

# Accuracy of real-time PCR assays for human papillomavirus using urine samples: a systematic review and meta-analysis

Byeong-Min Park,<sup>1</sup> Soohyun Kim,<sup>2</sup> Jieun Choi,<sup>3</sup> Yoonkyung Song,<sup>1</sup> Seungman Park<sup>1</sup>

**AUTHOR AFFILIATIONS** See affiliation list on p. 14.

**ABSTRACT** This study aims to assess the diagnostic accuracy of real-time PCR assays for detecting human papillomavirus (HPV) in urine samples through a systematic review and meta-analysis. A comprehensive search of PubMed, Embase, and Springer databases (2014–2024) was conducted. Studies comparing urine-based HPV tests with cervical samples as the reference standard were included. Diagnostic accuracy measures such as sensitivity, specificity, diagnostic odds ratio (DOR), likelihood ratios (LR+ and LR–), percent agreement, and Cohen’s kappa were calculated. Heterogeneity was assessed using the Higgins’  $I^2$  index, and subgroup analyses were performed based on HPV test type and urine volume. The study revealed that 15 studies met the inclusion criteria, with pooled sensitivity of urine-based HPV tests at 0.82 (95% CI, 0.78–0.86), specificity at 0.91 (95% CI, 0.87–0.94), positive LR at 9.5 (95% CI, 6.3–14.3), negative LR at 0.19 (95% CI, 0.16–0.24), DOR at 49 (95% CI, 32–75), and the area under the curve at 0.92 (95% CI, 0.90–0.94), with significant heterogeneity observed ( $I^2 > 50\%$ ), particularly in sensitivity and specificity, and subgroup analysis indicating that urine volumes  $\leq 20$  mL demonstrated higher sensitivity compared to those  $> 20$  mL, despite this finding being based on a limited number of studies. Results suggest that urine-based HPV testing shows strong diagnostic accuracy and could be a viable alternative to cervical swabs, with potential benefits for increasing screening accessibility, especially in areas with limited healthcare resources, despite some variability and limitations in the data.

**IMPORTANCE** This study is significant as it thoroughly evaluates the diagnostic accuracy of real-time PCR assays for human papillomavirus (HPV) detection in urine samples through a rigorous systematic review and meta-analysis. By integrating data from multiple databases and comparing urine-based HPV tests with the established cervical sample reference standard, the study provides valuable insights into the effectiveness and reliability of non-invasive HPV screening methods. The findings demonstrate that urine-based tests exhibit high sensitivity and specificity, offering a promising alternative to traditional cervical swabs. This advancement has significant implications for increasing accessibility to HPV screening, particularly in under-resourced settings, thereby potentially enhancing cervical cancer prevention efforts on a broader scale. The study not only fills a critical gap in HPV screening methodologies but also supports the development of more inclusive and practical public health strategies for combating cervical cancer.

**KEYWORDS** human papillomavirus, cervical cancer screening, real-time PCR, urine-based testing meta-analysis, diagnostic accuracy

Human papillomavirus (HPV) infection is the most significant risk factor for cervical cancer, and approximately 70% of cervical cancer cases are caused by HPV types 16 and 18 (1). Since cervical cancer can be effectively treated if detected early in precancer-

**Editor** Randall Hayden, St Jude Children's Research Hospital, Memphis, Tennessee, USA

Address correspondence to Byeong-Min Park, hermani@ncc.re.kr.

The authors declare no conflict of interest.

**Received** 29 August 2024

**Accepted** 26 February 2025

**Published** 31 March 2025

Copyright © 2025 Park et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

ous lesions, regular screenings are essential for both prevention and timely intervention (2).

The participation rate for cervical cancer screening exhibits a statistically significant decline from developed nations to developing regions (3). For example, only around 20% of women in lower-income countries have undergone cervical cancer screening (4), in stark contrast to the more than 60% participation rate observed in high-income countries (5, 6). Numerous factors contribute to the low uptake of cervical cancer screening in underdeveloped nations, with one of the primary challenges being the scarcity of medical personnel (7). Paradoxically, even in developed countries with relatively ample healthcare resources, approximately 20%–30% of women fail to engage in cervical cancer screening programs (8).

In addition to improving screening rates, increasing HPV vaccination coverage is crucial for reducing cervical cancer incidence, particularly in developing countries (9). Highly effective vaccines, such as Gardasil and Cervarix, prevent infections caused by HPV types 16 and 18, which account for the majority of cervical cancer cases (10). However, barriers such as limited healthcare access and affordability have resulted in low vaccination rates in many low- and middle-income countries (11). Integrating HPV vaccination programs with existing cervical cancer screening initiatives could significantly enhance prevention efforts and reduce the global burden of cervical cancer (12).

Although cervical swab samples are the standard method for HPV testing, studies on urine-based HPV tests have been increasingly reported due to advantages such as non-invasive techniques and the development of new test methods for convenience sampling (13, 14). Commercial real-time polymerase chain reaction (PCR) assays provide the advantage of automated batch-mode testing (15), providing faster test results than Pap smears (16, 17). Since the results are presented as quantitative outputs, such as cycle threshold values or viral copy numbers, they enable objective interpretation and facilitate standardized monitoring over time (18). Therefore, this study aims to evaluate the diagnostic accuracy of the real-time PCR assays for HPV detection in urine samples via a systematic review and meta-analysis.

## MATERIALS AND METHODS

### Literature search and criteria for selection/exclusion

A comprehensive search of electronic databases including PubMed, Embase, and Springer was conducted, covering publications from 2014 to 2024. Key search terms such as “HPV,” “human papillomavirus,” “cervical cancer screening,” “real-time PCR,” and “urine” were employed during the database search. Articles were excluded if they were duplicates, were unrelated, or lacked sufficient data (Fig. 1). Studies utilizing non-real-time PCR methods, such as transcription-mediated amplification (TMA)-based assays (e.g., Aptima HPV test) (19), were intentionally excluded to ensure methodological consistency and minimize heterogeneity in the pooled analysis. Although TMA is a reliable diagnostic tool for detecting HPV, its amplification method fundamentally differs from real-time PCR. This difference causes variations in interpretation, making direct comparisons difficult. Cervical samples were used as the reference standard, while urine samples served as the index test to construct  $2 \times 2$  contingency tables for data analysis. In all 15 included studies, both cervical and urine samples were analyzed using the same real-time PCR assays, ensuring methodological consistency across sample types. This approach enabled a direct and reliable comparison of HPV detection performance between cervical and urine samples. A summary of the qualitative characteristics of the selected studies is provided in Table 1.

### Quality assessment

The quality of the studies was assessed by two independent reviewers (B.M. and S.H.) using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool (35). This QUADAS-2 tool comprises four key domains: patient selection, index test, reference

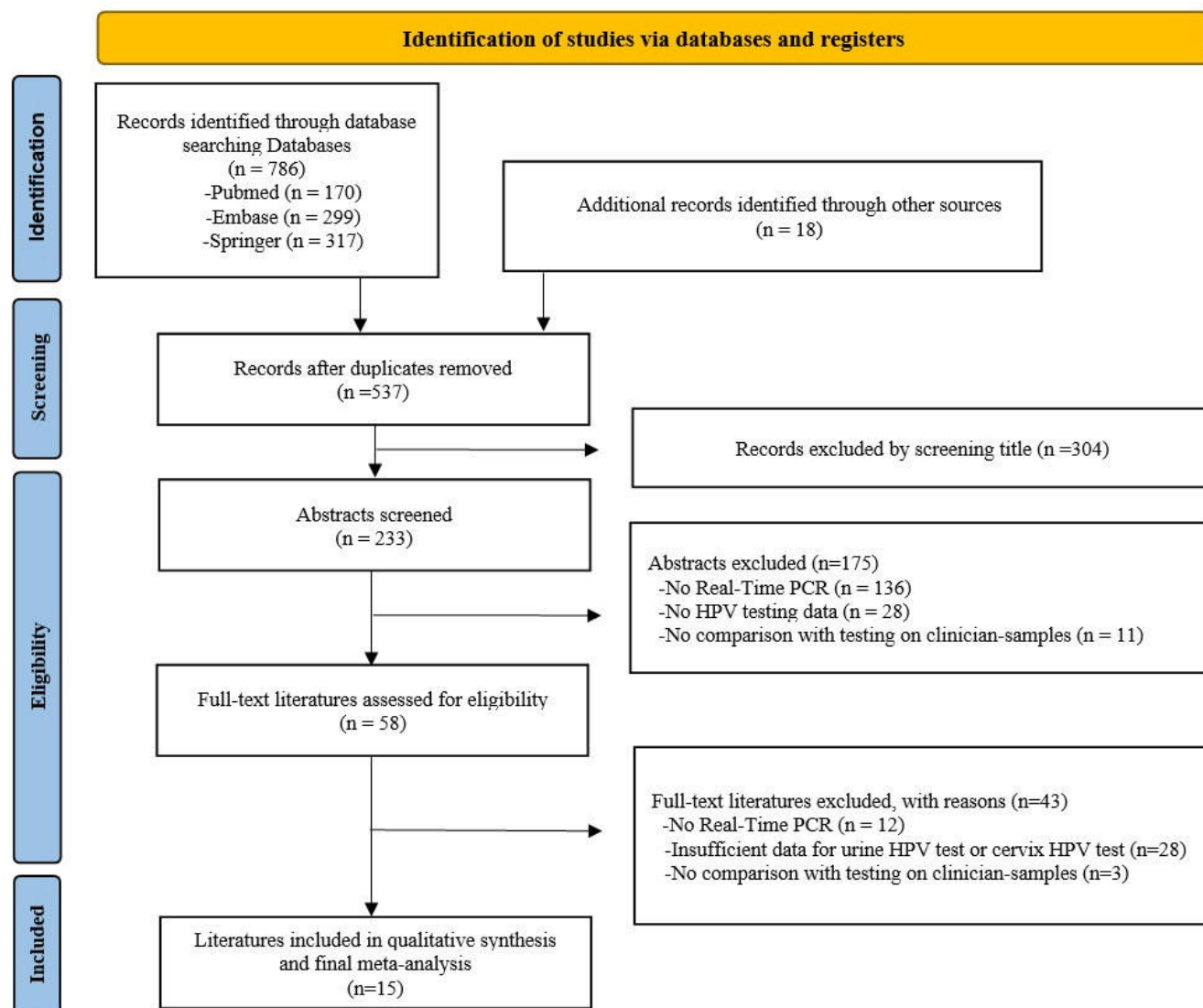


FIG 1 Flowchart illustrating the study selection process for eligible studies.

standard, and flow and timing. Each domain includes a systematic evaluation of both bias risk and concerns regarding applicability. For each detailed item within each domain, bias risk is categorized as either "yes," "no," or "unclear," while overall bias risk and applicability concerns are rated as "low," "high," or "unclear" for each domain. In cases of discordant evaluations, consensus was achieved through discussions between the reviewers.

