction among nursing students regarding clinical training in internal emergency departments [5]. Similarly, Bolverdi et al. found relative satisfaction among emergency medicine trainees but emphasized the need for educational redesign based on actual clinical demands [6]. However, while studies have evaluated general EMS training, no study has conducted a pilot evaluation of a female-oriented program using a structured framework like Kirkpatrick's from the participants' perspective. This gap underscores the need for focused evaluation to guide the early development of such targeted initiatives. Accordingly, this pilot study aimed to conduct an initial evaluation of the first female-oriented medical emergencies training program using Kirkpatrick's model. The primary objectives were to assess the feasibility, participant reaction (Level 1), and post-program knowledge and skill performance (Level 2) as descriptive outcomes to inform immediate program refinement and guide the planning of future larger-scale studies

Materials & Methods

Design and setting(s)

This descriptive cross-sectional study was conducted as part of an approved thesis at Shahid Beheshti University of Medical Sciences. The training program was designed and implemented by the corresponding author (A.I.) who serves as the instructor for women's emergency training at Iranshahr University of Medical Sciences, where the course was delivered. The study was conducted from October 31, 2025 to April 20, 2026.This was a pilot evaluation of the inaugural cohort of the first female-oriented medical emergencies training program at Iranshahr University of Medical Sciences, using Kirkpatrick's model. The census sample (N=6) allowed for a comprehensive, in-depth mixed-methods assessment of the program's initial implementation, providing foundational data for future iterations and larger-scale studies.

Participants and sampling

The study population consisted of all six female students admitted to the medical emergencies program through the national entrance examination at Iranshahr University of Medical Sciences. Given the limited cohort size, a census sampling approach was applied, including all eligible participants.

Tools/Instruments

Three instruments were used for data collection. Participants' reactions to the course were explored using a semi-structured interview guide with open-ended questions, corresponding to Kirkpatrick's Level 1. The

content validity of the guide was confirmed by three emergency medicine specialists. Learning at the cognitive level was assessed using a 100-item multiple-choice written examination developed based on the course objectives and validated through blueprinting and expert review. Practical learning at Kirkpatrick's Level 2 was assessed using five station-specific mini-OSCE checklists, including an additional station on childbirth management. The content validity of these checklists was confirmed by experts. Reliability coefficients were not calculated due to the small pilot sample.

Data collection methods

Following ethical approval (IR.SBMU.SME.REC.1404.040) and obtaining informed consent from all participants, data collection was conducted at two levels of the Kirkpatrick model. To assess Level 1 (reaction), semi-structured interviews lasting 20–30 minutes were conducted, audio-recorded, and transcribed verbatim. To assess Level 2 (learning), the written examination was administered at the end of the course under standardized conditions. In addition, practical performance was evaluated through a five-station mini-OSCE, in which trained assessors used structured checklists to rate participants' performance.

Data analysis

Data analysis was guided by Kirkpatrick's evaluation framework, focusing on the Reaction and Learning levels. Qualitative interview data were analyzed using Graneheim and Lundman's qualitative content analysis approach [7]. Transcripts were independently coded by two researchers (A.I., A.H.). Inter-coder agreement was calculated as 88% initially; disagreements were discussed and resolved through consensus. To enhance credibility, a member check was conducted: the extracted codes and categories were presented to all six participants, and all of them confirmed the accuracy and alignment of the findings with their experiences. Additionally, peer debriefing was performed by a third researcher (M.K.) who reviewed the coding and categorization process. Codes were compared, discussed, and categorized into themes through an iterative consensus process.

Results

Findings are presented according to Kirkpatrick's evaluation framework, focusing on Level 1 (Reaction) and Level 2 (Learning), and are intended to provide preliminary descriptive evidence regarding the feasibility and initial outcomes of the program.

Kirkpatrick Level 1: Reaction

Trainee Demographic Characteristics Analysis of demographic data showed that the mean age of trainees was 24 years. All participants were single and resided in counties covered by Iranshahr University of Medical Sciences. The mean grade point average (GPA) from the previous semester was 17.2 out of 20. None of the trainees had prior experience in similar emergency medical training programs. Regarding interest in the medical emergencies profession, assessed using a three-point Likert scale, 83.3% (n = 5) reported high interest, while 16.7% (n = 1) reported moderate interest.

Qualitative Analysis of Trainee Interviews All findings reported below were confirmed by participants during the member check process. Qualitative content analysis of semi-structured interviews yielded comprehensive insights into the program's feasibility and acceptability. For the main text, we highlight the most salient categories and supporting evidence (See Appendix for the complete qualitative framework). The analysis underscored high participant satisfaction with the program's relevance and teaching approach, and identified key areas for practical refinement, which are central objectives of a pilot study. Participants emphasized the contextual relevance of the curriculum and its alignment with job needs. One trainee (P3) stated, "Training in CPR techniques and hemorrhage management was exactly what we need at an accident scene."

