Accuracy of human papillomavirus tests on self-collected urine versus clinician-collected samples for the detection of cervical precancer: a systematic review and meta-analysis

Hyun-Woong Cho,1 Sung Ryul Shim,2 Jae Kwan Lee,1 Jin Hwa Hong1

1Department of Obstetrics and Gynecology, Korea University Guro Hospital, Korea University College of Medicine, Seoul, Korea
2Department of Preventive Medicine, Korea University College of Medicine, Seoul, Korea

ABSTRACT

Objective: The human papillomavirus (HPV) test is an effective screening tool to prevent cervical cancer. Urinary sampling for HPV detection improves the accessibility and participation of screening services and reduces the cost and burden on physicians. The clinical accuracy of urinary HPV test has yet to be determined via meta-analysis. This study assessed the clinical accuracy of these tests to detect cervical intraepithelial neoplasia (CIN) 2 or worse.

Methods: Relevant studies were identified using the PubMed, Embase, and Cochrane databases. Research eligibility was based on the clinical accuracy of HPV test on clinician-collected samples as a comparator test, and urine as an index test. The reference standard was the presence of CIN2 or worse. The pooled absolute, relative sensitivity, and specificity of the urinary HPV test versus clinician-collected samples were assessed using a bivariate model.

Results: The pooled sensitivity of urinary HPV test was significantly lower than that of clinician-collected samples (ratio=0.84, 95% confidence interval [CI]=0.78–0.91). However, some polymerase chain reaction (PCR)-based HPV test such as GP5+/6+ (relative sensitivity=0.98, 95% CI=0.91-1.05), SPF10 (relative sensitivity=0.98, 95% CI=0.88–1.08) and non GP5+/6+ PCR (relative sensitivity=1.00, 95% CI=0.88–1.14) showed similar sensitivity in both the urine and clinician-collected samples.

Conclusion: Our findings indicate that HPV test with some PCR-based assay on urine versus clinician-collected samples demonstrate similar clinical accuracy to detect CIN2 or worse. It suggests that urinary HPV test may present itself as a decent alternative screening tool for the detection of cervical pre-cancer.

Trial Registration: PROSPERO identifier: CRD42021227901

Keywords: Human Papillomavirus DNA Test; Urine; Cervical Intraepithelial Neoplasia; Uterine Cervical Neoplasms

https://ejgo.org
INTRODUCTION

The human papillomavirus (HPV) test is an effective screening tool for cervical cancer used as a primary screening tool in several countries. Previous randomized controlled trials demonstrate that screening with high-risk HPV (hrHPV) is more sensitive in detecting precancerous lesions as opposed to cytological screening methods although its specificity is relatively low [1]. In 2018, the United States Preventive Services Task Force recommended screening for hrHPV alone, once every 5 years, as a preferred strategy [2]. European guideline also recommends primary testing for HPV to improve cervical cancer prevention and control.

Despite these powerful screening tests, more than 570,000 women are diagnosed with cervical cancer, and 311,000 women have succumbed to this disease in 2018 worldwide [3]. Importantly, inequalities in access to screening tools result in ethnic, racial, and social disparities in the incidence and mortality of cervical cancer [4,5]. Low participation is also a barrier to cervical cancer screening, even in countries with well-established screening programs [6].

There is a major interest in showing that self-collected samples including urine for cervical screening is just as effective as clinical sampling both in developing countries due to its poor accessibility to screening units and developed countries due to low screening uptake rates, especially in light of the coronavirus disease 2019 pandemic. Self-sampling has been suggested to increase accessibility and uptake of screening services, reduce the burden on health workers, and save the costs of screening. In particular, urine sampling for HPV detection offers a more accessible and acceptable method [7-9]. Vaginal fluid containing exfoliated HPV infected cells are washed away with urine which allows identification of HPV DNA in urine [10]. In 2014, meta-analysis of 14 studies has shown that the urinary HPV test had a pooled sensitivity of 77% and specificity of 88% compared with clinician-collected cervical HPV test (cervical HPV test) [11]. However, this meta-analysis has not been updated for 6 years, and clinical accuracy was not assessed to detect cervical intraepithelial neoplasia (CIN) 2 or worse. Although several recent studies have demonstrated that the urinary HPV test shares similar accuracy in detecting cervical precancer compared with cervical HPV test, clinical outcomes for urinary HPV vary widely between studies [12-14]. As such, there is a need to update the meta-analysis in order to comprehensively evaluate the clinical performance of urinary HPV test for secondary prevention of cervical cancer.

This study evaluated relative accuracy of the urinary HPV test versus the cervical HPV test that uses clinician-collected cervical samples to detect CIN2 or worse. We also assessed the absolute clinical accuracy of the HPV test on urine. We included a subgroup meta-analysis to evaluate the relative accuracy of urinary HPV test according to the income status of country, hrHPV assay, sampling method, sampling device, and storage medium.

https://ejgo.org

https://doi.org/10.3802/jgo.2022.33.e4

2/16
MATERIALS AND METHODS

This study was registered with PROSPERO (CRD42021227901) and performed according to the standard Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement.

1. Search strategy and eligibility criteria
We conducted an electronic search for English-language literature in PubMed, Embase and the Cochrane databases from August 1, 1968 to July 15, 2020. The detailed study scheme and search strategies are described in the Fig. 1 and Table S1. Studies were deemed eligible for inclusion if they satisfied 3 key criteria: 1) the study assessed the clinical accuracy of HPV test on clinician-collected cervical and urine samples as an index test in women; 2) the reference was the presence of CIN2 or worse via colposcopic biopsy or conization; and 3) the study provided numbers of true positive, false positive, false negative, and true negative results, or this data were derived from the published results of studies.