All 15 included studies underwent quality assessment using the QUADAS-2, and no studies were excluded based on this assessment. The QUADAS-2 effectively identified potential biases across the included studies (36). Six studies lacked clarity in reporting patient selection, one study lacked clarity in reporting the index test, and nine studies lacked clarity in reporting flow and timing. Concerns regarding applicability were also flagged in five studies due to incomplete reporting of patient selection. These limitations were carefully considered during data synthesis and interpretation to ensure a balanced evaluation of the findings. Despite these flagged concerns, the overall reliability of the meta-analysis findings was supported by consistent diagnostic performance across studies, as reflected by high pooled sensitivity and specificity values (37). However, these flagged limitations underscore the importance of cautious interpretation in specific

TABLE 1 Qualitative characteristics of selected studies<sup>a</sup>

Year	Author (Ref)	TP	FP	FN	TN	Sample size	Population	Prevalence %	HPV detection assays	Sample preparation	Timing	Urine volume	Age (range)
2014	Bernal et al. (20)	57	9	6	53	125	Retrospective	50.4	Cobas 4800 HPV test	Pellet	First-voided urine	20 mL	Median 35.5 (21–65)
2015	Stanczuk et al. (21)	78	1	14	7	100	Retrospective	92.0	Cobas 4800 HPV test	Pellet	N/A	20 mL	Median 27.5 (21–60)
2016	Hagihara et al. (22)	98	5	22	115	240	General	50.0	Anyplex II HPV28 detection kit	Pellet	First-voided urine	N/A	Median 31 (19–58)
2016	Khunamornpong et al. (23)	24	6	11	82	123	Retrospective	28.5	Cobas 4800 HPV test	Pellet	First-voided urine	5 mL–50 mL	Mean 45.8 (± SD 8.1)
2017	Asciutto et al. (24)	141	27	8	39	215	Retrospective	69.3	Cobas 4800 HPV test	Unprocessed	First-voided urine	N/A	Mean 35.2 (19–71)
2017	Lim et al. (25)	57	3	21	119	200	Retrospective	39.0	Cobas 4800 HPV test	Pellet	First-voided, mid-stream urine	30 mL	N/A
2017	Lim et al. (25)	55	4	21	120	200	Retrospective	38.0	RealTime High Risk HPV test	Pellet	First-voided, mid-stream urine	30 mL	N/A
2018	Vergara et al. (26)	289	12	63	179	543	General	64.8	Non-commercial software	Pellet	First-voided urine	10 mL–15 mL	Mean 36 ± 11.1 (18–64)
2019	Sargent et al. (27)	45	3	11	20	79	Retrospective	70.9	RealTime High Risk HPV test	Unprocessed	First-voided urine	N/A	40% under 30 years
2019	Sargent et al. (27)	44	1	4	17	66	Retrospective	72.7	Cobas 4800 HPV test	Unprocessed	First-voided urine	N/A	N/A
2020	Tranberg et al. (28)	23	4	13	110	150	Retrospective	24.0	Cobas 4800 HPV test	Pellet	First-voided urine	10 mL–12 mL	30–59
2021	Cho et al. (29)	194	16	53	51	314	Retrospective	78.7	RealTime High Risk HPV test	Pellet	First-voided urine	30 mL	20–60
2021	Cho et al. (29)	164	13	66	71	314	Retrospective	73.2	Anyplex II HPV28 assay	Pellet	First-voided urine	30 mL	20–60
2021	Van Keer et al. (30)	247	36	46	164	493	Retrospective	59.4	RealTime High Risk HPV test	Unprocessed	First-voided urine	13 mL	Median 40 (19–72)
2022	Kim et al. (31)	19	7	7	177	210	General	12.4	PANA RealTyper HPV Screening Kit using CFX96	Pellet	Random voided urine	N/A	Mean 39.9 (20–85)
2022	Punyashthira et al. (32)	66	4	12	14	96	Retrospective	81.2	Anyplex II HPV high-risk	Unprocessed	First and random voided urine	10 mL–15 mL	Mean 47.5
2023	Martinelli et al. (33)	149	12	11	73	245	Retrospective	65.3	Anyplex II HPV28 assay	Unprocessed	N/A	20 mL	Mean 38.6 (17–67)
2024	Niliyanimit et al. (34)	34	4	4	198	240	General	15.8	Cobas 4800 HPV test	Pellet	First-voided urine	20 mL	Mean 46 (22–70)

<sup>a</sup>TP, true positive; FP, false positive; FN, false negative; TN, true negative; N/A, not available.

subgroups or contexts where generalizability may be limited. Figure 2 illustrates the quality assessment of the 15 included studies using the QUADAS-2.

## Statistical analysis

The primary objective was to summarize the overall diagnostic performance (38, 39). Diagnostic performance was evaluated in terms of sensitivity, specificity, diagnostic odds ratio (DOR), positive likelihood ratio (LR+), negative likelihood ratio (LR-), percent agreement, and Cohen's kappa, with results presented along with 95% confidence intervals (CIs). A bivariate random effects model was used to simultaneously consider sensitivity and specificity, and a summary receiver operating characteristic (SROC) curve was analyzed (40, 41). Results from different HPV detection tests within a single study were treated as separate studies. A forest plot was used to visually present the pooled diagnostic performance results. The Higgins' inconsistency index ( $I^2$ ) was calculated to assess heterogeneity among studies (42), and subgroup differences were tested using a random effects model to determine whether differences in HPV detection tests and urine volumes contributed to this heterogeneity. HPV detection tests were classified into four groups: Cobas 4800, Anyplex II, RealTime High-Risk, and others. Urine volume was classified as  $\leq 20$  mL and  $>20$  mL, and studies that did not report values or could not be classified were treated as missing. Publication bias was assessed using Deeks' funnel plot, which was developed to examine diagnostic performance (43). All statistical analyses were conducted using statistical software, including Stata version 12.0 (StataCorp, College Station, TX, USA) and R version 4.1.1 (R Foundation for Statistical Computing, Vienna, Austria), with statistical significance defined as  $P < 0.05$ .

## RESULTS

### Study selection and characteristics

A total of 804 articles were identified, 786 from primary databases (PubMed, Embase, and Springer) and 18 additional articles through other sources, including manual reference list searches and expert recommendations. After applying the inclusion and exclusion criteria, the final data set of 15 studies was derived exclusively from primary database searches, as none of the articles from other sources met the eligibility criteria. After removing duplicates, 537 records remained. These were systematically narrowed down through title and abstract screening, which excluded 479 records based on relevance and specific eligibility criteria. A subsequent full-text assessment further refined the selection, eliminating 43 studies that did not fully meet the inclusion criteria. Ultimately, 15 studies were chosen for qualitative synthesis and meta-analysis, forming the foundation of the review outcomes (20–34). Figure 1 illustrates the flow of the database search and literature selection process. Table 1 provides a summary of the selected studies' characteristics, including year of publication, author, true positive, false positive, false negative, true negative, sample size, population prevalence, HPV detection assays, sample preparation methods, urine volume, and participant age range. The risk of bias assessment identified concerns in patient selection, index test, reference standard, and flow and timing. Specifically, six studies lacked clarity in reporting patient selection, one study lacked clarity in reporting the index test, none of the studies lacked clarity in reporting the reference standard, and nine studies lacked clarity in reporting flow and timing. The assessment of concerns regarding applicability indicated that the study results might not be applicable to clinical practice situations in some cases, such as the five studies with incomplete reporting of patient selection (Fig. 2).

### Diagnostic performance and agreement in urine samples compared to cervical samples

The HPV real-time PCR assays using urine samples demonstrated a pooled sensitivity of 0.82 (95% CI, 0.78–0.86), specificity of 0.91 (95% CI, 0.87–0.94), LR+ of 9.5 (95% CI, 6.3–14.3), LR- of 0.19 (95% CI, 0.16–0.24), and diagnostic odds ratio of 49 (95% CI,

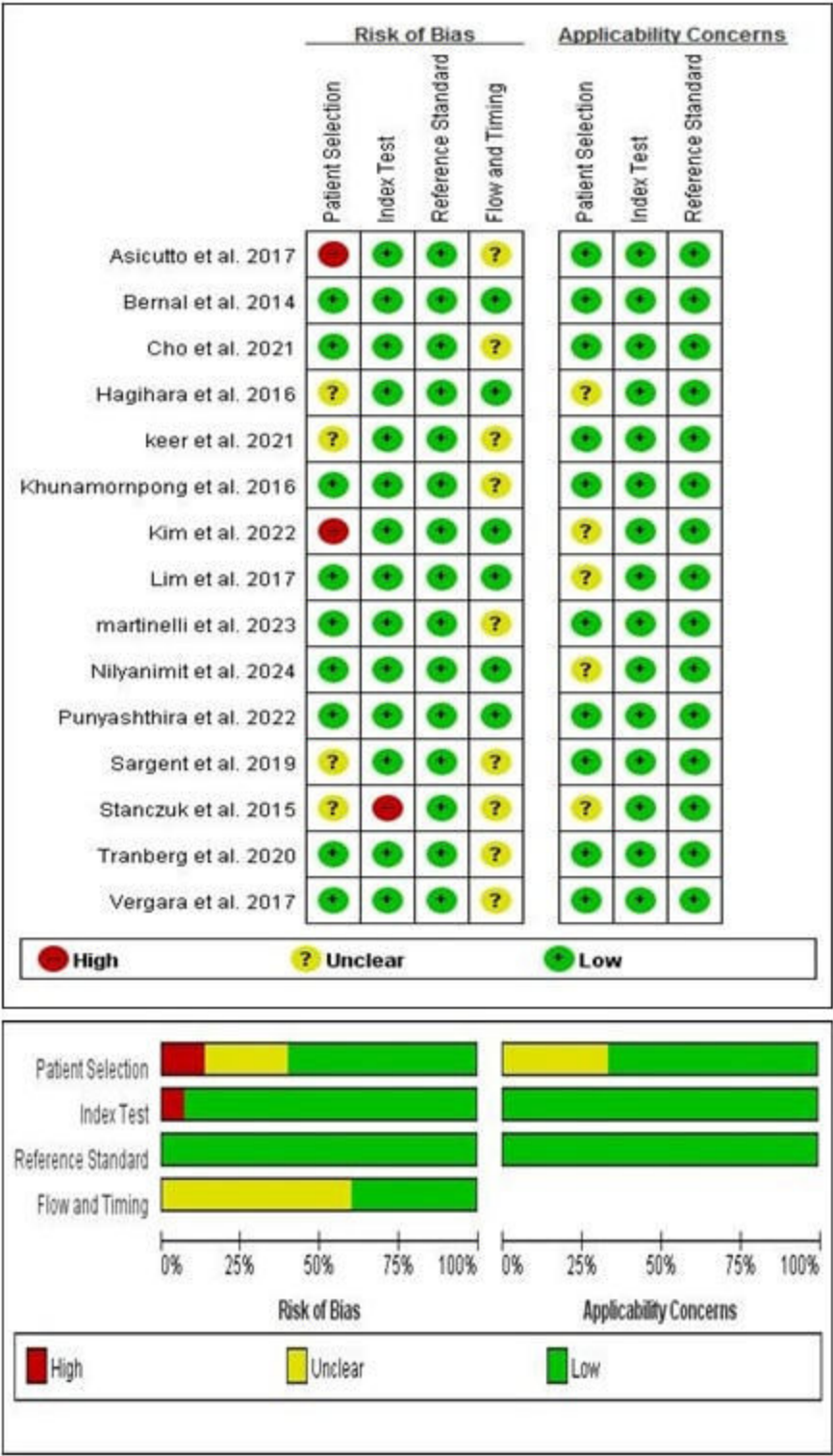


FIG 2 Quality assessment of the 15 included studies using QUADAS-2.

