Received: 28 March 2022 | Revised: 6 July 2022 | Accepted: 20 July 2022

DOI: 10.1002/jgf2.572

Journal of General and Family Medicine Open Access WILEY

ORIGINAL ARTICLE

Diagnostic accuracy of pelvic examination in pelvic inflammatory disease: A meta-analysis

Hiroyoshi Iwata MD, MSc, PhD[1,2] | Yoshifumi Sugiyama MD, PhD[1,3] | Yoshinao Satoi MD[1,4] | Naoko Sasamoto MD, MPH, PhD[5] | Takuya Aoki MD, PhD, MMA[1] | Masato Matsushima MD, MPH, PhD[1]

[1]Division of Clinical Epidemiology, Research Center for Medical Sciences, The Jikei University School of Medicine, Tokyo, Japan

[2]Department of Clinical Pharmacology and Therapeutics, Graduate School of Medicine, University of the Ryukyus, Nishihara, Japan

[3]Division of Community Health and Primary Care, Center for Medical Education, The Jikei University School of Medicine, Tokyo, Japan

[4]Kuji Clinic, Japanese Health and Welfare Co-operative Federation, Kawasaki, Japan

[5]Department of Obstetrics and Gynecology, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA

Correspondence
Hiroyoshi Iwata, Division of Clinical Epidemiology, Research Center for Medical Sciences, The Jikei University School of Medicine, 3-25-8, Nishi-Shimbashi, Minato-ku, Tokyo, Japan.
Email: hiroyosii@jikei.ac.jp

Funding information
This research was supported by The Jikei University Research Fund for Graduate Students, Grant/Award Number: 2019-502DJ

Abstract

Background: Pelvic inflammatory disease (PID) is not a mere transient infection. PID can lead to chronic pain, ectopic pregnancy, and infertility. Although the Centers for Disease Control and Prevention have established minimum diagnostic criteria, including pelvic examination, the diagnostic value of pelvic tenderness has recently garnered controversy. Our meta-analysis aimed to confirm whether pelvic tenderness, cervical motion tenderness, and adnexal tenderness can help diagnose PID.

Methods: We searched for studies reporting the diagnostic test accuracy of pelvic tenderness, cervical motion tenderness, and adnexal tenderness among female patients at risk for PID, using MEDLINE, EMBASE, CENTRAL, CINAHL, Google, and Google Scholar through May 25th, 2022. After quality assessment using QUADAS-2, we performed data synthesis using a bivariate random effect model and Bayesian hierarchical summary receiver operating characteristic model. We then conducted sensitivity analysis excluding studies with non-PID cases.

Results: The literature search produced 6769 articles. After quality assessment, 14 studies and their 2808 participants were eligible for synthesis on pelvic tenderness. Laparoscopy, either alone or in combination, was the most frequent reference standard. The main results for pelvic tenderness sensitivity and specificity were 0.81, 95% confidence interval (CI) [0.67–0.90] and 0.40, 95% CI [0.25–0.57], respectively. Sensitivity and specificity were 0.72, 95% CI [0.57–0.83] and 0.50, 95% CI [0.34–0.66], for cervical motion tenderness, and 0.87 [0.64–0.96] and 0.27, 95% CI [0.12–0.52] for adnexal tenderness, respectively.

Conclusions: Our meta-analysis suggests that pelvic tenderness assessed by pelvic examination may be useful for PID examination with moderate-to-high sensitivity, whereas clinicians should be aware of the diagnostic significance of pelvic tenderness.

KEYWORDS
epidemiology, gynecology, meta-analysis, pelvic inflammatory disease, sexually transmitted disease

1 | INTRODUCTION

Pelvic inflammatory disease (PID) is an inflammatory condition of the female upper reproductive tract, often caused by sexually transmitted diseases (STDs).[1] It is a syndrome including endometritis, salpingitis, adnexitis, oophoritis, pelvic peritonitis, pelvic cellulitis, and tubo-ovarian abscess.[1,2] Not a mere transient infection, PID may cause chronic pelvic pain, infertility, and ectopic pregnancy after its acute stage.[3] The Centers for Disease Control and Prevention (CDC) reports that one in eight women with a history of PID has difficulty getting pregnant.[4] Further, many specialties handle PID, including gynecology, emergency, internal medicine, infectious disease, and family medicine. Earlier precise diagnosis allows earlier effective treatment, which can improve preventable reproductive morbidity.[5]

PID diagnosis is challenging, partially because it is a syndrome of multi-site inflammation of the endometrium, fallopian tubes, ovaries, and pelvic peritoneum.[1,6] Furthermore, symptoms, severity, and clinical course are varied, making PID diagnosis complex, and there is no perfect single standard test to definitively diagnose PID.[1,6] Pelvic tenderness has significant value in PID diagnosis. Pelvic tenderness allows clinicians to detect PID, which helps prevent future infertility. This physical examination can be conducted without difficult preparation; pelvic examinations do not need expensive devices and can be implemented in areas that otherwise lack sufficient access to gynecological care, such as developing countries or rural areas.

