



Contents lists available at ScienceDirect

European Journal of Obstetrics & Gynecology and Reproductive Biology

journal homepage: www.journals.elsevier.com/european-journal-of-obstetrics-and-gynecology-and-reproductive-biology

Review article

Leadership training in obstetric care: A systematic review of educational interventions and outcomes

Sara Tameish^{a,b,c,d,*}, Pere Cavallé Busquets^{a,d}, Óscar Martínez Pérez^{b,c,e}^a Department of Obstetrics and Gynecology, Hospital Universitari Sant Joan de Reus, Reus, Spain^b Grupo Español de Seguridad Obstétrica (GESO), Spain^c Universidad Católica San Antonio de Murcia (UCAM), Murcia, Spain^d Universitat Rovira i Virgili, Tarragona, Spain^e Department of Obstetrics and Gynecology, Hospital HM Madrid Río, Madrid, Spain

ARTICLE INFO

Keywords:

Leadership
Obstetrics
Maternity care
Team performance
Medical education

ABSTRACT

Objective: To evaluate the structure, implementation and reported outcomes of leadership training interventions in obstetric care.

Methods: Following PRISMA guidelines and PROSPERO registration (CRD420251248483), a systematic search of PubMed, Web of Science, and Scopus was conducted for English-language studies published between 2005 and 2025. Eligible studies evaluated leadership training interventions involving healthcare professionals working in obstetric settings, including obstetricians, midwives, nurses, anesthesiologists, and residents. Randomized, non-randomized, and pre-post interventional studies were included. Outcomes were categorized according to Kirkpatrick's four-level evaluation model. Methodological quality was assessed using the Medical Education Research Study Quality Instrument (MERSQI). Due to heterogeneity, findings were synthesized descriptively.

Results: Ten studies met inclusion criteria. Interventions varied substantially in duration, format, and educational approach, with most incorporating simulation-based or multimodal strategies. Most interventions targeted individual learners (70%), with only two studies involving complete obstetric teams trained together. Outcome assessment relied predominantly on self-reported, non-validated questionnaires. Most studies reported improvements in perceived leadership knowledge, confidence, and skills. However, objective behavioral change was infrequently measured, and no study evaluated patient-level outcomes. The overall methodological quality was moderate (mean MERSQI score 10.1).

Conclusion: Leadership training in obstetric care is associated with improvements in self-reported competencies. However, current evidence is limited by methodological weaknesses, a predominance of subjective self-reported outcomes, scarce objective or behavioral assessments and absence of evaluation of patient-level outcomes. More rigorous, multidisciplinary, team-based programs are needed to establish clinical impact.

Introduction

Leadership is a fundamental component of effective team performance, ensuring clear communication, coordinated decision-making and a strong safety culture [1–4]. In healthcare, leadership is defined as the ability to influence a group of individuals to achieve a common goal by establishing a shared vision [1,5–8]. Ineffective leadership, particularly during transitions of care or critical obstetric emergencies, has been linked to an increased risk of medical errors and adverse maternal and neonatal outcomes [3,9]. Communication failures have

been identified as a leading cause of adverse perinatal outcomes, contributing to 72% of such events [10]. Reports from the “Mothers and Babies: Reducing Risk through Audits and Confidential Enquiries” (MBRRACE) in the United Kingdom highlight that poor interprofessional collaboration and deficiencies in leadership significantly contribute to maternal mortality and morbidity [11,12].

Leadership is increasingly recognized as a non-technical skill that can be developed through education and training [13,14]. Consequently, the integration of leadership development into both undergraduate and postgraduate medical education has become essential to

* Corresponding author at: Department of Obstetrics and Gynecology Hospital Universitari Sant Joan de Reus, Avinguda Josep Laporte 2, 43204 Reus (Tarragona), Spain.

E-mail address: sara.tameish@salutsantjoan.cat (S. Tameish).

<https://doi.org/10.1016/j.ejogrb.2026.115184>

Received 3 March 2026; Received in revised form 22 April 2026; Accepted 10 May 2026

Available online 19 May 2026

0301-2115/© 2026 Elsevier B.V. All rights reserved, including those for text and data mining, AI training, and similar technologies.

prepare clinicians to effectively lead teams, make high-stakes decisions and manage complex healthcare systems [5,7,15–19]. Leadership training programs have been implemented across various medical specialties and are linked to positive outcomes at both individual and team levels, including higher job satisfaction and retention, reduced staff burnout and fewer medication errors [19–21]. Leadership training in medicine has been associated with improved patient related outcomes, particularly in high-acuity settings such as stroke units and trauma care [22]. Leadership interventions have been shown to enhance communication, teamwork and overall organizational efficiency [14,18,19,21,22].

Despite its recognized importance, leadership training remains underemphasized in obstetric education. Most residency programs continue to prioritize technical and clinical competencies, often at the expense of non-technical skills such as leadership. As a result, many healthcare professionals of the obstetric field still develop leadership competencies informally, leading to variability in leadership practice and team dynamics [2,3,5,23].

This gap is particularly concerning given the unique demands of obstetric care. Obstetrics is a medical specialty in which multiple disciplines work closely together, including midwives, obstetricians, anesthesiologists, and nurses. It is also characterized by high cognitive workload, frequent emergencies, and the need for rapid, coordinated decision-making. In this context, effective leadership is essential to ensure high-quality care and maintain patient safety.

Yet, despite this critical need, it remains insufficiently understood how leadership training in obstetrics is currently being implemented; specifically, who these programs are designed for, which competencies they aim to develop, and how their effectiveness is evaluated. Furthermore, there is limited evidence linking these training initiatives to tangible outcomes, whether in terms of team performance or improvements in maternal and neonatal health.

Objective

The objective of this systematic review is to comprehensively analyze existing leadership development programs in obstetric care. This includes examining their structure and content, target populations, evaluation methodologies and reported outcomes. To the best of our knowledge, this is the first systematic review specifically focused on leadership training within the field of obstetrics.

Methods

This systematic review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [24] and is registered in the PROSPERO online database (PROSPERO Identifier CRD420251248483).

The quality of the included studies was assessed using the Medical Education Research Study Quality Indicator (MERSQI) [25]

a) Search strategy and screening

For this systematic review, we conducted a search in PubMed/MEDLINE, Scopus and Web of Science using a predefined strategy. Additionally, we performed a manual hand-searching of references from relevant leadership journals (*BMJ Leader*, *Journal of Healthcare Management* and *Leadership Quarterly*) as well as snowballing of reference lists from included studies. Gray literature; including conference abstracts, editorials, interviews, lectures, letters, news articles, guidelines, and methodology paper; was excluded from the review.

The search strategy employed a combination of controlled vocabulary (e.g., MeSH terms) and free-text keywords focused on leadership, obstetrics, maternity care, and team performance. The search was limited to English-language articles published between 2005 and 2025. A sample search for PubMed is available in Supplementary File 1.

All titles and abstracts were screened independently by two reviewers (S.T, P.C.B). Full-text articles of potentially eligible studies were

then retrieved and assessed based on predefined inclusion and exclusion criteria. Any discrepancies were resolved through consensus or by consulting a third reviewer (O.M.P). The study selection process is detailed in the PRISMA flow diagram (Fig. 1).

b) Inclusion criteria

Eligible studies included studies evaluating leadership training interventions involving health care professionals working in obstetric care settings such as labor and delivery units, antenatal or postnatal wards, or maternity services (including simulation centers). Study designs included randomized and non-randomized interventional studies and observational studies with pre-intervention and post-intervention assessments. The study population needed to consist of licensed or registered healthcare professionals actively involved in obstetric care, including obstetricians, midwives, nurses, anesthesiologists, or other relevant practitioners. Residents in those areas were also included. Additionally, eligible studies had to evaluate outcomes related to the intervention either satisfaction, leadership knowledge and/or skill adequation, team performance, such as communication, collaboration, leadership behavior, or decision-making efficiency, or to patient-level outcomes, including maternal or neonatal morbidity or mortality, safety metrics, or other indicators of quality of care.

