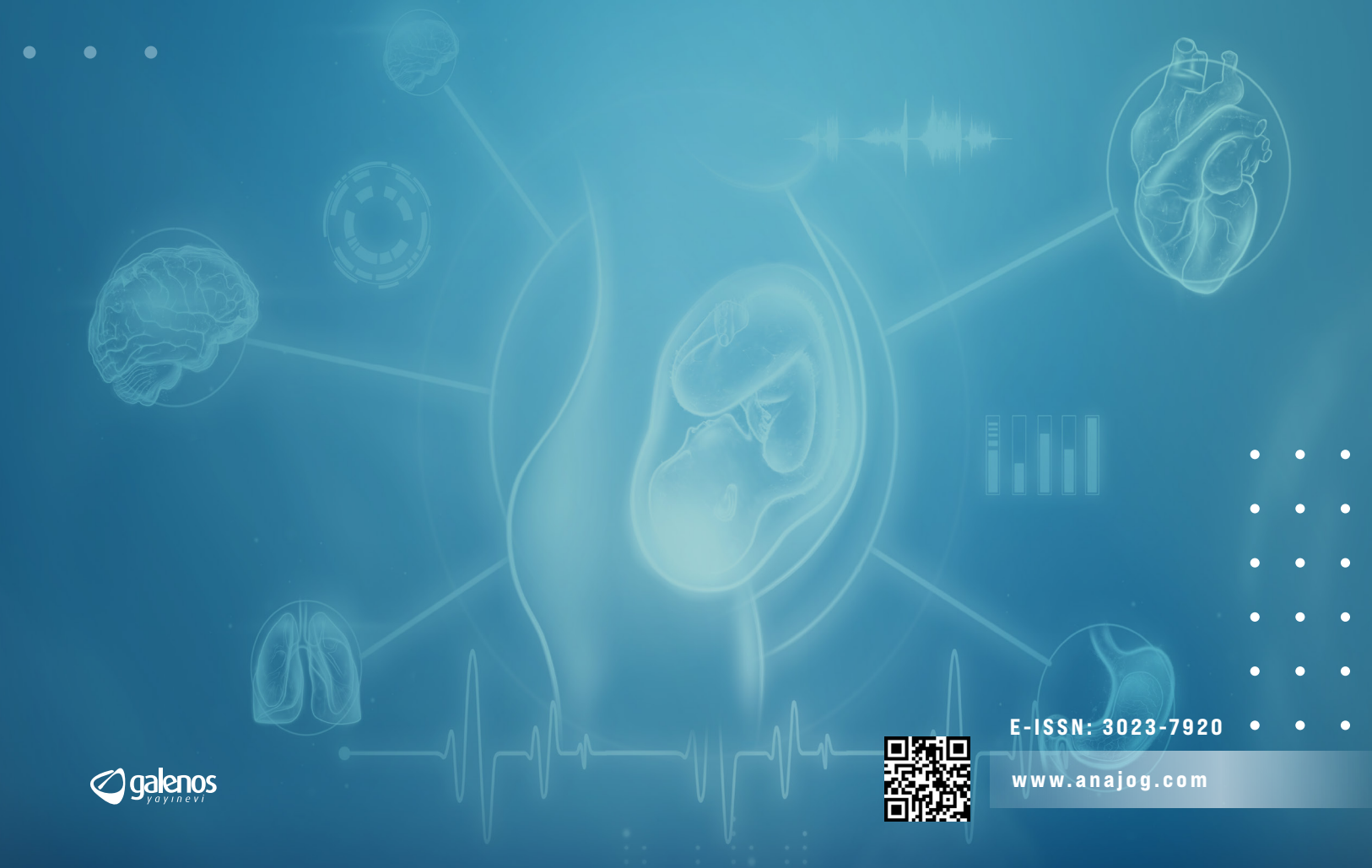
ANATOLIAN JOURNAL OF OBSTETRICS & GYNECOLOGY RESEARCH

Years | 2026 | Volume | 3 | Issue | 2 | Month | August



E-ISSN: 3023-7920

www.anajog.com

galenos yayınevi

ANATOLIAN JOURNAL OF
OBSTETRICS & GYNECOLOGY
RESEARCH

Years | 2026 | Volume | 3 | Issue | 2 | Month | August

ANATOLIAN JOURNAL OF
OBSTETRICS & GYNECOLOGY
RESEARCH

Years | 2026 | Volume | 3 | Issue | 2 | Month | August

ANATOLIAN JOURNAL OF
OBSTETRICS & GYNECOLOGY
RESEARCH

Years | 2026 | Volume | 3 | Issue | 2 | Month | August

Yusuf Üstün

University of Health Sciences Turkey, Ankara SUAM, Ankara/Turkey

E-mail: *ustunyusuf@gmail.com*

Erkut Attar

Yeditepe University Faculty of Medicine, İstanbul/Turkey

E-mail: *dreattar@gmail.com*

Batuhan Özmen

Ankara University Faculty of Medicine, Department of Obstetrics and Gynecology, Ankara/Turkey

E-mail: *drbatuhanozmen@gmail.com*

Ömer Tarık Yalçın

Eskişehir Osmangazi University, Eskişehir/Turkey

E-mail: *otyalcin@gmail.com*

Murat Ekin

Medipol Mega University Hospital, Department of Obstetrics and Gynecology, İstanbul/Turkey

E-mail: *drmuratekin@gmail.com*

Rukset Attar

Yeditepe University, İstanbul/Turkey

E-mail: *rukseta@gmail.com*

Mehmet Sakıncı

Private Clinic, Antalya/Turkey

E-mail: *mehmetsakinci@hotmail.com*

Yasemin Taşçı

Kütahya Health Sciences University, Kütahya/Turkey

E-mail: *yasemin.tasci@ksbu.edu.tr*

Murat Api

University of Health Sciences Turkey, Kartal Dr. Lutfi Kırdar Training and Research Hospital, Hamidiye International Faculty of Medicine, İstanbul/Turkey

E-mail: *muratapi@hotmail.com*

Taner Usta

Acıbadem Mehmet Ali Aydınlar University, Acıbadem Altunizade Hospital, Department of Obstetrics and Gynecology, İstanbul/Turkey

E-mail: *drtanerusta@gmail.com*

Cihan Kaya

İstanbul Aydın University, Department of Obstetrics and Gynecology, İstanbul/Turkey

E-mail: *drcihankaya@gmail.com*

Yüksel Arıkan Onaran

University of Health Sciences Turkey, Ankara Bilkent City Hospital, Ankara/Turkey

E-mail: *dryukselonaran@yahoo.com*

Erol Tavmergen

Ege University Faculty of Medicine, Department of Obstetrics and Gynecology, İzmir/Turkey

E-mail: *etavmergen@gmail.com*

Ilgın Türkçüoğlu

Gaziantep City Hospital, Gaziantep/Turkey

E-mail: *dr.ilgin@yahoo.com*

Deniz Balsak

Mardin Artuklu University Faculty of Medicine, Mardin/Turkey

E-mail: *denizbalsak@gmail.com*

Sefa Arlıer

University of Health Sciences Turkey, Adana City Hospital, Adana/Turkey

E-mail: *sefaarlier@gmail.com*

Hakan Nazik

Private Clinic, Department of Obstetrics and Gynecology, Adana/Turkey

E-mail: *drhakannazik@gmail.com*

Pınar Yalçın Bahat

Private Clinic, İstanbul/Turkey

E-mail: *dr_pinaryalcin@hotmail.com*

Ulun Uluğ

Haliç University Faculty of Medicine, Department of Obstetrics and Gynecology, İstanbul/Turkey

E-mail: *uulug@hotmail.com, ulunulug@halic.edu.tr*

Hatice Gönbe

Private Clinic, Ankara/Turkey

E-mail: *bilgi@haticeisik.com*

Fazıl Avcı

Ministry of Health Şırnak State Hospital, Şırnak/Turkey

E-mail: *fazilavci01@hotmail.com*

Yusuf Madendağ

Erciyes University Faculty of Medicine, Department of Obstetrics and Gynecology, Erciyes/Turkey

E-mail: *yusufmadendag@gmail.com*

ANATOLIAN JOURNAL OF
OBSTETRICS & GYNECOLOGY
RESEARCH

Years | 2026 | Volume | 3 | Issue | 2 | Month | August

Please refer to the journal's webpage (https://www.anajog.com/) for "Aims and Scope," "Instructions to Authors," and "Journal Policies." The editorial and publication processes of the Anatolian Journal of Obstetrics and Gynecology Research are shaped in accordance with the guidelines of ICMJE, WAME, CSE, COPE, EASE, and NISO. The journal is in conformity with the Principles of Transparency and Best Practice in Scholarly Publishing.

The Anatolian Journal of Obstetrics and Gynecology Research is currently indexed in **SOBIAD, Gale, Türk Medline, Türkiye Atıf Dizini, Idealonline, J-Gate, EBSCO,** and **TR Dizin**.

The journal is published online.

Owner: International Society of Gynecology and Obstetrics

Responsible Manager: Eray ÇALIŞKAN

Contact
Ömerağa Mahallesi, Alemdar Cad. Soydan İş Merkezi Kat: 4 Nasıl: 122/A İzmit, Kocaeli/Turkey
Phone: +90 530 339 50 55 x E-mail: ujod@ujod.org

Publisher Contact
Address: Molla Gürani Mah. Kaçamak Sk. No: 21/1 34093 İstanbul, Turkey
Phone: +90 (530) 177 30 97 / +90 (212) 621 99 25
E-mail: info@galenos.com.tr/yayin@galenos.com.tr
Web: www.galenos.com.tr

Publisher Certificate Number: 14521
Online Publication Date: September 2026
E-ISSN: 3023-7920
International periodical journal published three times in a year.

ANATOLIAN JOURNAL OF
OBSTETRICS & GYNECOLOGY
RESEARCH

Years | 2026 | Volume | 3 | Issue | 2 | Month | August

CONTENTS

ANATOLIAN JOURNAL OF
OBSTETRICS & GYNECOLOGY
RESEARCH

Years | 2026 | Volume | 3 | Issue | 2 | Month | August

CONTENTS

Guideline

DOI: 10.4274/anajog.galenos.2026.30306

Anat J Obstet Gynecol Res 2026;3(2):62-78

National Consensus Statement on Aesthetic Vaginoplasty A Modified Delphi Study by the Pelvic Floor and Cosmetic Gynecology Association (PETKOZ) of Turkey

Ozan Doğan[1], Murat Yassa[2], Aşkı Ellibeş Kaya[3], Eray Çalışkan[4], Pınar Kadiroğulları[5], Erhan Hüseyin Cömert[6], Cavide Ali[7], Ebru Alper[8], Kübra Bağcı[9], Melih Bestel[10], Pınar Birol İlter[11], Sinem Bostan[2], Sezen Bozkurt Köseoğlu[12], Yasin Ceylan[12], Telal Doğruel[13], Murat Emanetoğlu[5], Süleyman Eserdağ[14], Gökçe Gökkaya[2], Sevtap Hamdemir Kılıç[15], Ümran Karabulut Doğan[16], Ozan Karadeniz[17], Cihan Kaya[3], Özgür Leylek[18], Ali Mutlu[2], Nurullah Peker[19], Hanifi Şahin[20], Mehmet Baki Şentürk[21], Akın Sivaslıoğlu[22], İnci Süleymanova[2], Özgüç Takmaz[23], Elif Uçar[10]

[1]İstanbul Nişantaşı University Faculty of Medicine, Department of Obstetrics and Gynecology, İstanbul, Turkey
[2]Acıbadem Kartal Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Turkey
[3]İstanbul Aydın University Faculty of Medicine, Department of Obstetrics and Gynecology İstanbul, Turkey
[4]Private Clinic, Clinic of Obstetrics and Gynecology, Kocaeli, Turkey
[5]Private Clinic, Clinic of Obstetrics and Gynecology, İstanbul, Turkey
[6]Biruni University Faculty of Medicine, Department of Obstetrics and Gynecology, İstanbul, Turkey
[7]Acıbadem Kadıköy Dr. Şinasi Can Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Turkey
[8]VKV American Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Turkey
[9]Private Clinic, Clinic of Obstetrics and Gynecology, İzmir, Turkey
[10]İstanbul Esenyurt University Faculty of Health Sciences, Department of Midwifery, İstanbul, Turkey
[11]Kütahya City Hospital, Clini of Obstetrics and Gynecology, Kütahya, Turkey
[12]Private Clinic, Clinic of Obstetrics and Gynecology, Muğla, Turkey
[13]VM Medical Park Maltepe Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Turkey
[14]Altınbaş University Faculty of Medicine, Department of Gynecology and Obstetrics, İstanbul, Turkey
[15]Private Clinic, Clinic of Obstetrics and Gynecology, Ankara, Turkey
[16]Acıbadem Kayseri Hospital, Clinic of Obstetrics and Gynecology, Kayseri, Turkey
[17]University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital, Clinic of Gynecology and Obstetrics, İstanbul, Turkey
[18]Private Clinic, Female Aesthetic Genital Academy, Clinic of Obstetrics and Gynecology, İstanbul, Turkey
[19]Dicle University Faculty of Medicine, Department of Obstetrics and Gynecology, Diyarbakır, Turkey
[20]İstanbul Health and Technology University Faculty of Medicine, Department of Obstetrics and Gynecology, İstanbul, Turkey
[21]Tekirdağ Namık Kemal University, Department of Obstetrics and Gynecology, Tekirdağ, Turkey
[22]International Training Center on Urogynecology & Functional Female Genital Esthetics, İzmir, Turkey
[23]Acıbadem Mehmet Ali Aydınlar University Faculty of Medicine, Department of Obstetrics and Gynecology, İstanbul, Turkey

ABSTRACT

Purpose: Aesthetic vaginoplasty, defined as surgical vaginal tightening performed for cosmetic, sexual, and/or functional reasons without a primary medical indication, has become increasingly common in gynecological practice. Despite growing demand, no validated clinical guidelines exist to govern patient selection, surgical technique, perioperative management, or outcome assessment for these procedures. The Pelvic Floor and Cosmetic Gynecology (PETKOZ) Association of Turkey initiated this project to develop the first, expert-based consensus statement on aesthetic vaginoplasty.

Methods: A modified Delphi methodology was used across three successive rounds, with a 100% response rate in each round. Twenty-six nationally recognized experts in gynecology, urogynecology, pelvic floor surgery, and related disciplines participated. A steering committee developed an initial statement pool across 14 clinical domains: contraindications; preoperative assessment; indications; surgical principles and goals; technical selection; intraoperative management; revision surgery; complications; perioperative management; postoperative care; outcomes; guidelines and future directions; advanced technical topics; and ethics and psychosexual evaluation. Consensus was defined as ≥70% agreement. All rounds were conducted anonymously via Google Forms.

Address for Correspondence: Prof. Dr, Ozan Doğan, İstanbul Nişantaşı University Faculty of Medicine, Department of Obstetrics and Gynecology, İstanbul, Turkey

E-mail: ozandogan02@hotmail.com **ORCID ID:** orcid.org/0000-0002-0016-8749

Received: 20.05.2026 **Accepted:** 21.06.2026 **Publication Date:** 02.09.2026

Cite this article as: Doğan O, Yassa M, Ellibeş Kaya A, et al. National consensus statement on aesthetic vaginoplasty a modified delphi study by the Pelvic Floor and Cosmetic Gynecology Association (PETKOZ) of Turkey. Anat J Obstet Gynecol Res. 2026;3(2):62-78

ABSTRACT

Results: A total of 211 statements were evaluated across 14 domains. Consensus was reached on 198 statements (93.8%); 13 statements (6.2%) across six domains did not meet the threshold. Agreement rates ranged from 70% to 100%, with majority of statements reaching 88% or above. Strong consensus was obtained on comprehensive preoperative assessment, individualized technique selection, the importance of functional alongside aesthetic outcomes, standardized perioperative protocols, and the need for formal training programs. Statements that did not reach consensus included whether subjective vaginal laxity alone is a sufficient surgical indication, the role of partner expectations in decision-making, the use of local anesthesia, and suitability of same-day discharge.

Conclusion: This consensus statement, developed by PETKOZ through structured expert consultation, provides a practical guidance framework for the safe implementation of aesthetic vaginoplasty in clinical practice. Given the limitations of the current evidence base, this document is intended as a reference for clinicians until data from well-designed prospective studies become available.

Keywords: Vaginoplasty, vaginal tightening surgery, vaginal rejuvenation, perineoplasty, cosmetic gynecology, vaginal laxity, national guideline, pelvic floor, sexual function

INTRODUCTION

Aesthetic vaginoplasty is a surgical procedure aimed at tightening the vaginal canal and reconstructing the perineal body, most commonly performed in women reporting vaginal laxity following vaginal delivery, ageing, or menopause.[1,2] The procedure is decribed under various names in the literature, including vaginoplasty, perineoplasty, colpoperineoplasty, and vaginal recalibration, but all share the common goal of reducing vaginal diameter and restoring introital support.[3,4]

Demand for these procedures has increased considerably in recent years. In 2021, over 3,000 vaginoplasties were reported by plastic and cosmetic surgeons in the United States alone, representing a 374% increase from the previous year, and this figure does not include procedures performed by obstetricians and gynecologists.[5] Women sought vaginal tightening primarily to increase vaginal friction during intercourse, improve genital self-image, and address orgasmic concerns.[6,7] Despite growing patient demand and increasing surgical volume, the evidence base for aesthetic vaginoplasty remains limited, with available studies showing considerable variability in surgical techniques, outcome measures, and follow-up duration that prevents definitive conclusions on efficacy and safety.[8]

Currently, no country has established nationally endorsed clinical guidelines to govern patient selection, surgical indications, preoperative evaluation, intraoperative principles, or postoperative management for aesthetic vaginoplasty. This absence of structured guidance leaves clinicians without a common reference standard in a field where medicolegal, ethical, and functional considerations are substantial.

The objective of this study was to report an expert-based consensus statement developed by Pelvic Floor and Cosmetic Gynecology (PETKOZ) Association to help guide the adoption of aesthetic vaginoplasty into clinical practice with respect to 14 key domains. Until further high-level evidence becomes available, this statement is intended to provide appropriate guidance for the safe and standardized implementation of aesthetic vaginoplasty in Turkey, drawing on the collective experience of nationally recognized experts in this field.

METHODS

Purpose and Scope

This consensus statement was prepared to provide expert-based practical guidance for clinicians involved in the care of women presenting with vaginal laxity. The document covers patient assessment, selection criteria, surgical management, perioperative care, and outcome evaluation for aesthetic vaginoplasty, and was developed by the PETKOZ through a structured modified Delphi process. It is intended for use by gynecologists, urogynecologists, pelvic floor surgeons, cosmetic gynecologists, and other healthcare professionals who participate in the perioperative management of these patients.

The scope of this document is limited to elective aesthetic vaginoplasty performed for cosmetic, sexual, or functional reasons in women without a primary medical indication. Patient selection, preoperative evaluation, technique selection, intraoperative principles, perioperative management, complication prevention, and outcome assessment are all addressed. Procedures carried out for pelvic organ prolapse, stress urinary incontinence, or other urogynecological conditions with existing evidence-based guidelines are not within scope, nor are procedures for congenital anomalies, obstetric or traumatic injury, malignancy, or gender affirmation. Sexual dysfunction reported without demonstrable vaginal laxity does not constitute a standalone surgical indication; when sexual dysfunction accompanies laxity as a contributing complaint, psychosexual evaluation is required before any surgical decision is made.

Population and Setting

The recommendations apply to adult women aged 18 years or older with objectively or subjectively demonstrable vaginal laxity. Four clinical presentations are within scope. The first is postpartum vaginal laxity following obstetric trauma, perineal body injury, or loss of introital support after vaginal delivery. The second is laxity associated with menopausal changes, where estrogen deficiency leads to progressive loss of vaginal wall tone and tissue quality. The third presentation considered was functional laxity without concurrent pelvic organ prolapse, in

which women report reduced vaginal friction, altered sensation, or impaired sexual function that is attributed to laxity rather than a structural support defect. The fourth encompasses situations where aesthetic or psychosocial concerns are the principal reason for seeking surgery, in the absence of a primary gynecological or urogynecological indication. The guidance is relevant to outpatient clinic, ambulatory surgical, and hospital-based settings.

Background

Despite growing procedural volume, the evidence base for aesthetic vaginoplasty remains sparse. The most recent systematic review on the topic identified only 11 eligible studies with a combined 806 patients published between 2006 and 2024, and the available data showed considerable variability in surgical techniques, patient selection criteria, outcome measures, and follow-up duration, making it impossible to draw definitive conclusions on efficacy or safety.[8] Randomized controlled trial data are largely absent, and most published evidence comes from retrospective or prospective uncontrolled case series that are inherently susceptible to selection bias.

Alongside the scarcity of evidence, the field lacks any universally accepted procedural classification, standardized technical terminology, or shared decision-making framework. Procedures performed under the broad label of "aesthetic vaginoplasty" include perineoplasty, posterior colporrhaphy-based repairs, levator ani-plication, mucosal excision techniques, energy-based adjunctive modalities, and hybrid reconstructive combinations, yet these are routinely reported under inconsistent or interchangeable nomenclature.[3,4] No internationally endorsed guideline or clinical algorithm exists to guide patient selection, technique choice, perioperative management, or outcome assessment. As a result, practice is largely shaped by individual training and personal experience, and considerable variation persists in surgical approach, patient selection criteria, perioperative protocols, and the instruments used to measure outcomes.

It was this combination of limited evidence and fragmented practice that motivated PETKOZ to undertake the present project. The aim was to bring together a group of nationally recognized specialists and, through a structured Delphi process, identify where expert opinion converges, establish a shared reference framework for clinical decision-making, and produce recommendations that practitioners can rely on until prospective comparative data become available. This statement is not intended as a substitute for high-quality evidence, which all participants recognized as the field's most pressing need. Rather, it represents a first step toward a common clinical language and a shared set of principles from which future research, outcome registries, and evidence-based guidelines can develop.

Nomenclature

Several terms have been used interchangeably in both the scientific literature and clinical practice to describe surgical procedures for vaginal laxity, including "aesthetic vaginoplasty," "vaginal tightening surgery," and "vaginal rejuvenation." These terms are not conceptually or procedurally equivalent, and their interchangeable use has contributed to heterogeneity in patient selection criteria, outcome reporting, and evidence synthesis. In accordance with terminology frameworks endorsed by the International Urogynecological Association (IUGA), the International Continence Society, and the American College of Obstetricians and Gynecologists (ACOG), this manuscript adopts the term "aesthetic vaginoplasty" as the preferred nomenclature. This term encompasses the surgical correction of vaginal laxity for cosmetic, sexual, and/or functional reasons in the absence of a primary medical indication and reflects the procedural and functional dimensions of the surgery more precisely than the commercially derived term "vaginal rejuvenation," which lacks a standardized clinical definition and has been identified by ACOG as a non-specific and potentially misleading descriptor.[7] The term "vaginal tightening surgery" is used in this manuscript as a neutral descriptive synonym where appropriate.

Delphi Survey Design

This consensus statement was developed using a modified Delphi methodology incorporating three successive rounds.[9,10] The study adhered to the Guidance on Conducting and Reporting Delphi Studies reporting guidelines.[11] All three rounds were completed by all 26 participating panelists, giving a response rate of 100% in each round. Anonymity was maintained throughout the entire process: each participant received a unique, individually addressed survey link for every round, responses were submitted directly to a secure platform, and individual response data were not accessible to other panel members or to the steering committee until aggregate compilation was complete. Given the absence of universally accepted objective measurement systems for vaginal laxity, expert consensus methodology was considered particularly appropriate for this field, where clinical decision-making continues to rely substantially on accumulated specialist experience in the absence of high-quality comparative evidence. Ethics approval was not required for this study, as it did not involve the collection of patient data or direct clinical intervention with human subjects. Participation was voluntary, responses remained anonymous throughout all rounds, and all procedures complied with institutional ethical standards. Survey distribution and anonymous data storage were managed via Google Forms (Google LLC, Mountain View, CA, USA).

Questionnaire Development

A steering committee comprising senior members of PETKOZ with expertise in pelvic floor surgery, vaginal reconstructive surgery, and cosmetic gynecology was responsible for the conceptual development and overall direction of the project. To inform statement development, a structured literature search was conducted in PubMed/MEDLINE, EMBASE, and the Cochrane Central Register of Controlled Trials, with no restriction on publication start date and a search cutoff of May 2026. The following search terms were applied individually and in Boolean combination:"vaginoplasty," "vagina surgery," "vaginal laxity," "vaginal tightening," "vaginal rejuvenation,"

"vaginal relaxation," "vaginal recalibration," "perineoplasty," "perineum surgery," "colpoperineoplasty," "colporrhaphy," "cosmetic surgery," "gynecologic surgical procedures," "female genital cosmetic surgery," "sexual dysfunction," "sexual function," "dyspareunia," "pelvic floor surgery," "pelvic organ prolapse surgery," "urinary incontinence stress surgery," "reconstructive surgical procedures," "patient satisfaction," and "quality of life". Eligible publications included original research articles, systematic reviews, meta-analyses, and clinical practice guidelines published in English. Case reports, letters, editorials, conference abstracts, and non-peer-reviewed sources were excluded. Reference lists of all identified articles and relevant society position statements were hand-searched to capture additional studies not retrieved by the electronic search. The committee reviewed the retrieved evidence with a focus on patient selection, surgical technique, perioperative management, clinical outcomes, and safety. In view of the limited and heterogeneous nature of the available evidence, supplementary guidance was sought from position statements and committee opinions issued by major international organisations, including the ACOG, the Royal College of Obstetricians and Gynaecologists (RCOG), the American Urogynecologic Society (AUGS), and the IUGA. Based on this synthesis and on the collective clinical experience of the steering committee, an initial pool of candidate statements was developed and organized across 14 clinical domains: (1) contraindications; (2) preoperative assessment; (3) indications; (4) surgical principles and goals; (5) technical selection; (6) intraoperative management; (7) revision surgery; (8) complications; (9) perioperative management; (10) postoperative care; (11) clinical outcomes; (12) guidelines and future directions; (13) advanced technical and clinical topics; and (14) ethics and psychosexual evaluation. The ethics and psychosexual evaluation domain was developed as a supplementary module following completion of the initial 13-domain statement pool. Given the smaller number of candidate statements in this domain, it was administered to a subset of 19 panelists who indicated specific expertise in psychosexual medicine or female genital cosmetic surgery ethics. Participants rated their agreement with each statement on a 5-point Likert scale (strongly disagree, disagree, neutral, agree, strongly agree).

Expert Panel

Nationally recognized experts in gynecology, urogynecology, pelvic floor surgery, and related disciplines with documented clinical experience in vaginal reconstructive and/or aesthetic pelvic procedures were invited via email to participate. All experts provided informed consent. A total of 26 experts agreed to participate and completed all three rounds of the survey, giving a retention rate of 100%. Panel demographic characteristics are summarized in Table 1.

Modified Delphi Process and Consensus Development

The modified Delphi process is illustrated in Figure 1. Each participant received a unique anonymous survey link by email for every round. All 26 panelists completed all three rounds, with a response rate of 100% per round and no missing data.

Table 1. Demographic characteristics of the expert panel (n=26)

Variable	n	%
Age group		
30-39 years	4	15.4
40-49 years	14	53.8
50-59 years	3	11.5
≥60 years	5	19.2
Sex		
Male	16	61.5
Female	10	38.5
Academic title		
Specialist	9	34.6
Assistant professor	1	3.8
Associate professor	6	23.1
Professor	10	38.5
Primary clinical specialty		
General gynecology	9	34.6
Cosmetic gynecology	9	34.6
Urogynecology	5	19.2
Pelvic reconstructive surgery	2	7.7
Other	1	3.8
Additional subspecialty training[a]		
Aesthetic/cosmetic surgery	24	92.3
Laser and energy-based devices	15	57.7
Psychosexual medicine	5	19.2
Pelvic floor physiotherapy	2	7.7
None	2	7.7
Disciplines in clinical practice[a]		
Gynecology	26	100.0
Urogynecology	24	92.3
Pelvic floor surgery	21	80.8
Aesthetic and plastic surgery	15	57.7
Psychosexual medicine	7	26.9
Pelvic floor physiotherapy	5	19.2
Other	3	11.5
Total years of surgical experience		
5-10 years	5	19.2
11-20 years	10	38.5
>20 years	11	42.3
Annual vaginoplasty/perineoplasty volume		
10-25 procedures/year	6	23.1
26-50 procedures/year	10	38.5
>50 procedures/year	10	38.5
Institution type		
Private clinic or hospital	22	84.6
University hospital	3	11.5
Public hospital	1	3.8

Table 1. Continued		
Variable	**n**	**%**
Geographic region (Turkey)		
Marmara	15	57.7
Aegean	6	23.1
Black Sea	2	7.7
Central Anatolia	2	7.7
Southeastern Anatolia	1	3.8
Country of surgical training		
Turkey	26	100.0
SCI-indexed publications (pelvic surgery)		
0-5	12	46.2
6-15	7	26.9
16-30	4	15.4
>30	3	11.5
Prior Delphi/consensus study participation		
Yes	10	38.5
No	16	61.5
PETKOZ membership duration		
<2 years	1	3.8
2-5 years	14	53.8
6-10 years	5	19.2
Not a member	6	23.1

[a]Multiple responses permitted; percentages exceed 100%
PETKOZ: Pelvic Floor and Cosmetic Gynecology Association

Aggregate results from rounds 1 and 2 were compiled by the steering committee and circulated to all panelists as feedback before the subsequent round, enabling experts to reconsider their responses in light of group opinion. Between rounds, statements that remained below the consensus threshold were reviewed by the steering committee: wording was revised where panel comments indicated that ambiguity or lack of clarity had contributed to disagreement, while statements reflecting genuine substantive controversy were retained unmodified and resubmitted to the panel with a summary of the preceding round's distribution of responses. Response stability between consecutive rounds was monitored by comparing the proportion of panelists endorsing each statement across rounds; statements showing substantial shifts in agreement rates prompted targeted discussion at the virtual consensus meeting. Agreement scores for each statement were summarized using median Likert scores and the interquartile range (IQR), with the IQR serving as a measure of response dispersion and convergence across rounds. Prior to round 3, a virtual consensus meeting was held to allow structured discussion of statements that had not yet reached the predefined threshold. The meeting was moderated by a member of the steering committee who had not contributed to statement development, to preserve neutrality. Each borderline statement was presented with its round 2 aggregate response distribution, and panelists were invited to articulate their reasoning before revoting anonymously. Statements were revised in wording if the discussion identified specific language as the source of disagreement; statements where disagreement reflected divergent clinical opinion rather than interpretive ambiguity were retained unchanged and subjected to a final anonymous vote. Following this meeting, round 3 was administered anonymously. No sensitivity analyses were performed, as the 100% response rate across all rounds precluded missing-data scenarios that would have required such procedures. The final manuscript was drafted by the core group and reviewed and approved by all participating experts.

Consensus was defined as ≥70% agreement of the experts, corresponding to a response of agree or strongly agree on the 5-point Likert scale. A threshold of ≥70% was selected in accordance with previously published Delphi-based consensus methodologies in clinical guideline development,[9,10] reflecting a balance between the requirement for meaningful expert convergence and the recognition that absolute unanimity is rarely achievable in fields where evidence is limited and practice variation is substantial. To facilitate interpretation of agreement levels across domains, consensus strength was categorized as follows: very strong consensus for agreement rates above 95%, strong consensus for rates of 85 to 95%, and moderate consensus for rates of 70 to 85%. Statements achieving consensus were finalized and not reconsidered in subsequent rounds. Statements failing to reach the threshold after round 3 are reported separately.

RESULTS

Delphi Survey Rounds

Twenty-six experts completed all three rounds of the modified Delphi process, with a retention rate of 100%. A total of 211 statements were evaluated across 14 clinical domains. Of these, 198 statements (93.8%) reached the predefined consensus threshold of 70% or above (Supplementary Table 1). Thirteen statements (6.2%) did not reach consensus and are presented in Supplementary Table 2. Agreement rates for statements that achieved consensus ranged from 70.0% to 100%, with the majority reaching 88% or above. For interpretive purposes, consensus strength was categorized as very strong (agreement above 95%), strong (85 to 95%), or moderate (70 to 85%), as defined in the Methods section. Domain-level agreement statistics are summarized in Supplementary Table 3. Domain-level Likert response statistics and Kendall's coefficient of concordance are reported in Supplementary Table 4. Domain-level consensus rates are illustrated in Figure 2.

Consensus Outcomes

Across multiple domains, the panel consistently prioritized preservation of vaginal function, neurovascular integrity, tissue quality, and psychosexual wellbeing over excessive narrowing or purely cosmetic goals. This recurrent theme represents one of the strongest conceptual messages of the consensus and is described in the domain-specific findings below.

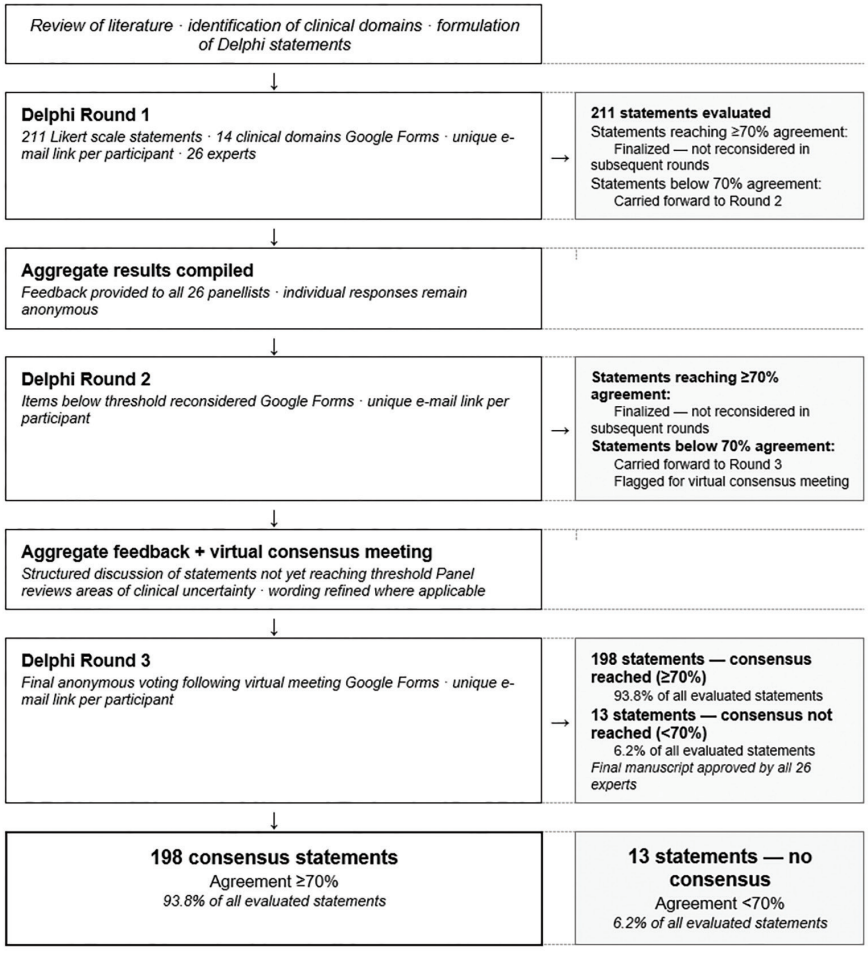


Figure 1. Modified Delphi consensus process for the development of the PETKOZ consensus statement on aesthetic vaginoplasty. A total of 211 statements across 14 clinical domains were evaluated over three successive rounds by 26 nationally recognized experts. Statements achieving the predefined threshold of ≥70% agreement at each round were finalized and not re-evaluated in subsequent rounds. A virtual consensus meeting was held prior to round 3 to allow structured discussion of statements that had not yet reached the threshold. The final manuscript was drafted by the core group and approved by all participating experts

PETKOZ: Pelvic Floor and Cosmetic Gynecology Association

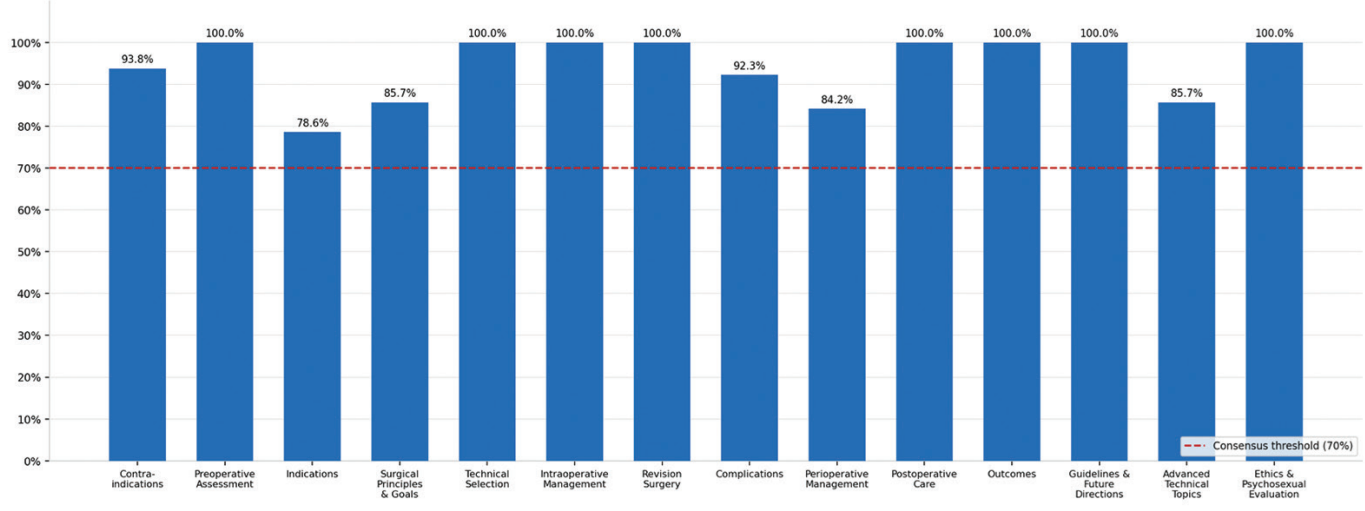


Figure 2. Consensus agreement per domain (%). n=26 experts; consensus threshold ≥70% (dashed red line); percentages represent proportion of statements reaching consensus within each domain

Contraindications to aesthetic vaginoplasty

Accurate patient selection requires clear identification of surgical contraindications, particularly in elective procedures where the risk-benefit balance must be carefully established before any surgical commitment. Consensus was reached on 15 of 16 evaluated statements. Absolute contraindications with unanimous or near-unanimous agreement included pregnancy and the early postpartum period, a history of pelvic radiotherapy, connective tissue disease with impaired wound healing, unrealistic patient expectations, active genital herpes, age under 18 years, and stage III or higher pelvic organ prolapse. Body dysmorphic disorder was accepted as an absolute contraindication (84.6%), and previous vaginal surgery was recognized as a risk-modifying factor requiring increased surgical caution. In cases of high-risk human papillomavirus positivity, surgery was agreed to be appropriate only following Pap smear and colposcopic evaluation. Whether condyloma acuminata constitutes an absolute contraindication did not reach consensus. Taken together, these findings reflect a conservative and ethically grounded approach to patient selection, with particular weight placed on psychological fitness and the exclusion of structural or systemic factors that would compromise the safety or meaningfulness of an elective procedure.

Preoperative assessment and patient preparation

Structured preoperative evaluation forms the foundation of safe surgical decision-making in aesthetic vaginoplasty, addressing both anatomical and psychosexual dimensions of the patient's complaint. Consensus was reached on all 26 evaluated statements. Core requirements endorsed with near-unanimous agreement included pelvic examination, perineal body assessment, documentation of functional symptoms, and systematic review of prior pelvic and vaginal surgeries. Use of the female sexual function index (FSFI) or pelvic organ prolapse/urinary incontinence sexual questionnaire-12 (PISQ-12), pelvic organ prolapse quantification (POP-Q) staging, and the vaginal laxity questionnaire all met the threshold. Pelvic floor muscle function assessment, psychosexual evaluation in selected patients, and hormonal preparation in postmenopausal or atrophic tissue were similarly endorsed. A standardized informed consent form was unanimously required, including explicit discussion of potential effects on sexual function; patients must be informed of the limited and heterogeneous evidence base prior to surgery (92.3%). These findings signal increasing recognition within the field that psychosexual and functional factors are as central to preoperative assessment as anatomical examination, and that informed consent in this setting carries an ethical obligation concerning transparency about the limits of the available evidence.

Surgical indications

Defining appropriate surgical candidacy in the absence of a validated objective measure of vaginal laxity is one of the most debated areas in aesthetic vaginoplasty, and this was reflected in the pattern of consensus across this domain. Consensus was reached on 11 of 14 evaluated statements. Objectively demonstrable vaginal laxity was accepted as a surgical indication, and its combination with subjective patient complaint was agreed to constitute a definitive indication. Perineal body defect, obstetric trauma and perineal damage, recurrent vaginal relaxation, and aesthetic concerns in appropriately selected patients were accepted as valid indications. Symptom severity and its impact on quality of life were agreed to be determinative criteria in the surgical decision (96.2%). Surgery must not be recommended to resolve relationship problems or with the expectation of improving partner fidelity, and the surgical decision must be based solely on patient autonomy (96.2%). Three statements did not reach consensus: subjective vaginal laxity alone as a sufficient indication (30.8%), sexual dysfunction or psychosexual distress alone as justification for surgery (57.7%), and partner expectation as a determinative factor in decision-making (19.2%). The failure to reach consensus on these three items is clinically and ethically significant. The panel's rejection of subjective laxity, unaccompanied psychosexual distress, and partner influence as standalone surgical justifications reflects a collective position that surgery should not proceed without objective clinical correlation and independently verified patient autonomy. These are among the strongest ethical contributions of the present consensus, establishing that aesthetic vaginoplasty must be grounded in clinical findings rather than subjective or externally driven concerns alone. The disagreement in these areas likely reflects ongoing variability in how individual practitioners weigh functional and psychological evidence, differing training backgrounds, and the absence of validated objective tools for measuring laxity, all of which make consistent threshold-setting difficult in current practice.

Surgical principles and goals

Agreement on core surgical principles reflects a shared conviction that functional preservation must guide aesthetic modification rather than be subordinate to it. Consensus was reached on 18 of 21 evaluated statements. Preservation of natural anatomical structures and vaginal function was unanimously identified as the primary surgical goal, and functional outcomes were agreed to carry equal weight to aesthetic outcomes. Excessive muscle plication was unanimously identified as a risk factor for postoperative dyspareunia, conservative mucosal resection was endorsed, and protection of neurovascular structures was unanimously required. Vaginal canal diameter and length preservation, vaginal axis maintenance, and combined perineoplasty in selected patients were similarly endorsed. Pelvic floor physiotherapy was supported as first-line treatment before surgical referral (92.3%). Three statements did not reach consensus: energy-based devices as a direct surgical alternative (26.9%), an approach focused solely on aesthetics without functional evaluation (19.2%), and routine levator ani plication in selected patients (69.2%). These three non-consensus items are conceptually coherent: the panel declined to endorse any approach that circumvents functional evaluation, whether by substituting energy devices for surgery, omitting functional assessment entirely, or applying levator

plication routinely without individualized indication. Taken together, the findings in this domain represent a clear departure from older tightening-centred surgical paradigms and signal a modern reconstructive philosophy in which neurovascular protection, conservative tissue handling, and preservation of vaginal compliance are non-negotiable priorities. The disagreement around energy-based devices and levator plication specifically reflects the limited comparative evidence for these modalities, which has not yet reached a level that would support their routine endorsement by a nationally representative expert panel.

Surgical technique selection

Technique selection in vaginal tightening surgery cannot be standardized across patients, as anatomical and functional profiles differ considerably, based on obstetric history, tissue quality, and the location and severity of the defect. Consensus was reached on all 11 evaluated statements. The degree and distribution of vaginal laxity were agreed to be determinative in technique selection, with a limited approach considered sufficient in isolated introital laxity. A technique based solely on distal tightening was not accepted as appropriate when proximal vaginal support defects are present. Combined assessment of the perineal body, posterior fourchette, and perineal fascia was agreed as required during planning. Scar tissue and tissue loss, prior pelvic or vaginal surgery, atrophic tissue, concomitant pelvic organ prolapse or incontinence, and the risk of excessive tightening were all recognized as technique-modifying factors. Individualized planning based on anatomical findings, functional goals, and patient expectations was unanimously endorsed. Based on these consensus findings, a procedural classification framework was developed that organizes aesthetic vaginoplasty into five standardized technique categories: (1) perineoplasty, indicated for isolated introital laxity and perineal body defects; (2) posterior colporrhaphy-based repair, for mid and distal vaginal laxity with conservative mucosal excision and fascial plication; (3) levator ani plication, reserved for selected cases with central support defects; (4) energy-based adjunctive procedures, including fractional CO_2 laser and radiofrequency, applied as adjuncts rather than primary surgical replacements; and (5) hybrid reconstructive approaches, combining the above techniques in patients with multi-compartment defects or concurrent pelvic organ prolapse. A schematic algorithm presenting this classification system and the technique selection pathway is provided in Figure 3. The panel strongly endorsed individualized reconstruction rather than standardized tightening approaches, recognizing that no single technique is universally appropriate and that reconstruction must be guided by the anatomical and functional heterogeneity of each patient's presentation.

Intraoperative management

Intraoperative decisions regarding tissue handling, suture technique, and the degree of tightening are direct determinants of both short-term complications and long-term functional outcomes. Consensus was reached on all 19 evaluated statements. Suture material selection was agreed to affect complication rates, layered closure to improve wound healing, and dorsal lithotomy to be the most appropriate patient position. Excessive electrocautery use was agreed to increase tissue damage and complication risk. Intraoperative assessment of vaginal canal diameter and symmetry was unanimously required, and the degree of tightening must be confirmed before wound closure to prevent dyspareunia risk. Closure of vaginal mucosa under excessive tension was agreed to adversely affect wound healing, correct anatomical plane identification during perineal body reconstruction was unanimously agreed to improve surgical success, and vaginal axis preservation throughout the procedure was unanimously endorsed as important for functional outcomes. The consistent emphasis across these findings on avoiding overcorrection and preserving vaginal compliance reflects an intraoperative philosophy oriented toward functional safety rather than maximal anatomical narrowing. Notably, suture size preference for fascial and muscular layers did not reach full consensus in the advanced topics domain, indicating that while the principle of appropriate suture selection is widely endorsed, specific material recommendations require further evidence before universal adoption.

Revision surgery

Revision vaginoplasty carries distinct clinical challenges that differ substantially from those of primary procedures, including a higher complication burden and greater technical complexity due to scarring and altered tissue planes. Consensus was reached on all 13 evaluated statements. Revision surgery was unanimously agreed to carry higher complication rates than primary vaginoplasty, and a minimum healing interval of 3 to 6 months before re-operation was endorsed. More careful evaluation of patient expectations in revision cases, detailed psychosexual assessment before proceeding, and analysis of the prior surgical technique were all unanimously required. Scar tissue was agreed to increase technical complexity, re-evaluation and reconstruction of fascial planes was agreed as necessary, and revision surgery was agreed to be indicated only in the presence of a functional problem or significant aesthetic concern (88.5%). The unanimous support for mandatory psychosexual reassessment before revision surgery is particularly important, as it signals that persistent dissatisfaction alone does not constitute sufficient grounds for reoperation and that the underlying expectations and psychological context must be re-evaluated before any further surgical commitment is made.

Complications

Dyspareunia and excessive tightening represent the most clinically significant functional consequences of vaginal tightening surgery, and their prevention is central to the safe delivery of these procedures. Consensus was reached on 12 of 13 evaluated statements. Dyspareunia was agreed to be among the most important complications following vaginoplasty, and excessive tightening was unanimously agreed to require active prevention at the time of surgery. Vaginal stenosis was accepted as a rare but serious complication, and suture line dehiscence as clinically significant. Hematoma and early-period bleeding,

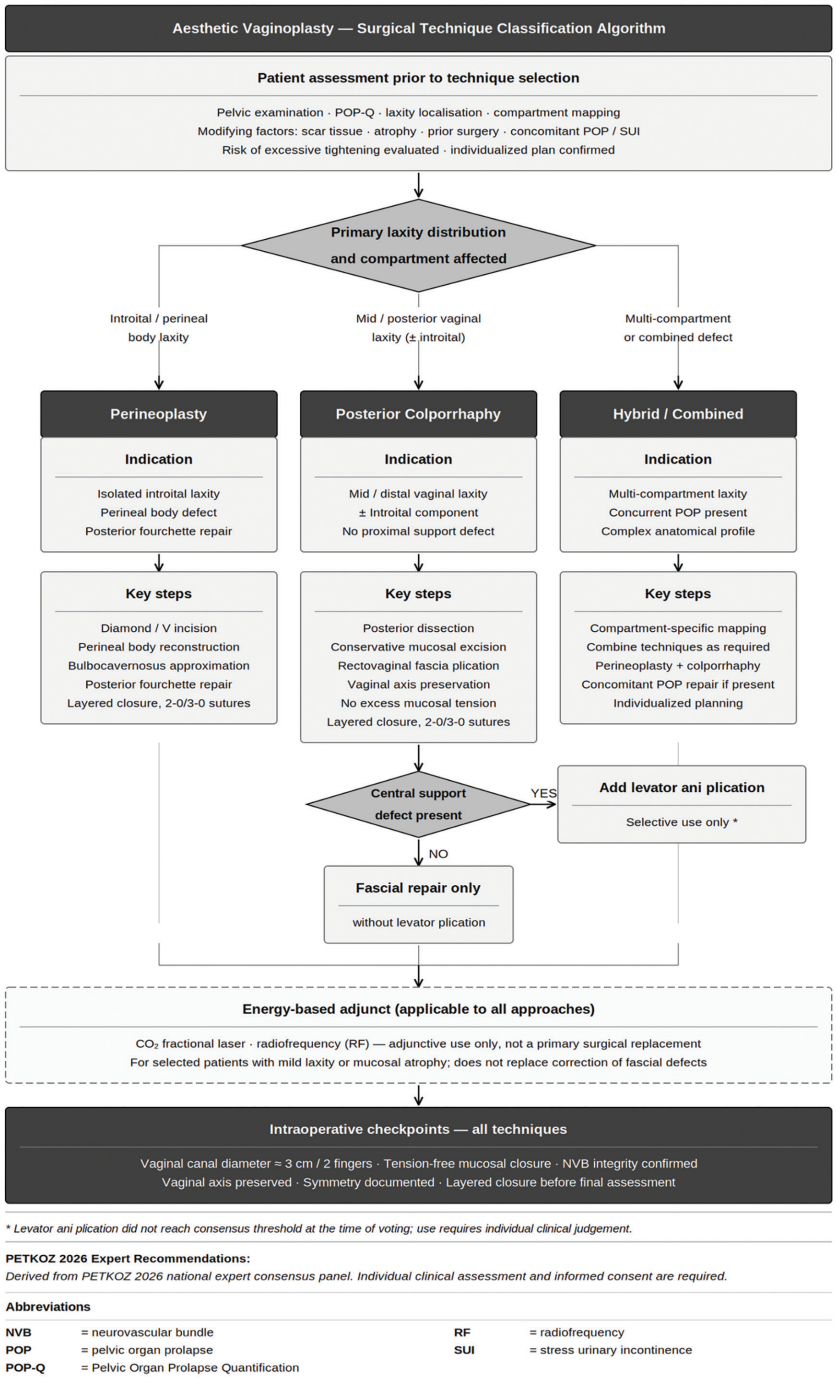


Figure 3. Surgical technique classification algorithm for aesthetic vaginoplasty, based on the PETKOZ 2026 national expert consensus panel

NVB: Neurovascular bundle, POP: Pelvic organ prolapse, POP-Q: Pelvic Organ Prolapse Quantification, RF: Radiofrequency, SUI: Stress urinary incontinence, PETKOZ: Pelvic Floor and Cosmetic Gynecology Association

low infection risk with appropriate technique, delayed wound healing in relation to technique and tissue quality, prolonged postoperative pain, and sensation changes as determinants of patient satisfaction all reached consensus. Whether patient dissatisfaction constitutes an important indicator of surgical success did not reach consensus (57.7%). The inability to reach agreement on patient dissatisfaction as a success indicator is worth noting and reflects genuine philosophical disagreement about whether subjective dissatisfaction in the absence of objective functional failure should carry clinical weight, particularly in an elective procedure with a high baseline satisfaction rate. More fundamentally, the panel's collective emphasis on dyspareunia prevention throughout this domain reframes this complication not merely as an adverse event but as a functional outcome failure, one that directly negates the sexual and quality-of-life benefits that patients seek from surgery and that must therefore be treated as a primary safety endpoint rather than a secondary concern.

Perioperative management

Standardized perioperative protocols are essential for ensuring reproducible outcomes and patient safety, and the panel identified several areas where current practice varies without clear justification. Consensus was reached on 16 of 19 evaluated statements. Routine antibiotic prophylaxis, individualized pain management, perioperative risk assessment, standardized patient preparation including antisepsis, smoking cessation counselling, and written patient information on the postoperative course were all endorsed, with several reaching unanimous agreement. Early mobilization was agreed to reduce complication risk, and the time to return to both sexual activity and intense physical activity should be standardized. Three statements did not reach consensus: whether surgery can be performed under local anesthesia (30.8%), whether routine urinary catheter and vaginal pack placement is indicated for all patients (57.7%), and whether same-day discharge is appropriate (50%). The failure to reach consensus on these three logistical items most likely reflects genuine heterogeneity in how the procedure is performed across different settings rather than disagreement on principle. Local anesthesia feasibility, catheter necessity, and discharge timing are all highly dependent on procedural complexity, institutional infrastructure, and the experience of individual surgeons, and the absence of procedure-specific published data on these management choices leaves practitioners with no shared reference standard to align against.

Postoperative care

The postoperative period is a critical determinant of wound healing, patient comfort, and functional recovery, and uniform management protocols reduce the risk of preventable complications. Consensus was reached on all 13 evaluated statements. Sexual intercourse restriction for at least 4 to 6 weeks, and scheduled follow-up visits on days 10 to 15 and day 40 for wound and suture monitoring, were unanimously endorsed. Postoperative wound care, patient education for early complication recognition, and postoperative hygiene recommendations all reached unanimous agreement. Cold application was agreed to be effective in edema management, physical activity restriction was recommended in the early period, and pelvic floor rehabilitation or dilator use in selected patients was accepted as an appropriate recommendation (73.1%). The uniformity of consensus across this domain reflects a well-developed shared understanding of postoperative management priorities and suggests that postoperative care protocols are among the most implementation-ready components of this consensus for direct clinical adoption.

Clinical outcomes

Accurate outcome assessment in aesthetic vaginoplasty requires validated measurement tools and sufficient follow-up duration, given the multifactorial nature of sexual satisfaction and the time required for complete tissue healing. Consensus was reached on all 13 evaluated statements. A minimum of 3 to 6 months was agreed to be required before evaluating final surgical outcomes, patient satisfaction was agreed to be high after vaginal tightening surgery, and long-term outcomes were agreed to be stable and sustainable. Sexual function was agreed to improve in the postoperative period (80.8%), and psychosexual improvement was agreed to be a clinically significant outcome. Concordance between preoperative expectations and postoperative outcomes was unanimously identified as the primary determinant of patient satisfaction, and functional outcomes were unanimously agreed to carry equal weight to aesthetic outcomes. Long-term follow-up data were unanimously agreed as necessary for proper efficacy assessment. Particularly noteworthy is the panel's unanimous position that patient satisfaction alone does not constitute a sufficient definition of surgical success; by placing expectation concordance and functional improvement at the center of outcome evaluation, the consensus moves toward a more rigorous and multidimensional framework for assessing what this surgery actually achieves.

Guidelines and future directions

The absence of standardized guidelines and prospective data infrastructure was unanimously identified as the most pressing limitation of current practice in aesthetic vaginoplasty. Consensus was reached on all 13 evaluated statements. A standardized clinical guideline was agreed as necessary, and a multidisciplinary approach incorporating gynecology, pelvic floor physiotherapy, and psychosexual support was unanimously endorsed as the gold standard of care. More high-quality prospective and randomized controlled studies, data registry systems for long-term outcome evaluation, a more clearly defined ethical and medicolegal framework, and integration of patient-reported outcome measures were all unanimously endorsed. Formal training and certification programs were agreed as needed (92.3%). The strength of consensus across this domain substantially increases the value of the present statement: when a panel unanimously endorses the need for multidisciplinary care, prospective research, long-term registries, and formal training, it simultaneously acknowledges the current limitations of practice and sets a clear institutional direction. These findings position the PETKOZ consensus not as a definitive endpoint but as a structured starting point for the systematic infrastructure development that the field requires.

Advanced technical and clinical considerations

Several technical aspects of aesthetic vaginoplasty remain without established standards, reflecting the limited evidence base and the heterogeneity of current practice in this area. Consensus was reached on 12 of 14 evaluated statements. Preoperative vaginal mapping to identify compartment-specific defects was agreed as appropriate, and a target postoperative vaginal diameter of approximately two examining fingers or 3 cm was agreed upon (84.6%). Absorbable 2-0 or 3-0 sutures for vaginal mucosal repair were unanimously endorsed, and Enhanced Recovery After Surgery protocol application following vaginoplasty was agreed as appropriate. The neurosensory properties of the vagina were agreed to require consideration during tightening procedures, anterior vaginal wall mucosal excisions were agreed to be kept minimal, proximal vaginal pathology was agreed to require evaluation

and management when indicated, and compartment-specific documentation of pelvic floor findings at each follow-up visit was unanimously endorsed. Two statements did not reach consensus: whether absorbable sutures of size 1 or 2 are preferred for fascial and muscle repairs (69.2%), and whether non-absorbable sutures represent an appropriate option for fascial and ligament repairs (69.2%). The failure to reach agreement on suture selection for deeper tissue layers is consistent with the broader pattern of non-consensus in this survey: it reflects not philosophical disagreement but rather the practical reality that no published comparative data exist on suture-specific outcomes in aesthetic vaginoplasty, making it impossible for even experienced surgeons to align on a shared recommendation.