2. Study selection, data extraction, and quality assessment
All titles and abstracts for relevant studies were reviewed. Two authors (H.W.C and S.R.S) had independently reviewed the full text for study selection, and data extraction. Information

Fig. 1. Flow chart of study identification and selection.
CIN, cervical intraepithelial neoplasia; HPV, human papillomavirus.
regarding the study participants, the setting, hrHPV assays, sampling devices, sampling, storage methods, and clinical outcomes (absolute numbers of true positive, false positive, false negative, and true negative results) were collected in a comprehensive table (Table 1). We accepted the cutoff proposed by the manufacturer to define HPV positivity by tests. Two authors independently evaluated the quality of the method using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool. Disagreements between the authors were settled by consensus involving third author if necessary.

3. Data analysis

Absolute sensitivity and specificity of HPV test on the urine and clinician-collected samples were calculated using a bivariate model using midas, a STATA module for the meta-analysis of diagnostic accuracy studies. The ratio of sensitivity and specificity of urinary HPV test compared with cervical HPV test using clinician-collected samples was estimated, and a bivariate random effect model with metan functions in STATA, was used to pool the meta-analysis. Sub-analysis of relative accuracy and meta-analysis was conducted to evaluate variations in the tests accuracy based on the study setting (primary screening or follow-up) and covariates including hrHPV assays, urine collection time, urinary stream, urine collecting device, and storage medium. We conducted meta-regression and sub-analysis using the covariate to identify significant factors that had an impact on the accuracy of urinary HPV test. The accuracy of diagnostic tests, including sensitivity, specificity, and diagnostic odds ratios (DORs), were described as point estimates with 95% confidence intervals (CIs). A DOR was defined as the ratio of the odds of the test being positive if the subject has a disease relative to the odds of the test being positive if the subject does not have the disease. Statistical heterogeneity between studies was evaluated using I² and the Cochrane Q test based on random-effects analysis. Publication bias was assessed using the Deeks’ funnel plot asymmetry test for pooled absolute accuracy measurements. Statistical tests were -sided, and statistical significance was defined as p<0.05. We used Meta-DiSc (version 1·4) for meta-regression analysis, whilst all other statistical analysis was undertaken using STATA (version 16·1).

RESULTS

The literature search process is described in Fig. 1; a total of 297 studies from PubMed, 506 studies from Embase, and 48 studies from Cochrane were identified. Following review, 21 studies were deemed eligible for inclusion. All included studies were observational studies. As some studies contained up to 2–4 combinations of HPV assays, urine collection method, device and preservative use [13,14,20,26,30,31], each of these combinations were analyzed as independent studies. Thus, data from 30 combinations for urine and 25 combinations for clinician-collected samples were included in the meta-analysis, totaling to 11,159 and 10,774 women for each collection method, respectively. For the urinary HPV test studies, generally healthy women were enrolled in 4 combinations from 3 studies [15,29,31], whilst women who were referred to colposcopic clinics were enrolled in 26 combinations from 17 studies [7,12-14,16-26,28,30]. Women undergoing conization for CIN were enrolled in only one study [27].

The summary of the study design, population, demographics of participants, and details of the hrHPV assay, sample collection, sampling device, and storage methods are available in the Table 1, Tables S2 and S3.
### Table 1. Characteristics of studies included in the meta-analysis

| Study                                | Study design | Population/setting | Study size (total number for analysis) | Age | HPV assays | Urine collection | Sampling device | Preservative for urine transport | Gold standard | Clinical outcome (number) |
|--------------------------------------|--------------|---------------------|----------------------------------------|-----|------------|------------------|-----------------|-------------------------------|---------------|---------------------------|
| Alameda et al. [15]                  | Cross-sectional | Primary screening | 50 mean, 36; range, 28–55 | Unknown | Urine: (PCR) using consensus primers (MY09/MY11) | Unknown | preservCyt | Colposcopic cervical biopsy | All participants | CIN2+ (11) |
| Asciutto et al. [16]                 | Cross-sectional | Follow-up (colposcopic clinic) | 218 Mean, 35.2; range, Clinician: Cobas | First stream | Plastic cup | Cobas PCR media | Colposcopic cervical biopsy | All participants | CIN2+ (112) |
| Asciutto et al. [17]                 | Cross-sectional | Follow-up (colposcopic clinic) | 209 Mean, 33.7; range, Clinician: Aptima | Initial stream urine | Plastic container | No (aptima transport media) | Colposcopic cervical biopsy | All participants | CIN2+ (67) |
| Bernal et al. [18]                   | Cross-sectional | Follow-up (colposcopic clinic) | 120 Median, 35.5; range, Clinician: Cobas | First stream | Sterile container | Unknown | Colposcopic cervical biopsy or conization | All participants | CIN2+ (20) |
| Buchegger et al. [19]                | Cross-sectional | Follow-up (colposcopic clinic) | 190 Median, 28; IQR, 15–75 | First stream | No | Sterile container (10% crystal violet) | Colposcopic cervical biopsy or conization | All participants | CIN2+ (61) |
| Cuzick et al. [20]                   | Cross-sectional | Follow-up (colposcopic clinic) | 501 Median, 30; IQR, 27–34 | Urine: (PCR) using consensus primers (MY09/MY11) | Unknown | Trovagene Clinician: Trovagene | Unknown | Preservative solution | Colposcopic cervical biopsy | All participants | CIN2+ (145) |
| Leeman et al. [13]                   | Cross-sectional | Follow-up (colposcopic clinic) | 91 Range, 18–60 | Urine: SPF10-DEIA-LIP25 assay, GPS5+/6+/EIA-LMNX | Morning first | Random first void | 4 mL of a buffered lithium dodecyl sulfate solution containing RNA preservative | Colposcopic cervical biopsy or conization | All participants | CIN2+ (66) |
| Arias et al. [20]                    | Cross-sectional | Follow-up (colposcopic clinic) | 433 Mean, 36; range, 21–74 | Urine: Aptima Clinician: Aptima | Unknown | No or aptima transport media (ATS) | Colposcopic cervical biopsy or conization | All participants | CIN2+ (19) |
| Padhy et al. [21]                    | Cross-sectional | Follow-up (colposcopic clinic) | 189 Median, 41 | Urine: Linear array Clinician: Linearray | Initial stream urine | Sterile container | No (aptima transport media) | Colposcopic cervical biopsy | All participants | CIN2+ (33) |
| Piyathilake et al. [22]              | Cross-sectional | Follow-up (colposcopic clinic) | 502/468 Mean, 38; range, 25–28 | Urine: Linear array Clinician: Linearray | Unknown | Unknown | Colposcopic cervical biopsy | All participants | CIN2+ (72) |
| Rohner et al. [23]                   | Cross-sectional | Follow-up (colposcopic clinic) | 434 Median, 36; IQR, 21–45 | Urine: Onclarity Clinician: Onclarity | Unknown | 0.2 mL of a proprietary preservative | Colposcopic cervical biopsy | All participants | CIN2+ (83) |
| Sahasrabuddhe et al. [24]           | Cross-sectional | Follow-up (colposcopic clinic) | 72/71 Median, 28; range, 24–34 | Urine: Linear array Clinician: Linearray | Random first void | Collection 10 mL EDTA cup | Colposcopic cervical biopsy | All participants | CIN2+ (26) |
| Sahasrabuddhe et al. [25]           | Cross-sectional | Follow-up (colposcopic clinic) | 72/71 Median, 28; range, 24–34 | Urine: Trovagene Clinician: Linearray | Random first void | Collection 10 mL EDTA cup | Colposcopic cervical biopsy | All participants | CIN2+ (26) |
| Sargent et al. [26]                  | Cross-sectional | Follow-up (colposcopic clinic) | 79 40% of patients between 25–29 | Urine: Cobas Clinician: Cobas | First stream | Sterile dry Unknown | Colposcopic cervical biopsy or conization | All participants | CIN2+ (18) |
| Sellors et al. [7]                   | Cross-sectional | Follow-up (colposcopic clinic) | 200 Mean, 31.5; SD, 9.4 | Urine: Hybrid Capture II Clinician: Hybrid Capture II | Plastic bottle | Unknown | Colposcopic cervical biopsy | All participants | CIN2+ (58) |