Studies were excluded if they focused exclusively on students or non-licensed personnel; assessed leadership in non-obstetric contexts; did not report measurable outcomes related to individual, team or patient care; were opinion articles, editorials, narrative commentaries or case reports lacking empirical data.

c) Data extraction and analysis

Data from the included studies were summarized using a standardized template. The data fields included: first author and year, study design, setting, specialty and number of participants, level of participant, multidisciplinary or individual training, intervention, leadership topics, training modality, trainer profile and assessment of outcomes (Table 1).

To classify the outcomes, we applied Kirkpatrick's four-level model of training evaluation, which distinguishes between: participants' reactions (Level 1), knowledge acquisition (Level 2), behavioral change (Level 3), and broader system or organizational impact (Level 4) [26]. Following guidance from previous reviews in the leadership development literature, we further divided these categories into subjective and objective outcomes [14,18]. See Table 3. As a result, we applied a seven-tiered classification system to map evaluation findings: (1) Reaction/satisfaction; (2A) Subjective knowledge gain; (2B) Objective knowledge acquisition; (3A) Self-reported behavior or leadership skill use; (3B) Observed or measured behavior change; (4A) Perceived system-level impact; and (4B). Measurable outcomes at the organizational or patient level (see Table 2).

d) Study quality appraisal

According to the recommendations from Geerts *et al* [19], Rosenman *et al* [21] and Lyons *et al* [14], we assessed study quality using the Medical Education Research Study Quality Indicator (MERSQI), a tool designed to measure the methodological quality of quantitative medical education research studies [25,29]. This tool consists of 10 items in six domains related to design, sampling, type of data collected, validity of evaluation, analysis and outcomes. Each domain is scored to a maximum of 3, for a total of 5–18 (Supplement 2). According to previous review using this score [14,19], a score of 12 or higher is considered a high reliability study. The study quality was evaluated by two reviewers (S.T and O.M.P). Any discrepancies were resolved by a third reviewer (P.C.B).

Results

a) Study selection

A total of 298 records were identified through the initial database search. After removing duplicates, 228 records remained for title and abstract screening. Of these, 197 were excluded for not meeting the

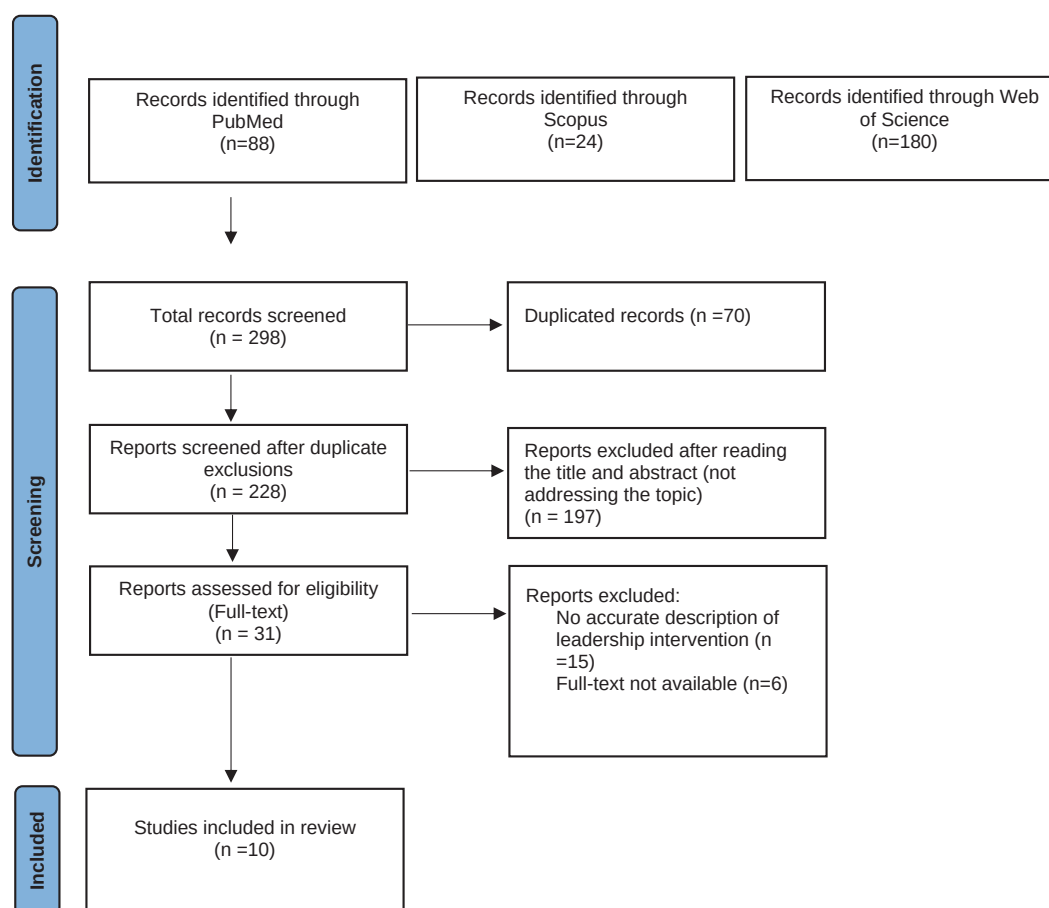


Fig. 1. PRISMA flow diagram.

inclusion criteria based on titles and abstracts. Full texts of the remaining 31 articles were then assessed for eligibility. Finally, 10 studies were included in this review, having fulfilled inclusion criteria (Fig. 1).

b) Study characteristics

The characteristics of the ten included studies are summarized in Table 1 and Table 2. The most common study designs were descriptive studies (n = 4; 40%) [30–33], followed by single group pre-posttest studies (n = 3; 30%) [32–34] and single group posttest studies (n = 2; 20%) [34,35]. One study used a randomized controlled trial design (n = 1; 10%) [37]. Most studies were conducted at a single institution (n = 7, 70%) [27,28,30,33,34,36], with only three studies being multicentric (30%) [29,30,37].

Participants and training structure

Participant numbers varied from small pilot studies (n = 23) [35] to larger, multicenter samples (n = 110) [37]. The majority of interventions targeted licensed healthcare professionals (n = 7) [27,28,31–33,35,36] followed by residents (n = 3; 30%) [30,34,37].

The majority of the studies (70%) were single-profession, including mainly obstetricians and gynecologists in their leadership interventions [27,28,30,34,36,37]. Only four studies focused on multidisciplinary participants including obstetricians, midwives and anesthesiologists or other medical specialties [32,33,35,37]. However, only three (30%) of these involved participants training together as a team [33,35,37]. Importantly, only two of these three teams involved real obstetric teams [33,35].

Participant selection varied across studies and included individuals who volunteered, applied to the program or were invited to participate.

Profile of trainers

Most studies provided limited information on the professional

background, qualifications, or experience of trainers, with trainer characteristics not reported in two of the ten included studies [34,35]. Where reported (8/10 studies, 70%), trainers represented a heterogeneous group of clinical and educational professionals.

External specialists were the most frequently described trainers, involved in six studies (60%). These included multidisciplinary experts in leadership education, physician leadership, communication, negotiation, simulation-based education, implicit bias and graduate medical education [27,28,32,33,36,37].

In another study, training was delivered by external clinicians, specifically obstetricians from the United Kingdom providing leadership training and mentorship in a low-resource setting (1/10, 10%) [31].

In-house clinical staff were less commonly reported as trainers. One program was facilitated by midwives (1/10, 10%) delivering a leadership workshop for obstetrics and gynecology residents [30].

Research setting and context

Simulation centers and teaching hospitals were frequent settings, particularly in studies using experiential or high-fidelity simulation-based training [33,34].

The majority of the included studies were conducted in the United States (n=7). The remaining studies were conducted in Australia (n=1) [33], Ghana (n=1) [32] and Malawi (n=1) [31].