Ethics and psychosexual evaluation

The ethics and psychosexual evaluation domain was developed as a supplementary module administered to 19 panelists with documented expertise in psychosexual medicine or female genital cosmetic surgery ethics. Consensus was reached on all six evaluated statements. Preoperative clinical screening for body dysmorphic disorder was endorsed in patients requesting surgery without a functional indication when risk indicators are present (94.7%), and psychosexual evaluation was recommended in cases identified through risk indicators or clinician judgement (94.7%). The potential influence of social media content on genital appearance perception was agreed to warrant individualized discussion within preoperative counselling (94.7%), and the possible shaping of surgical motivation by partner pressure was agreed to warrant thorough evaluation with deferral of the surgical decision when necessary (84.2%). The potential effects of commercial marketing content on patient expectations were agreed to be incorporated into the informed consent process (89.5%). Unanimously endorsement was achieved concerning the position that the purpose of psychosexual evaluation is not to exclude the patient from surgery but to ensure appropriate informed consent and expectation management (100.0%). The strong agreement across this domain reinforces the ethical framework established throughout the consensus: psychosexual evaluation is a clinical tool for optimizing patient benefit, not a gatekeeping mechanism, and its integration into the preoperative pathway reflects a commitment to patient-centered care, grounded in autonomy and transparency.

Synthesis of Consensus Findings

Review of consensus patterns across all 14 domains reveals several thematic consistencies that transcend individual clinical questions and reflect a shared conceptual framework underlying the panel's collective judgement. First, functional preservation was endorsed as the primary surgical objective across every domain in which it was addressed, from the prohibition of excessive plication in surgical principles to the mandatory intraoperative diameter check to the equal weighting of functional and aesthetic outcomes in clinical assessment. This recurrence across independent domains suggests not incidental agreement but a genuine paradigm shift away from tightening-centered surgical models toward a philosophy of anatomically respectful, function-preserving reconstruction. Second, patient autonomy and psychosocial integrity emerged as cross-cutting requirements: the rejection of partner-driven indications, the mandatory psychosexual evaluation, the requirement for body dysmorphic disorder screening, the prohibition of surgery to address relationship problems, and the unanimous requirement for psychosexual reassessment before revision surgery collectively define a coherent ethical framework in which surgical access is conditional on independently verified, clinically grounded patient motivation. Third, the distribution of non-consensus items is itself informative. All areas of disagreement involved either logistical variables dependent on institutional and procedural context, such as anesthetic approach, catheterization, and same-day discharge, or technical choices for which no comparative evidence exists, such as suture size for deep tissue repair, energy devices as primary alternatives, and routine levator plication. The absence of non-consensus on core safety, ethical, and functional principles indicates that disagreement was bounded to operationally variable domains rather than foundational ones. Taken together, these patterns position the present consensus as a clinically actionable framework grounded in shared principle, while clearly delineating the evidence gaps that future research must resolve.

DISCUSSION

This study presents the first national expert-based consensus statement on aesthetic vaginoplasty developed by PETKOZ, covering 14 clinical domains and 211 statements evaluated by 26 nationally recognized experts. The only exception was the ethics and psychosexual evaluation domain which was assessed by a subset of 19 specialists with relevant expertise.

Consensus was achieved on 198 statements (93.8%), a rate that reflects broad agreement across the domains of contraindications, preoperative assessment, surgical technique, intraoperative management, revision surgery, complications, postoperative care, clinical outcomes, guidelines for future practice, and ethics and psychosexual evaluation. The lowest consensus rates were observed in surgical indications and perioperative management, reflecting genuine clinical uncertainty in the absence of high-quality prospective data. Across all domains, the clearest areas of agreement were the primacy of functional outcome preservation over aesthetic modification, the absolute necessity of structured patient assessment and informed consent, and the need for standardized outcome measurement and prospective data collection. These findings provide a clinical reference framework for aesthetic vaginoplasty practice in Turkey while identifying the specific questions that require resolution through future research. The following sections contextualize these consensus findings within the available evidence base and discuss their implications for clinical practice and research.

A fundamental challenge in the vaginoplasty literature is the absence of a validated, objective definition of vaginal laxity. Garcia et al.,[3] in a systematic review specifically calling for

standardized outcome measures in cosmetic gynecology, identified this as the single most important barrier to evidence synthesis across the field. The Alavi-Arjas et al.[8] systematic review of sexual function outcomes after vaginal tightening surgery, the most comprehensive evidence synthesis to date, identified only 11 eligible studies with 806 patients over 18 years and concluded that definitive conclusions on efficacy were not possible due to heterogeneous techniques and unvalidated instruments. A multicenter cross-sectional study found no significant association between self-reported vaginal laxity and objective examination findings, questioning the reliability of subjective complaint as a standalone indicator for surgical candidacy.[12] This dissociation between subjective report and objective findings is not merely a measurement problem; it has direct clinical consequences, as it means that the indication for surgery in a substantial proportion of candidates rests on patient-reported perception alone, without an independently verifiable anatomical or functional correlate. Several emerging technologies hold promise for the development of objective laxity assessment. Tactile imaging devices, which reconstruct the mechanical properties of pelvic floor tissues through pressure-mapping probes, have demonstrated reproducibility in mapping fascial support defects and may provide a quantitative tissue compliance metric applicable to surgical candidacy assessment. Elastography, both ultrasound-based and magnetic resonance-based, offers a non-invasive means of quantifying tissue stiffness and has been applied to pelvic floor musculature with encouraging early results. Dynamic magnetic resonance imaging enables real-time assessment of pelvic floor descent, hiatal dimensions, and compartment-specific prolapse under Valsalva, providing structural information that static examination cannot capture. Three-dimensional ultrasound has been used to characterize levator ani morphology and hiatal area with good inter-rater reliability and may be adaptable to introital and posterior compartment assessment. Biomechanical modelling approaches, using finite element analysis derived from patient-specific imaging data, represent a longer-term research direction that could eventually enable individualized surgical planning based on tissue property prediction. None of these technologies is currently validated for routine clinical application in aesthetic vaginoplasty, and their integration into surgical decision-making will require prospective studies with standardized protocols. In this context, the expert panel's endorsement of the FSFI, PISQ-12, vaginal laxity questionnaire, and POP-Q as minimum assessment instruments provides a practical measurement framework for the near term, while the development of objective biomechanical tools remains the field's most pressing methodological priority.

The management of patient selection is further complicated by the medicolegal and psychosocial dimensions inherent to elective genital cosmetic procedures. Body dysmorphic disorder is a recognized risk factor in cosmetic surgery populations, and its relevance is particularly high in genital aesthetic procedures where the perceived defect is not apparent to others. Hostiuc et al.[13] demonstrated that patients with body dysmorphic disorder systematically fail to benefit from cosmetic surgery because the disorder distorts body perception rather than informing genuine preference and argued that autonomous decision-making is functionally impaired in this setting. The ACOG Committee Opinion similarly identifies body dysmorphic disorder screening as mandatory before any elective female genital cosmetic procedure and states that patient motivations must be autonomous and free from external pressure.[7] Agreement on body dysmorphic disorder as an absolute contraindication (84.6%), psychosexual evaluation in selected patients, and explicit counselling on the limited evidence base prior to surgery (92.3%) is consistent with these professional recommendations. Beyond clinical psychiatric diagnoses, the broader sociocultural context in which demand for aesthetic vaginoplasty arises warrants explicit attention. The exponential growth of social media platforms has created new channels through which idealized and frequently non-representative genital images are disseminated, shaping body image expectations among women who may have had no prior awareness of procedural options. Commercial marketing by aesthetic clinics, often conducted outside regulated advertising frameworks, can amplify perceived inadequacy and frame elective genital modification as a medical necessity rather than a discretionary cosmetic choice. Exposure to pornography-derived aesthetic norms, which bear no relationship to the anatomical diversity of the female pelvis, has been identified in qualitative research as a driver of genital dissatisfaction and is increasingly encountered in preoperative consultations for female genital cosmetic surgery. Partner influence, even when not rising to the level of coercion, can shape patient motivation in ways that compromise autonomous decision-making. The panel's rejection of partner expectation as a surgical indication (19.2%), its prohibition of surgery to address relationship problems (96.2%), and its unanimous requirement for patient-autonomous decision-making collectively establish that this consensus recognizes and guards against these vulnerabilities. Structured preoperative psychosexual assessment, conducted by practitioners with training in sexual medicine or clinical psychology, represents the most reliable available means of distinguishing clinically grounded surgical candidacy from requests driven by external social pressures, unrealistic expectations, or psychological vulnerability, and should be considered a standard component of the preoperative pathway rather than an optional adjunct.

The positions taken by the PETKOZ panel are broadly consistent with, and in some respects extend beyond, those of major international organizations. The RCOG has taken one of the most restrictive stances in this field, advising clinicians not to perform cosmetic vaginal procedures such as vaginal tightening, G-spot amplification, and clitoral hood reduction in response to a perceived demand, and emphasizing that the evidence base for these procedures is poor and that the risks have not been adequately established. The RCOG position explicitly raises concerns about the influence of marketing and media on patient decision-making and calls for robust informed consent that includes an honest discussion of the absence of evidence for benefit.[14] The AUGS has similarly cautioned that elective vaginal procedures are not supported by evidence of safety or efficacy and has called for rigorous

outcome data before widespread adoption. These stances highlight the extent to which the PETKOZ consensus, though developed in a permissive clinical context where demand for these procedures is growing, aligns with international caution: the panel's unanimous endorsement of structured informed consent including disclosure of the limited evidence base, mandatory body dysmorphic disorder screening, and the prohibition of surgery to resolve relationship problems collectively reflects the ethical safeguards that international bodies have identified as minimum requirements.

Where this statement advances the international discussion is in its provision of a procedural classification framework, a structured technique selection algorithm, and a minimum outcome measurement toolkit, none of which have been formally proposed by RCOG, AUGS, or International Society of Aesthetic Plastic Surgery in their current guidance documents. The International Society of Aesthetic Plastic Surgery recognizes vaginal tightening as an established aesthetic procedure and provides general patient-facing information on surgical options but stops short of issuing clinical practice recommendations on technique selection, patient assessment protocols, or outcome standards.[6] The present consensus therefore occupies a distinct position in the international literature; it accepts the legitimacy of these procedures in carefully selected patients while providing the operational framework that existing position statements have not.

Dyspareunia is the complication with the greatest direct impact on surgical outcomes in vaginal tightening surgery, as it specifically reverses the improvement in sexual function that patients seek. Ulubay et al.,[15] in a retrospective study of perineoplasty for the sensation of a wide vagina, reported an overall success rate of 87.9% at 6 months but found that 10% of patients reported introital dyspareunia at follow-up, even in a well-selected cohort treated by experienced surgeons. İnan et al.[16] found that perineoplasty produced significant improvements in sexual desire, arousal, lubrication, orgasm, and satisfaction on the FSFI, but did not yield a significant improvement in the pain domain, suggesting that dyspareunia following vaginoplasty may be resistant to surgical correction once established. Across multiple series in the Alavi-Arjas et al.[8] systematic review, *de novo* dyspareunia rates ranged from 10% to 61.1% depending on technique and patient population. These figures are not merely complication statistics; they represent a form of functional surgical failure. In a procedure performed to improve sexual function, the induction of persistent dyspareunia constitutes a direct negation of the therapeutic objective and must therefore be classified as a primary outcome failure, carrying the same clinical and medicolegal weight as any other major surgical complication. This reframing has practical consequences: it means that dyspareunia prevention cannot be relegated to a postoperative consideration but must be the organizing principle of every intraoperative decision, from suture selection and tissue plane management to the degree of plication and the final diameter confirmation. The panel's strong agreement that excessive muscle plication carries a risk of postoperative dyspareunia, that intraoperative control of tightening degree is required before wound closure, and that a target postoperative vaginal diameter of approximately two examining fingers or 3 cm should be maintained (84.6%) reflects an evidence-based precautionary approach that, in the context of this reframing, acquires the character of a non-negotiable safety standard rather than a technical preference.

Sexual function outcomes following vaginal tightening surgery are broadly positive in the published literature, but with clinically relevant domain-specific variation. Goodman et al.[17] conducted a large multicenter outcome study of female genital plastic surgery and reported high rates of sexual enhancement for both patients and their partners. Among studies using validated instruments, Jamali et al.[18] found significant FSFI improvement in all domains except pain and lubrication, and Fang et al.[19] reported significant improvement in desire, arousal, orgasm, and satisfaction following bilateral wall tightening without mucosal excision.[8] In contrast, Park and Whang[20] found statistically significant improvement only in the satisfaction domain, and Li et al.[21] reported no significant change in total FSFI score despite an isolated orgasm subscore benefit.[8] This variability suggests that sexual function improvement after vaginal tightening is not uniform across domains and is likely influenced by the match between patient expectations and the functional changes the procedure produces, which is consistent with the panel's unanimous position that concordance between preoperative expectations and postoperative outcomes is the primary determinant of patient satisfaction. This domain-specific variability also underscores a broader problem, namely the absence of a standardized outcome assessment framework. The current evidence base draws on a heterogeneous array of instruments, many of which are either unvalidated or applied inconsistently, making meaningful cross-study synthesis impossible. The panel's endorsement of the FSFI, PISQ-12, Vaginal Laxity Questionnaire, and POP-Q as minimum preoperative and postoperative assessment instruments represents an important step toward standardization. However, adoption of these tools must be operationalized through formal inclusion in study protocols, registry systems, and training curricula if the field is to generate guideline-quality evidence. Without this infrastructure, the scientific credibility of aesthetic vaginoplasty research will remain limited regardless of the volume of data produced.

An important and rapidly evolving domain in aesthetic vaginoplasty and non-surgical vaginal procedures is the use of energy-based devices, including fractional CO_2 laser and radiofrequency technologies, which are increasingly used as non-surgical alternatives for vaginal laxity and sexual dysfunction. An international multidisciplinary expert panel reviewing the available literature concluded that no formal clinical guidance exists for the use of these technologies and that the available randomized controlled trial data were insufficient to support definitive recommendations at the time of the review.[22] Although a small number of randomized controlled trials have since been published, most studies in this area remain prospective or retrospective case series without control groups, and comparative data between energy-based and surgical approaches are largely absent. A systematic review evaluating 74 studies of 15 different nonsurgical vulvovaginal

restoration devices found that improvement in symptoms was reported across all included studies. However, adverse events were identified across all device categories, with CO_2 laser devices associated with the highest frequency of reported complications, and the authors concluded that a substantial gap in level I evidence persists.[23] It must be stated clearly that the current evidence base does not support the use of energy-based devices as equivalents to or replacements for surgical correction in women with demonstrable structural vaginal laxity. Their mechanism of action, which relies on thermal stimulation of collagen remodeling, does not address the fascial support defects, perineal body disruption, or mucosal redundancy that typically underlie clinically significant laxity. Until randomized controlled trials with validated laxity measures, standardized patient selection criteria, and minimum 12-month follow-up are available, the clinical positioning of these devices should remain adjunctive. The panel's failure to reach agreement on energy-based devices as a direct surgical alternative is a scientific position grounded in this evidential reality. Notably, a distinct statement endorsing energy-based devices as a selective option in appropriately chosen patients, which was framed as an adjunctive modality rather than a direct surgical replacement, did reach consensus (88.5%), indicating that the panel's disagreement was specifically with designating these technologies as primary surgical equivalents in patients with demonstrable structural laxity, not with their targeted use in selected cases where structural defects are absent or minimal. As comparative data accumulate, a dedicated consensus exercise or evidence-based guideline on non-surgical aesthetic vaginal procedures will be needed.

A persistent structural limitation of the aesthetic vaginoplasty literature is the absence of a validated procedural classification system. The field currently encompasses a heterogeneous spectrum of interventions, including perineoplasty, posterior colporrhaphy-based repairs, levator ani plication techniques, mucosal excision approaches, energy-based adjunctive procedures, and hybrid reconstructive combinations that are frequently described under inconsistent or interchangeable nomenclature, without an agreed framework for categorizing surgical approaches or standardizing technical terminology.[3,4] This heterogeneity directly undermines evidence synthesis and cross-study comparability. As described in the results section, the present consensus addresses this gap by proposing a five-category procedural classification system based on anatomical indication and technical approach, with a corresponding technique selection algorithm. Adoption of this classification and the associated standardized terminology in future clinical research would enable technique-stratified outcome analysis, improve the reproducibility of surgical descriptions, and facilitate the generation of comparative evidence across procedural variants, all of which are essential preconditions for evidence-based guideline development in this field.

Several perioperative management questions produced the lowest consensus rates in the survey, with local anesthesia (30.8%), routine urinary catheter and vaginal pack placement (57.7%), and same-day discharge (50%) all below threshold. These items relate to logistical and institutional practice rather than to surgical principles or patient safety in a direct sense, and the published literature does not address them specifically for elective vaginal tightening surgery. Furnas et al.[24] noted that anesthetic approach, patient preparation, and postoperative monitoring in female genital plastic surgery should be tailored to institutional resources, patient comorbidities, and procedure complexity rather than applied as a fixed protocol. The absence of consensus on these points reflects variation in current practice and highlights the most operationally relevant questions for future prospective data collection.

Looking beyond the immediate horizon of this consensus, several directions will be critical in shaping the next generation of evidence in aesthetic vaginoplasty. The integration of artificial intelligence into surgical planning represents a particularly promising avenue. Machine learning algorithms trained on multimodal preoperative data, including three-dimensional pelvic imaging, validated patient-reported outcome scores, and tissue biomechanical parameters, could in principle generate individualized risk profiles, optimal technique predictions, and postoperative outcome estimates that exceed the accuracy of current expert-based heuristics. Personalized surgical algorithms, grounded in compartment-specific anatomical mapping and patient-specific functional goals rather than generalized technique preferences, are a natural extension of the individualized approach endorsed unanimously by this panel. Objective laxity scoring systems incorporating the biomechanical technologies discussed above would provide a standardized entry criterion that current surgical candidacy assessments lack, enabling patient stratification that is reproducible across centers and investigators. Long-term multicenter registries capturing standardized demographic, anatomical, procedural, and outcome data at minimum follow-up intervals of one, three, and five years are the infrastructure without which technique-stratified analysis, complication incidence estimation, and revision rate benchmarking cannot be performed. Neurofunctional outcome assessment, addressing the preservation of vaginal sensory properties, introital mechanoreceptor integrity, and orgasmic response following different tightening techniques, remains almost entirely uncharted and represents a research priority that directly intersects with the functional preservation philosophy endorsed throughout this consensus.

A further structural limitation of the present consensus, and of consensus methodology in aesthetic vaginoplasty more broadly, is its surgeon-centered perspective. The statements evaluated in this process were developed, adjudicated, and voted upon by surgical specialists; the priorities, concerns, and lived psychosexual experiences of the women who undergo these procedures were not directly represented. This gap matters because the definition of a successful outcome in elective genital surgery is ultimately patient-defined, and what surgeons regard as technically adequate results may diverge considerably from what patients regard as functionally and psychosexually meaningful improvements. Incorporating patient-reported priorities into future consensus development, through structured qualitative interviews, validated preference-elicitation instruments, or formal patient and public involvement panels integrated into the Delphi

methodology, would substantially increase the clinical validity of the recommendations produced. Patient-centered outcome models, in which the minimum clinically important difference is defined by patient preference rather than by statistically significant change in mean values on a (subjective) scale, should be the standard for efficacy assessment in future clinical trials of aesthetic vaginoplasty. The panel's unanimous agreement that concordance between preoperative expectations and postoperative outcomes is the primary determinant of patient satisfaction (100%) is a first step in this direction. Systematic patient involvement in defining what those expectations should reasonably encompass is the logical next step.

This study has several limitations. The expert panel reflects national practice in Turkey, and the recommendations may not be directly transferable to settings with different regulatory frameworks, resource environments, or procedural volumes. As with all Delphi studies, the consensus reflects structured expert opinion rather than data from clinical trials. Furthermore, the 70% agreement threshold identifies convergence of practice, not evidence of efficacy or safety. Patient perspectives were not included in the consensus development, which is an important limitation given that patient experience and reported outcomes should ultimately drive the evolution of these procedures. A further limitation relates to the composition of the expert panel itself: all 26 participating specialists were surgeons, and no representation was included from allied health professionals whose involvement is integral to safe and effective aesthetic vaginoplasty practice. Pelvic floor physiotherapists, whose role in preoperative assessment, postoperative rehabilitation, and the conservative management of dyspareunia is well established, were not part of the consensus process, nor were clinical psychologists, sexual medicine specialists, or gynecological nurses with expertise in perioperative care and patient support. Future iterations of this consensus should adopt a genuinely multidisciplinary panel composition to ensure that the full scope of the clinical pathway is represented, and that recommendations reflect the perspective of all professionals who contribute to patient outcomes rather than the surgical perspective alone.

Clinical Practice Recommendations

The following recommendations are derived from synthesis of consensus findings across 14 clinical domains and are intended to provide practical guidance for clinicians at all levels of experience in aesthetic vaginoplasty.

(1) Aesthetic vaginoplasty must not be offered on demand. Surgery is appropriate only when vaginal laxity is confirmed on pelvic examination, functional or psychosexual impairment attributable to laxity is documented, absolute contraindications are excluded, and the patient's decision is autonomous and free from partner or relational pressure. Surgery must not be recommended to resolve relationship problems or meet partner expectations.

(2) All candidates must undergo a structured preoperative assessment including pelvic examination with POP-Q staging, perineal body assessment, documentation of functional and sexual symptoms, and pelvic floor muscle evaluation. As a minimum, one validated patient-reported outcome instrument (FSFI, PISQ-12, or vaginal laxity questionnaire) must be administered before surgery and at postoperative follow-up. Informed consent must include explicit discussion of the limited evidence base.

(3) Psychosexual evaluation must be performed when the surgical request appears discordant with clinical findings, or where psychological vulnerability, unrealistic expectations, or external pressure are identified as primary drivers. Where the complaint is disproportionate to objective findings or consistent with a body image disturbance, formal psychiatric evaluation must precede any surgical planning. Body dysmorphic disorder is an absolute contraindication to surgery.

(4) Pelvic floor physiotherapy must be offered as first-line treatment before surgical referral when structural support defects are absent. Where postoperative dyspareunia or pelvic floor dysfunction is identified, physiotherapy-based rehabilitation must be initiated promptly. Access to pelvic floor physiotherapy is a standard component of the care pathway, not an optional referral.

(5) Technique selection must be individualized. Perineoplasty is indicated for isolated introital and perineal laxity; posterior colporrhaphy-based repair for mid- and distal-vaginal laxity; levator ani plication selectively for confirmed central support defects only; and hybrid approaches for multi-compartment defects. No single technique is universally appropriate. Selection must be guided by anatomical findings and documented accordingly.

(6) Dyspareunia prevention is the primary intraoperative safety principle. Excessive muscle plication, mucosal closure under tension, and tightening beyond functional optimum must be avoided. Vaginal canal diameter must be confirmed before wound closure, targeting approximately two examining fingers or 3 cm, and vaginal axis must be preserved. Persistent postoperative dyspareunia constitutes a functional surgical failure and must be managed accordingly.

(7) Energy-based devices (fractional CO_2 laser, radiofrequency) are adjuncts to surgical correction and must not be offered as equivalents to surgery in patients with demonstrable structural laxity or fascial defects. Patients must be informed that current evidence does not support these modalities as primary interventions and that formal guidelines do not endorse them in this role.

(8) Postoperative care must follow a standardized protocol: sexual intercourse restricted for at least four to six weeks; follow-up visits at days 10 to 15 and day 40; written instructions covering wound care, hygiene, and activity restrictions; formal outcome assessment with the preoperative validated instrument at three to six months postoperatively.

(9) Revision surgery must not be undertaken without review of the primary procedure, full psychosexual reassessment, and re-confirmation of patient expectations. A minimum healing interval of three to six months must be observed. Patient dissatisfaction without a demonstrable functional or anatomical problem is not a sufficient indication for reoperation.

(10) Aesthetic vaginoplasty should be delivered within a multidisciplinary framework incorporating access to pelvic floor physiotherapy, psychosexual medicine, and clinical psychology. Surgical volume, complication rates, and patient-reported outcomes must be documented systematically. Formal training with structured mentorship before independent practice is an ethical and professional requirement.

CONCLUSION

Aesthetic vaginoplasty is a rapidly growing field of gynecological practice that currently operates without validated clinical guidelines. This consensus statement by PETKOZ establishes the first national expert-based framework governing patient selection, surgical technique, perioperative management, complications, and outcome assessment for these procedures in Turkey. The central message of this consensus is that aesthetic vaginoplasty must be approached as a functional surgical procedure first and a cosmetic one second. The protection of neurovascular structures, the prevention of dyspareunia, the preservation of vaginal axis and diameter, and the alignment of surgical outcomes with patient expectations are non-negotiable standards regardless of the indication. No surgical decision should be made without structured preoperative evaluation. Standardized validated outcome instruments must replace the currently fragmented measurement landscape. This statement does not substitute for high-quality evidence, which the panel unanimously identified as the field's most urgent need. It is intended as a practical guide for clinicians implementing aesthetic vaginoplasty in Turkey until that evidence becomes available, and as a foundation for the standardization of practice that will make generating such evidence possible. The trajectory that this consensus defines, toward evidence-based patient selection, functional outcome preservation, ethical practice, and multidisciplinary care, is not merely a description of current best practice but a statement of direction. Aesthetic vaginoplasty will become a credible and internationally respected surgical discipline only when its practitioners commit to measuring what they achieve with the same rigor they apply to achieving it, when patients are protected by robust screening from procedures they would not benefit from, and when the surgical decision is guided by functional anatomy rather than aesthetic expectation. This consensus is offered as a step toward that standard.

Plans for Updating

This consensus statement is intended to be a living document, subject to formal review and revision as new evidence becomes available. Given the expectation of new data from ongoing prospective studies, national registry initiatives, and emerging comparative trials in aesthetic vaginoplasty, PETKOZ plans to conduct a formal update of this statement within three years of its initial publication date. An interim literature review will be conducted on the first anniversary of publication to identify newly published high-quality studies that may warrant expedited modification of specific recommendations before the scheduled full update. Areas most likely to be revised in the near term include the management of perioperative logistical variables that did not reach consensus in the present process, the role of energy-based adjunctive devices as evidence from randomized controlled trials matures, and the integration of objective laxity assessment tools if validated instruments become available. The updated statement will be subject to the same modified Delphi methodology, expert panel selection criteria, and consensus threshold applied in the present document. Efforts will be made to expand panel composition to include multidisciplinary representation from pelvic floor physiotherapy, sexual medicine, and clinical psychology in future iterations, and to incorporate patient-reported priorities through structured consultation with patient advocacy groups. Clinicians and institutions are encouraged to consult the PETKOZ Association for announcements regarding updated recommendations and supplementary guidance as they become available.

Footnotes

Authorship Contributions

Surgical and Medical Practices: O.D., M.Y., A.E.K., E.Ç., P.K., E.H.C., Concept: O.D., M.Y., Design: O.D., M.Y., Data Collection or Processing: O.D., M.Y., A.E.K., E.Ç., P.K., E.H.C., C.A., E.A., K.B., M.B., P.B.İ., S.B., S.B.K., Y.C., T.D., M.E., S.E., G.G., S.H.K., Ü.K.D., O.K., C.K., Ö.L., A.M., N.P., H.Ş., M.B.Ş., A.S., İ.S., Ö.T., E.U., Analysis or Interpretation: O.D., M.Y., A.E.K., E.Ç., P.K., E.H.C., Literature Search: O.D., M.Y., A.E.K., E.Ç., P.K., E.H.C., Writing: M.Y., P.K.

Conflict of Interest: Six authors of this article, Ozan Doğan, Murat Yassa, Eray Çalışkan, Sevtap Hamdemir Kılıç, Aşkı Ellibeş Kaya, Mehmet Baki Şentürk are members of the Editorial Board of the Anatolian Journal of Obstetrics and Gynecology Research. However, they did not involved in any stage of the editorial decision of the manuscript. The editors who evaluated this manuscript are from different institutions. All other authors declare that they have no conflict of interest.

Financial Disclosure: All those involved in the development of this guideline were unpaid volunteers and receive no direct funding for their work in producing the guideline.

Supplementary Materials 1-4: https://d2v96fxpocvxx.cloudfront.net/f6118de3-f656-440d-95f2-50c620149951/content-images/f497811d-b07d-40f6-8fb0-875b40d39962.pdf

REFERENCES

1. Krychman ML. Vaginal laxity issues, answers and implications for female sexual function. J Sex Med. 2016;13(10):1445-1447.
2. Barbara G, Facchin F, Buggio L, Alberico D, Frattaruolo MP, Kustermann A. Vaginal rejuvenation: current perspectives. Int J Womens Health. 2017;9:513-519.
3. Garcia B, Scheib S, Hallner B, Thompson N, Schiavo J, Peacock L. Cosmetic gynecology-a systematic review and call for standardized outcome measures. Int Urogynecol J. 2020;31(10):1979-1995.
4. Austin RE, Lista F, Vastis PG, Ahmad J. Posterior vaginoplasty with perineoplasty: a Canadian experience with vaginal tightening surgery. Aesthet Surg J Open Forum. 2019;1(4):ojz030.
5. Aesthetic Plastic Surgery National Databank statistics 2020-2021. Aesthet Surg J. 2022;42(Suppl 1):1-18.
6. International Society of Aesthetic Plastic Surgery (ISAPS). Vaginal Tightening: Vaginoplasty and Perineoplasty [Internet]. ISAPS; [cited 11 May 2026]. Available from: https://www.isaps.org/discover/

patients-home/procedures/body-procedures/vaginal-tightening-vaginoplasty-and-perineoplasty/

7. Elective female genital cosmetic surgery: ACOG Committee opinion, number 795. Obstet Gynecol. 2020;135(1):e36-e42.
8. Alavi-Arjas F, Goodman MP, Simbar M, Majd HA, Nahidi F, Rastegar F. The effect of vaginal tightening surgery on sexual function: a systematic review. Int Urogynecol J. 2025;36(1):11-24.
9. Jones J, Hunter D. Consensus methods for medical and health services research. BMJ. 1995;311(7001):376-380.
10. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs. 2000;32(4):1008-1015.
11. Jünger S, Payne SA, Brine J, Radbruch L, Brearley SG. Guidance on Conducting and Reporting Delphi studies (CREDES) in palliative care: recommendations based on a methodological systematic review. Palliat Med. 2017;31(8):684-706.
12. Polland A, Duong V, Furuya R, et al. Description of vaginal laxity and prolapse and correlation with sexual function (DeVeLoPS). Sex Med. 2021;9(6):100443.
13. Hostiuc S, Isailă OM, Rusu MC, Negoi I. Ethical challenges regarding cosmetic surgery in patients with body dysmorphic disorder. Healthcare (Basel). 2022;10(7):1345.
14. Royal College of Obstetricians and Gynaecologists. Ethical considerations in relation to female genital cosmetic surgery (FGCS). RCOG Ethics Committee Opinion Paper. London: RCOG; 2013. Available from: https://qna.files.parliament.uk/qna-attachments/788291/original/HL3083%20B.pdf
15. Ulubay M, Keskin U, Fidan U, et al. Safety, efficiency, and outcomes of perineoplasty: treatment of the sensation of a wide vagina. Biomed Res Int. 2016;2016:2495105.
16. İnan C, Ağır MÇ, Sağır FG, et al. Assessment of the effects of perineoplasty on female sexual function. Balkan Med J. 2015;32(3):260-265.
17. Goodman MP, Placik OJ, Benson RH 3rd, et al. A large multicenter outcome study of female genital plastic surgery. J Sex Med. 2010;7(4 Pt 1):1565-1577.
18. Jamali S, Abedi P, Rasekh A, Mohammadjafari R. The long term effect of elective colpoperineoplasty on sexual function in the reproductive aged women in Iran. Int Sch Res Notices. 2014;2014:912786.
19. Fang QQ, Yao JM, Xue YN, et al. Management of vaginal laxity through bilateral wall tightening without mucosal excision. J Plast Reconstr Aesthet Surg. 2023;80:28-35.
20. Park TH, Whang KW. Vaginal rejuvenation with gore-mycromesh. Aesthetic Plast Surg. 2015;39(4):491-494.
21. Li Y, Xia Z, Bai M, et al. New method for genital aesthetic surgery: an easy-to-learn 2-step approach with acellular dermal matrix. Aesthet Surg J. 2022;42(9):1045-1052.
22. Shobeiri SA, Kerkhof MH, Minassian VA, Bazi T; IUGA Research and Development Committee. IUGA Committee opinion: laserbased vaginal devices for treatment of stress urinary incontinence, genitourinary syndrome of menopause, and vaginal laxity. Int Urogynecol J. 2019;30(3):371-376.
23. Preminger BA, Kurtzman JS, Dayan E. A systematic review of nonsurgical vulvovaginal restoration devices: an evidence-based examination of safety and efficacy. Plast Reconstr Surg. 2020;146(5):552e-564e.
24. Furnas HJ, Canales FL, Pedreira RA, Comer C, Lin SJ, Banwell PE. The safe practice of female genital plastic surgery. Plast Reconstr Surg Glob Open. 2021;9(7):e3660.

Original Article

DOI: 10.4274/anajog.galenos.2026.77486

Anat J Obstet Gynecol Res 2026;3(2):79-85

Utilization Patterns and Clinical Outcomes Following Oocyte Cryopreservation in Turkey: A Multicenter Cohort Study Within a Nationally Regulated Setting

Gerçek Aydın[1,2], Aziz İhsan Tavuz[3], Hayal Uzelli Şimşek[4], Serap Simavlı[5], Emek Doğer[4]

[1]Eurofertil IVF Center, Bursa, Turkey
[2]İstanbul Health and Technology University Faculty of Medicine, Department of Obstetrics and Gynecology, İstanbul, Turkey
[3]Momart IVF Center, Clinic of Obstetrics and Gynaecology, İstanbul, Turkey
[4]Kocaeli University Faculty of Medicine, Department of Obstetrics and Gynecology, Kocaeli, Turkey
[5]Üskudar University Faculty of Medicine, Department of Obstetrics and Gynecology, İstanbul, Turkey

ABSTRACT

Purpose: To analyse the utilization patterns, patient characteristics, and reproductive outcomes of oocyte cryopreservation (OC) performed under the current fertility preservation (FP) framework in Turkey.

Methods: This retrospective multicenter cohort study analyzed data from Ministry of Health-accredited FP programs across Turkey between 2005 and 2026. The study included women who underwent OC for approved medical indications, predominantly diminished ovarian reserve and FP before gonadotoxic treatment. Primary outcomes were oocyte utilization rates and reproductive outcomes among women who returned to use their cryopreserved oocytes. Secondary outcomes included socio-demographic profiling and cycle-level characteristics.

Results: A total of 1,582 women were included. Mean age at cryopreservation was 33.5±7.2 years. Most women underwent OC for diminished ovarian reserve (92.3%), while 7.6% underwent FP before treatment for malignant disease. Most participants were university-educated (63.9%) and employed (65.7%), with obstetrics and gynecology specialists being the primary source of awareness (53.3%). Mean anti-Müllerian hormone was 0.6±0.6 ng/mL, and a mean of 3.6±3.0 mature oocytes were cryopreserved per cycle. During the observation period, only 39 women (2.4%) returned to utilize their cryopreserved oocytes. Among thaw cycles, the clinical pregnancy rate per transfer was 15.6%, and the implantation rate was 12.7%.

Conclusion: In this large multicenter cohort, OC was predominantly performed in women with diminished ovarian reserve, resulting in relatively limited oocyte yield and low observed utilization during the available follow-up period. These findings provide real-world data on the characteristics, utilization patterns, and reproductive outcomes of FP programs in Turkey and may help inform future clinical practice and policy discussions.

Keywords: Fertility preservation, oocyte cryopreservation, restrictions, utilization rate

INTRODUCTION

In today's world, the demands of professional life often lead women to consider motherhood at later stages, when reproductive potential may already be in decline. Advancing age is inherently associated with diminished oocyte quality and quantity,[1] limiting the chances of successful conception compared to younger women. In this context, preserving fertility emerges as a critical consideration in life planning for many individuals. The term "fertility preservation" (FP) relates to offering women and men the opportunity to extend their reproductive potential beyond limits imposed or expected because of medical or non-medical reasons. Due to enhanced access to information and growing patient awareness, FP has become a pivotal strategy in modern reproductive medicine.

One of the most significant milestones in this field was the introduction of vitrification techniques in in virto fertilization

Address for Correspondence: Lec, Gerçek Aydın, MD, İstanbul Health and Technology University Faculty of Medicine, Department of Obstetrics and Gynecology, İstanbul, Turkey
E-mail: gercekay@yahoo.com **ORCID ID:** orcid.org/0000-0002-5638-3675
Received: 18.06.2026 **Accepted:** 01.08.2026 **Publication Date:** 02.09.2026

Cite this article as: Aydın G, Tavuz Aİ, Şimşek HU, Simavlı S, Doğer E. Utilization patterns and clinical outcomes following oocyte cryopreservation in Turkey: a multicenter cohort study within a nationally regulated setting. Anat J Obstet Gynecol Res. 2026;3(2):79-85

(IVF) laboratories. Since the first live birth following human oocyte vitrification was reported in 1999,[2] the clinical use of this method has expanded considerably. The growing body of experience, coupled with evidence showing no increase in congenital anomaly risk in subsequent pregnancies, has led to a rise in the number of autologous oocyte cryopreservation (OC) cycles performed over the past decades.[3,4] This growing confidence has also contributed to the development of donor programs and banking strategies, further enhancing the clinical utility of oocyte vitrification.[5]

Autologous OC is currently performed for both medical and non-medical indications. Medical indications primarily include FP before potentially gonadotoxic treatments, whereas non-medical (elective) OC is generally pursued because of anticipated age-related fertility decline, reproductive life planning, or the absence of a suitable reproductive partner. In both settings, age at the time of oocyte retrieval remains one of the most important determinants of future reproductive success.[6] However, access to this technology remains uneven globally, often shaped by national legislation and cultural norms.[6-14] In Turkey, OC is regulated by the ministry of health and is currently permitted under specific medical indications. These include FP before potentially gonadotoxic treatments, as well as selected circumstances related to ovarian reserve assessment and reproductive risk. Elective OC for non-medical reasons is not currently permitted within the existing regulatory framework. The regulatory framework governing OC differs substantially across countries and may influence patient access, referral pathways, and utilization patterns. Understanding how FP programs function within different regulatory environments is therefore important for clinicians, patients, and healthcare policymakers.

Although OC has become increasingly established worldwide, relatively limited data are available regarding utilization patterns and long-term outcomes of FP programs operating under different regulatory frameworks. Thus, the aim of this multicenter cohort study was to describe patient characteristics, utilization patterns, and reproductive outcomes following OC in Turkey and to provide real-world data from a large multicenter cohort operating within a nationally regulated FP framework.

METHODS

Study Design and Setting

This retrospective multicenter cohort study was conducted at six FP programs licensed by the Turkish Ministry of Health for assisted reproductive technologies, located across different regions of Turkey. Participating centers included Kocaeli University IVF Center, Bursa Eurofertil IVF Center, Istanbul Momart IVF Center, Yüzyıl Hospital IVF Center, Bursa Pembe Mavi IVF Center, and collaborating private FP clinics. Variation in patient volume across centres reflects differences in programme size and duration rather than methodological inconsistency. Of note, the study does not compare outcomes between centres so this heterogeneity does not introduce analytical bias. Centres were recruited on the basis of voluntary participation, encompassing university-affiliated, public hospital-based, and private FP programmes across different regions of Turkey. Clinical data encompassing OC cycles performed between 2005 and 2026 were retrospectively collected from each center in anonymized form.

The primary aim of this study was descriptive: to characterise the patient profile, indications, stimulation outcomes, and utilization patterns of OC within a nationally regulated setting. The limited number of reproductive events observed during the follow-up period precluded multivariable modelling of predictors of reproductive outcomes. This represents an important direction for future prospective research in Turkey.

Study Population

Eligible participants were women who underwent OC at one of the participating centers during the study period and for whom complete clinical and laboratory records were available. Women were excluded if the cryopreservation cycle was cancelled before oocyte retrieval, if embryo cryopreservation rather than OC had been performed, or if essential clinical information was unavailable.

Data Collection

Standardized case report forms were used to extract demographic, clinical, and laboratory parameters from the electronic medical records of each center. Collected variables included age at first oocyte retrieval, body mass index (BMI), tobacco use, obstetric history, educational and occupational status, and indication for OC. Oncological status and treatment details were recorded where applicable. Ovarian stimulation data included the type and total dose of gonadotropins administered, stimulation protocol, trigger method for final oocyte maturation, and the use of adjuvant oral agents. Hormonal and ultrasound parameters included baseline serum follicle stimulating hormone (FSH), luteinizing hormone (LH), estradiol (E2), anti-Müllerian hormone (AMH), and antral follicle count (AFC). However, as serum FSH, LH, and E2 values were not uniformly documented across all participating centres and time periods and there would have been methodological and potentially normative range differences, these were not included in the primary results due to an insufficient completeness rate for meaningful summary reporting.

Oocyte retrieval outcomes comprised the total number of oocytes retrieved, the number of germinal vesicle and degenerated oocytes, the number of mature (MII) oocytes, vitrification method, storage location, and total duration of cryopreservation.

Outcomes

The primary outcome was the oocyte utilization rate, defined as the proportion of women who returned to use their cryopreserved oocytes during the observation period. Secondary outcomes included reproductive outcomes following thaw cycles (clinical pregnancy rate per transfer and implantation rate), demographic and socioeconomic characteristics, indications for OC, ovarian stimulation parameters, and oocyte yield per retrieval cycle.

Ethics Approval

This study was approved by the Kocaeli University Faculty of Medicine Non-Interventional Clinical Research Ethics Committee (approval no: KÜ GOKAEK-2025/17/18, date: 19.08.2025). The study was conducted in accordance with the ethical principles of the Declaration of Helsinki. Due to the retrospective nature of the study, the requirement for written informed consent was waived in accordance with national ethical guidelines. For patients younger than 18 years of age, all procedures were performed in accordance with national regulations governing FP in minors.

Statistical Analysis

Descriptive statistics were reported as means ± standard deviation (SD) or medians with minimum-maximum (min-max) values for continuous variables, and as frequencies with percentages for categorical variables. Between-group comparisons were performed using Student's t-test or the Mann-Whitney U test for continuous variables, and the chi-square or Fisher's exact test for categorical variables, as appropriate. Statistical significance was set at $p<0.05$. All analyses were performed using SPSS version 29.0 (IBM Corp., Armonk, NY, USA).

As this study enrolled all eligible patients from the participating centres during the study period no sample size calculation or power analysis was performed.

RESULTS

A total of 1,582 women underwent OC during the study period. The number of patients contributing data per centre was: Kocaeli University IVF Centre (n=710), Bursa Eurofertil IVF Centre (n=75), Istanbul Momart IVF Centre (n=84), Yüzyıl Hospital IVF Centre (n=75), Bursa Pembe Mavi IVF Centre (n=37), and four collaborating consultant centres (n=601). The mean age was 33.5±7.2 years (range: 14-50), and the mean BMI was 22.3±4.8 kg/m^2. Most cases (92.3%, n=1,461) were indicated for diminished ovarian reserve, while 7.6% (n=121) underwent FP following a diagnosis of malignant disease. Only seven women (0.4%) had a living child at the time of cryopreservation. Sixteen patients were aged 45 years or older, while 11 (0.7%) patients were under 18 years of age (range: 14-17 years), all of whom had been diagnosed with hematological or solid malignancies including leukemia, lymphoma, or sarcoma. Demographic characteristics and patient referral patterns are summarized in Table 1.

The majority of patients were university graduates (63.9%) and were employed (65.7%). Awareness of OC was most commonly attributed to an obstetrics and gynecology (ObGyn) specialist (53.3%), followed by media and internet sources (20.1%) and family or friends (18.2%). The remaining patients were referred by physicians other than ObGyn (8.2%), including oncologists, family practitioners, radiologists, general surgeons, pediatricians, and geneticists (Table 1).

Regarding ovarian stimulation, the antagonist protocol was used in 94.8% of cycles, while progesterone-primed and long agonist protocols were employed in 4.2% and 0.8% of cases, respectively. Random start cycles accounted for 2.7% of all stimulations. Letrozole and clomiphene citrate were co-administered with gonadotropins in 9.0% and 3.2% of cycles, respectively.

Oocyte retrieval was performed by transvaginal ultrasound guidance in 91.0% of cases, with transabdominal and laparoscopic approaches used in 4.8% and 4.1%, respectively. Mean AMH was 0.6±0.6 ng/mL and mean AFC was 5.5±3, consistent with the predominantly diminished ovarian reserve population. Mean stimulation duration was 9.9±1.4 days with a total gonadotropin dose of 3,252±992 IU. A mean of 4.5±4.0 oocytes were retrieved per cycle, of which 3.7±3.2 were mature (MII) and 3.6±3.0 were successfully cryopreserved. The ovarian hyperstimulation syndrome rate was 0.5%. Cycle characteristics, stimulation protocols, and oocyte yield are detailed in Table 2.

The mean duration of oocyte storage was 34.7±19.3 months (range: 1-101 months). A total of 137 women (8.6%) obtained ministry of health approval to extend storage beyond the legally mandated five-year limit. Among the 51 women whose cryopreservation was discontinued prior to utilization, reasons included cancer-related death (n=11), spontaneous conception (n=14), cancer recurrence or persistence precluding further treatment (n=12), and reaching the legal storage limit without wishing to extend (n=14).

Table 1. Demographic variables and patient awareness in oocyte cryopreservation patients

Variable	Number of patients (n=1582)
Age, years	33.5±7.2 (14-50)
BMI, kg/m^2	22.3±4.8 (13-47.9)
Tobacco use	311 (19.6%)
Children alive	7 (0.4%)
Education	
Secondary school	249 (15.7%)
High school	322 (20.3%)
University	1011 (63.9%)
Working status	
Employed	1040 (65.7%)
Unemployed	542 (34.2%)
Patient referral and awareness	
Obstetrics and gynecology specialist	844 (53.3%)
*Other doctors	131 (8.2%)
Family or friends	289 (18.2%)
Media-internet	318 (20.1%)
Indication for oocyte cryopreservation	
Low ovarian reserve	1461 (92.3%)
Fertility preservation after diagnosis of malignant tumors	121 (7.6%)

*Other doctors include oncologist, family practitioner, radiologist, general surgeon, pediatrician, geneticists
The data is presented as numbers and percentages or mean ± standard deviation of the mean and minimum-maximum values
BMI: Body mass index

During the study period, 39 women (2.4%) returned to utilize their cryopreserved oocytes. Of these, 27 (69.2%) had originally undergone OC because of diminished ovarian reserve and 12 (30.8%) for FP before cancer treatment. A mean of 3.3±2.8 mature oocytes were thawed per cycle. The wide SD observed for the fertilisation rate (84±42%) reflects genuine inter-cycle variability inherent to the small thaw cohort (n=39), in which cycles with a single or two oocytes thawed disproportionately influence summary statistics; no data-entry errors were identified upon re-verification of the source data.

Embryo transfer was performed in 32 of 39 thaw cycles (82.0%). A total of six pregnancies were achieved, including one biochemical pregnancy and five clinical pregnancies. The overall pregnancy rate per transfer was 18.7% (6/32), while the clinical pregnancy rate per transfer was 15.6% (5/32). The implantation rate was 12.7% (6/47 transferred embryos). Thaw cycle outcomes and reproductive results are presented in Table 3.

DISCUSSION

In this large multicenter cohort study, we evaluated utilization patterns and reproductive outcomes following OC in the setting of nationally regulated FP. Three principal findings emerged from our analysis. First, more than 90% of women undergoing OC had diminished ovarian reserve rather than oncological indications. Second, the cohort demonstrated a relatively low oocyte yield, consistent with the biological characteristics of the study population. Third, only a small proportion of women returned to use their cryopreserved oocytes during the statutory 5-year follow-up period. Collectively, these findings describe OC practice within a nationally regulated setting and highlight the importance of patient selection when interpreting utilization and reproductive outcomes.

While numerous studies have investigated FP outcomes, many have been limited by small sample sizes and short follow-up durations. However, recent studies with longer follow-up periods are beginning to provide more meaningful insights into clinical outcomes.[15,16] Against this background, our multicenter findings may contribute valuable insight into the current status of FP in Turkey and allow for comparison with international cohorts. Previous studies have reported substantial variability in utilization rates following OC, reflecting differences in patient populations, indications, follow-up duration, healthcare systems, and regulatory environments.[15-17] Consequently, utilization rates should be interpreted within the context of the characteristics of the population being studied. In our cohort, only 39 women (2.4%) returned to use their cryopreserved oocytes during the observation period. However, utilization rates require careful contextualisation, as there is no universally accepted follow-up duration for defining non-utilization after OC. Some women may return within a short period after storage, whereas others may not return for many years. Therefore, the utilization rate reported in the present study reflects observed utilization during the available follow-up period rather than the final cumulative utilization rate of the cohort. The mean storage duration of 34.7 months, with some patients exceeding eight years of storage, nevertheless provides valuable insight into real-world utilization patterns within a large multicenter cohort.

Table 2. Cycle characteristics, protocols and oocyte yield of oocyte cryopreservation cases

	Participants (n=1582)
Induction protocol	
Long agonist	14 (0.8%)
Antagonist cycle	1501 (94.8%)
Progesterone primed	67 (4.2%)
Random start	43 (2.7%)
Oral agents used with gonadotropins	
Clomiphene	52 (3.2%)
Letrozole	143 (9%)
Oocyte retrieval method	
Transvaginal ultrasound guided	1440 (91%)
Transabdominal ultrasound guided	77 (4.8%)
Laparoscopic	65 (4.1%)
AMH (ng/mL)	0.6±0.6 (0-5)
AFC	5.5±3 (1-19)
Stimulation duration (days)	9.9±1.4 (5-20)
Total gonadotropin dose (IU)	3252±992 (600-9000)
Retrieved total oocytes	4.5±4 (1-49)
Mature (metaphase II) oocytes (MII)	3.7±3.2 (1-26)
Cryopreserved oocytes	3.6±3.0 (1-26)
Ovarian hyperstimulation syndrome	9 (0.5)

The data is presented as numbers and percentages or mean ± standard deviation of the mean and minimum-maximum values
AMH: Anti-Müllerian hormone, AFC: Antral follicle count

Table 3. Cycles for reuse of oocyte cryopreservation cases

Variable	**Thaw cycles of the ocyte crypreserved cases (n=39)**
Number of M^2 oocytes thawed	3.3±2.8
*Fertilization rate (%)	84±42%
Embryo Transfer	32/39 (82%)
Day 3	23/32 (71.8%)
Day 5	9/32 (28.1%)
Embryo transfer Single	17/32 (53.1%)
Double	15/32 (46.8%)
Total number of embryos transferred (n)	47
Pregnancy rate per transfer (%)	6/32 (18.7%)
Biochemical pregnancy per transfer (%)	1/32 (3.1%)
Clinical pregnancy rate per transfer (%)	5/32 (15.6%)
Implantation rate (%)	6/47 (12.7%)

A total of 39 (2.4%) cases returned for the use of oocyte cryopreserved oocytes. Twelve of these were cryopreserved for oncological treatments and 27 of them were cryopreserved for low ovarian reserve
*Fertilization rate was calculated on a per-cycle basis

Some of these women may have conceived spontaneously and therefore no longer needed their frozen oocytes. Others may have experienced delays in marriage or encountered ambivalence, either personally or from their partners, about pursuing parenthood later in life. Understanding the reasons beyond non-utilization from a sociodemographic perspective could offer valuable guidance for improving program design and implementation.

The demographic profile of women enrolled in our program also warrants discussion. The predominance of university graduates (63.9%) and employed women (65.7%) among applicants suggests that awareness of and access to OC remains concentrated among more educated and economically independent segments of society. The predominance of university graduates and employed women may suggest disparities in awareness, access, or utilization of FP services across different socio-economic groups.

The striking imbalance in indications, with 92.3% of cases driven by diminished ovarian reserve and only 7.6% by oncological reasons, alongside the finding that non-ObGyn physicians accounted for fewer referrals than media or personal networks, suggests a systemic gap in interdisciplinary awareness. Oncologists, surgeons, and other specialists who routinely encounter women of reproductive age facing potentially gonadotoxic treatment represent an underutilised referral source, and targeted awareness initiatives could meaningfully address this deficit.

Maternal age at the time of oocyte retrieval remains one of the most important determinants of future reproductive success following OC. Previous studies have consistently demonstrated superior reproductive outcomes when oocytes are cryopreserved at younger ages.[17-19] In the present cohort, the clinical pregnancy rate per transfer was 15.6% and the implantation rate was 12.7%. These outcomes should be interpreted in the context of the study population, which was characterized by low AMH levels, low AFC values, and a limited number of cryopreserved oocytes per cycle. Collectively, these findings suggest that baseline ovarian characteristics and oocyte yield are important considerations when evaluating reproductive outcomes after OC.

International approaches to FP reveal a stark contrast between many western nations and countries with stricter regulations. In the United States, Spain, Canada, Belgium, Brazil, Australia, and Israel, elective egg freezing is liberally permitted for all adult women regardless of marital status or age, reflecting a progressive stance on reproductive autonomy.[7-10,20] The United Kingdom also has permissive policies, allowing access without age or relationship restrictions.[11] France even offers partial public funding,[7] while countries like Austria and China maintain significant restrictions, limiting access to only medical indications or married women.[12]

In Muslim-majority countries, policies on elective OC vary considerably, shaped by cultural, legal, and religious frameworks. While some nations have recently moved toward more permissive regulations, others continue to impose strict limitations. Among Islamic scholars, the procedure is broadly regarded as makruh-permissible but discouraged-which continues to influence both policy and individual decision-making across these settings.[13,14]

The findings of this study should be interpreted in the context of the patient population studied. Unlike many published elective OC cohorts, 92% of women in the present study presented with diminished ovarian reserve, which is likely to have influenced oocyte yield, utilisation patterns, and reproductive outcomes, independent of healthcare system and regulatory factors.[7-14,20] Furthermore, the relatively low AMH levels, low AFC values, and limited number of cryopreserved oocytes observed in our cohort are consistent with the characteristics of women presenting with diminished ovarian reserve. Therefore, direct comparisons with elective OC cohorts from unrestricted settings are of limited validity, given the substantial differences in indication, ovarian reserve status, and reproductive prognosis between these populations.[6,17]

Future studies focusing on referral pathways, access patterns, patient decision-making, and long-term utilization outcomes may further clarify how different healthcare systems and regulatory environments influence FP practice and outcomes across diverse patient populations.

Study Limitations

This study has several limitations that should be considered when interpreting the findings. First, the retrospective design may be associated with selection bias and incomplete data capture despite the use of standardized data collection procedures across participating centers. Second, the study period spans 21 years (2005-2026), during which substantial advances occurred in vitrification technology, ovarian stimulation protocols, laboratory practices, and regulatory policy in Turkey. These changes may have influenced oocyte yield, cryopreservation efficiency, and utilisation patterns in a non-uniform manner across the cohort. However, as year-of-procedure data were not uniformly available across all participating centres, temporal subgroup analyses could not be performed, and the overall results should therefore be interpreted as reflecting aggregate practice across the entire study period rather than current clinical standards. Third, utilization rates should be interpreted as observed utilization during the available follow-up period rather than ultimate lifetime utilization rates. Women may return to use their cryopreserved oocytes after highly variable storage durations, and no universally accepted follow-up threshold exists for defining non-utilization. Fourth, only a limited number of women returned to utilize their cryopreserved oocytes, resulting in a relatively small number of thaw cycles and reproductive events. Consequently, reproductive outcome estimates should be interpreted cautiously. Detailed post-warming oocyte survival data were not uniformly available across participating centers and therefore could not be analyzed.

Although two indication groups were identifiable, a formal subgroup comparison was not performed, as the oncological group was numerically small (n=121), clinically heterogeneous across centres, and the limited reproductive events precluded meaningful statistical comparison. Finally, complete follow-up data on live births, pregnancy loss, ectopic pregnancy,

and obstetric outcomes were not uniformly available across participating centres. This represents a significant limitation, as live birth rate per transfer, rather than clinical pregnancy rate, is the most clinically meaningful endpoint in assisted reproduction. Future studies with prospective follow-up and standardized outcome reporting are strongly encouraged to address this gap.

The present study also has several important strengths. To the best of our knowledge, it represents one of the largest multicenter cohorts evaluating OC practice in Turkey. The inclusion of 1,582 women from multiple FP programs provides a comprehensive overview of patient characteristics, indications, stimulation outcomes, and observed utilization patterns within a real-world clinical setting. In addition, the study includes both women undergoing FP before potentially gonadotoxic treatment and women undergoing OC because of diminished ovarian reserve, thereby reflecting the spectrum of patients currently accessing these services in routine clinical practice in Turkey. Finally, the availability of follow-up data extending beyond eight years in some patients provides valuable insight into long-term utilization behavior following OC.

