

Diagnostic Accuracy of Transvaginal Ultrasonography and Serum β -hCG in the Diagnosis of Ectopic Pregnancy: A Systematic Review

I Made Wira Saputra^{1*}, Ni Putu Abdi Mahadewi²

¹⁻² General Practitioner, Wangaya General Hospital, Denpasar, Indonesia, ² Faculty Medicine

Udayana University, Denpasar, Indonesia

*Corresponding Email: wirasaputra9898@gmail.com

Abstract: Ectopic pregnancy (EP) remains an important cause of first-trimester maternal death, and accurate early diagnosis depends heavily on transvaginal ultrasonography (TVS) and serum β -human chorionic gonadotropin (β -hCG), yet recent studies report conflicting figures on how accurate these tests really are, particularly when β -hCG is low. This review aimed to synthesize recent evidence (2020–2026) on the diagnostic accuracy of TVS and serum β -hCG, alone and combined, for diagnosing EP, searching PubMed, Google Scholar, and ScienceDirect for cross-sectional, cohort, and case-control studies verified against surgical, histopathological, or clinical follow-up reference standards, with quality appraised using QUADAS-2. Nine studies published between 2020 and 2026 met the inclusion criteria. TVS alone showed highly variable sensitivity, from as low as 15.8% in women with low β -hCG and subtle tubal signs to over 90% in symptomatic women with higher β -hCG, while its specificity remained consistently high, generally above 90%. Serum β -hCG alone performed moderately, with sensitivity around 81% and specificity around 75%, and combining TVS with β -hCG, and in some studies serum progesterone, consistently raised both sensitivity and specificity above 90%. Heterogeneity in reference standards and β -hCG cutoffs precluded formal statistical pooling, so findings are presented narratively. Neither TVS nor serum β -hCG alone is sufficiently accurate at low hormone concentrations, and combining both tests, ideally with serial measurement, remains the most reliable non-invasive strategy for diagnosing EP.

Keywords: Ectopic pregnancy, Transvaginal ultrasonography, β -hCG, Diagnostic accuracy.

INTRODUCTION

Ectopic pregnancy (EP), the implantation of a fertilized ovum outside the endometrial cavity, occurs in approximately 1–2% of all pregnancies and continues to account for a substantial share of first-trimester maternal deaths worldwide, with mortality estimated at roughly 3.6 per 10,000 cases in the United Kingdom and considerably higher in low-resource settings (Wu et al., 2025). More than nine in ten ectopic pregnancies implant within the fallopian tube, while a smaller proportion present as uterine variants such as cesarean scar pregnancy, interstitial pregnancy, or cervical pregnancy, and a minority occur as heterotopic pregnancy alongside a coexisting intrauterine gestation (Zorilă et al., 2026).

For more than four decades, the combination of transvaginal ultrasonography (TVS) and quantitative serum β -human chorionic gonadotropin (β -hCG) has served as the principal non-invasive diagnostic strategy for EP, largely replacing diagnostic laparoscopy and curettage. Central to this strategy is the discriminatory concept, the serum β -hCG threshold above which an intrauterine gestational sac should reliably be visible in a normal pregnancy. Recent evidence, however, indicates that this threshold is not as fixed as once assumed: a meaningful proportion of confirmed viable intrauterine pregnancies present without a visible sac at β -hCG concentrations well above classic cutoffs (Park et al., 2023).

Diagnostic performance also appears to vary considerably depending on how and when TVS is applied. Recent studies restricted to women with low β -hCG concentrations report markedly lower sensitivity of TVS for detecting subtle ectopic signs such as tubal dilation, compared with cohorts of symptomatic women presenting with higher hormone concentrations (Wu et al., 2025); (Shafiei et al., 2024). Comparative studies evaluating serum β -hCG and TVS side by side, and studies adding serum progesterone or novel serum biomarkers as additional markers, have reported further gains in accuracy when tests are combined rather than used in isolation (Zeynizadeh Jeddi et al., 2020); (Ma et al., 2022); (Hashemi et al., 2023).

Given the continued publication of new, and at times conflicting, diagnostic accuracy data over the past several years, this review was undertaken to systematically synthesize recent evidence on the diagnostic accuracy of TVS and serum β -hCG, individually and in combination, for diagnosing EP, and to clarify how this accuracy is influenced by β -hCG concentration and diagnostic strategy.

METHODS

Study Design

This article is a systematic review guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework. The review focused on the diagnostic accuracy of TVS and serum β -hCG, alone and in combination, for diagnosing EP, restricted to studies published within the past five years (2020–2026) to reflect contemporary ultrasound technology and laboratory assay standards.