Training modalities

The leadership interventions identified across the included studies were notably heterogeneous in design, timing and implementation. The analysis of the included studies revealed a wide range of training modalities, including simulation-based exercises, interactive workshops, coaching, online modules and mentoring programs (Table 2).

A mixed-method intervention was the most frequent training modality (60% of the studies). Six studies employed comprehensive

Table 1

Overview of study characteristics, interventions and outcomes related to leadership in obstetric care.

Source (First Author, Year)	Country	Study design	Setting	Speciality of learner and number of participants	Level of learner	Intervention Description	Topics	Training modality	Duration	Trainer profile	Assessment of training	Long-term follow-up	Main findings by Kirkpatrick level	MERSQI score
Hansen, 2022	USA	Double-blind, multicenter, randomized controlled trial.	Multicenter (5 academic medical centers from the United States)	OB-GYN and emergency medicine residents (n=110)	Residents	Modified TeamSTEPPS or LEADS (online program of 60 minutes and 3 standardized simulations with debriefing) compared to standard simulation without leadership intervention.	Directing communication, assign roles, developing a shared mental model, prioritizing task, and providing summaries and anticipatory guidance	Mixed intervention. Simulation-based + interactive web-based modules	60 minutes of online course plus 1 day simulation training	Experts in leadership education, simulation, implicit bias, and graduate medical education	Leadership performance score at 0 and 3–6 months (Clinical Teamwork Scale) Secondary outcomes detailed leadership evaluation (DLE) with Leadership Behavior Description Questionnaire (LBDQ) measured on 5-point Likert scales	At 3 and 6 months after curriculum exposure: new simulation and evaluation of outcomes	2b. Both programs significantly improved leadership skills over control, sustained over 3–6 months	17/18
Steinhardt, 2018	USA	Descriptive study	Single center	OB-GYN (number of participants not mentioned)	Residents (3 rd year)	2.5-hour workshop (creative exercises + simulation + reflection)	Time management, goals and priorities, leadership styles and skills, and conducting effective meetings	Mixed intervention: small group sessions (reflection, discussion) and simulation	2,5 hours	Midwives	Self-perceived confidence in feedback, communication, leadership insight. 5-point Likert scale for quantitative evaluation and qualitative comments	No	1. Almost 100% numerical rating of 5/5 Likert-scale 1.Satisfaction 2a. Improved reported leadership skills	7/18
Pfeiffer, 2019	Ghana	Descriptive multicenter pilot study	Multicenter	Frontline healthcare workers at Greater Accra Regional Hospital: Physicians, nurses and midwives (n=140)	Licensed provider	18-month multimodal leadership training program. Multiple 2 day-workshops (individual coaching, participatory discussions, role plays, and didactic sessions)	Leadership styles, emotional intelligence, accountability, compassionate care, interpersonal communication, institutional logistics, compliance, efficiency and staff attitudes	Mixed intervention: individual coaching, participatory discussions, role plays, and didactic sessions	18 months	External leadership consultant specialized in experiential coaching methods	5-point Likert scale for quantitative evaluation and qualitative comments	NA	1.High satisfaction. 2b. Self-perceived improvement of leadership skills. 3a. Self-reported implementation of improvement projects using acquired leadership skills	11/18
Broski, 2023	USA	Single group posttest	Single center	Interprofessional teams in OB-GYN (n=160)	Licensed provider	EXPAND Leadership (Distributed co-leadership by nurses and anesthesiologists during simulation); 3-hour training: orientation,	Team response to critical events, role clarity, psychological safety, and leadership role transitions.	Simulation-based	3 hours	NA	Qualitative analysis of debriefing transcripts from videos.	No	1. Positive participant reactions and perceived usefulness of training. 2a. Shift in attitudes toward leadership roles and hierarchy	8,5/18

(continued on next page)

Table 1 (continued)

Source (First Author, Year	Country	Study design	Setting	Speciality of learner and number of participants	Level of learner	Intervention Description	Topics	Training modality	Duration	Trainer profile	Assessment of training	Long-term follow-up	Main findings by Kirkpatrick level	MERSQI score
						simulation warm-up, didactics, two practice cases, and debriefing sessions							(nurses transitioning to leadership). 2b. Improved non-technical skills (situational awareness, delegation, coordination). 3a. Participants reported intent to apply distributed leadership in practice.	
Ellard, 2014	Malawi	Descriptive study	Multicenter	Clinicians in Malawi (n=79)	Licensed provider	30-month program of knowledge and skills training including mentoring practice.	Behaviors and leadership skills, leading change, service improvement and problem identification.	Theory workshop and posterior mentoring.	30 months	UK Clinicians (Obstetricians)	Qualitative evaluation with semistructured interviews about skill implementation, leadership behavior change and peer knowledge transfer	NA	1. Participants perception on the course was overall positive. 2a and 2b. Increase in leadership knowledge and skills 3a. Implementation of knowledge in clinical practice. 4a: perceived improvements in service quality, team effectiveness, and their ability to lead change.	11/18
Pliego, 2008	USA	Prospective pilot study Single group posttest	Single center	Obstetricians and gynecologists residents (n=23)	Residents	One day Ob/Gyn Boot Camp using high-fidelity human simulators. No specific leadership intervention.	Confidence in leadership role	Simulation-based	1 day	NA	5-point Likert scale for quantitative evaluation on perceived competency, leadership confidence and stress hardness	No	1. High satisfaction and positive reaction (82%–91% positive ratings). 2a. Increased interest in developing leadership skills 2b. No improvement of leadership skills.	7/18
Fernandez , 2024	USA	Retrospective pre- and post-test.	Single center	Obstetricians and gynecologists) (n=55)	Licensed-provider	ACOG cefalo leadership training program comparing a virtual cohort	Organizational culture, leading and empowering, communication, motivation,	Online module vs mixed approach (Interactive	Virtual cohort: 4 days (~24.5 h) In person	Physician leadership experts, negotiators, and	6-month post-program data on 14 competencies for knowledge gains and skills	6-months post-program changes in	1. 100% of respondents indicated that training was beneficial for	9.5/18

(continued on next page)

Table 1 (continued)

Source (First Author, Year	Country	Study design	Setting	Speciality of learner and number of participants	Level of learner	Intervention Description	Topics	Training modality	Duration	Trainer profile	Assessment of training	Long-term follow-up	Main findings by Kirkpatrick level	MERSQI score
						(2021, n = 22) to a in-person cohort (2022, n = 33).	advocacy, media, negotiation skills, health policy.	lectures, breakout groups, reflection, coaching and 360° feedback).	cohort: 3.5 days (~31 h)	communications specialists.	use using 5-point Likert scale for quantitative evaluation. Qualitative evaluation with open ended questions.	competency and skills.	their practice 2b. Posttest scores were significantly higher in knowledge and skills retention in both versions. 3a. Self-reported changes in communication and leadership approaches in the workplace (100%). 3a. 100% both groups felt that the course prepared them for their new position and/or leadership position.	
9 Fernandez, 2016	USA	Retrospective pre- and post- test method.	Single center	Obstetricians (Junior fellows, young physicians, senior fellows, not otherwise specified) (n=37)	Licensed- provider	ACOG Robert C. Cefalo National Leadership program. Interactive skills- building workshops, series of leadership and psychological assessment tools, including a 360- degree assessment. Formal coach debriefs of assessment, small and large group sessions	Organizational culture, leading and empowering, communication, motivation, advocacy, media, negotiation skills, health policy	Mixed (interactive lectures, breakout groups, reflection, coaching) 360° feedback	3,5 days	Physician leadership experts, negotiators, and communications specialists.	Retrospective pre- and posttest of perceived skill levels in 10 targeted competency areas. 6 months follow-up. Quantitative evaluation using 5-point Likert- scale and qualitative evaluation with open ended questions.	6 months post- program evaluation pf perceived skills.	1.100% of respondents indicated that they would recommend the course to colleagues 2b. Posttest scores were significantly higher than pretest (p<.0001) in all 10 targeted leadership skills, 0 and 6 months post completion. Mean differences ranged from 0.8 to 1.81 (5-point scale from unskilled to highly skilled) 3a. Self-reported application of skills at job and perceived expansion of leadership roles.	9/18