(continued to the next page)
1. Quality assessment

Fig. S1 shows the QUADAS-2 assessment in the meta-analysis. The overall quality of studies appeared to be adequate; all studies used consecutive enrollments of participants. The risk of bias for patient selection was considered low in 3/30 (10%) patients. The risk of bias for the index/comparator test and reference test was low across all studies (100%). Risk of flow and timing was considered low in 24 combinations as index, comparator and reference test were done on the same day (80%). Time interval between test was unknown in 6 (20%). Most studies were evaluated as having low applicability for patient selection (29/30, 96·7%), index/comparator test (28/30, 93·3%), and reference tests (25/30, 83·3%).

2. Relative clinical accuracy of HPV test on urine versus clinician-collected samples

The pooled relative sensitivity and specificity of HPV test on urine versus clinician-collected samples is demonstrated in Fig. 2. The overall relative sensitivity and specificity of HPV test on urine versus clinician-collected samples was 0.84 (95% CI=0.78–0.91) and 1.06 (1.03–1.10), respectively. Although the variation based on the clinical setting was not substantial, significant heterogeneity was observed in the sensitivity between studies (p<0.001). Table 2 and Fig. S2 shows that relative sensitivity of urinary versus cervical HPV are widely varied by type of hrHPV assays used. Some PCR-based HPV test such as GP5+/6+ and SPF10 showed

---

Table 1. (Continued) Characteristics of studies included in the meta-analysis

| Study                  | Study design | Population/setting | Study size (total/number for analysis) | Age | HPV assays | Urine collection | Sampling device | Preservative for urine transport | Gold standard | Clinical outcome (number) |
|------------------------|--------------|---------------------|----------------------------------------|-----|------------|-----------------|-----------------|---------------------------------|---------------|---------------------------|
| Senkomago et al. [14]  | Cross-sectional | Follow-up (colposcopic clinic) | 37/37 | Median, 42; range, 30–63 | Urine: Trovagene Clinician: Aptima | Morning first void Initial stream Mid-stream | Collection cup | 8 mL EDTA | Colposcopic cervical biopsy | All participants | CIN2+ (11) |
| Sorbi et al. [27]      | Cross-sectional | Follow-up (conization) | 134 | Unknown | Urine: Linear array Clinician: Linear array | First void Unknown Unknown | Unknown | Colposcopic cervical biopsy or conization | All participants | CIN2+ (103) |
| Stanczuk et al. [28]   | Cross-sectional | Follow-up (colposcopic clinic) | 100/100 | Median, 27.5; range, 21–60 | Urine: Cobas Clinician: Cobas | Unknown | Universal container | Colposcopic cervical biopsy | All participants | CIN2+ (65) |
| Stanczuk et al. [29]   | Cross-sectional | Primary screening | 5,318/5,003 | Median, 41.3; range, 18–76 | Urine: Cobas Clinician: Cobas | Random void | Universal container | 6 mL was mixed with 3 mL of Roche PCR media (Roche Molecular Systems) | Colposcopic cervical biopsy | High-grade abnormalities | 2 low-grade or 3 borderline smears | 3 consecutive unsatisfactory or a subsequent abnormal smear | cytol−/hrHPV+ women if HPV 16 and/or 18 positive | All participants | CIN2+ (130) | CIN3+ (68) |
| Tshomo et al. [30]     | Cross-sectional | Follow-up (colposcopic clinic) | 89 | Median, 39; 5–95%, 30–54 | Urine: E7 MPG, GP5+/6+ Clinician: E7 MPG, GP5+/6+ | First void | Colli-pee | 7 mL urine preservative medium | Colposcopic cervical biopsy or conization | All participants | CIN2+ (5) | CIN3+ (5) |
| Xu et al. [31]         | Cross-sectional | Primary screening | 1,952/989,983 | 48.2±7.3 yr | Urine: Cobas, CareHPV Clinician: Cobas | Random void | Sample bottle | 10 mL EDTA | Colposcopic cervical biopsy Triage: HPV (+) and or abnormal cytology | All participants | CIN2+ (20) |