Under the CDC's proposed diagnostic criteria, clinicians should start presumptive antibiotic treatment if a sexually active woman at risk of STDs is suspected to have PID based on the minimum clinical criteria: cervical motion tenderness, or adnexal tenderness, or uterine tenderness.[6] These three types of tenderness are components of pelvic tenderness, which is detected through pelvic examination, also referred to as gynecological examination, or bimanual examination. The extremely high sensitivity of 95% for adnexal tenderness supports these criteria.[7] However, a recent prospective study found low sensitivity (37.9%) for pelvic tenderness, and that it did not increase the sensitivity or specificity of diagnosis of PID or cervicitis.[8] Although Farrukh's study included non-PID patients, this conclusion questioning the utility of gynecological examination has sparked discussion, with some clinicians reemphasizing the value of pelvic examinations.[9]

Therefore, our study was conducted to accurately assess the diagnostic value of pelvic examinations in women at risk for PID. There were several obstacles to conducting this meta-analysis: PID comprises multiple diseases, there is no established single perfect reference standard, and there were some inconsistencies among the original candidate studies eligible for meta-analysis. We thus implemented meta-analysis using methods intended to cope with these issues.

2 | MATERIALS AND METHODS

We registered our study protocol on PROSPERO (the International Prospective Register of Systematic Reviews) on March 29, 2019, under registration ID CRD42019122527.[10] This study followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) statement and its flow chart[11] and DTA checklist.[12] (Table S1) Because our study is a meta-analysis, the informed consent statement was not applicable.

Inclusion conditions were: (1) original studies giving sufficient data to fill two-by-two tables (true-positive, false-positive, true-negative, and false-negative) for diagnosis of acute PID and its related diseases; (2) clinical trials, cross-sectional, case–control, cohort, and diagnostic studies, which were performed to confirm PID.

Our exclusion criteria sought to minimize bias due to inappropriate study design or inferior report quality.[13] Studies with inappropriate reference standards, review studies, letters, editorials, gray literature, duplicate or series publications, and nonhuman studies were excluded. Study participants were outpatients, inpatients, and emergency patients suspected of PID. Additionally, PID can coexist with near-site infections. Cervicitis is technically not PID; however, we included studies whose participants potentially had both PID and cervicitis at the first stage of synthesis. Index tests were pelvic tenderness including cervical motion tenderness and adnexal tenderness. Considering the complexity of PID diagnosis and lack of a gold standard diagnostic test, our study included several kinds of reference standards mentioned by the CDC.[6] The reference standard tests described by the CDC include: histopathologic evidence of endometritis on biopsy, transvaginal sonography or magnetic resonance imaging showing enlarged fluid-filled tubes with or without free pelvic fluid or tubo-ovarian complex, and PID consistent inflammation of the endometrium, fallopian tubes, ovaries, and pelvic peritoneum confirmed by laparoscopy, which were mentioned as specific tests for PID.[2,6] We also included abnormal cervical or vaginal mucopurulent discharge, the presence of white blood cells on saline microscopy of vaginal secretions, laboratory documentation of a cervical infection with *Neisseria gonorrhoeae* or *Chlamydia trachomatis* in urine, which were also CDC additional criteria and were implemented as reference standards in some studies.[2,6] However, these reference standards alone use may cause bias. We also included studies using a clinical method combining the above and other additional information including fever or laboratory tests, which is described by CDC to increase diagnostic certainty, i.e., sensitivity and specificity[6] Thus, we excluded any study that used the CDC minimum criteria as its only reference standard or as its primary reference standard, given the similarities between the index test and reference standard.

We performed a literature search with MEDLINE (PubMed), EMBASE, CINAHL, and CENTRAL, followed by a manual search using Google and Google Scholar covering dates up to May 25th, 2022. (Table S2) We did not contact any authors. Research collaborators carefully examined search terms, and librarians in the academic information center at the Jikei University School of Medicine provided support. We used a combination of PID disease terms: pelvic inflammatory disease, endometritis, salpingitis, adnexitis, oophoritis, pelvic peritonitis pelvic cellulitis, tubo-ovarian abscess, and their thesaurus matches; and diagnostic terms: diagnosis, examination,

symptom, pain, tenderness, and their thesaurus matches. Two independent teams [HI, NS and YSa, YSu] performed the first screening. Considering a report that restricting search scope to English language has little effect on outcomes in systematic reviews, we excluded languages other than English in order to allow more precise quality assessment.[14] At the second stage, we did full text reviews. Original studies were selected based on the inclusion and exclusion criteria. Furthermore, NS judged eligible studies from the perspective of gynecology. Discrepancies of inclusion or exclusion were discussed and resolved by all authors.

We performed quality assessment of the eligible studies using Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2), evaluating applicability and assessing risk of bias.[15] At the data synthesis stage, if an article did not include data for pelvic tenderness including pelvic examination, bimanual examination, and bimanual tenderness, we used cervical motion tenderness as a proxy for pelvic tenderness. Similarly, we used adnexal tenderness as a proxy for pelvic tenderness if cervical motion tenderness data were unavailable. For data synthesis, while there were various reference standards for diagnosis of PID, we defined them as the same reference standard because each method was mentioned by the CDC.[6] For eligible studies that did not show true positive, false positive, true negative, and false negative rates, which are necessary for statistical synthesis, we calculated them based on related information such as sensitivity and specificity. The sensitivity and specificity values shown in Table 1 are derived from those true positive, false positive, true negative, and false negative rates.