The creation of a safe and supportive learning environment was a particularly valued aspect of program implementation. One trainee (P6) illustrated this: "As a woman, I felt more comfortable asking questions, even basic ones, from our female instructor." This directly speaks to the program's acceptability and feasibility within its specific cultural context. Participants also reported significant gains in personal confidence and professional motivation. One trainee (P5) explained, "Before this course, I could only watch during emergencies. Now I feel I have the knowledge and courage to intervene effectively." Concurrently, the analysis provided clear direction for improvement, identifying challenges related to logistical resources and assessment methods. For instance, a participant (P1) noted, "During some practical sessions, we lost significant time due to outdated equipment, which reduced training efficiency." These findings are crucial

for informing the refinement phase following this pilot evaluation.

Kirkpatrick Level 2: Learning

Final Written Examination Results Quantitative findings related to post-program knowledge performance are summarized in Table 1. All trainees successfully passed the 100-item multiple-choice written examination. The mean score was 74.80 out of 100 (SD = 5.20), indicating acceptable theoretical knowledge performance at the end of the course (**Table 1**).

Table 1. Distribution of Trainees' Written Examination Scores

Statistical Indicator	Value
Number of Participants	6
Mean Score	74.80
Standard Deviation	5.20
Median	74.00
Minimum Score	68.00
Maximum Score	83.00
Passing Score	50.00
Passing Rate	100%

The highest score (83/100) was obtained by Participant 2, while the lowest score (68/100) belonged to Participant 5, reflecting a narrow score range among trainees.

Practical Examination (mini-OSCE) Results : Post-program practical skill performance demonstrated adequate mastery across evaluated competencies. The mean mini-OSCE score was 82.30 out of 100 (SD = 4.80), indicating relatively consistent performance among trainees (**Table 2**).

Table 2. Overall Trainee Performance in mini-OSCE

Statistical Indicator	Score (/100)
Mean	82.30
Standard Deviation	4.80
Minimum Score	76.00
Maximum Score	87.00

Analysis of station-level performance revealed the highest mean scores in cardiopulmonary resuscitation (17.5/20) and the lowest in childbirth management (15.9/20), suggesting a specific area for potential refinement of practical training emphasis—a key output of a pilot study (**Table 3**).

Table 3. Trainee Performance Across Different OSCE Stations

Station	Mean Score (/20)
Cardiopulmonary Resuscitation	17.50
Injections	17.10
Bandaging and Splinting	16.80
Airway Management	16.20
Childbirth Management	15.90

Discussion

This pilot evaluation of the first female-oriented EMS training cohort provides preliminary descriptive findings regarding program feasibility, high learner acceptability, and satisfactory post-program knowledge and skill performance. The qualitative findings indicated high learner satisfaction with the relevance of course content, teaching quality, and the supportive learning environment. The importance of such a safe and supportive environment for female trainees is well-documented [8], and participants' reported increase in self-confidence suggests the program appears to have fostered an empowering climate. Acceptable post-program knowledge and skill acquisition were observed in written and practical examination results, which is consistent with the active, simulation-based training strategies employed.

The program also addresses a recognized systemic need for female emergency responders in culturally sensitive settings. The findings, while preliminary, suggest such targeted initiatives may be feasible and valuable within the national context, aligning with recommendations for gender-responsive training pathways [9]. Participant-identified challenges, such as limited simulation equipment and high content density, offer direct, actionable insights for refining the curriculum and logistical planning in future iterations.

This study has several limitations, most notably the lack of a pretest. Without a pretest, we cannot draw causal conclusions about learning or attribute post-program performance to the training itself; our findings only reflect post-program performance. Additionally, the evaluation of a single, small cohort (N=6) limits statistical generalizability but was ideal for an in-depth, mixed-methods analysis of this first iteration. The small sample size means that these findings are preliminary and require replication in larger cohorts before any broader generalization can be considered. The findings are also limited to immediate post-course outcomes (Kirkpatrick Levels 1 & 2). Furthermore, social desirability bias may influence self-reported interview data. However, these limitations are balanced by the strengths of a comprehensive pilot evaluation, including rich qualitative data and detailed assessment of all initial participants, which provides a robust foundation for program development. Future work should build on this foundation by extending evaluation to Kirkpatrick Levels 3 (Behavior) and 4 (Results) to assess long-term skill transfer and organizational impact, and by including pretest assessments and larger sample sizes.