EDTA, ethylenediaminetetraacetic acid; HPV, human papillomavirus; IQR, interquartile range; PCR, polymerase chain reaction; SD, standard deviation.
similar levels of sensitivity and specificity in both the urine and clinician-collected samples. In contrast, Aptima and HC2 tests were less sensitive, although more specific in terms of urine compared to clinician-collected samples. Although there was no significant difference, the urinary HPV test was less sensitive than the cervical HPV test when Trovagene, Abbott, Linear array, Cobas, and CareHPV were used.

Table 3 and Table S4 shows the relative accuracy of hrHPV assays on urine compared to that of the clinician-collected samples; accuracy was dependent on urine collection time, urinary stream, sampling device and the type of preservative used for storage. When morning first urine (ratio of sensitivity=0.97, 95% CI=0.88-1.07/ratio of specificity 1.02, 95% CI=0.77-1.34) or Colli-pee (ratio of sensitivity=0.98, 95% CI=0.91-1.05/ratio of specificity 1.03, 95% CI=0.90-1.19) was used, there was no significant difference of sensitivity and specificity between urine and clinician-collected samples. In studies of middle- and low-income
The sensitivity (ratio=0.93 [95% CI=0.84–1.02]) and specificity (ratio=1.01 [95% CI=0.99–1.04]) of urinary HPV test were comparable with cervical HPV.

Table 4 shows the results of multivariate meta-regression to evaluate the moderator effect of HPV assays and covariates. It suggests that the sensitivity of urinary HPV versus cervical HPV differ significantly based on the hrHPV assay type. Aside from these factors, there were no other covariates that significantly impacted the clinical accuracy of the HPV test. In addition, preservative use of urine storage was significant factor to negatively affect the specificity of urinary HPV (p=0.001).
Table 2. Relative sensitivity and specificity of urinary versus cervical HPV test, using HPV assays, for the detection of cervical intraepithelial neoplasia 2 or worse

| Type of Test          | No. of combinations (study) | Relative sensitivity (95% CI) | Relative specificity (95% CI) |
|-----------------------|-----------------------------|------------------------------|-------------------------------|
| PCR-based assays      |                             |                              |                               |
| PCR GPS+/6+           | 4 (3)                       | 0.98 (0.91–1.05)             | 1.03 (0.90–1.18)              |
| PCR-SPF10*            | 2 (1)                       | 0.98 (0.88–1.08)             | 0.92 (0.65–1.29)              |
| Other non GPS+/6+ PCR | 3 (3)                       | 1.00 (0.88–1.14)             | 1.00 (0.88–1.14)              |
| Trovagene‡            | 5 (3)                       | 0.94 (0.89–1.00)             | 0.94 (0.77–1.15)              |
| Linear Array†          | 3 (3)                       | 0.84 (0.69–1.01)             | 1.27 (1.11–1.45)              |
| Cobas§                | 6 (6)                       | 0.80 (0.64–1.10)             | 1.03 (1.01–1.04)              |
| Others                |                             |                              |                               |
| CareHPV∥              | 1 (1)                       | 0.87 (0.59–1.27)             | 1.02 (0.98–1.06)              |
| Aptima¶               | 4 (3)                       | 0.52 (0.41–0.67)             | 1.39 (1.07–1.81)              |
| Hybrid Capture II**   | 1 (1)                       | 0.46 (0.34–0.61)             | 1.34 (1.16–1.62)              |
| Abbott††              | 6 (6)                       | 0.94 (0.72–1.22)             | 1.33 (0.86–2.07)              |

CI, confidence interval; HPV, human papillomavirus; PCR, polymerase chain reaction.
*SPF10 PCR-DEIA-LIPAX version 1 (Labo Bio-medical Products, Rijswijk, the Netherlands); †Trovagene (Trovagene Inc., San Diego, CA, USA); ‡Linear Array (Roche Molecular Systems, Pleasanton, CA, USA); §Cobas 4800 HPV (Roche Molecular Systems, Pleasanton, CA, USA); ∥CareHPV (Qiagen Corporation, Germantown, MD, USA); ¶Aptima (Gen-Probe Inc., San Diego, CA, USA); **Hybrid capture II HPV (Digene Corporation, Gaithersburg, MD, USA); ††Abbott RT PCR hrHPV (Abbott Molecular Inc., Des Plaines, IL, USA).

Table 3. Relative accuracy of urinary versus cervical human papillomavirus test, by other covariates, for detection of cervical intraepithelial neoplasia 2 or worse

| Variables            | No. of combinations (study) | Relative sensitivity (95% CI) | Relative specificity (95% CI) |
|----------------------|-----------------------------|------------------------------|-------------------------------|
| Urine collection time|                             |                              |                               |
| Morning first        | 3 (3)                       | 0.97 (0.88–1.07)             | 1.02 (0.77–1.34)              |
| Others               | 27 (18)                     | 0.83 (0.77–0.90)             | 1.07 (1.03–1.11)              |
| Urinary stream       |                             |                              |                               |
| Initial stream       | 23 (16)                     | 0.83 (0.76–0.91)             | 1.10 (1.03–1.18)              |
| Others               | 7 (6)                       | 0.87 (0.75–1.02)             | 1.03 (1.00–1.05)              |
| Collecting device    |                             |                              |                               |
| Colli-pee            | 6 (2)                       | 0.98 (0.91–1.05)             | 1.03 (0.90–1.19)              |
| Others               | 24 (19)                     | 0.82 (0.75–0.89)             | 1.07 (1.03–1.11)              |
| Preservative for storage|                          |                              |                               |
| Yes                  | 18 (11)                     | 0.91 (0.85–0.97)             | 1.03 (1.01–1.04)              |
| No or unknown        | 12 (10)                     | 0.74 (0.63–0.88)             | 1.16 (1.06–1.27)              |
| National income status|                            |                              |                               |
| Middle and low       | 5 (3)                       | 0.93 (0.84–1.02)             | 1.01 (0.99–1.04)              |
| High                 | 25 (18)                     | 0.84 (0.78–0.92)             | 1.11 (1.04–1.18)              |
| Population/setting   |                             |                              |                               |
| Primary screening    | 4 (3)                       | 0.79 (0.61–1.03)             | 1.03 (1.01–1.04)              |
| Follow-up/high-risk  | 26 (18)                     | 0.85 (0.79–0.92)             | 1.11 (1.04–1.18)              |