When combining the diagnostic data among the eligible studies, we implemented the bivariate random effect model and Bayesian hierarchical summary receiver operating characteristic model (HSROC).[16,17] The bivariate random model is suitable for estimating summary diagnostic values, including sensitivity and specificity. On the contrary, the HSROC model is feasible for SROC curve estimation to evaluate the test accuracy and for grasping how the curve's position and shape vary.[18] As mentioned above, while PID can coexist with cervicitis, cervicitis is technically not PID and anatomically may not cause adnexal or uterine tenderness. Thus, on sensitivity analysis, we restricted to eligible studies whose participants had only PID and excluded studies of patients with non-PID diseases such as cervicitis.

Heterogeneity was assessed descriptively, visually and with the I^2 method.[19] We refrained from statistical evaluation of publication biases. Data synthesis was done with Revman 5.3,[20] STATA version 13.1,[21] and midas and metandi packages.[19,22]

3 | RESULTS

A flow chart showing our study selection process is presented in Figure 1. After removing obvious duplications, there were 6769 potentially eligible articles, with seven articles added through a manual Google and Google Scholar search. We then screened the 6776 articles using title and abstract. We excluded 6511 articles based on our criteria. Of the remaining 258 articles, 244 were ineligible after full text reviews for insufficient data, not using original data, primarily including cases of other diseases, or incorrect index or comparison of PID and control. We excluded a study whose participants were chronic PID including tuberculosis as inappropriate participants.[23]

Finally, 14 studies were eligible for quality assessment. We used QUADAS-2 assessment to carefully evaluate risk of bias and applicability (Figure 2). Overall, high risk proportions were less than 15% in all contents, and no study was excluded based on its quality: risk of bias and applicability. As a result, 14 studies were appropriate for our meta-analysis. (Table S3).

The characteristics and results of individual studies and the study data of pelvic tenderness synthesis are shown in Table 1. The 14 studies were published between 1985 and 2021. Study settings included the United States, Zimbabwe, Pakistan, Columbia, Sweden, and Spain. The most common disease is PID, followed by salpingitis, endometritis and cervicitis. The methodology of reference standard varies from sole use of laparoscopy, to urine STD test, to mixed method. Among these 14 studies, the ranges of mean sensitivity and mean specificity for pelvic tenderness were 0.38–1.00 and 0.13–0.74, respectively. (Figure S1) We performed synthesis using 14 studies, where the synthesis of pelvic tenderness was from all 14 studies; cervical motion tenderness was from 10 studies, ID numbers 1, 4–6, 8–10, and 12–14; and adnexal tenderness was from 10 studies, ID numbers 1, 3–6, 8–10, and 12–13. The pelvic tenderness, cervical motion tenderness, and adnexal tenderness results are shown in Table 2. (Table S4).

On pelvic tenderness synthesis, the mean and 95% confidence interval (CI) of sensitivity and specificity were 0.81 95% CI [0.67–0.90] and 0.40 95% CI [0.25–0.57], respectively. The AUC was 0.66 95% CI [0.62–0.70]. The I^2 statistic of sensitivity and specificity were 96.8, 95% CI [95.9–97.7] and 95.9, 95% CI [94.6–97.2], respectively.

The other cervical motion tenderness and adnexal tenderness syntheses results are shown in Table 2. Furthermore, Figure 3 above shows the HSROC model of pelvic tenderness, cervical motion tenderness, and adnexal tenderness in the main.

As shown in Table 1, the studies by Farrukh include cases of cervicitis, which was not PID. We analyzed the data excluding it and performed synthesis again. The results are shown in Table 2 and Figure 3.

On sensitivity analysis, the 13 studies limited to the PID study (ID 1–12, 14). The results show sensitivity and specificity of 0.83 [0.70–0.91] and 0.38 [0.22–0.57], respectively. The AUC was 0.69.

4 | DISCUSSION

Our meta-analysis implies that pelvic tenderness has moderate-high sensitivity, around 80%, and low specificity, at around 40%. Thus, pelvic examination appears to be useful for screening suspected PID patients, while specificity and AUC are low. Clinicians should know that PID cannot be conclusively excluded by the absence of pelvic tenderness.