(continued on next page)

Table 1 (continued)

Source (First Author, Year	Country	Study design	Setting	Speciality of learner and number of participants	Level of learner	Intervention Description	Topics	Training modality	Duration	Trainer profile	Assessment of training	Long-term follow-up	Main findings by Kirkpatrick level	MERSQI score
Janssens, 2019	Australia	Descriptive. Simulation- based video analysis study	Single center	Multidisciplinaryclinicians (midwifery, obstetric and anesthetic staff) (n=16 teams)	Licensed provider	16 teams attending 8-hour multidisciplinary simulation training. 60 min prebriefing and orientation, skills stations and six fully immersive simulations scenarios each followed by debriefing. A 60 min interactive module including trigger videos and small group discussions, focusing on crisis resource management principles and emphasizing the importance of singular, hands-- off leadership was delivered following the first scenario and prior to study scenarios.	Shared leadership	Simulation- based	8 hours	Multidisciplinary clinician educators with advanced training in simulation- based education.	Quantitative evaluation. Distribution of leadership utterances between team members using the Gini coefficient. Teamwork rated with Auckland Team Behavior Tool. Time to critical intervention. Video analysis.	No	2b. Leadership was widely shared across team members; no improvement in teamwork scores; faster time to critical intervention in low leadership- sharing teams. 3b. Observed leadership behaviors showed dominant leaders accounted for ~58% of leadership utterances; leadership sharing does not appear to be associated with better team performance.	12/18
Fernandez, 2023	USA	Retrospective pre- and post- test.	Single center	Obstetricians and gynecologists) (n= 71)	Licensed- provider	ACOG cefalo leadership training program comparing a virtual cohort (2021, n = 32) to a in-person cohort (2022, n = 39)	Organizational culture, leading and empowering, communication, advocacy, media, negotiation skills, health policy.	Online module vs mixed approach (interactive lectures, breakout groups, reflection, coaching and 360° feedback).	Virtual: 4 days (24,5 hours) In-person: 3.5 days (31 hours).	Physician leadership experts, negotiators, and communications specialists.	Pre- and post- program evaluation of knowledge and skills gain using 7-point Likert scale for quantitative evaluation. Qualitative evaluation with open ended questions.	No	1. Overall high satisfaction with both programs. 2b. Posttest scores were significantly higher than pretest in knowledge and ability scores in both programs. Some reported higher gains in the virtual training.	9/18

OB-GYN = Obstetrics and Gynecology;

ACOG = American College of Obstetricians and Gynecologists;

CRM = Crew Resource Management;

PROMPT = Practical Obstetric Multi-Professional Training;

SBAR = Situation, Background, Assessment, Recommendation;

LEADS = Leadership Education Advanced During Simulation;

DLE = Detailed Leadership Evaluation;

Table 2

Characteristics of the included studies (n=10).

Characteristic	No. (%)
Single or multi-institution	
Single institution	7 (70%)
Multicenter	3 (30%)
Study design	
Descriptive	4 (40%)
Single-group posttest	2 (20%)
Single-group pre-post test	3 (30%)
Randomized controlled trial	1(10%)
Level of learner	
Residents	3 (30%)
Licensed provider	7 (70%)
Category	No. (%)
Multidisciplinary participants (not necessarily team-based)	4 (40%)
Team-based multidisciplinary training	3 (30%)
Real obstetric team** training	2 (20%)
Individual /not team-based training	7 (70%)
Training modality	
Mixed-method (simulation + didactic/coaching/feedback)	6 (60%)
Simulation with debriefing only	2 (20%)
Simulation only	1 (10%)
Incorporate online learning	4 (40%)
Virtual-only	2 (20%) *
Theory-based with mentoring	1 (10%)
Participant reaction/satisfaction/, confidence. (Level 1)	8 (80%)
Change in attitudes (Level 2a)	4 (40%)
Knowledge or skills acquisition (Level 2b)	9 (90%)
Behavior change, self-reported (Level 3a)	5 (50%)
Behavior change, observed (Level 3b)	1 (10%)
Patient- or system-level outcomes, self-reported (Level 4a)	1 (10%)
Patient- or system-level outcomes, observed (Level 4b)	0

*Virtual cohort of Fernandez et al. [27,28].

** A real obstetric team was defined as a team composed of obstetricians, midwives, and/or anesthesiologists.

Table 3

Kirkpatrick's framework for training evaluation. Adapted from Lyons et al. [20].

Kirkpatrick Level	Description
Level 1: Reaction	Participants' satisfaction with the learning experience, organization, content and teaching methods.
Level 2A: Change in Attitudes	Changes in participants' attitudes or perceptions toward leadership, management or administration.
Level 2B: Change in Knowledge or Skills	Acquisition of concepts, procedures, thinking, problem-solving, motor or social skills.
Level 3A: Behavioral Change (Self-Reported)	Application of learning at work, as reported by the participants themselves.
Level 3B: Behavioral Change (Observed)	Workplace changes noted by others or reflected in promotions.
Level 4A: Results (Self-Reported)	Participants' perception of organizational impact and team effectiveness.
Level 4B: Results (Observed)	Tangible organizational outcomes, such as cost reduction, improved quality and safety or project impact.

educational programs that integrated simulation with interactive workshops, small-group discussions, coaching and/or 360-degree feedback [27,28,30,32,36,37]. Simulation-based training alone was used by only one study which used high-fidelity simulation as a stand-alone intervention without complementary theoretical or coaching components [34]. Simulation with debriefing and structured observation was used in two studies [33,35]. Four studies incorporated a virtual platform for online learning [27,28,36,37]. A single study implemented a long-term mentoring and theory-based curriculum without simulation [31].

The duration of leadership training interventions showed notable variability, although short-format programs predominated. Five studies delivered training within a single day or less, including sessions lasting only a few hours [30,33–35,37]. Three studies implemented short intensive courses lasting between 1 and 4 days [27,28,36]. In contrast,

LBDQ = Leadership Behavior Description Questionnaire;
MSF = Multisource Feedback;
TOAsT = Temporal Observational Analysis of Teamwork;
NA = Not applicable.
MBRRACE-UK: Mothers and Babies: Reducing Risk through Audits and Confidential Enquiries- United Kingdom.
NHS: National Health System.
MERSQJ score: Medical Education Research Study Quality Instrument.

only two studies described long-term, longitudinal programs exceeding 12 months in duration [31,32].

Long-term follow-up of the improvement from the reported outcomes was described in three studies (30%) [27,36,37] and ranged between 3–6 months.

Assessment and impact

The evaluation of outcomes across the included studies primarily focused on the lower to mid-levels of the Kirkpatrick model [26]. Participant reaction, satisfaction and confidence (Level 1) were assessed in 8 studies [27,28,30–32,34–36], representing 80% of the sample. Changes in attitudes (Level 2a) was less frequently measured, reported in 4 studies [30,31,34,35]. The majority of studies, 9 out of 10 (90%), evaluated knowledge or skills acquisition (Level 2b), typically using self-assessments, written tests or direct observation during simulation activities [27,28,30–37]. Fewer studies evaluated behavioral change, with 50% reporting self-reported behavior change (Level 3a) [27,31,32,35,36] and only one study (10%) assessing observed behavioral outcomes (Level 3b) [33]. Notably, only one study evaluated system-level outcomes (Level 4a) [31]. No study evaluated patient-related outcomes.

Evaluation methods

The evaluation methods across the 10 included studies were varied and frequently combined both quantitative and qualitative outcomes. Five studies used mixed methods, incorporating quantitative analysis with Likert-scale surveys alongside open-ended questions or other qualitative data such as debriefing transcripts and interviews. Three studies relied exclusively on quantitative assessments, using structured scales or video-coded metrics [33,34,37], while two studies (20%) used only qualitative analysis [31,35].