Conclusion

This pilot study provides preliminary descriptive findings that the first cohort of a women-specific medical emergency training program was feasible and well-received among this small cohort at Kirkpatrick's Reaction and Learning levels. The positive preliminary findings, including high learner satisfaction and acceptable post-program knowledge and skill performance, show preliminary promise and support the continued development of this training program. They highlight specific areas for refinement (e.g., equipment access, childbirth training emphasis).

Crucially, this pilot establishes a methodology and baseline for future work, which should involve larger cohorts, comparative designs, pretest-posttest designs, and longitudinal follow-up to assess skill retention and clinical impact (Kirkpatrick Levels 3 & 4). Due to the lack of a pretest and the small sample size, these findings should be interpreted as descriptive and hypothesis-generating rather than confirmatory of effectiveness.

Ethical considerations

Ethical approval for this study was obtained from the Shahid Beheshti University of Medical Sciences ethics committee (Code: IR.SBMU.SME.REC.1404.040). The study was conducted in accordance with the principles of the Declaration of Helsinki.

The purpose of the study was thoroughly explained to all participants, and written informed consent was obtained from each individual. Participants were assured of the confidentiality of their data and their right to withdraw from the study at any time without any consequence.

Artificial intelligence utilization for article writing

The authors declare that no generative AI tools (e.g., ChatGPT) were used in the writing of the intellectual content of this manuscript. All authors take full responsibility for the integrity and accuracy of the work presented.

Acknowledgment

The researchers would like to express their sincere gratitude to all the female participants in the medical emergencies training course for their valuable time and cooperation. We also extend our thanks to the officials and staff of Iranshahr University of Medical Sciences for their support in conducting this course.

Conflict of interest statement

The authors declare that there is no conflict of interest.

Author contributions

Alireza Irandegani, Amin Habibi, and Masomeh Kalantarion contributed to the conceptualization and investigation of the study, with Amin Habibi and Masomeh Kalantarion specifically focusing on the methodology. Alireza Irandegani was responsible for data curation, while formal analysis was carried out by Alireza Irandegani and Amin Habibi. The manuscript's original draft was prepared by Alireza Irandegani and Amin Habibi, and subsequently reviewed and edited by all authors. Project administration was overseen by Masomeh Kalantarion. All authors have read and approved the final version of the manuscript

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

The datasets generated and analyzed during the current study are not publicly available due to the confidentiality agreements and ethical considerations related to the participants' privacy. However, anonymized data are available from the corresponding author upon reasonable request and with the permission of the relevant ethics committee.

References

1. Lerner EB, Moscati RM. The golden hour: scientific fact or medical “urban legend”? *Acad Emerg Med*. 2001;8(7):758–760. https://doi.org/10.1111/j.1553-2712.2001.tb00201.x
2. Rothschild SK. Cross-cultural issues in primary care medicine. *Dis Mon*. 1998;44(7):293–319. https://doi.org/10.1016/s0011-5029(98)90010-4
3. Alsyaibany RR. The evaluation of effectiveness on education and training program. *J Educ Learn Stud*. 2019;2(1):56. https://doi.org/10.32698/0662
4. Kirkpatrick JD, Kirkpatrick WK. *Kirkpatrick’s four levels of training evaluation*. Alexandria (VA): Association for Talent Development; 2016. https://books.google.com/books?id=mo--DAAAQBAJ
5. Kheirkhah D, Haji Hadian Z, Dahim A, Yavari M, Azadchehar MJ, Sharif MR, et al. Medical interns’ satisfaction with the quality of clinical training in the

internal emergency department of Shahid Beheshti Hospital in Kashan, Iran. *Dev Strateg Med Educ*. 2023;9(4):62–71. https://doi.org/10.32598/DSME.9.4.515.1

6. Bolvardi E, Moosavi MS, Mousavi SM, Niroumand S, Zarghi S, Roudi F, et al. The satisfaction level of Mashhad University of Medical Sciences graduates with emergency medicine training. *Horizon Med Educ Dev*. 2026;17(2):28–37. https://doi.org/10.22038/hmed.2025.83977.1438
7. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Educ Today*. 2004;24(2):105–112. https://doi.org/10.1016/j.nedt.2003.10.001
8. McCann-Pineo M, Keating M, McEvoy T, Schwartz M, Schwartz RM, Washko J, et al. The female emergency medical services experience: a mixed methods study. *Prehosp Emerg Care*. 2024;28(4):626–634. https://doi.org/10.1080/10903127.2024.2306248
9. World Health Organization. Fair share for health and care: gender and the undervaluation of health and care work [Internet]. Geneva: WHO; 2024 [cited 2025 Aug 12]. Available from: https://www.who.int/publications/i/item/9789240082854