Search Strategy

Searches were conducted in PubMed, Google Scholar, and ScienceDirect using combinations of the keywords ("ectopic pregnancy" OR "tubal pregnancy") AND ("transvaginal ultrasonography" OR "transvaginal sonography") AND ("beta-hCG" OR "human chorionic gonadotropin") AND ("sensitivity" OR "specificity" OR "diagnostic accuracy"), restricted to articles published from 2020 onward. Reference lists of retrieved articles were also examined for additional eligible studies.

Inclusion and Exclusion Criteria

Studies were included if they (1) were cross-sectional, cohort, or case-control studies, or comparative diagnostic studies; (2) evaluated TVS and/or serum β -hCG for diagnosing EP; (3) used a surgical, laparoscopic, histopathological, or confirmed clinical follow-up reference standard; (4) reported extractable sensitivity, specificity, or predictive values; and (5) were published in English between 2020 and 2026. Studies were excluded if they (1) were case reports, conference abstracts, or narrative commentaries without original diagnostic accuracy data; (2) evaluated a specialized imaging technique, such as contrast-enhanced ultrasound, as the sole index test without reporting comparative TVS data; or (3) did not report data sufficient to determine diagnostic accuracy.

Article Selection Process

Records identified through the search were screened by title and abstract, and potentially eligible articles were then assessed in full text against the inclusion and exclusion criteria above. Articles that duplicated a population already captured by a larger or more complete report, or that could not be accessed in full text, were excluded. This screening process yielded nine studies that satisfied all inclusion criteria and were carried forward to data extraction and quality appraisal.

Data Extraction

For each included study, data were extracted on first author and year, country, study design, sample size, index test(s) evaluated, reference standard, applied β -hCG cutoff where relevant, and reported sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Extracted data were cross-checked against the original source before being tabulated.

Quality Assessment

The methodological quality of included studies was appraised using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool, which evaluates risk of bias and applicability across four domains: patient selection, index test, reference standard, and flow and timing. QUADAS-2 was chosen instead of cohort-based instruments such as the Newcastle-Ottawa Scale because it is purpose-built for diagnostic test accuracy studies rather than exposure-outcome association studies.

Data Synthesis

Given the marked heterogeneity across included studies in reference standards, applied β -hCG cutoffs, and study populations, ranging from unselected symptomatic women to women specifically selected for low β -hCG concentrations, this review was designed from the outset as a narrative synthesis rather than a quantitative pooling exercise. Findings are therefore presented as a structured narrative synthesis of reported sensitivity, specificity, and predictive values, supplemented by a descriptive comparison of accuracy metrics across studies.

RESULTS AND DISCUSSION

Result

Study Characteristics

Nine studies published between 2020 and 2026 met the inclusion criteria, conducted across Iran, Pakistan, China, the United States, and Romania, and comprising prospective and retrospective cohort studies, cross-sectional studies, and one comparative diagnostic-pathway study. The characteristics of the included studies are summarized in Table 1.

Table 1. Characteristics of studies included in the review

Author (Year)	Country	Design / Index test	Reference standard	Key finding
Zeynizadeh Jeddi et al. (2020)	Iran	Cross-sectional; TVS vs serum β -hCG	Surgical / clinical confirmation	Combining TVS with serum β -hCG improved diagnostic classification of EP compared with either marker used alone in symptomatic women.
Ma et al. (2022)	China	Case-control; serum biomarker panel incl. β -hCG	Clinical / surgical confirmation	A biomarker panel (GSTO1, ECM-1, β -hCG) reached an AUC of 0.93, sensitivity 88.9% and specificity 86.1%, outperforming β -hCG alone.
Hashemi et al. (2023)	Iran	Cross-sectional; serum β -hCG (n=765)	Clinical follow-up	Serum β -hCG 16 days post-embryo transfer was markedly lower in EP (369.7 IU/L) than ongoing pregnancy (1592.4 IU/L); a single cutoff could not fully discriminate outcomes.