CI, confidence interval.

Table 4. Effects of covariates on the relative accuracy of urinary versus cervical human papillomavirus test (multivariate meta-regression)

| Variables            | Relative sensitivity | Relative specificity |
|----------------------|----------------------|----------------------|
|                      | Coefficient          | SE       | p-value | Coefficient | SE       | p-value |
| Urine collection time| Morning first vs. others | 0.168   | 0.658 | 0.802 | 0.011 | 0.212 | 0.960 |
| Collecting device    | Colli-pee vs. others | −0.111 | 0.705 | 0.876 | −0.018 | 0.216 | 0.965 |
| Preservative for urine storage| Yes vs. no/unknown | 0.173 | 0.283 | 0.548 | −0.473 | 0.134 | 0.002 |
| Income status of country| Middle and low vs. high | 0.094 | 0.574 | 0.872 | −0.066 | 0.159 | 0.684 |
| hrHPV assays         | PCR-base vs. others | 0.737 | 0.330 | 0.037 | −0.028 | 0.122 | 0.819 |
| Population/setting   | Follow-up/high-risk vs. primary screening | 0.148 | 0.659 | 0.828 | −0.133 | 0.216 | 0.544 |

hrHPV, high-risk human papillomavirus; PCR, polymerase chain reaction; SE, standard error. *p-value from random effects meta-regression using restricted maximum likelihood.
3. Absolute clinical accuracy of HPV test on urine and clinician-collected samples

In primary screening studies, the pooled sensitivity of HPV test was 76% (95% CI=56–95) in urine, while it was 97% (95% CI=0.93–1.00) in clinician-collected samples for detection of CIN2 or worse. The pooled specificity of HPV test to exclude CIN2 or worse was similar between urine (87%, 95% CI=0.83–0.89) and clinician-collected samples (85%, 95% CI=0.79–0.91). In the follow-up studies, pooled sensitivity for detection of CIN2 or worse was 79% (95% CI=0.72–0.86) and 93% (95% CI=0.89–0.96) in urine and clinician-collected samples, respectively. The specificity in urine (48%, 95% CI=0.42–0.54) and clinician-collected samples (42%, 95% CI=0.36–0.48) were comparable. Fig. S3 presents a forest plot of absolute clinical accuracy for HPV test across all studies included in the meta-analysis.

Fig. S4 shows the summary receiver operating characteristics curves for the detection of CIN2 or worse. It suggests that the area under the curve were 0.86 (95% CI=0.83–0.89) for cervical HPV and 0.74 (95% CI=0.69–0.77) for urinary HPV. The pooled DOR was 14 (95% CI=7–30) for clinician-taken samples and 5 (95% CI=3–7) for urine samples.

4. Publication bias

The Deeks’ funnel plot asymmetry test indicated that there is no significant publication bias for urinary HPV test (p=0.24) and clinician-collected samples (p=0.40) (Fig. S5).

DISCUSSION

This study demonstrates that urinary HPV test is capable of detecting ≥79% of CIN2 or worse and urinary HPV test using a some PCR-based assay such as GP5+/6+, SPF10 and non GP5+/6+ PCR had comparable sensitivity to that of clinician-collected samples. In the primary screening and follow-up setting, the sensitivity of urinary HPV was 21% and 14% lower than that of cervical HPV test using clinician-collected samples, respectively. The specificity to exclude CIN2 or worse was 87% for primary screening studies, and 48% for follow-up studies. A 2% and 6% higher specificity of urinary HPV test was observed compared to the HPV test using clinician-collected samples, respectively. In terms of relative accuracy, the sensitivity of the urinary HPV test was significantly lower than that of its clinician-collected counterpart, whilst the specificity of the urinary HPV test was slightly higher than that of the HPV test on clinician-collected samples.

This study demonstrated that the type of HPV assay used is the most important factor affecting tests accuracy. In addition, meta-regression suggests that only the type of HPV test is a significant factor impacting on sensitivity of urinary HPV compared with cervical HPV. Overall, the HPV assay based on target amplification by PCR showed a higher accuracy in urine than the HPV assay based on signal amplification or messenger RNA (mRNA) tests with Apta. Some PCR amplifications such as GP5+/6+, other non-GP5+/6+, and SPF10 shared a similar sensitivity and specificity in both urine samples and clinician-collected samples. For Abbott or CareHPV, sensitivity was not significantly lower in urine; however, both hrHPV assays were only used in only one study [26,31]. The HC2 and APTIMA urinary HPV showed considerably lower sensitivity than that of the HPV test on clinician-collected samples. These findings are consistent with studies on HPV test on self-collected vaginal samples. In meta-analyses, the PCR based HPV assays showed similar levels of sensitivity for self-collected samples compared to that of clinician-collected samples, whereas sensitivity was significantly
lower when HPV test was based on signal amplification or mRNA tests [4,32]. In a meta-analysis of urinary HPV test, the influence of the HPV assay type had not been assessed, as PCR-based hrHPV assays were used in most studies [11]. The lower sensitivity of urinary HPV test based on signal amplification or mRNA tests may be attributable to the relatively low viral load of hrHPV from urine. As urine may contain insufficient exfoliated cervical cells, variations in sensitivity between hrHPV assays may have a significant impact on test accuracy. HPV assays based on target amplification, such as a PCR-based HPV test, are more sensitive for the detection of hrHPV compared with a HPV assay based on signal amplification [33]. In addition, several technical factors which are varied by HPV assays, including primer and DNA extraction, may contribute to the variability of urine test accuracy even among PCR-based hrHPV assays [10,34]. Moreover, whether whole urine was processed instead of urine supernatant to prevent the loss of cell-free DNA may affect the outcome of urinary HPV in DNA extraction [10]. It was not assessed due to lack of information in this study.