TABLE 1 The characteristics and data of the 14 studies eligible for synthesis

ID	Author	Year	Country	Design	Study type	Disease	Method of PID diagnosis	Participants	TP	FP	FN	TN	Sensitivity [95%CI]	Specificity [95%CI]	PT	CMT	AT
1	Bongard et al.	1985	US	Pros	NR	PID	Laparoscopy, laparotomy or clinical course	81	37	10	8	26	0.82 [0.68, 0.92]	0.72 [0.55, 0.86]	–	◯	◯
2	Najem et al.	1985	US	Retro	NR	PID	Laparotomy	145	33	79	6	27	0.85 [0.69, 0.94]	0.25 [0.18, 0.35]	◯	–	–
3	Muylder et al.	1986	Zimbabwe	NR	NR	PID	Mixed method, including laparoscopy	158	101	30	0	27	1.00 [0.96, 1.00]	0.47 [0.34, 0.61]			◯
4	Morcos et al.	1993	US	NR	NR	PID	Laparoscopy	176	107	29	27	13	0.80 [0.72, 0.86]	0.31 [0.18, 0.47]	–	◯	◯
5	Webster et al.	1993	US	Retro	NR	PID	Mixed method, including laparoscopy	126	57	23	11	35	0.84 [0.73, 0.92]	0.60 [0.47, 0.73]	–	◯	◯
6	Stacey et al.	1994	UK	NR	NR	PID	Laparoscopy	81	6	42	5	28	0.55 [0.23, 0.83]	0.40 [0.28, 0.52]	–	◯	◯
7	Avan et al.	2001	Pakistan	Retro	CC	PID	Laparoscopy	129	21	25	22	55	0.49 [0.33, 0.65]	0.64 [0.53, 0.74]	◯	–	–
8	Peipert et al.	2001	US	Pros	CS	Endometritis	Endometrial biopsy	651	285	297	26	43	0.92 [0.88, 0.94]	0.13 [0.09, 0.17]		◯	◯
9	Hernando et al.	2002	Colombia	NR	CS	PID	Mixed method, including laparoscopy	61	20	18	11	12	0.65 [0.45, 0.81]	0.40 [0.23, 0.59]	–	◯	◯
10	Eckert et al.	2002	US	Pros	CS	Endometritis Salpingitis	Endometrial stroma or endometrial surface epithelium pathology or laparoscopy	168	70	11	56	31	0.56 [0.46, 0.64]	0.74 [0.58, 0.86]	–	◯	◯
11	Simms et al.	2003	Sweden	NR	NR	PID	Laparoscopy	623	489	128	5	1	0.99 [0.98, 1.00]	0.01 [0,00, 0.04]	◯	–	–
12	Romsan et al.	2013	Sweden	Pros	NR	Endometritis Salpingitis	Endometrial histology with laparoscopy	52	19	15	7	6	0.73 [0.52, 0.88]	0.29 [0.11, 0.52]		◯	◯
13	Farrukh et al.	2018	US	Pros	NR	PID cervicitis	Urine test for chlamydia and gonorrhea	288	30	80	49	129	0.38 [0.27, 0.50]	0.62 [0.55, 0.68]	◯	◯	◯
14	Munros et al.	2021	Spain	Pros	NR	PID	Mixed method, including laparoscopy	75	44	4	18	9	0.71 [0.58, 0.82]	0.69 [0.39, 0.91]	◯[a]	◯	◯[a]

Abbreviations: AT, adnexal tenderness; CC, case–control; CMT, cervical motion tenderness; CS, cross-sectional; FN, false negative; FP, false positive; NR, not reported; PID, pelvic inflammatory disease; Pros, prospective; PT, pelvic tenderness; Retoro, retrospective; TN, true negative; TP, true positive.

[a]Cervical motion tenderness or uterine tenderness.