CONCLUSION

In this large multicenter cohort, OCin Turkey was performed predominantly in women with diminished ovarian reserve, resulting in relatively limited oocyte yield and low observed utilization during the available follow-up period. The reproductive outcomes observed among women who returned to use their cryopreserved oocytes should be interpreted in the context of the underlying ovarian reserve characteristics of the study population and the limited number of thaw cycles.

This large multicenter cohort provides insight into patient characteristics, utilisation patterns, and reproductive outcomes following OC in a regulated FP setting. The predominance of women presenting with diminished ovarian reserve highlights the potential importance of timely FP counseling, improved public awareness, and effective interdisciplinary referral pathways when designing future FP strategies. Future prospective studies in Turkey with longer follow-up and complete live birth reporting are needed to better define long-term utilization behavior and the ultimate effectiveness of this nationally regulated FP program.

Ethics

Ethics Committee Approval: This study was approved by the Kocaeli University Faculty of Medicine Non-Interventional Clinical Research Ethics Committee (approval no: KÜ GOKAEK-2025/17/18, date: 19.08.2025).

Informed Consent: Not applicable. The manuscript does not contain any individual person's data in any form (including individual details, images, or videos). Due to the retrospective nature of the study and the use of anonymized data, the requirement for informed consent to participate was waived by the Non-Interventional Clinical Research Ethics Committee of Kocaeli University Faculty of Medicine.

Authorship Contributions

Surgical and Medical Practices: G.A., A.İ.T., H.U.Ş., S.S., E.D., Concept: G.A., Design: G.A., Data Collection or Processing: G.A., A.İ.T., H.U.Ş., S.S., E.D., Analysis or Interpretation: G.A., Literature Search: G.A., Writing: G.A.

Footnotes

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

Declaration on the Use of Artificial Intelligence (AI): Artificial intelligence has been used to assist with "language correction".

REFERENCES

1. Zhu Z, Xu W, Liu L. Ovarian aging: mechanisms and intervention strategies. Med Rev (2021). 2022;2(6):590-610.
2. Kuleshova L, Gianaroli L, Magli C, Ferraretti A, Trounson A. Birth following vitrification of a small number of human oocytes: case report. Hum Reprod. 1999;14(12):3077-3079.
3. Johnston M, Richings NM, Leung A, Sakkas D, Catt S. A major increase in oocyte cryopreservation cycles in the USA, Australia and New Zealand since 2010 is highlighted by younger women but a need for standardized data collection. Hum Reprod. 2021;36(3):624-635.
4. Noyes N, Porcu E, Borini A. Over 900 oocyte cryopreservation babies born with no apparent increase in congenital anomalies. Reprod Biomed Online. 2009;18(6):769-776.
5. Cobo A, Garrido N, Pellicer A, Remohí J. Six years' experience in ovum donation using vitrified oocytes: report of cumulative outcomes, impact of storage time, and development of a predictive model for oocyte survival rate. Fertil Steril. 2015;104(6):1426-1434.e1-e8.
6. Cobo A, García-Velasco JA, Remohí J, Pellicer A. Oocyte vitrification for fertility preservation for both medical and nonmedical reasons. Fertil Steril. 2021;115(5):1091-1101.
7. Chehimi M, Boursier A, Lefebvre-Khalil V, et al. Sociodemographic profile of women seeking elective egg freezing when financed by the public healthcare system. Reprod Biomed Online. 2025;50(5):104705.
8. Varlas VN, Bors RG, Albu D, et al. Social freezing: pressing pause on fertility. Int J Environ Res Public Health. 2021;18(15):8088.
9. Wafi A, Nekkebroeck J, Blockeel C, De Munck N, Tournaye H, De Vos M. A follow-up survey on the reproductive intentions and experiences of women undergoing planned oocyte cryopreservation. Reprod Biomed Online. 2020;40(2):207-214.
10. Baldwin K. Conceptualising women's motivations for social egg freezing and experience of reproductive delay. Sociol Health Illn. 2018;40(5):859-873.
11. Gürtin ZB, Tiemann E. The marketing of elective egg freezing: a content, cost and quality analysis of UK fertility clinic websites. Reprod Biomed Soc Online 2021;12:56-68.
12. Zhao Y, Fu H. Social egg freezing for single women in China: legal and ethical controversies. Risk Manag Healthc Policy. 2023;16:2379-2389.
13. Chin AHB, Saifuddeen SM. Is social egg freezing (oocyte cryopreservation) for single women permissible in Islam? A perspective from Singapore. The New Bioethics. 2022;28(2):116-126.

14. Chin AHB, Muhsin SM, Ahmad MF. Islamic perspectives on elective ovarian tissue freezing by single women for non-medical or social reasons. ABR. 2023;15(3):335-349.
15. Blakemore JK, Grifo JA, DeVore SM, Hodes-Wertz B, Berkeley AS. Planned oocyte cryopreservation—10-15-year follow-up: return rates and cycle outcomes. Fertil Steril. 2021;115(6):1511-1520.
16. Leung AQ, Baker K, Vaughan D, et al. Clinical outcomes and utilization from over a decade of planned oocyte cryopreservation. Reprod Biomed Online. 2021;43(4):671-679.
17. Hirsch A, Hirsh Raccah B, Rotem R, Hyman JH, Ben-Ami I, Tsafrir A. Planned oocyte cryopreservation: a systematic review and meta-regression analysis. Hum Reprod Update. 2024;30(5):558-568. Erratum in: Hum Reprod Update. 2024;30(5):651.
18. Homburg R, van der Veen F, Silber SJ. Oocyte vitrification--women's emancipation set in stone. Fertil Steril. 2009;91(Suppl4):1319-1320.
19. Williams RS, Ellis DD, Wilkinson EA, Kramer JM, Datta S, Guzick DS. Factors affecting live birth rates in donor oocytes from commercial egg banks vs. program egg donors: an analysis of 40,485 cycles from the Society for Assisted Reproductive Technology registry in 2016-2018. Fertil Steril. 2022;117(2):339-348.
20. Hammarberg K, Kirkman M, Pritchard N, et al. Reproductive experiences of women who cryopreserved oocytes for non-medical reasons. Human Reproduction. 2017;33(3):575-581.

Original Article

DOI: 10.4274/anajog.galenos.2026.40085

Anat J Obstet Gynecol Res 2026;3(2):86-97

Genital Self-Image, Genital Shame, and Core Sexual Complaints in Women: A Cross-Sectional Online Survey from Turkey

Aşkı Ellibeş Kaya[1], Ebru Zülfikaroğlu[2]

[1]İstanbul Aydın University Faculty of Medicine, Department of Obstetrics and Gynecology, İstanbul, Turkey
[2]Eva Women Health Clinic, Clinic of Obstetrics and Gynecology, Ankara, Turkey

ABSTRACT

Purpose: To evaluate the relationship between selected genital self-appraisal indicators and core sexual complaints in women from Turkey, with particular emphasis on genital satisfaction, genital shame, and comfort with a partner seeing the genital area.

Methods: This cross-sectional online survey represents a focused secondary analysis of an ethically approved parent dataset collected in Turkey. The analytic subsample comprised 330 women and reflected a digitally reached, predominantly highly educated, non-probability sample. Three selected single-item genital self-appraisal indicators were examined rather than the full female genital self-image scale total score. Core sexual complaints included low sexual desire, arousal problems, and orgasm-related complaints. Distressing complaints were defined using a revised nested definition requiring both symptom presence and at least moderate problem severity. Bivariate analyses and multivariable logistic regression models were performed. Sensitivity analyses used the original 1-5 ordinal Likert scores.

Results: Among women with valid genital self-appraisal data, 53.4% reported high genital satisfaction, 63.4% reported high partner genital comfort, and 23.0% met the threshold for genital shame. Overall, 37.3% had at least one core sexual complaint and 29.2% had at least one distressing core sexual complaint using the revised nested definition. In bivariate analyses, lower genital satisfaction score, genital shame, and lower partner genital comfort score were associated with greater complaint burden. In the adjusted model for distressing complaints, high genital satisfaction and genital shame showed directionally consistent but non-significant associations. Genital shame remained independently associated with orgasm-related complaints (odds ratio 2.12, 95% confidence interval 1.14-3.94, $p=0.018$). Sensitivity analyses using ordinal scores supported the overall direction of the findings.

Conclusion: Selected genital self-appraisal indicators, particularly genital shame, were associated with sexual complaint patterns in this sample. The findings should be interpreted cautiously because of the cross-sectional design, selected single-item measurement approach, and non-probability online sampling.

Keywords: Body image, genital self-image, sexual dysfunctions, psychological, sexual health, women, Turkey

INTRODUCTION

Genital self-image is an increasingly recognized component of women's sexual health, yet it remains relatively underexplored in routine gynecologic and sexual medicine research.[1,2] Beyond appearance-related concerns alone, genital self-image reflects how a woman perceives, evaluates, and emotionally relates to her genital area. In clinical terms, this construct may influence sexual confidence, comfort with intimacy, body-related anxiety, and willingness to engage in sexual activity. Negative genital self-perceptions, including shame and embarrassment, may therefore have implications that extend beyond body image and into sexual well-being itself.[1-3] Early studies by Berman and colleagues identified genital self-image as an important component of female sexual health and demonstrated associations with sexual function, sexual satisfaction, and quality of life.[1,4] In recent years, attention has shifted toward understanding genital self-image not merely as a cosmetic or psychological issue, but as a dimension that may intersect with sexual functioning. Existing evidence suggests

Address for Correspondence: Prof. Dr. Aşkı Ellibeş Kaya, İstanbul Aydın University Faculty of Medicine, Department of Obstetrics and Gynecology, İstanbul, Turkey
E-mail: askiellibes@hotmail.com **ORCID ID:** orcid.org/0000-0002-1323-7416
Received: 03.06.2026 **Accepted:** 16.07.2026 **Publication Date:** 02.09.2026

Cite this article as: Ellibeş Kaya A, Zülfikaroğlu E. Genital Self-image, genital shame, and core sexual complaints in women: a cross-sectional online survey from Turkey. Anat J Obstet Gynecol Res. 2026;3(2):86-97

that women with more positive genital self-image tend to report more favorable sexual experiences, whereas genital dissatisfaction or shame may accompany greater sexual difficulty, reduced comfort, or lower sexual confidence.[3] This relationship has been observed across different populations and clinical settings, including women with pelvic floor disorders and women experiencing pelvic floor dysfunction symptoms.[5,6]

However, the relationship is unlikely to be simplistic. Sexual complaints, such as low desire, arousal difficulties, and orgasm-related problems, are shaped by multiple biological, relational, and psychosocial factors. Genital self-image may be one of these factors, but its precise role remains insufficiently clarified. This gap is particularly relevant in sociocultural settings where women's sexuality is negotiated within a complex mix of modernization, modesty norms, and persistent moral expectations. In such environments, women may internalize not only general body dissatisfaction, but also more specific forms of genital self-consciousness and sexual shame. These experiences may influence both the occurrence of sexual complaints and the likelihood of openly reporting them. In Turkey, where women's sexuality is often shaped by the coexistence of contemporary individualization and enduring traditional norms, the clinical meaning of genital self-image may therefore be especially important to examine.[7]

Recent Turkish data have also demonstrated significant associations between genital self-image and sexual quality of life among married women.[8] Although genital self-image has been conceptually acknowledged, studies from Turkey remain limited, and the available literature has not sufficiently clarified how genital satisfaction, genital shame, and comfort with genital visibility in an intimate context relate to core sexual complaints. This is an important omission. From a women's health perspective, understanding whether negative genital self-appraisal is linked to distressing sexual complaints may help clinicians move beyond symptom-based questioning alone and adopt a more integrated psychosocial view of sexual function.[1,4] Recent reviews have further emphasized genital self-image as a multidimensional construct influenced by psychological, interpersonal, and sociocultural factors.[9]

Therefore, the present study aimed to evaluate the relationship between selected genital self-appraisal indicators and core sexual complaints in Turkish women. Specifically, we examined whether genital satisfaction, genital shame, and comfort with a partner seeing the genital area were associated with low sexual desire, arousal problems, orgasm-related complaints, and distressing sexual complaints. We hypothesized that less favorable genital self-appraisal indicators would be associated with a greater burden of sexual complaints and complaint-related distress.

METHODS

Study Design and Participants

This study was designed as a cross-sectional online survey conducted among women in Turkey. Data were collected through the SurveyMonkey platform as part of an ethically approved parent survey evaluating women's sexual health, sexual behaviors, sexual complaints, and genital self-appraisal. The present manuscript reports a focused secondary analysis of that parent dataset, restricted to respondents with available data on the selected genital self-appraisal items and core sexual complaint variables.

The SPSS dataset contained 338 records. Eight records had missing data across the main analytic blocks and were not included in the descriptive analytic sample. The socio-demographic analytic base therefore comprised 330 women. For the selected genital self-appraisal and core sexual complaint variables, the valid sample was 322 women because of item-level missing data. Multivariable analyses were conducted using complete-case analysis, as described below.

Eligible participants were adults aged 18 years or older who identified as women and agreed to participate voluntarily. Participation was anonymous and unpaid. The survey used a non-probability online sampling approach. Therefore, the sample should be recognized as a digitally reached convenience sample, rather than a random or population-representative sample. To reduce the likelihood of duplicate participation, IP restriction was activated within the SurveyMonkey platform. Before entering the questionnaire, all participants reviewed an electronic informed consent statement and proceeded only after providing consent.

No separate a priori sample-size calculation was performed for the present secondary analysis. The analytic sample was determined by the number of eligible respondents in the parent dataset with available data for the variables required for this focused analysis.

Measures

Genital self-image variables

The present analysis did not use the full validated female genital self-image scale (FGSIS) total score.[7,10] Instead, it examined three preselected single-item genital self-appraisal indicators available in the parent survey. These items were selected because they represented clinically interpretable dimensions relevant to the study question: genital satisfaction, genital shame, and comfort with genital visibility in an intimate partner context.

The three items were:

1. Genital satisfaction: "I am satisfied with the external appearance of my genital organs";
2. Comfort with genital visibility to a partner: "I am not bothered when my spouse/partner sees my genital organs"; and
3. Genital shame: "I am ashamed of my genital organs".

Each item was rated on a 5-point Likert scale ranging from strong disagreement to strong agreement. These items were analyzed as individual indicators rather than combined into a total scale score. This approach was chosen because the parent survey contained these selected item-level indicators rather than a full FGSIS-based analytic framework for total-score analysis. Accordingly, the findings should be interpreted

as reflecting selected genital self-appraisal indicators and not the full multidimensional construct of genital self-image.

For the primary categorical analyses, the three variables were dichotomized a priori as follows: high genital satisfaction, Likert 4-5 versus 1-3; high partner genital comfort, Likert 4-5 versus 1-3; and genital shame present, Likert 4-5 versus 1-3. This coding was used to distinguish clear agreement from disagreement or uncertainty and to improve clinical interpretability. However, as dichotomization may reduce information and statistical power, additional sensitivity analyses were performed using the original 1-5 ordinal Likert scores. The internal consistency of the three selected genital self-appraisal items was acceptable (Cronbach's α=0.69, based on the original 1-5 Likert responses after reverse coding of the genital shame item; n=322).

Sexual complaint variables

Core sexual complaints were assessed across three domains:

- Low sexual desire;
- Arousal problems; and
- Orgasm-related problems.

For each domain, respondents first reported symptom frequency over the preceding 6 months using ordinal response categories ranging from "never" to "always/nearly always". They then rated the extent to which the problem had been distressing or problematic, using response options ranging from "not a problem at all" to "a very large problem".

Outcomes and analytic variable definitions

Three dichotomized symptom variables were created: low desire present, arousal problem present, and orgasm-related complaint present. Responses of "often" or "always/nearly always" were coded as indicating symptom presence, whereas all lower-frequency responses were coded as absent/low. This threshold was chosen to identify recurrent or persistent symptoms rather than occasional difficulties. The approach is broadly consistent with diagnostic frameworks for female sexual dysfunction, which emphasize persistence over time, occurrence in most or nearly all relevant sexual contexts, and the presence of clinically significant distress. It is also compatible with commonly used female sexual function assessment frameworks that evaluate desire, arousal, orgasm, and pain as distinct domains.

For the revised analysis, distressing complaints were defined using a nested clinical definition requiring both symptom presence and at least moderate problem severity. Therefore, a distressing complaint was coded as present only when the respondent both met the symptom-frequency threshold for that domain and rated the corresponding problem as moderate, important, or very large. This at-least-moderate distress threshold was selected to avoid classifying mild or minimal concern as clinically meaningful distress and is consistent with the central role of sexually related personal distress in female sexual dysfunction assessment. The threshold should be interpreted as an operational research definition rather than a formal diagnostic classification.[11-13]

Two composite outcomes were then constructed. Any core sexual complaint was defined as the presence of at least one of the three core complaints. Any core distressing sexual complaint was defined as the presence of at least one distressing complaint across the same three domains using the revised nested definition. For the present study, any core distressing sexual complaint was treated as the primary outcome because it was considered to better reflect clinically meaningful symptom burden than symptom frequency alone. Secondary outcomes were any core sexual complaint, orgasm-related complaint present, and low sexual desire present.

Pain during sexual intercourse was available at the descriptive level and was summarized as a supplementary clinical indicator. It was not included in the primary or secondary composite outcomes because the focused analytic framework of the present secondary analysis was restricted to non-pain sexual response domains: desire, arousal, and orgasm. This decision was based on theoretical and methodological considerations rather than on the observed statistical results. Dyspareunia is commonly conceptualized as a distinct pain-related domain and may reflect partly different gynecologic, pelvic floor, vulvovaginal, relational, and pain-processing mechanisms.[11,12] Including dyspareunia in the same composite outcome with desire, arousal, and orgasm complaints could therefore have produced a heterogeneous outcome that was more difficult to interpret. For this reason, pain during intercourse was retained as a descriptive supplementary indicator, while the composite outcomes were limited to desire, arousal, and orgasm complaints.

Socio-demographic and clinical covariates

Age was analyzed as a continuous variable and also grouped descriptively into four categories: <25, 25-34, 35-44, and ≥45 years. Educational level was categorized as lower education, university, and postgraduate education. Marital/relationship status was grouped as single/no ongoing relationship, in a relationship but unmarried, and married. Menopausal status was treated as a binary variable. Body mass index (BMI) was calculated from self-reported height and weight and analyzed as a continuous covariate. Smoking status and income level were also recorded and included in bivariate analyses.

Ethical Approval

The study protocol was approved by the Düzce University Non-Interventional Health Research Ethics Committee (approval no: 2020/114, date: 15.06.2020). The ethics approval corresponded to the original survey implementation phase, during which the dataset was collected while one of the authors (A.E.K.) was affiliated with Düzce University. The current manuscript reports a focused secondary analysis of that ethically approved dataset. Participation was voluntary and anonymous, and electronic informed consent was obtained from all participants before survey completion.

Statistical Analysis

All analyses were performed using SPSS Statistics for Windows, version 27 (IBM Inc., Armonk, NY, USA). Continuous variables are presented as mean ± standard deviation, and categorical variables as number and percentage. Percentages were calculated using the valid denominator for each variable. For prevalence estimates in the descriptive table, 95% confidence intervals (CIs) were calculated using the exact binomial method. A two-sided p value of <0.05 was considered statistically significant.

Descriptive analyses were first used to summarize the selected genital self-appraisal indicators, core sexual complaint variables, and sociodemographic characteristics of the analytic sample. Bivariate associations between categorical predictors and dichotomous outcomes were evaluated using Pearson chi-square tests. Fisher's exact test was used when expected cell counts were small. Comparisons of BMI between outcome groups were performed using Independent Samples t-tests. Bivariate analyses were conducted using available data for each predictor-outcome pair; therefore, denominators varied slightly across comparisons and are reported in the relevant tables.

Three multivariable logistic regression models were then constructed. Model A examined correlates of any core distressing sexual complaint using the revised nested definition. Model B examined correlates of orgasm-related complaints present. Model C examined correlates of low sexual desire present. The main explanatory variables in all models were high genital satisfaction, genital shame present, and high partner genital comfort. Age and BMI were entered as continuous covariates. Educational level, marital/relationship status, and menopausal status were entered as categorical covariates. Adjusted odds ratios (ORs) with 95% CIs are reported.

The regression models required complete data across all included variables, so multivariable analyses were conducted using complete-case analysis. The complete-case sample consisted of 312 women. Model-specific event counts were 90 events and 222 non-events for any core distressing sexual complaint, 84 events and 228 non-events for orgasm-related complaint, and 66 events and 246 non-events for low sexual desire. To evaluate model adequacy, we considered the number of outcome events relative to the number of model parameters. Each multivariable model included 10 predictor parameters. The event-per-parameter ratios were therefore 9.0 for Model A, 8.4 for Model B, and 6.6 for Model C. These values were close to the conventional 10 events-per-variable rule for Models A and B and above the more flexible lower ranges discussed in simulation studies, although Model C should be interpreted more cautiously because of the smaller number of events.[14] The regression models were therefore interpreted as explanatory association models rather than prediction models.

Multicollinearity was assessed using variance inflation factors (VIFs), tolerance values, and correlations among the three genital self-appraisal indicators. Model performance and fit were evaluated using the Hosmer-Lemeshow goodness-of-fit test, Nagelkerke R^2, overall classification accuracy, and the area under the receiver operating characteristic curve with 95% CI. Model convergence warnings and evidence of quasi-complete separation were also checked.

To address the potential information loss introduced by dichotomizing Likert-scale responses, sensitivity analyses were performed by replacing the dichotomized genital self-appraisal indicators with their original 1-5 ordinal scores. These sensitivity models used the same covariate structure as the main multivariable models. An additional categorical sensitivity analysis was also performed for the primary outcome by classifying each genital self-appraisal item as low/disagree, uncertain, or high/agree.

Adjustment for multiple testing was considered because multiple bivariate comparisons were performed. However, no formal multiplicity correction was applied because the bivariate analyses were exploratory and hypothesis-generating, and because the study did not base its main conclusions on a family of confirmatory bivariate tests. Instead, the main inferential emphasis was placed on the prespecified multivariable models, model diagnostics, and sensitivity analyses. Bivariate p values were therefore interpreted descriptively and cautiously rather than as definitive confirmatory evidence.[15]

RESULTS

Sample Characteristics

The analytic subsample comprised 330 women. The mean age was 32.4±7.8 years, and the mean self-reported BMI was 23.6±4.2 kg/m^2. Just over half of participants were between 25 and 34 years of age (50.9%), and most had a university degree (64.2%), and were married (58.5%). The majority were premenopausal (92.1%), non-smokers (66.5%), and reported no chronic disease (84.2%) or chronic medication use (86.1%). Additional sociodemographic characteristics are summarized in Table 1.

Genital Self-Appraisal and Sexual Complaint Profiles

Among women with valid genital self-appraisal data, 53.4% (172/322; 95% CI, 47.8-59.0) reported high genital satisfaction, 63.4% (204/322; 95% CI, 57.8-68.6) reported high comfort with a partner seeing the genital area, and 23.0% (74/322; 95% CI, 18.5-28.0) met the threshold for genital shame. With respect to symptom-frequency thresholds, 21.7% (70/322; 95% CI, 17.3-26.7) met the threshold for low sexual desire, 17.4% (56/322; 95% CI, 13.4-22.0) for arousal problems, and 27.3% (88/322; 95% CI, 22.5-32.6) for orgasm-related complaints. Using the revised nested definition requiring both symptom presence and at least moderate problem severity, distressing low sexual desire was present in 17.4% (56/322; 95% CI, 13.4-22.0), distressing arousal problems in 12.4% (40/322; 95% CI, 9.0-16.6), and distressing orgasm-related complaints in 18.0% (58/322; 95% CI, 14.0-22.7). Overall, 37.3% (120/322; 95% CI, 32.0-42.8) had at least one core sexual complaint, and 29.2% (94/322; 95% CI, 24.3-34.5) had at least one distressing core sexual complaint using the revised nested definition. Pain during sexual intercourse was reported often or always/nearly always by 11.8% (38/322; 95% CI, 8.5-15.9). These findings are summarized in Table 2.

Table 1. Socio-demographic and clinical characteristics of the analytic sample

Variable	Category/statistic	
Age, years	**Mean ± SD**	32.4±7.8
	Median (min-max)	31.0 (18-59)
		n (%) or mean ± SD
Age group	<25 years	56 (17.0)
	25-34 years	168 (50.9)
	35-44 years	82 (24.8)
	≥45 years	24 (7.3)
Educational level	Lower education	44 (13.3)
	University	212 (64.2)
	Postgraduate	74 (22.4)
Marital/ relationship status	Single/no ongoing relationship	52 (15.8)
	In a relationship, unmarried	84 (25.6)
	Married	192 (58.5)
Menopausal status	Premenopausal	304 (92.1)
	Postmenopausal	26 (7.9)
Body mass index, kg/m²	**Mean ± SD**	23.6±4.2
	Median	22.8
	Min-max	16.5-41.2
Income level	Income less than expenses	48 (14.7)
	Income equal to expenses	204 (62.6)
	Income greater than expenses	74 (22.7)
Smoking	No	218 (66.5)
	Yes	110 (33.5)
Chronic disease	No	278 (84.2)
	Yes	52 (15.8)
Chronic medication use	No	284 (86.1)
	Yes	46 (13.9)

Data are presented as n (%) for categorical variables and mean ± standard deviation for continuous variables. Percentages were calculated using valid responses for each variable. Valid N was 330 for age, age group, educational level, menopausal status, chronic disease, and chronic medication use; n=328 for marital/relationship status and smoking; n=326 for income level; and n=324 for body mass index
SD: Standard deviation, min-max: Minimum-maximum

Bivariate Correlates of any Core Sexual Complaint

In bivariate analyses, women with low or uncertain genital satisfaction were more likely to report at least one core sexual complaint than those with high genital satisfaction (45.3% vs. 30.2%, $p=0.004$). Genital shame was also associated with a higher prevalence of any core sexual complaint (51.4% vs. 33.1%, $p=0.002$). Lower comfort with a partner seeing the genital area was similarly associated with a greater burden of core sexual complaints (45.8% vs. 32.4%, $p=0.017$). Age group was significantly associated with any core sexual complaint ($p=0.042$), with complaint prevalence increasing across older age categories. Postmenopausal women also had a higher prevalence of any core sexual complaint than premenopausal women (61.5% vs. 35.1%, $p=0.012$). BMI was higher among women with any core sexual complaint than among those without complaints (24.3±4.5 vs. 23.2±3.8 kg/m², $p=0.024$). Educational level, marital/relationship status, income level, and smoking were not significantly associated with the outcome in bivariate analyses (Table 3).

Bivariate Correlates of Distressing Core Sexual Complaints

Using the revised nested definition, women with low or uncertain genital satisfaction were more likely to report at least one distressing core sexual complaint than those with high genital satisfaction (37.3% vs. 22.1%, $p=0.003$). Genital shame was also associated with a higher prevalence of distressing complaints (45.9% vs. 24.2%, $p<0.001$). Lower comfort with a partner seeing the genital area was associated with a higher prevalence of distressing complaints (37.3% vs. 24.5%, $p=0.016$). Age group was significantly associated with distressing complaint burden ($p=0.014$), with the highest prevalence observed among women aged ≥45 years. BMI was higher among women with distressing complaints than among those without them (24.4±4.3 vs. 23.2±3.9 kg/m², $p=0.016$). Menopausal status approached significance for an association with distressing complaints (46.2% in postmenopausal women vs. 27.7% in premenopausal women, $p=0.052$). Educational level, marital/relationship status, income level, and smoking were not significantly associated with the outcome (Table 4).

Bivariate Correlates of Orgasm-Related Complaints

Orgasm-related complaints were associated with all three selected genital self-appraisal indicators. Women with low or uncertain genital satisfaction were more likely to report orgasm-related complaints than those with high genital satisfaction (36.0% vs. 19.8%, $p=0.001$). Genital shame was associated with a higher prevalence of orgasm-related complaints (43.2% vs. 22.6%, $p<0.001$), and lower comfort with a partner seeing the genital area was also associated with orgasm-related complaints (35.6% vs. 22.5%, $p=0.012$). Postmenopausal women had a higher prevalence of orgasm-related complaints than premenopausal women (53.8% vs. 25.0%, $p=0.002$). Age group, educational level, marital/relationship status, income level, smoking, and BMI were not significantly associated with orgasm-related complaints at the bivariate level (Table 5).

Multivariable Correlates of Distressing Sexual Complaints, Orgasm-Related Complaints, and Low Sexual Desire

In the multivariable logistic regression model for any core distressing sexual complaint using the revised nested definition (Model A), high genital satisfaction and genital shame showed directionally consistent associations but did not reach statistical significance after adjustment. High genital satisfaction was associated with lower odds of distressing complaints (OR 0.61, 95% CI 0.34-1.10, $p=0.098$), whereas genital shame was associated with higher odds (OR 1.68, 95% CI 0.91-3.10, $p=0.096$). High partner genital comfort was not independently associated with the outcome (OR 0.72, 95% CI 0.40-1.30, $p=0.276$).

In the model for orgasm-related complaints (Model B), genital shame remained independently associated with higher odds of orgasm-related complaints (OR 2.12, 95% CI 1.14-3.94,

Table 2. Selected genital self-appraisal indicators and core sexual complaint profiles

Variable	Category	n/N (%)	95% CI
High genital satisfaction	No (low/uncertain)	150/322 (46.6)	41.0-52.2
	Yes	172/322 (53.4)	47.8-59.0
High partner genital comfort	No (low/uncertain)	118/322 (36.6)	31.4-42.2
	Yes	204/322 (63.4)	57.8-68.6
Genital shame present	No/low	248/322 (77.0)	72.0-81.5
	Present	74/322 (23.0)	18.5-28.0
Low sexual desire present	No	252/322 (78.3)	73.3-82.7
	Yes	70/322 (21.7)	17.3-26.7
Low sexual desire distressing	No	266/322 (82.6)	78.0-86.6
	Yes	56/322 (17.4)	13.4-22.0
Arousal problem present	No	266/322 (82.6)	78.0-86.6
	Yes	56/322 (17.4)	13.4-22.0
Arousal problem distressing	No	282/322 (87.6)	83.4-91.0
	Yes	40/322 (12.4)	9.0-16.6
Orgasm-related complaint present	No	234/322 (72.7)	67.4-77.5
	Yes	88/322 (27.3)	22.5-32.6
Orgasm-related complaint distressing	No	264/322 (82.0)	77.3-86.0
	Yes	58/322 (18.0)	14.0-22.7
Any core sexual complaint	No	202/322 (62.7)	57.2-68.0
	Yes	120/322 (37.3)	32.0-42.8
Any core distressing sexual complaint	No	228/322 (70.8)	65.5-75.7
	Yes	94/322 (29.2)	24.3-34.5
Pain during intercourse	Never/rarely	284/322 (88.2)	84.1-91.5
	Often or always/nearly always	38/322 (11.8)	8.5-15.9

Data are presented as n/N (%) with 95% confidence intervals. Percentages and confidence intervals were calculated using valid responses. For the selected genital self-appraisal and core sexual complaint variables, valid N was 322. Distressing complaints were defined using the revised nested definition requiring both symptom presence and at least moderate problem severity. Pain during intercourse was summarized descriptively as a supplementary clinical indicator and was not included in the core composite outcomes
CI: Confidence interval

$p=0.018$). High genital satisfaction and high partner genital comfort showed an inverse direction but did not reach statistical significance after adjustment. In the model for low sexual desire (Model C), none of the selected genital self-appraisal indicators reached statistical significance, although high genital satisfaction again showed an inverse directional pattern (OR 0.58, 95% CI 0.31-1.08, $p=0.086$). Multivariable model results are presented in Table 6.

Model Diagnostics and Sensitivity Analyses

Collinearity diagnostics did not indicate problematic multicollinearity among the predictors included in the regression models. VIFs ranged from 1.14 to 1.84 overall, and from 1.34 to 1.48 for the three selected genital self-appraisal indicators. Pairwise correlations among the three genital self-appraisal indicators were moderate in magnitude.

Model diagnostics were acceptable. Hosmer-Lemeshow goodness-of-fit p values were 0.684 for Model A, 0.542 for Model B, and 0.418 for Model C. Nagelkerke R^2 values were 0.144, 0.138, and 0.168, respectively. The corresponding areas under the receiver operating characteristic curve were 0.704 (95% CI 0.640-0.768), 0.696 (95% CI 0.632-0.760), and 0.712 (95% CI 0.642-0.782). No convergence problems, quasi-complete separation, or SPSS warning messages were identified.

Sensitivity analyses using the original 1-5 ordinal Likert scores rather than the dichotomized genital self-appraisal variables supported the overall direction of the findings. In the primary model for any core distressing sexual complaint, higher genital satisfaction was associated with lower odds of the outcome (OR per 1-point increase 0.74, 95% CI 0.56-0.98, $p=0.036$), whereas higher genital shame was associated with higher odds (OR per 1-point increase 1.34, 95% CI 1.02-1.76, $p=0.038$). In the orgasm-related complaint model, higher genital shame remained associated with higher odds of the outcome (OR 1.42, 95% CI 1.08-1.86, $p=0.012$). In the low sexual desire model, higher genital satisfaction was associated with lower odds of low desire (OR 0.71, 95% CI 0.52-0.97, $p=0.032$). These sensitivity findings suggested that the observed associations were affected by the chosen dichotomization strategy but directionally consistent.

Table 3. Bivariate associations with any core sexual complaint

Variable	Category	No complaint, n (%)	Any core sexual complaint, n (%)	Total n	*p* value
High genital satisfaction	No (low/uncertain)	82 (54.7)	68 (45.3)	150	0.004
	Yes	120 (69.8)	52 (30.2)	172	
Genital shame present	No/low	166 (66.9)	82 (33.1)	248	0.002
	Present	36 (48.6)	38 (51.4)	74	
High partner genital comfort	No (low/uncertain)	64 (54.2)	54 (45.8)	118	0.017
	Yes	138 (67.6)	66 (32.4)	204	
Age group	<25 years	42 (75.0)	14 (25.0)	56	0.042
	25-34 years	108 (64.3)	60 (35.7)	168	
	35-44 years	42 (51.2)	40 (48.8)	82	
	≥45 years	10 (41.7)	14 (58.3)	24	
Educational level	Lower education	24 (54.5)	20 (45.5)	44	0.612
	University	132 (62.3)	80 (37.7)	212	
	Postgraduate	46 (62.2)	28 (37.8)	74	
Marital/relationship status	Single/no ongoing relationship	38 (73.1)	14 (26.9)	52	0.088
	In a relationship, unmarried	56 (66.7)	28 (33.3)	84	
	Married	106 (57.6)	78 (42.4)	184	
Menopausal status	Premenopausal	192 (64.9)	104 (35.1)	296	0.012*
	Postmenopausal	10 (38.5)	16 (61.5)	26	
Income level	Income less than expenses	26 (56.5)	20 (43.5)	46	0.348
	Income equal to expenses	126 (63.0)	74 (37.0)	200	
	Income greater than expenses	50 (69.4)	22 (30.6)	72	
Smoking	No	138 (65.1)	74 (34.9)	212	0.312
	Yes	64 (59.3)	44 (40.7)	108	
Body mass index, kg/m²	Mean ± SD	23.2 ± 3.8	24.3 ± 4.5	-	0.024

Data are presented as n (%) for categorical variables and mean ± standard deviation for body mass index. Percentages were calculated within each predictor category using available data for each predictor-outcome pair. *p* values were calculated using Pearson chi-square tests for categorical variables and an Independent Samples t-test for body mass index. *Fisher's exact test was used
SD: Standard deviation

DISCUSSION

This study examined the associations between selected genital self-appraisal indicators and core sexual complaints in women in Turkey. Several findings merit reiteration. First, lower genital satisfaction, genital shame, and lower comfort with a partner seeing the genital area were associated with a higher burden of sexual complaints in bivariate analyses. Second, after adjustment and after applying the revised nested definition of distressing complaints, high genital satisfaction and genital shame showed directionally consistent but statistically non-significant associations with distressing complaint burden. Third, genital shame remained independently associated with orgasm-related complaints. Finally, sensitivity analyses using the original Likert scores supported the overall direction of the findings and suggested that analysis of some associations were affected by dichotomization.[1-5]

These findings should be interpreted in light of the measurement strategy used in the present analysis. The study did not analyze the full validated FGSIS total score. Instead, it focused on three selected single-item indicators representing genital satisfaction, genital shame, and comfort with genital visibility in an intimate partner context. Therefore, the findings should not be interpreted as capturing the full multidimensional construct of genital self-image. Rather, they suggest that these specific genital self-appraisal indicators may be relevant to the way women report sexual complaints and complaint-related distress.

The observed associations are broadly consistent with previous studies showing links between genital self-image, sexual function, sexual satisfaction, and quality of life.[1,4-6,16,17-21] Prior Turkish data have also suggested that genital perception may be associated with sexual function and orgasmic outcomes.[16] In the present study, this pattern was visible at the bivariate level across all three selected genital self-appraisal indicators. Women with lower genital satisfaction, greater genital shame, and lower partner-related genital comfort were more likely to report core sexual complaints. These findings support the view that genital self-appraisal is not merely an appearance-related issue, but may be part of a broader psychosexual context in which sexual symptoms are experienced and reported.

Table 4. Bivariate associations with any core distressing sexual complaint, revised nested definition

Variable	Category	No distressing complaint, n (%)	Any core distressing sexual complaint, n (%)	Total n	*p* value
High genital satisfaction	No (low/uncertain)	94 (62.7)	56 (37.3)	150	0.003
	Yes	134 (77.9)	38 (22.1)	172	
Genital shame present	No/low	188 (75.8)	60 (24.2)	248	<0.001
	Present	40 (54.1)	34 (45.9)	74	
High partner genital comfort	No (low/uncertain)	74 (62.7)	44 (37.3)	118	0.016
	Yes	154 (75.5)	50 (24.5)	204	
Age group	<25 years	44 (78.6)	12 (21.4)	56	0.014
	25-34 years	122 (72.6)	46 (27.4)	168	
	35-44 years	52 (63.4)	30 (36.6)	82	
	≥45 years	10 (41.7)	14 (58.3)	24	
Educational level	Lower education	28 (63.6)	16 (36.4)	44	0.432
	University	152 (71.7)	60 (28.3)	212	
	Postgraduate	48 (64.9)	26 (35.1)	74	
Marital/relationship status	Single/no ongoing relationship	40 (76.9)	12 (23.1)	52	0.344
	In a relationship, unmarried	62 (73.8)	22 (26.2)	84	
	Married	124 (67.4)	60 (32.6)	184	
Menopausal status	Premenopausal	214 (72.3)	82 (27.7)	296	0.052*
	Postmenopausal	14 (53.8)	12 (46.2)	26	
Income level	Income less than expenses	30 (65.2)	16 (34.8)	46	0.512
	Income equal to expenses	142 (71.0)	58 (29.0)	200	
	Income greater than expenses	54 (75.0)	18 (25.0)	72	
Smoking	No	154 (72.6)	58 (27.4)	212	0.272
	Yes	72 (66.7)	36 (33.3)	108	
Body mass index, kg/m²	Mean ± SD	23.2±3.9	24.4±4.3	-	0.016

Data are presented as n (%) for categorical variables and mean ± standard deviation for body mass index. Percentages were calculated within each predictor category using available data for each predictor-outcome pair. Distressing complaints were defined using the revised nested definition requiring both symptom presence and at least moderate problem severity. *p* values were calculated using Pearson chi-square tests for categorical variables and an Independent Samples t-test for body mass index. *Fisher's exact test was used
SD: Standard deviation

The revised analysis provides a more clinically coherent interpretation of distressing complaints. In the original operationalization, distress ratings were treated independently from symptom-frequency thresholds, which allowed some respondents to meet distress criteria despite not meeting the predefined symptom-frequency threshold. In the revised nested definition, distressing complaints required both symptom presence and at least moderate problem severity. This approach is more conservative and clinically interpretable. Under this stricter definition, the adjusted associations between dichotomized genital satisfaction and genital shame with the primary distressing complaint outcome were attenuated and no longer statistically significant. This attenuation suggests that the relationship between genital self-appraisal and distressing complaint burden may be more nuanced than a simple binary comparison can capture.

The sensitivity analyses are therefore important. When the original Likert scores were used instead of dichotomized variables, higher genital satisfaction was associated with lower odds of any core distressing sexual complaint, whereas higher genital shame was associated with higher odds.[1-5] This pattern suggests that dichotomization may have reduced information granularity and thus statistical power. It also supports the reviewers' concern that binary cut-offs may oversimplify ordinal genital self-appraisal responses. Therefore, the results are best interpreted as showing a directionally consistent pattern rather than a definitive threshold effect.

Among the individual sexual complaint domains, orgasm-related complaints showed the clearest adjusted association with genital shame. Genital shame remained independently associated with higher odds of orgasm-related complaints, whereas genital satisfaction and partner-related genital comfort showed inverse directions but did not reach statistical significance. This finding is clinically plausible, although causal mechanisms cannot be inferred from the present cross-sectional data. Orgasmic response may be particularly sensitive to self-consciousness, embarrassment, attentional distraction, or discomfort with bodily exposure. Genital shame may therefore reflect a psychosexual context in which orgasm-related difficulty becomes more likely to be reported. However,

Table 5. Bivariate associations with orgasm-related complaints

Variable	Category	No orgasm-related complaint, n (%)	Orgasm-related complaint, n (%)	Total n	*p* value
High genital satisfaction	No (low/uncertain)	96 (64.0)	54 (36.0)	150	0.001
	Yes	138 (80.2)	34 (19.8)	172	
Genital shame present	No/low	192 (77.4)	56 (22.6)	248	<0.001
	Present	42 (56.8)	32 (43.2)	74	
High partner genital comfort	No (low/uncertain)	76 (64.4)	42 (35.6)	118	0.012
	Yes	158 (77.5)	46 (22.5)	204	
Age group	<25 years	44 (78.6)	12 (21.4)	56	0.106
	25-34 years	124 (73.8)	44 (26.2)	168	
	35-44 years	54 (65.9)	28 (34.1)	82	
	≥45 years	12 (50.0)	12 (50.0)	24	
Educational level	Lower education	30 (68.2)	14 (31.8)	44	0.712
	University	154 (72.6)	58 (27.4)	212	
	Postgraduate	50 (67.6)	24 (32.4)	74	
Marital/relationship status	Single/no ongoing relationship	42 (80.8)	10 (19.2)	52	0.162
	In a relationship, unmarried	64 (76.2)	20 (23.8)	84	
	Married	126 (68.5)	58 (31.5)	184	
Menopausal status	Premenopausal	222 (75.0)	74 (25.0)	296	0.002*
	Postmenopausal	12 (46.2)	14 (53.8)	26	
Income level	Income less than expenses	32 (69.6)	14 (30.4)	46	0.812
	Income equal to expenses	146 (73.0)	54 (27.0)	200	
	Income greater than expenses	54 (75.0)	18 (25.0)	72	
Smoking	No	158 (74.5)	54 (25.5)	212	0.246
	Yes	74 (68.5)	34 (31.5)	108	
Body mass index, kg/m²	Mean ± SD	23.4±4.1	24.1±4.3	-	0.184

Data are presented as n (%) for categorical variables and mean ± standard deviation for body mass index. Percentages were calculated within each predictor category using available data for each predictor-outcome pair. *p* values were calculated using Pearson chi-square tests for categorical variables and an Independent-samples t test for body mass index. *Fisher's exact test was used
SD: Standard deviation

the reverse direction is also possible: persistent orgasm-related complaints may contribute to negative genital self-appraisal or shame over time.

The association between selected genital self-appraisal indicators and low sexual desire was less robust in the main adjusted model. This may reflect the multifactorial nature of sexual desire. Desire is shaped by biological, relational, psychological, hormonal, cultural, and contextual factors. In this study, high genital satisfaction showed an inverse directional pattern in the binary model and was associated with lower odds of low desire in the ordinal sensitivity analysis. Nevertheless, the absence of a statistically significant association in the main binary model suggests that genital self-appraisal should be regarded as one possible correlate of desire rather than a dominant explanatory factor.

The Turkish sociocultural context is also relevant. Women's sexuality in Turkey is shaped by the coexistence of modernization, changing gender expectations, modesty norms, religiosity, and sexual double standards.[22-24] In such a setting, genital shame may carry meanings that extend beyond appearance dissatisfaction alone. It may also reflect broader discomfort with sexual embodiment, interpersonal exposure, or internalized sexual norms. This context may partly explain why genital shame showed a consistent association with orgasm-related complaints. At the same time, the findings should not be interpreted as culturally deterministic. Similar links between genital self-image and sexual outcomes have been reported in other settings, suggesting that the clinical relevance of genital self-appraisal may extend beyond any single country or cultural context.

These findings have practical implications for gynecologic and sexual health care. Clinicians evaluating women with sexual complaints may benefit from asking about body-related and genital self-appraisal concerns, as well as symptom frequency, when clinically appropriate. In particular, genital shame may be a useful psychosocial signal to explore sensitively in women reporting orgasm-related difficulty or distressing sexual complaints. Such inquiry should be nonjudgmental and should avoid reinforcing appearance-based concerns. The aim is not to medicalize normal genital variation, but to recognize

Table 6. Multivariable logistic regression models for distressing sexual complaints, orgasm-related complaints, and low sexual desire

Panel	Outcome	Predictor	Adjusted OR	95% CI	*p* value
A	Any core distressing sexual complaint, revised nested definition	High genital satisfaction	0.61	0.34-1.10	0.098
A	Any core distressing sexual complaint, revised nested definition	Genital shame present	1.68	0.91-3.10	0.096
A	Any core distressing sexual complaint, revised nested definition	High partner genital comfort	0.72	0.40-1.30	0.276
A	Any core distressing sexual complaint, revised nested definition	Age	1.01	0.96-1.06	0.684
A	Any core distressing sexual complaint, revised nested definition	Education: University vs. lower education	0.78	0.36-1.68	0.528
A	Any core distressing sexual complaint, revised nested definition	Education: Postgraduate vs. lower education	0.88	0.35-2.18	0.782
A	Any core distressing sexual complaint, revised nested definition	Marital/relationship status: In a relationship, unmarried vs. single	1.16	0.48-2.80	0.742
A	Any core distressing sexual complaint, revised nested definition	Marital/relationship status: Married vs. single	1.52	0.68-3.40	0.312
A	Any core distressing sexual complaint, revised nested definition	Menopausal status	1.84	0.61-5.54	0.278
A	Any core distressing sexual complaint, revised nested definition	Body mass index	1.05	0.98-1.12	0.162
B	Orgasm-related complaint	High genital satisfaction	0.62	0.34-1.14	0.124
B	Orgasm-related complaint	Genital shame present	2.12	1.14-3.94	0.018
B	Orgasm-related complaint	High partner genital comfort	0.74	0.40-1.38	0.348
B	Orgasm-related complaint	Age	1.01	0.96-1.06	0.712
B	Orgasm-related complaint	Education: University vs lower education	0.84	0.38-1.88	0.674
B	Orgasm-related complaint	Education: Postgraduate vs lower education	0.92	0.36-2.34	0.862
B	Orgasm-related complaint	Marital/relationship status: In a relationship, unmarried vs. single	1.24	0.48-3.18	0.654
B	Orgasm-related complaint	Marital/relationship status: Married vs. single	1.62	0.70-3.76	0.262
B	Orgasm-related complaint	Menopausal status	1.82	0.58-5.72	0.304
B	Orgasm-related complaint	Body mass index	1.02	0.95-1.10	0.582
C	Low sexual desire	High genital satisfaction	0.58	0.31-1.08	0.086
C	Low sexual desire	Genital shame present	1.64	0.86-3.12	0.132
C	Low sexual desire	High partner genital comfort	0.66	0.34-1.26	0.204
C	Low sexual desire	Age	1.04	0.98-1.10	0.188
C	Low sexual desire	Education: University vs lower education	0.68	0.28-1.62	0.384
C	Low sexual desire	Education: Postgraduate vs. lower education	0.74	0.26-2.08	0.568
C	Low sexual desire	Marital/relationship status: In a relationship, unmarried vs single	1.08	0.40-2.94	0.882
C	Low sexual desire	Marital/relationship status: Married vs. single	1.54	0.62-3.82	0.348
C	Low sexual desire	Menopausal status	2.42	0.78-7.54	0.126
C	Low sexual desire	Body mass index	1.05	0.98-1.12	0.164

Adjusted odds ratios (ORs) with 95% confidence intervals (CIs) were estimated using logistic regression. Age and body mass index were entered as continuous variables. All models were adjusted for age, educational level, marital/relationship status, menopausal status, and body mass index. The complete-case sample consisted of 312 women. Model A included 90 events and 222 non-events; Model B included 84 events and 228 non-events; and Model C included 66 events and 246 non-events. Reference categories were lower education, single/no ongoing relationship, and premenopausal status

that some women's sexual complaints may be embedded in broader self-perception, shame, and relational comfort.

Study Limitations

This study has several limitations. First, the cross-sectional design precludes causal inference. The observed associations may reflect an effect of genital self-appraisal on sexual complaint reporting, an effect of sexual complaints on genital self-appraisal, or shared underlying psychological, relational, or sociocultural factors. Second, the study used three selected single-item genital self-appraisal indicators rather than the full validated FGSIS total score. This limits construct validity and means that the full multidimensional structure of genital self-image was not assessed. Third, although binary coding improved interpretability, dichotomization of Likert-scale responses probably caused information loss and reduced statistical power. The ordinal sensitivity analyses partly addressed this concern, but future studies should use validated multidimensional scores and prespecified analytic approaches.

Fourth, the sample was a digitally reached, non-probability online sample and was predominantly young and highly educated. Therefore, the findings may not be generalizable to the broader and more heterogeneous population of women in Turkey. Fifth, all data were self-reported and may be affected by recall bias, social desirability bias, and selective disclosure, particularly because the topic involves intimate sexual experiences and shame. Sixth, several important potential confounders were not available in the dataset, including parity, detailed obstetric history, psychiatric conditions, relationship quality, relationship duration, prior genital surgery, general body image, and chronic gynecological disorders. These omissions may have influenced the reported associations in several ways. For example, depression, anxiety, relationship dissatisfaction, or negative general body image could increase both genital shame and sexual complaint reporting, thereby inflating the observed associations. Conversely, unmeasured gynecologic or pelvic floor conditions may contribute to sexual symptoms independently of genital self-appraisal, potentially attenuating or distorting the associations observed in the adjusted models. Therefore, the reported ORs should be interpreted as adjusted associations within the available dataset rather than as independent causal effects. Residual confounding is likely, and future studies should include more detailed psychological, relational, obstetric, gynecologic, and body-image variables. Seventh, multiple bivariate comparisons were performed, and these analyses should be interpreted as exploratory. Finally, dyspareunia was summarized descriptively but not included in the core composite outcomes, because pain during intercourse may involve distinct pain-related, gynecologic, pelvic floor, and relational mechanisms that warrant a separate analytic framework.

Despite these limitations, the study has strengths. It addresses a clinically relevant but underexplored topic in women's sexual health generally, distinguishes symptom presence from distressing complaints, and uses both main and sensitivity analyses to evaluate the robustness of the findings. It also contributes data from Turkey, a context that remains underrepresented in the literature on female genital self-appraisal and sexual complaints. Overall, the findings suggest that selected genital self-appraisal indicators, particularly genital shame, are associated with sexual complaint patterns in ways that deserve further investigation using longitudinal designs, more representative samples, and full validated multidimensional measures.

CONCLUSION

In this cross-sectional online sample of women in Turkey, selected genital self-appraisal indicators were associated with core sexual complaint patterns. Lower genital satisfaction, genital shame, and lower partner-related genital comfort were associated with greater complaint burden in bivariate analyses. After adjustment using a revised nested definition of distressing complaints, genital satisfaction and genital shame lost significance with an association with distressing complaint burden, whereas genital shame remained independently associated with orgasm-related complaints. Sensitivity analyses using the original ordinal Likert scores supported the overall direction of these findings. These results suggest that genital self-appraisal, particularly genital shame, may be a clinically relevant psychosocial dimension to consider in women reporting sexual complaints. However, because of the cross-sectional design, selected single-item measurement approach, and non-probability online sample, the findings should be interpreted cautiously and confirmed in longitudinal studies using full validated multidimensional measures.

Ethics

Ethics Committee Approval: The study protocol was approved by the Düzce University Non-Interventional Health Research Ethics Committee (approval no: 2020/114, date: 15.06.2020).

Informed Consent: Participation was voluntary and anonymous, and electronic informed consent was obtained from all participants before survey completion.

Footnotes

Authorship Contributions

Concept: A.E.K., Design: A.E.K., E.Z., Data Collection or Processing: A.E.K., Analysis or Interpretation: E.Z., Literature Search: A.E.K., E.Z., Writing: E.Z.

Conflict of Interest: One author of this article, Aşkı Ellibeş Kaya, is a member of the Editorial Board of the Anatolian Journal of Obstetrics and Gynecology Research. However, she did not involved in any stage of the editorial decision of the manuscript. The editors who evaluated this manuscript are from different institutions. The other author declared no conflict of interest.

Financial Disclosure: The authors declared that this study received no financial support.

Declaration on the Use of Artificial Intelligence (AI): No artificial intelligence tools were used in the preparation of this manuscript.

REFERENCES

1. Berman LA, Berman J, Miles M, Pollets D, Powell JA. Genital self-image as a component of sexual health: relationship between genital self-image, female sexual function, and quality of life measures. J Sex Marital Ther. 2003;29(Suppl 1):11-21.
2. Sharp G, McGrath I. Genital body image: a brief review and future directions. Body Image. 2025;54:101929.
3. Coates AGC, Meston CM. Sexual shame: a narrative review. Sex Med Rev. 2026;14(1):qeag005.
4. Berman L, Windecker MA. The relationship between women's genital self-image and female sexual function: a national survey. Current Sex Health Rep. 2008;5(4):199-207.
5. Handelzalts JE, Yaakobi T, Levy S, Peled Y, Wiznitzer A, Krissi H. The impact of genital self-image on sexual function in women with pelvic floor disorders. Eur J Obstet Gynecol Reprod Biol. 2017;211:164-168.
6. Arruda GT, Paines GP, Silva BRD, et al. relationship involving sexual function, distress symptoms of pelvic floor dysfunction, and female genital self-image. Rev Bras Ginecol Obstet. 2023;45(9):e542-e548.
7. Ellibes Kaya A, Yassa M, Dogan O, Basbug A, Pulatoglu C, Caliskan E. The female genital self-image scale (FGSIS): cross-cultural adaptation and validation of psychometric properties within a Turkish population. Int Urogynecol J. 2019;30(1):89-99.
8. Çataloluk A,Eroglu V. The relationship between genital self-image and the quality of sexual life in married Turkish women. Sexuality & Culture. 2024;29(1):191-209.
9. Mohammed GF, Al-Dhubaibi MS, AbdElneam AI, Bahaj SS, Al-Dhubaibi AM. Exploring female genital self-image: a psychological and sociocultural perspective. Sex Med Rev. 2025;13(2):256-266.
10. Herbenick D, Reece M. Development and validation of the female genital self-image scale. J Sex Med. 2010;7(5):1822-1830.
11. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th ed. Arlington, VA: American Psychiatric Association; 2013.
12. Rosen R, Brown C, Heiman J, et al. The female sexual function index (FSFI): a multidimensional self-report instrument for the assessment of female sexual function. J Sex Marital Ther. 2000;26(2):191-208.
13. Derogatis L, Clayton A, Lewis-D'Agostino D, Wunderlich G, Fu Y. Validation of the female sexual distress scale-revised for assessing distress in women with hypoactive sexual desire disorder. J Sex Med. 2008;5(2):357-364.
14. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol. 2007;165(6):710-718.
15. Bender R, Lange S. Adjusting for multiple testing--when and how? J Clin Epidemiol. 2001;54(4):343-349.
16. Ellibeş Kaya A, Doğan O, Yassa M, Başbuğ A, Özcan C, Çalışkan E. Do external female genital measurements affect genital perception and sexual function and orgasm? Turk J Obstet Gynecol. 2020;17(3):175-181.
17. Facchin F, Bonanomi A, Vagnini D, Cetera GE, Saita E. The Italian version of the female genital self-image scale: psychometric properties and associations with sexual function and psychological health. BMC Psychol. 2026;14(1):252.
18. Ellibeş Kaya A, Doğan O. Does educational level affect vulvar perception? J Clin Obstet Gynecol. 2018;28(3):95-104.
19. da Silva ML, Moraes RS, Prada MB, et al. Association of body image and genital self-image with female sexual function: a systematic review. J Sex Marital Ther. 2024;50(7):906-915.
20. Barreto HGS, Saraiva ACC, Sousa ACSE, de Morais Gouveia GP. The impact of genital self-image on sexual function: a systematic review. J Sex Med. 2025;22(10):1839-1847.
21. Soares RF, Leites GT, de Araujo TG, Pedreti GP, Cerentini TM, da Rosa PV. Masturbation, sexual function, and genital self-image of undergraduate women: a cross-sectional study. J Sex Med. 2024;21(3):211-216.
22. Sümer ZH. Gender, religiosity, sexual activity, sexual knowledge, and attitudes toward controversial aspects of sexuality. J Relig Health. 2015;54(6):2033-2044.
23. Eşsizoğlu A, Yasan A, Yildirim EA, Gurgen F, Ozkan M. Double standard for traditional value of virginity and premarital sexuality in Turkey: a university students case. Women Health. 2011;51(2):136-150.
24. Sönmez T, Koç Ö, Alver E. The effect of genital self-image on sexual satisfaction and stress in women after vaginal delivery. Rev Assoc Med Bras (1992). 2024;70(10):e20240692.