Author (Year)	Country	Design / Index test	Reference standard	Key finding
Park et al. (2023)	United States	Retrospective cohort; TVS + serum β -hCG	Confirmed viable intrauterine pregnancy	A meaningful proportion of viable intrauterine pregnancies had β -hCG above classic discriminatory thresholds before a sac was visible on TVS.
Javed et al. (2024)	Pakistan	Cross-sectional (n=211); transabdominal ultrasound (TAS)	Surgical findings	TAS achieved sensitivity 97.4%, specificity 89%, PPV 96.3%, NPV 92.3%, accuracy 95.2%, indicating good accuracy even for abdominal imaging.
Shafiei et al. (2024)	Iran	Prospective cohort; TVS in β -hCG <1000 mIU/mL	Surgery / serial β -hCG follow-up	Sensitivity and specificity of TVS were only 50% and 66%, with PPV 78.3% and NPV 35.7%, at low β -hCG concentrations.
Malik et al. (2024)	Pakistan	Cross-sectional; Doppler-enhanced TVS	Surgical / clinical confirmation	Doppler-enhanced TVS added diagnostic value over grey-scale imaging alone, most apparent in hemodynamically stable patients.
Wu et al. (2025)	China	Retrospective cohort (n=224 TVUS; n=21 CEUS subset)	Laparoscopy / curettage	TVUS sensitivity was only 15.8% (specificity 100%) for tubal dilation at β -hCG <3,500 mIU/mL, versus 94.7% for contrast-enhanced ultrasound.
Zorilă et al. (2026)	Romania	Retrospective cohort (n=60)	Histopathology	TVS enabled accurate site-specific localization of EP (63.3% tubal, 33.3% uterine variants, 3.3% heterotopic); β -hCG decline $\geq 15\%$ by day 7 predicted successful methotrexate therapy.

TVS: transvaginal ultrasonography; TVUS: transvaginal ultrasound; TAS: transabdominal ultrasound; CEUS: contrast-enhanced ultrasound; β -hCG: beta-human chorionic gonadotropin; EP: ectopic pregnancy; PPV: positive predictive value; NPV: negative predictive value.

Quality Assessment

Quality appraisal using QUADAS-2 indicated overall acceptable methodological quality across the nine included studies, though several recurring concerns were identified. Studies that deliberately restricted enrollment to women with low β -hCG concentrations (Shafiei et al., 2024); (Wu et al., 2025) carried some concern for patient-selection bias, since their reported

accuracy reflects a more diagnostically difficult subgroup rather than the full clinical spectrum of suspected EP. Retrospective single-center designs (Wu et al., 2025); (Zorilă et al., 2026) carried some concern regarding flow and timing, given the potential for selective case retrieval. Studies with a clearly pre-specified reference standard applied consistently to all participants, such as surgical or histopathological confirmation (Javed et al., 2024); (Zorilă et al., 2026), were judged at low risk of bias for the reference-standard domain. A summary of the quality appraisal is presented in Table 2.

Table 2. Quality appraisal of included studies (QUADAS-2)

Author (Year)	Patient Selection	Index Test	Reference Standard	Flow and Timing
Zeynizadeh Jeddi et al. (2020)	Low concern	Low concern	Low concern	Some concerns
Ma et al. (2022)	Some concerns	Low concern	Low concern	Low concern
Hashemi et al. (2023)	Some concerns	Low concern	Some concerns	Low concern
Park et al. (2023)	Some concerns	Low concern	Low concern	Some concerns
Javed et al. (2024)	Low concern	Low concern	Low concern	Low concern
Shafiei et al. (2024)	Some concerns	Low concern	Low concern	Low concern
Malik et al. (2024)	Some concerns	Some concerns	Low concern	Some concerns
Wu et al. (2025)	Some concerns	Low concern	Low concern	Some concerns
Zorilă et al. (2026)	Low concern	Low concern	Low concern	Some concerns

QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies-2.

Diagnostic Accuracy of Transvaginal Ultrasonography

Three studies reported the diagnostic accuracy of TVS used alone. Its performance varied markedly according to serum β -hCG concentration and the specific sonographic sign relied upon. In women with β -hCG below 1,000 mIU/mL, TVS achieved a sensitivity of only 50% and a specificity of 66% (Shafiei et al., 2024). In women with β -hCG below 3,500 mIU/mL, the sensitivity of TVS for detecting tubal dilation fell even further, to 15.8%, although specificity remained at 100% (Wu et al., 2025). By contrast, in a broader symptomatic population evaluated by transabdominal ultrasound, a form of sonography closely related to but distinct from TVS, sensitivity reached 97.4% with a specificity of 89% (Javed et al., 2024). Doppler-enhanced TVS provided incremental diagnostic value over grey-scale TVS alone, particularly in hemodynamically stable women with an adequately visualized adnexal mass (Malik et al., 2024).