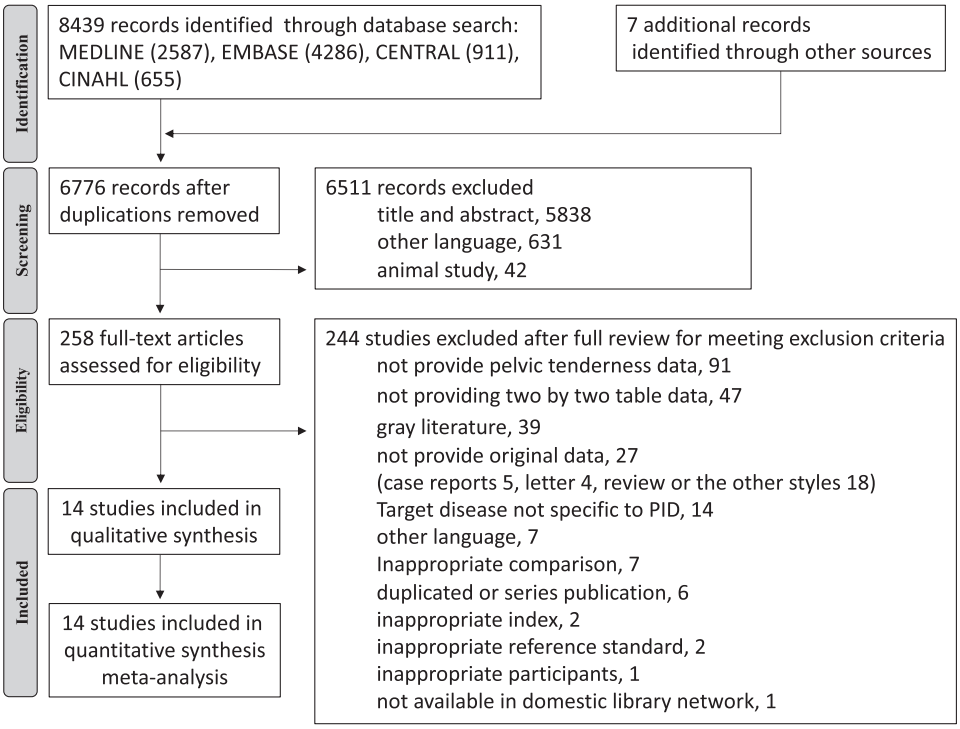

FIGURE 1 PRISMA literature search flowchart

The main and additional analyses show that pelvic tenderness, cervical motion tenderness, and adnexal tenderness have moderate-high sensitivity ranging from 73% to 88%. However, all three negative likelihood ratios were around 0.4–0.5, underscoring the fact that clinical usage of all three types of tenderness should be approached with caution. High sensitivity alone cannot be used to rule out the target disease; a low negative likelihood ratio can allow clinicians to rule out the target disease, but high sensitivity does not necessarily imply a low negative likelihood ratio, as in our findings. Negative results on a test with a relatively high negative likelihood ratio can decrease posttest probability only when the pretest probability was low. Thus, when the pretest probability of PID is relatively high, a negative pelvic tenderness test cannot be used clinically to exclude the possibility of PID. Hence, it is conceivable that pelvic tenderness can aid in exclusion when PID prevalence is low, but it may not be otherwise useful in clinical practice.

Our results support the moderate-to-high sensitivity of pelvic tenderness and conflict with the extremely low sensitivity found in the studies by Farrukh.[8] One significant reason for their reported low sensitivity seems to be that their studies included patients with cervicitis. Generally, cervicitis alone tends not to cause abdominal and pelvic tenderness as a symptom.[24] Further, cervicitis causes cervical motion tenderness and adnexal tenderness at lower rates than PID.[25] However, on the meta-synthesis, there was a concern that presence of cervicitis patients may strongly affect the main results. We therefore conducted the additional analyses excluding non-PID diseases, resulting in moderate-to-high sensitivity and the similar diagnostic values in all three examinations. (Table 2, Figure 3.) Furthermore, because laparoscopy alone is not a perfect gold reference standard for PID diagnosis, we conducted sensitivity analysis restricting to mixed methods including laparoscopy. The results of our sensitivity analysis support the high sensitivity of pelvic tenderness. However, the existence of various reference standards still remains a concern.

We took various measures to cope with the unique difficulties of conducing a meta-synthesis of PID studies. Because PID represents a diverse group of diseases, we carefully conducted our literature search and evaluation of candidate studies. First, we obtained support from two experienced librarians to conduct a systematic literature search for PID using multiple literature search tools, including thesaurus matches. Second, the literature search process was conducted according to the PRISMA flow chart, and we followed the PRISMA DTA guidelines. Furthermore, bias assessment was performed with QUADAS-2. As a result, critical bias in quality assessment was not implied, and no publication bias was confirmed.

Furthermore, we employed a bivariate random effect model and a hierarchical summary receiver operating characteristic model, accounting for heterogeneity and incomplete reference standards to the greatest extent possible. Furthermore, the HSROC model allows meta-analysis of diagnostic accuracy studies to address within- and between-study variability, and the HSROC model is applicable without a gold standard reference test.[11,26,27] However, we would not cope with heterogeneity and imperfect standards completely. In general, diagnostic studies tend to have different study settings, leading to strong heterogeneity between studies.[18] Similarly, given the fact that there is no established single reference standard,