Original Article

DOI: 10.4274/anajog.galenos.2026.36855

Anat J Obstet Gynecol Res 2026;3(2):98-104

Wedge Versus Linear Labiaplasty: Comparative Effects on Sexual Function, Genital Self-Image, and Patient Satisfaction

Yasin Ceylan[1], Özlem Yüksel Aybek[2], Berkay Yüksel[3], Engin Yurtçu[4]

[1]Private Clinic, Clinic of Obstetrics and Gynecology, Muğla, Turkey
[2]University of Health Sciences Turkey, Bağcılar Training and Research Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Turkey
[3]Private Clinic, Clinic of Obstetrics and Gynecology, Antalya, Turkey
[4]Private Clinic, Clinic of Obstetrics and Gynecology, Düzce, Turkey

ABSTRACT

Purpose: Labiaplasty has become one of the most frequently performed procedures in female genital cosmetic surgery due to increasing aesthetic awareness, functional complaints, and concerns related to sexual well-being. Although wedge and linear (trim) resections are the most commonly used techniques, comparative data addressing their outcomes, particularly sexual function and genital self-image, remain limited. Thus, the aim was to compare clinical, psychosexual, and patient-reported outcomes following wedge or linear labiaplasty and to identify factors associated with postoperative satisfaction and complications.

Methods: This prospective comparative study included women who underwent labiaplasty at a tertiary referral center between 2025 and 2026. Patients were allocated to either the wedge or linear technique group according to surgeon preference and anatomical suitability of the labial morphology. Demographic characteristics, operative variables, and postoperative complications were recorded. Genital self-image and sexual function were assessed preoperatively and at the 6-month postoperative follow-up using validated instruments: the female genital self-image scale (FGSIS) and the female sexual function index (FSFI). Within-group and between-group comparisons were performed using appropriate parametric statistical tests.

Results: The cohort consisted of 94 women, split equally between the wedge (n=47) and linear (n=47) groups. Both techniques resulted in significant postoperative improvements in genital self-perception and sexual function. FGSIS and FSFI scores increased significantly in both groups (all $p<0.001$). Operative time was longer in the wedge group, whereas the linear technique was associated with shorter procedures but higher rates of aesthetic dissatisfaction. Wound dehiscence occurred more frequently following wedge resection. No significant differences were observed between groups regarding postoperative total FGSIS or FSFI scores.

Conclusion: Both wedge and linear labiaplasty provided meaningful functional and psychosexual benefits. However, each technique was associated with distinct complication profiles and aesthetic trade-offs. Surgical success appeared to depend more on individualized patient selection and tailored operative planning than on the surgical technique itself. A patient-centered, anatomy-based approach is essential to optimize safety, satisfaction, and long-term outcomes.

Keywords: Labiaplasty, vulva, body image, sexual behavior, female genitalia

Address for Correspondence: Özlem Yüksel Aybek, MD, University of Health Sciences Turkey, Bağcılar Training and Research Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Turkey
E-mail: yukselozlem89@gmail.com **ORCID ID:** orcid.org/0000-0003-0332-496X
Received: 23.05.2026 **Accepted:** 07.07.2026 **Publication Date:** 02.09.2026

Cite this article as: Ceylan Y, Yüksel Aybek Ö, Yüksel B, Yurtçu E. Wedge versus linear labiaplasty: comparative effects on sexual function, genital self-image, and patient satisfaction. Anat J Obstet Gynecol Res. 2026;3(2):98-104

INTRODUCTION

Female genital cosmetic surgery has become increasingly prevalent over the past two decades, paralleling the growing awareness of genital aesthetics, the desire to address functional complaints, and the heightened emphasis on sexual well-being. Among these procedures, labiaplasty represents one of the most frequently performed female genital cosmetic surgical interventions worldwide. Women commonly seek labiaplasty due to labia minora hypertrophy, asymmetry, recurrent vulvar or labial infections, discomfort during daily activities or sexual intercourse, and dissatisfaction with genital appearance.[1]

Previous studies have demonstrated that dissatisfaction with genital appearance may negatively influence body image, sexual self-confidence, and overall sexual satisfaction. Although numerous surgical techniques for labiaplasty have been described, wedge resection and linear (trim) resection remain the most commonly utilized approaches.[2-4]

The linear technique is technically simpler and aims to reduce postoperative complication rates but it may disrupt the natural contour and edge of the labia.[5] In contrast, the wedge technique preserves the natural labial margin, pigmentation, and anatomical contour, thereby potentially providing superior aesthetic and functional outcomes. Nevertheless, wedge resection has been associated with specific complications, particularly wound dehiscence. This method involves excision of a wedge-shaped segment of labial tissue, allowing restoration of a more physiological and natural contour compared with straight-line resections.[4]

Despite the widespread use of both techniques, studies evaluating their effects on sexual function and genital self-image using validated assessment tools remain limited. The current literature stresses that genital aesthetic surgery should not be evaluated solely on anatomical outcomes but rather within a multidimensional framework of sexual health.[3] This recommendation highlights the importance of assessing a combination of individual patient sexual well-being, psychosocial changes and relational outcomes following genital surgery. In line with this, current research has demonstrated a strong correlation between genital self-image and sexual function, showing that post-operative improvements in aesthetic satisfaction significantly enhance various dimensions of sexual health, including desire and arousal.[6]

Furthermore, concomitant clinical and individual factors have been reported to adversely influence postoperative outcomes, highlighting appropriate patient selection as a critical determinant of surgical success and reduced complication rates.[7] Therefore, comprehensive preoperative counseling should include detailed discussions of potential complications, anticipated functional and aesthetic benefits, and normal anatomical variations, with the aim of establishing realistic patient expectations.[8] Careful planning and individualization of the surgical technique are also essential to minimize perioperative and postoperative complications, optimize postoperative sexual function, and enhance overall patient satisfaction.

Given this background, the aim of the present study was to comparatively evaluate postoperative genital self-image, sexual function, psychosexual outcomes, and complication profiles following wedge or linear labiaplasty. A further aim was to identify clinical and demographic factors associated with surgical outcomes.

METHODS

This prospective comparative study included women who underwent labiaplasty at a single tertiary referral center between 2025 and 2026. The study protocol was approved by the İstanbul Medipol University Non-Interventional Clinical Research Ethics Committee (approval number: 1131, date: 11.09.2025). All procedures were conducted in accordance with the principles of the Declaration of Helsinki. Written informed consent was obtained from all participants prior to enrollment.

Patients were allocated to either the wedge or linear technique group according to surgeon preference and anatomical suitability of the labial morphology. Therefore, this study should be considered a prospective comparative cohort study rather than a randomized controlled trial. All procedures were performed by two surgeons with extensive experience in genital aesthetic surgery; one surgeon had more than 5 years of experience and had performed over 100 labiaplasty procedures, while the other had more than 10 years of experience and had performed over 200 labiaplasty procedures. Women undergoing labiaplasty for functional and/or aesthetic indications were eligible for inclusion. Patients with incomplete clinical records or missing preoperative or postoperative questionnaire data were excluded from the analysis.

Demographic and baseline clinical characteristics were systematically recorded, including age, gravida, parity, smoking status, and the presence of systemic comorbidities. Obstetric history and mode of delivery were documented. Preoperative genital examination included evaluation of labial anatomy, classification of labial morphology, and assessment of asymmetry. Surgical variables comprised operation duration, primary indication for surgery (functional complaints, aesthetic concerns, or hypertrophy), concomitant procedures such as hoodoplasty or labia majora interventions, detailed description of the surgical technique, suture material, and energy modality used.

Postoperative outcomes were assessed during scheduled follow-up visits. Representative pre- and postoperative (first month) clinical photographs of four patients are presented in Figures 1 and 2. Complications were defined and recorded as surgical site infection, wound dehiscence, hematoma, scar contracture, and patient-reported dissatisfaction. Patient-reported outcome measures were used to evaluate both genital self-perception and sexual function. Female genital self-image was assessed using the female genital self-image scale (FGSIS),[9,10] a validated questionnaire measuring women's perceptions of genital appearance, comfort, and confidence, with higher scores reflecting a more positive self-image. Sexual function was evaluated using the female

sexual function index (FSFI),[11,12] a validated multidimensional instrument assessing six domains: desire; arousal; lubrication; orgasm; satisfaction; and pain. Both domain-specific and total scores are calculated. Identical validated versions of both instruments were administered preoperatively and at the sixth month postoperatively to ensure consistency and comparability of outcomes; FGSIS and FSFI scores were evaluated and compared during these time periods. No patients were lost to follow-up, and all participants completed the 6-month postoperative assessment.

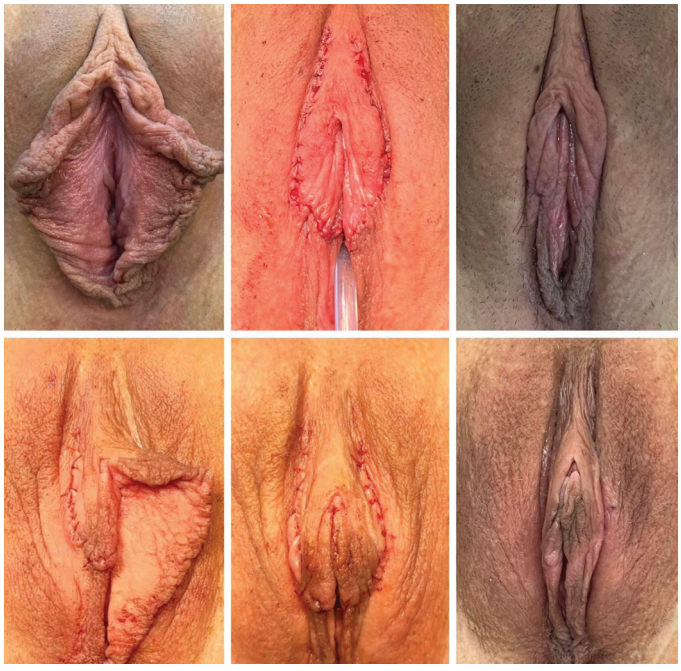


Figure 1. Preoperative, Intraoperative and postoperative (first month) photographs of wedge labiaplasty patients

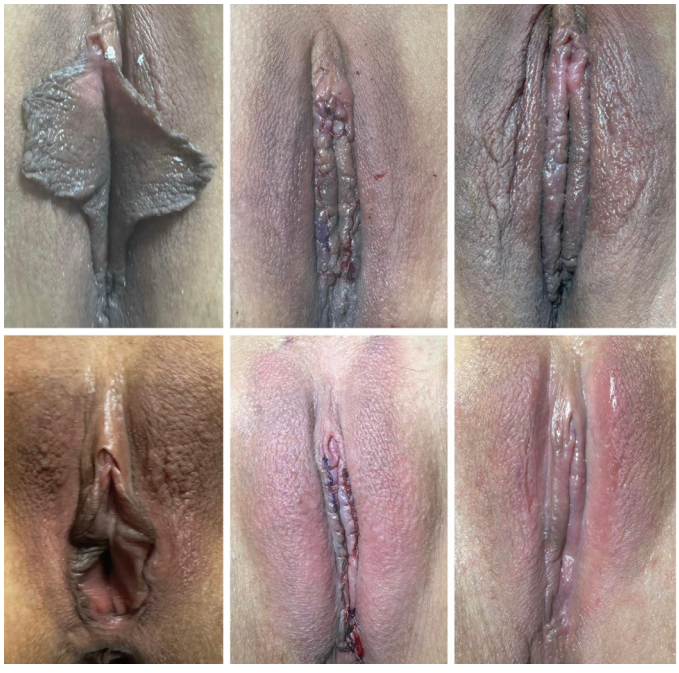


Figure 2. Preoperative, Intraoperative and postoperative (first month) photographs of linear labiaplasty patients

Statistical Analysis

Sexual function and genital self-perception were evaluated using validated, standardized patient-reported outcome instruments. The FSFI is a 19-item multidimensional questionnaire assessing six domains of female sexual function: desire, arousal, lubrication, orgasm, satisfaction, and pain. Domain scores are weighted and summed to generate a total score ranging from 2 to 36, with higher scores indicating better sexual function. Consistent with previously validated thresholds, a total FSFI score below 26.55 was considered indicative of sexual dysfunction.

Genital self-image was assessed using the FGSIS, a validated 7-item instrument designed to evaluate women's perceptions, comfort, and satisfaction regarding their genital appearance. Each item is rated on a 4-point Likert scale ranging from strongly agree to strongly disagree, yielding a total score between 7 and 28. Higher scores reflect a more positive genital self-image.

Sample size calculation was performed using G*Power software (version 3.1, Heinrich Heine University, Düsseldorf, Germany). Based on an expected large effect size (Cohen's d=0.8), a statistical power of 95%, and a significance level of 0.05, the required total sample size was calculated as 84 patients (42 per group).

All statistical analyses were performed using IBM SPSS Statistics version 25 (IBM Corp., Armonk, NY, USA). Categorical variables are summarized as frequencies and percentages, whereas continuous variables are presented as means with standard deviations. Changes between preoperative and postoperative FSFI and FGSIS scores were evaluated using paired-samples t-tests. All statistical tests were two-tailed, and a p value of <0.05 was considered statistically significant.

RESULTS

A total of 94 women were included with 47 in both the wedge and linear groups, thus ensuring adequate statistical power. The demographic and clinical characteristics of the study population are summarized in Table 1. The mean age was 39.3±5.6 years in the wedge labiaplasty group and 34.7±7.7 years in the linear labiaplasty group. The prevalence of smoking was 31.9% in the wedge group and 36.2% in the linear group. According to the labia minora morphological classification, type 1 morphology was bilaterally predominant in the wedge group, whereas type 2 morphology was the most frequently observed pattern in the linear group. Labial asymmetry was identified in 46.8% of patients undergoing wedge labiaplasty and 38.3% in those undergoing the linear procedure (Table 1).

Comparison of surgical characteristics demonstrated that the mean operative time was longer in the wedge group compared to the linear group (94.6±26.9 minutes vs. 80.7±31.7 minutes). With respect to surgical indications, aesthetic concerns constituted the primary motivation in the wedge resection group (80.8%), whereas labial hypertrophy was the most commonly reported indication in the linear resection group (85.1%). Among patients undergoing concomitant labia majoraplasty, hyaluronic acid injection was the most

frequently preferred adjunctive treatment in the wedge group (85.1%), while lipid-based filler application predominated in the linear labiaplasty group (56.7%). Lateral hoodoplasty was the most commonly performed hoodoplasty technique in both groups. Regarding surgical instrumentation, wedge resections were typically performed using scissors or a scalpel, whereas linear resections were most commonly carried out using electrocautery.

An analysis of the surgical complications reveals that the most frequent issue in the wedge labiaplasty group was dehiscence (wound separation), occurring in 10.6% of cases (n=5). In addition, this group experienced infection, hematoma, and patient dissatisfaction, each at a rate of 4.2% (n=2). In contrast, the linear labiaplasty group reported no instances of dehiscence or infection. However, complications in the linear group included contracture (2.1%, n=1), hematoma (2.1%, n=1), and dissatisfaction (4.2%, n=2), along with a 4.2% (n=2) rate of complications related to labia majoraplasty, such as edema, infection, or hematoma. Despite these numerical variations, statistical analysis indicated that there were no significant differences between the two techniques regarding infection ($p=0.5$), dehiscence ($p=0.1$), hematoma ($p=0.6$), or patient dissatisfaction ($p=0.9$) (Table 2).

Functional and psychosexual outcomes demonstrated significant postoperative improvements in both cohorts. Total scores of the FGSIS increased significantly from 10.9±1.9 preoperatively to 25.6±1.3 postoperatively in the Wedge group and from 11.5±1.8 to 25.7±2.1 in the linear group (both $p<0.001$). Similarly, the FSFI total scores showed marked improvement, rising from 14.9±3.1 to 28.3±3.0 in the wedge group and from 12.3±2.4 to 27.4±4.1 in the linear group (both $p<0.001$).

Analysis of FSFI domain scores revealed statistically significant improvements across all subdomains in both groups when preoperative and postoperative values were compared ($p<0.05$) (Table 3). Between-group comparisons demonstrated no significant differences in either preoperative or postoperative FGSIS total scores. However, preoperative FSFI total scores differed significantly between the groups. At the domain level, significant differences were observed in all preoperative parameters except satisfaction and pain, whereas postoperatively only the satisfaction domain showed a statistically significant difference between the groups (Table 3).

Additional analyses of postoperative change scores demonstrated that the mean ΔFGSIS scores were comparable between the wedge and linear groups (14.7±2.1 vs. 14.2±2.5, respectively; $p=0.30$). In contrast, the improvement in sexual function, as assessed by the ΔFSFI total score, was significantly greater in the linear group than in the wedge group (15.1±4.4 vs. 13.3±3.7, respectively; $p=0.03$).

Overall analyses further demonstrated that patients with a previous cesarean delivery achieved significantly greater improvements in FSFI total scores than those with a history of vaginal delivery (ΔFSFI: 15.1±4.3 vs. 12.9±3.6, $p=0.007$), whereas ΔFGSIS scores did not differ significantly according to mode of delivery (14.6±2.3 vs. 14.2±2.3, $p=0.43$).

Table 1. Demographic data of patients

		Wedge labiaplasty group (n=47)	Linear labiaplasty group (n=47)	*p*
		Mean ± SD	**Mean ± SD**	
Age		39.3±5.6	34.7±7.7	0.01*
Gravida		2.1±1.0	1.6±1.1	0.02*
Parity		1.9±0.9	1.4±1.0	0.02*
		(%)	**(%)**	
Mode of delivery	**Vaginal birth**	63.8	21.3	
	C-section	36.2	78.7	<0.001*
	None			
Smoking		31.9	36.2	0.38
Left labium minor classification	**Type 1**	53.1	29.8	
	Type 2	42.5	53.2	0.02*
	Type 3	4.2	17	
Right labium minor classification	**Type 1**	55.3	29.8	
	Type 2	38.2	57.4	0.02*
	Type 3	7.5	12.8	
Presence of asymmetry	**Present/ observed**	46.8	38.3	0.07
Hypertension		4.2	0	0.26
Obesity		4.2	0	0.26

*Chi-square test, statistically significant $p<0.05$
SD: Standard deviation

Table 2. Surgical characteristics of the groups

		Wedge labiaplasty group (n=47)		Linear labiaplasty group (n=47)		*p*
		Mean ± SD	(%)	Mean ± SD	(%)	
Operation duration (min)		94.6±26.9		80.7±31.7		0.02
Reason for operation	**Hypertrophy**	36	76.5	40	85.1	0.9
	Functional	26	55.3	30	63.8	
	Aesthetic concer	38	80.8	36	76.5	
L. majoraplasty technique	**Hyaluronic acid**	40	85.1	12	26.7	
	Lipid filler	6	12.7	27	56.7	<0.001
	Surgery	1	2.1	8	16.7	
Hoodoplasty technique	**Lateral hoodoplasty**	45	95.7	40	83.3	
	Central hoodoplasty	2	4.3	7	16.7	0.11
Suture material	**4.0**	5	10.7	21	44.6	
	5.0	42	89.3	26	55.4	<0.001
Energy modality used	**Cautery**	5	10.7	39	82.9	
	Cold	42	89.3	8	17.1	<0.001
Complications observed in labium minoraplasty	**Infection**	2	4.2	0	0	0.5
	Dehiscence	5	10.6	0	0	0.1
	Contracture	0	0	1	2.1	0.9
	Hematoma	2	4.2	1	2.1	0.6
	Dissatisfaction	2	4.2	2	4.2	0.9
Complications observed in L. majoraplasty	**Edema, infection, hematoma**	0	0	2	4.2	0.8

SD: Standard deviation

Table 3. Evaluation of FGSIS and FSFI scores before and after surgery

		Wedge labiaplasty group (n=47)				Linear labiaplasty group (n=47)			
		Before surgery mean ± SD	**After surgery mean ± SD**	**Within group *p****	**Between groups before surgery *p*****	**Before surgery mean ± SD**	**After surgery mean ± SD**	**Within group *p****	**Between groups after surgery *p*****
FGSIS total score		10.9±1.9	25.6±1.3	<0.001*	0.14	11.5±1.8	25.7±2.1	<0.001*	0.74
FSFI	**Desire**	2.2±0.6	4.7±0.8	<0.001*	0.02**	1.9±0.5	4.3±1.3	<0.001*	0.06
	Arousal	2.2±0.5	4.7±0.7	<0.001*	<0.001**	1.7±0.6	4.4±1.1	<0.001*	0.1
	Lubrication	2.3±0.7	4.4±0.8	<0.001*	0.005**	1.8±0.9	4.5±0.9	<0.001*	0.7
	Orgasm	2.1±0.9	4.6±0.8	<0.001*	<0.001**	1.9±0.6	4.8±0.8	<0.001*	0.2
	Satisfaction	2.4±0.9	5.1±0.6	<0.001*	0.2	2.1±1	4.5±0.9	<0.001*	<0.01**
	Pain	3.6±1.2	4.6±0.7	<0.001*	0.9	3.7±2.4	4.4±1.4	0.04*	0.5
	FSFI total score	14.9±3.1	28.3±3	<0.001*	<0.001**	12.3±2.4	27.4±4.1	<0.001*	0.2

*Paired samples t test, statistically significant ($p<0.05$), **Independent Samples t test, statistically significant ($p<0.05$)
FGSIS: Female genital self-image scale, FSFI: Female sexual function index

Similarly, patients who underwent cold-cutting techniques demonstrated significantly greater improvements in sexual function than those treated with electrocautery (ΔFSFI: 15.8±4.2 vs. 12.7±3.5, $p<0.001$). However, changes in genital self-image were comparable between the two energy modalities (ΔFGSIS: 14.7±2.5 vs. 14.0±2.7, $p=0.10$). Furthermore, one-way ANOVA demonstrated that suture caliber was significantly associated with postoperative improvements in both FGSIS ($p=0.02$) and FSFI ($p=0.001$), with thinner suture materials being associated with greater score improvements.

To identify independent predictors of postoperative improvement, a forward stepwise linear regression analysis was performed including mode of delivery, suture size, energy modality (cold cutting vs. electrocautery), concomitant

hoodoplasty, labia majoraplasty, and labia minora surgical technique. For ΔFSFI total score, the use of thinner suture material (β=0.27, p=0.01) and cold-cutting technique (β=0.22, p=0.04) emerged as independent predictors of greater postoperative improvement. In contrast, for ΔFGSIS, thinner suture material was the only independent predictor of greater improvement (β=0.24, p=0.01).

At baseline, all patients had total FSFI scores below the validated cutoff value of 27.2, indicating female sexual dysfunction. Postoperatively, a substantial improvement was observed, with only 14 patients (29.7%) in the wedge group and 17 patients (36.1%) in the Linear group remaining below this threshold. Consequently, 33 patients (70.3%) in the wedge group and 30 patients (63.9%) in the linear group achieved FSFI scores consistent with normal sexual function following surgery.

Similarly, preoperative genital self-image was generally poor in both groups. In each group, 45 patients (95.7%) had FGSIS scores between 7 and 14 points, while the remaining two patients (4.3%) had scores between 15 and 21 points. Following surgery, a marked shift toward higher FGSIS categories was observed. No patients remained within the lower score ranges, and all patients (100%) achieved FGSIS scores between 22 and 28 points, indicating a substantial improvement in genital self-image.

DISCUSSION

In recent years, increasing aesthetic awareness and heightened functional expectations regarding female genital appearance have led to a marked rise in the demand for female genital cosmetic procedures, with labiaplasty emerging as one of the most frequently performed operations worldwide. National databank statistics demonstrate a continuous growth in labiaplasty rates, underscoring its evolving role within modern aesthetic and reconstructive gynecology.[13]

Parallel to this growing demand, contemporary surgical practice has diversified substantially, with the development of multiple operative techniques aimed at addressing patient-specific anatomical characteristics, functional complaints, and aesthetic goals. Consequently, the selection of an appropriate surgical approach should not be regarded merely as a technical preference but rather as a multidimensional clinical decision-making process. Comprehensive preoperative evaluation, including anatomical variations, symptom severity, functional expectations, cosmetic concerns, safety considerations, and anticipated complication risks, is essential for optimizing both surgical success and patient satisfaction. Despite ongoing refinements in technique, no single procedure has yet demonstrated universal superiority across all patient populations. Therefore, current evidence supports an individualized, patient-centered strategy rather than a standardized approach.[6,14]

In the present study, both wedge and linear (trim) labiaplasty techniques resulted in significant improvements in genital self-image and sexual function. Substantial postoperative increases in total FGSIS and FSFI scores were observed in both subgroups, highlighting the positive psychosexual and quality-of-life benefits associated with labiaplasty, irrespective of the surgical method employed.

The wedge resection technique remains one of the most widely used approaches in clinical practice and has been consistently associated with favorable aesthetic outcomes. Its principal advantage lies in the precise control of excision depth and margins, allowing individualized tissue reduction while preserving the natural labial edge, pigmentation gradients, and contour. By maintaining anatomical continuity, this technique has been reported to result in a more physiologic postoperative appearance and less conspicuous scarring.[15] Consistent with these theoretical benefits, patients undergoing wedge resection in our cohort demonstrated marked postoperative improvements in both FGSIS and FSFI scores.

However, wedge resection is not without limitations. Previous studies have reported an increased risk of wound dehiscence, likely attributable to wider excision areas and elevated tension along newly created suture lines.[15] Similarly, in our series, wound dehiscence represented the most frequent complication in the wedge group, occasionally necessitating revision surgery. To mitigate these risks, modified wedge techniques that preserve central neurovascular structures have been introduced, with reports of reduced dehiscence rates and improved healing profiles. Such refinements suggest that careful preservation of vascular supply may play a critical role in minimizing postoperative complications.[16]

Conversely, the linear (trim) technique is frequently preferred due to its technical simplicity, short learning curve, and shorter operative time. By employing a straightforward longitudinal excision of redundant tissue, this method offers a predictable and efficient surgical workflow.[17] In accordance with these characteristics, operative times were significantly shorter in the linear resection group in our study. Nevertheless, because this approach involves direct removal of the native labial edge, it may compromise natural morphology and pigmentation, potentially resulting in less physiologic aesthetic outcomes. Postoperative dissatisfaction related to cosmetic appearance was the most commonly reported concern among patients undergoing linear resection in our cohort. In addition, scar contracture along the incision line represented a recurrent issue, occasionally contributing to local tightness or contour irregularities. Thus, although technically straightforward and reliable, the trim technique may present aesthetic trade-offs in selected patients.[18]

Study Limitations

This study has several limitations that should be acknowledged. Its single-center design and relatively limited sample size may restrict the generalizability of the findings. In addition, the six-month follow-up period may not fully capture long-term aesthetic and functional outcomes. As patients were allocated according to anatomical suitability and surgeon preference, some baseline differences between groups emerged (Table 1) and a degree of selection bias cannot be excluded. In particular, significant baseline differences in age and preoperative FSFI scores were present between the groups. Since adjusted analyses were not performed to

account for these imbalances, residual confounding may have influenced the comparative results and should be considered when interpreting the findings. Nevertheless, the prospective collection of data, the use of standardized surgical protocols performed by experienced genital aesthetic surgeons, and the assessment of outcomes using validated instruments strengthen the reliability of the present findings.

Taken together, both techniques provided significant functional and psychosexual benefits, yet each demonstrated distinct complication profiles and aesthetic considerations.[19] Furthermore, multivariable regression analysis demonstrated that technical aspects of the procedure, particularly the use of thinner suture material and cold-cutting rather than electrocautery, were the strongest independent predictors of greater postoperative improvement in sexual function. In contrast, thinner suture material was the only independent predictor of greater improvement in genital self-image. Importantly, no significant differences were observed between groups in postoperative FGSIS or FSFI total scores, suggesting that improvements in genital self-image and sexual function are influenced more by appropriate patient selection and individualized surgical planning than by the specific surgical technique itself.

CONCLUSION

Neither wedge nor linear labiaplasty can be considered universally superior in this cohort. Instead, optimal outcomes appeared to depend on tailoring the surgical approach to each patient's anatomy, expectations, and risk profile. An individualized, patient-centered strategy remains fundamental for maximizing both safety and aesthetic success. Future large-scale, prospective, and multicenter studies are warranted to further clarify technique-specific indications and to establish evidence-based guidelines for patient-specific surgical selection.

Ethics

Ethics Committee Approval: The study was approved by the İstanbul Medipol University Non-Interventional Clinical Research Ethics Committee (approval number: 1131, date: 11.09.2025).

Informed Consent: Written informed consent was obtained from all participants prior to enrollment.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Y.C., B.Y., Concept: Y.C., E.Y., Design: Y.C., E.Y., Data Collection or Processing: Y.C., B.Y., Analysis or Interpretation: Ö.Y.A., E.Y., Literature Search: Ö.Y.A., B.Y., Writing: Y.C., Ö.Y.A.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

Declaration on the Use of Artificial Intelligence (AI): No artificial intelligence tools were used in the preparation of this manuscript.

REFERENCES

1. Ortega-Sánchez I, Lucha-López MO, Monti-Ballano S. Motivational factors for labiaplasty: a systematic review of medical research. J Clin Med. 2025;14(8):2686.
2. Veale D, Eshkevari E, Ellison N, et al. Psychological characteristics and motivation of women seeking labiaplasty. Psychol Med. 2014;44(3):555-566.
3. Goodman MP. Female genital cosmetic and plastic surgery: a review. J Sex Med. 2011;8(6):1813-1825.
4. Alinsod R, Banwell PE, Berreni N, et al. Chapter 5 Labia Minora Reduction Surgery: Curved Linear Resection. Hamori CA, Banwell PE, Alinsod R, eds. Female Cosmetic Genital Surgery: Concepts, Classification and Techniques. Stuttgart: Thieme; 2017;59-74.
5. Wiśniewska-Ś Lepaczuk K, Pieczykolan A, Grzesik-Ga Sior J, Wdowiak A. A review of aesthetic gynecologic procedures for women. Plast Aesthet Nurs (Phila). 2022;42(4):226-237.
6. Ucar E, Bestel M, Ucar BH, Dogan O. The effect of technique selection in labiaplasty surgery: analysis of aesthetic and functional outcomes. J Clin Med. 2025;14(24):8923.
7. Köle E, Doğan O, Arslan G, Köle MÇ, Aslan E, Çalışkan E. Labiaplasty outcomes and complications in Turkish women: a multicentric study. Int Urogynecol J. 2024;35(5):1045-1050.
8. Azmoude E, Zagami SE, Hooshmand E, Taheri E, Shoorab NJ. Female genital cosmetic procedures from the perspective of health practitioners: a systematic review of knowledge, attitude and practice studies. BMC Womens Health. 2024;24(1):604.
9. Herbenick D, Reece M. Development and validation of the female genital self-image scale. J Sex Med. 2010;7(5):1822-1830.
10. Ellibes Kaya A, Yassa M, Dogan O, Basbug A, Pulatoglu C, Caliskan E. The female genital self-image scale (FGSIS): cross-cultural adaptation and validation of psychometric properties within a Turkish population. Int Urogynecol J. 2019;30(1):89-99.
11. Rosen R, Brown C, Heiman J, et al. The female sexual function index (FSFI): a multidimensional self-report instrument for the assessment of female sexual function. J Sex Marital Ther. 2000;26(2):191-208.
12. Oksuz E, Malhan S. Reliability and validity of the female sexual function index in Turkish population. SENDROM. 2005;17(7):54-60.
13. Aesthetic plastic surgery national databank statistics 2022. Aesthetic Surgery Journal. 2023;43(Suppl 2):1-19.
14. Géczi AM, Varga T, Vajna R, et al. Comprehensive assessment of labiaplasty techniques and tools, satisfaction rates, and risk factors: a systematic review and meta-analysis. Aesthet Surg J. 2024;44(11):NP798-NP808.
15. Sinnott CJ, Glickman LT, Natoli NB, Dobryansky M, Veeramachaneni HB. Outcomes, techniques, and risk factors for dehiscence in central wedge labiaplasty. Ann Plast Surg. 2020;85(S1 Suppl 1):S68-S75.
16. Ju M, Wang W, Ma N, Chen W. Reduction of hypertrophic labia minora by posterior-lateral wedge resection with preservation of the central blood vessels and nerve bundle. Aesthetic Plast Surg. 2019;43(3):742-749.
17. Lista F, Mistry BD, Singh Y, Ahmad J. The safety of aesthetic labiaplasty: a plastic surgery experience. Aesthet Surg J. 2015;35(6):689-695.
18. Propst K, Ferrando CA. Outcomes of labiaplasty in women with labial hypertrophy. Int Urogynecol J. 2021;32(5):1247-1251.
19. Lange M, Hage JJ, Karim RB, Amant F. An algorithm for labia minora reduction based on a review of anatomical, configurational, and individual considerations. Arch Plast Surg. 2023;50(1):17-25.

Original Article

DOI: 10.4274/anajog.galenos.2026.41736

Anat J Obstet Gynecol Res 2026;3(2):105-111

Effect of Preoperative Pelvic Floor Muscle Training on Transobturator Tape Outcomes in Women with Stress-Predominant Mixed Urinary Incontinence: A Retrospective Cohort Study

Özgür Aslan, Tuğba Demirtaş, Mülayim Tetik, Özlem Yüksel Aybek, Hakan Güraslan

University of Health Sciences Turkey, Bağcılar Training and Research Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Turkey

ABSTRACT

Purpose: Pelvic floor muscle training is the first-line conservative therapy for stress urinary incontinence, but its value as a preoperative adjunct to mid-urethral sling surgery remains uncertain. Our purpose was to determine whether structured six-week preoperative pelvic floor muscle training (PFMT) improved objective and subjective cure rates in women undergoing transobturator tape (TOT) surgery for stress-predominant mixed urinary incontinence (MUI).

Methods: In this single-center retrospective cohort study, we reviewed the records of women who underwent TOT for stress-predominant MUI between January 2020 and December 2024 and who attended an in-person follow-up with a cough stress test at 12 months or later. Patients were stratified by completion of a structured preoperative PFMT program of at least six weeks (PFMT group) versus proceeding directly to surgery (non-PFMT group). The primary outcome was objective cure, defined as the absence of leakage on a standardized cough stress test; the secondary outcome was subjective cure, defined as an improvement of at least 4 points on the international consultation on incontinence questionnaire-urinary incontinence short form (ICIQ-UI SF). Between-group and within-group comparisons were performed using non-parametric tests.

Results: Of 137 women, 82 completed preoperative PFMT and 55 proceeded directly to surgery; baseline characteristics were comparable. Objective cure occurred in 87.8% of the PFMT group versus 81.8% of the non-PFMT group ($p=0.468$), and subjective cure in 92.7% versus 85.5% ($p=0.280$); no between-group differences were significant. ICIQ-UI SF scores improved markedly in both groups, confirming effective surgery regardless of PFMT.

Conclusion: In this retrospective cohort, a six-week preoperative PFMT program was not associated with significantly higher objective or subjective cure rates after TOT in women with stress-predominant MUI. Given the high cure rate of surgery alone, this does not preclude a true benefit of PFMT. Whether such a benefit exists, particularly in women with greater baseline urgency, warrants confirmation in adequately powered prospective studies.

Keywords: Urinary incontinence stress, suburethral slings, pelvic floor, exercise therapy, treatment outcome

INTRODUCTION

Urinary incontinence is common among adult women. It is divided into three phenotypes: stress, urgency, and mixed incontinence. Stress urinary incontinence (SUI) refers to involuntary urine leakage during physical exertion, effort, sneezing, or coughing.[1,2] Recent data from the 2015-2018 National Health and Nutrition Examination Survey indicated that more than 60% of adult women experience some form of urinary incontinence, and over one in five report moderate to severe symptoms.[3] Among affected women, SUI is the most common subtype.[3] Despite this prevalence and its well-documented impact on quality of life, urinary incontinence remains underdiagnosed and undertreated worldwide.[3,4]

The therapeutic management of SUI follows a stepwise algorithm, prioritizing conservative modalities as first-line therapy, whereas surgical intervention is indicated for cases refractory to conservative management.[5,6] Pelvic floor muscle

Address for Correspondence: Özgür Aslan, MD, University of Health Sciences Turkey, Bağcılar Training and Research Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Turkey

E-mail: ozgraslan92@gmail.com **ORCID ID:** orcid.org/0000-0002-7048-2100

Received: 01.06.2026 **Accepted:** 15.07.2026 **Publication Date:** 02.09.2026

Cite this article as: Cite this article as: Aslan Ö, Demirtaş T, Tetik M, Yüksel Aybek Ö, Güraslan H. Effect of preoperative pelvic floor muscle training on transobturator tape outcomes in women with stress-predominant mixed urinary incontinence: a retrospective cohort study. Anat J Obstet Gynecol Res. 2026;3(2):105-111

training (PFMT), commonly known as Kegel exercises, is the first-line conservative therapy for women with SUI.[5] A recent Cochrane systematic review of 31 trials concluded that women with SUI who undergo PFMT are approximately eight times more likely to report cure than untreated controls. Symptoms and quality of life also improved significantly.[5] When surgical intervention is required, the mid-urethral sling (MUS) is the standard procedure. The two principal approaches are the retropubic and transobturator routes.[6,7]

Short-term subjective cure rates with MUS range between 62% and 98%, and the safety profile remains favorable.[6] Surgical success, however, is not uniform. Some patients still face persistent or recurrent symptoms afterward. In a large observational cohort, older age, elevated body mass index (BMI), and previous incontinence surgery were independent predictors of failure following transobturator sling placement.[8] Recurrence is also frequent over the long term. In a retrospective analysis, more than one third of women who were initially cured after transobturator tape (TOT) surgery developed recurrent incontinence within three years.[9] These findings have prompted investigation into whether perioperative interventions, particularly PFMT, may improve surgical outcomes.

The value of PFMT as an adjunct to MUS surgery remains controversial. Randomized and observational studies have diverged on both subjective and objective outcomes.[10,11] In a multicenter trial of women with mixed urinary incontinence (MUI), adding perioperative behavioral and PFMT to the sling improved continence over surgery alone, though the difference was statistically rather than clinically significant.[12] The more detailed evidence is no more encouraging. A retrospective cohort and a randomized controlled trial each reported limited and inconsistent effects of PFMT on objective continence after surgery.[10,11]

A more recent randomized trial focused on the preoperative window. In women awaiting surgery, PFMT yielded significant improvements in pelvic floor muscle strength and reductions in incontinence severity, although postoperative outcomes were not evaluated.[13] Whether this benefit persists after surgery is unknown. No study has examined the question directly in women undergoing TOT MUS for stress-predominant MUI. To address this gap, we conducted a retrospective cohort study comparing women who completed preoperative PFMT with those who proceeded directly to surgery. We examined whether this preoperative intervention was associated with postoperative cough stress test results and international consultation on incontinence questionnaire-urinary incontinence short form (ICIQ-UI SF) scores compared to the direct to surgery sub-group.

METHODS

Study Design and Setting

This single-center, retrospective, cohort study was conducted in the urogynecology unit of a tertiary referral center. Medical records of women who underwent TOT surgery for stress-predominant MUI between January 2020 and December 2024 were retrospectively reviewed. The study protocol was approved by the University of Health Sciences Turkey, Bağcılar Training and Research Hospital Non-Interventional Clinical Research Ethics Committee (approval no: 2025/08/12/075, date: 07.08.2025), and the requirement for informed consent was waived owing to the retrospective design of the study. All procedures conformed to the principles of the Declaration of Helsinki.

Participants and Eligibility Criteria

Women were eligible for inclusion if they had undergone TOT as a first-line surgical treatment for stress-predominant MUI and had completed an in-person, follow-up assessment, including a cough stress test at 12 months or later after surgery. Stress predominance was determined based on patient-reported symptom prominence, with stress symptoms being more bothersome than urgency symptoms. Exclusion criteria comprised a history of prior incontinence or pelvic organ prolapse surgery, concomitant pelvic organ prolapse surgery, stage ≥2 cystocele, neurogenic bladder, interstitial cystitis, mesh-related complications requiring reoperation, and missing preoperative or postoperative ICIQ-UI SF data.

Women were classified according to whether they underwent the structured, supervised preoperative PFMT program before surgery; those who proceeded directly to surgery without this program constituted the non-PFMT group. The PFMT group (n=82) comprised women who completed a structured preoperative PFMT program for at least six weeks before surgery. The non-PFMT group (n=55) comprised women who declined preoperative PFMT, had contraindications to PFMT, or proceeded directly to surgery owing to clinical urgency or patient preference.

Data Collection and Baseline Variables

Baseline demographic and clinical data were extracted from medical records using a standardized data-capture form. Recorded characteristics included age, BMI, parity, menopausal status, presence of diabetes mellitus, and the preoperative ICIQ-UI SF score.

Surgical Procedure

All procedures were performed using a standardized transobturator MUS technique. Sling placement followed an outside-in or inside-out transobturator approach with tension-free positioning at the mid-urethra after vaginal dissection. After positioning the tape, it was adjusted to achieve the desired tension. The protective sheath was subsequently withdrawn, excess mesh was trimmed, and the incisions were closed using sutures.

Pelvic Floor Muscle Training Protocol

Preoperative PFMT was prescribed and supervised by the attending urogynecologists. The protocol consisted of an initial supervised instruction session, during which correct pelvic floor muscle activation was confirmed by digital vaginal palpation, followed by a daily structured home-exercise program of six weeks duration before surgery. Patients were provided with a standardized institutional written instruction

sheet and instructed to perform three exercise sets per day, each comprising slow-twitch contractions (6-10 seconds of contraction followed by an equivalent relaxation period, 8-12 repetitions) and fast-twitch contractions (1-second contraction and relaxation, 8-10 repetitions). Biofeedback and neuromuscular electrical stimulation were not used. Adherence to the prescribed regimen was assessed at the preoperative visit by the supervising urogynecologist and documented in the clinical record; only patients confirmed to have completed the program were included in the PFMT group. However, as the study was retrospective, PFMT completion was based on routine clinical documentation and was not independently verified; a validated exercise diary and quantitative compliance metrics were not available.

Postoperative Care and Follow-Up

Postoperative management followed institutional standard care. Patients were scheduled for in-person follow-up at 12 months or later postoperatively, at which the cough stress test was performed and the ICIQ-UI SF was administered again. The interval between surgery and the follow-up assessment was recorded for each patient.

Outcome Assessment

In accordance with contemporary recommendations for the assessment of SUI surgery outcomes, both objective and subjective endpoints were used.[14] The primary outcome was objective cure at the postoperative follow-up. Objective cure was defined as the absence of urine leakage on a standardized cough stress test performed at a bladder volume of 300 mL in the lithotomy position. The secondary outcome was subjective cure, assessed using the Turkish-validated ICIQ-UI SF,[15] administered preoperatively and at the postoperative follow-up. The ICIQ-UI SF is a widely used patient-reported instrument that captures the frequency and amount of urine leakage and its impact on quality of life, generating a total score from 0 to 21, with higher scores indicating greater symptom burden. Symptom change was expressed as ΔICIQ-UI SF (preoperative minus postoperative score), and subjective cure was defined a priori as ICIQ-UI SF ≥4, in accordance with the established minimum clinically important difference for this instrument.[16]

Statistical Analysis

Statistical analyses were performed using jamovi (version 2.6.19; The Jamovi project, Sydney, Australia). The normality of continuous variables was assessed with the Shapiro-Wilk test. As all continuous variables deviated from normal distribution, descriptive statistics are presented as median and [interquartile range (IQR): 25th-75th percentile]; categorical variables are presented as frequency (n) and percentage. Between-group comparisons of continuous variables used the Mann-Whitney U test. Categorical variables were compared using the chi-square test with continuity correction. Within-group pre- to postoperative changes in ICIQ-UI SF scores were assessed descriptively with the Wilcoxon signed-rank test. Treatment effects were summarized as odds ratios (ORs) and absolute risk differences with 95% confidence intervals (CIs). To address confounding arising from non-random treatment allocation, a propensity score for preoperative PFMT completion was estimated by logistic regression on age, BMI, parity, menopausal status, presence of diabetes mellitus, and baseline ICIQ-UI SF score, and covariate balance was assessed with standardized mean differences; the association between PFMT and each cure outcome was then re-estimated using Firth's penalized logistic regression adjusted for the propensity score, with a direct multivariable model and inverse-probability-of-treatment weighting as sensitivity analyses; full model estimates and covariate balance are provided in Supplementary Tables 1 and 2.[17] As the study was retrospective, no a priori sample-size calculation was performed; instead, a post-hoc minimal-detectable-effect analysis was conducted, which is a design-based precision analysis rather than a post-hoc observed-power calculation. Using a two-proportion comparison with the observed group sizes (82 versus 55) and the observed non-PFMT cure rate as the reference, at $\alpha=0.05$ (two-sided) and 80% power, the smallest detectable PFMT cure rates were approximately 96% (objective) and 98% (subjective); the study was therefore powered to detect only large between-group differences. All hypotheses were tested two-sided, and a p value <0.05 was considered statistically significant.

RESULTS

A total of 197 women were initially screened for eligibility. Sixty patients were excluded for one or more of the following reasons, with some meeting more than one exclusion criterion: concomitant pelvic organ prolapse surgery (n=43), stage ≥2 cystocele on Pelvic Organ Prolapse Quantification examination (stage 2, n=23; stage 3, n=17), missing preoperative or postoperative ICIQ-UI SF data (n=12), mesh-related complications requiring reoperation (n=7), and a history of prior incontinence or pelvic organ prolapse surgery (n=6). The final analytic cohort comprised 137 women, of whom 82 (59.9%) completed structured preoperative PFMT for at least six weeks before surgery (PFMT group) and 55 (40.1%) proceeded directly to surgery (non-PFMT group) (Figure 1). The median postoperative follow-up interval was 15.0 months (IQR: 14.0-16.0).

Baseline demographic and clinical characteristics of the study population, stratified by group, are summarized in Table 1. Age, BMI, and parity were similar between the PFMT and non-PFMT groups (all $p>0.05$). The distribution of menopausal status and diabetes mellitus did not differ significantly between groups (all $p>0.05$). Preoperative ICIQ-UI SF scores were comparable between groups [median 19.0 (IQR 18.0-20.0) vs. 19.0 (IQR 17.0-21.0); $p = 0.936$].

Postoperative outcomes are presented in Table 2. The primary outcome, objective cure on the postoperative cough stress test, was achieved in 72 of 82 patients (87.8%) in the PFMT group and 45 of 55 patients (81.8%) in the non-PFMT group, with no significant between-group difference ($p=0.468$). The secondary outcome, subjective cure, defined as a pre- to postoperative ICIQ-UI SF improvement ≥4, was observed in 76 of 82 patients (92.7%) in the PFMT group and 47 of 55 patients

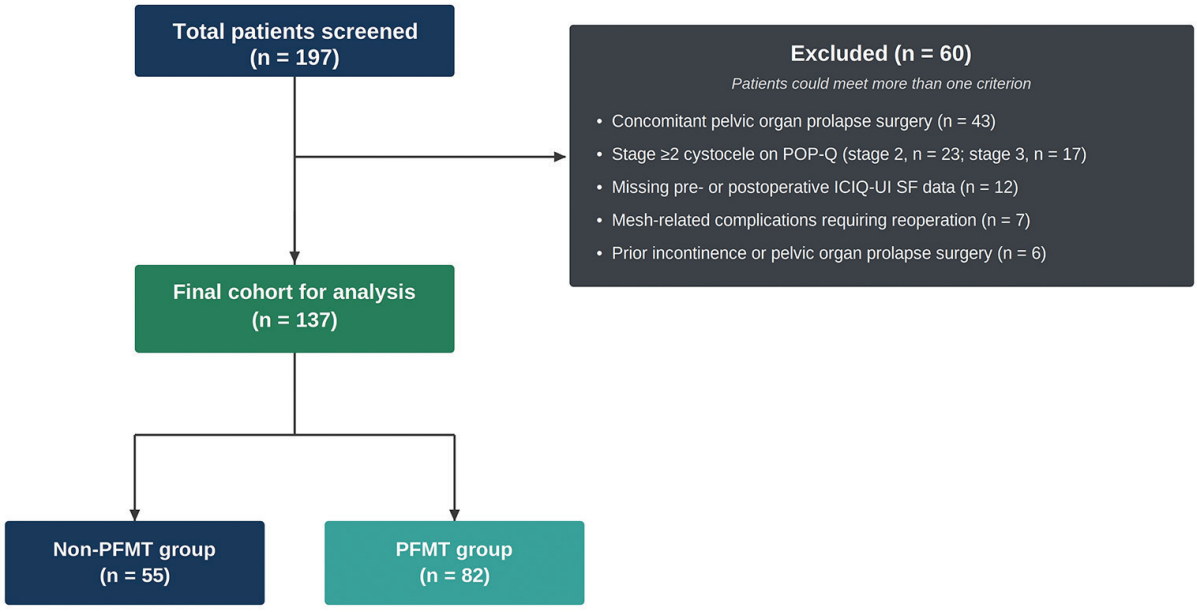


Figure 1. Flow diagram of patient selection through the study. Patients could meet more than one exclusion criterion

ICIQ-UI SF: International Consultation on Incontinence Questionnaire-Urinary Incontinence Short Form, PFMT: Pelvic floor muscle training, POP-Q: Pelvic Organ Prolapse Quantification

Table 1. Baseline demographic and clinical characteristics of the study cohort

Parameter	Non-PFMT group (n=55)	PFMT group (n=82)	*p* value
Age (years), median [IQR]	48.0 [44.5-57.5]	49.0 [45.0-53.8]	0.963
BMI (kg/m²), median [IQR]	29.0 [28.0-33.0]	29.5 [27.0-33.0]	0.897
Parity, median [IQR]	3.0 [2.0-4.0]	3.0 [2.0-4.0]	0.998
Menopausal status, n (%)	24 (43.6)	37 (45.1)	0.999
Diabetes mellitus, n (%)	18 (32.7)	20 (24.4)	0.382
Preoperative ICIQ-UI SF, median [IQR]	19.0 [17.0-21.0]	19.0 [18.0-20.0]	0.936
Follow-up (months), median [IQR]	15.0 [14.0-16.0]	15.0 [14.0-16.0]	0.768

Values are presented as median interquartile range, (IQR) for continuous variables and n (%) for categorical variables. Between-group comparisons were performed using the Mann-Whitney U test for continuous variables and the Pearson chi-square test or Fisher's exact test for categorical variables, as appropriate.
BMI: Body mass index, ICIQ-UI SF: International Consultation on Incontinence Questionnaire-Urinary Incontinence Short Form, PFMT: Pelvic floor muscle training

(85.5%) in the non-PFMT group ($p=0.280$). Postoperative ICIQ-UI SF scores were 5.5 (IQR 1.0-10.0) in the PFMT group and 8.0 (IQR 1.0-11.0) in the non-PFMT group ($p=0.473$). The median pre- to postoperative ICIQ-UI SF improvement was 13.0 points (IQR 8.0-17.0) in the PFMT group and 11.0 points (IQR 6.0-16.0) in the non-PFMT group ($p=0.304$). Baseline covariates were well balanced between the groups (all standardized mean differences below 0.10 except diabetes mellitus, 0.18, which remained below the 0.25 threshold for important imbalance). For objective cure the crude OR favoured PFMT but was not statistically significant (OR 1.60, 95% CI 0.62-4.15; absolute risk difference +6.0 percentage points, 95% CI -6.4 to +18.4) and was essentially unchanged after propensity-score adjustment (adjusted OR 1.63, 95% CI 0.63-4.21). Corresponding estimates for subjective cure were OR 2.16 (95% CI 0.70-6.60; risk difference +7.2 percentage points, 95% CI -3.7 to +18.1) and adjusted OR 2.22 (95% CI 0.75-6.89). The width of these intervals indicates limited precision.

As a descriptive check, ICIQ-UI SF scores fell markedly from baseline to follow-up within each group (both $p<0.0001$; Table 3), confirming that surgery was effective in both arms; these within-group comparisons are descriptive and do not address the study question, which is answered by the between-group analysis above. A floor effect was evident on the postoperative ICIQ-UI SF: 19.7% of patients (PFMT 17.1%, non-PFMT 23.6%) achieved the best possible score of zero.

Table 2. Comparison of postoperative outcomes between the PFMT and non-PFMT groups

Outcome measure	Non-PFMT group (n=55)	PFMT group (n=82)	*p* value	Crude OR (95% CI)	Adjusted OR (95% CI)
ICIQ-UI SF, postoperative, median [IQR]	8.0 [1.0-11.0]	5.5 [1.0-10.0]	0.473	-	-
ICIQ-UI SF, change score, median [IQR]	11.0 [6.0-16.0]	13.0 [8.0-17.0]	0.304	-	-
Objective cure (negative CST), n (%)	45 (81.8)	72 (87.8)	0.468	1.60 (0.62-4.15)	1.63 (0.63-4.21)
Subjective cure (ΔICIQ-UI SF ≥4), n (%)	47 (85.5)	76 (92.7)	0.280	2.16 (0.70-6.60)	2.22 (0.75-6.89)

Values are presented as median interquartile range (IQR) for continuous variables and n (%) for categorical variables. Between-group comparisons were performed using the Mann-Whitney U test for continuous variables and the chi-square test with continuity correction for categorical variables. Objective cure was defined as the absence of urine leakage on a standardized cough stress test performed at a bladder volume of 300 mL in the lithotomy position. Subjective cure was defined as a pre- to postoperative improvement of ≥4 points in the ICIQ-UI SF score, corresponding to the established minimum clinically important difference for this instrument
CST: Cough stress test, ICIQ-UI SF: International Consultation on Incontinence Questionnaire-Urinary Incontinence Short Form, ΔICIQ-UI SF: Pre- to postoperative ICIQ-UI SF score difference, OR: Odds ratio, CI: Confidence interval, adjusted OR: Propensity-score-adjusted firth penalized estimate, PFMT: Pelvic floor muscle training

Table 3. Within-group comparison of preoperative and postoperative ICIQ-UI SF scores

Group	Preoperative ICIQ-UI SF, median [IQR]	Postoperative ICIQ-UI SF, median [IQR]	*p* value
Non-PFMT group (n=55)	19.0 [17.0-21.0]	8.0 [1.0-11.0]	<0.0001
PFMT group (n=82)	19.0 [18.0-20.0]	5.5 [1.0-10.0]	<0.0001

Values are presented as median interquartile range, (IQR). Within-group comparisons were performed using the Wilcoxon signed-rank test
ICIQ-UI SF: International Consultation on Incontinence Questionnaire-Urinary Incontinence Short Form, PFMT: Pelvic floor muscle training

DISCUSSION

In this retrospective cohort study, the addition of a structured six-week preoperative PFMT program to TOT surgery was not associated with a significantly higher objective or subjective cure rate in women with stress-predominant MUI ($p=0.468$ and $p=0.280$, respectively). Although the absolute cure rates were numerically higher in the PFMT group for both outcomes, the between-group differences did not reach statistical significance. Symptom scores fell markedly from baseline in both groups, confirming that this absence of a between-group difference was not attributable to ineffective surgery in either arm.

Most randomized and observational studies that have asked this question have reached the same answer. ESTEEM, the largest trial in this area, found a statistically significant improvement with combined therapy that nevertheless did not reach the trial's pre-specified clinically important threshold.[12] McLean et al.[11] reported a higher cure rate when physiotherapy was added to surgery, but the advantage held only for patient-reported symptoms and disappeared on pad testing and bladder diary. Nauman et al.,[10] in a retrospective cohort closely matching ours in clinical profile, drew the same conclusion; preoperative training before retropubic sling did not meaningfully improve stress incontinence resolution (79.2% versus 69.4%). Üzelpasaci et al.[13] recently confirmed that an intensive six-week preoperative protocol can substantially improve pelvic floor muscle strength and symptom severity before surgery. Whether these preoperative gains carry into surgical outcomes remains an open question, since that study did not include postoperative follow-up.[13] A secondary analysis of the ESTEEM cohort offered a plausible explanation for the broader pattern, which was that most treatment failures after a MUS involved persistent urgency rather than recurrent stress incontinence, and further stress incontinence treatment was needed in only a small minority of women.[18] Once a MUS has effectively addressed the stress component, adjunctive PFMT appears to have limited room to add measurable benefit in stress-predominant populations like ours.

PFMT and a MUS treat the same continence failure by different mechanisms. PFMT works gradually, building levator plate bulk, increasing resting muscle tone, and training a voluntary pre-contraction that occludes the urethra before exertion, whereas a sling provides passive mechanical support at the mid-urethral level from day one.[5] Muscle adaptations take time. Most randomised trials showing clinical benefit used supervised protocols of eight to twelve weeks; a six-week preoperative course may not be long enough for full strength gains to develop.[5,19] Even when a short preoperative course does produce measurable strength gains,[13] those gains add little once the sling has corrected the stress component. The remaining value of PFMT is probably in residual urgency. This matches the failure pattern seen after surgery; in a secondary analysis of the ESTEEM cohort, most women who needed further treatment after a MUS were treated for persistent urgency, not recurrent stress incontinence.[18]

MUS surgery offers high, durable cure rates for SUI.[6] Serati and colleagues documented this prospectively across multiple centres: at ten years after tension-free vaginal tape-obturator in women with pure SUI, objective and subjective cure rates remained at 92% and 97%, with no meaningful deterioration over follow-up.[20] In our cohort, even the non-PFMT group achieved an objective cure rate of 81.8%, leaving little head-room to a complete response. Consistent with this, the postoperative symptom measure showed a pronounced floor effect: 19.7% of patients reached the best possible ICIQ-UI SF

score of zero, exceeding the 15% threshold conventionally used to define such an effect,[21] so there was limited capacity to detect further improvement. The numerically higher but non-significant 6.0 percentage-point difference is therefore consistent with a ceiling effect of effective surgery rather than evidence that PFMT confers no benefit; alternative explanations are considered among the limitations below.