Diagnostic Accuracy of Serum β -hCG

Serum β -hCG measured in isolation demonstrated moderate diagnostic performance. In a large cross-sectional study of women following embryo transfer, β -hCG concentrations 16 days after transfer were markedly and significantly lower in women who ultimately had EP than in those with ongoing pregnancies, though a single absolute cutoff could not fully discriminate between the two groups (Hashemi et al., 2023). A separate biomarker study reported that β -hCG alone achieved lower diagnostic accuracy than a composite biomarker panel that included β -hCG alongside two novel serum proteins, with the composite panel reaching a sensitivity of 88.9% and specificity of 86.1% (Ma et al., 2022). Reinforcing the limitations of relying on a single β -hCG threshold, a retrospective cohort of women with confirmed viable intrauterine pregnancies found that a meaningful proportion had β -hCG concentrations exceeding classic discriminatory cutoffs before an intrauterine sac became sonographically visible (Park et al., 2023).

Diagnostic Accuracy of Combined TVS and Serum β -hCG

Across the included studies, combining TVS with serum β -hCG, and in some cases with additional markers such as serum progesterone or novel serum proteins, consistently improved diagnostic accuracy compared with either test used in isolation (Hognert et al., 2022). A comparative study in symptomatic Iranian women reported that classification of EP was more accurate when TVS findings were interpreted together with the serum β -hCG level than when either marker was used alone (Jeddi et al., 2020). Similarly, contrast-enhanced ultrasound, used as an adjunct to conventional TVS in women with low β -hCG and an inconclusive initial scan, raised sensitivity for detecting tubal dilation from 15.8% to 94.7% while maintaining a specificity of 100% (Wu et al., 2025). Beyond diagnosis alone, TVS combined with serial β -hCG monitoring also proved useful in guiding management, with a decline in β -hCG of at least 15% by treatment day 7 predicting successful methotrexate therapy in women with confirmed EP (Zorilă et al., 2026).

Discussion

This review of nine recent studies confirms that neither TVS nor serum β -hCG alone is a reliably accurate stand-alone test for diagnosing EP, and that the accuracy of both varies substantially depending on the clinical scenario in which they are applied. The most consistent and clinically important pattern to emerge was the marked decline in TVS sensitivity at low serum β -hCG concentrations, falling to as low as 15.8% in one recent cohort restricted to women with subtle tubal signs (Wu et al., 2025), compared with sensitivity above 90% in less hormonally restricted populations (Javed et al., 2024). This pattern is consistent with the underlying biology: at low β -hCG concentrations, any gestation, whether intrauterine or ectopic, may simply be too small to visualize reliably, regardless of operator skill or equipment quality.

A second important finding concerns the discriminatory β -hCG threshold itself. Although this concept has long guided clinical decision-making, recent evidence indicates that it is considerably less absolute than previously assumed, with a meaningful proportion of viable intrauterine pregnancies presenting without a visible gestational sac at β -hCG concentrations above classic cutoffs (Park et al., 2023). Applied rigidly, over-reliance on a single β -hCG threshold could lead a clinician to presume a diagnosis of EP, or of a failing pregnancy, in a woman who is, in fact, carrying a viable intrauterine gestation, particularly given that TVS itself performs poorly in exactly this low- β -hCG range.

These findings together support a consistent clinical message: TVS and serum β -hCG should be interpreted jointly, and ideally serially, rather than as independent, one-off tests.

Every study in this review that directly compared a combined strategy with either test alone reported a meaningful gain in accuracy from combination, whether through pairing TVS with quantitative β -hCG (Jeddi et al., 2020), adding serum progesterone or novel biomarkers to β -hCG (Ma et al., 2022), or supplementing an inconclusive TVS scan with contrast-enhanced ultrasound (Wu et al., 2025). Where TVS findings remain inconclusive and β -hCG is below the discriminatory range, serial β -hCG measurement, rather than a single reading, appears to be the more defensible next step, consistent with its demonstrated value in guiding methotrexate management once EP is confirmed (Zorilă et al., 2026).

It is also worth distinguishing between transvaginal and transabdominal approaches. While TVS is generally regarded as the first-line modality because of its superior resolution for pelvic structures, transabdominal ultrasound achieved comparably high accuracy in at least one recent cohort (Javed et al., 2024), suggesting it retains clinical value where transvaginal access is limited, declined by the patient, or unavailable, provided the sonographer is adequately experienced.