32–75) (Table 2). The area under the curve (AUC) of the SROC curve was 0.92 (95% CI, 0.90–0.94), and as a result of visually inspecting the heterogeneity of the SROC curve, the symmetry of the curve and the spread of individual studies were appropriate, leading to



TABLE 2 Summary of diagnostic performance for HPV testing on urine in the 15 studies reviewed<sup>a</sup>

Parameter	Estimate	95% CI
Sensitivity	0.82	0.78–0.86
Specificity	0.91	0.87–0.94
Positive likelihood ratio	9.5	6.3–14.3
Negative likelihood ratio	0.19	0.16–0.24
Diagnostic odds ratio	49	32–75

<sup>a</sup>95% CI, 95% confidence interval.

no suspicion of heterogeneity among the included studies (Fig. 3). The pooled kappa was 0.68 (95% CI, 0.62–0.73), and the percent agreement was 0.87 (95% CI, 0.84–0.90) (Fig. 4).

Heterogeneity, meta-regression, and publication bias

The Higgins'  $I^2$  index, used to quantify the degree of heterogeneity in a meta-analysis (42), showed values of 82.2% for sensitivity, 89.8% for specificity, 78.8% for kappa, and 87.2% for percent agreement (Fig. 4 and 5).  $I^2$  values greater than 50% were considered to

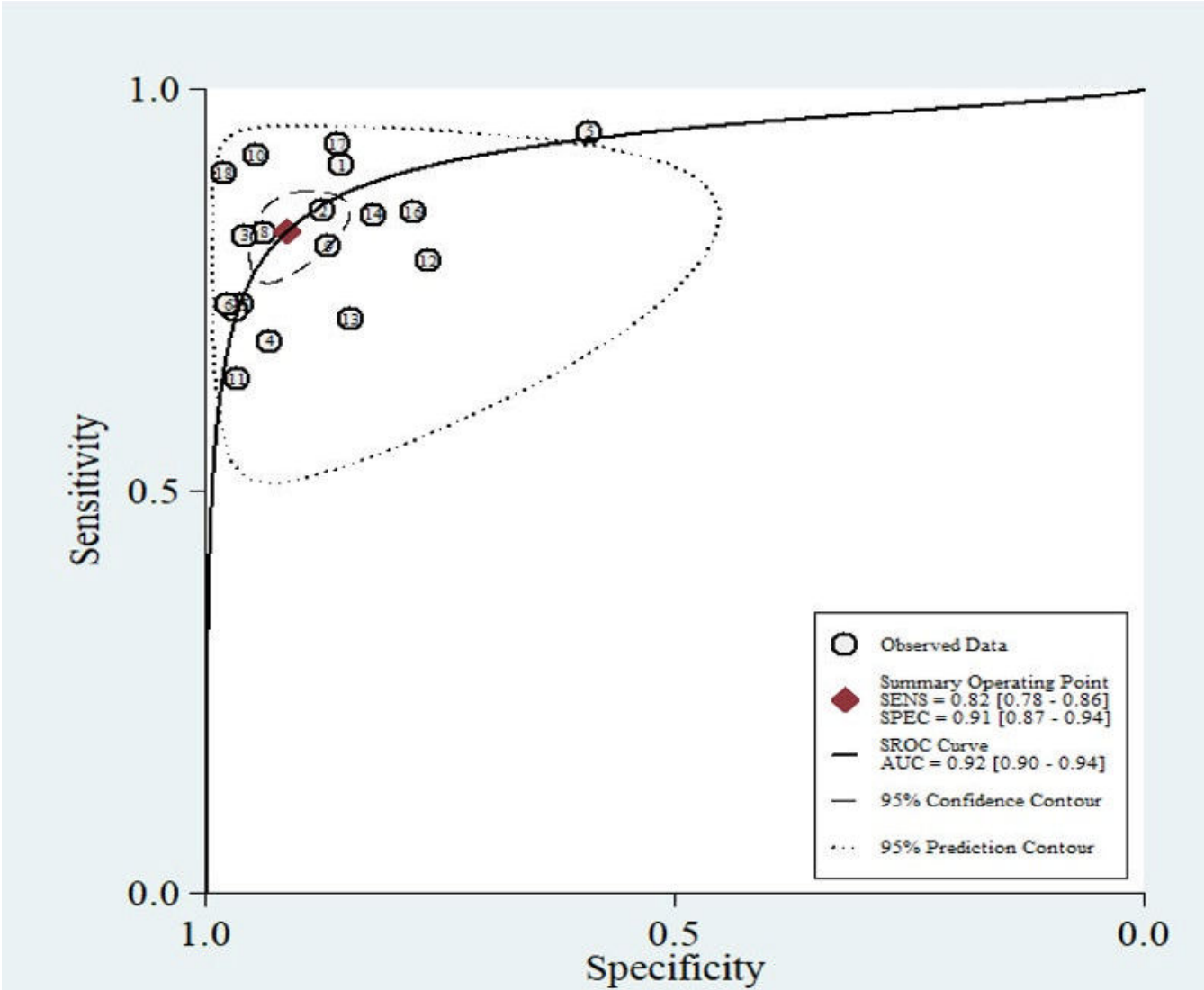


FIG 3 SROC curve with 95% confidence and prediction regions around the mean operating sensitivity and specificity point.

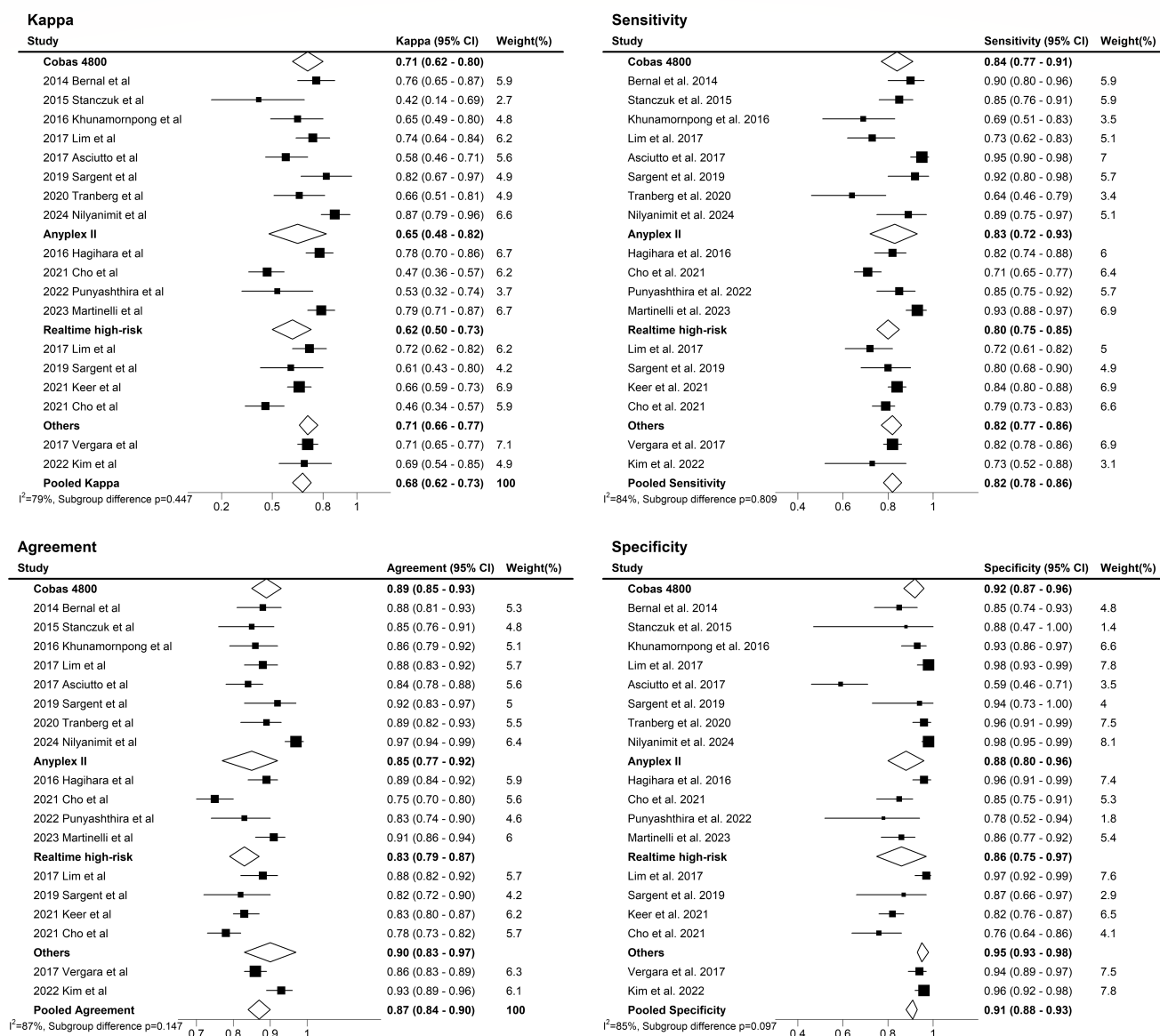


FIG 4 Meta-analysis assessing heterogeneity in HPV detection performance and agreement across commercial real-time PCR systems in urine.

indicate substantial heterogeneity. In cases of significant heterogeneity in meta-analysis, it is possible to explore the causes of heterogeneity (44). Accordingly, the subgroup analyses using a random effects model were conducted to assess the impact of HPV detection test type and urine volume based on the study characteristics.