Likert-type questionnaires were the most frequently used quantitative tool, appearing in 7 of the 10 studies (70%) [27,28,30,32,34,36,37]. These instruments typically measured self-perceived leadership abilities, confidence, communication, and knowledge across 5 or 7-point scales. For qualitative evaluation, open-ended questions, semi-structured interviews or comments, were the preferred format in 6 of the 10 studies [27,28,30–32,36].

Most evaluations (70%) relied on self-assessment to measure learning outcomes and perceived changes in leadership skills. Fewer studies (30%) included external perspectives, such as expert observers or video analysis [33,35,37]. In two studies, structured tools like the Gini coefficient, Clinical Teamwork Scale, or Auckland Team Behavior Tool were used to evaluate leadership behaviors and team function objectively [33,37].

c) Methodological quality assessment

The methodological quality of the included studies was evaluated using the Medical Education Research Study Quality Instrument (MERSQI), which assesses six domains: study design, sampling, type of data, instrument validity, data analysis, and outcome level. Each domain is scored from 0 to 3, with a maximum total score of 18 points [25,29]. See Supplement 2. Studies with a score of 12 or higher were categorized as high reliability studies [14,19].

The mean MERSQI score was 10.1 (SD 2.9). The scores ranged from 7 to 17 out of 18, reflecting low to high methodological quality, but with substantial variability between studies. See results in Table 1. Only 2 of the 10 included studies [33,37] achieved high methodological quality (MERSQI ≥ 12), while the majority (8/10) were of low to moderate quality [27,28,30–32,34–36]. The highest score (17/18) was achieved by a single study, likely reflecting its randomized controlled design, multicenter setting, use of validated instruments, objective outcome measures and appropriate statistical analysis [37].

Discussion

a) Principal findings and comparison with existing literature

This systematic review provides the first focused synthesis of leadership development programs in obstetric care. Our aim was to analyze

their structure, target populations, training methods and reported outcomes, as well as to assess the methodological quality of the available evidence. Despite growing recognition of the importance of leadership as a non-technical skill essential for patient safety and effective team functioning, our findings reveal that formal leadership training in obstetrics remains insufficiently developed and inconsistently applied in this field. Across the ten included studies, we observed substantial heterogeneity in study design, intervention formats, participant profiles and evaluation strategies.

Compared with the broader literature on leadership development programs in healthcare; where interventions in areas such as emergency medicine, trauma and intensive care have shown associations with improved outcomes; programs in obstetrics remain limited in both number and scope [19,22]. Most of the included studies focused on short-term, local initiatives and evaluated primarily subjective outcomes.

An important finding in this review is the limited inclusion of multidisciplinary obstetric teams, despite strong evidence supporting the importance of interprofessional team training in improving outcomes in obstetric care [33,35,37]. An umbrella review has shown that leadership programs are more effective when they include multidisciplinary cohorts and team-based approaches [38].

Importantly, multidisciplinary participation and team-based training represent were distinct concepts. While some programs included participants from different professional backgrounds, this did not necessarily imply that training was delivered in an integrated team-based format [32].

In our review, most interventions targeted physicians only [27,28,30,31,34,36], with only four studies (40%) including multidisciplinary participants [32,33,35,37]. Among these, just three (30%) [33,35,37] implemented team-based training, and only two [33,35] involved complete obstetric teams training together as they would function in clinical practice. This represents a relevant gap, given that obstetric emergencies are inherently managed by coordinated multidisciplinary teams, where effective leadership depends on team interaction, communication, and shared decision-making. Training leadership skills in isolation, without the team context in which they are applied, may therefore limit the transferability of these interventions to real clinical settings. Consequently, leadership training programs in obstetrics should ideally incorporate team-based components that reflect the complexity and interdependence of clinical care. This limited integration reduces the applicability of many interventions to real clinical team dynamics.

Similar patterns have been reported in previous systematic reviews of medical leadership training programs. Rosenman et al. and Frich et al. found that only 4% and 24.5% of studies, respectively, included multidisciplinary participants [18,21]. Notably, these reviews primarily focused on participant composition and did not distinguish between multidisciplinary participation and true team-based training, which may further underestimate the gap between current training approaches and the collaborative nature of real clinical practice.

We observed that only 30% of the studies in our review targeted residents, representing a missed opportunity to introduce leadership competencies early in clinical training [30,34,37]. Prior research in medical education supports the value of integrating leadership skills at this formative stage, as it contributes to long-term behavioral change and culture development within healthcare teams [5,17]. In comparison, Lyons et al. found that 60% of the leadership interventions they reviewed were targeted at residents, suggesting a greater emphasis in other fields on integrating leadership training into graduate medical education [14].

In terms of training formats, most studies (60%) employed modal approaches (mixed-methods) that combined simulation, interactive workshops, coaching, and theoretical instruction [27,28,30,32,36,37]. This is in accordance with the results of systematic reviews which concluded that leadership development programs were

more effective when multiple learning methods were employed [18,38]. A simulation-based component was present in 9 of the 10 included studies, but only one study relied exclusively on simulation without accompanying theoretical or reflective components [34]. A similar preference for simulation-based approaches was observed in other systematic reviews. Rosenman et al. reported that 81% of the interventions they analyzed used simulation as a primary training method [21].

The most frequent format was a face-to-face training. Online-only modules were present in 4 cohorts and were associated with improved knowledge retention and self-reported behavior change [27,28]. However, the isolated use of virtual learning without interactive or experiential components, as reported by Philipson et al in their umbrella review, is less associated with effectiveness in improving outcomes than those with at least some degree of face-to-face learning [38].

One study employed a longitudinal mentoring model combined with theoretical instruction [31]. In this study, despite the absence of simulation, participants reported leadership knowledge gains and application of skills in practice, suggesting that sustained engagement and mentorship may serve as a viable alternative in some settings.

With respect to geographical distribution and intervention duration, the included studies demonstrated a concentration in specific settings and training formats. Most studies (70%) were conducted in the United States, with limited representation from other countries. Leadership training interventions were predominantly delivered as short, intensive formats (≤ 4 days), accounting for 80% of the included studies, and often consisting of single-session or brief course-based approaches. In contrast, only a small number of studies (20%) implemented long-term programs extending beyond 12 months.

In terms of outcomes, positive outcomes related to leadership training were reported in 90% of the included studies. This is in line with results of other systematic reviews focusing on healthcare leadership development programs [38,39]. The only study that did not demonstrate measurable improvement involved a simulation-only intervention, conducted without any accompanying debriefing, coaching, or preparatory instruction [34]. This finding suggests that simulation alone may be insufficient to achieve meaningful learning outcomes, particularly in leadership domains, and highlights the importance of incorporating reflective and feedback-based components into training design.

The outcome measures reported across the included studies were predominantly subjective in nature, relying on participants' self-assessment of leadership skills, perceived confidence and self-reported behavioral change. These outcomes were commonly measured through post-intervention Likert-scale surveys, semistructured interviews or descriptive feedback forms [27,30,32,36]. While these tools have numerical data, they remain subjective, as they reflect the perceptions of the participant rather than externally validated changes. Other systematic reviews similarly demonstrate a predominance of subjective outcomes, primarily assessed through self-reported questionnaires [14,19,21,38,39].

In our review, most studies evaluated outcomes at the lower levels of Kirkpatrick's model [26], focusing mainly on immediate perceptions and knowledge acquisition. Participant satisfaction and confidence (Level 1) were assessed in 80% of studies. Attitudinal change (Level 2a) was examined in 40% while nearly all studies (90%) measured knowledge or skill acquisition (Level 2b), typically through self-assessment tools or observation in simulation settings. Self-reported behavioral change (Level 3a) was reported in five studies (50%) [27,31,32,35,36], whereas only one study (10%) assessed objectively observed behavioral change (Level 3b) [33]. Overall, six studies (60%) evaluated behavioral change at any level. These results are in line with previous reviews. Lyons et al. found Level 3 outcomes in 57% of studies and Level 4 in 24% [14]. Rosenman et al. reported that 18% of studies evaluated system-level outcomes [21]. Frich et al. also reported system-level evaluations in 6 of 45 studies (13%) with measurable organizational improvements, such as enhanced disease management, leadership advancement and increased patient satisfaction [18]. An umbrella review from 2025 also

highlighted a predominance of outcomes based on Kirkpatrick level 1, 2 and 3 and less level 4 based on organizational or patient outcomes [38]. Overall, leadership training in obstetric care tends to emphasize subjective learning outcomes, with limited evidence of real-world behavioral or organizational impact. This highlights a key gap in the field and a need for more rigorous, higher-level outcome evaluations.