Lack of statistical significance in our cohort does not mean PFMT plays no role. Two further analyses of the ESTEEM cohort suggest its benefit is concentrated rather than evenly spread. A planned interaction analysis showed that the benefit varied with baseline urgency. In the highest Urogenital Distress Inventory (UDI)-irritative quartile, sling alone carried a substantially higher failure risk; combined therapy did not.[18] A parallel cost-effectiveness analysis reached the same conclusion. In the overall MUI population, combined treatment was not good value. It became cost-effective in women with more bothersome baseline urgency, and cost-saving in the most severe stratum from the societal perspective.[22] Our cohort was stress-predominant, so by definition it contains fewer women in this responsive subgroup. The non-significant difference we observed fits this picture. The implication for practice is selective rather than universal use. Structured preoperative PFMT is unlikely to help across the board, but may be worth offering to women whose baseline urgency burden is greatest. This interpretation rests on external evidence rather than on our own data. This was because a validated measure of baseline urgency burden was not available, and so we could not test the urgency-stratified hypothesis directly in this cohort. It is therefore best regarded as hypothesis-generating. Prospective studies should stratify by baseline urgency and be powered for the subgroup most likely to benefit.

The cohort reflects routine clinical practice, and all patients underwent the same surgical procedure (TOT), removing variability across sling types. The preoperative PFMT protocol was standardized across exposed patients, six weeks of structured supervised training. Outcomes were assessed both objectively (postoperative cough stress test) and through a validated patient-reported instrument (ICIQ-UI SF), and the findings from these two measures were consistent with each other.

Study Limitations

Several limitations qualify these findings. The study is subject to selection bias because treatment allocation was not randomized but was determined by patient preference and clinical judgement; the non-PFMT group included women who declined training, had contraindications, or proceeded directly to surgery. Although baseline characteristics, including symptom severity, were comparable and the treatment effect was unchanged after propensity-score adjustment, unmeasured factors such as motivation and treatment expectations could not be captured from the medical record, and residual confounding cannot be excluded. Baseline urgency burden was not quantified with a validated instrument (such as the UDI-irritative subscale or an overactive-bladder symptom score), and urgency incontinence episode counts, bladder diaries, and urodynamic parameters were unavailable. Consequently, the analysis could not be stratified by baseline urgency severity, and a subgroup benefit of PFMT concentrated in women with greater urgency burden may have gone undetected. Adherence to preoperative PFMT was ascertained from the preoperative clinical notes rather than quantified with a validated diary, and reliable measurement of PFMT adherence remains an unresolved methodological problem across the literature;[23] any residual misclassification of PFMT status would most likely be non-differential and would bias the treatment estimate toward the null, potentially masking a true benefit. Outcome assessment was not blinded. The CIs around the treatment estimates were wide and the sample was powered to detect only large between-group differences, so a type II error cannot be excluded. We reiterate that the non-significant result should not be read as evidence that PFMT confers no benefit, and limited power, residual confounding, and a possibly insufficient six-week training regime cannot be distinguished from a ceiling effect of effective surgery as explanations for it. The primary outcomes were captured at a single timepoint roughly 15 months after surgery, leaving the possibility of later recurrences undetected.

CONCLUSION

In this retrospective cohort of women undergoing TOT for stress-predominant MUI, a six-week structured preoperative PFMT program did not significantly improve either objective or subjective cure rates. This pattern is consistent with a ceiling effect, in which the high cure rate of surgery alone, together with a floor effect on the symptom measure, limits both the room for and the power to detect an incremental benefit and it does not establish that PFMT is without effect. Combined with the broader evidence, any incremental gain from preoperative PFMT, if present, may concentrate in women with the heaviest baseline urgency and because urgency was not formally measured, this is best regarded as hypothesis-generating and would be most directly tested by future, adequately powered, urgency-stratified prospective studies.

Supplementary Tables Link: https://d2v96fxpocvxx.cloudfront.net/bda9171a-fae8-4995-8276-2138323f1e16/content-images/8d77713d-181a-4bd3-8050-ceb9e4f346bd.pdf

Ethics

Ethics Committee Approval: The study was approved by the University of Health Sciences Turkey, Bağcılar Training and Research Hospital Non-Interventional Clinical Research Ethics Committee (approval no: 2025/08/12/075, date: 07.08.2025).

Informed Consent: Owing to the retrospective design of the study, the requirement for informed consent was waived by the ethics committee.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Ö.A., T.D., M.T., Ö.Y.A., H.G., Concept: Ö.A., H.G., Design: Ö.A., H.G., Data Collection or Processing: Ö.A., M.T., Ö.Y.A., Analysis or Interpretation: Ö.A., Ö.Y.A., Literature Search: T.D., M.T., Writing: Ö.A., H.G.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

Declaration on the Use of Artificial Intelligence (AI): No artificial intelligence tools were used in the preparation of this manuscript.

REFERENCES

1. Haylen BT, de Ridder D, Freeman RM, et al. An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for female pelvic floor dysfunction. Int Urogynecol J. 2010;21(1):5-26.
2. Abrams P, Cardozo L, Fall M, et al.; Standardisation Sub-committee of the International Continence Society. The standardisation of terminology of lower urinary tract function: report from the Standardisation Sub-committee of the International Continence Society. Neurourol Urodyn. 2002;21(2):167-178.
3. Patel UJ, Godecker AL, Giles DL, Brown HW. Updated prevalence of urinary incontinence in women: 2015-2018 national population-based survey data. Female Pelvic Med Reconstr Surg. 2022;28(4):181-187.
4. Luber KM. The definition, prevalence, and risk factors for stress urinary incontinence. Rev Urol. 2004;6(Suppl 3):S3-S9.
5. Dumoulin C, Cacciari LP, Hay-Smith EJC. Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women. Cochrane Database Syst Rev. 2018;10(10):CD005654.
6. Ford AA, Rogerson L, Cody JD, Ogah J. Mid-urethral sling operations for stress urinary incontinence in women. Cochrane Database Syst Rev. 2015;(7):CD006375.
7. Kobashi KC, Albo ME, Dmochowski RR, et al. Surgical treatment of female stress urinary incontinence: AUA/SUFU guideline. J Urol. 2017;198(4):875-883.
8. Sabadell J, Montero-Armengol A, Salicrú S, Catalán-Martínez M, Gil-Moreno A, Poza JL. Long-term outcomes of transobturator suburethral tapes for female stress urinary incontinence. Neurourol Urodyn. 2022;41(1):391-398.
9. Jun T, Yoon HS, Kim HS, Lee JW, Bae J, Lee HW. Recurrence rate of stress urinary incontinence in females with initial cure after transobturator tape procedure at 3-year follow-up. Investig Clin Urol. 2017;58(1):54-60.
10. Nauman K, Stolzle A, Owens L, Frilot CF 2nd, Gomelsky A. Impact of preoperative pelvic floor muscle training (pretraining) on urinary storage and emptying symptoms in women undergoing sling surgery. Int Urogynecol J. 2022;33(3):673-679.
11. McLean L, Charette M, Varette K, et al. Pelvic floor muscle training as an adjunct to a midurethral sling: a single-blind randomised controlled trial. Int Urogynecol J. 2022;33(4):809-819.
12. Sung VW, Borello-France D, Newman DK, et al. Effect of behavioral and pelvic floor muscle therapy combined with surgery vs surgery alone on incontinence symptoms among women with mixed urinary incontinence: the ESTEEM randomized clinical trial. JAMA. 2019;322(11):1066-1076.
13. Üzelpasaci E, Özkutlu Ö, Demir E, Dülger A, Ün Yildirim N, Özdemir Ö. Effectiveness of intensive preoperative pelvic floor muscle training in women with stress urinary incontinence awaiting surgery: a randomized controlled trial. Int Urogynecol J. 2026;37(6):1743-1754.
14. Carmel ME, Deng DY, Greenwell TJ, Zimmern PE. Definition of success after surgery for female stress incontinence or voiding dysfunction: an attempt at standardization. Eur Urol Focus. 2016;2(3):231-237.
15. Çetinel B, Özkan B, Can G. The validation study of ICIQ-SF Turkish version. Turk J Urol. 2004;30(3):332-338.
16. Lim R, Liong ML, Lim KK, Leong WS, Yuen KH. The minimum clinically important difference of the International Consultation on Incontinence Questionnaires (ICIQ-UI SF and ICIQ-LUTSqol). Urology. 2019;133:91-95.
17. Heinze G, Schemper M. A solution to the problem of separation in logistic regression. Stat Med. 2002;21(16):2409-2419.
18. Sung VW, Richter HE, Moalli P, et al.; Eunice Kennedy Shriver National Institute of Child Health and Human Development Pelvic Floor Disorders Network*. Characteristics associated with treatment failure 1 year after midurethral sling in women with mixed urinary incontinence. Obstet Gynecol. 2020;136(3):482-491. Retraction in: Obstet Gynecol. 2021;138(2):312.
19. Bø K, Hilde G. Does it work in the long term?—A systematic review on pelvic floor muscle training for female stress urinary incontinence. Neurourol Urodyn. 2013;32(3):215-223.
20. Serati M, Braga A, Athanasiou S, et al. Tension-free vaginal tape-obturator for treatment of pure urodynamic stress urinary incontinence: efficacy and adverse effects at 10-year follow-up. Eur Urol. 2017;71(4):674-679.
21. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34-42.
22. Harvie HS, Sung VW, Neuwahl SJ, et al.; Eunice Kennedy Shriver National Institute of Child Health and Human Development Pelvic Floor Disorders Network. Cost-effectiveness of behavioral and pelvic floor muscle therapy combined with midurethral sling surgery vs surgery alone among women with mixed urinary incontinence: results of the effects of surgical treatment enhanced with exercise for mixed urinary incontinence randomized trial. Am J Obstet Gynecol. 2021;225(6):651.e1-651.e26.
23. Sacomori C, Berghmans B, Mesters I, de Bie R, Cardoso FL. Strategies to enhance self-efficacy and adherence to home-based pelvic floor muscle exercises did not improve adherence in women with urinary incontinence: a randomised trial. J Physiother. 2015;61(4):190-198.

Original Article

DOI: 10.4274/anajog.galenos.2026.70783

Anat J Obstet Gynecol Res 2026;3(2):112-117

Adjuvant Nonavalent HPV Vaccination is Associated with Improved One-Year HPV Clearance and Cervical Cytological Outcomes in Women with Histopathologically Confirmed LSIL/CIN1: A Retrospective Comparative Cohort Study

Mesut Ali Halisçelik[1], Süleyman Cemil Oğlak[2], Erdal Şeker[3], Sevda Yeleç[2], Mustafa Fırat Aydın[2], Salih Burçin Kavak[4]

[1]University of Health Sciences Turkey, Diyarbakır Gazi Yaşargil Research and Training Hospital, Clinic of Obstetrics and Gynecology, Division of Gynecologic Oncology, Diyarbakır, Turkey
[2]University of Health Sciences Turkey, Diyarbakır Gazi Yaşargil Research and Training Hospital, Clinic of Obstetrics and Gynecology, Diyarbakır, Turkey
[3]Private Clinic, Clinic of Obstetrics and Gynecology, Ankara, Turkey
[4]Fırat University Faculty of Medicine, Department of Obstetrics and Gynecology, Elazığ, Turkey

ABSTRACT

Purpose: To compare one-year human papilloma virus (HPV) deoxyribonucleic acid and cervical cytology outcomes between women with histopathologically confirmed low-grade squamous intraepithelial lesions/cervical intraepithelial neoplasia grade 1 (LSIL/CIN1) who received adjuvant nonavalent HPV vaccination and a contemporaneous unvaccinated control cohort, and to identify baseline predictors of persistent disease.

Methods: This single-center retrospective comparative cohort study included women with histopathologically confirmed LSIL/CIN1, divided into vaccinated and unvaccinated (control) subgroups, who were treated between 2020 and 2025. The primary outcomes were one-year HPV clearance and cytological normalization. Complete regression was defined as concurrent HPV negativity and normal cervical cytology, whereas persistent disease was defined as persistent HPV positivity together with abnormal cervical cytology. Univariable logistic regression analysis was performed to identify predictors of persistent disease in the vaccinated cohort.

Results: Of the study cohort of 233 women, 153 (65.7%) women completed the three-dose nonavalent HPV vaccination schedule, while 80 (34.3%) women served as controls. Baseline demographic and clinical characteristics were comparable between the groups. At 12 months, HPV clearance was significantly more prevalent in the vaccinated group than in the control group (70.6% vs. 52.5%, $p=0.009$). Cytological normalization was also more frequent in vaccinated women (72.5% vs. 26.2%, $p<0.001$). Complete regression occurred in 53.6% of vaccinated women compared with 16.2% of controls ($p<0.001$), and consequently persistent disease was significantly less common in the vaccinated group (10.5% vs. 37.5%, $p<0.001$). Although HPV16, HPV18, other high-risk HPV types, and multiple HPV infections decreased in both groups, genotype-specific differences were not statistically significant. Baseline HPV18 positivity was the only significant predictor of persistent disease (odds ratio: 3.97, 95% confidence interval: 1.29-12.24; $p=0.017$).

Conclusion: Adjuvant nonavalent HPV vaccination was associated with higher HPV clearance, improved cervical cytological outcomes, higher complete regression, and lower persistent disease rates in women with histopathologically confirmed LSIL/CIN1. Prospective multicenter studies are warranted to confirm these findings and further define the role of adjuvant HPV vaccination in the conservative management of LSIL/CIN1.

Keywords: LSIL/CIN1, human papillomavirus, nonavalent HPV vaccine, HPV clearance, cervical cytology, persistent disease

Address for Correspondence: Erdal Şeker MD, Private Clinic, Clinic of Obstetrics and Gynecology, Ankara, Turkey
E-mail: erdalseker84@gmail.com **ORCID ID:** orcid.org/0000-0001-9818-0414
Received: 23.05.2026 **Accepted:** 08.07.2026 **Publication Date:** 02.09.2026

Cite this article as: Halisçelik MA, Oğlak SC, Şeker E, Yeleç S, Aydın MF, Kavak SB. Adjuvant nonavalent HPV vaccination is associated with improved one-year HPV clearance and cervical cytological outcomes in women with histopathologically confirmed LSIL/CIN1: a retrospective comparative cohort study. Anat J Obstet Gynecol Res. 2026;3(2):112-117

INTRODUCTION

Although cervical cancer is largely preventable, it continues to pose a major global health burden and remains one of the most frequently diagnosed malignancies in women worldwide.[1,2] Persistent infection with high-risk human papilloma virus (HPV) is the main etiological contributor to cervical carcinogenesis, with HPV-16 and HPV-18 genotypes responsible for approximately 70% of cases.[2,3] Persistent high-risk HPV infection may lead to the development of cervical intraepithelial neoplasia (CIN), which can progress from CIN1 to CIN2 and CIN3 and, if left untreated, eventually to invasive cervical cancer.[4] low-grade squamous intraepithelial lesions (LSIL) or CIN1 typically represent the early phase of HPV infection and, although a substantial proportion of these lesions undergo spontaneous regression, some cases may progress to persistent infection and disease progression.[4,5]

Prophylactic vaccines developed against HPV play a crucial role in the prevention of cervical cancer, and bivalent, quadrivalent, and nonavalent HPV vaccines are currently in clinical use.[1,6] The nonavalent HPV vaccine provides broader protection against oncogenic HPV types by covering HPV-6, HPV-11, HPV-16, HPV-18, as well as HPV-31, HPV-33, HPV-45, HPV-52, and HPV-58.[6] In recent years, evidence has emerged suggesting that HPV vaccination may not only serve a primary preventive role but also enhance viral clearance in HPV-positive women and reduce the risk of persistence or recurrence of HPV-related lesions. However, evidence regarding the adjuvant use of HPV vaccination in women with histologically confirmed LSIL/CIN1 remains limited, particularly from comparative studies including unvaccinated control groups.[7,8]

The aim of this retrospective comparative cohort study was to compare one-year HPV deoxyribonucleic acid (DNA) and cervical cytology outcomes between women with histopathologically confirmed LSIL/CIN1 who received adjuvant nonavalent HPV vaccination and a contemporaneous unvaccinated control cohort. Secondary objectives were to evaluate complete regression, persistent disease, and baseline clinical factors associated with persistent disease.

METHODS

This single-center retrospective comparative cohort study was conducted at the Department of Obstetrics and Gynecology, University of Health Sciences Turkey, Diyarbakır Gazi Yaşargil Research and Training Hospital, between 2020 and 2025. Women aged 18-65 years with histopathologically confirmed LSIL/CIN1 and no previous history of HPV vaccination were eligible for inclusion.

Patients in the vaccination group received the nonavalent HPV vaccine (Gardasil 9®, Merck Sharp & Dohme LLC, Rahway, NJ, USA) according to the standard three-dose schedule (0, 2, and 6 months). The control cohort included contemporaneous women diagnosed during the same study period who met the same inclusion and exclusion criteria but did not receive HPV vaccination.

Patients who were pregnant or postpartum, receiving immunosuppressive therapy, previously vaccinated against HPV, had undergone prior cervical excisional treatment, were diagnosed with HSIL/CIN2-3 or invasive cervical cancer, or had incomplete clinical follow-up were excluded from cohorts.

Ethical approval for the study was obtained from the Clinical Research Ethics Committee of University of Health Sciences Turkey, Diyarbakır Gazi Yaşargil Training and Research Hospital (approval no: 737, date: 21.11.2025), and the study was conducted in accordance with the principles of the Declaration of Helsinki.

Demographic and clinical data were obtained from the hospital's electronic medical record system. These included age, body mass index (BMI), gravida, parity, smoking status, contraceptive method, baseline cytology and HPV test results, HPV genotype, and histopathological findings from colposcopy-guided biopsies.

Cervical cytology samples were collected under standard conditions, processed using liquid-based cytology, and evaluated according to the Bethesda system.[9] HPV samples were collected during the same visit and analyzed using the Abbout, İstanbul, Turkey real time high-risk HPV assay. HPV-16 and HPV-18 genotypes were reported separately, while other high-risk HPV types were reported collectively.

Indications for colposcopy were determined according to the 2019 risk-based guidelines of the American Society for Colposcopy and Cervical Pathology.[10] All patients included in the study underwent colposcopy-guided cervical biopsy before enrollment, and only women with histopathologically confirmed LSIL/CIN1 were included in the final analysis. Endocervical curettage was performed when clinically indicated.[10]

Histopathological evaluation was performed using hematoxylin-eosin staining, and p16 immunohistochemistry was applied when necessary. Diagnoses were classified according to the World Health Organization and Lower Anogenital Squamous Terminology criteria.[11,12]

Colposcopies were performed by a second-year gynecological oncology fellow or a gynecologist with three years of experience.

Adjuvant nonavalent HPV vaccination was offered to patients diagnosed with LSIL/CIN1.[5] The primary outcomes were one-year HPV clearance and regression of cytological abnormalitiesin the vaccinated and unvaccinated cohorts. Complete regression was defined as the presence of both a negative HPV test and normal cervical cytology at the 12-month follow-up. Persistent disease was defined as persistent HPV positivity together with abnormal cervical cytology at the same follow-up visit. These outcome measures were used to compare the clinical outcomes between the two cohorts.

Statistical Analysis

Statistical analyses were performed using SPSS, version 22.0 (IBM Corp., Armonk, NY, USA). Continuous variables are expressed as mean ± standard deviation, while categorical variables are presented as frequencies and percentages.

Comparisons between groups were performed using the independent samples t-test for continuous variables. Categorical variables were compared using the chi-square test, and Fisher's exact test was applied when the expected cell count was less than five.

Changes in HPV status and cervical cytology between baseline and the 12-month follow-up were evaluated both within and between groups using these methods. Differences between groups in the primary outcomes, including HPV clearance, cytological normalization, complete regression, and persistent disease, were assessed using the chi-square test.

Univariable logistic regression analysis was performed in the vaccinated cohort to identify potential predictors of persistent disease. Age, BMI, gravida, parity, smoking status, and baseline HPV18 positivity were included in the model. The results are reported as odds ratios (ORs) with 95% confidence intervals (95% CIs). However, because the number of events per variable (EPV=2.7) was relatively low, the regression findings were interpreted as exploratory. A two-sided p value of <0.05 was considered statistically significant for all analyses.

RESULTS

A total of 233 women with histopathologically confirmed LSIL/CIN1 were included in the study, comprising 153 (65.7%) women who received adjuvant nonavalent HPV vaccination and 80 (34.3%) unvaccinated controls. The groups were similar in terms of age, BMI, gravida, parity, smoking status, and contraceptive method (all $p>0.05$), indicating good baseline comparability between cohorts (Table 1).

At baseline, all patients in both groups tested positive for HPV. At the 12-month follow-up, the HPV clearance rate was 70.6% (108/153) in the vaccinated group compared with 52.5% (42/80) in the unvaccinated control group ($p=0.009$). Although reductions in HPV16 and HPV18 positivity, as well as in other high-risk HPV types and multiple HPV infections, were observed in both groups, no differences were found between the groups in these genotype-specific comparisons (all $p>0.05$). The detailed distribution of HPV genotypes and HPV clearance is presented in Table 2.

Evaluation of cervical cytology demonstrated a significantly higher rate of cytological normalization in the vaccinated group at the 12-month follow-up. The proportion of women

Table 1. Baseline demographic and clinical characteristics of the study cohorts

Variable	Vaccinated group (n=153)	Unvaccinated control group (n=80)	*p* value
Continuous variables, mean ± SD			
Age (years)	43.62±9.08	44.06±8.60	0.720
BMI (kg/m²)	29.27±2.87	29.34±2.74	0.859
Gravidy	3.42±2.21	3.48±2.32	0.855
Parity	3.23±2.09	3.08±1.73	0.572
Categorical variables, n (%)			
Smoking (yes)	56 (36.6%)	27 (33.8%)	0.773
Contraceptive method, n (%)			
None	52 (34.0%)	37 (46.2%)	0.088
Condom	32 (20.9%)	18 (22.5%)	0.867
Intrauterine device	52 (34.0%)	19 (23.8%)	0.134
Oral contraceptive pills	17 (11.1%)	6 (7.5%)	0.490

Continuous variables were compared using the independent samples t-test, and categorical variables were compared using Fisher's exact test. A p value <0.05 was considered statistically significant
SD: Standard deviation, BMI: Body mass index

Table 2. Comparison of HPV status at baseline and 12 months within and between the study groups

HPV status	Vaccinated group (n=153)		Unvaccinated control group (n=80)		*p* value*
	Baseline, n (%)	12 months, n (%)	Baseline, n (%)	12 months, n (%)	
HPV negative	0 (0.0%)	108 (70.6%)	0 (0.0%)	42 (52.5%)	**0.009**
HPV16	47 (30.7%)	26 (17.0%)	28 (35.0%)	20 (25.0%)	0.199
HPV18	24 (15.7%)	12 (7.8%)	13 (16.2%)	10 (12.5%)	0.358
Other HPV types	68 (44.4%)	5 (3.3%)	37 (46.2%)	8 (10.0%)	0.068
Co-infection (≥2 HPV types)	14 (9.2%)	2 (1.3%)	2 (2.5%)	0 (0.0%)	0.780

*p value: chi-square test comparing HPV status between the vaccinated and unvaccinated groups at the 12-month follow-up. At baseline, all participants in both groups were HPV-positive
HPV: Human papilloma virus

with normal cytology increased from 17.6% at baseline to 72.5% at 12 months in the vaccinated group, compared with an increase from 12.5% to 26.2% in the unvaccinated control group ($p<0.001$). Likewise, the frequencies of atypical squamous cells of undetermined significance (ASC-US) and LSIL decreased significantly in the vaccinated group compared with the control group (both $p<0.001$). In contrast, no significant difference was observed between the groups in the rate of HSIL, which was uncommon at baseline ($p=0.517$). A detailed comparison of cervical cytology results is presented in Table 3.

At the 12-month follow-up, the vaccinated group demonstrated significantly better outcomes than the unvaccinated control group across all primary outcome measures. All primary outcomes are summarized in Table 4. Complete regression, defined as the simultaneous presence of a negative HPV test and normal cervical cytology, was observed in 53.6% of the vaccinated group compared with only 16.2% of the control group at the 12 month follow-up ($p<0.001$). In contrast, the rate of persistent disease, defined as persistent HPV positivity together with abnormal cervical cytology, was significantly lower in the vaccinated group than in the control group (10.5% vs. 37.5%, $p<0.001$).

Univariable logistic regression analysis was performed in the vaccinated cohort (n=153) to identify baseline predictors of persistent disease, which occurred in only 16 patients (10.5%). The detailed results of the logistic regression analysis are presented in Table 5. Age, BMI, gravida, parity, and smoking status were not significantly associated with the development of persistent disease. In contrast, baseline HPV18 positivity was identified as a significant predictor of persistent disease (OR: 3.97; 95% CI: 1.29-12.24; $p=0.017$). Given the relatively low number of EPV (EPV=2.7), these findings should be interpreted as exploratory and require confirmation in larger studies.

Table 3. Within-group and between-group comparison of cervical cytology results at baseline and 12-month follow-up

Cytology	Vaccinated group (n=153)		Unvaccinated control group (n=80)		*p* value*
	Baseline, n (%)	12 months, n (%)	Baseline, n (%)	12 months, n (%)	
Negative	27 (17.6%)	111 (72.5%)	10 (12.5%)	21 (26.2%)	**<0.001**
ASC-US	49 (32.0%)	19 (12.4%)	30 (37.5%)	28 (35.0%)	**<0.001**
LSIL	67 (43.8%)	20 (13.1%)	39 (48.8%)	31 (38.8%)	**<0.001**
HSIL	9 (5.9%)	3 (2.0%)	1 (1.2%)	0 (0.0%)	0.517

*p value: Chi-square test comparing cervical cytology results between the vaccinated and unvaccinated groups at the 12-month follow-up
ASC-US: Atypical squamous cells of undetermined significance, LSIL: Low-grade squamous intraepithelial lesion, HSIL: High-grade squamous intraepithelial lesion

Table 4. Between-group comparison of clinical outcomes at the 12-month follow-up

Outcome	Vaccinated group (n=153), n (%)	Unvaccinated control group (n=80), n (%)	*p* value
HPV clearance (HPV-negative at 12 months)	108/153 (70.6%)	42/80 (52.5%)	**0.009**
Cytological regression (normal pap smear at 12 months)	111/153 (72.5%)	21/80 (26.2%)	**<0.001**
Complete regression† (HPV-negative and normal cytology)	82/153 (53.6%)	13/80 (16.2%)	**<0.001**
Persistent disease‡ (HPV-positive and abnormal cytology)	16/153 (10.5%)	30/80 (37.5%)	**<0.001**

†Complete regression: defined as both HPV negativity and normal cervical cytology at the 12-month follow-up. ‡Persistent disease (strict definition): Defined as concurrent persistence of HPV positivity and abnormal cervical cytology at the 12-month follow-up. *P* values were calculated using the chi-square test
HPV: Human papilloma virus

Table 5. Univariate logistic regression analysis of factors associated with persistent disease in the vaccinated group (n=153)

Variable	Comparison	OR	95% CI	*p* value
Age	Per 1-year increase	0.98	0.92-1.04	0.504
BMI	Per 1 kg/m^2 increase	1.00	0.83-1.20	0.979
Gravidity	Per 1-unit increase	0.98	0.77-1.24	0.839
Parity	Per 1-unit increase	0.94	0.72-1.22	0.642
Smoking	Yes vs. no	0.77	0.25-2.33	0.639
Baseline HPV18	HPV18 vs. other HPV types	3.97	1.29-12.24	**0.017**

Outcome: Persistent disease (concurrent HPV positivity and abnormal cervical cytology at the 12-month follow-up; n=16, 10.5%). All predictor variables represent baseline values. The 12-month HPV18 variable used in the previous analysis was excluded from the model because of circularity
Events per variable: 16 events/6 variables =2.7. Because the EPV was below the recommended threshold, the findings should be considered exploratory and interpreted with caution
BMI: Body mass index, HPV: Human papilloma virus, OR: Odds ratio, CI: Confidence interval

DISCUSSION

In this retrospective comparative cohort study, we evaluated one-year HPV DNA and cervical cytology outcomes in women with histopathologically confirmed LSIL/CIN1 who received adjuvant nonavalent HPV vaccination compared with an unvaccinated control cohort. Women who received adjuvant HPV vaccination were significantly more likely to have achieved HPV clearance ($p=0.009$), higher rates of normal cervical cytology ($p<0.001$), higher rates of complete regression ($p<0.001$), and lower rates of persistent disease ($p<0.001$) than unvaccinated women. Overall, these findings suggest that adjuvant nonavalent HPV vaccination was associated with significantly better one-year virological and cytological outcomes in women with LSIL/CIN1. However, because of the retrospective observational design, a causal relationship cannot be established.

The natural history of LSIL/CIN1 is generally favorable, with most lesions (57-80%) regressing spontaneously without treatment. Persistent high-risk HPV infection remains the principal determinant of disease progression and represents the main target of HPV vaccination.[1,2] Ostör[13] reported spontaneous regression in approximately 57% of CIN1 lesions, whereas 32% persisted, 11% progressed to CIN3, and only 1% progressed to invasive cervical cancer. Similarly, Gardella et al.[5] reported spontaneous regression rates of 60-80% for LSIL/CIN1 and demonstrated that HPV vaccination was associated with faster HPV clearance and a lower risk of disease progression. In the present study, HPV clearance occurred in 52.5% of unvaccinated women after one year, reflecting the expected natural course of LSIL/CIN1. In contrast, vaccinated women achieved a significantly higher rate of HPV clearance (70.6%, $p=0.009$), supporting the potential role of adjuvant HPV vaccination in promoting viral elimination.

Our findings are consistent with previous studies evaluating adjuvant HPV vaccination in HPV-positive women and those with cervical intraepithelial lesions. In a systematic review including 19,414 women, Pruski et al.[14] reported complete HPV remission in up to 72.4% of vaccinated women compared with 45.7% of unvaccinated controls. Likewise, Palumbo et al.[3] found significantly lower persistent HPV positivity among vaccinated women with CIN1 than among unvaccinated controls (18% vs. 38%, $p=0.0169$). Gardella et al.[5] also demonstrated that vaccination shortened the time to HPV clearance (OR: 1.59, 95% CI: 1.22-2.07), while De Vincenzo et al.[15] reported significantly faster HPV clearance after completion of the three-dose nonavalent vaccination schedule (hazard ratio: 7.80, 95% CI: 4.83-12.60; $p<0.0001$).[5] Together, these findings support the association between adjuvant HPV vaccination and improved viral clearance.

In addition to enhanced HPV clearance, vaccinated women in our study demonstrated significantly better cytological outcomes than unvaccinated controls. At the 12-month follow-up, normal cervical cytology was observed in 72.5% of vaccinated women compared with 26.2% of controls ($p<0.001$). Similarly, ASC-US and LSIL were significantly less frequent in vaccinated women (12.4% vs. 35.0% and 13.1% vs. 38.8%, respectively; both $p<0.001$), whereas no significant difference was observed in HSIL frequency ($p=0.517$). Vaccinated women also had higher complete regression rates (53.6% vs. 16.2%, $p<0.001$) and lower rates of persistent disease (10.5% vs. 37.5%, $p<0.001$). Collectively, these findings suggest that the potential benefit of adjuvant HPV vaccination extends beyond viral clearance and may also be reflected by improved cytological recovery. Similar observations have been reported by Gardella et al.,[5] who demonstrated lower rates of CIN persistence and progression among vaccinated women, and by Pruski et al.,[14] whose systematic review concluded that adjuvant HPV vaccination improved clinical outcomes by enhancing HPV clearance and reducing HPV-related disease progression.[5] Although most previous studies have primarily focused on HPV clearance, our findings indicate that favorable cytological outcomes may also be associated with adjuvant vaccination in women with histologically confirmed LSIL/CIN1.

Baseline HPV18 positivity was the only significant predictor of persistent disease in our cohort (OR: 3.97, 95% CI: 1.29-12.24; $p=0.017$). Persistent HPV18 infection has long been recognized as an important risk factor for cervical carcinogenesis because of its strong association with cervical adenocarcinoma.[16-19] However, the wide CI and the limited number of outcome events indicate that this finding should be interpreted cautiously. Moreover, the low EPV ratio limits the stability of the regression model, and the predictive value of baseline HPV18 positivity should therefore be considered exploratory until confirmed in larger prospective studies.

Study Limitations

This study has several strengths, including the inclusion of only histopathologically confirmed LSIL/CIN1 cases, the use of a contemporaneous unvaccinated control group, and a clinically meaningful definition of persistent disease based on concurrent HPV positivity and abnormal cervical cytology. Nevertheless, several limitations should be acknowledged. The retrospective single-center design precludes causal inference, the 12-month follow-up does not permit assessment of long-term outcomes, and histological confirmation was not available at follow-up. In addition, important potential confounders, including sexual behavior, new sexual partners, condom use, immunological status, and concomitant sexually transmitted infections, could not be evaluated because these variables were not consistently documented in the medical records. These factors should be considered when interpreting the observed associations between adjuvant HPV vaccination and clinical outcomes.

CONCLUSION

Adjuvant nonavalent HPV vaccination was associated with higher HPV clearance, improved cervical cytological outcomes, higher complete regression, and lower persistent disease rates than in unvaccinated women with histopathologically confirmed LSIL/CIN1. These findings support the potential role of adjuvant HPV vaccination as an adjunct to conservative management. Prospective multicenter studies with longer follow-up are needed to confirm these findings.

Ethics

Ethics Committee Approval: Ethical approval for the study was obtained from the Clinical Research Ethics Committee of University of Health Sciences Turkey, Diyarbakır Gazi Yaşargil Training and Research Hospital (approval no: 737, date: 21.11.2025).

Informed Consent: This study was conducted as a retrospective observational case series.

Footnotes

Authorship Contributions

Surgical and Medical Practices: M.A.H., E.Ş., Concept: M.A.H., E.Ş., Design: M.A.H., E.Ş., S.Y., S.B.K., Data Collection or Processing: M.A.H., E.Ş., Analysis or Interpretation: M.A.H., S.C.O., E.Ş., S.Y., M.F.A., S.B.K., Literature Search: M.A.H., E.Ş., Writing: M.A.H., E.Ş., S.Y., M.F.A.

Conflict of Interest: One author of this article, Süleyman Cemil Oğlak, is a member of the Editorial Board of the Anatolian Journal of Obstetrics and Gynecology Research. However, he did not involved in any stage of the editorial decision of the manuscript.

Financial Disclosure: The authors received no financial support for the research, authorship, and/or publication of this article

REFERENCES

1. Illah O, Olaitan A. Updates on HPV vaccination. Diagnostics (Basel). 2023;13(2):243.
2. Kamolratanakul S, Pitisuttithum P. Human papillomavirus vaccine efficacy and effectiveness against cancer. Vaccines (Basel). 2021;9(12):1413.
3. Palumbo M, Lavitola G, Di Filippo C, et al. Impact of human papillomavirus 9-valent vaccine on viral clearance after surgical treatment: a single-center retrospective observational study. Eur J Obstet Gynecol Reprod Biol. 2025;310:113994.
4. Baasland I, Bjørge T, Engesæter B, Tropé A, Opdahl S. Cervical intraepithelial neoplasia grade 1 and long-term risk of progression and treatment. PLoS One. 2025;20(4):e0320739.
5. Gardella B, Dominoni M, Pasquali MF, et al. Low-grade cervical intraepithelial neoplasia (CIN1) evolution: analysis of opportunistic preventive vaccination role. Vaccines (Basel). 2023;11(2):284.
6. Dvořák V, Petráš M, Dvořák V, et al. Reduced risk of CIN2+ recurrence in women immunized with a 9-valent HPV vaccine post-excision: retrospective cohort study. Hum Vaccin Immunother. 2024;20(1):2343552.
7. Pruski D, Millert-Kalińska S, Jach R, Przybylski M. Effect of vaccination against HPV in the HPV-positive patients not covered by primary prevention on the disappearance of infection. Sci Rep. 2025;15(1):12642.
8. Sutcuoglu BM. Efficacy of the HPV vaccine in the secondary prevention of cervical dysplasia in patients undergoing surgery. Eurasian J Med Investig. 2023;7(3):181-187.
9. Nayar R, Wilbur DC, editors. The bethesda system for reporting cervical cytology: definitions, criteria, and explanatory notes. 3rd ed. Cham: Springer; 2015.
10. Perkins RB, Guido RS, Castle PE, et al.; 2019 ASCCP Risk-Based Management Consensus Guidelines Committee. 2019 ASCCP risk-based management consensus guidelines for abnormal cervical cancer screening tests and cancer precursors. J Low Genit Tract Dis. 2020;24(2):102-131. Erratum in: J Low Genit Tract Dis. 2020;24(4):427.
11. Darragh TM, Colgan TJ, Thomas Cox J, et al.; Members of the LAST Project Work Groups. The lower anogenital squamous terminology standardization project for HPV-associated lesions: background and consensus recommendations from the College of American Pathologists and the American Society for Colposcopy and Cervical Pathology. Int J Gynecol Pathol. 2013;32(1):76-115. Erratum in: Int J Gynecol Pathol. 2013;32(4):432. Erratum in: Int J Gynecol Pathol. 2013;32(2):241.
12. Höhn AK, Brambs CE, Hiller GGR, May D, Schmoeckel E, Horn LC. 2020 WHO Classification of Female Genital Tumors. Geburtshilfe Frauenheilkd. 2021;81(10):1145-1153.
13. Ostör AG. Natural history of cervical intraepithelial neoplasia: a critical review. Int J Gynecol Pathol. 1993;12(2):186-192.
14. Pruski D, Millert-Kalińska S, Jach R, Żurawski J, Przybylski M. Impact of vaccinating adult women who are HPV-positive or with confirmed cervical SIL with the 9-valent vaccine-a systematic review. Viruses. 2025;17(10):1377.
15. De Vincenzo R, Caporale N, Bertoldo V, et al. HPV and cytology testing in women undergoing 9-valent HPV opportunistic vaccination: a single-cohort follow up study. Vaccines (Basel). 2021;9(6):643.
16. IARC Working Group on the evaluation of carcinogenic risks to humans. Biological agents. IARC Monogr Eval Carcinog Risks Hum. 2012;100(Pt B):1-441.
17. Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. Lancet. 2007;370(9590):890-907.
18. Katki HA, Schiffman M, Castle PE, et al. Five-year risks of CIN 3+ and cervical cancer among women with HPV testing of ASC-US Pap results. J Low Genit Tract Dis. 2013;17(5 Suppl 1):S36-42.
19. Oğlak SC, Obut M. Comparison of pap-smear and colposcopy in the absence of HPV test for the diagnosis of premalignant and malignant cervical lesions. Eastern J Med. 2020;25(2):299-304.

Original Article

DOI: 10.4274/anajog.galenos.2026.36349

Anat J Obstet Gynecol Res 2026;3(2):118-123

Internet-Related Health Anxiety in Women Diagnosed with Endometrial Intraepithelial Neoplasia: A Cross-Sectional Study

Hande Nur Öncü[1], Oğuz Kaan Köksal[1], Şahin Kaan Baydemir[1], Candost Hanedan[1], Çağanay Soysal[2], Vakkas Korkmaz[3]

[1]University of Health Sciences Turkey, Ankara Etlik City Hospital, Clinic of Gynecologic Oncology, Ankara, Turkey
[2]University of Health Sciences Turkey, Ankara Etlik City Hospital, Clinic of Obstetrics and Gynecology, Ankara, Turkey
[3]Lokman Hekim University Faculty of Medicine, Department of Obstetrics and Gynecology, Ankara, Turkey

ABSTRACT

Purpose: To evaluate the level of internet-related health anxiety and the factors associated with it in women diagnosed with endometrial intraepithelial neoplasia (EIN).

Methods: Women diagnosed with EIN were included in the study. Socio-demographic and clinical characteristics of the participants were recorded. Internet-related health anxiety was assessed using the cyberchondria severity scale-12 (CSS-12) with higher scores suggesting increased anxiety. The Mann-Whitney U test was used for between-group comparisons. Variables associated with the total CSS-12 score at $p<0.10$ in univariate analysis were included in the multivariable linear regression model.

Results: A total of 78 women participated with a median age of 44 (26-69) years, and median BMI of 32.1 kg/m^2 (21.8-47.6). The median total CSS-12 score was 31 (12-60), with higher scores indicating greater internet-related health anxiety. In univariate analysis, women younger than 40 years had significantly higher CSS-12 scores than women aged 40 years or older (37 vs. 31, $p=0.004$). CSS-12 scores were also significantly higher in nulliparous women than in multiparous women (45 vs. 32, $p=0.003$), and in women with fertility desire than in those without fertility desire (42 vs. 29, $p<0.001$). No significant association was found between educational level and CSS-12 score ($p=0.200$). In multivariable linear regression analysis, age younger than 40 years [B=5.41, 95% confidence interval (CI): 2.18-8.64, $p=0.001$], nulliparity (B=4.83, 95% CI: 1.56-8.10, $p=0.005$), and fertility desire (B=7.68, 95% CI: 3.94-11.42, $p<0.001$) were independently associated with the total CSS-12 score.

Conclusion: Among women diagnosed with EIN, greater levels of internet-related health anxiety was associated with age <40 years, nulliparity, and ongoing fertility desire. Accurate post-diagnosis information, guidance toward reliable online resources, and individualized counseling may help reduce internet-related health anxiety in this patient group.

Keywords: Cyberchondria, cyberchondria severity scale-12, endometrial intraepithelial neoplasia, internet-related health anxiety, online

INTRODUCTION

Endometrial intraepithelial neoplasia (EIN) is a premalignant lesion of the endometrium and is considered a direct precursor of endometrial adenocarcinoma.[1,2] Among patients diagnosed with EIN who undergo hysterectomy, endometrial cancer is detected in the hysterectomy specimen in approximately 30% to 50% of cases.[3-5] Although hysterectomy is the standard treatment approach for EIN, fertility-sparing treatment options may also be considered in women who have not yet given birth or who wish to preserve fertility. The uncertainty inherent in this premalignant condition and the complexity of treatment decisions may create substantial psychological burden and anxiety in women diagnosed with EIN.[2,6]

Address for Correspondence: Hande Nur Öncü, MD, University of Health Sciences Turkey, Ankara Etlik City Hospital, Clinic of Gynecologic Oncology, Ankara, Turkey
E-mail: handekayahan@gmail.com **ORCID ID:** orcid.org/0000-0002-9407-2704
Received: 09.06.2026 **Accepted:** 22.07.2026 **Publication Date:** 02.09.2026

Cite this article as: Öncü HN, Köksal OK, Baydemir ŞK, Hanedan C, Soysal Ç, Korkmaz V. Internet-related health anxiety in women diagnosed with endometrial intraepithelial neoplasia: a cross-sectional study. Anat J Obstet Gynecol Res. 2026;3(2):118-123

Today, the internet has become one of the most common and easily accessible sources of health information. As of 2023, the number of internet users worldwide has reached 5.3 billion, and with the widespread use of smartphones and digital platforms, the internet has become one of the most frequently used sources for health-related information. Individuals with an existing illness or a newly diagnosed condition tend to search online for their symptoms, diagnoses, and treatment options before or after consulting healthcare professionals. However, this search behavior does not always serve an informative function; on the contrary, it may sometimes turn into a cycle that reinforces and sustains anxiety. The increasing use of the internet to access health-related information is thought to contribute to anxiety and predispose individuals to cyberchondria, particularly due to the presence of conflicting and sometimes excessively detailed information.[7,8]

Cyberchondria is characterized by increased health anxiety and worry resulting from excessive or repetitive online searches for health-related information.[9] To assess this concept, McElroy and Shevlin[10] developed the 33-item cyberchondria severity scale (CSS) in 2014. Subsequently, the CSS-12, a shorter and more practical version for clinical and research settings, was introduced.[11] The Turkish validity and reliability study of the CSS-12 was conducted by Yorgancıoğlu Tarcan et al.[12] in 2023, and the internal consistency coefficient of the scale was reported as 0.80.

Women diagnosed with EIN may experience intense worry due to the possibility of hysterectomy, the risk of fertility loss, and the potential presence of concurrent cancer. In this patient group, the clinical factors associated with online health information-seeking behavior and internet-related health anxiety have not yet been adequately investigated. Identifying high-risk patient groups in advance is important for healthcare professionals to provide accurate information and individualized counseling. In this study, we aimed to evaluate internet-related health anxiety using the CSS-12 and to determine the clinical and demographic factors associated with this anxiety in women diagnosed with EIN.

METHODS

This cross-sectional study was conducted between August 2025 and June 2026 after obtaining approval from the University of Health Sciences Turkey, Ankara Etlik City Hospital Scientific Research Ethics Committee (approval no: AEŞH-BADEK2-2025-019, date: 22.07.2025). The study included women diagnosed with EIN based on histopathological examination. Women aged 18 years or older with histopathologically confirmed EIN who reported using the internet for at least one hour per day and who self-reported seeking health-related information online were included in the study. The criterion of at least one hour of daily internet use was selected to ensure regular exposure to online information sources. Patients who did not use the internet daily, had a known psychiatric disorder or regular psychiatric medication use, had missing responses in the questionnaire form, or did not agree to participate were excluded. All participants were informed about the study, and written informed consent was obtained before administration of the questionnaire.

Internet-related health anxiety was assessed using the CSS-12. This scale, developed by McElroy et al.,[11] was adapted into Turkish and validated in 2023 by Yorgancıoğlu Tarcan et al.[12] The CSS-12 is a 5-point Likert-type scale consisting of 12 items, each scored from 1 to 5. The scale includes four subscales: compulsion, excessiveness, distress, and reassurance seeking. The compulsion subscale evaluates the extent to which online health searches interfere with daily functioning; the excessiveness subscale evaluates repetitive and difficult-to-control health information-seeking behavior; the distress subscale evaluates negative emotional responses after online health searches; and the Reassurance seeking subscale evaluates the tendency to repeatedly seek confirmation and reassurance from healthcare professionals or other sources. Each subscale score ranges from 3 to 15, while the total CSS-12 score ranges from 12 to 60. Higher scores indicate higher levels of internet-related health anxiety.[11]

Data about participants' age, body mass index (BMI), parity, menopausal status, fertility desire, and educational level were recorded. Age was evaluated in two groups: younger than 40 years and 40 years or older. Parity status was classified as nulliparous or multiparous. Educational level was categorized as primary-middle school and high school or above. The associations of these variables with the total CSS-12 score and subscale scores were analyzed.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics for Mac, version 22.0 (IBM Corp., Armonk, NY, USA). The distribution of continuous variables was assessed using the Kolmogorov-Smirnov test. Continuous variables that did not show normal distribution are presented as median (minimum-maximum), while categorical variables are presented as numbers and percentages.

The Mann-Whitney U test was used for pairwise group comparisons of the total CSS-12 score and subscale scores. Variables associated with the total CSS-12 score at $p<0.10$ in univariate analyses were included in the multivariable linear regression model using the enter method. Accordingly, age younger than 40 years, nulliparity, and fertility desire were entered into the multivariable linear regression model. The total CSS-12 score was considered a continuous outcome variable and was used as the dependent variable in the multivariable model.

Since the study was conducted over a predefined period, a post-hoc power analysis was performed using G*Power 3.1.9.7 to evaluate the adequacy of the final sample size. For multiple linear regression with three predictors, assuming a medium effect size ($f^2=0.15$), $\alpha=0.05$, and 80% power, the minimum required sample size was 77 participants. The final sample exceeded this value.

Before multivariable linear regression analysis, model assumptions were evaluated in terms of residual distribution and multicollinearity. Multicollinearity was assessed using variance inflation factor and tolerance values. Linear regression analysis was considered appropriate because the model residuals showed an acceptable distribution and no

significant multicollinearity was detected. Regression results were reported as B coefficients, 95% confidence intervals (CIs), and p values. A p value of <0.05 was considered statistically significant.

RESULTS

The median age of the 78 participants was 44 years (26-69), the median BMI was 32.1 kg/m^2 (21.8-47.6), and the median parity was 1 (0-6). Of the participants, 56 (71.8%) were premenopausal and 22 (28.2%) were postmenopausal. In terms of parity status, 19 (24.4%) were nulliparous and 59 (75.6%) were multiparous. Fertility desire was reported by 17 patients (21.8%) and absent in 61 patients (78.2%). Educational level was primary-middle school in 26 patients (33.3%) and high school or above in 52 patients (66.7%). The median total CSS-12 score was 31 (12-60) (Table 1).

In the univariate analysis, the total CSS-12 score was 37 (15-58) in women younger than 40 years and 31 (12-50) in women aged 40 years or older ($p=0.004$). Women younger than 40 years also had significantly higher scores in the excessiveness, distress, and compulsion subscales, whereas the difference in the reassurance subscale did not reach statistical significance. The total CSS-12 score was 45 (19-60) in nulliparous women and 32 (12-49) in multiparous women ($p=0.003$), with significantly higher scores in all CSS-12 subscales among nulliparous women. No significant difference was observed in total CSS-12 score based on menopausal status with premenopausal women scoring 32 (12-60) vs. postmenopausal 29 (13-56), ($p=0.176$), and none of the CSS-12 subscale scores differed significantly according to menopausal status. The total CSS-12 score was 42 (18-60) in women with fertility desire and 29 (12-49) in who did not express a desire for fertility ($p<0.001$), with significantly higher scores in all CSS-12 subscales among women with fertility desire. The total CSS-12 score was 32 (12-50) in women with primary-middle school education and 33 (13-57) in those with high school or above education ($p=0.200$), and no significant differences were found in CSS-12 subscale scores according to educational level (Table 2).

Table 1. Demographic and clinical characteristics of the study population (n=78)

Variable	Value
Age (years), median (min-max)	44 (26-69)
BMI (kg/m^2), median (min-max)	32.1 (21.8-47.6)
Parity, median (min-max)	1 (0-6)
Total CSS-12 score, median (min-max)	31 (12-60)
Age group, n (%)	
<40 years	27 (34.6)
≥40 years	51 (65.4)
Menopausal status, n (%)	
Premenopausal	56 (71.8)
Postmenopausal	22 (28.2)
Parity status, n (%)	
Nulliparous	19 (24.4)
Multiparous	59 (75.6)
Fertility desire, n (%)	
Present	17 (21.8)
Absent	61 (78.2)
Educational level, n (%)	
Primary-middle school	26 (33.3)
High school or above	52 (66.7)

BMI: Body mass index, CSS-12: Cyberchondria severity scale-12

In the multivariable linear regression analysis, age younger than 40 years (B=5.41, 95% CI: 2.18-8.64, $p=0.001$), nulliparity (B=4.83, 95% CI: 1.56-8.10, $p=0.005$), and fertility desire (B=7.68, 95% CI: 3.94-11.42, $p<0.001$) were independently associated with the total CSS-12 score. The model explained 42% of the variance in the total CSS-12 score ($R^2=0.42$; adjusted $R^2=0.39$) (Table 3). No significant multicollinearity was detected among the variables included in the multivariable model (VIF range: 1.18-1.74; tolerance range: 0.57-0.85).

DISCUSSION

This study evaluated internet-related health anxiety using the CSS-12 scale in women diagnosed with EIN, a population for which evidence on cyberchondria remains limited. In our study, the median total CSS-12 score was 31 (12-60). As no established cut-off values exist for the CSS-12, this score cannot be categorized as low, moderate, or high. In the multivariable analysis, age younger than 40 years, nulliparity, and fertility desire were identified as three factors independently associated with the total CSS-12 score. Our findings highlight the importance of clinicians developing awareness of internet-related health anxiety in the management of women diagnosed with EIN.

Age younger than 40 years was independently associated with the total CSS-12 score ($p=0.001$), and CSS-12 scores were also higher in women younger than 40 years in the univariate analysis ($p=0.004$). This finding is consistent with the existing literature.[13-15] In a general population study by Aleyeidi et al.,[14] younger women were shown to have higher cyberchondria scores compared with older women. In the study conducted by Güleşen and Beydağ[16] among women with heart disease, cyberchondria levels were also found to be higher in women younger than 50 years compared with those aged 50 years or older. In our study, receiving a premalignant diagnosis, such as EIN, at a young age may add an additional perceived threat to reproductive health and may further increase online information-seeking behavior.

The strongest independent predictor was the presence of fertility desire (B=7.68, 95% CI: 3.94-11.42, $p<0.001$), with a markedly higher median CSS-12 score in women with fertility desire compared with those without (42 vs. 29, $p<0.001$). Nulliparity was also identified as an independent predictor (B=4.83, 95% CI: 1.56-8.10, $p=0.005$), with higher median CSS-12 scores in nulliparous women compared with multiparous women (45 vs. 32, $p=0.003$). The substantial overlap between these two variables, and the fact that both were identified as independent predictors, suggest

Table 2. Univariate analysis of total CSS-12 and subscale scores

Variable	Total CSS-12	*p* value	Excessiveness	*p* value	Distress	*p* value	Reassurance	*p* value	Compulsion	*p* value
Age group										
<40 years, n=27	37 (15-58)	0.004	11 (5-15)	0.006	10 (4-15)	0.008	9 (3-14)	0.068	8 (3-14)	0.038
≥40 years, n=51	31 (12-50)		8 (3-15)		7 (3-12)		7 (3-12)		6 (3-12)	
Parity status										
Nulliparous, n=19	45 (19-60)	0.003	12 (6-15)	0.006	11 (5-15)	0.007	9 (4-15)	0.048	8 (3-15)	0.046
Multiparous, n=59	32 (12-49)		8 (3-13)		7 (3-12)		7 (3-12)		6 (3-12)	
Menopausal status										
Premenopausal, n=56	32 (12-60)	0.176	9 (3-15)	0.232	8 (3-15)	0.189	7 (3-15)	0.448	7 (3-15)	0.275
Postmenopausal, n=22	29 (13-56)		8 (3-14)		7 (3-14)		7 (3-13)		6 (3-15)	
Fertility desire										
Present, n=17	42 (18-60)	<0.001	12 (6-15)	0.003	11 (5-15)	0.001	10 (4-15)	0.003	9 (3-15)	0.008
Absent, n=61	29 (12-49)		8 (3-13)		7 (3-12)		7 (3-12)		6 (3-12)	
Educational level										
Primary-middle school, n=26	32 (12-50)	0.200	8 (3-13)	0.168	7 (3-12)	0.224	7 (3-12)	0.328	6 (3-13)	0.128
High school or above, n=52	33 (13-57)		9 (4-15)		8 (3-14)		8 (3-14)		8 (3-14)	

Values are presented as median (minimum-maximum). Mann-Whitney U test was used for group comparisons
CSS-12: Cyberchondria severity scale-12

Table 3. Multivariable linear regression analysis for factors associated with total CSS-12 score

Variable	B	95% CI	*p* value
Age <40 years	5.41	2.18 to 8.64	0.001
Nulliparity	4.83	1.56 to 8.10	0.005
Fertility desire	7.68	3.94 to 11.42	<0.001

Dependent variable: total CSS-12 score. R^2=0.42; adjusted R^2=0.39. Reference categories were age ≥40 years, multiparity, and absence of fertility desire
CSS-12: Cyberchondria severity scale-12, B: Unstandardized regression coefficient, CI: Confidence interval

a common psychological background. Stroeken et al.[17] described common experiences among young women with gynecologic cancer as feeling trapped between achieving the best oncological outcome and preserving fertility, as well as experiencing time pressure. They also showed that young age at diagnosis, time pressure, and inadequate counseling may increase reproductive concerns and lead to long-term psychological distress.[17] Similarly, Deng et al.[18] demonstrated that fertility-related concerns may outweigh disease-related concerns and that worries about fertility loss may negatively affect treatment adherence and quality of life. In a study conducted among patients with cervical cancer, fertility concerns were reported to be more pronounced in nulliparous women and to mediate the relationship between depressive symptoms and fear of recurrence.[19] The psychological burden caused by social pressure associated with motherhood in childless women has also been emphasized in the literature.[20]

In our society, where the notion of children and family are highly valued, having a premalignant diagnosis like EIN may trigger a similar psychological burden mechanism in women with fertility desire. These women have to face the possibility of hysterectomy or cancer and the uncertainties of fertility-sparing treatment at the same time; this may create an ongoing cycle of anxiety and fuel repeated online information-seeking behavior.

Although nulliparity and fertility desire are conceptually related, they do not completely represent the same clinical construct. Nulliparity reflects reproductive history, whereas fertility desire reflects ongoing reproductive intention and future-oriented concern. Therefore, the independent association of both variables with the total CSS-12 score may suggest that internet-related health anxiety in women with EIN is influenced both by not having children and by the uncertainty surrounding future fertility plans.