This review has several limitations. The included studies were markedly heterogeneous in their reference standards, applied β -hCG cutoffs, and study populations, which precluded formal quantitative pooling and limited the review to a narrative and descriptive synthesis; the accuracy figures reported here should therefore be interpreted as illustrative of a consistent pattern rather than as precise pooled estimates. Several included studies were also based on comparatively small or single-center samples, which may limit generalizability. Few of the identified studies were conducted in Southeast Asia, including Indonesia, where access to high-resolution transvaginal equipment and rapid serial β -hCG testing may differ from the higher-resource settings represented in most included studies; prospective multicenter studies from this region using standardized diagnostic criteria are needed.

CONCLUSION

Transvaginal ultrasonography and serum β -hCG each contribute meaningfully to the diagnosis of ectopic pregnancy, but neither is sufficiently accurate on its own, particularly when β -hCG concentrations are low. Diagnostic accuracy improves substantially when the two tests are interpreted together, and serial β -hCG measurement combined with follow-up TVS, rather than reliance on a single ultrasound scan or a single β -hCG threshold, remains the most reliable non-invasive strategy currently supported by the recent literature for diagnosing ectopic pregnancy.

REFERENCES

- Hashemi, S., Arabipoor, A., Kohestani, S., Zolfaghary, Z., Vesali, S., & Ghaffari, F. (2023). The predictive value of serum β -hCG levels in the detection of ectopic pregnancy sixteen days after embryo transfer: A cross-sectional study. *International Journal of Fertility & Sterility*, 17(3), 181–186.
- Jar-Allah, T., Hognert, H., Köcher, L., Berggren, L., Fiala, C., Milsom, I., & Gemzell-Danielsson, K. (2022). Detection of ectopic pregnancy and serum beta hCG levels in women undergoing very early medical abortion: a retrospective cohort study. *The European Journal of Contraception & Reproductive Health Care*, 27(3), 240–246.
- Javed, M. A., Niaz, A., & Asim, U. (2024). Accuracy of transabdominal ultrasonography in diagnosing ectopic pregnancy, taking surgical findings as gold standard. *Pakistan Armed Forces Medical Journal*, 74(Suppl-2), 334–338.
- Ma, D., Yang, R., Chen, Y., Huang, Z., Shen, Y., He, C., & Zhao, L. (2022). Identification of noninvasive diagnostic biomarkers for ectopic pregnancy using data-independent

- acquisition (DIA) proteomics: A pilot study. *Scientific Reports*, 12(1), 1–8.
- Malik, N., Ahmed, T. M., Shehnaz, S., Aftab, S. B., Noor, R., & Ahmad, H. (2024). Diagnostic accuracy of Doppler ultrasound in ectopic pregnancy. *Annals of Pakistan Institute of Medical Sciences*, 21(1), 115–119.
- Park, K. E., Latack, K. R., Vestal, N. L., Ingles, S. A., Paulson, R. J., & Awadalla, M. S. (2023). Association of hCG level with ultrasound visualization of the gestational sac in early viable pregnancies. *Reproductive Sciences*, 30(12), 3623–3628.
- Shafiei, P., Sohbati, S., Mansouri, G., Nikpour, N., & Karami Robati, F. (2024). Diagnostic value of transvaginal ultrasonography in the evaluation of ectopic pregnancy with β -hCG levels lower than 1000 mIU/mL. *The Iranian Journal of Obstetrics, Gynecology and Infertility*, 27(8), 1–8.
- Wu, J., Chen, X., Wu, Q., Wei, N., An, G., Yu, H., & Wang, N. (2025). Contrast enhanced ultrasound – a useful method for diagnosing tubal ectopic pregnancy with low level β -HCG. *The Ultrasound Journal*, 17, Article.
- Zeynizadeh Jeddi, S., Mobaraki, N., & Pirzadeh, A. (2020). Transvaginal ultrasound compared with serum β -hCG level for diagnosis of ectopic pregnancy in symptomatic patients. *International Journal of Reproduction, Contraception, Obstetrics and Gynecology*, 9(3), 1228–1233.
- Zorilă, M. V., Iliescu, D.-G., Zorilă, G.-L., Pirici, D., Istrate-Ofițeru, A.-M., Roșu, C.-G., Busuioc, C. J., Mogoantă, L., Nimigean, V. R., Căpitănescu, R. G., Berbecaru, E. I. A., Drăgușin, R.-C., Comănescu, M.-C., & Paitici, S. (2026). From suspicion to confirmation: An original study on a complete diagnostic pathway for ectopic pregnancy. *Journal of Clinical Medicine*, 15(2).