In the HPV detection test subgroup, the subgroup difference test  $P$ -values for sensitivity, specificity, kappa, and percent agreement were 0.809, 0.097, 0.447, and 0.147, respectively, indicating no statistically significant difference. Although no statistically significant differences were observed in sensitivity and specificity, the Cobas 4800 test exhibited numerically higher diagnostic values, indicating potential clinical significance (Fig. 4). Further subgroup analysis based on urine volume showed no significant differences in specificity, kappa, and percent agreement, with  $P$ -values of 0.953, 0.175, and 0.135, respectively. However, a statistically significant difference in sensitivity was observed between urine volume subgroups (0.857 vs 0.746,  $P < 0.001$ ), with higher sensitivity observed in samples with low urine volume (Fig. 6).



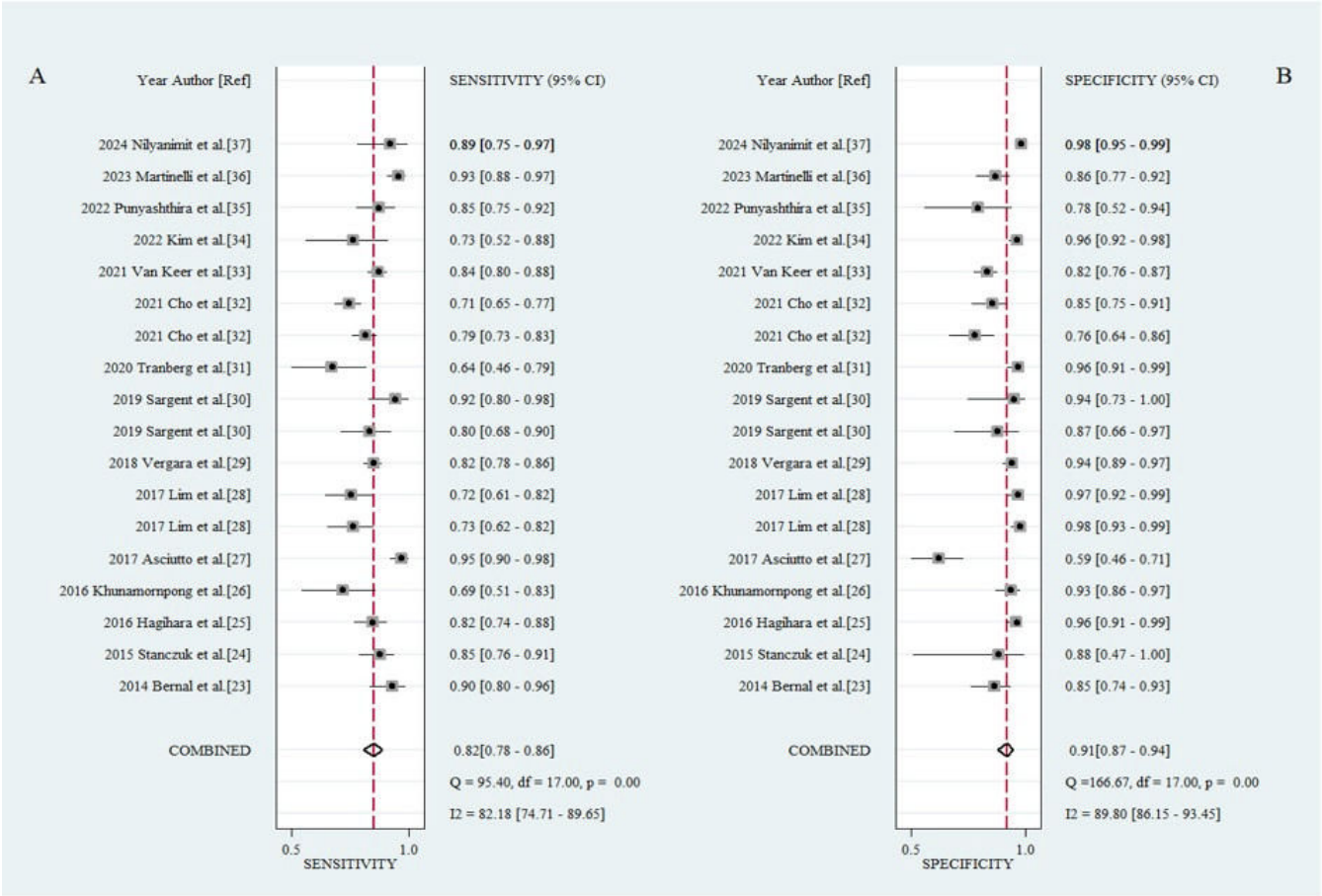


FIG 5 Forest plot illustrating (A) sensitivity and (B) specificity of real-time PCR assays for HPV detection in urine samples.

A funnel plot was employed to assess the presence of publication bias in the reviews (45). Deeks' test for publication bias (43) yielded a  $P$ -value of 0.62, indicating a low likelihood of publication bias in this meta-analysis (Fig. 7).

DISCUSSION

HPV is spread through sexual contact, and approximately 311,000 women died from cervical cancer in 2018 (46), with more than 85% of these deaths occurring in low- and middle-income countries (47), underscoring the urgent need for improved screening strategies. Numerous studies have been conducted to support women in medically vulnerable groups, particularly in cervical cancer screening (48). Among the promising alternatives, urine specimens have been widely studied as potential replacements for cervical samples (49, 50). Urine samples are particularly attractive due to their ease of collection in a non-invasive manner. Notably, first-void urine samples have demonstrated higher detection rates of high-risk HPV (51); however, they may be susceptible to vulvar contamination, which could limit their accuracy in detecting cervical HPV infection (52). This limitation must be considered when interpreting urine-based HPV testing results and highlights the need for further research to clarify the exact relationship between HPV detection in urine and cervical infection. Despite the potential advantages of urine-based testing, the current standard screening methods are the Pap test and the HPV test (53). These tests are considered the gold standard for detecting cervical cell abnormalities and HPV presence. Generally, in meta-analyses, the sensitivity and specificity of the HPV test using DNA PCR with liquid-based cytology have been found to be higher than the Pap test, which involves sampling cells directly from the cervix (54).

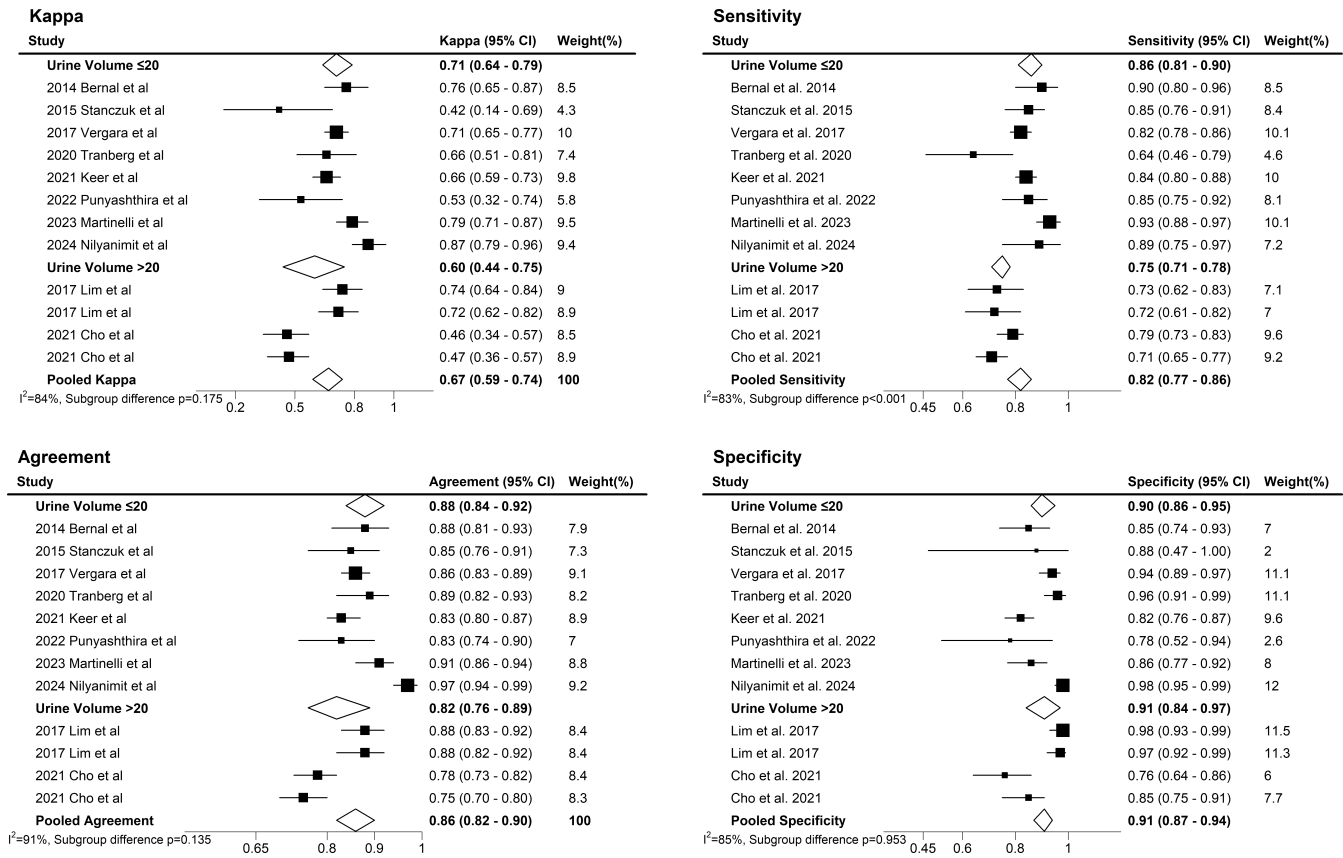


FIG 6 Effect of urine volume thresholds ( $\leq 20$  mL vs  $> 20$  mL) on HPV concordance and heterogeneity in the meta-analysis.