The findings of this study are consistent with the limited but growing literature on cyberchondria in the field of gynecologic oncology.[15,21-23]. In studies evaluating the relationship between human papillomavirus (HPV) positivity, which may be associated with fear of malignant transformation, and cyberchondria, HPV 16/18 positivity was also found to have a significant effect on cyberchondria levels.[21,22] Corrales et al.[15] found that the use of the internet for health information was more common among younger and more highly educated patients attending a gynecologic oncology clinic. In a study conducted among patients undergoing oncologic surgery, a significant positive relationship was found between cyberchondria and surgical fear, and a high level of cyberchondria was reported to be the most

important independent variable determining fear of surgery.[23] This finding indicates that cyberchondria may not be merely a passive accompanying condition in oncologic processes, but rather an active factor that may influence clinical outcomes.

In the present study no significant association was found between educational level and the total CSS-12 score ($p=0.200$). The relationship between educational level and cyberchondria is not consistent in the literature. Some studies have reported that higher educational level is associated with lower cyberchondria scores, and this has been explained by better health literacy, greater ability to evaluate information sources, and a more critical interpretation of digital content.[24,25] In contrast, other studies have reported that individuals with higher educational levels may demonstrate more frequent health information-seeking behavior because they use the internet more often and have greater access to health-related information.[15,26] These differing results suggest that cyberchondria cannot be explained solely by formal educational level.

From a clinical perspective, these findings highlight the importance of structured counseling from the time of diagnosis in patients diagnosed with EIN, particularly in young, nulliparous women and those with ongoing fertility desire. The risk of concurrent malignancy, treatment options, and fertility-sparing approaches in eligible patients should be clearly explained on an individual basis. In addition, guiding patients toward reliable online resources and providing individualized patient education may help reduce repetitive and anxiety-provoking online searches.

One of the strengths of this study is that it evaluated internet-related health anxiety in a specific patient group with a premalignant condition that is clinically important in terms of fertility. To our knowledge, no previously published study has specifically evaluated cyberchondria or internet-related health anxiety in women diagnosed with EIN. Available evidence in gynecologic populations has mainly focused on HPV-positive women and gynecologic oncology patients.[15,21,22] In addition, the use of the validated and reliable CSS-12 scale enabled internet-related health anxiety to be assessed using a standardized method. Demonstrating clinically meaningful variables such as age, nulliparity, and fertility desire as independent predictors in multivariable analysis is also one of the important strengths of the study.

Study Limitations

Nevertheless, this study has some limitations. First, the study has a single-center and cross-sectional design and therefore, causal relationships between variables cannot be established. The limited sample size requires caution when interpreting subgroup analyses. Additional psychometric scales evaluating general health anxiety, depression, or anxiety levels were not used in the study. In addition, participants' internet use and online health information-seeking behaviors were based on self-report. Another limitation is that only women who regularly used the internet and reported online health information-seeking behavior were included. This may have introduced selection bias, as women without routine internet use, limited digital access, or lower digital health literacy were not represented in the study population. Therefore, the findings may not be generalizable to all women diagnosed with EIN. The type, quality, and reliability of the online sources accessed by participants were not objectively evaluated. Finally, the absence of a control group prevents comparison of EIN-specific internet-related health anxiety with benign gynecologic conditions or the general population.

CONCLUSION

Internet-related health anxiety was associated with age younger than 40 years, nulliparity, and expressed fertility desire among women diagnosed with EIN. These findings indicate that, in the management of EIN, not only oncologic treatment decisions but also patients' post-diagnosis information-seeking behaviors and psychological needs should be taken into consideration. Accurate information, guidance toward reliable sources, and individualized counseling may help reduce internet-related health anxiety, particularly in young, nulliparous patients with fertility desire. Prospective, multicenter studies including control groups and more comprehensive psychosocial assessments are needed to address this issue more broadly.

Ethics

Ethics Committee Approval: Approval from the University of Health Sciences Turkey, Ankara Etlik City Hospital Scientific Research Ethics Committee (approval no: AEŞH-BADEK2-2025-019, date: 22.07.2025).

Informed Consent: All participants were informed about the purpose and procedures of the study, and written informed consent was obtained prior to completion of the questionnaire. Participant confidentiality and privacy were strictly maintained.

Footnotes

Authorship Contributions

Surgical and Medical Practices: O.K.K., Ş.K.B., C.H., Concept: H.N.Ö., V.K., Design: H.N.Ö., V.K., Data Collection or Processing: H.N.Ö., O.K.K., Ş.K.B., C.H., Ç.S., Analysis or Interpretation: H.N.Ö., Ç.S., V.K., Literature Search: H.N.Ö., Writing: H.N.Ö.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

Declaration of Generative AI and AI-assisted Technologies in the Writing Process: During the preparation of this work, the author(s) used ChatGPT to improve "readability" and "language". After using this tool/service, the author(s) reviewed and edited the content as needed and took full responsibility for the content of the publication.

REFERENCES

1. Sherman ME. Theories of endometrial carcinogenesis: a multidisciplinary approach. Mod Pathol. 2000;13(3):295-308.
2. Management of endometrial intraepithelial neoplasia or atypical endometrial hyperplasia: ACOG clinical consensus No. 5. Obstet Gynecol. 2023;142(3):735-744.

3. Trimble CL, Kauderer J, Zaino R, et al. Concurrent endometrial carcinoma in women with a biopsy diagnosis of atypical endometrial hyperplasia: a Gynecologic Oncology Group study. Cancer. 2006;106(4):812-819.
4. Dolanbay M, Kutuk MS, Uludag S, et al. Concurrent endometrial carcinoma in hysterectomy specimens in patients with histopathological diagnosis of endometrial hyperplasia in curettage specimens. Ginekol Pol. 2015;86(10):753-758.
5. Doherty MT, Sanni OB, Coleman HG, et al. Concurrent and future risk of endometrial cancer in women with endometrial hyperplasia: a systematic review and meta-analysis. PLoS One. 2020;15(4):e0232231.
6. Rose J, Baandrup L, Hertzum-Larsen R, Aarestrup J, Baker JL, Kjær SK. Trends in the incidence of endometrial intraepithelial neoplasia and endometrial cancer by type, age, and socioeconomic status in Denmark 2005-2022. Maturitas. 2026;209:108959.
7. Li H, Li D, Zhai M, Lin L, Cao Z. Associations among online health information seeking behavior, online health information perception, and health service utilization: cross-sectional study. J Med Internet Res. 2025;27:e66683.
8. Bahadir O, Dundar C. The impact of online health information source preference on intolerance to uncertainty and cyberchondria in a youthful generation. Indian J Psychiatry. 2024;66(4):360-366.
9. Starcevic V, Berle D. Cyberchondria: towards a better understanding of excessive health-related Internet use. Expert Rev Neurother. 2013;13(2):205-213.
10. McElroy E, Shevlin M. The development and initial validation of the cyberchondria severity scale (CSS). J Anxiety Disord. 2014;28(2):259-265.
11. McElroy E, Kearney M, Touhey J, Evans J, Cooke Y, Shevlin M. The CSS-12: development and validation of a short-form version of the cyberchondria severity scale. Cyberpsychol Behav Soc Netw. 2019;22(5):330-335.
12. Yorgancıoğlu Tarcan G, Karahan A, Sebik NB. Validity and reliability of the short form cyberchondry (CSS-12) severity scale: a specific application for health informatics. Hacettepe Sağlık İdaresi Dergisi. 2023;26(1):207-218.
13. Kobryn M, Duplaga M. Does health literacy protect against cyberchondria: a cross-sectional study? Telemed J E Health. 2024;30(4):e1089-e1100.
14. Aleyeidi N, Fayed A, Alnowaiser G, et al. Cyberchondria, somatic symptoms, and internet-based self-diagnosis among the Saudi population. Sci Rep. 2025;15(1):18208.
15. Corrales DM, Wells AE, Radecki Breitkopf C, et al. Internet use by gynecologic oncology patients and its relationship with anxiety. J Health Commun. 2018;23(3):299-305.
16. Güleşen A, Beydağ KD. Cryberchondria level in women with heart disease and affecting factors. Arch Health Sci Res. 2020;7(1):1-7.
17. Stroeken Y, Hendriks F, Beltman J, Ter Kuile M. Quality of life and psychological distress related to fertility and pregnancy in ayas treated for gynecological cancer: a systematic review. Cancers. 2024;16(20):3456.
18. Deng J, Chen J, Guo X, Xie C. Factors influencing fertility distress in reproductive-aged gynecologic cancer patients: a narrative overview. J Clin Med. 2025;14(23):8512.
19. He XY, Huang Y, Qiao CP, et al. Fertility anxiety partially mediates depression and recurrence fear in reproductive-age cervical cancer patients: a cross-sectional study. World J Clin Oncol. 2025;16(9):110031.
20. Foti FL, Karner-Huţuleac A, Maftei A. The value of motherhood and psychological distress among infertile women: the mediating role of coping strategies. Front Public Health. 2023;11:1024438.
21. Turhan Cakir A. Cyberchondria levels in women with human papilloma virus. J Obstet Gynaecol Res. 2022;48(10):2610-2614.
22. Öncü HN, Ege G, Sucu ST, Öztürk N, Hanedan C, Korkmaz V. The impact of human papillomavirus 16/18 positivity on cyberchondria levels in women undergoing colposcopy. Rev Assoc Med Bras (1992). 2025;71(5):e20241692.
23. Polat H, Kapıkıran G, Kartal M. The effect of cyberchondria on anxiety and surgical fear levels in surgical oncology patients. Curr Psychol. 2025;44(7):6531-6540.
24. Peng XQ, Chen Y, Zhang YC, et al. The status and influencing factors of cyberchondria during the COVID-19 epidemic. a cross-sectional study in nanyang city of China. Front Psychol. 2021;12:712703.
25. Özbaş C, Yıldız ET, Tüzün H, Kara Çiğdem AG, Uğraş Dikmen A. Health literacy and sociodemographic determinants of cyberchondria: a cross-sectional study among outpatients of a university hospital in Turkey. Healthcare (Basel). 2025;13(19):2445.
26. Lee CJ, Gray SW, Lewis N. Internet use leads cancer patients to be active health care consumers. Patient Educ Couns. 2010;81 Suppl(0):S63-S69.

Original Article

DOI: 10.4274/anajog.galenos.2026.69775

Anat J Obstet Gynecol Res 2026;3(2):124-130

Prediction of Surgical Resectability after Neoadjuvant Chemotherapy in Advanced Ovarian Cancer: Comparison of MRI-Based and Laparoscopy-Based Scoring Systems

Süleyman Özen[1], Eda Güner Özen[2], Gülin Özuyar Şimşek[1], İnci Başkır İşlek[1], Muzaffer Sancı[1]

[1]University of Health Sciences Turkey, İzmir City Hospital, Clinic of Gynecologic Oncology Surgery, İzmir, Turkey
[2]University of Health Sciences Turkey, İzmir City Hospital, Clinic of Obstetrics and Gynecology, İzmir, Turkey

ABSTRACT

Purpose: To compare the predictive performance of magnetic resonance imaging (MRI)-based and laparoscopy (LS)-based scoring systems for determining surgical resectability after neoadjuvant chemotherapy (NACT) in advanced ovarian cancer initially considered unsuitable for primary cytoreductive surgery.

Methods: This retrospective observational study included consecutive patients with advanced epithelial ovarian, fallopian tube, or primary peritoneal carcinoma treated at a tertiary gynecological oncology center between December 2023 and April 2026. All patients were initially deemed unresectable, received platinum-based NACT, and subsequently underwent post-NACT MRI followed by diagnostic LS for reassessment of surgical feasibility. MRI-peritoneal cancer index (PCI), MRI-Fagotti, LS-PCI, and LS-Fagotti were evaluated. The primary endpoint was successful cytoreduction, defined as complete or optimal interval cytoreductive surgery. Receiver operating characteristic curve analysis was used to assess predictive performance.

Results: A total of 42 patients were included. Complete or optimal cytoreduction was achieved in 25 patients (59.5%), while 17 patients (40.5%) underwent suboptimal cytoreduction or were considered unresectable. MRI-based scores showed limited predictive performance, with area under the curve (AUC) values of 0.595 [95% confidence interval (CI), 0.401-0.782 <9 for MRI-PCI and 0.542 (95% CI, 0.362-0.721) for MRI-Fagotti. In contrast, LS-based scores demonstrated substantially higher predictive accuracy. LS-PCI yielded an AUC of 0.913 (95% CI, 0.810-0.989), whereas LS-Fagotti achieved an AUC of 0.991 (95% CI, 0.966-1.000). An LS-Fagotti cutoff value of ≤8 predicted successful cytoreduction with 96.0% sensitivity and 100.0% specificity.

Conclusion: In patients undergoing reassessment after NACT for advanced ovarian cancer, LS-based scoring systems demonstrated superior predictive performance compared with MRI-based scoring systems. Diagnostic LS remains a valuable tool for selecting candidates for interval cytoreductive surgery and may improve surgical decision-making in the post-NACT setting.

Keywords: Ovarian neoplasms, neoadjuvant therapy, laparoscopy, cytoreduction surgical procedures, magnetic resonance imaging

INTRODUCTION

Epithelial ovarian cancer remains the most lethal gynecological malignancy worldwide and is frequently diagnosed at an advanced stage, characterized by extensive peritoneal dissemination.[1] Despite advances in systemic therapy and surgical techniques, complete cytoreduction remains the most important prognostic factor associated with progression-free and overall survival.[2] Consequently, accurate preoperative assessment of disease extent and surgical resectability is essential for selecting the most appropriate treatment strategy and avoiding unnecessary laparotomies.[3]

For patients presenting with extensive tumor burden or poor performance status, neoadjuvant chemotherapy (NACT) followed by interval debulking surgery has become an accepted alternative to primary cytoreductive surgery.[4]

Address for Correspondence: Süleyman Özen, MD, University of Health Sciences Turkey, İzmir City Hospital, Clinic of Gynecologic Oncology Surgery, İzmir, Turkey
E-mail: drsuleymanozen@gmail.com **ORCID ID:** orcid.org/0000-0002-4870-2840
Received: 03.06.2026 **Accepted:** 17.07.2026 **Publication Date:** 02.09.2026

Cite this article as: Özen S, Güner Özen E, Özuyar Şimşek G, Başkır İşlek, Sancı M. Prediction of surgical resectability after neoadjuvant chemotherapy in advanced ovarian cancer: comparison of mri-based and laparoscopy-based scoring systems. Anat J Obstet Gynecol Res. 2026;3(2):124-130

Several randomized trials have demonstrated comparable survival outcomes between these approaches when optimal cytoreduction can be achieved.[5,6] However, patient selection remains challenging, particularly during post-NACT reassessment, where the decision to proceed with interval surgery depends on the likelihood of achieving complete or optimal cytoreduction.[7]

Various imaging modalities have been investigated to predict surgical resectability in advanced ovarian cancer. Computed tomography, magnetic resonance imaging (MRI), diffusion-weighted MRI, and positron emission tomography/computed tomography have all been used to estimate tumor burden and peritoneal dissemination.[8] More recently, imaging-based adaptations of the peritoneal cancer index (PCI) and Fagotti scoring system have been proposed as objective tools for preoperative surgical triage.[9,10] Although several studies have reported promising diagnostic performance for MRI-derived PCI and radiological Fagotti scores, the accuracy of these approaches after NACT remains uncertain.[11,12] Treatment-related fibrosis, inflammatory changes, and residual microscopic disease may limit the ability of conventional imaging to accurately reflect the true intra-abdominal disease burden.[13,14]

Diagnostic laparoscopy (LS) has emerged as a valuable method for direct assessment of peritoneal disease and remains widely used for evaluating surgical feasibility in advanced ovarian cancer.[11,15] Laparoscopic PCI and the Fagotti predictive index value have demonstrated strong associations with cytoreductive outcomes in both primary and interval surgical settings.[16,17] Nevertheless, data directly comparing MRI-based and LS-based scoring systems in patients undergoing post-NACT reassessment remain limited.[10,18]

The present study was designed to evaluate the ability of MRI-based and LS-based scoring systems to predict successful interval cytoreduction after NACT in patients with advanced ovarian cancer initially deemed unresectable. Specifically, we compared the predictive performance of MRI-PCI, MRI-Fagotti, LS-PCI, and LS-Fagotti scores for identifying patients likely to achieve complete or optimal cytoreduction.

METHODS

Study Design and Patient Selection

This retrospective observational study was conducted at the Department of Gynecologic Oncology, University of Health Sciences Turkey, İzmir City Hospital, Turkey. Consecutive patients with advanced epithelial ovarian, fallopian tube, or primary peritoneal carcinoma who were initially considered unsuitable for primary cytoreductive surgery and subsequently received NACT between December 2023 and April 2026 were screened for eligibility.

Patients were included if they met all of the following criteria: (1) age ≥18 years; (2) histopathologically confirmed epithelial ovarian, fallopian tube, or primary peritoneal carcinoma; (3) Federation of Gynecology and Obstetrics (FIGO) stage IIIC-IVB disease at diagnosis; (4) documentation at multidisciplinary tumor board review of unresectable disease or anticipated inability to achieve complete or optimal primary cytoreduction; (5) administration of platinum-based NACT; (6) availability of post-NACT abdominal and pelvic MRI; (7) diagnostic LS performed after completion of NACT for reassessment of surgical resectability; and (8) complete clinical, operative, and pathological records.

Patients with borderline ovarian tumors, non-epithelial ovarian malignancies, recurrent disease, absence of post-NACT MRI, incomplete imaging studies, missing LS assessment, incomplete operative documentation, or concurrent malignancies affecting treatment strategy were excluded.

The study was conducted in accordance with the ethical principles of the Declaration of Helsinki and was approved by the University of Health Sciences Turkey İzmir City Hospital Non-Interventional Ethics Committee (approval no: 2026/224, date: 08.04.2026). Due to the retrospective nature of the study, the requirement for informed consent was waived by the Ethics Committee.

Magnetic Resonance Imaging Assessment

Following completion of NACT, all patients underwent abdominal and pelvic MRI using a 1.5-T scanner equipped with dedicated phased-array pelvic coils. The imaging protocol included axial and sagittal T1-weighted sequences, multiplanar T2-weighted sequences, diffusion-weighted imaging (DWI) with apparent diffusion coefficient maps, and dynamic contrast-enhanced T1-weighted imaging following intravenous gadolinium administration.

MRI examinations were interpreted in routine institutional practice by radiologists experienced in gynecological oncology imaging. For the present retrospective study, MRI-based PCI and MRI-based Fagotti scores were reconstructed from the finalized MRI examinations and radiology reports. MRI-PCI was calculated according to the Sugarbaker methodology.[19] MRI-Fagotti was adapted from the original predictive index value system and translated into seven MRI-defined domains: peritoneal carcinomatosis (diffuse nodular or miliary peritoneal disease), omental cake (confluent omental tumor or extension toward the greater curvature), diaphragmatic involvement (focal or diffuse diaphragmatic implants), mesenteric disease (mesenteric nodularity, mesenteric root involvement, or retraction), bowel infiltration (serosal nodules or bowel wall involvement), stomach infiltration (gastric wall or serosal involvement), and liver surface metastases (any capsular implant). MRI-Fagotti scoring followed the same 0/2 weighting scheme as the original laparoscopic predictive index value, resulting in a total score ranging from 0 to 14. This radiologic adaptation was based on previously published MRI-based applications of the Fagotti system.[18]

Radiological assessment additionally included ascites, peritoneal implants, omental disease, diaphragmatic involvement, mesenteric infiltration, bowel serosal involvement, liver surface lesions, splenic hilar disease, and lymph node enlargement. Lymph nodes were considered suspicious when the short-axis diameter was ≥10 mm, when irregular margins were present, or when restricted diffusion was observed on DWI sequences. Unfortunately, formal interobserver

agreement analysis could not be performed because separate independent reader-level datasets were not available in this retrospective cohort.

Baseline serum cancer antigen 125 (CA-125) was defined as the value obtained before initiation of NACT. When only a single CA-125 measurement was available, this value was accepted as the baseline level.

Diagnostic Laparoscopy and Surgical Management

All patients underwent diagnostic LS after completion of MRI assessment. LS evaluation was performed by gynecological oncologists to determine the feasibility of interval cytoreductive surgery. Under institutional practice, the decision to proceed to exploratory laparotomy was based on the overall LS reassessment rather than on MRI scores alone.

Disease extent was quantified using LS-PCI and LS-Fagotti. LS-PCI was calculated according to the standard 13-region Sugarbaker classification.[19] LS-Fagotti was determined using the original Predictive Index Value scoring system.[18]

Patients considered suitable candidates for interval cytoreductive surgery proceeded to exploratory laparotomy and cytoreductive surgery. Surgical procedures were individualized according to disease distribution and included total abdominal hysterectomy, bilateral salpingo-oophorectomy, omentectomy, peritonectomy procedures, bowel resection, diaphragmatic stripping, splenectomy, and other upper abdominal procedures, when indicated.

Patients in whom complete or optimal cytoreduction was considered unlikely after LS assessment were classified as having unresectable disease and were managed according to institutional multidisciplinary treatment protocols.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows, version 30.0 (IBM Corp., Armonk, NY, USA). All statistical tests were two-sided, and a p value <0.05 was considered statistically significant.

Continuous variables were assessed for normality using the Shapiro-Wilk test. Normally distributed variables are presented as mean ± standard deviation and were compared using the independent-samples t-test. Non-normally distributed variables are reported as median and interquartile range (IQR) and were compared using the Mann-Whitney U test.

Categorical variables are expressed as frequencies and percentages and were compared using the chi-square test or Fisher's exact test, as appropriate.

Receiver operating characteristic (ROC) curve analysis was performed to evaluate the predictive performance of MRI-PCI, MRI-Fagotti, LS-PCI, and LS-Fagotti for successful cytoreduction. Area under the curve (AUC) with 95% confidence intervals (CIs) were calculated. Optimal cutoff values were determined using the Youden index. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall accuracy were calculated for each model. Comparisons between ROC curves were performed using bootstrap resampling methods. As secondary analyses, complete cytoreduction (R0) was evaluated separately, and multivariable binomial logistic regression was performed to examine whether score-based variables remained associated with favorable cytoreduction after adjustment for age, FIGO stage (IIIC vs stage IV), and log-transformed CA-125. As quasi-complete separation was observed with LS-Fagotti, the parsimonious multivariable model included LS-PCI.

RESULTS

Patient Characteristics

A total of 42 patients met the eligibility criteria and were included in the final analysis. All patients were initially considered unsuitable for primary cytoreductive surgery, received NACT, and subsequently underwent post-treatment MRI followed by diagnostic LS for reassessment of surgical resectability; the median interval between MRI and LS was 9 days (IQR, 6-13 days; range, 2-18 days).

High-grade serous carcinoma represented the predominant histological subtype (35/42, 83.3%), followed by endometrioid carcinoma (4/42, 9.5%) and clear cell carcinoma (3/42, 7.1%). FIGO stage distribution consisted of stage IIIC in 23 patients (54.8%), stage IVA in 7 (16.7%), and stage IVB in 12 (28.6%). Thirty-nine patients (92.9%) were postmenopausal. Baseline clinicopathological and perioperative characteristics are summarized in Table 1.

Complete or optimal cytoreduction was achieved in 25 patients (59.5%), whereas 17 patients (40.5%) underwent suboptimal cytoreduction or were considered unsuitable for cytoreductive surgery. Residual disease assessment demonstrated R0 resection in 15 patients (35.7%), optimal residual disease in 10 (23.8%), suboptimal residual disease in 10 (23.8%), and unresectable disease in 7 patients (16.7%).

Comparison According to Surgical Outcome

Patients in the unfavorable outcome group were significantly older than those in the favorable outcome group [69 years (IQR, 66-75) vs. 55 years (IQR, 52-64), $p=0.004$]. Baseline CA-125 levels were higher in the unfavorable group, although the difference did not reach statistical significance [4575 U/mL (IQR, 3161-5456) vs. 3106 U/mL (IQR, 2062-5441), $p=0.106$].

MRI-derived scores did not differ significantly between outcome groups. Median MRI-PCI was 12 (IQR, 7-14; range, 2-16) in the favorable group and 10 (IQR, 7-18) in the unfavorable group ($p= 0.302$). Median MRI-Fagotti scores were 6 (IQR, 2-6; range, 0-8) and 4 (IQR, 4-8; range, 0-8), respectively ($p=0.646$).

LS-based scores differed significantly according to surgical outcome. Median LS-PCI was 18 (IQR, 15-20) in the favorable group and 24 (IQR, 22-31) in the unfavorable group ($p<0.001$). Median LS-Fagotti scores were 8 (IQR, 8-8) and 12 (IQR, 10-12), respectively ($p<0.001$).

Mean operative time was 335.3±63.5 minutes in the favorable group and 397.1±36.5 minutes in the unfavorable group ($p=0.001$). Estimated blood loss was 1058.0±345.4 mL and 1487.6±262.6 mL, respectively ($p<0.001$).

Predictive Performance of MRI-Based and LS-Based Scores

ROC analysis results are presented in Table 2 and Figure 1. MRI-based scoring systems demonstrated limited predictive performance for successful cytoreduction. MRI-PCI yielded an AUC of 0.595 (95% CI, 0.401-0.782); the optimal cutoff value was ≤16, corresponding to a sensitivity of 100.0% and a specificity of 35.3%. MRI-Fagotti yielded an AUC of 0.542 (95% CI, 0.362-0.721), with an optimal cut-off value of ≤2, providing a sensitivity of 36.0% and a specificity of 82.4%. In contrast, LS-based scoring systems demonstrated superior discriminatory ability. LS-PCI yielded an AUC of 0.913 (95% CI, 0.810-0.989); a cut-off value of ≤20 resulted in a sensitivity of 88.0% and a specificity of 88.2%. LS-Fagotti achieved the highest predictive performance, with an AUC of 0.991 (95% CI, 0.966-1.000); a cut-off value of ≤8 was associated with a sensitivity of 96.0%, specificity of 100.0%, PPV of 100.0%, and NPV of 94.4%. Bootstrap comparison demonstrated that LS-Fagotti significantly outperformed MRI-Fagotti (ΔAUC 0.448, $p<0.001$) and that LS-PCI significantly outperformed MRI-PCI (ΔAUC 0.318, $p=0.015$). LS-Fagotti also showed marginally higher discrimination than LS-PCI (ΔAUC 0.078, $p=0.049$).

On multivariable binomial logistic regression adjusted for age, FIGO stage (IIIC vs stage IV), and log-transformed CA-125, LS-PCI remained independently associated with favorable cytoreduction (adjusted odds ratio (OR) 0.55, 95% CI 0.33-0.94, $p=0.027$); age also remained significant (adjusted OR 0.82, 95% CI 0.69-0.98, $p=0.024$) (Table 3). A corresponding model including LS-Fagotti was not stable because of quasi-complete separation.

Table 1. Baseline characteristics according to cytoreductive outcome

Variable	Overall (n=42)	Complete/optimal (n=25)	Suboptimal/not feasible (n=17)	*p*
Age, years	61.1±11.2	56.4±10.5	67.9±8.4	0.004
CA-125, U/mL	3654.0 (2449.5-5452.2)	3106.0 (2062.0-5441.0)	4575.0 (3161.0-5456.0)	0.106
MRI PCI	11.0 (7.0-14.0)	12.0 (7.0-14.0)	10.0 (7.0-18.0)	0.302
MRI Fagotti	5.0 (2.0-8.0)	6.0 (2.0-6.0)	4.0 (4.0-8.0)	0.646
LS PCI	20.0 (16.0-24.0)	18.0 (15.0-20.0)	24.0 (22.0-31.0)	<0.001
LS Fagotti	8.0 (8.0-10.0)	8.0 (8.0-8.0)	12.0 (10.0-12.0)	<0.001
Operation time, min	360.3±61.8	335.3±63.5	397.1±36.5	0.001
Blood loss, mL	1231.9±377.3	1058.0±345.4	1487.6±262.6	<0.001
FIGO stage: IIIC	23 (54.8%)	14 (56.0%)	9 (52.9%)	0.583
IVA	7 (16.7%)	3 (12.0%)	4 (23.5%)	
IVB	12 (28.6%)	8 (32.0%)	4 (23.5%)	
Histology: high-grade serous carcinoma	35 (83.3%)	20 (80.0%)	15 (88.2%)	0.762
Endometrioid carcinoma	4 (9.5%)	3 (12.0%)	1 (5.9%)	
Clear cell carcinoma	3 (7.1%)	2 (8.0%)	1 (5.9%)	
Menopause: postmenopausal	39 (92.9%)	22 (88.0%)	17 (100.0%)	0.260
Premenopausal	3 (7.1%)	3 (12.0%)	0 (0.0%)	
Bowel resection: yes	21 (50.0%)	12 (48.0%)	9 (52.9%)	1.000
No	21 (50.0%)	13 (52.0%)	8 (47.1%)	
Any major complication: yes	7 (16.7%)	4 (16.0%)	3 (17.6%)	1.000
No	35 (83.3%)	21 (84.0%)	14 (82.4%)	

Data are presented as mean ± standard deviation, median (interquartile range), or n (%), as appropriate
CA-125: Cancer antigen 125, MRI: Magnetic resonance imaging, PCI: Peritoneal cancer index, LS: Laparoscopic

Table 2. Predictive performance of MRI- and laparoscopy-based scores

Score	AUC (95% CI)	Optimal cut-off	Sensitivity, %	Specificity, %	PPV, %	NPV, %	Accuracy, %
MRI-based PCI	0.595 (0.401-0.782)	≤16	100.0	35.3	69.4	100.0	73.8
MRI-based Fagotti	0.542 (0.362-0.721)	≤2	36.0	82.4	75.0	46.7	54.8
Laparoscopic PCI	0.913 (0.810-0.989)	≤20	88.0	88.2	91.7	83.3	88.1
Laparoscopic Fagotti	0.991 (0.966-1.000)	≤8	96.0	100.0	100.0	94.4	97.6

Pairwise bootstrap comparisons of AUCs: LS-Fagotti vs MRI-Fagotti, $p<0.001$; LS-PCI vs MRI-PCI, $p=0.015$; LS-Fagotti vs LS-PCI, $p=0.049$
AUC: Area under the curve, CI: Confidence interval, PPV: Positive predictive value, NPV: Negative predictive value, MRI: Magnetic resonance imaging, PCI: Peritoneal cancer index

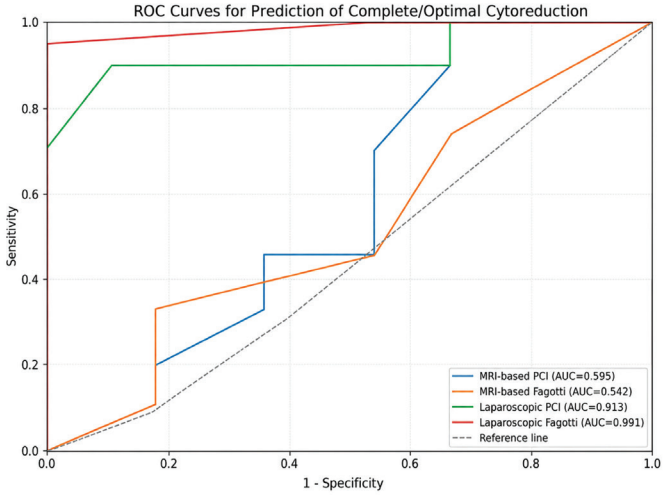


Figure 1. Receiver operating characteristic (ROC) curves of MRI-based and LS-based scoring systems for prediction of complete or optimal cytoreduction after neoadjuvant chemotherapy. LS-Fagotti demonstrated the highest predictive performance (AUC=0.991), followed by LS-PCI (AUC=0.913), whereas MRI-based scores showed limited discriminatory ability

MRI: Magnetic resonance imaging, LS: Laparoscopic, AUC: Area under the curve, PCI: Peritoneal cancer index

Table 3. Multivariable binomial logistic regression for favorable cytoreduction

Variable	Adjusted OR	95% CI	*p*
Age	0.82	0.69-0.98	0.024
FIGO stage IV vs. IIIC	2.11	0.10-44.54	0.630
log CA-125	4.19	0.18-95.99	0.370
LS-PCI	0.55	0.33-0.94	0.027

Model adjusted for age, FIGO stage (IIIC vs. stage IV), log-transformed CA-125, and LS-PCI. A separate LS-Fagotti model was not stable because of quasi-complete separation
OR: Odds ratio, CI: Confidence interval, FIGO: Federation of Gynecology and Obstetrics, PCI: Peritoneal cancer index, LS: Laparoscopic, CA-125: Cancer antigen 125

Secondary Analysis for Complete Cytoreduction (R0)

Complete cytoreduction (R0) was achieved in 15 of 42 patients (35.7%). In the secondary ROC analysis using R0 as the endpoint, MRI-based scores showed limited discriminatory ability. MRI-PCI yielded an AUC of 0.548, with an optimal cutoff value of ≤16, corresponding to 100.0% sensitivity, 22.2% specificity, 41.7% PPV, 100.0% NPV, and 50.0% accuracy. MRI-Fagotti yielded an AUC of 0.507, with an optimal cutoff value of ≤2, corresponding to 33.3% sensitivity, 74.1% specificity, 41.7% PPV, 66.7% NPV, and 59.5% accuracy. In contrast, laparoscopic scores retained higher discriminatory ability. LS-PCI yielded an AUC of 0.759, with an optimal cutoff value of ≤20, corresponding to 86.7% sensitivity, 59.3% specificity, 54.2% PPV, 88.9% NPV, and 69.0% accuracy. LS-Fagotti yielded an AUC of 0.807, with an optimal cut-off value of ≤8, corresponding to 93.3% sensitivity, 63.0% specificity, 58.3% PPV, 94.4% NPV, and 73.8% accuracy.

DISCUSSION

The present study compared the predictive performance of MRI-based and LS-based scoring systems for determining the feasibility of interval cytoreductive surgery after NACT in patients with advanced ovarian cancer who had initially been considered unsuitable for primary cytoreductive surgery. The main finding was that LS assessment showed substantially higher predictive accuracy than MRI-based scoring systems for identifying patients likely to achieve complete or optimal interval cytoreduction.

Diagnostic LS has become an established component of preoperative assessment in advanced ovarian cancer and remains one of the most reliable methods for evaluating resectability.[20] In our cohort, an LS-Fagotti cut-off value of 8 predicted successful cytoreduction with excellent diagnostic performance, consistent with the original predictive index value model described by Fagotti et al.[21] Similarly, the predictive value of LS-PCI observed in our study is in agreement with previous reports demonstrating that PCI values above 20 are associated with incomplete cytoreduction and poorer outcomes.[17,22,23]

In contrast, MRI-based scoring systems showed limited discriminatory ability. MRI-PCI and MRI-Fagotti demonstrated substantially lower predictive performance than their LS counterparts. These findings differ from recent studies reporting high diagnostic accuracy for MRI-derived PCI and radiological Fagotti scores.[12,18] Other investigations incorporating diffusion-weighted MRI and morphologic imaging features have similarly reported strong correlations between imaging-based disease assessment and surgical outcomes.[8,24] Furthermore, prospective multicenter studies have demonstrated that advanced imaging modalities can reliably predict non-resectability in selected patient populations.[10,25]

Several factors may explain the discrepancy between our findings and previous reports. Unlike most published studies, which evaluated treatment-naïve patients, our cohort consisted exclusively of patients reassessed after NACT. Post-treatment fibrosis, inflammatory changes, and chemotherapy-related tissue alterations may obscure the distinction between viable tumor and treatment response on conventional MRI. In addition, small-volume peritoneal implants and miliary carcinomatosis remain challenging to detect radiologically, despite advances in imaging technology.[3,13,26,27] Although emerging techniques such as whole-body diffusion-weighted MRI have shown promising results in the post-NACT setting, these approaches are not routinely available in daily clinical practice [14]. Importantly, the 95% CIs of MRI-PCI and MRI-Fagotti both included an AUC of 0.50, reinforcing the limited standalone discriminatory value of MRI-based prediction in this post-NACT assessment cohort.

Study Limitations

The present study has several limitations. Its retrospective design, single-center setting, and relatively limited sample size may restrict the generalizability of the findings. Diagnostic LS was not only evaluated as a predictive tool but also contributed to the decision to proceed to exploratory laparotomy. Therefore,

a degree of incorporation bias may have favored LS-based models. In addition, separate reader-level MRI datasets were not available, precluding formal interobserver agreement analysis. The near-perfect discrimination observed for LS-Fagotti should also be interpreted cautiously, as the small sample size and single-center design may have contributed to overestimation. Nevertheless, the study evaluated a homogeneous post-NACT population and provided a direct comparison between radiological and laparoscopic scoring systems within the same cohort, thereby offering clinically relevant data for surgical decision-making.

CONCLUSION

In patients with advanced ovarian cancer undergoing reassessment after NACT, laparoscopic scoring systems demonstrated superior predictive performance compared with MRI-based scoring systems for determining the feasibility of successful interval cytoreduction. LS-Fagotti and LS-PCI were strongly associated with surgical outcome, whereas MRI-derived scores showed limited predictive value in the post-NACT setting. These findings support the continued use of diagnostic LS as an important component of preoperative evaluation before interval debulking surgery.

Ethics

Ethics Committee Approval: Approved by the University of Health Sciences Turkey İzmir City Hospital Non-Interventional Ethics Committee (approval no: 2026/224, date: 08.04.2026).

Informed Consent: The requirement for informed consent was waived due to the retrospective design of the study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: S.Ö., G.Ö.Ş., İ.B.İ., M.S., Concept: S.Ö., E.G.Ö., Design: S.Ö., E.G.Ö., Data Collection or Processing: S.Ö., E.G.Ö., İ.B.İ., M.S., Analysis or Interpretation: S.Ö., G.Ö.Ş., Literature Search: S.Ö., E.G.Ö., G.Ö.Ş., Writing: S.Ö., E.G.Ö., M.S.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

Declaration on the Use of Artificial Intelligence (AI): No artificial intelligence tools were used in the preparation of this manuscript.

REFERENCES

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229-263.
2. Vergote I, Coens C, Nankivell M, et al. Neoadjuvant chemotherapy versus debulking surgery in advanced tubo-ovarian cancers: pooled analysis of individual patient data from the EORTC 55971 and CHORUS trials. Lancet Oncol. 2018;19(12):1680-1687. Erratum in: Lancet Oncol. 2019;20(1):e10.
3. Miceli V, Gennarini M, Tomao F, et al. Imaging of Peritoneal Carcinomatosis in Advanced Ovarian Cancer: CT, MRI, radiomic features and resectability criteria. Cancers (Basel). 2023;15(24):5827.
4. Kehoe S, Hook J, Nankivell M, et al. Primary chemotherapy versus primary surgery for newly diagnosed advanced ovarian cancer (CHORUS): an open-label, randomised, controlled, non-inferiority trial. Lancet. 2015;386(9990):249-257.
5. Reuss A, du Bois A, Harter P, et al. TRUST: trial of radical upfront surgical therapy in advanced ovarian cancer (ENGOT ov33/AGO-OVAR OP7). Int J Gynecol Cancer. 2019;29(8):1327-1331.
6. Classe JM, Ferron G, Ouldamer L, et al. CHRONO: randomized trial of the CHROnology of surgery after neoadjuvant chemotherapy for ovarian cancer. Int J Gynecol Cancer. 2022;32(8):1071-1075.
7. Coleridge SL, Bryant A, Kehoe S, Morrison J. Chemotherapy versus surgery for initial treatment in advanced ovarian epithelial cancer. Cochrane Database Syst Rev. 2021;2(2):CD005343.
8. Li H, Lu J, Deng L, et al. Diffusion-weighted magnetic resonance imaging and morphological characteristics evaluation for outcome prediction of primary debulking surgery for advanced high-grade serous ovarian carcinoma. J Magn Reson Imaging. 2023;57(5):1340-1349.
9. Moro F, Pinto P, Chiappa V, et al. Prediction of nonresectability using the updated predictive index value model assessed by imaging and surgery in tubo-ovarian cancer: a prospective multicenter ISAAC study. Am J Obstet Gynecol. 2024;231(6):632.e1-632.e14.
10. Fischerova D, Pinto P, Alcázar Zambrano JL, et al.; Collaborators. Comparison of ultrasound with computed tomography and whole-body diffusion-weighted MRI in prediction of surgical outcome using ESMO-ESGO criteria in patients with tubo-ovarian carcinoma: prospective ISAAC study. Ultrasound Obstet Gynecol. 2026;67(2):207-219.
11. Fagotti A, Ferrandina G, Fanfani F, et al. Prospective validation of a laparoscopic predictive model for optimal cytoreduction in advanced ovarian carcinoma. Am J Obstet Gynecol. 2008;199(6):642.e1-6.
12. Konidari M, Gourtsoyianni S, Thomakos N, et al. Prediction of resectability of peritoneal disease in ovarian cancer patients using the peritoneal cancer index (PCI) and Fagotti score on MRI. Cancers (Basel). 2026;18(1):165.
13. Tsili AC, Alexiou G, Tzoumpa M, Siempis T, Argyropoulou MI. Imaging of peritoneal metastases in ovarian cancer using MDCT, MRI, and FDG PET/CT: a systematic review and meta-analysis. Cancers (Basel). 2024;16(8):1467.
14. Vandecaveye V, Dresen RC, Baert T, et al. Value of whole-body diffusion-weighted MRI for the prediction of surgical and clinical outcome after neoadjuvant chemotherapy in advanced ovarian cancer. Br J Cancer. 2026;134(11):1605-1613.
15. Dalaty M, Nicodin MA, Nicodin OV, PuşcaÃ Iu L, Căpîlnă ME. The role of laparoscopic triage in the management of advanced ovarian cancer: a review of literature. Chirurgia (Bucur). 2025;120(5):511-518.
16. Saner FAM, Ruggeri G, Siegenthaler F, Wampfler J, Imboden S, Mueller MD. Change of Fagotti score is associated with outcome after neoadjuvant chemotherapy for ovarian cancer. Int J Gynecol Cancer. 2023;33(10):1595-1601.
17. Climent MT, Serra A, Gilabert-Estellés J, Gilabert-Aguilar J, Llueca A. Comparison of peritoneal carcinomatosis scoring methods in predicting resectability and prognosis in gynecologic malignancies. J Clin Med. 2021;10(12):2553.
18. Siegenthaler F, Zurbriggen S, Wyss C, et al. Diagnostic performance of a radiological Fagotti score assessed on diffusion-weighted magnetic resonance imaging for predicting tumor resectability in ovarian cancer patients: a feasibility study. Front Oncol. 2025;15:1680992.
19. Harmon RL, Sugarbaker PH. Prognostic indicators in peritoneal carcinomatosis from gastrointestinal cancer. Int Semin Surg Oncol. 2005;2(1):3.

20. Rutten MJ, Gaarenstroom KN, Van Gorp T, et al. Laparoscopy to predict the result of primary cytoreductive surgery in advanced ovarian cancer patients (LapOvCa-trial): a multicentre randomized controlled study. BMC Cancer. 2012;12:31.
21. Fagotti A, Ferrandina G, Fanfani F, et al. A laparoscopy-based score to predict surgical outcome in patients with advanced ovarian carcinoma: a pilot study. Ann Surg Oncol. 2006;13(8):1156-1161.
22. Fagan PJ, Gomes N, Heath OM, et al. The peritoneal cancer index as a predictor of complete cytoreduction at primary and interval cytoreductive surgery in advanced ovarian cancer. Int J Gynecol Cancer. 2023;33(11):1757-1763.
23. Ashmore AA, Al-Majali J, Chui SK, et al. The peritoneal cancer index as a predictor of cytoreductive surgery outcomes and heatmapping of ovarian cancer distribution: a retrospective analysis. Cancers (Basel). 2025;17(17):2790.
24. Engbersen MP, Van' T Sant I, Lok C, et al. MRI with diffusion weighted imaging to predict feasibility of complete cytoreduction with the peritoneal cancer index (PCI) in advanced stage ovarian cancer patients. Eur J Radiol. 2019:146-151.
25. Pinto P, Moro F, Alcázar JL, et al. Prediction of non-resectability in tubo-ovarian cancer patients using peritoneal cancer index - a prospective multicentric study using imaging (ISAAC study). Gynecol Oncol. 2024;191:132-142.
26. Kandemir H, Sözen H, Kartal MG, Özkan ZG, Topuz S, Salihoğlu MY. An assessment of the effectiveness of preoperative imaging modalities (MRI, CT, and 18F-FDG PET/CT) in determining the extent of disease spread in epithelial ovarian-tubal-peritoneal cancer (EOC). Medicina (Kaunas). 2025;61(2):199.
27. Meucci R, Prosperi D, Lauri C, et al. Peritoneal carcinomatosis of malignant gynecological origin: a systematic review of imaging assessment. J Clin Med. 2024;13(5):1254.

Original Article

DOI: 10.4274/anajog.galenos.2026.55265

Anat J Obstet Gynecol Res 2026;3(2):131-138

Serial Progesterone Kinetics after Methotrexate Treatment for Persistent Pregnancy of Unknown Location and Presumed Ectopic Pregnancy: A Retrospective Cohort Study

Cemal Ünlü[1], Hüseyin Aksoy[2], Harika Göçer[2], Merve Genco[2], Emrah Göçer[2], Gonca Türker Ergün[3]

[1]Private Clinic, Clinic of Obstetrics and Gynecology, Kayseri, Turkey
[2]University of Health Sciences Turkey, Kayseri City Hospital, Clinic of Obstetrics and Gynecology, Kayseri, Turkey
[3]University of Health Sciences Turkey, Ankara Bilkent City Hospital, Clinic of Obstetrics and Gynecology, Ankara, Turkey

ABSTRACT

Purpose: Medical treatment with methotrexate (MTX) is an established option for some patients with ectopic pregnancies or persistent pregnancy of unknown location (PUL). This study evaluated whether serial progesterone kinetics provide clinically useful information beyond beta-human chorionic gonadotropin (β-hCG) kinetics during early follow-up after MTX treatment for persistent PUL or presumed ectopic pregnancy.

Methods: In this retrospective cohort study, hemodynamically stable patients treated with systemic MTX between 1 January 2025 and 1 September 2025 were screened. Serum β-hCG and progesterone were measured on day 1, day 4, and day 7, with day 1 defined as the day of MTX administration. The primary outcome was insufficient early biochemical response, defined as a β-hCG decrease of less than 15% from day 4 to day 7. The main secondary outcome was actual second-dose MTX administration.

Results: Among 73 assessed patients, 50 (68.5%) met the inclusion criteria. Seven (14%) patients had sonographic findings compatible with tubal ectopic pregnancy, and 43 (86%) were managed as persistent PUL or presumed ectopic pregnancy. Twelve patients (24.0%) had an insufficient day 4-day 7 β-hCG response, and 9 (18.0%) received a second MTX dose. Day 7 β-hCG was higher in patients with insufficient early response, whereas day 1, day 4, and day 7 progesterone concentrations did not differ significantly between response groups. In exploratory receiver operating characteristic analysis for actual second-dose MTX administration, day 7 β-hCG and lower day 4-day 7 β-hCG decline showed greater discrimination than progesterone-derived variables. Five clinically stable patients with subthreshold β-hCG decline refused additional MTX and achieved biochemical cure with close follow-up.

Conclusion: In this predominantly PUL/presumed ectopic pregnancy cohort, early β-hCG kinetics appeared more informative than serial progesterone kinetics for identifying insufficient response and second-dose MTX administration. Serial progesterone measurement was feasible but showed limited additional discriminatory value. These preliminary findings require validation in larger prospective cohorts.

Keywords: Ectopic pregnancy, methotrexate, pregnancy of unknown location, progesterone, beta-human chorionic gonadotropin

INTRODUCTION

Medical treatment with methotrexate (MTX) is an established option for hemodynamically stable patients with selected ectopic pregnancies when close follow-up is feasible.[1] Early studies of the single-dose MTX regimen evaluated patients with unruptured ectopic pregnancy and established this approach as a conservative treatment option in appropriately selected cases.[2,3] Single-dose MTX has also been evaluated in selected patients with persistent pregnancy of unknown location (PUL).[4] Comparative evidence and larger clinical series have also supported its surgery-sparing role in appropriate patients.[5,6] Safe use requires structured biochemical surveillance.[1]

Address for Correspondence: Cemal Ünlü, MD, Private Clinic, Clinic of Obstetrics and Gynecology, Kayseri, Turkey
E-mail: cunlu06@gmail.com **ORCID ID:** orcid.org/0009-0008-3242-025X
Received: 22.05.2026 **Accepted:** 06.07.2026 **Publication Date:** 02.09.2026

Cite this article as: Ünlü C, Aksoy H, Göçer H, Genco M, Göçer E, Türker Ergün G. Serial progesterone kinetics after methotrexate treatment for persistent pregnancy of unknown location and presumed ectopic pregnancy: a retrospective cohort study. Anat J Obstet Gynecol Res. 2026;3(2):131-138

Post-MTX follow-up relies mainly on serial serum beta-human chorionic gonadotropin (β-hCG). In single-dose protocols, β-hCG is measured on day 1, day 4, and day 7. A decrease of at least 15% from day 4 to day 7 is accepted as indicating an adequate early response, whereas a smaller decrease, plateau, or increase prompts reassessment and possible additional treatment.[2,3,6] Predictive and validation studies support this day 4-day 7 β-hCG rule.[7,8] The National Institute for Health and Care Excellence recommends β-hCG measurements on days 4 and 7 after MTX, followed by weekly monitoring until a negative result; plateauing or rising β-hCG should prompt review.[1]

Although useful, the day 4-day 7 β-hCG rule is not a direct biological measure of trophoblastic regression. An adequate decrease has high positive predictive value for treatment success, but a subthreshold decrease does not always exclude cure without further MTX.[7,8] Alternative early β-hCG patterns and variation in time to resolution after MTX also support interpreting early β-hCG kinetics within the broader clinical course.[9-11] Additional biomarkers may therefore be relevant in stable patients with an intermediate response.

Progesterone differs biologically from β-hCG and has been studied as a marker of early pregnancy viability, pregnancy failure, and MTX response.[12,13] Earlier studies suggested that pretreatment progesterone may be associated with MTX success and may fall to low levels during ectopic pregnancy resolution.[12,13] However, serial progesterone kinetics measured at the same follow-up points as β-hCG remain poorly defined.

This study evaluated serial progesterone concentrations measured concurrently with β-hCG after MTX treatment in a cohort of patients with persistent PUL and presumed ectopic pregnancy. The primary objective was to assess whether early progesterone kinetics were correlated with insufficient day 4-day 7 β-hCG response. Secondary objectives were to compare progesterone patterns by second-dose MTX administration and to describe clinically stable patients with subthreshold β-hCG decline who achieved biochemical cure after declining additional MTX.

METHODS

Study Design and Setting

This retrospective cohort study was conducted in the Department of Obstetrics and Gynecology of a tertiary referral hospital in Turkey. Patients managed with systemic MTX for persistent PUL or presumed ectopic pregnancy between 1 January 2025 and 1 September 2025 were retrospectively identified from the electronic medical records and re-reviewed.

The study was approved by the University of Health Sciences Turkey, Kayseri City Hospital Non-Interventional Clinical Research Ethics Committee (approval no: 634, date: 04.11.2025), and the requirement for informed consent was waived because anonymized retrospective data were used.

Patients were eligible if they were hemodynamically stable, were initially managed with systemic MTX for persistent PUL or presumed ectopic pregnancy, and had serum β-hCG and progesterone measurements on day 1, day 4, and day 7. Day 1 was defined as the day of MTX administration. In this manuscript, day 1 refers to the calendar day on which MTX was administered; this corresponds to day 0 in some single-dose MTX protocols. We retained the day 1 terminology because it reflected the nomenclature used in our institutional records and follow-up forms. The cohort included patients with sonographic findings compatible with tubal ectopic pregnancy and patients with persistent PUL managed as presumed ectopic pregnancy. Patients were excluded if they were hemodynamically unstable, had suspected or confirmed rupture, underwent primary surgery, had ectopic fetal cardiac activity, had a contraindication to MTX, lacked day 1, day 4, or day 7 β-hCG/progesterone data, or were lost to follow-up before cure. Cesarean scar, cervical, interstitial or cornual, ovarian, and heterotopic pregnancies were excluded.

Diagnosis and treatment decisions were based on symptoms, serial β-hCG findings, transvaginal ultrasonography and, when indicated, β-hCG response after uterine evacuation. Persistent PUL was defined as no intrauterine or extrauterine gestational sac on transvaginal ultrasonography with an abnormal β-hCG pattern. After diagnostic uterine evacuation, persistent PUL or presumed ectopic pregnancy was diagnosed when the β-hCG decrease was less than 15% at follow-up. Complete biochemical cure was defined as serum β-hCG below 10 mIU/mL.

Treatment and Follow-Up Protocol

MTX was administered intramuscularly at 50 mg/m^2. Body surface area was calculated using the Mosteller formula.[14] Serum β-hCG and progesterone were measured on day 1, day 4, and day 7. After day 7, patients were followed weekly with serum β-hCG measurement until biochemical cure. Progesterone was also measured at the same available follow-up points as β-hCG. Measurements were performed in the hospital central laboratory using electrochemiluminescence immunoassays on a Cobas 8000 modular analyzer series (Roche Diagnostics, Mannheim, Germany). β-hCG was reported as mIU/mL and progesterone as ng/mL. A previously published analytical process evaluation from the same hospital central laboratory reported two-level internal quality-control coefficients of variation of 3.68% and 3.21% for β-hCG and 3.10% and 4.47% for progesterone.[15] Patient analyte concentration values were extracted from electronic medical records.

An adequate early biochemical response was defined as at least a 15% decrease in β-hCG from day 4 to day 7. A smaller decrease, plateau, or increase was considered insufficient and prompted reassessment. A second MTX dose at the same body-surface-area-based dose was recommended based on clinical judgement in patients with insufficient early response and was also considered when an initially adequate day 4-day 7 decrease was followed by plateauing or rising β-hCG during weekly follow-up. Final management incorporated hemodynamic stability, symptoms, ultrasonographic findings when available, follow-up reliability, and patient preference. Clinically stable patients

who declined a recommended second dose were followed closely with serial clinical and biochemical assessment.

Variables and Outcomes

Baseline variables were age, gravidity, parity, abortus, previous ectopic pregnancy, previous tubal surgery, assisted reproduction, intrauterine device use, vaginal bleeding, abdominal or pelvic pain, admission hemoglobin level, MTX dose, body weight, height, body surface area, day 1 β-hCG, and day 1 progesterone.

The primary outcome was insufficient early biochemical response, defined as a β-hCG decrease of less than 15% from day 4 to day 7. The main secondary outcome was actual second-dose MTX administration. A descriptive subgroup analysis included clinically stable patients with insufficient early response who declined a recommended second MTX dose and later achieved biochemical cure with close surveillance.

Progesterone variables were day 1, day 4, and day 7 concentrations; percentage changes from day 1 to day 4 and from day 4 to day 7; and day 7 progesterone below 1.5 ng/mL. Percentage change was calculated as the later value minus the earlier value divided by the earlier value and multiplied by 100; negative values indicate a decrease. The 1.5 ng/mL threshold was selected because previous work reported this level as associated with resolution of ectopic gestation.[13]

Statistical Analysis

Continuous variables are summarized as median and interquartile range (IQR), and categorical variables as number and percentage. Due to the small sample size and non-normal biomarker distributions, Mann-Whitney U and Fisher's exact tests were used. Spearman correlation assessed associations between β-hCG and progesterone at corresponding time points and between early percentage changes.

Exploratory receiver operating characteristic (ROC) analyses were restricted to actual second-dose MTX administration. ROC analysis was not performed for the primary outcome using day 4-day 7 β-hCG decline or closely related variables, because the primary outcome was defined by this same interval change. Area under the curve (AUC) values were reported with bootstrap 95% confidence intervals using 2000 resamples. ROC findings were interpreted only as exploratory and hypothesis-generating because actual second-dose MTX administration was a management-dependent endpoint. Multivariable regression was not performed because the limited number of outcome events would make adjusted estimates unstable. Across Tables 1 and 2, 44 unadjusted group comparisons were performed. No adjustment was made for multiple comparisons; therefore, isolated statistically significant findings from unadjusted analyses should be interpreted cautiously because of the risk of false-positive results. Two-sided *p* values below 0.05 were considered statistically significant in unadjusted analyses.

RESULTS

During the study period, 73 potentially eligible patients were identified. Eight were excluded because they underwent primary surgical management without MTX, seven because day 1, day 4, or day 7 β-hCG/progesterone data were missing, and eight because they discontinued follow-up before biochemical cure. The final cohort included 50 (68.5%) patients (Figure 1).

Of these 50, seven had sonographic findings compatible with tubal ectopic pregnancy, while 43 were managed as persistent PUL or presumed ectopic pregnancy; six of these were diagnosed after an insufficient β-hCG decrease following uterine evacuation. Nine patients (18.0%) received a second MTX dose, and 41 (82.0%) achieved biochemical cure after a single dose. Twelve patients (24.0%) had an insufficient day 4-day 7 β-hCG response; seven received a second MTX dose, whereas five clinically stable patients declined additional MTX and achieved biochemical cure with close follow-up. The remaining two second-dose patients initially had an adequate day 4-day 7 β-hCG decrease but later developed plateauing or rising β-hCG during weekly follow-up. Among the final 50 patients followed until biochemical cure, no patient developed tubal rupture, hemodynamic instability, or needed emergency surgery or hospitalization for clinical deterioration during post-MTX follow-up.

Baseline and early biomarker characteristics according to the primary outcome are shown in Table 1. Patients with insufficient early β-hCG response had higher day 7 β-hCG values than those with adequate response. Day 1, day 4, and day 7 progesterone concentrations were not significantly different between groups. Both absolute and percentage progesterone changes are presented in Table 1. Day 1-day 4 progesterone percentage change differed between groups, with a greater early progesterone decrease in patients with adequate β-hCG response; however, this exploratory finding should be interpreted cautiously because multiple unadjusted comparisons were performed.