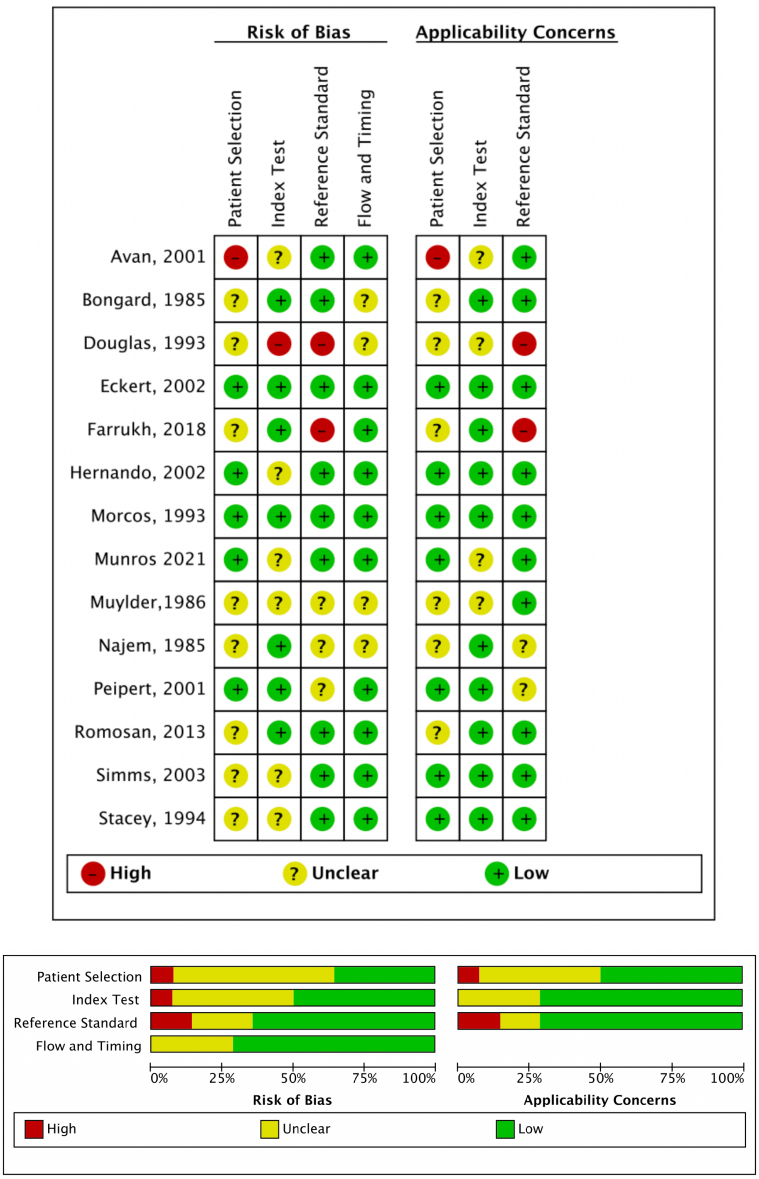

FIGURE 2 QUADAS-2 results; risk of bias and applicability concerns summary table

multiple methods are likely to be used to confirm PID in clinical practice, suggesting strong heterogeneity within studies. There are some novel Bayesian methods, such as latent class analysis, which address imperfect reference standards.[28] However, this method is not well established as of 2021, and managing heterogeneity remains a pressing obstacle to be solved.

We restricted our literature search to English language publications, which is a potential limitation of our study. While we included 14 studies from all continents except Australia, most are from European countries and the United States of America. If we had included non-English studies, additional results may have appeared. Next, we did not statistically evaluate publication bias and admit that our study selection process is biased toward selecting studies with positive results. Thus, our results may be overestimated.

As a meta-analysis, our study is limited by the quality of the included studies. PID diagnosis is complex and lacks a gold standard; some studies used only laparoscopy as a reference standard, and reference standards varied among studies. Although laparoscopy is important in confirming the diagnosis of PID, it is not a perfect diagnostic gold standard; it is an invasive procedure, which cannot confirm or exclude inapparent endometritis, endometriosis, or mild inflammation of the fallopian tube and may not be available in emergency situations. We therefore restricted the studies that implemented composite methods including laparoscopy as a sensitivity analysis. Further, there were some informational gaps regarding study type (prospective or retrospective) and quality assessment. Diagnostic meta-analyses often include studies with various designs, populations, and reference standards,[19] and we had to accommodate the informational gaps as a part of our analysis.

5 | CONCLUSION

Our meta-analysis strongly supports the moderate-to-high sensitivity and low specificity of pelvic examinations for PID, implying that pelvic tenderness can be useful in gynecological exams when PID or

TABLE 2 Results for pelvic tenderness, cervical motion tenderness, and adnexal tenderness before and after limiting to PID cases

Main analyses				Sensitivity analyses			
Physical examination	**Parameter**	**Estimate**	**95% CI**	**Physical examination**	**Parameter**	**Estimate**	**95% CI**
Pelvic tenderness	Sensitivity	0.81	[0.67–0.90]	Pelvic tenderness	Sensitivity	0.83	[0.70, 0.91]
	Specificity	0.4	[0.25–0.57]		Specificity	0.38	[0.22–0.57]
	Positive likelihood ratio	1.35	[1.08–1.68]		Positive likelihood ratio	1.35	[1.07–1.70]
	Negative likelihood ratio	0.48	[0.30–0.76]		Negative likelihood ratio	0.44	[0.27–0.70]
	AUC	0.66	[0.62–0.70]		AUC	0.69	[0.64–0.72]
Cervical motion tenderness	Sensitivity	0.72	[0.57–0.83]	Cervical motion tenderness	Sensitivity	0.78	[0.71–0.84]
	Specificity	0.5	[0.34–0.66]		Specificity	0.46	[0.31–0.62]
	Positive likelihood ratio	1.43	[1.11–1.85]		Positive likelihood ratio	1.45	[1.13–1.87]
	Negative likelihood ratio	0.57	[0.39–0.82]		Negative likelihood ratio	0.47	[0.35–0.64]
	AUC	0.65	[0.03–0.99]		AUC	0.73	[0.03–1.00]
Adnexal tenderness	Sensitivity	0.87	[0.64–0.96]	Adnexal tenderness	Sensitivity	0.9	[0.76–0.97]
	Specificity	0.27	[0.12–0.52]		Specificity	0.22	[0.10–0.42]
	Positive likelihood ratio	1.2	[1.00–1.50]		Positive likelihood ratio	1.2	[1.00–1.40]
	Negative likelihood ratio	0.48	[0.22–1.05]		Negative likelihood ratio	0.44	[0.19–1.03]
	AUC	0.6	[0.05–0.98]		AUC	0.62	[0.05–0.98]

Abbreviations: AUC, Area under the curve; CI, Confidence interval.