In this study, HPV real-time PCR assays using urine samples demonstrated a pooled sensitivity of 0.82 (95% CI, 0.78–0.86) and a specificity of 0.91 (95% CI, 0.87–0.94), indicating a high concordance between urine and cervical samples for HPV detection, while suggesting potential discrepancies that require further investigation. Additionally, to establish the diagnostic utility of a general screening test, the combined value of sensitivity and specificity should be at least 1.5 (150%), a threshold referenced in evidence-based literature, as a practical benchmark for balancing sensitivity and specificity, minimizing false positives and negatives, and ensuring reliable diagnostic performance (55, 56).

However, the pooled specificity of 0.91 suggests that HPV was detected in urine samples but not in cervical samples, indicating potential false positives (57). Most studies included in this meta-analysis did not conduct comprehensive discrepancy analyses to confirm these results. Future research should incorporate rigorous discrepancy resolution protocols, such as repeat testing, longitudinal follow-up, and co-testing with Pap smears, to gain a more accurate understanding of urine-based HPV testing.

In this study, the agreement between tests, as indicated by the percent agreement and kappa, was measured at 0.87 (95% CI, 0.84–0.90) and 0.68 (95% CI, 0.62–0.73), respectively. When assessing the consistency between tests, particularly with categorical data, researchers commonly employ statistical measures such as kappa or percent agreement (58). Generally, inter-rater reliability data for categorical variables are considered inadequate if percent agreement falls below 80% or if kappa is 0.6 or lower (59). The data were reviewed again due to a few studies reporting low kappa values; however, the results remained consistent. Notably, in the study by Stanczuk et al., the gold standard exhibited a distribution of 92% HPV-positive versus 8% HPV-negative cases. This uneven distribution likely contributed to the lower calculated kappa value (60, 61). These observations clearly highlight the inherent limitations of using kappa.

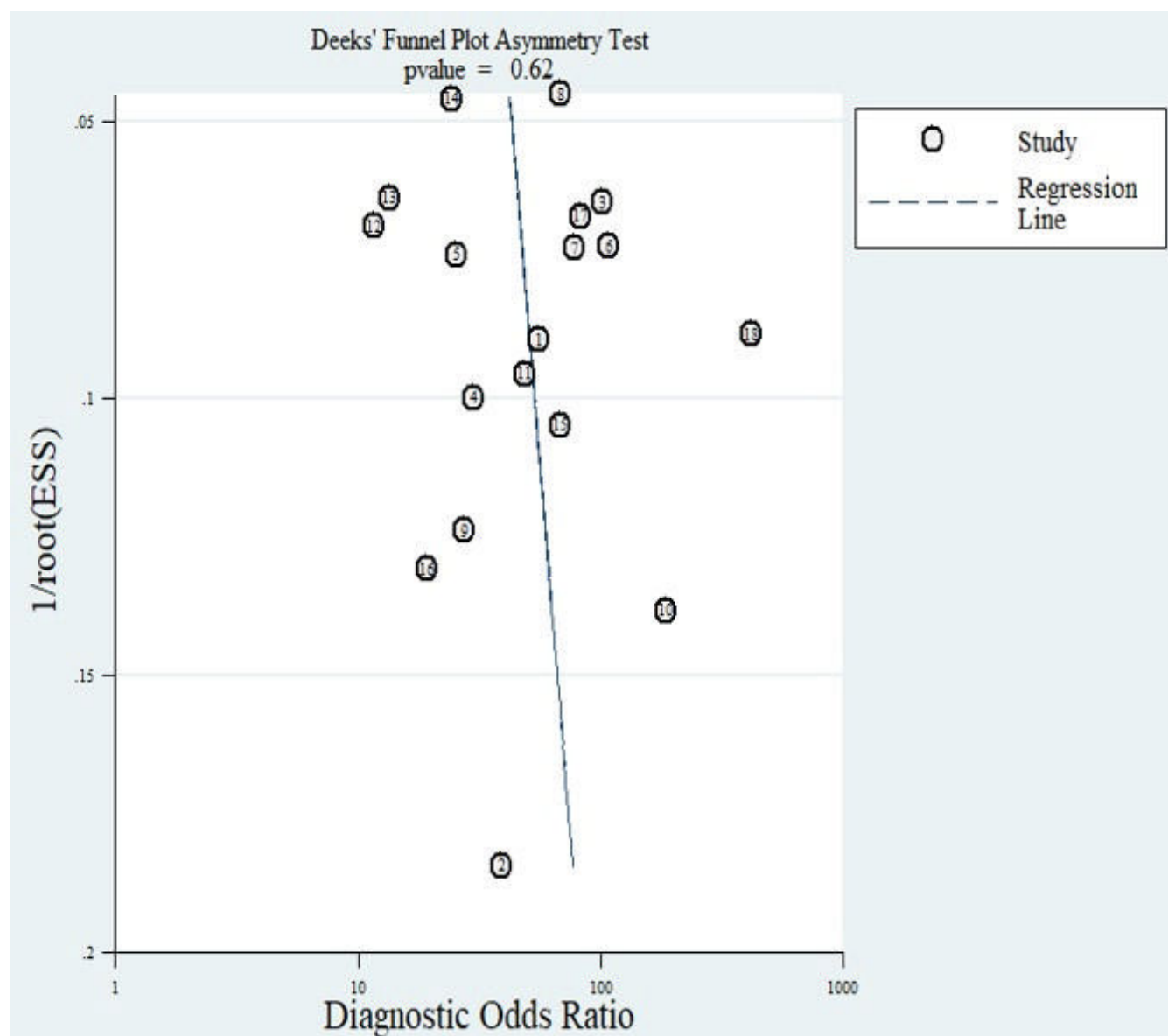


FIG 7 Deeks' funnel plot for assessing publication bias.

Specifically, a notable limitation of kappa is its sensitivity to variations in evaluator assignment, regardless of agreement (62). The degree of homogeneity in the data distribution can significantly influence kappa values, with greater uniformity leading to higher kappa values, and when marginal homogeneity is satisfied (i.e., when off-diagonal cells are equal,  $A1 = B1$ ), resulting in lower kappa values. In essence, the observed percent agreement may remain consistent in this study, while kappa can vary substantially depending on the data distribution. Although slightly below the commonly accepted threshold of 90% (63), we propose that percent agreement remains a more suitable measure for this study, given its relatively high overall value (0.87) and reduced susceptibility to prevalence-related bias observed in kappa values from skewed data distributions (64).

The Higgins'  $I^2$  values range from 0 to 100% and represent the proportion of inter-study variability as heterogeneity. Values of 25%, 50%, and 75% are interpreted as indicating low, moderate, and high estimates of heterogeneity, respectively (65). In this study,  $I^2$  values of 82.18 for sensitivity and 89.80 for specificity indicate high

heterogeneity according to these criteria. The included studies demonstrated variability in the sample preparation methods, with most utilizing centrifugation to concentrate urine and analyze the pellet. Some studies used unprocessed samples, and none independently tested the supernatant. This variability likely influenced assay sensitivity, as previous research has shown that preparation methods such as centrifugation can significantly increase HPV DNA concentration (66), thereby enhancing diagnostic accuracy. As shown in Table 1, prevalence rates ranged from 12.4% to 92.0%, with retrospective studies reporting higher rates due to selection biases in high-risk populations, whereas prospective studies targeting general populations reported lower prevalence (67). These differences, along with variability in sample preparation methods, urine volumes, and study population characteristics (51), substantially contributed to the observed heterogeneity (sensitivity:  $I^2 = 82.18\%$ ; specificity:  $I^2 = 89.80\%$ ; Fig. 5) in this meta-analysis. Rigorous protocols, including repeat testing, longitudinal follow-up, and co-testing with Pap smears, have been suggested in previous literature to improve diagnostic accuracy and address potential discrepancies (68). Future research should focus on the standardization of preparation methods, stratifying results by factors such as vaccination status and screening history, and prioritizing general screening populations to enhance the generalizability and comparability of findings. To this end, it is important to identify the causes of heterogeneity and conduct subgroup and sensitivity analyses (69, 70). Moreover, diagnostic tests commonly exhibit inherent variability due to factors such as diverse patient populations, variations in techniques, and other contributing factors (71, 72).

Previous studies have indicated that there is no statistically significant difference in HPV detection rates for urine volumes below 20 mL, and that first-void urine is more suitable than midstream urine (51, 73). In this meta-analysis, samples with urine volumes below 20 mL demonstrated significantly higher sensitivity compared to those with volumes exceeding 20 mL. This suggests that dilution effects may reduce the sensitivity of assays when larger urine volumes are used. In this meta-analysis, most of the included studies employed centrifugation to concentrate urine and analyze the pellet, a method known to enhance HPV DNA detection rates by increasing its concentration (74). However, as summarized in Table 1, only a few studies analyzed unprocessed urine samples, and none examined the supernatant independently. A limitation of this meta-analysis is that the fraction of urine analyzed was often unclear or randomly collected. Consequently, due to the limited number of studies that exclusively analyzed first-void urine, we included results from different sampling times when comparing urine volumes. Further studies are required to systematically evaluate the diagnostic impact of dilution and concentration methods.

The pooled sensitivity of 82% (95% CI, 0.78–0.86) for urine-based HPV testing, although lower than those of cervical sampling methods, should be interpreted in the context of its high specificity of 91% (95% CI, 0.87–0.94) and its potential to improve screening accessibility (Table 2; Fig. 5) (51). Subgroup analysis revealed that urine volumes  $\leq 20$  mL demonstrated significantly higher sensitivity compared to volumes  $> 20$  mL (0.857 vs 0.746,  $P < 0.001$ ; Fig. 6). This underscores the critical importance of standardized urine collection protocols in optimizing diagnostic accuracy (75). While a detailed discrepancy analysis of false-positive results was not feasible in this meta-analysis due to data limitations, potential sources of false positives, including contamination and assay variability, have been discussed in previous studies (76). Urine-based testing offers a promising complementary approach for HPV screening, particularly in settings with limited healthcare resources. Future studies should prioritize standardizing urine collection methods, improving assay specificity, and conducting large-scale prospective studies to validate the diagnostic performance of urine-based HPV testing (77).

Another limitation of this study is that the included studies did not consistently differentiate between screening-naïve versus previously screened populations or vaccinated versus unvaccinated individuals. Previous research has shown that vaccinated individuals typically exhibit lower HPV prevalence and are less likely to test positive for

HPV (78), while individuals with regular screening are more likely to identify cervical abnormalities at earlier stages, leading to improved diagnostic outcomes (79). The absence of stratification for these critical factors is likely to have contributed to variability in the pooled diagnostic accuracy. Future studies should systematically report these variables to enable comprehensive subgroup analyses and improve our understanding of HPV test performance across diverse populations.