The role and profile of those responsible for delivering leadership training programs remain poorly defined in the included articles. Where reported, trainers ranged from in-house clinical staff, such as midwives and multidisciplinary clinician educators with simulation expertise, to external clinicians and specialists in leadership education, communication, negotiation, and implicit bias. Overall, the articles provide little detail on the professional background, qualifications, or experience of trainers. This gap is also acknowledged in a systematic review from 2020 that highlighted the absence of research on program facilitators [15]. An umbrella review of systematic reviews in leadership training programs concluded that interventions that combined internal and external faculty were theorized to be more effective than those led solely by internal or external faculty [38]. Particularly, one systematic review reported that training was most often delivered by in-house faculty (36%) or by a combination of in-house and external faculty (32%), with mixed-faculty programs showing the strongest impact on organizational outcomes [17].

Training providers should adapt approaches to specific participant needs. Probably, the most effective approach in obstetrics would be a multidisciplinary training team, combining external experts in leadership and training methodologies with obstetricians, midwives, and/or anesthesiologists who not only have leadership knowledge but also understand the specific demands of the delivery room, obstetric emergencies and the unique dynamics of maternity care. Future studies specifically addressing this gap are crucial to advance understanding and improve practice.

The methodological quality of included studies, assessed using the MERSQI tool, ranged from 7 to 17 (mean 10.1) and only two studies [33,37] met criteria for high methodological quality (MERSQI ≥ 12). Common limitations included a lack of control groups, non-validated instruments and incomplete reporting of evaluation design. These findings are consistent with prior systematic reviews on leadership training in healthcare, which have similarly reported moderate methodological rigor. Kingsley-Smith et al reported a mean MERSQI score of 10.5, Rosenman et al. a score of 11, Geerts et al. a score of 10, and Lyons et al. a lower average of 8.5, suggesting that the overall quality of leadership intervention studies across specialties remains limited and in need of improvement [14,19,21,39]. In an umbrella review of systematic reviews from 2025 which included 86 systematic reviews with a total of 1659 individual articles; quality score generally fell in the low-moderate reliability range [38].

In our review, the two studies [33,37] which met criteria for high methodological quality shared some characteristics that distinguish them from lower-quality studies. First, they employed more robust study designs, including a double-blind, multicenter, randomized controlled trial or a well-structured pre-post design, in contrast to the predominantly single-group or descriptive approaches observed in lower quality studies. Second, they incorporated more rigorous outcome assessment strategies, including the use of objective or validated measurement tools (Auckland Team Behavior Tool, Clinical Teamwork scale, detailed leadership evaluation, Behavior Description Questionnaire) whereas lower-quality studies largely relied on subjective, self-reported outcomes.

Additionally, high-quality studies demonstrated greater methodological rigor in data analysis and reporting, including clearer description of sampling strategies and large or multicenter samples. They were also more likely to include follow-up assessments, allowing evaluation of the sustainability of leadership skills over time. In contrast, most of the lower quality studies lacked longitudinal follow-up and focused primarily on immediate post-intervention perceptions. On the other side,

lower quality studies that incorporated follow-up [27,36], achieved moderate methodological quality (mean MERSQI of 9), suggesting that longitudinal assessment contributes to greater methodological rigor, although it is not sufficient in isolation to meet high-quality standards. To facilitate comparison, the differences between high- and lower-quality studies are summarized in Table 4.

Summary of key observations and gaps

Although the number of included studies was limited and findings should be interpreted with caution, several observations can be identified across our data. A summary of our findings is provided in Tables 5 and 6. We observed that leadership training programs were predominantly conducted in the United States and were most often delivered as short, intensive interventions lasting between 1 and 4 days. The majority of programs targeted individual participants rather than multidisciplinary teams (70%), with only a minority adopting a truly multidisciplinary team-based approach (20%). Similarly, relatively few interventions focused on residents (30%), with most programs directed at licensed healthcare professionals.

In terms of educational design, multimodal training formats were most common (60%), typically combining simulation with additional components such as workshops, coaching, or reflective activities. Most studies reported positive outcomes (90%); however, these were largely based on subjective, self-reported measures, and were primarily evaluated at the lower to intermediate levels of the Kirkpatrick model (Levels 1-3).

The profile of trainers was inconsistently reported, although when described, programs were delivered by a heterogeneous group including clinician educators, simulation experts, and external leadership specialists.

Overall, the methodological quality of the included studies was

Table 4
Methodological differences between high- and lower-quality studies based on MERSQI assessment.

Methodological characteristic	High-quality (HQ) studies (n = 2)	Lower-quality studies aligned with HQ characteristics (n=5)	Lower-quality studies not aligned (n=3)
Study design	Randomized or well-structured observational designs (randomized controlled trial, structured video-analysis studies)	Pre-post designs without control groups	Predominantly single-group posttest or descriptive designs
Outcome measures	Objective and validated tools (Clinical Teamwork Scale, Auckland Team Behavior Tool)	Structured but non-validated tools	Predominantly self-reported outcomes (Likert scales, qualitative feedback)
Data collection	Objective or externally assessed data (video analysis, validated tools)	Combination of objective and self-reported data	Exclusively self-reported or qualitative data
Data analysis	Clearly described and appropriate statistical or quantitative analysis	Basic statistical analysis	Descriptive or qualitative analysis only
Sampling	Larger or multicenter samples	Moderate sample sizes, typically single-center	Small, single-center samples
Follow-up assessment	Present in high-quality studies (limited evidence)	Included in selected studies	Not reported

MERSQI: Medical Education Research Study Quality Instrument. High-quality studies were defined as those with a MERSQI score ≥ 12 [14].

Table 5
Summary of leadership training programs in obstetric care. Table adapted and expanded from Geerts et al., based on the synthesis of included studies in this review [13].

Domain	Element	Most frequent findings
Program Design	Target population	Predominantly licensed healthcare professionals (70%); limited inclusion of residents (30%)
	Multidisciplinary approach	Predominantly individual training (70%); 40% included multidisciplinary participants, and only 30% involved team-based multidisciplinary training
	Training objectives	Focus on individual leadership skills (communication, decision-making, role allocation)
	Duration	Highly variable; most programs short-format (≤ 4 days, 80%), with few long-term longitudinal programs (> 12 months, 20%)
	Structure	Primarily short courses or modular programs; few longitudinal curricula
Delivery Methods	Educational approach	Predominantly multimodal (60%), combining simulation with workshops, coaching or feedback
	Simulation use	Present in 90% of studies; rarely used as a standalone method (10%)
	Complementary methods	Workshops, small-group discussions, coaching, mentoring, and 360° feedback
	Learning format	Primarily face-to-face; some incorporation of online/virtual components (40%), including fully virtual formats (20%)
Content	Trainer profile	Heterogeneous; mainly external experts; limited reporting across studies
	Core competencies	Communication, team coordination, leadership styles, situational awareness
	Advanced topics	Emotional intelligence, negotiation, health systems and change management (less consistently addressed)
Evaluation	Outcome type	Predominantly subjective self-reported outcomes (70%)
	Measurement tools	Mainly Likert-scale questionnaires; limited use of validated or objective instruments
	Kirkpatrick level	Primarily Level 1 (80%) and Level 2b (90%); fewer studies assessed Level 3a (50%) and only one assessed Level 3b (10%); minimal evaluation at Level 4 (10%)
	Follow-up	Limited (30%), typically short-term (3–6 months)

Abbreviations: *MERSQI score: Medical Education Research Study Quality Instrument.