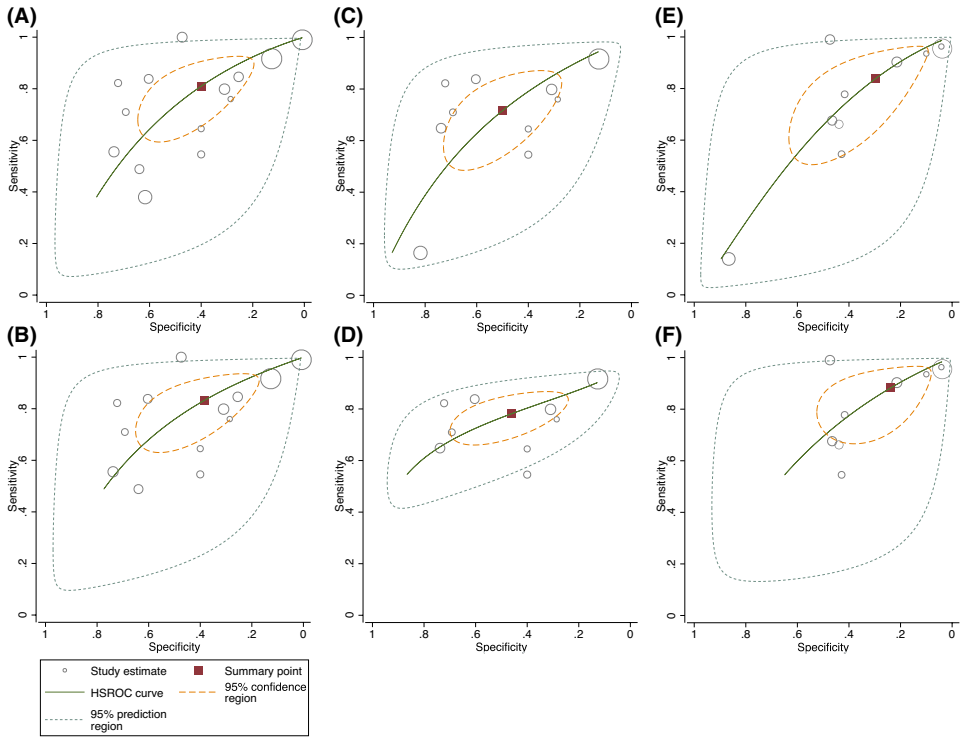

FIGURE 3 Bayesian hierarchical summary receiver operating characteristic model of pelvic tenderness, cervical motion tenderness, and adnexal tenderness. (A) pelvic tenderness main analysis, (B) pelvic tenderness additional analysis, (C) cervical motion tenderness main analysis, (D) cervical motion tenderness additional analysis, (E) adnexal tenderness main analysis, (F) adnexal tenderness additional analysis

STDs are suspected among a cohort whose presumed prevalence of PID is low, i.e., screenings at nonemergency or regular clinic visits. Our study thus provides an answer to the larger question of the clinical diagnostic value of pelvic examinations.

ACKNOWLEDGMENTS

The authors would like to thank Shinichi Abe and Izumi Osaki at the Academic Information Center of the Jikei University School of Medicine for systematic literature search support and Allen Paul Heffel for English language editing. This research was supported by The Jikei University Research Fund for Graduate Students.

CONFLICT OF INTEREST

HI, YSu, YSa, NS, and TA have no conflicts of interest relevant to the content of this research. MM received lecture fees and lecture travel fees from the Center for Family Medicine Development of the Japanese Health and Welfare Co-operative Federation. MM is an adviser for the Center for Family Medicine Development Practice-Based Research Network and a program director of the Jikei Clinical Research Program for Primary-care. MM's son-in-law worked at IQVIA Services Japan K.K., which is a contract research organization and a contract sales organization. MM's son-in-law works at SYNEOS HEALTH CLINICAL K.K., which is a contract research organization and a contract sales organization. The above descriptions do not match the COI contents of the JGFM.

ORCID

Hiroyoshi Iwata https://orcid.org/0000-0002-6949-0795
Takuya Aoki https://orcid.org/0000-0002-8232-2155

REFERENCES

1. Brunham RC, Gottlieb SL, Paavonen J. Pelvic inflammatory disease. N Engl J Med. 2015;372(21):2039–48. https://doi.org/10.1056/NEJMra1411426
2. Crossman SH. The challenge of pelvic inflammatory disease. Am Fam Physician. 2006;73(5):859–64.
3. Kreisel K, Torrone E, Bernstein K, Hong J, Gorwitz R. Prevalence of pelvic inflammatory disease in sexually experienced women of reproductive age - United States, 2013–2014. Morb Mortal Wkly Rep. 2017;66(3):80–3. https://www.cdc.gov/mmwr/volumes/66/wr/mm6603a3.htm
4. Centers for Disease Control and Prevention. Pelvic Inflammatory Disease (PID) - CDC Fact Sheet. c2015. Available from: https://www.cdc.gov/std/pid/stdfact-pid.htm. Accessed 12 Sep 2020.
5. Haggerty CL, Ness RB. Newest approaches to treatment of pelvic inflammatory disease: a review of recent randomized clinical trials. Clin Infect Dis. 2007;44(7):953–60. https://doi.org/10.1086/512191
6. Centers for Disease Control and Prevention. Pelvic Inflammatory Disease (PID). c.2015. Available from: https://www.cdc.gov/std/tg2015/pid.htm. Accessed 12 Sep 2019.
7. Peipert JF, Ness RB, Blume J, Soper DE, Holley R, Randall H, et al. Clinical predictors of endometritis in women with symptoms and signs of pelvic inflammatory disease. Am J Obstet Gynecol. 2001;184(5):856–63. https://doi.org/10.1067/mob.2001.113847