Our findings did not demonstrate any clear moderator effects of urine collection time, urinary stream, sampling device, and preservative use for storage. When the initial urinary stream in the morning was collected or a Colli-pee or preservative solution was used, a higher sensitivity in the urinary HPV test was generally observed. However, meta-regression suggested that these covariates are not significant factors to affect the test accuracy after controlling effect of hrHPV assays. Previous meta-analysis on self-collected vaginal sample did not show an effect from the sampling device on relative sensitivity and specificity, which is in line with our findings [4,32]. However, finding that the relative sensitivity did not vary with urine collection time had not been anticipated. Previous comparative studies showed the favorable results for morning first or initial stream of urine. The meta-analysis demonstrated that the accuracy of HPV test on first-void urine was significantly higher than that of the midstream or random urine sample [11]. Senkomago et al. [14] compared the HPV test accuracy between the morning-first void at home, initial stream, and midstream at a clinic, finding that the hrHPV positivity rate was similar in the morning-first void and initial stream, and lower for midstream samples. Leeman et al. [13] demonstrated that HPV test in very first-void urine in the morning and another first void urine at clinic showed high sensitivity for detecting CIN2 or worse without any significant differences. Although there is no significant impact of urine sampling time in our findings, we should be cautious about drawing conclusions from these results. There is a considerable variance in the study design and settings in few comparisons of urine collection time from the included studies. Similarly, there were no comparative studies on the use of sampling devices or preservative for hrHPV DNA test and one studies on storage medium [20]. Therefore, it is difficult to conclude that the urine collection time, urinary stream, sampling device, or use of preservatives does not affect the test results.

This study suggested that sensitivity of urinary HPV versus cervical HPV was slightly higher in low- and mid-income countries than high-income countries, but national income status was not a meaningful factor that impact clinical accuracy of HPV test. Higher sensitivity of urine HPV in low- and mid-income countries may be due to high prevalence of HPV infection or multiple HPV infection [35-37]. However, as only 3 clinical trials of urinary HPV test in middle and low-income countries were obtained, there is no certainty in determining as to whether the incomes of these countries materially affect the accuracy of urinary HPV test.

Urine HPV may be useful in specific population who have cultural preference or acceptance. A Thailand study demonstrated that the 3 most common reasons of women for not undergoing screening are no symptoms, fear of pain, and embarrassment [38]. In a Korean study evaluating
satisfaction with cervical cancer screening modalities, satisfaction of urine HPV was significantly higher than that of cervical HPV [8]. Urine HPV may help to improve the uptake of cervical cancer screening by eliminating emotional or cultural barriers due to clinician sampling.

There are some limitations to this study; as it included primary screening and follow-up studies, the absolute sensitivity and specificity varied based on the study setting. For this reason, the relative accuracy of comparing urinary tests versus tests clinician-collected samples was assessed, and it was found that there was no substantial variance based on the study setting. There were limitations in identifying clinical factors that affect test accuracy due to the lack of detailed data from the studies that were included. Meta-regression failed to show a significant effect from urine collection, sampling device or storage medium. A well-designed clinical trial to evaluate the effect of sample handling or storage is urgently needed. Another limitation is relatively small number (21) and small sample sizes (<500) of the included studies. Especially, since there were only 3 studies in primary screening setting, further research will be needed to determine the clinical accuracy of HPV test on urine for screening purpose. In addition, the effect of sample postal conditions and clinical accuracy of HPV test on urine for detecting CIN3 were not assessed due to lack of information from included studies. Finally, there is a minor confusion regarding the definition of first void. While it generally refers to the initial stream of urine, some studies refer to it as morning first urine. To prevent misunderstandings arising from this difference, we regarded first void as initial stream of urine and if there is no mention that urine was collected in the morning, we did not regard first void as morning first urine.

Nevertheless, this study has several strengths. To the best of our knowledge, this study is the first meta-analysis that assesses the clinical performance of urinary HPV test to detect CIN2 or worse. Most previous systematic reviews have been on self-collected vaginal sample and one meta-analysis assessed only virological accuracy of urinary HPV test. In addition, we tried to examine the impact of various factors such as study population, national income status, urine collection time, urinary stream, sampling device, and preservative use for storage on clinical relative accuracy of urinary HPV test.

Our findings suggest that urinary HPV test with appropriate HPV assays, collection methods, sampling devices, and storage may serve as an alternative strategy for cervical cancer screening. This is particularly the case in certain populations that cannot be used to screen for cervical cancer and have a heavy burden of cervical cancer. This study confirmed the need for a randomized non-inferiority trial to evaluate the clinical accuracy of urinary HPV test within an organized screening setting using optimal HPV assays, collection methods, and storage. In terms of HPV test on self-collected samples, a meta-analysis, the randomized trial (IMPROVE study), confirmed that the clinical performance of HPV self-sampling is not inferior to that of clinician-collected samples to detect CIN2 or worse [39]. Furthermore, well-designed comparative studies to assess the effect of urine collection methods, collecting devices, and storage media are necessary to improve the clinical performance of HPV test in urine.