In other research findings, meta-regression identified the first urine as a statistically significant source of heterogeneity, showing higher levels of high-risk HPV compared to other urine fractions (80, 81). Another study corroborated its potential as a urinary biomarker for cervical cancer screening (75). In this study, the pooled kappa and agreement for heterogeneity analysis of commercial PCR systems for HPV detection were highest for the Roche Cobas 4800, followed by the Seegene Anyplex II HPV28 detection and Anyplex II HPV HR detection, and then the Abbott RealTime High-Risk HPV. However, the test *P*-values comparing differences between subgroups, including those for non-commercial PCR systems, did not show statistically significant differences. Given that a limited number of studies can lead to inaccuracies in estimates of heterogeneity (82), it will be essential to collect more data to accurately assess and understand the underlying causes of heterogeneity in future investigations.

The LR indicates how much more likely a test result is to occur in patients with the disease compared to those without it (83). Likelihood ratios calculated as integrated estimates of sensitivity and specificity reveal crucial diagnostic insights (84). An LR of 1 indicates that the test is not diagnostically useful, while LR values above 10 for positive LR and below 0.1 for negative LR signify a test that is definitely clinically useful, with the potential to alter clinical decisions. Positive LR values above 5 and negative LR values below 0.2 provide strong diagnostic evidence (85). In this study, urine-based HPV testing demonstrated a positive LR of 9.5 (95% CI, 6.3–14.3) and a negative LR of 0.19 (95% CI, 0.16–0.24).

The DOR is another measure of diagnostic accuracy that is calculated by combining the likelihood ratios (LR+ and LR–). It ranges from 0 to infinity, with higher values indicating better discriminatory test performance (86). Tests with DOR values exceeding 20 are generally considered potentially useful, although this threshold can vary depending on the test's sensitivity and specificity (87). In other words, as sensitivity and specificity increase and false positive and false negative rates decrease, the DOR is likely to increase. In this study, the DOR value was calculated at 49 (95% CI, 32–75), demonstrating good diagnostic accuracy.

ROC analysis is a valuable tool for assessing the performance of diagnostic tests, and the ROC curve illustrates the trade-off between sensitivity and specificity calculated based on varying cutoff values of test results measured on a continuous scale (88). The AUC, which represents the area under the ROC curve, serves as an overall summary of diagnostic accuracy, with a value of 1 indicating a perfect diagnostic test and a value below 0.5 suggesting the test is ineffective (89). In this meta-analysis, the AUC was calculated to be 0.92, indicating strong diagnostic performance for urine-based HPV detection.

Cervical cancer screening rates remain suboptimal worldwide due to various barriers. Limited healthcare access is a major factor in resource-constrained settings, but cultural beliefs, feelings of shame, and the stigma surrounding cervical examinations act as significant deterrents across diverse populations (90, 91). These sociocultural factors often discourage women from undergoing traditional Pap smears. Non-invasive alternatives, such as urine-based HPV testing, offer a promising solution by eliminating invasive procedures, reducing stigma, and enabling private sample collection (92, 93). Our meta-analysis highlights the reliable diagnostic accuracy of urine-based HPV testing (pooled sensitivity: 0.82, 95% CI: 0.78–0.86; specificity: 0.91, 95% CI: 0.87–0.94) (Table 2). This suggests that urine-based HPV testing could serve as a viable alternative to cervical swabs. By addressing cultural barriers and prioritizing patient comfort, these approaches



can encourage greater participation in screening programs and promote early diagnosis worldwide.

## Conclusion

Urine-based HPV detection may be less effective than cervical sampling due to the etiology of cervical cancer. However, the high sensitivity, specificity, and strong concordance with cervical samples suggest that urine-based real-time PCR assays offer a viable, non-invasive alternative. These findings support the potential role of urine-based HPV testing in expanding screening accessibility, particularly in settings with limited healthcare resources, ultimately contributing to improved public health outcomes.

## ACKNOWLEDGMENTS

We sincerely appreciate statistician Eun-Young Park for her valuable contributions to the statistical analysis.

## AUTHOR AFFILIATIONS

<sup>1</sup>Department of Laboratory Medicine, National Cancer Center, Goyang-si, Gyeonggi-do, South Korea

<sup>2</sup>Department of Health Promotion, National Health Insurance Service Ilsan Hospital, Goyang-si, Gyeonggi-do, South Korea

<sup>3</sup>Department of Laboratory Medicine, National Children's Medical Center, Tashkent, Uzbekistan

## AUTHOR ORCID*s*

Byeong-Min Park  <http://orcid.org/0000-0002-9590-6862>

## AUTHOR CONTRIBUTIONS

Byeong-Min Park, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing | Soohyun Kim, Data curation, Formal analysis | Jieun Choi, Writing – original draft, Writing – review and editing | Yoonkyung Song, Supervision | Seungman Park, Conceptualization, Supervision

## REFERENCES

- Ramakrishnan S, Patricia S, Mathan G. 2015. Overview of high-risk HPV's 16 and 18 infected cervical cancer: pathogenesis to prevention. *Biomed Pharm* 70:103–110. <https://doi.org/10.1016/j.biopha.2014.12.041>
- Mishra GA, Pimple SA, Shastri SS. 2011. An overview of prevention and early detection of cervical cancers. *Indian J Med Paediatr Oncol* 32:125–132. <https://doi.org/10.4103/0971-5851.92808>
- Cohen PA, Jhingran A, Oaknin A, Denny L. 2019. Cervical cancer. *Lancet* 393:169–182. [https://doi.org/10.1016/S0140-6736\(18\)32470-X](https://doi.org/10.1016/S0140-6736(18)32470-X)
- Matenge TG, Mash B. 2018. Barriers to accessing cervical cancer screening among HIV positive women in Kgatleng district, Botswana: A qualitative study. *PLoS One* 13:e0205425. <https://doi.org/10.1371/journal.pone.0205425>
- Alo H. 2023. *Cancer screening programs: quarterly data*. AIHW, Canberra.
- Aitken CA, Kaljouw S, Siebers AG, Bron M, Morssink A, van Kemenade FJ, de Kok IMCM. 2021. Investigating the decrease in participation in the Dutch cervical cancer screening programme: the role of personal and organisational characteristics. *Prev Med Rep* 22:101328. <https://doi.org/10.1016/j.pmedr.2021.101328>
- Mantula F, Toefy Y, Sewram V. 2024. Barriers to cervical cancer screening in Africa: a systematic review. *BMC Public Health* 24:525. <https://doi.org/10.1186/s12889-024-17842-1>
- Akinlotan M, Bolin JN, Helduser J, Ojinnaka C, Lichorad A, McClellan D. 2017. Cervical cancer screening barriers and risk factor knowledge among uninsured women. *J Community Health* 42:770–778. <https://doi.org/10.1007/s10900-017-0316-9>
- Brisson M, Kim JJ, Canfell K, Drolet M, Gingras G, Burger EA, Martin D, Simms KT, Bénard É, Boily M-C, Sy S, Regan C, Keane A, Caruana M, Nguyen DTN, Smith MA, Laprise J-F, Jit M, Alary M, Bray F, Fidarova E, Elsheikh F, Bloem PJN, Broutet N, Hutubessy R. 2020. Impact of HPV vaccination and cervical screening on cervical cancer elimination: a comparative modelling analysis in 78 low-income and lower-middle-income countries. *The Lancet* 395:575–590. [https://doi.org/10.1016/S0140-6736\(20\)30068-4](https://doi.org/10.1016/S0140-6736(20)30068-4)
- Harper DM, DeMars LR. 2017. HPV vaccines - a review of the first decade. *Gynecol Oncol* 146:196–204. <https://doi.org/10.1016/j.ygyno.2017.04.004>
- Ali HA, Hartner A-M, Echeverria-Londono S, Roth J, Li X, Abbas K, Portnoy A, Vynnycky E, Woodruff K, Ferguson NM, Toor J, Gaythorpe KA. 2022. Vaccine equity in low and middle income countries: a systematic review and meta-analysis. *Int J Equity Health* 21:82. <https://doi.org/10.1186/s12939-022-01678-5>
- Wirtz C, Mohamed Y, Engel D, Sidibe A, Holloway M, Bloem P, Kumar S, Brotherton J, Reis V, Morgan C. 2022. Integrating HPV vaccination programs with enhanced cervical cancer screening and treatment, a systematic review. *Vaccine (Auckl)* 40 Suppl 1:A116–A123. <https://doi.org/10.1016/j.vaccine.2021.11.013>