Baseline characteristics and early β-hCG/progesterone kinetics according to actual second-dose MTX administration are presented in Table 2. Median age, baseline day 1 β-hCG, and day 1 progesterone were similar between the single-dose cure and second-dose groups. Abortus count differed between groups, but this isolated finding should be interpreted cautiously because of the small second-dose group and exploratory comparisons. The second-dose group had higher day 7 β-hCG values and a markedly lower day 4-day 7 β-hCG decline. In contrast, progesterone concentrations and early progesterone percentage changes did not differ significantly between groups.

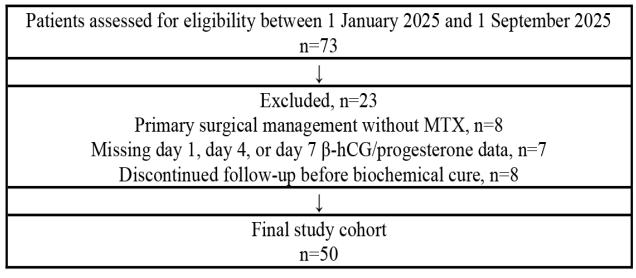


Figure 1. Flow diagram of patient selection

β-hCG: Beta-human chorionic gonadotropin, MTX: Methotrexate

Table 1. Baseline and early biomarker characteristics according to early β-hCG response

Variable	Adequate early β-hCG response (n=38)	Insufficient early β-hCG response (n=12)	*p* value
Age, years	32 (25-33)	28 (25-34)	0.802
Gravidity	3 (2-3)	2 (2-4)	0.735
Parity	1 (0-2)	1 (0-2)	0.629
Abortus	0 (0-1)	0 (0-1)	0.430
Admission hemoglobin, g/dL	12.6 (12.1-13.3)	12.6 (11.7-13.3)	0.682
MTX dose, mg	85 (81-92)	84 (77-86)	0.198
Body surface area, m^2	1.74 (1.63-1.89)	1.72 (1.57-1.85)	0.481
Day 1 β-hCG, mIU/mL	730 (112-1912)	838 (472-2094)	0.329
Day 4 β-hCG, mIU/mL	420 (97-1703)	1071 (576-2238)	0.143
Day 7 β-hCG, mIU/mL	111 (44-736)	1206 (642-2408)	0.002
Day 1 progesterone, ng/mL	3.73 (1.00-8.10)	3.91 (2.80-4.71)	0.982
Day 4 progesterone, ng/mL	2.41 (0.76-6.62)	4.22 (2.42-6.09)	0.358
Day 7 progesterone, ng/mL	1.80 (0.33-4.29)	2.97 (1.70-4.05)	0.195
Day 1-day 4 progesterone absolute change, ng/mL	-0.40 (-2.42 to 0.71)	0.03 (-0.78 to 1.19)	0.192
Day 4-day 7 progesterone absolute change, ng/mL	-0.65 (-2.37 to 0.11)	-0.92 (-1.72 to -0.47)	0.420
Day 1-day 4 progesterone change, %	-44.8 (-70.8 to 16.3)	2.0 (-17.2 to 29.9)	0.042
Day 4-day 7 progesterone change, %	-31.7 (-55.7 to 25.9)	-29.1 (-36.6 to -19.6)	0.776
Day 7 progesterone <1.5 ng/mL	17/38 (44.7)	2/12 (16.7)	0.100
Actual second-dose MTX	2/38 (5.3)	7/12 (58.3)	<0.001

Data are median (IQR) or n/N (%). *p* values were calculated with the Mann-Whitney U test or Fisher's exact test. Progesterone absolute change was calculated as the later value minus the earlier value. Progesterone percentage change was calculated as the later value minus the earlier value divided by the earlier value and multiplied by 100; negative values indicate a decrease from the earlier to the later time point
IQR: Interquartile range, MTX: Methotrexate, β-hCG: Beta-human chorionic gonadotropin

Table 3 presents three clinical subgroups: adequate early β-hCG response with single-dose cure; insufficient response with second-dose refusal and subsequent cure; and actual second-dose MTX. The observed insufficient-response subgroup had higher day 7 β-hCG than the adequate-response single-dose group but still achieved biochemical cure without further MTX. Day 7 progesterone was not uniformly low in this subgroup.

Exploratory ROC analyses for actual second-dose MTX administration are shown in Table 4. Day 7 β-hCG showed moderate discrimination, whereas a smaller or negative day 4-day 7 β-hCG decline showed good discrimination for actual second-dose MTX administration. In contrast, progesterone-derived variables showed weak or limited discrimination. Once again, these ROC findings should be interpreted cautiously because only nine (18%) patients received a second MTX dose and because second-dose administration was influenced by clinical judgment, follow-up findings, and patient preference rather than by biomarker values alone.

Spearman analysis showed moderate positive correlations between β-hCG and progesterone measured at the same time points on day 4 (rho=0.52, $p<0.001$) and day 7 (rho=0.55, $p<0.001$), whereas the day 1 correlation was weak and was not statistically significant in unadjusted analysis (rho=0.27, $p=0.055$). In contrast, day 4-day 7 β-hCG decline was not meaningfully correlated with day 4-day 7 progesterone change (rho=0.11, $p=0.459$). day 7 β-hCG was also not meaningfully correlated with day 4-day 7 progesterone change (rho=0.07, $p=0.641$).

DISCUSSION

In this retrospective cohort of 50 patients treated with MTX for persistent PUL or presumed ectopic pregnancy, early β-hCG kinetics appeared more informative than serial progesterone measurements for identifying patients with insufficient early biochemical response and those who ultimately received a second MTX dose. Although a greater day 1-day 4 progesterone decline was observed in patients with an adequate day 4-day 7 β-hCG response, progesterone concentrations at days 1, 4, and 7 and the day 4-day 7 progesterone change showed limited discriminatory performance overall. In practical terms, these findings suggest that progesterone may reflect some aspects of endocrine regression after MTX, but it does not currently provide decision-making value comparable to β-hCG during routine post-treatment surveillance.

Several physiological mechanisms may explain why progesterone and β-hCG kinetics were not closely parallel after MTX. β-hCG is produced by trophoblastic tissue and is therefore more directly linked to trophoblastic burden and post-treatment trophoblastic regression. Progesterone, in contrast, reflects corpus luteum function as well as pregnancy viability, and its secretion during early pregnancy

Table 2. Baseline characteristics and early β-hCG/progesterone kinetics according to actual second-dose MTX administration

Variable	Single-dose cure (n=41)	Second-dose MTX (n=9)	*p*
Age, years	31 (25-33)	33 (25-34)	0.310
Gravidity	3 (2-3)	3 (2-4)	0.217
Parity	1 (0-2)	1 (0-2)	0.865
Abortus	0 (0-1)	1 (0-2)	0.019
Admission hemoglobin, g/dL	12.6 (11.9-13.3)	12.8 (12.0-13.7)	0.527
MTX dose, mg	84 (81-92)	84 (74-86)	0.210
Body weight, kg	65 (61-80)	68 (60-72)	0.527
Body surface area, m^2	1.70 (1.62-1.90)	1.74 (1.65-1.81)	0.640
Day 1 β-hCG, mIU/mL	785 (133-1892)	676 (595-2169)	0.189
Day 1 progesterone, ng/mL	3.66 (1.30-8.11)	4.29 (2.99-5.23)	0.743
Previous ectopic pregnancy	6/41 (14.6)	2/9 (22.2)	0.623
Previous tubal surgery	3/41 (7.3)	0/9 (0.0)	1.000
Assisted reproduction/IVF	6/41 (14.6)	1/9 (11.1)	1.000
Intrauterine device use	4/41 (9.8)	0/9 (0.0)	1.000
Vaginal bleeding	29/41 (70.7)	9/9 (100.0)	0.092
Abdominal/pelvic pain	32/41 (78.0)	9/9 (100.0)	0.183
Day 4 β-hCG, mIU/mL	520 (110-1716)	878 (674-2239)	0.207
Day 7 β-hCG, mIU/mL	157 (44-739)	1041 (770-2409)	0.016
Day 4-day 7 β-hCG decline, %	50.8 (26.1-59.9)	-7.6 (-15.9 to 12.5)	<0.001
Insufficient day 4-day 7 β-hCG response (<15% decline)	5/41 (12.2)	7/9 (77.8)	<0.001
Day 4 progesterone, ng/mL	3.20 (0.80-7.62)	4.10 (2.12-5.90)	1.000
Day 7 progesterone, ng/mL	2.45 (0.47-5.50)	2.49 (1.64-3.85)	0.950
Day 1-day 4 progesterone change, %	-31.9 (-68.3 to 26.3)	0.7 (-35.6 to 18.9)	0.350
Day 4-day 7 progesterone change, %	-30.5 (-52.0 to 2.7)	-34.7 (-38.4 to -26.5)	0.900
Day 7 progesterone <1.5 ng/mL	17/41 (41.5)	2/9 (22.2)	0.452

Data are median (IQR) or n/N (%). *p* values were calculated with the Mann-Whitney U test or Fisher's exact test. Negative β-hCG decline values indicate an increase from day 4 to day 7. Two patients in the second-dose MTX group had an adequate day 4-day 7 β-hCG decrease but received a second dose because of plateauing or rising β-hCG during subsequent weekly follow-up
IQR: Interquartile range, MTX: Methotrexate, β-hCG: Beta-human chorionic gonadotropin, IVF: In vitro fertilization

Table 3. Clinical subgroups according to early β-hCG response and treatment course

Variable	Adequate response with single-dose cure (n=36)	Insufficient response, observed after second-dose refusal (n=5)	Actual second-dose MTX (n=9)
Day 1 β-hCG, mIU/mL	628 (103-1899)	1040 (340-1083)	676 (595-2169)
Day 4 β-hCG, mIU/mL	420 (93-1582)	1264 (283-1716)	878 (674-2239)
Day 7 β-hCG, mIU/mL	111 (44-733)	1371 (257-1479)	1041 (770-2409)
Day 4-day 7 β-hCG decline, %	52.6 (42.2-64.3)	9.2 (7.6-10.7)	-7.6 (-15.9 to 12.5)
Day 1 progesterone, ng/mL	3.73 (1.08-8.18)	3.53 (2.22-4.54)	4.29 (2.99-5.23)
Day 4 progesterone, ng/mL	2.41 (0.79-6.62)	4.46 (3.20-8.59)	4.10 (2.12-5.90)
Day 7 progesterone, ng/mL	1.80 (0.33-4.67)	4.03 (2.75-7.55)	2.49 (1.64-3.85)
Day 4-day 7 progesterone change, %	-31.7 (-57.8 to 17.8)	-14.1 (-21.4 to -12.1)	-34.7 (-38.4 to -26.5)
Day 7 progesterone <1.5 ng/mL	16/36 (44.4)	1/5 (20.0)	2/9 (22.2)

Data are median (IQR) or n/N (%). The observed insufficient-response subgroup included clinically stable patients who declined an additional MTX dose and achieved complete biochemical cure with close surveillance
IQR: Interquartile range, MTX: Methotrexate, β-hCG: Beta-human chorionic gonadotropin

Table 4. Exploratory ROC analysis for actual second-dose MTX administration

Predictor	AUC (95% CI)
Day 4 β-hCG	0.64 (0.46-0.80)
Day 7 β-hCG	0.76 (0.57-0.91)
Lower day 4-day 7 β-hCG decline	0.88 (0.67-1.00)
Day 7 progesterone	0.51 (0.33-0.67)
Higher day 1-day 4 progesterone change	0.60 (0.43-0.78)
Higher day 4-day 7 progesterone change	0.51 (0.35-0.68)

AUC values are exploratory because of the small number of second-dose MTX events. Progesterone change was calculated as the later value minus the earlier value divided by the earlier value and multiplied by 100; higher values therefore indicate less decline or an increase
MTX: Methotrexate, β-hCG: Beta-human chorionic gonadotropin, ROC: Receiver operating characteristic, AUC: Area under the curve, CI: Confidence interval

is supported by hCG. Thus, progesterone is not biologically independent of β-hCG. Nevertheless, serum progesterone has been reported to reach low levels earlier than β-hCG during ectopic pregnancy resolution.[13] Previous studies suggesting prognostic value for progesterone provided the rationale for evaluating whether its serial kinetics could add information to routine β-hCG monitoring.[12,13] In the present study, however, serial progesterone changes showed limited additional discriminatory value, supporting β-hCG as the more informative marker for post-MTX surveillance.

The most important comparator for the present study is the prospective Human Reproduction study by Kirk et al.,[8] which remains one of the most informative post-MTX biomarker studies in this field. In that study, serum β-hCG and progesterone were measured on days 1, 3, 4, 5, and 7 in 69 women treated with single-dose MTX, and the conventional day 4-day 7 β-hCG rule correctly predicted outcome in 90.3% of cases, with sensitivity 93.0% and specificity 84.2%.[8] Importantly, Kirk et al.[8] also showed that progesterone concentrations at several follow-up time points differed between successful and unsuccessful outcomes, whereas progesterone changes between days did not. Our findings are directionally consistent with that observation. In our cohort, progesterone showed only a limited exploratory signal, mainly in the day 1-day 4 interval, but serial progesterone kinetics did not outperform β-hCG and did not clearly distinguish patients who required additional MTX. Taken together, the two studies suggest that progesterone may have a biological association with the treatment course, yet its serial changes appear less robust than β-hCG kinetics for clinical decision-making after MTX.

The comparison with Kirk et al.[8] is also informative because our cohort differs in clinically relevant ways. Their population included tubal ectopic pregnancies, interstitial ectopic pregnancies, persisting PUL, and persistent trophoblastic tissue, whereas most patients in our cohort were managed as persistent PUL or presumed ectopic pregnancy rather than sonographically confirmed tubal ectopic pregnancy. In addition, our analyses focused not only on inadequate day 4-day 7 β-hCG response but also on actual second-dose MTX administration, a pragmatic endpoint that reflects real-world management after reassessment. This distinction is important, because in routine care the decision to repeat MTX is influenced not only by biochemical behavior but also by symptoms, ultrasound findings, later β-hCG trends, follow-up reliability, and patient preference. Progesterone may therefore be too indirect a marker to function as a stand-alone determinant of second-dose treatment in this setting.

Kirk et al.[8] also highlighted a clinically relevant gray zone that parallels our own data. In their cohort, five women had successful single-dose outcomes without a documented >15% day 4-day 7 β-hCG decline; two of these women had suboptimal declines of 8% and 13%, refused a second MTX dose, and still resolved successfully.[8] Our study reproduced this clinically important gray zone in a contemporary real-world setting: five clinically stable patients with a subthreshold day 4-day 7 β-hCG decline did not accept additional MTX and nevertheless achieved biochemical cure with close follow-up. However, these patients represented a highly selected subgroup. They were clinically stable, had no evidence of rupture or hemodynamic compromise, were considered suitable for close surveillance by the treating clinicians, and were willing to comply with follow-up. Therefore, this observation should not be interpreted as evidence that a subthreshold day 4-day 7 β-hCG decline is generally benign. Rather, it supports interpreting the 15% threshold as a safety-oriented trigger for reassessment, counseling, and individualized management, not as an absolute biological boundary proving that cure without further MTX is impossible.

The broader literature also supports keeping β-hCG, rather than progesterone, at the center of post-MTX follow-up. Earlier studies suggested that pretreatment progesterone may be associated with MTX success and that progesterone may fall to a low threshold during ectopic pregnancy resolution.[12,13] However, those studies addressed a somewhat different question from the present one, being more informative about baseline prognosis or endocrine resolution than about treatment escalation during the first post-MTX week. By contrast, more recent work has focused on refining early β-hCG -based prediction. Skubisz et al.[9] and Mackenzie et al.[11] examined whether earlier post-treatment β-hCG changes could predict MTX success before day 7, whereas Davenport et al.[10] showed that time to resolution after MTX remains variable even among tubal ectopic pregnancies. Considered together with Kirk et al.,[8] these studies suggest that the field has continued to improve β-hCG -based risk stratification, but no alternative biomarker has clearly superseded the conventional day 4-day 7 framework. Our findings extend that message by showing that serial progesterone, at least when measured on days 1, 4, and 7, does not yet fill that gap.

The clinical implication of the present study is therefore not to dilute established MTX surveillance algorithms, but to refine how their results are interpreted. An insufficient day 4-day 7 β-hCG decline should still prompt reassessment and consideration of further treatment, because this remains the most widely validated early safety signal.[7,8] However, the final decision should continue to integrate the absolute β-hCG level, the direction of subsequent β-hCG change, symptoms,

ultrasonographic findings, and the patient's ability and willingness to comply with close follow-up. In this intermediate zone, progesterone remains a research marker rather than a practice-changing test. Conversely, an initially adequate day 4-day 7 β-hCG decline should not end follow-up, because some patients may later plateau or rise and still require additional MTX, as observed in our cohort.

Study Limitations

This study has several limitations. Its retrospective single-center design and modest sample size limited precision, particularly for the insufficient-response, second-dose MTX, and refusal subgroups. The low number of outcome events also limited statistical power. Therefore, the absence of significant differences in most progesterone-related parameters may reflect a type II error rather than a definite absence of association. As most patients had persistent PUL or presumed ectopic pregnancy, the findings should be interpreted within this mixed and predominantly PUL/presumed ectopic pregnancy population and may not translate directly to cohorts composed predominantly of sonographically confirmed tubal ectopic pregnancies. In addition, baseline β-hCG concentrations were relatively low, which may have contributed to the high overall treatment success rate and may limit generalizability to higher-risk patients with greater trophoblastic burden. We also could not perform robust multivariable adjustment for baseline β-hCG level, ultrasound characteristics, pre-treatment β-hCG trajectory, previous ectopic pregnancy, clinical presentation, or other potentially relevant confounders. Actual second-dose MTX administration was clinically meaningful but partly management-dependent, because it was influenced by symptoms, ultrasound findings, later β-hCG trends, follow-up reliability, clinician judgment, and patient preference. Finally, multiple unadjusted comparisons were performed without adjustment for multiple comparisons; therefore, isolated statistically significant findings from unadjusted analyses, including the day 1-day 4 progesterone change, should be considered exploratory.

Nevertheless, the study has an important strength: progesterone was measured at the same scheduled time points as routine β-hCG monitoring, allowing a direct head-to-head comparison of serial biomarker kinetics in real practice. These data support the central role of β-hCG, confirm that the <15% subgroup is clinically heterogeneous, and suggest that serial progesterone still lacks sufficient discriminatory performance to modify routine post-MTX management.

CONCLUSION

In conclusion, in this predominantly persistent PUL/presumed ectopic pregnancy cohort treated with MTX, early β-hCG kinetics appeared more informative than progesterone kinetics for identifying insufficient response and second-dose MTX administration. Serial progesterone measurement was feasible but did not show clear additional discriminatory value in this small retrospective cohort. However, because of the limited sample size, low number of outcome events, PUL-predominant population, and limited statistical power, these findings should be considered preliminary and hypothesis-generating. Larger prospective studies, particularly in cohorts with sonographically confirmed tubal ectopic pregnancy and higher baseline β-hCG levels, are needed before progesterone can be recommended as a supportive marker in post-MTX surveillance.

Ethics

Ethics Committee Approval: The study was approved by the University of Health Sciences Turkey, Kayseri City Hospital Non-Interventional Clinical Research Ethics Committee (approval no: 634, date: 04.11.2025).

Informed Consent: Due to the retrospective design and the use of anonymized clinical data, the requirement for informed consent was waived by the ethics committee. The study was conducted in accordance with the principles of the Declaration of Helsinki.

Footnotes

Authorship Contributions

Surgical and Medical Practices: C.Ü., H.A., H.G., Concept: C.Ü., H.A., Design: C.Ü., H.A., G.T.E., Data Collection or Processing: C.Ü., H.G., M.G., E.G., Analysis or Interpretation: C.Ü., H.A., G.T.E., Literature Search: C.Ü., G.T.E., Writing: C.Ü., H.A., H.G., M.G., E.G., G.T.E.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

Declaration of Generative AI and AI-assisted Technologies in the Writing Process: During the preparation of this work, the author(s) used ChatGPT to improve "readability" and "language". After using this tool/service, the author(s) reviewed and edited the content as needed and took full responsibility for the content of the publication.

REFERENCES

1. 2019 exceptional surveillance of ectopic pregnancy and miscarriage: diagnosis and initial management (NICE guideline NG126) [Internet]. London: National Institute for Health and Care Excellence (NICE); 2019.
2. Stovall TG, Ling FW, Gray LA. Single-dose methotrexate for treatment of ectopic pregnancy. Obstet Gynecol. 1991;77(5):754-757.
3. Stovall TG, Ling FW. Single-dose methotrexate: an expanded clinical trial. Am J Obstet Gynecol. 1993;168(6 Pt 1):1759-1765.
4. Wu J, Ludlow JP, De Vries B, Black K, Beale P. Single-dose methotrexate treatment for ectopic pregnancy and pregnancy of unknown location and progesterone as a predictor of success. Aust N Z J Obstet Gynaecol. 2014;54(5):469-474.
5. Barnhart KT, Gosman G, Ashby R, Sammel M. The medical management of ectopic pregnancy: a meta-analysis comparing single-dose and multidose regimens. Obstet Gynecol. 2003;101(4):778-784.
6. Lipscomb GH, Bran D, McCord ML, Portera JC, Ling FW. Analysis of three hundred fifteen ectopic pregnancies treated with single-dose methotrexate. Am J Obstet Gynecol. 1998;178(6):1354-1358.

7. Lipscomb GH, McCord ML, Stovall TG, Huff G, Portera SG, Ling FW. Predictors of success of methotrexate treatment in women with tubal ectopic pregnancies. N Engl J Med. 1999;341(26):1974-1978.
8. Kirk E, Condous G, Van Calster B, et al. A validation of the most commonly used protocol to predict the success of single-dose methotrexate in the treatment of ectopic pregnancy. Hum Reprod. 2007;22(3):858-863.
9. Skubisz MM, Dutton P, Duncan WC, Horne AW, Tong S. Using a decline in serum hCG between days 0-4 to predict ectopic pregnancy treatment success after single-dose methotrexate: a retrospective cohort study. BMC Pregnancy Childbirth. 2013;13:30.
10. Davenport MJ, Lindquist A, Brownfoot F, Pritchard N, Tong S, Hastie R. Time to resolution of tubal ectopic pregnancy following methotrexate treatment: a retrospective cohort study. PLoS One. 2022;17(5):e0268741.
11. Mackenzie SC, Moakes CA, Doust AM, et al. Early (days 1-4) post-treatment serum hCG level changes predict single-dose methotrexate treatment success in tubal ectopic pregnancy. Hum Reprod. 2023;38(7):1261-1267.
12. Ransom MX, Garcia AJ, Bohrer M, Corsan GH, Kemmann E. Serum progesterone as a predictor of methotrexate success in the treatment of ectopic pregnancy. Obstet Gynecol. 1994;83(6):1033-1037.
13. Saraj AJ, Wilcox JG, Najmabadi S, Stein SM, Johnson MB, Paulson RJ. Resolution of hormonal markers of ectopic gestation: a randomized trial comparing single-dose intramuscular methotrexate with salpingostomy. Obstet Gynecol. 1998;92(6):989-994.
14. Mosteller RD. Simplified calculation of body-surface area. N Engl J Med. 1987;317(17):1098.
15. Öztürk KN, Ekici Günay N, Koçer D. Six Sigma and analytical process in biochemistry laboratory of Kayseri City Hospital. J Turk Clin Biochem. 2024;22(2):72-82.

Original Article

DOI: 10.4274/anajog.galenos.2026.18189

Anat J Obstet Gynecol Res 2026;3(2):139-148

Early Psychological Distress after Ectopic Pregnancy Loss in Women in Their First Pregnancy: Associations with Treatment Modality, Social Support, and Childbearing Motivation

Sait Erbey[1], Mehmet Alican Sapmaz[1], Murat Polat[1], Talat Sarıkavak[2], Çağanay Soysal[1]

[1]University of Health Sciences Turkey, Ankara Etlik City Hospital, Clinic of Obstetrics and Gynecology, Ankara, Turkey
[2]Atlas University Faculty of Medicine, Department of Psychiatry, İstanbul, Turkey

ABSTRACT

Purpose: This study examined associations of treatment modality, perceived social support, and childbearing motivation with depression, anxiety, and stress in women in their first pregnancy after ectopic pregnancy loss.

Methods: Women in their first pregnancy with ectopic pregnancy were included. Psychological assessments were administered mostly on post-treatment days 2-3 (all within 1 week) using the [depression anxiety stress scale-21 (DASS-21); standard 0-42 metric], multidimensional scale of perceived social support, childbearing motivation scale, and pregnancy psychosocial health assessment scale. Group comparisons, Pearson correlations, and hierarchical multivariable regression analyses were conducted.

Results: A total of 146 women were included (methotrexate, n=83; salpingectomy, n=63). Surgical management was associated with higher depression (20.63±6.17 vs. 6.82±4.90, $p<0.001$, $d=2.52$), anxiety ($d=1.88$), and stress scores ($d=1.94$), and lower perceived social support ($d=2.14$). According to standard DASS-21 cut-offs, 54.0% of surgically managed women fell into the severe or extremely severe depression category, compared with 2.4% in the medical group. Social support was negatively correlated with distress in bivariate analyses (r between -0.49 and -0.53, all $p<0.001$). In the hierarchical model for depression, surgical treatment ($\beta=0.75$), anxiety ($\beta=0.18$), and, after adjustment, social support ($\beta=0.17$) were positively associated with depressive symptoms, whereas childbearing motivation was negatively associated ($\beta=-0.10$). The final model explained 64.5% of the variance (adjusted $R^2=0.645$).

Conclusion: Ectopic pregnancy loss was associated with substantial psychological burden, particularly after surgical management. However, because treatment was clinically indicated rather than randomized, these associations should not be interpreted as causal. Surgically managed women also had markedly greater clinical severity (higher beta human chorionic gonadotropin and frequent tubal rupture) so these associations are best regarded as exploratory and hypothesis-generating rather than as evidence of a causal effect of surgical treatment itself. Findings support integrating psychosocial assessment into routine care after ectopic pregnancy loss, especially when it is a first pregnancy.

Keywords: Ectopic pregnancy, depression, anxiety, social support, pregnancy loss

Address for Correspondence: Sait Erbey, MD, University of Health Sciences Turkey, Ankara Etlik City Hospital, Clinic of Obstetrics and Gynecology, Ankara, Turkey

E-mail: saiterbey@gmail.com **ORCID ID:** orcid.org/0009-0007-1950-9041

Received: 25.05.2026 **Accepted:** 09.07.2026 **Publication Date:** 02.09.2026

Cite this article as: Erbey S, Sapmaz MA, Polat M, Sarıkavak T, Soysal Ç. Early psychological distress after ectopic pregnancy loss in women in their first pregnancy: associations with treatment modality, social support, and childbearing motivation. Anat J Obstet Gynecol Res. 2026;3(2):139-148

INTRODUCTION

Ectopic pregnancy is a clinically important obstetric condition characterized by the implantation of a fertilized ovum outside the uterine cavity, with the potential to cause severe morbidity and mortality in women of reproductive age.[1,2] Although most ectopic pregnancies implant in the fallopian tubes, ectopic implantation may occur in other pelvic locations, such as the ovary, cervix, or interstitial portion of the tube.[3,4] In developed countries, ectopic pregnancies constitute approximately 1.5% of all pregnancies and remain a leading cause of maternal mortality in the first trimester.[5]

Treatment selection for ectopic pregnancy is guided by clinical status, beta human chorionic gonadotropin (β-hCG) levels, ultrasonographic findings, and hemodynamic stability.[1,3] The most common treatment approaches are medical therapy with methotrexate and surgical management, primarily salpingectomy.[2,4,6] Treatment decisions require consultation regarding risks, benefits, follow-up, and the patient's future fertility desires.[7] In cases of uncontrolled hemorrhage or significant tubal damage, salpingectomy becomes the preferred option.[7]

However, surgical management involving removal of the affected tube can be psychologically challenging, particularly for women who have not previously had a successful pregnancy or who have no living children.[1,5,8,9] The loss itself triggers an acute grief reaction.[9] A lack of social support may intensify feelings of loneliness and guilt,[10,11] and partners may attribute blame to each other for the loss, leading to relational strain and asymmetry.[12] When concerns about future fertility are added, a wide range of psychiatric outcomes, from post-traumatic stress disorder to anxiety disorders and depression, may emerge.[13]

Prior studies have identified two major psychological dimensions following pregnancy loss.[14] The first involves grief related to the lost pregnancy, frequently manifesting as depressive and post-traumatic stress symptoms.[1,12,13,15] The second encompasses anxiety specifically associated with future fertility concerns and uncertainty regarding the possibility of having children.[8,9] This anxiety often includes persistent worry, fear of recurrent loss, and distress about potential difficulty conceiving.[13,16] Together, these reactions illustrate the complexity and depth of the emotional impact of early pregnancy loss.

Two psychosocial factors have been highlighted as central modulators of these reactions. Social support is particularly relevant to grief-related depression and post-traumatic responses.[16,17] Adequate support, including social rituals and empathic interaction, facilitates healthy progression through grief and mitigates feelings of guilt and existential questioning, such as "Why did this happen to me?".[18,19] Anxiety regarding future fertility, in contrast, is predominantly linked to the individual's motivation and desire to have children.[20] Higher motivation to become a parent may amplify anxiety and stress about future reproductive outcomes, particularly concerns about difficulty conceiving or recurrent loss.

Despite growing evidence of the psychological distress after pregnancy loss, fewer studies have examined how treatment modality, perceived social support, and childbearing motivation are jointly associated with depression, anxiety, and stress after ectopic pregnancy, particularly among women in their first pregnancy. This population may be especially vulnerable, because the loss may be experienced not only as an acute medical event but also as a threat to future fertility and anticipated motherhood. The present study therefore examined the associations of treatment modality (medical vs. surgical), perceived social support, childbearing motivation, and retrospectively assessed psychosocial health during pregnancy with post-treatment depression, anxiety, and stress in women in their first pregnancy with ectopic pregnancy.

METHODS

Study Design and Setting

This single-center, cross-sectional study included patients diagnosed clinically with ectopic pregnancy on the basis of transvaginal ultrasonography and laboratory findings, admitted to the gynecology unit of a tertiary care hospital. The protocol was approved by the institutional University of Health Sciences Turkey, Ankara Etlik City Hospital Scientific Research Ethics Committee (approval number: AEŞH-BADEK-2025-0278, date: 26.03.2025), all participants provided written informed consent, and the study was conducted in accordance with the Declaration of Helsinki.

Eligibility Criteria

Inclusion criteria: Were women aged 18 years or older who were nulligravid prior to the index pregnancy (experiencing their first pregnancy, with no prior history of pregnancy, miscarriage, live birth, ectopic pregnancy, or abortion, and without any living children), with a gestational age of <6 weeks.

Exclusion criteria: Included patients who declined to provide informed consent; those with previous pregnancy experiences, malignancies, major psychiatric or neurological disorders, diabetes, hypertension, or other chronic systemic diseases; patients younger than 18 years; and pregnancies achieved through assisted reproductive technologies, such as in vitro fertilization or intrauterine insemination.

Medical Management

Medical treatment consisted of a single-dose methotrexate regimen administered intramuscularly at 50 mg/m^2 based on body surface area, in accordance with the institutional protocol. Body surface area was calculated using the Mosteller formula (typically ~75 mg for an average 1.5 m^2). Patients were followed with serial β-hCG measurements until complete biochemical resolution, which typically required several weeks of outpatient follow-up beyond the in-hospital observation period. In accordance with the institutional protocol, medical management with methotrexate was offered to hemodynamically stable patients without evidence of tubal rupture or significant hemoperitoneum and with a serum β-hCG level below 5,000 mIU/mL, a threshold above

which methotrexate failure rates increase substantially and surgical management is generally preferred.[7] Consistent with this protocol, no woman in the medical group had a β-hCG level ≥5,000 mIU/mL, whereas 22 women (34.9%) in the surgical group did (Table 1). Whereas surgical management was undertaken in the presence of hemodynamic instability, suspected or confirmed tubal rupture, significant hemoperitoneum, a contraindication to methotrexate, failure of medical treatment, or when medical management was otherwise deemed unsuitable. The choice between the two approaches was therefore made on clinical grounds rather than at random.

Surgical Management

Surgical management consisted of salpingectomy (laparoscopic or open) in hemodynamically unstable patients, in cases of suspected tubal rupture, or when conservative treatment was not feasible. Future fertility potential, including contralateral tubal status, was evaluated as part of routine clinical care. All surgical patients underwent salpingectomy; tube-conserving salpingostomy was not performed in this cohort. All patients who experienced failed medical treatment were subsequently operated and so in all analyses these patients were classified within the surgical group according to the definitive treatment they received. Contralateral tubal status was not recorded in a standardized format and is therefore not reported, and no assumptions regarding subsequent fertility prognosis were made on the basis of these data.

Psychological Assessment Procedure

Psychological assessments were administered after treatment initiation and clinical stabilization, after surgery or methotrexate administration, once hemodynamic stability had been achieved and acute symptoms had resolved sufficiently for the patient to complete self-report measures. Questionnaires were completed mostly on post-treatment days 2-3 and in all cases within 1 week. In the medical group this preceded complete biochemical (β-hCG) resolution, which typically takes several weeks; the window therefore captured acute psychological responses rather than reactions after definitive resolution. It also reflects routine practice at the study site, where women are monitored for several days after salpingectomy or methotrexate administration. All measures were administered in a single session during the same visit, with questionnaire order kept constant across participants. Each participant completed the questionnaires only once. In the 11 women whose medical treatment failed and who subsequently underwent salpingectomy, the assessments were administered after salpingectomy (mostly on post-operative days 2-3 and in all cases within 1 week of surgery). These women therefore contributed a single set of measurements, obtained after their definitive treatment, consistent with their classification in the surgical group.

Measures

Pregnancy psychosocial health assessment scale (PPHAS): The PPHAS is a 46-item instrument that evaluates overall psychosocial well-being during pregnancy, including emotional state, partner relationship, and social support, with items rated on a 5-point Likert scale. Higher scores indicate better psychosocial health. The original Turkish validation study reported good internal consistency (Cronbach's $\alpha>0.80$), and internal consistency was similarly high in the current sample ($\alpha=0.938$).[21] In this study, the PPHAS was administered retrospectively shortly after diagnosis of ectopic pregnancy, asking participants to evaluate their psychosocial state during the recent pregnancy that was diagnosed as ectopic. The retrospective use of the PPHAS in this acute post-loss context may be susceptible to recall bias and mood-congruent recall, and its psychometric properties have not been previously established for this specific application; these considerations are addressed in the limitations section.

Depression anxiety stress scale-21 (DASS-21): The DASS-21 short form consists of 21 items, yielding three 7-item subscales assessing symptoms of depression, anxiety, and stress over the past week. Items are rated on a 4-point scale from 0 ("did not apply to me at all") to 3 ("applied to me very much, or most of the time"), and subscale sums are multiplied by 2 to obtain final scores on the standard 0-42 metric, which is comparable to the DASS-42 full version and is required for interpreting the published severity cut-offs (depression: normal 0-9, mild 10-13, moderate 14-20, severe 21-27, extremely severe 28+; anxiety: normal 0-7, mild 8–9, moderate 10-14, severe 15-19, extremely severe 20+; stress: normal 0-14, mild 15-18, moderate 19-25, severe 26-33, extremely severe 34+). The Turkish short form has demonstrated satisfactory reliability and validity, and internal consistency coefficients in the present sample were high for depression ($\alpha=0.819$), anxiety ($\alpha=0.808$), and stress ($\alpha=0.755$).[22]

Multidimensional scale of perceived social support (MSPSS): The MSPSS includes 12 items assessing perceived support from family, friends, and a significant other, rated on a 7-point Likert scale from 1 ("very strongly disagree") to 7 ("very strongly agree"). Higher total scores indicate greater perceived social support. Prior Turkish studies have reported robust psychometric properties, and internal consistency in the present sample was excellent ($\alpha=0.89$).[23]

Childbearing motivation scale (CBMS): The CBMS assesses the strength of motivation and desire to have children. The Turkish adaptation[24] of the original scale developed by Guedes et al.[25] consists of 35 items rated on a 5-point response scale, with higher scores indicating stronger childbearing motivation. Both the original and the Turkish version have shown acceptable internal consistency and construct validity, and internal consistency in the current sample was $\alpha=0.916$.

Statistical Analysis

Statistical analyses were performed using jamovi, version 2.6 (The Jamovi project, Sydney, Australia). Normality was assessed using skewness and kurtosis; values within ±1.5 were considered approximately normal. Between-group differences (medical vs. surgical treatment) were examined using Student's t-test for normally distributed variables and the Mann-Whitney U test for non-normally distributed variables, with effect sizes reported as Cohen's d. Categorical variables

(including the categorical distribution of DASS-21 severity classifications using the standard ×2 cut-offs described above) were compared using the chi-square test where appropriate.

Associations between age, perceived social support, childbearing motivation, psychosocial health, and psychological outcomes (depression, anxiety, stress) were examined using Pearson's correlation coefficients. To explore the extent to which these variables were statistically associated with psychological outcomes, hierarchical multiple linear regression analyses were conducted for depression, with treatment modality (coded 0= medical, 1= surgical), childbearing motivation, retrospectively assessed psychosocial health during pregnancy, and anxiety entered in Model 1, and perceived social support added in Model 2 to evaluate its incremental contribution. For anxiety and stress, multiple linear regression analyses including treatment modality and the same psychosocial covariates were also conducted as confirmatory analyses. Multicollinearity was assessed using the variance inflation factor (VIF) and tolerance statistics (acceptance criteria: VIF <5, tolerance >0.2); all VIF values in the regression models were below 4.2. Of note, because treatment allocation was non-randomized, we additionally examined whether the association between surgical management and depression persisted after adjustment for baseline clinical severity. Among the clinical-severity variables, serum β-hCG was the only one available for both treatment groups; tubal rupture, hemoperitoneum, and surgical route were, by definition, specific to the surgical group and could not be entered as covariates in a model spanning both groups. As a sensitivity analysis, the multivariable depression model was therefore repeated with baseline β-hCG (per 1000 mIU/mL) added as a covariate. A two-tailed *p* value of <0.05 was considered statistically significant. Given the cross-sectional design and the non-randomized allocation of treatment, all analyses were interpreted as indicating associations rather than causal effects.

The sample size was determined by feasibility, given the low incidence of ectopic pregnancy and the restrictive eligibility criteria (women in their first pregnancy with gestational age <6 weeks and no previous pregnancies). Over the recruitment period, 146 women met all inclusion criteria and agreed to participate. This exceeds the commonly recommended minimum of 10-15 participants per predictor for regression models with up to five predictors and is consistent with previous psychological studies in this population.[13,16] The study was therefore considered to have adequate power to detect medium-to-large associations between treatment modality, psychosocial factors, and psychological outcomes.

RESULTS

During the study period, 245 women with a clinical diagnosis of ectopic pregnancy were evaluated, of whom 172 (70.2%) met the eligibility criteria. Eligible patients were approached in person by a clinician who explained the study and invited participation. Of these, 146 (84.9%) provided written informed consent and completed the questionnaires. The most common reasons for non-participation were distress or fatigue and a preference not to discuss psychological reactions during hospitalization. Treatment distribution in the study cohort was: 83 women (56.8%) received medical management with methotrexate; and 63 (43.2%) underwent salpingectomy. Eleven patients (17.5% of the surgical group) had initially

Table 1. Participant characteristics and psychological measures by treatment group

Variable	Medical (n=83)	Surgical (n=63)	Total (n=146)	*p* value	Cohen's d
Age (years)	25.27±4.11	25.38±3.71	25.3±3.9	0.86	-0.03
Perceived social support	56.76±13.70	28.79±12.16	44.7±19.0	<0.001	2.14
Childbearing motivation	119.61±26.45	123.29±30.18	121.2±28.1	0.44	-0.13
Psychosocial health	159.28±30.67	98.63±30.80	133.1±43.0	<0.001	1.97
Depression (DASS-21)	6.82±4.90	20.63±6.17	12.78±8.77	<0.001	-2.52
Anxiety (DASS-21)	6.99±4.37	15.27±4.45	10.56±6.01	<0.001	-1.88
Stress (DASS-21)	12.75±6.48	25.11±6.24	18.08±8.84	<0.001	-1.94
β-hCG (mIU/mL)	2170.9±1032.4	4500.4±2068.3	3176.1±1942.2	<0.001	-1.49
β-hCG category, n (%)					
<2000	43 (51.8)	6 (9.5)	49 (33.6)		
2000-4999	40 (48.2)	35 (55.6)	75 (51.4)		
≥5000	0 (0.0)	22 (34.9)	22 (15.1)		
Hemoglobin (g/dL)	11.9±1.2	11.8±1.2	11.9±1.2	0.72	0.06
Length of hospital stay (days)	7.3±1.6	3.3±1.3	5.6±2.5	<0.001	2.61

Values are mean ± standard deviation. DASS-21 subscale scores are reported on the standard 0-42 metric (raw subscale sum ×2). Group comparisons performed using Independent-samples t-test. Negative Cohen's d values indicate higher scores in the surgical group. β-hCG and length of hospital stay were compared using the Mann-Whitney U test owing to skewed distributions, and the corresponding medians [IQR] are reported in the text; the positive Cohen's d for length of hospital stay reflects a longer stay in the medical group, related to the inpatient methotrexate monitoring protocol rather than to greater clinical severity. β-hCG category values are within-group column percentages and are reported descriptively
DASS-21: Depression anxiety stress scale-21, IQR: Interquartile range, β-hCG: Beta human chorionic gonadotropin

received methotrexate, did not respond, and subsequently underwent salpingectomy; these patients are included in the surgical group only for analysis.

The mean age of the total sample was 25.3±3.9 years, and gestational age was not systematically recorded beyond the <6-week inclusion criterion. Total-sample means for all psychosocial and DASS-21 measures are reported in Table 1.

As shown in Table 1, there was no significant difference in age between the medical and surgical treatment groups (25.27±4.11 vs. 25.38±3.71 years, $p=0.86$), or in childbearing motivation (119.61±26.45 vs. 123.29±30.18, $p=0.44$). In contrast, women who underwent surgical treatment reported markedly lower perceived social support (28.79±12.16 vs. 56.76±13.70, $p<0.001$) and lower retrospectively assessed psychosocial health scores (98.63±30.80 vs. 159.28±30.67, $p<0.001$). Depression, anxiety, and stress scores were all substantially higher in the surgical group than in the medical group, with very large effect sizes (Cohen's d ranging from 1.88 to 2.52; all $p<0.001$). The magnitude of these between-group differences is illustrated in Figure 1, Panel A.

Baseline clinical characteristics also differed between the two groups in a manner consistent with confounding by indication (Table 1). Serum β-hCG was substantially higher in the surgical group than in the medical group [median 4078 (interquartile range 3134-5694) vs. 1935 (1449-2886) mIU/mL; mean 4500±2068 vs. 2171±1032; Mann-Whitney $p<0.001$, Cohen's d=-1.49]. Within the surgical group, tubal rupture was present in 41 patients (65.1%) and hemoperitoneum in 53 (84.1%), confirming that surgically managed women represented a clinically more severe and more acute subgroup rather than simply a different treatment modality; salpingectomy was performed laparoscopically in 60 patients (95.2%) and by laparotomy in 3 (4.8%), and 11 patients (17.5%) underwent surgery after failed methotrexate. Hemoglobin did not differ between groups (11.9±1.2 vs. 11.8±1.2 g/dL, $p=0.72$). Length of hospital stay was longer in the medical group [median 7 (6-8) vs. 3 (2-4) days, $p<0.001$], reflecting the institutional practice of inpatient serial β-hCG monitoring after methotrexate administration rather than greater clinical severity in that group. When initial serum β-hCG was grouped by clinically relevant categories (<2000, 2000-4999, and ≥5000 mIU/mL), the same severity gradient was evident: in the medical group, 43 women (51.8%) had β-hCG <2000 mIU/mL and none (0.0%) reached ≥5000 mIU/mL, whereas in the surgical group 22 women (34.9%) had β-hCG ≥5000 mIU/mL (Table 1). These categories were not prespecified and are reported descriptively. As initial β-hCG is itself one of the clinical criteria guiding treatment selection, this categorization is not independent of treatment allocation.

When DASS-21 scores were categorized according to the standard severity cut-offs (Figure 1, Panel B), the distribution of severity differed markedly between the two groups (χ^2 tests: depression $\chi^2(4)=94.34$, anxiety $\chi^2(4)=66.82$, stress $\chi^2(4)=66.93$; all $p<0.001$). Most medically managed women were classified as normal, whereas severe-to-extremely-severe categories were far more common in the surgical group: 54.0% for depression, 57.1% for anxiety, and 54.0% for stress, compared with 2.4%, 8.4%, and 3.6% in the medical group, respectively.

Pearson correlation analyses (Table 2) indicated several significant associations among the study variables. Perceived social support was negatively correlated with depression ($r=-0.49$), anxiety ($r=-0.50$), and stress ($r=-0.53$), and positively correlated with psychosocial health during pregnancy ($r=0.50$;

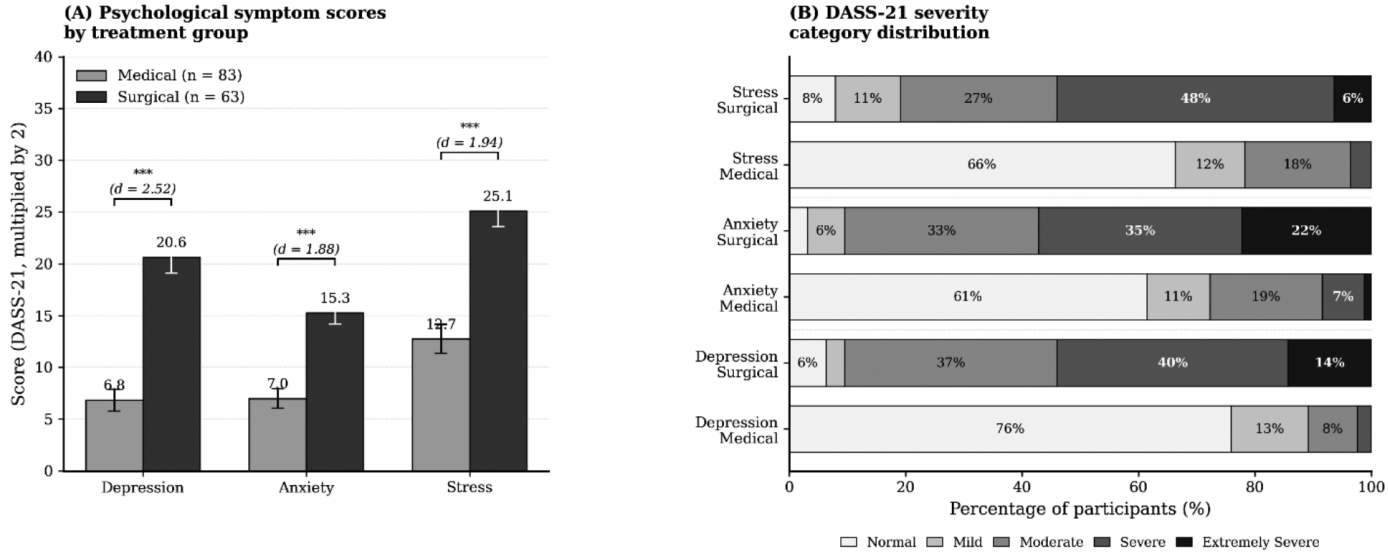


Figure 1. Psychological outcomes by treatment group. **Panel A:** Mean depression, anxiety, and stress scores (DASS-21, on the standard 0-42 metric after multiplication by 2) for medical (n=83) and surgical (n=63) groups. Error bars represent 95% confidence intervals; effect sizes (Cohen's d) and significance levels (***indicates p<0.001) are shown above each comparison. **Panel B:** Distribution of DASS-21 severity categories by group, based on the standard published cut-offs. Percentages are shown for each segment greater than 5%. Note the marked shift toward severe and extremely severe categories in the surgical group across all three outcomes

DASS-21: Depression anxiety stress scale-21

Table 2. Pearson correlations among age, psychosocial measures, and psychological outcomes (n=146)

Variable	1	2	3	4	5	6	7
Age	-						
Childbearing motivation	0.02	-					
Social support	-0.10	-0.06	-				
Psychosocial health	0.07	-0.09	0.50***	-			
Depression	-0.01	-0.07	-0.49***	-0.57***	-		
Anxiety	0.08	-0.04	-0.50***	-0.45***	0.64***	-	
Stress	0.07	-0.05	-0.53***	-0.52***	0.57***	0.52***	-

***$p<0.001$. Values are Pearson's correlation coefficients (r). DASS-21 subscale scores are on the standard 0-42 metric
DASS-21: Depression anxiety stress scale-21

all $p<0.001$), indicating moderate-to-strong relationships. Retrospectively assessed psychosocial health during pregnancy was also negatively correlated with depression ($r=-0.57$), anxiety ($r=-0.45$), and stress ($r=-0.52$; all $p<0.001$). Childbearing motivation did not show significant bivariate correlations with the psychological outcomes, whereas the three DASS-21 subscales were moderately-to-strongly intercorrelated (r between 0.52 and 0.64, all $p<0.001$). To further characterize the interrelationships among predictors, point-biserial correlations between treatment modality (coded 0= medical, 1= surgical) and the main study variables were also examined. Treatment modality was strongly positively correlated with depression (0.78), anxiety (0.68), and stress (0.70); strongly negatively correlated with perceived social support (-0.73) and retrospectively assessed psychosocial health (-0.70); and essentially uncorrelated with childbearing motivation (0.07) and age (0.02). Treatment modality was therefore strongly associated with both depression and perceived social support, whereas perceived social support itself showed only a moderate negative bivariate correlation with depression (-0.49).

Two hierarchical models were estimated for depressive symptoms (Table 3). Model 1 entered treatment modality, childbearing motivation, retrospectively assessed psychosocial health during pregnancy, and anxiety as predictors, without perceived social support, and explained 63.4% of the variance (adjusted $R^2=0.634$, $F(4,141)=63.91$, $p<0.001$). Surgical treatment ($\beta=0.63$, $p<0.001$) and higher anxiety ($\beta=0.18$, $p=0.009$) were positively associated with depressive symptoms, whereas higher childbearing motivation was associated with lower depressive symptoms ($\beta=-0.10$, $p=0.04$). The contribution of retrospectively assessed psychosocial health during pregnancy was not statistically significant in this model.

In Model 2 (Table 3), perceived social support was added to the predictors, and the explained variance increased modestly to 64.5% (adjusted $R^2=0.645$; $\Delta R^2=0.013$, F change $(1,140)=5.24$, $p=0.024$). In this full model, surgical treatment remained a strong positive correlate of depression ($\beta=0.75$, $p<0.001$), and both higher anxiety ($\beta=0.18$, $p=0.008$) and higher perceived social support ($\beta=0.17$, $p=0.024$) were positively associated with depressive symptoms after adjustment, while childbearing motivation showed a small negative association ($\beta=-0.10$, $p=0.043$). VIF values for all predictors were below 4.2, indicating acceptable multicollinearity; however, multicollinearity statistics alone do not fully explain the change in the sign of the social-support coefficient between bivariate and multivariable analyses, which is interpreted further in the discussion in the context of statistical suppression.

In a sensitivity analysis that added baseline β-hCG (per 1000 mIU/mL) to the multivariable depression model, the association of surgical treatment with depression was essentially unchanged ($\beta=0.79$, $p<0.001$), and β-hCG was not independently associated with depression after accounting for treatment and the other covariates ($\beta=-0.06$, $p=0.33$); all VIFs remained below 4.8. The associations of surgical treatment with anxiety and stress were likewise unaffected by adjustment for β-hCG. However, because β-hCG represents only one dimension of clinical severity, and the other severity features (tubal rupture, hemoperitoneum, surgical route) were specific to the surgical group, this analysis reduces but does not eliminate the possibility of residual confounding by clinical severity and the surrounding emergency context.

In confirmatory multivariable regression analyses (Table 3) including treatment modality, perceived social support, childbearing motivation, and psychosocial health during pregnancy as predictors, surgical treatment was the dominant correlate of both anxiety ($\beta=0.72$, $p<0.001$; adjusted $R^2=0.462$) and stress ($\beta=0.61$, $p<0.001$; adjusted $R^2=0.481$). The remaining psychosocial covariates did not retain independent significance in these models, suggesting that their bivariate associations with anxiety and stress were largely accounted for by the treatment modality.

Taken together, these models indicate that treatment modality, perceived social support, childbearing motivation, and psychosocial health show substantial statistical associations with depression, anxiety, and stress in women following ectopic pregnancy in their first pregnancy, without implying causal effects.

DISCUSSION

In this study of women in their first pregnancy with ectopic pregnancy, surgical management was associated with substantially higher depression, anxiety, and stress than

Table 3. Hierarchical and multivariable linear regression analyses for depression, anxiety, and stress (n=146)						
Predictor	**β (std)**	**B**	**SE**	**t**	**p**	**95% CI**
Outcome: Depression-Model 1 (adjusted R^2=0.634, F(4,141)=63.91, p<0.001)						
Surgical treatment	0.63	11.07	1.53	7.23	<0.001	8.04 to 14.10
Anxiety	0.18	0.267	0.101	2.64	0.009	0.07 to 0.47
Childbearing motivation	-0.10	-0.033	0.016	-2.05	0.042	-0.064 to -0.001
Psychosocial health	-0.05	-0.011	0.014	-0.74	0.458	-0.039 to 0.018
Outcome: Depression-Model 2 (adjusted R^2=0.645, F(5,140)=53.71, p<0.001; ΔR^2=0.013, p=0.024)						
Surgical treatment	0.75	13.23	1.78	7.43	<0.001	9.71 to 16.76
Anxiety	0.18	0.269	0.100	2.69	0.008	0.07 to 0.47
Social support	0.17	0.076	0.033	2.29	0.024	0.010 to 0.142
Childbearing motivation	-0.10	-0.032	0.016	-2.05	0.043	-0.063 to -0.001
Psychosocial health	-0.05	-0.010	0.014	-0.70	0.482	-0.038 to 0.018
Outcome: Anxiety-multivariable model (adjusted R^2=0.462, F(4,141)=32.12, p<0.001)						
Surgical treatment	0.72	8.77	1.31	6.70	<0.001	6.18 to 11.36
Childbearing motivation	-0.08	-0.017	0.013	-1.32	0.190	-0.043 to 0.009
Psychosocial health	0.06	0.008	0.012	0.67	0.504	-0.016 to 0.032
Social support	-0.01	-0.002	0.028	-0.08	0.940	-0.058 to 0.054
Outcome: Stress-multivariable model (adjusted R^2=0.481, F(4,141)=34.59, p<0.001)						
Surgical treatment	0.61	10.94	1.89	5.79	<0.001	7.20 to 14.68
Childbearing motivation	-0.09	-0.030	0.019	-1.58	0.117	-0.067 to 0.008
Psychosocial health	-0.08	-0.016	0.017	-0.95	0.344	-0.051 to 0.018
Social support	-0.04	-0.019	0.041	-0.47	0.636	-0.100 to 0.061

Treatment modality was coded 0= medical, 1= surgical. DASS-21 outcomes are on the standard 0-42 metric, so unstandardized coefficients are scaled accordingly
β: Standardized regression coefficient, B: Unstandardized coefficient, SE: Standard error, CI: Confidence interval, DASS-21: Depression anxiety stress scale-21

medical treatment, while perceived social support and childbearing motivation showed independent, and partly counterintuitive, associations with depressive symptoms in multivariable analysis. The between-group differences (Cohen's d between 1.88 and 2.52, Figure 1A) and the explained variance in depression (64.5%) indicate that treatment modality and accompanying psychosocial factors are strongly intertwined with acute post-treatment outcomes. All psychological assessments were obtained within the first post-treatment week (mostly on days 2-3) so the symptom levels reported here are best understood as acute, peri-treatment psychological responses rather than as established or longer-term psychiatric outcomes, and it cannot be determined from the present cross-sectional design whether these early reactions persist, attenuate, or progress to clinically significant disorders. These findings are consistent with prior work indicating substantial psychological burden after early pregnancy loss[13,14,26] and extend it by jointly characterizing treatment modality, perceived social support, and childbearing motivation in a homogeneous cohort of women in their first pregnancy. It must be stressed that treatment was not randomized and therefore these associations should not be read as causal effects of surgery; they reflect treatment modality together with the clinical circumstances that led to it. The unusually large between-group effect sizes observed here (most notably for depression) should likewise be interpreted in this light: a meaningful part of their magnitude is likely attributable to residual confounding and to the substantial differences in acute clinical severity between the surgical and medical groups, rather than to the modality of treatment itself.