In conclusion, although clinician-collected sampling should be recommended in screening program using HPV test, urinary HPV test may be offered as an additional test method for women who do not participate in a cervical cancer screening program if validated PCR-based assays are used. Some PCR-based HPV test share similar accuracy on urine and clinician-collected samples to detect CIN2 or worse. Urinary HPV test based on signal amplification or mRNA tests should not be recommended.
ACKNOWLEDGEMENTS

We thank Korea national health insurance cooperation for providing database.

SUPPLEMENTARY MATERIALS

**Table S1**
Search strategies

Click here to view

**Table S2**
Used HPV assay

Click here to view

**Table S3**
Used device for urine collection and preservative for urine transport/storage.

Click here to view

**Table S4**
Relative sensitivity of specificity of urinary versus cervical HPV test in urine collection time, urinary stream, collecting device, preservative use by type of HPV assay

Click here to view

**Fig. S1**
Quality assessment (Quality Assessment of Diagnostic Accuracy Studies-2).

Click here to view

**Fig. S2**
Relative sensitivity (A) and specificity (B) of urinary versus cervical HPV test, by HPV assays, for detection of CIN2 or worse.

Click here to view

**Fig. S3**
Forest plot of absolute sensitivity and specificity of (A) urinary HPV test and (B) cervical HPV test.

Click here to view

**Fig. S4**
Summary of receiver operating characteristics curves of (A) urinary HPV test and (B) cervical HPV test for the detection of cervical intraepithelial neoplasia 2 or worse.

Click here to view
Fig. S5
Funnels for publication bias in studies of (A) urinary HPV and (B) cervical HPV.

Click here to view

SUPPLEMENTARY REFERENCES

Click here to view

REFERENCES

1. Melnikow J, Henderson JT, Burda BU, Senger CA, Durbin S, Weyrich MS. Screening for cervical cancer with high-risk human papillomavirus testing: updated evidence report and systematic review for the us preventive services task force. JAMA 2018;320:687-705. PUBMED | CROSSREF
2. US Preventive Services Task Force, Curry SJ, Krist AH, Owens DK, Barry MJ, Caughey AB, et al. Screening for cervical cancer: US preventive services task force recommendation statement. JAMA 2018;320:674-86. PUBMED | CROSSREF
3. Arbyn M, Weiderpass E, Bruni L, de Sanjosé S, Saraiya M, Ferlay J, et al. Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. Lancet Glob Health 2020;8:e191-203. PUBMED | CROSSREF
4. Arbyn M, Smith SB, Temin S, Sultana F, Castle P; Collaboration on Self-Sampling and HPV Testing. Detecting cervical precancer and reaching underscreened women by using HPV testing on self samples: updated meta-analyses. BMJ 2018;363:k4823. PUBMED | CROSSREF
5. Ginsburg O, Paskett ED. Ethnic and racial disparities in cervical cancer: lessons from a modelling study of cervical cancer prevention. Lancet Public Health 2018;3:e8-9. PUBMED | CROSSREF
6. Peto J, Gilham C, Fletcher O, Matthews FE. The cervical cancer epidemic that screening has prevented in the UK. Lancet 2004;363:249-56. PUBMED | CROSSREF
7. Sellors JW, Lorincz AT, Mahony JB, Mielzynska I, Lytvyn A, Roth P, et al. Comparison of self-collected vaginal, vulvar and urine samples with physician-collected cervical samples for human papillomavirus testing to detect high-grade squamous intraepithelial lesions. CMAJ 2000;163:513-8. PUBMED
8. Shin HY, Lee B, Hwang SH, Lee DO, Sung NY, Park JY, et al. Evaluation of satisfaction with three different cervical cancer screening modalities: clinician-collected Pap test vs. HPV test by self-sampling vs. HPV test by urine sampling. J Gynecol Oncol 2019;30:e76. PUBMED | CROSSREF
9. Chamot E, Mulambia C, Kapambwe S, Shrestha S, Parham GP, Macwan'gi M, et al. Preference for human papillomavirus-based cervical cancer screening: results of a choice-based conjoint study in Zambia. J Low Genit Tract Dis 2015;19:119-23. PUBMED | CROSSREF
10. Van Keer S, Pattyn J, Tjalma WAA, Van Ostade X, Ieven M, Van Damme P, et al. First-void urine: a potential biomarker source for triage of high-risk human papillomavirus infected women. Eur J Obstet Gynecol Reprod Biol 2017;216:141. PUBMED | CROSSREF
11. Pathak N, Dodds J, Zamora J, Khan K. Accuracy of urinary human papillomavirus testing for presence of cervical HPV: systematic review and meta-analysis. BMJ 2014;349:g5264. PUBMED | CROSSREF
12. Cuzick J, Cadman L, Ahmad AS, Ho L, Terry G, Kleeman M, et al. Performance and diagnostic accuracy of a urine-based human papillomavirus assay in a referral population. Cancer Epidemiol Biomarkers Prev 2017;26:1053-9. PUBMED | CROSSREF
13. Leeman A, Del Pino M, Molijn A, Rodriguez A, Torné A, de Koning M, et al. 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 2017;124:1356-63.

14. Senkomago V, Des Marais AC, Rahangdale L, Vibat CR, Erlander MG, Smith JS. Comparison of urine specimen collection times and testing fractions for the detection of high-risk human papillomavirus and high-grade cervical precancer. J Clin Virol 2016;74:26-31.

15. Alameda F, Bellosillo B, Fusté P, Musset M, Mariñoso ML, Mancebo G, et al. Human papillomavirus detection in urine samples: an alternative screening method. J Low Genit Tract Dis 2007;11:5-7.

16. Asciutto KC, Henningsson AJ, Borgfeldt H, Darlin L, Borgfeldt C. Vaginal and urine self-sampling compared to cervical sampling for HPV-testing with the Cobas 4800 HPV test. Anticancer Res 2017;37:4183-7.