13. Cho HW, Ouh YT, Hong JH, Min KJ, So KA, Kim TJ, Paik ES, Lee JW, Moon JH, Lee JK. 2019. Comparison of urine, self-collected vaginal swab, and cervical swab samples for detecting human papillomavirus (HPV) with roche cobas HPV, Anyplex II HPV, and RealTime HR-S HPV assay. *J Virol Methods* 269:77–82. <https://doi.org/10.1016/j.jviromet.2019.04.012>
14. Tshomo U, Franceschi S, Tshokei T, Tobgay T, Baussano I, Tenet V, Snijders PJF, Gheit T, Tommasino M, Vorsters A, Clifford GM. 2017. Evaluation of the performance of human papillomavirus testing in paired urine and clinician-collected cervical samples among women aged over 30 years in Bhutan. *Viol J* 14:74. <https://doi.org/10.1186/s12985-017-0744-2>
15. Moberg M, Gustavsson I, Gyllensten U. 2003. Real-time PCR-based system for simultaneous quantification of human papillomavirus types associated with high risk of cervical cancer. *J Clin Microbiol* 41:3221–3228. <https://doi.org/10.1128/JCM.41.7.3221-3228.2003>
16. Cuzick J, Arbyn M, Sankaranarayanan R, Tsu V, Ronco G, Mayrand M-H, Dillner J, Meijer CJLM. 2008. Overview of human papillomavirus-based and other novel options for cervical cancer screening in developed and developing countries. *Vaccine (Auckl)* 26 Suppl 10:K29–41. <https://doi.org/10.1016/j.vaccine.2008.06.019>
17. Yu K, Park S, Chang Y, Hwang D, Kim G, Kim J, Kim S, Kim E-J, Lee D. 2019. Evaluation of commercial complementary DNA synthesis kits for detecting human papillomavirus. *Korean J Clin Lab Sci* 51:309–315. <https://doi.org/10.15324/kjcls.2019.51.3.309>
18. Watzinger F, Ebner K, Lion T. 2006. Detection and monitoring of virus infections by real-time PCR. *Mol Aspects Med* 27:254–298. <https://doi.org/10.1016/j.mam.2005.12.001>
19. Haedicke J, Iftner T. 2016. A review of the clinical performance of the aptima HPV assay. *J Clin Virol* 76 Suppl 1:S40–S48. <https://doi.org/10.1016/j.jcv.2015.10.027>
20. Bernal S, Palomares JC, Artura A, Parra M, Cabezas JL, Robles A, Martín Mazuelos E. 2014. Comparison of urine and cervical samples for detecting human papillomavirus (HPV) with the Cobas 4800 HPV test. *J Clin Virol* 61:548–552. <https://doi.org/10.1016/j.jcv.2014.10.001>
21. Stanczuk GA, Currie H, Baxter G, Foster A, Gibson L, Graham C, Cuschieri K. 2015. Cobas 4800 HPV detection in the cervical, vaginal and urine samples of women with high-grade CIN before and after treatment. *J Clin Pathol* 68:567–570. <https://doi.org/10.1136/jclinpath-2014-202851>
22. Hagiwara M, Yamagishi Y, Izumi K, Miyazaki N, Suzuki T, Kato H, Nishiyama N, Koizumi Y, Suematsu H, Mikamo H. 2016. Comparison of initial stream urine samples and cervical samples for detection of human papillomavirus. *J Infect Chemother* 22:559–562. <https://doi.org/10.1016/j.jiac.2016.05.009>
23. Khunamornpong S, Settakorn J, Sukpan K, Lekawanvijit S, Katruang N, Siriaunkgul S. 2016. Comparison of human papillomavirus detection in urine and cervical samples using high-risk HPV DNA testing in Northern Thailand. *Obstet Gynecol Int* 2016:6801491. <https://doi.org/10.1155/2016/6801491>
24. Asciutto KC, Henningsson AJ, Borgfeldt H, Darlin L, Borgfeldt C. 2017. Vaginal and urine self-sampling compared to cervical sampling for HPV-testing with the cobas 4800 HPV Test. *Anticancer Res* 37:4183–4187. <https://doi.org/10.21873/anticancer.11807>
25. Lim MC, Lee DH, Hwang SH, Hwang NR, Lee B, Shin HY, Jun JK, Yoo CW, Lee DO, Seo SS, Park SY, Joo J. 2017. Comparison of the abbot realtime high risk HPV test and the Roche cobas 4800 HPV test using urine samples. *J Virol Methods* 243:74–79. <https://doi.org/10.1016/j.jviromet.2017.01.020>
26. Vergara N, Balandá M, Hidalgo W, Martín HS, Aceituno A, Roldán F, Villalón T, Hott M, Espinoza G, Quiero A, Valenzuela MT, Ramírez E. 2018. Detection and genotyping of HPV in urine samples from Chilean women attending primary health care centers. *Med Microbiol Immunol* 207:95–103. <https://doi.org/10.1007/s00430-017-0530-1>
27. Sargent A, Fletcher S, Bray K, Kitchener HC, Crosbie EJ. 2019. Cross-sectional study of HPV testing in self-sampled urine and comparison with matched vaginal and cervical samples in women attending colposcopy for the management of abnormal cervical screening. *BMJ Open* 9:e025388. <https://doi.org/10.1136/bmjopen-2018-025388>
28. Tranberg M, Jensen JS, Bech BH, Andersen B. 2020. Urine collection in cervical cancer screening - analytical comparison of two HPV DNA assays. *BMC Infect Dis* 20:926. <https://doi.org/10.1186/s12879-020-05663-7>
29. Cho HW, Hong JH, Min KJ, Ouh YT, Seong SJ, Moon JH, Cho SH, Lee JK. 2021. Performance and diagnostic accuracy of human papillomavirus testing on self-collected urine and vaginal samples in a referral population. *Cancer Res Treat* 53:829–836. <https://doi.org/10.4143/crt.2020.1165>
30. Van Keer S, Peeters E, Vanden Broeck D, De Sutter P, Donders G, Doyen J, Tjalma WAA, Weyers S, Vorsters A, Arbyn M. 2021. Clinical and analytical evaluation of the realtime high risk HPV assay in colli-pee collected first-void urine using the VALHUDES protocol. *Gynecol Oncol* 162:575–583. <https://doi.org/10.1016/j.ygyno.2021.06.010>
31. Kim DH, Jin H, Lee KE. 2022. Analysis of HR-HPV infection concordance rates in cervical and urine specimens; proposal of additional cervical screening process for women who refuse invasive cervical sampling. *JPM* 12:1949. <https://doi.org/10.3390/jpm12121949>
32. Punyashthira A, Horthongkham N, Jaishuen A, Jareemit N, Achariyapota V. 2022. A comparison of high-risk human papillomavirus DNA detection between urine and cervical sample testing in women with abnormal pap smears. *J Obstet Gynaecol Res* 48:448–455. <https://doi.org/10.1111/jog.15095>
33. Martinelli M, Giubbi C, Di Meo ML, Perdoni F, Musumeci R, Leone BE, Fruscio R, Landoni F, Cocuzza CE. 2023. Accuracy of human papillomavirus (HPV) testing on urine and vaginal self-samples compared to clinician-collected cervical sample in women referred to colposcopy. *Viruses* 15:1889. <https://doi.org/10.3390/v15091889>
34. Nilyanimit P, Chaithongwongwatthana S, Oranratanaphan S, Poudyal N, Excler JL, Lynch J, Vongpunsawad S, Poovorawan Y. 2024. Comparable detection of HPV using real-time PCR in paired cervical samples and concentrated first-stream urine collected with Colli-Pee device. *Diagn Microbiol Infect Dis* 108:116160. <https://doi.org/10.1016/j.diagmicrobio.2023.116160>
35. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MMG, Sterne JAC, Bossuyt PMM, QUADAS-2 Group. 2011. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 155:529–536. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>
36. Yang B, Mallett S, Takwoingi Y, Davenport CF, Hyde CJ, Whiting PF, Deeks JJ, Leeflang MMG, Bossuyt PMM, Brazzell MG, Dinnes J, Gurusamy KS, Jones HE, Lange S, Langendam MW, Macaskill P, McInnes MDF, Reitsma JB, Rutjes AWS, Sinclair A, de Vet HCW, Virgili G, Wade R, Westwood ME, QUADAS-C Group. 2021. QUADAS-C: a tool for assessing risk of bias in comparative diagnostic accuracy studies. *Ann Intern Med* 174:1592–1599. <https://doi.org/10.7326/M21-2234>
37. Campbell JM, Klugar M, Ding S, Carmody DP, Hakonsen SJ, Jadotte YT, White S, Munn Z. 2015. Diagnostic test accuracy: methods for systematic review and meta-analysis. *JBIM Evidence Implementation* 13:154–162. <https://doi.org/10.1097/XEB.0000000000000061>
38. Shim SR, Kim S-J, Lee J. 2019. Diagnostic test accuracy: application and practice using R software. *Epidemiol Health* 41:e2019007. <https://doi.org/10.4178/epih.e2019007>
39. Shim S, Shin I, Bae J. 2015. Brief Communication: Meta-analysis of Diagnostic Tests Accuracy using STATA Software. *J Health Info Stat* 40:190–200.
40. Chu H, Cole SR. 2006. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *J Clin Epidemiol* 59:1331–1332. <https://doi.org/10.1016/j.jclinepi.2006.06.011>
41. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JAC. 2007. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 8:239–251. <https://doi.org/10.1093/biostatistics/kxl004>
42. Higgins JPT, Thompson SG. 2002. Quantifying heterogeneity in a meta-analysis. *Stat Med* 21:1539–1558. <https://doi.org/10.1002/sim.1186>
43. Deeks JJ, Macaskill P, Irwig L. 2005. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 58:882–893. <https://doi.org/10.1016/j.jclinepi.2005.01.016>
44. Thompson SG, Higgins JPT. 2002. How should meta-regression analyses be undertaken and interpreted? *Stat Med* 21:1559–1573. <https://doi.org/10.1002/sim.1187>
45. Song F, Khan KS, Dinnes J, Sutton AJ. 2002. Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int J Epidemiol* 31:88–95. <https://doi.org/10.1093/ije/31.1.88>
46. Arbyn M, Weiderpass E, Bruni L, de Sanjosé S, Saraiya M, Ferlay J, Bray F. 2020. Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. *Lancet Glob Health* 8:e191–e203. [https://doi.org/10.1016/S2214-109X\(19\)30482-6](https://doi.org/10.1016/S2214-109X(19)30482-6)
47. Hull R, Mbele M, Makhafola T, Hicks C, Wang S-M, Reis RM, Mehrotra R, Mkhize-Kwitshana Z, Kibiki G, Bates DO, Dlamini Z. 2020. Cervical cancer