Table 6
Key gaps in leadership training programs in obstetric care.

Domain	Element	Most frequent findings
Key Gaps	Multidisciplinary	Limited inclusion of multidisciplinary training; few interventions involved participants training together as integrated obstetric teams
	Outcome evaluation	Predominant reliance on self-reported outcomes, with minimal assessment of objectively measured behavioral change and absence of patient-level outcomes
	Standardization	Substantial heterogeneity in program design, duration, and evaluation methods, limiting comparability across studies
	Assessment methods	Limited use of objective outcome measures and validated assessment tools; most studies relied on non-validated self-reported questionnaires
	Long-term impact	Limited use of follow-up assessments, restricting evaluation of the sustainability of training effects over time

predominantly low to moderate, with only a small number meeting criteria for high methodological quality (20%). Higher-quality studies were more likely to incorporate objective or validated outcome measures, whereas lower quality studies predominantly relied on subjective self-reported outcomes. The inclusion of follow-up assessments appeared to be associated with higher methodological quality, although this alone was not sufficient to meet high-quality criteria.

Shorter and less structured programs, such as simulation-based or short-course interventions [27,28,33–36] appeared to focus primarily on individual-level outcomes and immediate post-intervention perceptions. In contrast, more comprehensive or structured interventions [31,32,37] included broader evaluation approaches, such as behavioral outcomes or implementation of leadership skills in practice. Notably, none of the included studies assessed long-term patient-level outcomes, which may reflect an important gap in the current literature.

Taken together, these findings highlight several key gaps in the current evidence base, including limited implementation of multidisciplinary team-based training, a predominant reliance on self-reported outcomes, lack of standardized program design and evaluation approaches, and insufficient longitudinal assessment of outcomes.

b) Strengths and limitations

To the best of our knowledge, this systematic review offers the first focused synthesis of leadership development programs specifically within obstetric care. One of its main strengths lies in the comprehensive and methodological approach, following PRISMA guidelines and applying the MERSQI framework for quality assessment. The review highlights key structural and conceptual elements of leadership interventions, mapping them across Kirkpatrick's levels of evaluation and identifying trends in training design, implementation, and outcomes. Another strength is that it exposes gaps in the current literature, such as the lack of multidisciplinary approaches, the predominance of subjective outcome measures, and the limited integration of leadership training into formal obstetric education. We also provide directions for future research and practice.

However, several limitations should be noted. The number of eligible studies was small and characterized by substantial heterogeneity in design, population, and evaluation methods. Most studies relied on self-reported, non-validated instruments and short-term outcomes, limiting the strength of the evidence.

Furthermore, the predominance of studies conducted in the United States (70%) may limit the generalizability of these findings to other healthcare systems and settings with different organizational and resource contexts. Additionally, three of the ten studies evaluated variations of the same program, which may have reduced the overall diversity of the evidence [27,28,36].

Most studies reported positive findings, which raises the possibility of publication bias. Given the limited number of studies included in this review, caution is warranted in interpreting these findings, and the observed patterns should be viewed as preliminary and hypothesis-generating rather than definitive.

Implications for future research

The findings of our systematic review highlight several priorities for future research in leadership training within obstetric care. There is a need for more standardized and multidisciplinary programs that better reflect the team-based nature of clinical practice, and that actively include trainees such as residents, given their key role in clinical teams and future leadership development.

Future studies should incorporate more robust evaluation strategies, including the use of objective and validated outcome measures, such as standardized assessment tools and validated rating scales. Whenever feasible, these evaluations should be complemented by third-party observations rather than relying exclusively on self-reported subjective outcomes. Greater emphasis should also be placed on higher levels of the Kirkpatrick model, particularly behavioural (level 3) and patient-level outcomes (level 4), including clinically meaningful maternal and neonatal outcomes where feasible.

In addition, future research should explore the sustainability and long-term transfer of leadership skills into clinical practice through longitudinal follow-up designs, and consider the influence of organizational context and safety culture on training effectiveness. Stronger study designs, such as controlled and multicenter studies, are needed to improve the overall quality and generalizability of the evidence. Improved reporting of intervention characteristics, especially the profile and role of trainers, is also essential to enhance reproducibility.

Conclusions

Leadership training programs in obstetrics are associated with improvements in participants' self-assessed knowledge and skills. However, the current evidence base is limited by methodological weaknesses, a predominance of subjective outcome measures, and a lack of robust evaluation of long-term and clinical effects.

As a result, it remains unclear whether these perceived improvements translate into meaningful changes in clinical practice or patient outcomes. Future initiatives should focus on the development of standardized, multidisciplinary training programs integrated into formal education, and on their evaluation using objective, validated, and clinically relevant outcome measures.

Funding

This research received no external funding.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the author(s) used ChatGPT (OpenAI, GPT-4) in order to support the revision of English grammar, spelling and phrasing. The tool was employed solely for language editing purposes. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

CRedit authorship contribution statement

Sara Tameish: Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis, Data curation, Conceptualization. **Pere Cavallé Busquets:** Writing – review & editing, Formal analysis, Data curation, Conceptualization. **Óscar Martínez Pérez:** Writing – review & editing, Supervision, Methodology, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ejogrb.2026.115184>.

References

- [1] Frost J, Hillard T. Clinical leadership and service delivery in obstetrics and gynaecology. In: Professional development skills for obstetricians and gynaecologists. Cambridge University Press; 2018. p. 119–33. <https://doi.org/10.1017/9781316831748.013>.
- [2] Cornthwaite K, Edwards S, Siassakos D. Reducing risk in maternity by optimising teamwork and leadership: an evidence-based approach to save mothers and babies. *Best Pract Res Clin Obstet Gynaecol* 2013 Aug;27(4):571–81. <https://doi.org/10.1016/j.bpobgyn.2013.04.004>. PubMed PMID: 23647702.
- [3] Ashmore AA, Kanga K, Kaur-Desai T, Thorman K, Archer N. Building leadership capabilities in maternity. *BMJ Leader* 2022 Mar 1;6(1):10–4. <https://doi.org/10.1136/leader-2021-000449>. PubMed PMID: 35537013.
- [4] Nielsen P, Mann S. Team function in obstetrics to reduce errors and improve outcomes. *Obstet Gynecol Clin North Am* 2008;81–95. <https://doi.org/10.1016/j.ogc.2007.12.011>. PubMed PMID: 18319130.