Socio-demographic characteristics were largely comparable across treatment groups: participants were predominantly young women in their first pregnancy, and the medical and surgical groups did not differ significantly in age or childbearing motivation. This homogeneity reduces the likelihood that group differences in psychological outcomes are driven by baseline demographic disparities, although unmeasured confounders cannot be excluded. The higher distress in the surgical group aligns with previous studies reporting increased psychological distress after early pregnancy loss,[5,9,13] particularly in more invasive treatment settings. Surgical removal of the affected tube, especially in women without a prior live birth, may be experienced as both a medical emergency and a symbolic loss, intensifying grief and fears about future fertility.[6] Unfortunately, fertility-related concerns were not directly assessed in the present study and thus any interpretation invoking such concerns remains speculative and would require confirmation using dedicated, validated measures of fertility-related distress. It should be emphasized, however, that women managed surgically typically presented with greater clinical severity, such as hemodynamic instability or

suspected tubal rupture requiring emergency intervention; the elevated distress in this group therefore most plausibly reflects the combined impact of clinical severity, the acute emergency treatment setting, and the surgical procedure itself, rather than an effect of surgery alone. The categorical analysis of DASS-21 severity classes (Figure 1B) underscores the clinical relevance of these findings: in the surgical group, more than half of women fell into severe or extremely severe categories for depression, anxiety, and stress, indicating that many may meet thresholds at which clinical screening or referral would be appropriate.

Perceived social support was significantly lower among women who underwent surgical treatment. The present data cannot establish whether these lower scores represent a pre-existing difference between the groups, a psychological reaction to a more severe and acute clinical event, or a mismatch between the support needed and the support received during an emergency because perceived support was assessed only in the early post-treatment period. They indicate only that surgically managed women reported lower perceived social support during the early post-treatment period, and should not be interpreted as evidence that these women actually received less support.[1,14] Women requiring surgery may have experienced greater emotional or physical distress and thus needed more support than was met, increasing their sense of isolation.[3,6,7,9] In addition, the construct measured was perceived rather than objective support so these findings highlight the subjective nature of support and its relevance to psychological outcomes.[8,26]

Perceived social support was negatively correlated with depression, anxiety, and stress and positively correlated with psychosocial health during pregnancy, with coefficients generally around 0.5. These findings are consistent with prior studies showing that greater social support reduces maternal distress, including depression and anxiety,[12,17,18] and underscore the role of supportive social environments in fostering emotional resilience following pregnancy loss.[10,11] The potential benefits of structured peer groups and facilitated shared experiences further highlight the therapeutic value of intentionally designed support environments.[8,10,18]

A noteworthy and counterintuitive finding was that, although higher perceived social support was associated with lower depression, anxiety, and stress in bivariate analyses, it became positively associated with depressive symptoms in the multivariable model for depression after accounting for anxiety, childbearing motivation, and treatment modality. Statistically, this pattern is most consistent with a classical suppression effect: when multiple predictors that share variance with the outcome are entered together, the partial association of one predictor may reverse direction relative to its zero-order correlation. This reversal should therefore be interpreted with considerable caution. It most plausibly reflects the statistical interrelationships among the correlated predictors entered into the model rather than a genuine or harmful effect of social support, and it should not be read as implying that greater support worsens depression after ectopic pregnancy loss. In these data, perceived social support and treatment modality are strongly negatively correlated (the surgical group reported much lower support), and treatment modality carries most of the predictive variance for depression. Once treatment is in the model, the residual variance in social support may capture qualitative aspects of support rather than its overall level, and VIF values below 4.2 indicate this was not a problem of unstable estimation. Consistent with this perspective, the large variance shared by treatment modality and perceived social support (see the correlations reported above) is what produces the change in sign once both predictors are entered together. Several non-mutually exclusive mechanisms could contribute beyond this putative statistical mechanism. First, greater distress may mobilize more visible support from family and partners, so that higher support is reported by women already experiencing more severe depressive symptoms, a form of support reactivity.[27,28] Second, perceived support is not identical to received support; subjective appraisals may be shaped by expectations, relational tensions, or feelings of inadequacy.[28] Third, some forms of support may inadvertently communicate fragility, dependence, or reproductive pressure, particularly in sociocultural settings in which motherhood carries strong symbolic value,[29,30] potentially reinforcing helplessness or implicit expectations around fertility. The quality and meaning of support may therefore be at least as important as its mere presence, and perceived support after a reproductive loss should be characterized in both qualitative and quantitative dimensions. It should be emphasized, however, that these proposed mechanisms were not directly measured in the present study and therefore remain speculative. They are offered as hypotheses to be tested in future work rather than as established explanations, and they are conceptually distinct from the purely statistical (suppression) mechanism of the change in the direction of the social-support coefficient.

Childbearing motivation showed a small but significant negative association with depressive symptoms in both multivariable models, but did not retain independent significance for anxiety or stress. A strong desire for motherhood may preserve future-oriented meaning and hope after pregnancy loss, buffering some aspects of depressive withdrawal.[20] At the same time, this finding should be interpreted cautiously. The same motivational investment may, under different circumstances, intensify grief, perceived failure, or fertility-related fear.[15] Childbearing motivation may therefore not be uniformly protective and its role likely depends on how future fertility is understood, anticipated, and socially supported.[1,19,31]

From a clinical perspective, the strong association between surgical treatment and all three psychological outcomes underscores the need for greater attention to the mental health of women undergoing surgical management of ectopic pregnancy.[10,18] Given that many surgically managed women fell into severe or extremely severe DASS-21 categories, these findings argue for integrating brief psychosocial assessment and structured support into preoperative and postoperative care pathways.[8,32] Depression was related not only to treatment modality but also to anxiety, perceived social support, and childbearing motivation, so interventions addressing these domains in parallel may be beneficial. Brief psychoeducation

on fertility prognosis and treatment outcomes may contain catastrophic fears and reduce anxiety, while focused work on the quality and meaning of social support may minimize unintentionally invalidating or pressure-laden responses from partners and family. Supporting women's childbearing motivation in a realistic yet hopeful manner may help maintain future-oriented meaning after loss. Structured peer support or early follow-up contact could be especially relevant for women with pronounced distress after surgical management.

Study Limitations

Despite the strengths of the present study, which include a homogeneous cohort of women in their first pregnancy, the joint examination of treatment modality with psychosocial factors, and the use of validated Turkish-language instruments with high internal consistency, several limitations should be acknowledged. First, the single-center, cross-sectional design limits generalizability and precludes causal inference. Second, treatment was determined by clinical indications rather than by randomization. This is a classical confounding-by-indication situation: the surgical group differs from the medical group not only with respect to treatment but also in acute clinical severity, physical trauma (rupture, hemoperitoneum, urgency of surgery), and likely immediate fertility-related concerns following loss of the affected tube. We were able to retrieve and report several clinical-severity variables, including initial serum β-hCG, hemoglobin, and length of hospital stay for both groups, and surgical route, tubal rupture, and hemoperitoneum for the surgical group, which confirmed that the surgically managed women constituted a substantially more severe and more acute subgroup (for example, roughly twofold higher β-hCG and tubal rupture in 65% of cases). Initial β-hCG was further incorporated into a sensitivity analysis, in which the association between surgical treatment and depression remained essentially unchanged. Nevertheless, important severity dimensions, including adnexal mass size, pain severity, and contralateral tubal status, remained unavailable, and tubal rupture, hemoperitoneum, and surgical route were by definition specific to the surgical group and therefore could not be entered as covariates in models spanning both treatment groups. The strong association between surgical treatment and psychological outcomes should therefore not be interpreted as a direct effect of surgery itself, but as the combined effect of treatment modality and the clinical context that determined it. Third, assessments were obtained within a narrow window (mostly post-treatment days 2-3, all within one week); in the medical group this preceded complete β-hCG resolution. The design therefore captures acute responses but cannot inform on longer-term trajectories, nor on whether symptom courses diverge once medical management is completed. Fourth, all measures were self-reported and may be subject to recall or social-desirability bias; the retrospective PPHAS in particular is susceptible to mood-congruent recall, and its psychometric properties have not been established for this retrospective use. The DASS-21 is a screening rather than a diagnostic instrument. Moreover, because anxiety and depression were assessed concurrently using the same instrument (the DASS-21) within a cross-sectional design, their association cannot be interpreted as directional or predictive; in this study anxiety is best regarded as statistically associated with depressive symptoms rather than as a cause or temporal predictor of them. Fifth, unmeasured confounders such as prior psychiatric vulnerability, trauma history, the partner's reaction to the loss, or familial attitudes toward motherhood were not systematically controlled. Finally, cultural factors specific to the study setting may influence the meaning ascribed to pregnancy loss and to social support, limiting cross-cultural comparability.

Notwithstanding these limitations, the study offers important insights into the acute psychological response after ectopic pregnancy loss and highlights a need for individualized support, particularly in surgical contexts. Future longitudinal and multicenter studies, ideally with standardized capture of clinical-severity indicators (β-hCG, tubal rupture, hospital stay), received as well as perceived support, and longer-term trajectories, are needed to clarify temporal relationships, distinguish treatment effects from clinical context, and test whether these patterns replicate when such variables are explicitly modeled.

CONCLUSION

Ectopic pregnancy loss was associated with substantial psychological distress in this sample of women in their first pregnancy, particularly among those who underwent surgical management. However, these associations should not be interpreted as causal effects of surgery itself because treatment was clinically indicated rather than randomized. Given the substantially greater clinical severity of the surgically managed group, these findings are best understood as exploratory associations between treatment modality and acute distress that require confirmation in prospective, severity-adjusted studies. Perceived social support and childbearing motivation were also related to depressive symptoms, though these associations were complex and should be read in the context of broader psychosocial functioning and treatment circumstances. The findings support integrating brief, structured psychosocial assessment into the care of women with ectopic pregnancy, especially in women in their first pregnancy, in the early post-treatment period and in surgically managed patients, and indicate that targeted, multidomain interventions warrant evaluation in longitudinal and multicenter studies.

Ethics

Ethics Committee Approval: The protocol was approved by the institutional University of Health Sciences Turkey, Ankara Etlik City Hospital Scientific Research Ethics Committee (approval number: AEŞH-BADEK-2025-0278, date: 26.03.2025), and the study was conducted in accordance with the Declaration of Helsinki.

Informed Consent: All participants provided written informed consent.

Footnotes

Authorship Contributions

Surgical and Medical Practices: S.E., M.A.S., Concept: S.E., M.A.S., Design: S.E., T.S., Ç.S., Data Collection or Processing: M.A.S., M.P., T.S., Analysis or Interpretation: M.A.S., T.S., Ç.S., Literature Search: S.E., M.P., T.S., Writing: S.E., M.A.S., M.P., T.S., Ç.S.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

Declaration on the Use of Artificial Intelligence (AI): No artificial intelligence tools were used in the preparation of this manuscript.

REFERENCES

1. Yazawa H, Yazawa R, Matsuoka R, Ohara M. Surgical outcomes and trends in incidence of ectopic pregnancy. Gynecol Minim Invasive Ther. 2024;13(2):111-118.
2. Quenby S, Gallos ID, Dhillon-Smith RK, et al. Miscarriage matters: the epidemiological, physical, psychological, and economic costs of early pregnancy loss. Lancet. 2021;397(10285):1658-1667.
3. Al Naimi A, Moore P, Brüggmann D, Krysa L, Louwen F, Bahlmann F. Ectopic pregnancy: a single-center experience over ten years. Reprod Biol Endocrinol. 2021;19(1):79.
4. ESHRE Guideline Group on RPL; Bender Atik R, Christiansen OB, Elson J, et al. ESHRE guideline: recurrent pregnancy loss: an update in 2022. Hum Reprod Open. 2023;2023(1):hoad002.
5. Zhang S, Liu J, Yang L, Li H, Tang J, Hong L. Global burden and trends of ectopic pregnancy: an observational trend study from 1990 to 2019. PLoS One. 2023;18(10):e0291316.
6. Hao HJ, Feng L, Dong LF, Zhang W, Zhao XL. Reproductive outcomes of ectopic pregnancy with conservative and surgical treatment: a systematic review and meta-analysis. Medicine (Baltimore). 2023;102(17):e33621.
7. Xiao C, Shi Q, Cheng Q, Xu J. Non-surgical management of tubal ectopic pregnancy: a systematic review and meta-analysis. Medicine (Baltimore). 2021;100(50):e27851.
8. Iwanowicz-Palus G, Mróz M, Bień A. Quality of life, social support and self-efficacy in women after a miscarriage. Health Qual Life Outcomes. 2021;19(1):16.
9. Cuenca D. Pregnancy loss: consequences for mental health. Front Glob Womens Health. 2023;3:1032212. Erratum in: Front Glob Womens Health. 2023;4:1266931.
10. Conroy C, Jain T, Mody SK. Interest in peer support persons among patients experiencing early pregnancy loss. BMC Pregnancy Childbirth. 2023;23(1):506.
11. Lundsberg LS, Cutler AS, Stanwood NL, Yonkers KA, Gariepy AM. Association of pregnancy contexts with depression and low social support in early pregnancy. Perspect Sex Reprod Health. 2020;52(3):161-170.
12. Norman R. What happens after the loss? Review of early pregnancy loss information and support resources from the American College of Obstetricians and Gynecologists (ACOG) Women's Health. J Consum Health Internet. 2024;28(4):374-381.
13. Farren J, Mitchell-Jones N, Verbakel JY, Timmerman D, Jalmbrant M, Bourne T. The psychological impact of early pregnancy loss. Hum Reprod Update. 2018;24(6):731-749.
14. van den Akker OB. The psychological and social consequences of miscarriage. Expert Rev Obstet Gynecol. 2011;6(3):295-304.
15. Manninen BA. Grief after miscarriage and abortion: a pro-choice response. Womens Reprod Health. 2024;11(3):537-554.
16. Gudnadottir U, Du J, Hugerth LW, et al. Pre-pregnancy complications - associated factors and wellbeing in early pregnancy: a Swedish cohort study. BMC Pregnancy Childbirth. 2023;23(1):153.
17. Tu L. Lessons in grief: public health and healing in community. Am J Public Health. 2025;115(6):856-857.
18. Meyer MD. The paradox of time post-pregnancy loss: three things not to say when communicating social support. Health Commun. 2016;31(11):1426-1429.
19. Mullin A. Early pregnancy losses: multiple meanings and moral considerations. J Soc Philos. 2015;46(1):27-43.
20. Miller WB, Sable MR, Beckmeyer JJ. Preconception motivation and pregnancy wantedness: pathways to toddler attachment security. J Marriage Fam. 2009;71(5):1174-1192.
21. Yıldız H. Development of the pregnancy psychosocial health assessment scale. Maltepe Univ. 2011;4(1):63-74.
22. Yılmaz Ö, Boz H, Arslan A. Validity and reliability study of the Turkish short form of the depression anxiety stress scale (DASS-21). Finans Ekon Sos Arastir Derg. 2017;2(2):78-91.
23. Eker D, Arkar H, Yaldız H. Factor structure, validity and reliability of the revised form of the multidimensional scale of perceived social support. Turk Psikiyatri Derg. 2001;12(1):17-25.
24. Şimşek AH. Adaptation of the childbearing motivations scale into Turkish: validity and reliability study. J Soc Sci. 2017;4(13):395-408.
25. Guedes M, Pereira M, Pires R, Carvalho P, Canavarro MC. Childbearing motivations scale: construction of a new measure and its preliminary psychometric properties. J Child Fam Stud. 2015;24(1):180-194.
26. Bilardi JE, Temple-Smith M. We know all too well the significant psychological impact of miscarriage and recurrent miscarriage: so where is the support? Fertil Steril. 2023;120(5):937-939.
27. Palant A, Himmel W. Are there also negative effects of social support? A qualitative study of patients with inflammatory bowel disease. BMJ Open. 2019;9(1):e022642.
28. Williamson JA, Oliger C, Wheeler A, O'Hara MW. More social support is associated with more positive mood but excess support is associated with more negative mood. J Soc Pers Relat. 2019;36(11-12):3588-3610.
29. Gariépy G, Honkaniemi H, Quesnel-Vallée A. Social support and protection from depression: systematic review of current findings in Western countries. Br J Psychiatry. 2016;209(4):284-293.
30. Carter D, Misri S, Tomfohr L. Psychologic aspects of early pregnancy loss. Clin Obstet Gynecol. 2007;50(1):154-165.
31. Botero N, de Mezerac I, Ducard D, Caeymaex L. What motivates parents to continue a pregnancy after a life-limiting fetal diagnosis: a qualitative study of parents. Adv Pediatr Neonatol Care: APNC-114. 2020;3:114.
32. Palmer AD HBSW, Murphy-Oikonen J MSW, PhD. Social work intervention for women experiencing early pregnancy loss in the emergency department. Soc Work Health Care. 2019;58(4):392-411.

Original Article

DOI: 10.4274/anajog.galenos.2026.02411

Anat J Obstet Gynecol Res 2026;3(2):149-156

Comparison of Placental Histopathology in Early-Onset and Late-Onset Preeclampsia: A Prospective Cohort Study

Zeynep Gedik Özköse[1], Burak Özköse[1], Alev Atış Aydın[1], Nermin Gündüz[2]

[1]University of Health Sciences Turkey, Kanuni Sultan Süleyman Training and Research Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Turkey
[2]University of Health Sciences Turkey, Kanuni Sultan Süleyman Training and Research Hospital, Clinic of Pathology, İstanbul, Turkey

ABSTRACT

Purpose: To compare histopathological findings associated with reduced uteroplacental perfusion between [early-onset preeclampsia (EOPE), <34 weeks] and [late-onset preeclampsia (LOPE), ≥34 weeks] placentas, and to evaluate whether placental morphology may contribute to understanding differences between these subtypes.

Methods: This prospective observational cohort study was performed with the participation of women diagnosed with preeclampsia according to 2013 American College of Obstetricians and Gynecologists criteria who were divided into EOPE and LOPE subgroups. Placenta samples were collected after delivery, fixed in formalin, and evaluated after hematoxylin-eosin staining. Histopathological findings were assessed according to Amsterdam Placental Workshop Group Consensus criteria and reviewed by a blinded perinatal histopathologist. Comparisons between groups were performed using Student's t-test, Mann-Whitney U test, chi-square test, Fisher's exact test, and binary logistic regression where appropriate. A p value <0.05 was considered statistically significant.

Results: A total of 75 women were recruited, 33 (44%) with EOPE and 42 (56%) with LOPE. Villous infarction and cytotrophoblast cell proliferation were significantly more prevalent in the EOPE group ($p=0.026$ vs. $p=0.043$). Fibrinoid necrosis showed a non-significant tendency toward higher prevalence in EOPE. Syncytiotrophoblastic knotting was present in over 93% of cases in both groups ($p=0.852$). Perivillous fibrin deposits were present in all cases. Hemolysis, elevated liver enzymes, low platelets syndrome was observed only in the EOPE group (24.24% vs. 0%; $p=0.001$). Intrauterine growth restriction was significantly more common in EOPE (odds ratio 4.70) (48.48% vs. 16.67%; $p=0.003$). Mean birth weight, placental weight and Apgar scores at 1 minute and 5 minutes were all significantly lower in the EOPE group.

Conclusion: Placental defects revealing reduced uteroplacental perfusion included villous infarction and cytotrophoblastic proliferation which were significantly more prevalent in EOPE. Syncytiotrophoblastic knotting and perivillous fibrin deposition were highly prevalent in both groups, suggesting shared ischemic mechanisms. These findings support the hypothesis that EOPE and LOPE may differ in the degree and pattern of placental involvement while sharing several underlying pathophysiological mechanisms. Further studies are needed before routine implementation can be recommended in all preeclamptic pregnancies.

Keywords: Histopathology, preeclampsia, placenta, trophoblasts, placenta diseases

INTRODUCTION

Preeclampsia is a pregnancy-specific disorder characterized by new-onset hypertension after 20 weeks of gestation, with or without proteinuria, and it may lead to multi-organ dysfunction and adverse maternal and fetal outcomes. It affects approximately 2-8% of pregnancies worldwide and continues to be a major cause of maternal and perinatal mortality.[1,2] Despite extensive research, its underlying mechanisms are still not fully understood, and in daily practice management remains largely symptomatic. In most cases, delivery is still the only definitive treatment.[3]

Rather than a single disease entity, preeclampsia is increasingly considered a heterogeneous syndrome. It is commonly divided into [early-onset preeclampsia (EOPE), <34 weeks] and [late-onset preeclampsia (LOPE), ≥34 weeks] forms, which differ in

Address for Correspondence: Zeynep Gedik Özköse, MD, University of Health Sciences Turkey, Kanuni Sultan Süleyman Training and Research Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Turkey
E-mail: zeyneptfk@hotmail.com **ORCID ID:** orcid.org/0000-0001-6662-8042
Received: 17.04.2026 **Accepted:** 22.06.2026 **Publication Date:** 02.09.2026

Cite this article as: Gedik Özköse Z, Özköse B, Atış Aydın A, Gündüz N. Comparison of placental histopathology in early-onset and late-onset preeclampsia: a prospective cohort study. Anat J Obstet Gynecol Res. 2026;3(2):149-156

several important aspects.[4] EOPE is generally linked to impaired trophoblast invasion and defective spiral artery remodeling. This results in reduced uteroplacental perfusion, angiogenic imbalance, indicated typically by increased soluble fms-like tyrosine kinase-1 (sFlt-1) and decreased placental growth factor (PlGF) concentrations, and oxidative stress. LOPE, on the other hand, seems to be more closely associated with the maternal cardiovascular and metabolic background, although placental lesions and maternal vascular malperfusion (MVM) findings may also be observed in this phenotype.[4]

These differences are also reflected at the biomarker level. For example, increased sFlt-1/PlGF ratios are more closely related to placental insufficiency, even in cases that are clinically classified as late-onset.[5] This supports the suggestion that gestational age alone may not fully capture the biological diversity of the condition.

The placenta has a central role in the development of preeclampsia. Many of the characteristic lesions are now grouped under the concept of MVM, as defined by the Amsterdam Consensus Criteria.[6] This includes findings such as placental hypoplasia, infarction, retroplacental hemorrhage, and distal villous hypoplasia. Studies using this framework have shown that these lesions tend to be more frequent and more severe in early-onset disease.[7,8] Still, interpretation is not always straightforward. Interobserver variability remains an issue, and not all features are equally reproducible. Interestingly, infarction appears to be one of the more consistent findings across observers.[9] Even with these benefits, direct comparisons of placental histopathology between EOPE and LOPE are still relatively limited. In particular, it is not entirely clear how useful routine placental examination is in distinguishing between these subtypes in clinical practice.

In this study, we aimed to compare a broad range of histopathological features between EOPE and LOPE cases and to explore whether these differences may contribute to understanding the phenotypic variation between EOPE and LOPE. We hypothesize that EOPE is associated with a higher burden of MVM-related lesions compared with late-onset disease, reflecting more severe primary placental dysfunction.

METHODS

Study Design and Participants

We conducted a prospective observational cohort study in line with the STROBE guidelines at İstanbul Kanuni Sultan Süleyman Training and Research Hospital between April 2014 and January 2015. Ethical approval for this study was obtained from the University of Health Sciences Turkey, Kanuni Sultan Süleyman Training and Research Hospital Clinical Research Ethics Committee (approval no: 3, date: 26.09.2014). Written informed consent was obtained from all participants prior to enrollment. The study was conducted in accordance with the Declaration of Helsinki and applicable national regulations governing human research.

Women diagnosed with preeclampsia between 24 and 42 weeks of gestation were eligible for inclusion. Preeclampsia was defined according to the 2013 American College of Obstetricians and Gynecologists criteria,[10] based on new-onset hypertension (≥140/90 mmHg on two occasions at least four hours apart after 20 weeks of pregnancy) accompanied by proteinuria (≥300 mg/24 h). In the absence of proteinuria, diagnosis required evidence of end-organ involvement, including thrombocytopenia, renal insufficiency, impaired liver function, pulmonary edema, or new-onset neurological symptoms.

For the purpose of analysis, patients were grouped into early-onset (<34 weeks) or late-onset (≥34 weeks) preeclampsia. Gestational age was estimated based on the last menstrual period and confirmed by first-trimester crown-rump length measurements.

Women were excluded if they had chronic hypertension, pre-existing diabetes mellitus, autoimmune disease, chronic renal disease, were active smokers, or if fetal chromosomal or structural abnormalities, or intrauterine infection was present. Cases with insufficient placental material were also not included in the final analysis.

Clinical Data Collection

Maternal demographic and obstetric characteristics, including age, gravidity, parity, history of abortion, number of living children, gestational age at delivery, mode of delivery, birth weight, fetal sex, and Apgar scores were prospectively recorded. Laboratory findings at the time of diagnosis were also collected.

Intrauterine growth restriction (IUGR) was defined as an estimated fetal weight below the 10th percentile for gestational age on ultrasound. Hemolysis, elevated liver enzymes, low platelets (HELLP) syndrome was diagnosed using standard clinical and laboratory criteria, including hemolysis, elevated liver enzymes, and low platelet count.

Placental Histopathological Evaluation

Placental specimens were obtained immediately after delivery. After removal of membranes and umbilical cord, each placenta was weighed and fixed in 10% neutral buffered formalin for 48 hours. Tissue sampling included both central and peripheral regions, along with any visible lesions. Sections were embedded in paraffin, cut at a thiclness of 3-4 μm, and stained with hematoxylin and eosin. All slides were evaluated by an experienced histopathologist who was unaware of the clinical grouping. Histopathological evaluation was performed according to the Amsterdam Placental Workshop Group Consensus Statement criteria, with specific reference to MVM-related lesions. Villous infarction was defined as areas of ischemic villous necrosis characterized by loss of villous architecture and ghost-like villi with karyorrhectic debris. Fibrinoid necrosis was defined as deposition of eosinophilic fibrin-like material replacing villous stroma. Syncytiotrophoblastic knotting was assessed semi-quantitatively based on the proportion of terminal villi showing knot formation and categorized as absent (<10%), mild (10-30%), moderate (30-70%), or severe (>70%). Perivillous fibrin deposition was graded according to estimated percentage

of intervillous space involvement (<20%, 20-30%, >30%). Cytotrophoblastic proliferation was defined as persistence or increased prominence of cytotrophoblast cells along villous surfaces beyond expected gestational age-related regression. This was also assessed semi-quantitatively and categorized as absent, mild (focal increase in ≤10% of villi), moderate (10-30%), or severe (>30% of villi showing prominent proliferation).

Statistical Analysis

Statistical analyses were performed using NCSS 2007 software (NCSS, LLC, USA). Continuous variables are presented as mean ± standard deviation. Depending on distribution, comparisons were made using either Student's t-test or the Mann-Whitney U test. Standard binary logistic regression using maximum likelihood estimation was performed to explore factors associated with disease phenotype (EOPE versus LOPE). Disease phenotype was entered as the dependent variable, while IUGR, villous infarction, and cytotrophoblastic proliferation were included as independent variables based on their clinical relevance and univariate associations. HELLP syndrome was excluded from multivariable modeling due to zero-event distribution in the LOPE group, which precluded stable estimation of effect size. Given the limited sample size and sparse event distribution for some variables, the multivariable analysis should be considered exploratory and hypothesis-generating.[11] A p value <0.05 was considered statistically significant. No formal sample size calculation was performed. All eligible patients meeting the inclusion criteria during the study period were consecutively recruited. Therefore, the study should be considered exploratory. Given the close relationship between gestational age and preeclampsia phenotype (EOPE versus LOPE), gestational age was not included in the multivariable model. Therefore, the observed associations should be interpreted with caution, as residual confounding by gestational age cannot be excluded. The outcome coding used in the logistic regression analysis was clarified in the revised manuscript. In the logistic regression model, LOPE was coded as the outcome category (LOPE=1, EOPE=0). Therefore, odds ratios (ORs) below 1 indicate variables more strongly associated with EOPE. For ordinal histopathological variables, overall group comparisons were performed using chi-square tests for trend or Pearson's chi-square tests based on categorical distribution. Binary variables were analyzed using Pearson's chi-square test or Fisher's exact test where appropriate. All p values represent overall comparisons across categories rather than pairwise subgroup comparisons. Liver amniotransferase levels were compared using the Mann-Whitney U test because of non-normal distribution and extreme values.

RESULTS

Demographic, Clinical, and Laboratory Characteristics

Seventy-five women with preeclampsia were included, 33 (44%) with EOPE and 42 (56%) with LOPE. No significant differences were observed between groups in maternal age, gravidity, parity, number of abortions, or number of living children. Mean gestational age at delivery was markedly lower in the EOPE group (30.42±2.93 vs. 37.45±2.10 weeks; p<0.0001), as was mean birth weight (1480.61±504.63 g vs. 2760.36±685.24 g; p<0.0001). Systolic blood pressure at diagnosis was significantly higher in EOPE cases (165.00±13.91 vs. 153.81±14.18 mmHg; p=0.001). Diastolic blood pressure did not differ significantly between groups (102.94±10.79 vs. 100.19±9.91 mmHg; p=0.255). Apgar scores at 1 minute (5.93±2.36 vs. 7.51±1.86; p=0.002) and 5 minutes (8.03±1.30 vs. 9.10±1.16; p=0.001) were significantly lower in the EOPE group. Mode of delivery did not differ significantly between groups (cesarean: 78.79% vs. 59.52%; p=0.076). Female fetal sex was more prevalent in LOPE (66.67% vs. 42.42%; p=0.036).

HELLP syndrome occurred exclusively in the EOPE group (24.2% vs 0%, p=0.001). IUGR was identified in 48.48% of EOPE versus 16.67% of LOPE cases (p=0.003; OR 4.70, 95% CI 1.63-13.59). Eclampsia rates did not differ significantly (6.06% vs. 2.38%; p=0.420). Platelet counts were significantly lower in EOPE (198.28±87.51 vs. 236.28±73.43 $\times 10^3/\mu L$; p=0.045) and lactate dehydrogenase was significantly elevated (382.33±303.02 vs. 255.90±63.54 U/L; p=0.008). Placental weight was significantly lower in EOPE (319.24±120.86 vs. 393.57±129.42 g; p=0.013). Hemoglobin, hematocrit, urea, creatinine, aspartate aminotransferase, alanine aminotransferase, and spot urinary protein did not differ significantly between groups (p>0.05 for all; Table 1).

Histopathological Findings

All placental histopathological parameters are detailed in Table 2A and B. Villous infarction was significantly more prevalent in the EOPE group (48.48% vs. 23.81%; p=0.026; OR 3.01, 95% CI 1.12-8.07) (Figure 1). Infarcts were focal in 33.33% and multifocal in 15.15% of EOPE cases, versus 16.67% and 7.14% in LOPE, respectively.

Cytotrophoblastic cell proliferation was also significantly elevated in EOPE (15.15% vs. 2.38%; p=0.043). Fibrinoid necrosis was present in 69.70% of EOPE versus 47.62% of LOPE cases (p=0.065). Diffuse fibrinoid necrosis was observed in 6.06% of EOPE and 7.14% of LOPE cases.

Syncytiotrophoblastic knotting was detected in 93.94% of EOPE and 92.86% of LOPE placentas, with no significant difference between groups (p=0.852 for presence; p=0.617 for severity grade) (Figure 2). Perivillous fibrin deposits were universal (100% in both groups), with no significant difference in severity grading (p=0.374). Severe perivillous fibrin deposition (>30%, p=0.374) was noted in 15.15% of EOPE versus 7.14% of LOPE cases.

Villous hypovascularity was present in 78.79% of EOPE versus 69.05% of LOPE cases (p=0.293). Moderate-to-severe hypovascularity was more frequent in EOPE (54.54% vs. 33.33%). Villous fibrosis was present in 54.55% of EOPE and 47.62% of LOPE cases (p=0.551). Subcorionic hematoma was observed in 54.55% of EOPE and 71.43% of LOPE cases (p=0.131). No statistically significant differences were found for acute chorioamnionitis (Figure 3), fetal vessel acute inflammation, extravillous trophoblast proliferation, villous capillary congestion, thickening of villous vessel walls, fetal

Table 1. Demographic, clinical, and laboratory characteristics by group

Variable	LOPE (n=42)	EOPE (n=33)	*p* value
Maternal age (years), mean ± SD	29.40±6.78	29.12±5.99	0.851
Gravidity, mean ± SD	2.60±2.15	2.97±2.26	0.275
Parity, mean ± SD	1.10±1.50	1.39±1.44	0.173
Abortus, mean ± SD	0.43±1.25	0.45±0.83	0.301
Living children, mean ± SD	0.07±0.46	0.12±0.42	0.217
Gestational age at delivery (wk), mean ± SD	37.45±2.10	30.42±2.93	<0.0001*
Birth weight (g), mean ± SD	2760.36±685.24	1480.61±504.63	<0.0001*
Placental weight (g), mean ± SD	393.57±129.42	319.24±120.86	0.013*
Systolic BP at diagnosis (mmHg)	153.81±14.18	165.00±13.91	0.001*
Diastolic BP at diagnosis (mmHg)	100.19±9.91	102.94±10.79	0.255
Apgar score-1st minute, mean ± SD	7.51±1.86	5.93±2.36	0.002*
Apgar score-5th minute, mean ± SD	9.10±1.16	8.03±1.30	0.001*
SVD, n (%)	17 (40.48%)	7 (21.21%)	0.076
Cesarean delivery, n (%)	25 (59.52%)	26 (78.79%)	
Fetal sex-female, n (%)	28 (66.67%)	14 (42.42%)	0.036*
Fetal sex-male, n (%)	14 (33.33%)	19 (57.58%)	
IUGR, n (%)	7 (16.67%)	16 (48.48%)	0.003*
HELLP syndrome, n (%)	0 (0.00%)	8 (24.24%)	0.001*
Eclampsia, n (%)	1 (2.38%)	2 (6.06%)	0.420
Hemoglobin (g/dL), mean ± SD	11.99±1.91	11.95±1.76	0.929
Hematocrit (%), mean ± SD	36.07±4.75	34.61±4.47	0.180
Platelet Count ($\times 10^3/\mu$L), mean ± SD	236.28±73.43	198.28±87.51	0.045*
Urea (mg/dL), mean ± SD	20.81±6.89	23.85±8.92	0.101
Creatinine (mg/dL), mean ± SD	0.61±0.16	0.61±0.13	0.950
AST (U/L), mean ± SD	20.50±8.13	41.67±69.15	0.979
ALT (U/L), mean ± SD	12.55±7.58	21.88±33.64	0.607
LDH (U/L), mean ± SD	255.90±63.54	382.33±303.02	0.008*
Spot urinary protein (dipstick), mean ± SD	1.93±0.81	2.27±0.87	0.124
+1, n (%)	15 (36.59%)	7 (26.92%)	
+2, n (%)	14 (34.15%)	5 (19.23%)	
+3, n (%)	12 (29.27%)	14 (53.85%)	

*$p<0.05$. Data presented as mean ± SD or n (%). For delivery mode and fetal sex, *p* value is shown on the first category row. Spot protein subgroup denominators exclude 8 patients with 24-hour urine collection
LOPE: Late-onset preeclampsia, EOPE: Early-onset preeclampsia, SD: Standard deviation, BP: Blood pressure, wk: Weeks, SVD: Spontaneous vaginal delivery, IUGR: Intrauterine growth restriction, HELLP: Hemolysis, elevated liver enzymes, low platelets, AST: Aspartate aminotransferase, ALT: Alanine aminotransferase, LDH: Lactate dehydrogenase

vascular thrombosis, chronic villitis, intervillous hemorrhage, or placental calcification ($p>0.05$ for all).

In multivariable logistic regression analysis, IUGR ($p=0.041$) and cytotrophoblastic proliferation ($p=0.042$) remained significantly associated with disease phenotype, whereas villous infarction did not retain statistical significance ($p=0.395$) (Table 3). Since LOPE was coded as the outcome category (LOPE=1, EOPE=0), ORs below 1 in the multivariable model indicate stronger associations with EOPE.

DISCUSSION

In this study, we compared placental histopathological findings between early- and LOPE and observed both differences and substantial overlap between the two groups. Some features appeared to cluster with early-onset disease, whereas others were common across the spectrum. Overall, our findings support the concept that preeclampsia represents a heterogeneous disorder with overlapping but not identical placental phenotypes.

Table 2A. Histopathological findings in EOPE and LOPE ordinal histopathological findings

	LOPE n=42 (%)	EOPE n=33 (%)
Villous infarction		
Absent	32 (76.19%)	17 (51.52%)
Focal	7 (16.67%)	11 (33.33%)
Multifocal	3 (7.14%)	5 (15.15%)
Overall *p* value		0.026*
Cytotrophoblastic proliferation		
Absent	41 (97.62%)	28 (84.85%)
Mild	1 (2.38%)	4 (12.12%)
Severe	0 (0.00%)	1 (3.03%)
Overall *p* value		0.043*
Fibrinoid necrosis		
Absent	22 (52.38%)	10 (30.30%)
Focal	17 (40.48%)	21 (63.64%)
Diffuse	3 (7.14%)	2 (6.06%)
Overall *p* value		0.065
Syncytiotrophoblastic knotting		
Absent (<10%)	3 (7.14%)	2 (6.06%)
Mild (10-30%)	20 (47.62%)	11 (33.33%)
Moderate (30-70%)	16 (38.10%)	17 (51.52%)
Severe (>70%)	3 (7.14%)	3 (9.09%)
Overall *p* value		0.617
Perivillous fibrin deposition		
Mild (<20%)	18 (42.86%)	10 (30.30%)
Moderate (20-30%)	21 (50.00%)	18 (54.55%)
Severe (>30%)	3 (7.14%)	5 (15.15%)
Overall *p* value		0.374
Villous hypovascularity		
Absent	13 (30.95%)	7 (21.21%)
Mild	15 (35.71%)	8 (24.24%)
Moderate	12 (28.57%)	14 (42.42%)
Severe	2 (4.76%)	4 (12.12%)
Overall *p* value		0.293
Villous fibrosis		
Absent	22 (52.38%)	15 (45.45%)
Mild	16 (38.10%)	10 (30.30%)
Moderate	4 (9.52%)	7 (21.21%)
Severe	0 (0.00%)	1 (3.03%)
Overall *p* value		0.321
Intervillous hemorrhage		
Absent	11 (26.19%)	9 (27.27%)
Mild	20 (47.62%)	12 (36.36%)
Moderate	9 (21.43%)	11 (33.33%)
Severe	2 (4.76%)	1 (3.03%)
Overall *p* value		0.642

*$p<0.05$. Overall *p* values for ordinal variables were calculated using chi-square test for trend or Pearson's chi-square test based on distribution across all categories of each lesion, rather than pairwise comparisons
LOPE: Late-onset preeclampsia; EOPE: Early-onset preeclampsia

Table 2B. Binary histopathological findings

Feature	LOPE n=42	EOPE n=33	*p* value
Subcorionic hematoma, n (%)	30 (71.43%)	18 (54.55%)	0.131
Acute chorioamnionitis, n (%)	11 (26.19%)	8 (24.24%)	0.847
Fetal vessel acute ınflammation, n (%)	3 (7.14%)	1 (3.03%)	0.431
Extravillous trophoblast prolif., n (%)	28 (66.67%)	21 (63.64%)	0.784
Villous capillary congestion, n (%)	12 (28.57%)	9 (27.27%)	0.901
Villous vessel wall thickening, n (%)	5 (11.90%)	8 (24.24%)	0.161
Fetal vascular thrombosis, n (%)	0 (0.00%)	2 (6.06%)	0.106
Chronic villitis, n (%)	2 (4.76%)	1 (3.03%)	0.704
Placental calcification, n (%)	17 (40.48%)	12 (36.36%)	0.717

LOPE: Late-onset preeclampsia; EOPE: Early-onset preeclampsia

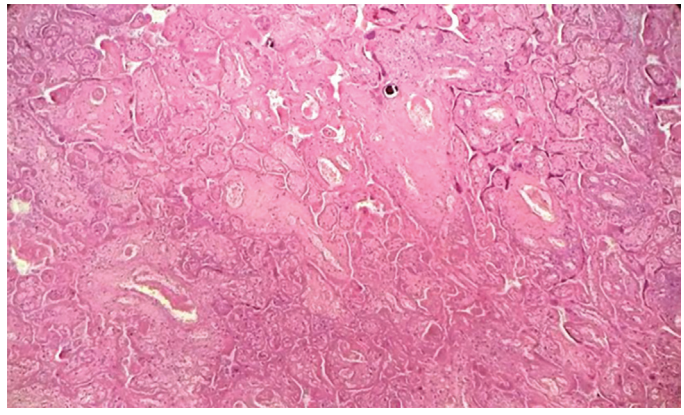


Figure 1. Villous infarction in early-onset preeclampsia placenta. Histopathological section of a placenta from a patient with early-onset preeclampsia showing villous infarction characterized by extensive areas of coagulative necrosis involving terminal villi. Loss of villous architecture, increased fibrinoid deposition, and collapse of intervillous spaces are evident. The lesion is consistent with maternal vascular malperfusion [Hematoxylin and eosin (H&E) stain, original magnification ×100]

Villous infarction was more frequent in EOPE,[12,13] which is consistent with impaired uteroplacental perfusion. This finding is biologically plausible and consistent with the current understanding of placental insufficiency in EOPE. In our cohort, this finding paralleled clinical severity, as indicated by lower birth weight, lower Apgar scores, and a higher rate of IUGR which were all seen in the EOPE group.[14] Rather than being a purely descriptive histological feature, infarction seems to reflect functionally relevant placental damage.

A similar pattern was observed for cytotrophoblastic proliferation. Normally, these cells become less prominent as gestation progresses. Their persistence, therefore, likely

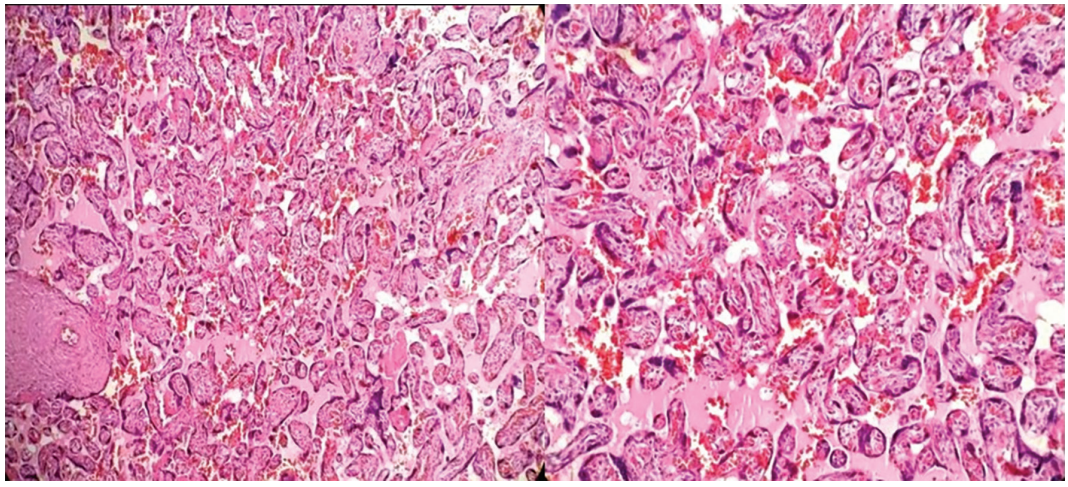


Figure 2. Syncytiotrophoblastic knotting in preeclamptic placenta. Placental tissue demonstrating increased syncytiotrophoblastic knot formation characterized by aggregation of syncytial nuclei at the villous surface. This finding is observed in both early-onset and late-onset preeclampsia cases and reflects villous maturation disturbance and placental stress (H&E stain, original magnification ×200)

H&E: Hematoxylin and eosin

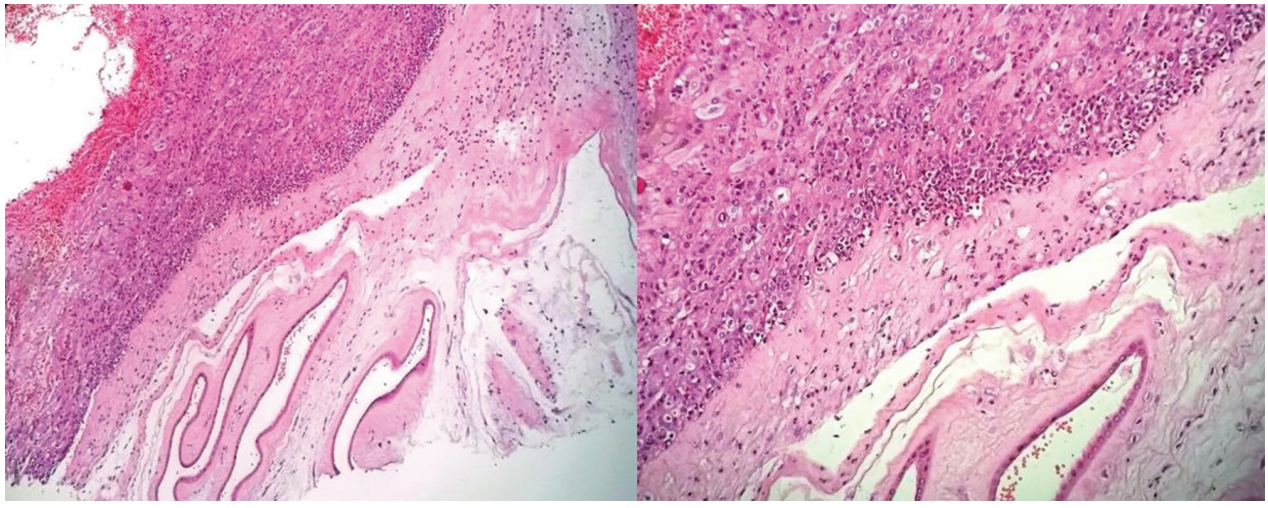


Figure 3. Acute chorioamnionitis. Histological section showing acute chorioamnionitis characterized by neutrophilic infiltration within the chorionic plate and fetal membranes. The inflammatory cells are predominantly located in the subchorionic space, consistent with an acute inflammatory response (H&E stain, original magnification ×100)

H&E: Hematoxylin and eosin

Table 3. Binary logistic regression analysis of key parameters

Parameter	B	SE	OR	95% CI	*p* value
IUGR	-1.24	0.61	0.29	0.09-0.95	0.041*
Villous infarction	-0.51	0.60	0.60	0.19-1.93	0.395
Cytotrophoblastic proliferation	-2.29	1.18	0.10	0.01-1.02	0.042*

*$p < 0.05$
B: Regression coefficient, SE: Standard error, OR: Odds ratio, CI: Confidence interval, IUGR: Intrauterine growth restriction

indicates an ongoing response to stress, possibly hypoxia-related.[15] We found that this feature remained significant even after multivariate analysis, which suggests that it may be associated with phenotypic differences between EOPE and LOPE, at least in the context of early-onset disease.

Not all findings behaved in the same way. Syncytiotrophoblastic knotting and perivillous fibrin deposition were present in almost all cases, regardless of subtype.[16] At first glance, this might suggest that they are important markers. However, their near-universal presence makes them less useful for distinguishing between EOPE and LOPE.

This point becomes clearer when our results are considered alongside previous work. A retrospective study by Atigan et al.[17] reported that syncytial knots and perivillous fibrin deposition were significantly increased in preeclamptic placentas compared to controls, but these findings were not associated with maternal or neonatal outcomes. In other words, the lesions were present, but they did not seem to carry prognostic information. Our data aligns with that observation but also extends it. While we observed high rates of these features, we found that other lesions, particularly villous infarction and cytotrophoblastic proliferation, were more closely linked to clinically relevant differences. This suggests that not all histopathological changes should be interpreted in the same way.

From a mechanistic perspective, many of the lesions that were more prevalent in EOPE, particularly villous infarction and cytotrophoblastic proliferation, are consistent with the spectrum of MVM.[18] These findings support the concept of inadequate spiral artery remodeling.[19] and reduced uteroplacental perfusion as central components of early-onset disease.[20] Such placental abnormalities may be consistent with angiogenic imbalance described in previous studies of preeclampsia; unfortunately, angiogenic biomarkers were not assessed in the present study.[21] In contrast, the relative absence of marked placental pathology in many LOPE cases supports the hypothesis that late-onset disease may be influenced to a greater extent by maternal cardiovascular, metabolic, and inflammatory factors rather than severe primary placental dysfunction. Under normal physiological conditions, cytotrophoblastic activity decreases with advancing gestation and so its persistence or reappearance in late gestation suggests ongoing placental stress. The independent association observed in adjusted models further suggests that this feature may represent a more specific marker of placental dysfunction in early-onset disease. In contrast, syncytiotrophoblastic knotting and perivillous fibrin deposition were present in nearly all cases, regardless of disease subtype. While these findings are consistent with placental stress, their lack of discriminatory value between EOPE and LOPE suggests that they reflect a common terminal pathway of placental injury rather than disease-specific mechanisms. This interpretation is supported by previous studies reporting similar lesions in preeclamptic placentas without clear correlation with clinical severity or outcomes. These observations support a model in which different histopathological lesions reflect different levels of biological specificity. Lesions such as infarction and cytotrophoblastic proliferation appear more closely linked to uteroplacental malperfusion and disease severity, whereas syncytial knots and fibrin deposition may represent nonspecific responses to chronic placental stress. Several studies have demonstrated that MVM-related lesions are not exclusive to EOPE and may also be frequently observed in late-onset disease, although often with a lower overall burden. This overlap suggests that placental malperfusion represents a shared pathological mechanism across the preeclampsia spectrum rather than a feature restricted to a single clinical phenotype. Previous reports using the Amsterdam Placental Workshop Group criteria have similarly shown that LOPE placentas may exhibit infarction, distal villous hypoplasia, and other MVM features, particularly in cases with more severe clinical presentation or fetal growth restriction.[22] These findings support a continuum model of placental involvement, in which EOPE and LOPE differ primarily in severity and extent of MVM rather than in completely distinct pathological mechanisms.

Study Limitations

There are several limitations to consider. First, the sample size was relatively modest, which may have limited statistical power to detect differences in less common placental lesions and increased the possibility of type II error. Second, the substantial difference in gestational age between EOPE and LOPE represents an important potential confounder, as some placental histopathological features vary physiologically throughout gestation. Therefore, part of the observed differences may reflect gestational age-related rather than disease-specific effects. In addition, the relatively low number of events for some variables, particularly cytotrophoblastic proliferation, may have affected the stability of the multivariable logistic regression model.

The wide confidence intervals observed for several regression estimates indicate limited statistical precision and support cautious interpretation of the multivariable findings. Therefore, the multivariable findings should be interpreted as exploratory and hypothesis-generating rather than confirmatory. Larger studies are needed to validate these associations. Third, all histopathological evaluations were performed by a single histopathologist, preventing assessment of interobserver reproducibility but strengthening observer consistency. Finally, molecular markers of placental dysfunction and angiogenic imbalance were not available for correlation with histopathological findings. Doppler velocimetry data were not available in this cohort, which represents a limitation of the study. The absence of Doppler velocimetry data limits the ability to distinguish placental insufficiency-related growth restriction from constitutionally small fetuses. In addition, multiple statistical comparisons were performed across clinical and histopathological variables. However, because no formal adjustment for multiple testing was applied, the statistically significant findings should be interpreted cautiously and considered exploratory and hypothesis-generating.

Although the study cohort was collected between 2014 and 2015, the principal outcomes relate to placental histopathology, which is unlikely to be materially affected by subsequent changes in clinical management. Moreover, the diagnostic criteria used for preeclampsia remain largely comparable to those currently applied. Nevertheless, evolving obstetric practices may influence the contemporary clinical profile of affected patients and should be considered when interpreting the results.

Despite these limitations, our findings contribute to the understanding of the heterogeneity of preeclampsia by demonstrating that not all placental lesions carry equal biological or clinical significance. While some changes reflect generalized placental stress, others appear more closely linked to disease severity and phenotype. In the present study,

villous infarction was more common in EOPE in univariate analyses, whereas cytotrophoblastic proliferation remained independently associated with disease phenotype in the exploratory multivariable model. These findings suggest that different placental lesions may vary in their ability to distinguish between preeclampsia subtypes.[23]

CONCLUSION

Our findings indicated that certain histopathological features, especially villous infarction and cytotrophoblastic proliferation, were more prominent in EOPE and may contribute to understanding phenotypic differences between EOPE and LOPE. Other features, although common, appear to reflect general placental stress rather than subtype-specific differences.

These findings suggest that EOPE and LOPE may differ in the degree and pattern of placental involvement while sharing several underlying pathophysiological mechanisms. A more systematic evaluation of placental findings may contribute to a better understanding of disease heterogeneity and placental involvement in preeclampsia. Further studies are needed to determine the potential clinical utility of histological examination of delivered placentas in preeclapmtic women.

Ethics

Ethics Committee Approval: This study was approved by the University of Health Sciences Turkey, Kanuni Sultan Süleyman Training and Research Hospital Clinical Research Ethics Committee (approval no: 3, date: 26.09.2014).

Informed Consent: All participants provided written informed consent prior to enrollment.

Acknowledgment

We are grateful to our patients for participating in this study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Z.G.Ö., B.Ö., Concept: Z.G.Ö., A.A.A., Design: Z.G.Ö., B.Ö., A.A.A., N.G., Data Collection or Processing: Z.G.Ö., B.Ö., Analysis or Interpretation: Z.G.Ö., B.Ö., Literature Search: Z.G.Ö., B.Ö., Writing: Z.G.Ö.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Duley L. The global impact of pre-eclampsia and eclampsia. Semin Perinatol. 2009;33(3):130-137.
2. Lee PL, Gudino P, Umana CD, Tufail MU, Gallippi V, Jim B. Updates on preeclampsia: pathogenesis, biomarkers, diagnosis, and management. Cardiol Rev. 2025.
3. Karrar SA, Martingano DJ, Hong PL. Preeclampsia. StatPearls [Internet]. Treasure Island, FL: StatPearls Publishing; Updated 2024.
4. Hernandez Pacheco JA, Pacheco Gutierrez IE, Iracheta Hernandez YA, et al. Late vs. early preeclampsia: mechanistic pathways and management strategies. Int J Mol Sci. 2025;26(22):11091.
5. Giardini V, Santagati AA, Marelli E, et al. Decoding placental dysfunction with a new angiogenic classification of PlGF and sFlt-1. Eur J Obstet Gynecol Reprod Biol. 2025;312:114539.
6. Khong TY, Mooney EE, Ariel I, et al. Sampling and definitions of placental lesions: Amsterdam Placental Workshop Group consensus statement. Arch Pathol Lab Med. 2016;140(7):698-713.
7. Dhruv J, Maitra N, Shindegalwekar SR, et al. Placental histopathology according to Amsterdam criteria and correlation with maternal and neonatal outcomes: a case-control study. Placenta Reprod Med. 2024;3.
8. Davis DL, Lechner AC, Chapel DB, et al. Outcome-based risk stratification model for the diagnosis of placental maternal vascular malperfusion. Mod Pathol. 2024;37(1):100370.
9. Rocha ML, Menter T, Zekic Tomas S, et al. Reproducibility of the Amsterdam consensus criteria for maternal vascular malperfusion (MVM): a multicenter evaluation of perinatal pathologists. J Perinat Med. 2026;54(4):811-815.
10. Hypertension in pregnancy. Report of the American College of Obstetricians and Gynecologists' task force on hypertension in pregnancy. Obstet Gynecol. 2013;122(5):1122-1131.
11. Altman DG, Gore SM, Gardner MJ, Pocock SJ. Statistical guidelines for contributors to medical journals. Br Med J (Clin Res Ed). 1983;286(6376):1489-1493.
12. van der Merwe JL, Hall DR, Wright C, Schubert P, Grové D. Are early and late preeclampsia distinct subclasses of the disease--what does the placenta reveal? Hypertens Pregnancy. 2010;29(4):457-467.
13. Zeek PM, Assali NS. The formation, regression, and differential diagnosis of true infarcts of the placenta. Am J Obstet Gynecol. 1952;64(6):1191-1200.
14. Battistelli M, Burattini S, Pomini F, Scavo M, Caruso A, Falcieri E. Ultrastructural study on human placenta from intrauterine growth retardation cases. Microsc Res Tech. 2004;65(3):150-158.
15. Narasimha A, Vasudeva DS. Spectrum of changes in placenta in toxemia of pregnancy. Indian J Pathol Microbiol. 2011;54(1):15-20.
16. Loukeris K, Sela R, Baergen RN. Syncytial knots as a reflection of placental maturity: reference values for 20 to 40 weeks' gestational age. Pediatr Dev Pathol. 2010;13(4):305-309.
17. Atigan A, Kilic D, Guler T, Karakaya YA. Can placental histopathological lesions be a guide to maternal and neonatal outcomes in patients with preeclampsia? J Ist Faculty Med. 2022;85(3):425-432.
18. Devisme L, Merlot B, Ego A, Houfflin-Debarge V, Deruelle P, Subtil D. A case-control study of placental lesions associated with pre-eclampsia. Int J Gynaecol Obstet. 2013;120(2):165-168.
19. Khong TY, De Wolf F, Robertson WB, Brosens I. Inadequate maternal vascular response to placentation in pregnancies complicated by pre-eclampsia and by small-for-gestational age infants. Br J Obstet Gynaecol. 1986;93(10):1049-1059.
20. Salafia CM, Pezzullo JC, Lopez-Zeno JA, Simmens S, Minior VK, Vintzileos AM. Placental pathologic features of preterm preeclampsia. Am J Obstet Gynecol. 1995;173(4):1097-1105.
21. Ezeigwe CO, Okafor CI, Eleje GU, Udigwe GO, Anyiam DC. Placental peripartum pathologies in women with preeclampsia and eclampsia. Obstet Gynecol Int. 2018 Sep 20;2018:9462938.
22. Mifsud W, Sebire NJ. Placental pathology in early-onset and late-onset fetal growth restriction. Fetal Diagn Ther. 2014;36(2):117-128.
23. Mongraw-Chaffin ML, Cirillo PM, Cohn BA. Preeclampsia and cardiovascular disease death: prospective evidence from the child health and development studies cohort. Hypertension. 2010;56(1):166-171.