- in low and middle-income countries. *Oncol Lett* 20:2058–2074. <https://doi.org/10.3892/ol.2020.11754>
48. Al Sad S, Pandit R, Alhashim N, Abdel-Rasoul M. 2023. Primary care providers' approaches to cervical cancer screening in muslim females. *Prev Med Rep* 32:102126. <https://doi.org/10.1016/j.pmedr.2023.102126>
  49. Kim DH, Jin H, Lee KE. 2021. Analysis of HPV prevalence in pairs of cervical and urine samples from the same woman. *BSL* 27:28–34. <https://doi.org/10.15616/BSL.2021.27.1.28>
  50. Munoz M, Camargo M, Soto-De Leon SC, Sanchez R, Pineda-Peña AC, Perez-Prados A, Patarroyo ME, Patarroyo MA. 2013. Classical molecular tests using urine samples as a potential screening tool for human papillomavirus detection in human immunodeficiency virus-infected women. *J Clin Microbiol* 51:3688–3693. <https://doi.org/10.1128/JCM.01302-13>
  51. Pathak N, Dodds J, Zamora J, Khan K. 2014. Accuracy of urinary human papillomavirus testing for presence of cervical HPV: systematic review and meta-analysis. *BMJ* 349:g5264. <https://doi.org/10.1136/bmj.g5264>
  52. Vorsters A, Micalessi I, Bilcke J, Ieven M, Bogers J, Damme P. 2012. Detection of human papillomavirus DNA in urine. a review of the literature. *Eur J Clin Microbiol Infect Dis* 31:627–640. <https://doi.org/10.1007/s10096-011-1358-z>
  53. Huh WK, Ault KA, Chelmon D, Davey DD, Goulart RA, Garcia FA, Kinney WK, Massad LS, Mayeaux EJ, Saslow D, Schiffman M, Wentzensen N, Lawson HW, Einstein MH. 2015. Use of primary high-risk human papillomavirus testing for cervical cancer screening: interim clinical guidance. *J Low Genit Tract Dis* 19:91–96. <https://doi.org/10.1097/LGT.000000000000103>
  54. Arbyn M, Ronco G, Anttila A, Meijer CJLM, Poljak M, Ogilvie G, Koliopoulos G, Naucier P, Sankaranarayanan R, Peto J. 2012. Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. *Vaccine (Auckl)* 30 Suppl 5:F88–99. <https://doi.org/10.1016/j.vaccine.2012.06.095>
  55. Power M, Fell G, Wright M. 2013. Principles for high-quality, high-value testing. *Evid Based Med* 18:5–10. <https://doi.org/10.1136/eb-2012-100645>
  56. Piao J, Yoo C, Kim S, Whang Y-W, Choi CU, Shin S. 2021. Performance comparison of the PFA-200 and Ansys-200: assessment of bleeding risk screening in cardiology patients. *Clin Hemorheol Microcirc* 79:445–454. <https://doi.org/10.3233/CH-211185>
  57. Kamal SA, Sultan F, Jamil SS. 2013. Sensitivity and specificity of screening tests. *The First*
  58. Gisev N, Bell JS, Chen TF. 2013. Interrater agreement and interrater reliability: key concepts, approaches, and applications. *Res Social Adm Pharm* 9:330–338. <https://doi.org/10.1016/j.sapharm.2012.04.004>
  59. McHugh ML. 2012. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 22:276–282.
  60. Park CU, Kim HJ. 2015. Measurement of inter-rater reliability in systematic review. *Hanyang Med Rev* 35:44. <https://doi.org/10.7599/hmr.2015.35.1.44>
  61. Gwet K. 2002. Inter-rater reliability: dependency on trait prevalence and marginal homogeneity. In *Statistical methods for inter-rater reliability assessment series 2:9*
  62. Sim J, Wright CC. 2005. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 85:257–268.
  63. Agarwal R. 2013. Measurement of errors in clinical laboratories. *Indian J Clin Biochem* 28:227–234. <https://doi.org/10.1007/s12291-013-0314-3>
  64. Hoehler FK. 2000. Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity. *J Clin Epidemiol* 53:499–503. [https://doi.org/10.1016/s0895-4356\(99\)00174-2](https://doi.org/10.1016/s0895-4356(99)00174-2)
  65. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. 2003. Measuring inconsistency in meta-analyses. *BMJ* 327:557–560. <https://doi.org/10.1136/bmj.327.7414.557>
  66. Cuzick J, Cadman L, Ahmad AS, Ho L, Terry G, Kleeman M, Lyons D, Austin J, Stoler MH, Vibat CRT, Dockter J, Robbins D, Billings PR, Erlander MG. 2017. Performance and diagnostic accuracy of a urine-based human papillomavirus assay in a referral population. *Cancer Epidemiol Biomarkers Prev* 26:1053–1059. <https://doi.org/10.1158/1055-9965.EPI-16-0960>
  67. Ebrahimi N, Yousefi Z, Khosravi G, Malayeri FE, Golabi M, Askarzadeh M, Shams MH, Ghezelbash B, Eskandari N. 2023. Human papillomavirus vaccination in low- and middle-income countries: progression, barriers, and future prospective. *Front Immunol* 14:1150238. <https://doi.org/10.3389/fimmu.2023.1150238>
  68. Cheng L, Wang R, Yan J. 2024. A review of urinary HPV testing for cervical cancer management and HPV vaccine surveillance: rationale, strategies, and limitations. *Eur J Clin Microbiol Infect Dis* 43:2247–2258. <https://doi.org/10.1007/s10096-024-04963-z>
  69. Spineli LM, Pandis N. 2020. Exploring heterogeneity in meta-analysis: meta-regression analysis. *Am J Orthod Dentofacial Orthop* 158:623–625. <https://doi.org/10.1016/j.jajodo.2020.07.002>
  70. Thompson SG. 1994. Why sources of heterogeneity in meta-analysis should be investigated. *BMJ* 309:1351–1355. <https://doi.org/10.1136/bmj.309.6965.1351>
  71. Bae J-M. 2014. An overview of systematic reviews of diagnostic tests accuracy. *Epidemiol Health* 36:e2014016. <https://doi.org/10.4178/epih/e2014016>
  72. Ramsubeik K, Motilal S, Sanchez-Ramos L, Ramrattan LA, Kaeley GS, Singh JA. 2020. Diagnostic accuracy of salivary gland ultrasound in sjögren's syndrome: a systematic review and meta-analysis. *Ther Adv Musculoskelet Dis* 12:1759720X20973560. <https://doi.org/10.1177/1759720X20973560>
  73. Téblick L, Van Keer S, De Smet A, Van Damme P, Laeremans M, Rios Cortes A, Beyers K, Vankerckhoven V, Matheeußen V, Mandersloot R, Floore A, Meijer CJLM, Steenbergen RDM, Vorsters A. 2021. Impact of collection volume and DNA extraction method on the detection of biomarkers and HPV DNA in first-void urine. *Molecules* 26:1989. <https://doi.org/10.3390/molecules26071989>
  74. Vorsters A, Van den Bergh J, Micalessi I, Biesmans S, Bogers J, Hens A, De Coster I, Ieven M, Van Damme P. 2014. Optimization of HPV DNA detection in urine by improving collection, storage, and extraction. *Eur J Clin Microbiol Infect Dis* 33:2005–2014. <https://doi.org/10.1007/s10096-014-2147-2>
  75. Van Keer S, Pattyn J, Tjalma WAA, Van Ostade X, Ieven M, Van Damme P, Vorsters A. 2017. First-void urine: A potential biomarker source for triage of high-risk human papillomavirus infected women. *European J Obst Gynec Reprod Bio* 216:1–11. <https://doi.org/10.1016/j.ejogrb.2017.06.036>
  76. Arbyn M, Verdoodt F, Snijders PJF, Verhoef VMJ, Suonio E, Dillner L, Minozzi S, Bellisario C, Banzi R, Zhao F-H, Hillemanns P, Anttila A. 2014. Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: a meta-analysis. *Lancet Oncol* 15:172–183. [https://doi.org/10.1016/S1470-2045\(13\)70570-9](https://doi.org/10.1016/S1470-2045(13)70570-9)
  77. Leeman A, Del Pino M, Molijn A, Rodriguez A, Torné A, de Koning M, Ordi J, van Kemenade F, Jenkins D, Quint W. 2017. HPV testing in first-void urine provides sensitivity for CIN2+ detection comparable with a smear taken by a clinician or a brush-based self-sample: cross-sectional data from a triage population. *BJOG* 124:1356–1363. <https://doi.org/10.1111/1471-0528.14682>
  78. Kavanagh K, Pollock KGJ, Potts A, Love J, Cuschieri K, Cubie H, Robertson C, Donaghy M. 2014. Introduction and sustained high coverage of the HPV bivalent vaccine leads to a reduction in prevalence of HPV 16/18 and closely related HPV types. *Br J Cancer* 110:2804–2811. <https://doi.org/10.1038/bjc.2014.198>
  79. Perkins RB, Wentzensen N, Guido RS, Schiffman M. 2023. Cervical cancer screening: a review. *JAMA* 330:547–558. <https://doi.org/10.1001/jama.2023.13174>
  80. Bober P, Firment P, Sabo J. 2021. Diagnostic test accuracy of first-void urine human papillomaviruses for presence cervical HPV in women: systematic review and meta-analysis. *Int J Environ Res Public Health* 18:13314. <https://doi.org/10.3390/ijerph182413314>
  81. Vorsters A, Van Damme P, Clifford G. 2014. Urine testing for HPV: rationale for using first void. *BMJ* 349:g6252. <https://doi.org/10.1136/bmj.g6252>
  82. von Hippel PT. 2015. The heterogeneity statistic I2 can be biased in small meta-analyses. *BMC Med Res Methodol* 15:1–8. <https://doi.org/10.1186/s12874-015-0024-z>
  83. Deeks JJ. 2001. Systematic reviews of evaluations of diagnostic and screening tests. *BMJ* 323:157–162.
  84. Deeks JJ, Altman DG. 2004. Diagnostic tests 4: likelihood ratios. *BMJ* 329:168–169. <https://doi.org/10.1136/bmj.329.7458.168>
  85. Jaeschke R, Guyatt GH, Sackett DL, Guyatt G, Bass E, Brill-Edwards P, Browman G, Cook D, Farkouh M, Gerstein H. 1994. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The evidence-based medicine working group. *JAMA* 271:703–707. <https://doi.org/10.1001/jama.271.9.703>

86. Zhu L, Liu Y, Chen G. 2014. Diagnostic value of mesothelin in pancreatic cancer: a meta-analysis. *Int J Clin Exp Med* 7:4000–4007.
87. Fischer JE, Bachmann LM, Jaeschke R. 2003. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Med* 29:1043–1051. <https://doi.org/10.1007/s00134-003-1761-8>
88. Hajian-Tilaki K. 2013. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian J Intern Med* 4:627–635.
89. Eusebi P. 2013. Diagnostic accuracy measures. *Cerebrovasc Dis* 36:267–272. <https://doi.org/10.1159/000353863>
90. Farajimakin O. 2024. Barriers to cervical cancer screening: a systematic review. *Cureus* 16:e65555. <https://doi.org/10.7759/cureus.65555>
91. Marques P, Gama A, Geraldles M, Silva J, Vaccari N, Dias S. 2021. Sociocultural factors of cervical cancer screening uptake among migrant women: a mixed methods study. *Eur J Public Health* 31:611. <https://doi.org/10.1093/eurpub/ckab164.611>
92. Petersen Z, Jaka A, Ginindza TG, Maseko G, Takatshana S, Ndlovu P, Zondi N, Zungu N, Varghese C, Hunting G, Parham G, Simelela P, Moyo S. 2022. Barriers to uptake of cervical cancer screening services in low-and-middle-income countries: a systematic review. *BMC Womens Health* 22:486. <https://doi.org/10.1186/s12905-022-02043-y>
93. Bøje RB, Bardou M, Mensah K, Rico Berrocal R, Giorgi Rossi P, Bonvicini L, Auzzi N, Taut D, Jiboc N, Tisler A, Reintamm K, Uusküla A, Teixeira M, Firmino-Machado J, Amorim M, Baia I, Lunet N, Michaylova R, Panayotova Y, Kotzeva T, Andersen B, Kirkegaard P, CBIG-SCREEN Consortium. 2024. What are the barriers towards cervical cancer screening for vulnerable women? A qualitative comparative analysis of stakeholder perspectives in seven European countries. *BMJ Open* 14:e079921. <https://doi.org/10.1136/bmjopen-2023-079921>