17. Asciutto KC, Ernstson A, Forslund O, Borgfeldt C. Self-sampling with HPV mRNA analyses from vagina and urine compared with cervical samples. J Clin Virol 2018;101:69-73.

18. Bernal S, Palomares JC, Artura A, Parra M, Cabezas JL, Robles A, et al. Comparison of urine and cervical samples for detecting human papillomavirus (HPV) with the Cobas 4800 HPV test. J Clin Virol 2014;61:548-52.

19. Buchegger K, Viscarra T, Andana A, Lloret J, Zanella L, et al. Detection and genotyping of human papillomavirus virus (HPV): a comparative analysis of clinical performance in cervical and urine samples in Chilean women. Int J Exp Pathol 2018;11:5413-21.

20. Arias M, Jang D, Dockter J, Ratnam S, Shah A, Elit L, et al. Treatment of first-void urine with Aptima Transfer Solution increases detection of high-risk HPV E6/E7 mRNA. J Virol Methods 2019;267:48-52.

21. Padhy RR, Davidov A, Madrigal L, Alcide G, Spahiu A. Detection of high-risk human papillomavirus RNA in urine for cervical cancer screening with HPV 16 & 18/45 genotyping. Heliyon (Lond) 2020;6:e03745.

22. Piyathilake CJ, Badiga S, Chambers MM, Brill IK, Matthews R, Partridge EE. Accuracy of urinary human papillomavirus testing for the presence of cervical human papillomaviruses and higher grades of cervical intraepithelial neoplasia. Cancer 2016;122:2836-44.

23. Rohner E, Rahangdale L, Sanusi B, Knittel AK, Vaughan L, Chesko K, et al. Test accuracy of human papillomavirus in urine for detection of cervical intraepithelial neoplasia. J Clin Microbiol 2020;58:e00443-19.

24. Sahasrabuddhe VV, Gravitt PE, Dunn ST, Brown D, Allen RA, Eby YJ, et al. Comparison of human papillomavirus detections in urine, vulvar, and cervical samples from women attending a colposcopy clinic. J Clin Microbiol 2014;52:187-92.

25. Sahasrabuddhe VV, Gravitt PE, Dunn ST, Robbins D, Brown D, Allen RA, et al. Evaluation of clinical performance of a novel urine-based HPV detection assay among women attending a colposcopy clinic. J Clin Microbiol 2014;60:414-7.

26. Sargent A, Fletcher S, Bray K, Kitchener HC, Crosbie EI. 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 2019;9:e025388.

27. Sorbi F, Malentacchi F, Bussani C, Fantappie G, Pavone D, Turrini I, et al. HPV-DNA genotyping in women with cervical intraepithelial neoplasia undergoing to conization: comparison between cervical cytobrush and corresponding self-collected urine samples. Int J Gynecol Cancer 2017;27:423.

28. Stanczuk GA, Currie H, Baxter G, Foster A, Gibson L, Graham C, et al. Cobas 4800 HPV detection in the cervical, vaginal and urine samples of women with high-grade CIN before and after treatment. J Clin Pathol 2015;68:567-70.
29. Stanczuk G, Baxter G, Currie H, Lawrence J, Cuschieri K, Wilson A, et al. Clinical validation of hrHPV testing on vaginal and urine self-samples in primary cervical screening (cross-sectional results from the Papillomavirus Dumfries and Galloway-PaVDaG study). BMJ Open 2016;6:e010660.

PUBMED | CROSSREF

30. Tshomo U, Franceschi S, Tshokey T, Tobgay T, Baussano I, Tenet V, et al. Evaluation of the performance of human papillomavirus testing in paired urine and clinician-collected cervical samples among women aged over 30 years in Bhutan. Virol J 2017;14:74.

PUBMED | CROSSREF

31. Xu H, Yu Y, George W, Smith JS, Hu S, Dang L, et al. Comparison of the performance of paired urine and cervical samples for cervical cancer screening in screening population. J Med Virol 2020;92:234-40.

PUBMED | CROSSREF

32. Arbyn M, Verdoort F, Snijders PJ, Verhoef VM, Suonio E, Dillner L, et al. Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: a meta-analysis. Lancet Oncol 2014;15:172-83.

PUBMED | CROSSREF

33. Snijders PJ, van den Brule AJ, Meijer CJ. The clinical relevance of human papillomavirus testing: relationship between analytical and clinical sensitivity. J Pathol 2003;201:1-6.

PUBMED | CROSSREF

34. Enerly E, Olofsson C, Nygård M. Monitoring human papillomavirus prevalence in urine samples: a review. Clin Epidemiol 2013;5:67-79.

PUBMED | CROSSREF

35. Atkinson AE, Mandujano CAM, Bejarano S, Kennedy LS, Tsongalis GJ. Screening for human papillomavirus in a low- and middle-income country. J Glob Oncol 2019;5:JGO1800233.

PUBMED | CROSSREF

36. de Martel C, Plummer M, Vignat J, Franceschi S. Worldwide burden of cancer attributable to HPV by site, country and HPV type. Int J Cancer 2017;141:664-70.

PUBMED | CROSSREF

37. Brenner H, Gefeller O. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. Stat Med 1997;16:981-91.

PUBMED | CROSSREF

38. Chaowawanit W, Tangjittgamol S, Kantathavorn N, Phoolcharoen N, Kittisiam T, Khunnarong J, et al. Knowledge, attitudes and behavior of Bangkok metropolitan women regarding cervical cancer screening. Asian Pac J Cancer Prev 2016;17:945-52.

PUBMED | CROSSREF

39. Polman NJ, Ebisch RMF, Heideman DAM, Melchers WJG, Bekkers RLM, Molijn AC, et al. Performance of human papillomavirus testing on self-collected versus clinician-collected samples for the detection of cervical intraepithelial neoplasia of grade 2 or worse: a randomised, paired screen-positive, non-inferiority trial. Lancet Oncol 2019;20:229-38.

PUBMED | CROSSREF