- [5] Blumenthal DM, Bernard K, Bohnen J, Bohmer R. Addressing the leadership gap in medicine: residents' need for systematic leadership development training. *Acad Med* 2012;87(4):513–22. <https://doi.org/10.1097/ACM.0b013e31824a0c47>.
- [6] Sundar E, Sundar S, Pawlowski J, Blum R, Feinstein D, Pratt S. Crew resource management and team training. *Anesthesiol Clin* 2007;283–300. <https://doi.org/10.1016/j.ancin.2007.03.011>. PubMed PMID: 17574191.
- [7] Levine S. Are all physicians leaders? the opinions of permanent physician-leaders. *Perm J* 2002;6.
- [8] Singh PSSKVTM. Navigating healthcare leadership: theories, challenges, and practical insights for the future. *J Postgrad Med* 2024 Dec;70(4):232–41.
- [9] The Joint Commission. Department of corporate communications. The essential role of leadership in developing a safety culture [internet]. 2021 Jun. Available from: www.jointcommission.org.
- [10] The Joint Commission. http://www.jointcommission.org/sentinel_event_alert_issue_30_preventing_infant_death_and_injury_during_delivery/. Accessed 13 January 2025.
- [11] Knight M. Saving lives, improving mothers' care maternal, newborn and infant clinical outcome review programme. 2019.
- [12] Felker A, Patel R, Kotnis R, Kenyon S, Knight M. Newborn and infant clinical outcome review programme saving lives, improving mothers' care [internet]. 2024. Available from: www.hqip.org.uk/national-programmes.
- [13] Geerts JM. Maximizing the impact and ROI of leadership development: a theory- and evidence-informed framework. *Behavioral Sciences. Multidisciplinary Digital Publishing Institute (MDPI)* 2024. <https://doi.org/10.3390/bs14100955>.
- [14] Lyons O, George R, Galante JR, Mafi A, Fordwoh T, Frich J, et al. Evidence-based medical leadership development: a systematic review. *BMJ Leader. BMJ Publishing Group*; 2021. p. 206–13. doi:10.1136/leader-2020-000360 PubMed PMID: 37850339.
- [15] Till A, McKimm J, Swanwick T. The importance of leadership development in medical curricula: a UK perspective (stars are aligning). *J Healthcare Leadership. Dove Medical Press Ltd.* 2020;19–25. <https://doi.org/10.2147/JHL.S210326>. PubMed PMID: 32214861.
- [16] O'Keeffe DF. Leadership. *Current Opinion in Obstetrics and Gynecology*. 2012. p. 436–9. doi:10.1097/GCO.0b013e31828359b605 PubMed PMID: 23090618.
- [17] Kumar B, Sweet ML, Suneja M. Leadership training programs in graduate medical education: a systematic review. *BMC Med Educ*. 2020 Jun 2;20(1). doi:10.1186/s12909-020-02089-2 PubMed PMID: 32487056.
- [18] Frich JC, Brewster AL, Cherlin EJ, Bradley EH. Leadership development programs for physicians: a systematic review. *J Gen Intern Med* 2015 May 1;30(5):656–74. <https://doi.org/10.1007/s11606-014-3141-1>. PubMed PMID: 25527339.
- [19] Geerts JM, Goodall AH, Agius S. Evidence-based leadership development for physicians: a systematic literature review. *Social science and medicine. Elsevier Ltd*; 2020. doi:10.1016/j.socscimed.2019.112709 PubMed PMID: 31887629.
- [20] Wong CA, Cummings GG, Ducharme L. The relationship between nursing leadership and patient outcomes: a systematic review update. *J Nurs Manag* 2013 Jul;21(5):709–24. <https://doi.org/10.1111/jonm.12116>. PubMed PMID: 23865924.
- [21] Rosenman ED, Shandro JR, Ilgen JS, Harper AL, Fernandez R. Leadership training in health care action teams: a systematic review. *Acad Med* 2014;1295–306. <https://doi.org/10.1097/ACM.0000000000000413>. PubMed PMID: 25006708.
- [22] Husebø SE, Akerjordet K. Quantitative systematic review of multi-professional teamwork and leadership training to optimize patient outcomes in acute hospital settings. *J Adv Nursing. Blackwell Publishing Ltd* 2016;2980–3000. <https://doi.org/10.1111/jan.13035>. PubMed PMID: 27240316.
- [23] Friedman CL, Boulos ME, Garber AB, Craig C, Mueller V. Learning to lead: an exploration of leadership development in obstetrics and gynaecology. *J Obstetrics Gynaecol Canada. Elsevier Inc.*; 2024. doi:10.1016/j.jogc.2024.102407 PubMed PMID: 38365114.
- [24] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLoS Med* 2009. <https://doi.org/10.1371/journal.pmed.1000100>. PubMed PMID: 19621070.
- [25] Cook DA, Reed DA. Appraising the quality of medical education research methods: the medical education research study quality instrument and the newcastle-ottawa scale-education. *Acad Med* 2015 Aug 31;90(8):1067–76. <https://doi.org/10.1097/ACM.0000000000000786>. PubMed PMID: 26107881.
- [26] Kirkpatrick DL. Techniques for evaluating training programs. *Train Dev J* 1960; 92:178.
- [27] Fernandez CSP, Adatsi G, Hays CN, Noble CC, Abel-Shoup M, Connolly A. Immersive leadership training for physicians: comparing retrospective pre-and post-test virtual vs in-person 6-month follow-up of learning. *J Healthc Leadersh*. 2024;16:235–54. <https://doi.org/10.2147/JHL.S455105>.
- [28] Fernandez CSP, Hays CN, Adatsi G, Noble CC, Abel-Shoup M, Connolly A. Comparing virtual vs in-person immersive leadership training for physicians. *J Healthc Leadersh* 2023;15:139–52. <https://doi.org/10.2147/JHL.S411091>.
- [29] Reed DA, Mmdacmmjbm, et al. Association between funding and quality of published medical education research. *J Am Med Assoc* 2007;298(9):1002–9.
- [30] Steinhardt L. Workshop for new leaders: innovative midwifery teaching for obstetrics and gynecology residents. *J Midwifery Womens Health* 2015;60(3): 313–7. <https://doi.org/10.1111/jmwh.12292>. PubMed PMID: 25959237, May 1.
- [31] David R Ellard 1 Wanangwa Chimwaza,2 David Davies,3 Joseph Paul O'Hare,4 Francis Kamwendo,5 Siobhan Quenby,6 Frances Griffiths,7 On behalf of The ETATMBA Study Group. Can training in advanced clinical skills in obstetrics, neonatal care and leadership, of non-physician clinicians in Malawi impact on clinical services improvements (the ETATMBA project): a process evaluation. *BMJ Open*. 2014;4(11). doi:10.1136/bmjopen-2014 PubMed PMID: 25377009.
- [32] Pfeiffer E, Owen M, Pettitt-Schieber C, Van Zeijl R, Srofenyoh E, Olufolabi A, et al. Building health system capacity to improve maternal and newborn care: a pilot leadership program for frontline staff at a tertiary hospital in Ghana. *BMC Med Educ*. 2019 Feb 11;19(1). doi:10.1186/s12909-019-1463-8 PubMed PMID: 30744625.
- [33] Janssens S, Simon R, Barwick S, Beckmann M, Marshall S. Leadership sharing in maternity emergency teams: a retrospective cohort study in simulation. *BMJ Simul Technol Enhanc Learn* 2020 May 1;6(3):135–9. <https://doi.org/10.1136/bmjstel-2018-000409>.
- [34] Pliego JF, Wehbe-Janek H, Rajab MH, Browning JL, Fothergill RE. Ob/Gyn Boot Camp using high-fidelity human simulators: enhancing residents' perceived competency, confidence in taking a leadership role, and stress hardness. *Simul Healthc* 2008 Jun;3(2):82–9. <https://doi.org/10.1097/SIH.0b013e3181658188>. PubMed PMID: 19088646.
- [35] Broski J, Tarver S, Krase K, Petersen S, Wolverton A, Winchester M, et al. Integrating simulation and interpretive description to explore operating room leadership: critical event continuing education. *Adv Health Sci Educ* 2023;28(4): 1211–44. <https://doi.org/10.1007/s10459-023-10212-3>. PubMed PMID: 37022534, Oct 1.
- [36] Fernandez CSP, Noble CC, Jensen ET, Chapin J. Improving leadership skills in physicians: a 6-month retrospective study. *J Leadersh Stud* 2016;9(4):6–19. <https://doi.org/10.1002/jls.21420>. Dec 1.
- [37] Hansen M, Harrod T, Bahr N, Schoonover A, Adams K, Kornegay J, et al. The effects of leadership curricula with and without implicit bias training on graduate medical education: a multicenter randomized trial. *Acad Med* 2022 May 1;97(5):696–703. <https://doi.org/10.1097/ACM.0000000000004573>. PubMed PMID: 34966032.
- [38] Phillipson J, Pinto AC, Kingsley-Smith H, Krachler N, McGivern G, Lyons O. Leadership training in healthcare: a systematic umbrella review. *BMJ Leader. BMJ Publishing Group*; 2025. doi:10.1136/leader-2025-001269.
- [39] Harry Kingsley-Smith CEFDFKKKMSSAMS. Leadership development programmes in healthcare research: a systematic review, meta-analysis and meta-aggregation. *BMJ Leader* 2024;8:283–92.