8. Farrukh S, Sivitz AB, Onogul B, Patel K, Tejani C. The additive value of pelvic examinations to history in predicting sexually transmitted infections for young female patients with suspected cervicitis or pelvic inflammatory disease. Ann Emerg Med. 2018;72(6):703–12. https://doi.org/10.1016/j.annemergmed.2018.05.004
9. Mealey K, Braverman PK, Koenigs LMP. Why a pelvic exam is needed to diagnose cervicitis and pelvic inflammatory disease. Ann Emerg Med. 2019;73(4):424–5. https://doi.org/10.1016/j.annemergmed.2018.11.028
10. National Institute for Health Research. PROSPERO (international prospective register of systematic reviews). c.2011–2020. Available from: https://www.crd.york.ac.uk/prospero/. Accessed 12 Sep 2020.
11. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009;62(10):1006–12. https://doi.org/10.1016/j.jclinepi.2009.06.005
12. McInnes MDF, Moher D, Thombs BD, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies the PRISMA-DTA statement. JAMA - J Am Med Assoc. 2018;319(4):388–96. https://doi.org/10.1001/jama.2017.19163
13. Kim KW, Lee J, Choi SH, Huh J, Park SH. Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers-part I. general guidance and tips. Korean J Radiol. 2015;16(6):1175–87. https://doi.org/10.3348/kjr.2015.16.6.1175
14. Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M, et al. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. Int J Technol Assess Health Care. 2012;28(2):138–44. https://doi.org/10.1017/S0266462312000086
15. Whiting PF, Rutjes AWS, Westwood ME, et al. Quadas-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529–36. https://doi.org/10.1186/1471-2288-3-25
16. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med. 2001;20(19):2865–84.
17. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005;58(10):982–90.
18. Lee J, Kim KW, Choi SH, Huh J, Park SH. Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers-part II. Statistical methods of meta-analysis. Korean J Radiol. 2015;16(6):1188–96.
19. Dwamena B, Sylvester R, Carlos R. midas: stata module for meta-analytical integration of diagnostic test accuracy studies. Statistical Software Components. Boston College Department of Economics; 2007. http://www.bc.edu/repec/bocode/m/midas.pdf
20. Collaboration NCCTC. Review manager (RevMan) version 5.3. Copenhagen: The Nordic Cochrane Centre. The Cochrane Collaboration; 2014.
21. StataCorp. Stata: Release 13. [CD-ROM]. College Station, TX: StataCorp LP; 2013.
22. Harbord R. Metandi: stata module for meta-analysis of diagnostic accuracy. Statistical software components. Boston College Department of Economics; 2008. http://fmwww.bc.edu/RePEc/bocode/m/metandi.html
23. Krishna UR, Sheth SS, Motashaw ND. Place of laparoscopy in pelvic inflammatory disease. J Obs Gynaecol India. 1979;29(3):505–10.
24. Powel AM, Nyirjesy P. Acute cervicitis. UpToDate, In: Post, T.W. (ed.) UpToDate. Waltham, MA: UpToDate Inc.; 2019. https://www.uptodate.com/contents/acute-cervicitis
25. Woods JL, Bailey SL, Hensel DJ, Scurlock AM. Cervicitis in adolescents: do clinicians understand diagnosis and treatment? J Pediatr Adolesc Gynecol. 2011;24(6):359–64. https://doi.org/10.1016/j.jpag.2011.06.006
26. Dendukuri N, Schiller I, Joseph L, Pai M. Bayesian meta-analysis of the accuracy of a test for tuberculous Pleuritis in the absence of a gold standard reference. Biometrics. 2012;68(4):1285–93. https://doi.org/10.1111/j.1541-0420.2012.01773.x
27. Liu Y, Chen Y, Chu H. A unification of models for meta-analysis of diagnostic accuracy studies without a gold standard. Biometrics. 2015;71(2):538–47. https://doi.org/10.1111/biom.12264
28. Chikere CMU, Umemneku Chikere CM, Wilson KJ, Joy Allen A, Vale L. Comparative diagnostic accuracy studies with an imperfect reference standard - a comparison of correction methods. BMC Med Res Methodol. 2021;21:67. https://doi.org/10.1186/s12874-021-01255-4

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Iwata H, Sugiyama Y, Satoi Y, Sasamoto N, Aoki T & Matsushima M. Diagnostic accuracy of pelvic examination in pelvic inflammatory disease: A meta-analysis. J Gen Fam Med. 2022;23:384–392. https://doi.org/10.1002/jgf2.572