

Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2021 April 01; 30(4): 661–668.

doi:10.1158/1055-9965.EPI-20-1226.

A randomised comparison of different vaginal self-sampling devices and urine for human papillomavirus testing – Predictors

5.1

Louise Cadman^{1,†}, Caroline Reuter¹, Mark Jitlal¹, Michelle Kleeman^{1,3}, Janet Austin¹, Tony Hollingworth¹, Anna L. Parberry², Lesley Ashdown-Barr¹, Deepali Patel^{1,3}, Belinda Nedjai¹, Attila T. Lorincz¹, Jack Cuzick¹

¹Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University of London, Charterhouse Square, London

²Colposcopy Department, Royal London Hospital, London

³NIHR Biomedical Research Centre, Guy's and St Thomas' NHS Foundation Trust and King's College, London

Abstract

Background—HPV-based screening is rapidly replacing cytology as the cervical screening modality of choice. In addition to being more sensitive than cytology, it can be done on self-collected vaginal or urine samples. This study will compare the hrHPV positivity rates and sensitivity of self-collected vaginal samples using four different collection devices and a urine sample.

Methods—620 women referred for colposcopy were invited to provide an initial stream urine sample collected with the Colli-Pee® device and take two vaginal self-samples, using either a dry flocked swab (DF) and a Dacron swab (WD), or a HerSwab (HS) and Qvintip (QT) device. HPV testing was performed by the BD Onclarity™ HPV Assay.

Correspondence to: Jack Cuzick.

Corresponding Author: Professor Jack Cuzick: Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University of London, Charterhouse Square, London, EC1M 6BQ, United Kingdom, **Tel:** +44 20 7882 3518 **Fax:** +44 20 7882 3890 j.cuzick@qmul.ac.uk.

[†]Deceased

Disclosure of potential conflicts of interest

Jack Cuzick – Support to my institution from Becton Dickinson, Qiagen and Genefirst has been provided for this, and for related studies from Genera Biosystems, Hologic, and Trovagene. I have no personal conflicts to disclose. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: J. Cuzick, L. Cadman, C. Reuter, J. Austin, A. Lorincz

Development of methodology: J. Cuzick, L. Cadman, C. Reuter, J. Austin, A. Parberry, A. Lorincz.

Acquisition of data: L. Cadman, J. Austin, D. Patel, L. Ashdown-Barr, T. Hollingworth

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): J. Cuzick, M. Jitlal, C. Reuter, L. Cadman

Writing, review, and/or revision of the manuscript: L. Cadman, J. Cuzick, C. Reuter, M. Jitlal, M. Kleeman, J. Austin, T. Hollingworth, A. Parberry, L. Ashdown-Barr, D. Patel, B. Nedjai, A. Lorincz

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): L. Cadman, C. Reuter, M. Jitlal, J. Austin, M. Kleeman, D. Patel, L. Ashdown-Barr, T. Hollingworth

Study supervision: L. Cadman, J. Cuzick

Results—600 vaginal sample pairs were suitable for analysis and 505 were accompanied by a urine sample. Similar positivity rates and sensitivities for CIN2+ and CIN3+ were seen for DF, WD and urine, but lower values were seen for QT and HS. No clear user preferences were seen between devices, but women found urine easiest to collect, and were more confident they had taken the sample correctly. The lowest confidence in collection was reported for HS.

Conclusions—Urine, a dry flocked swab and wet dacron swab all performed well and were well received by the women, whereas the Qvintip and HerSwab devices were less satisfactory.

Impact—This is the first study to compare five self-sampling methods in the same women taken at the same time. It supports wider use of urine or vaginal self-sampling for cervical screening.

Keywords

Cervical cancer screening; human papillomavirus; HPV; self-sampling; urine

Introduction

Self-sampling for high risk Human Papillomavirus (hrHPV) has been recognised as an alternative to clinician collected samples for cervical screening. In the UK the uptake of cervical screening has been declining in recent years. The most recent quarterly figures (31 March 2019) show 70.2% of women in England aged 25-49 years were screened adequately within the previous 3.5 years, and 76.4% for women aged 50 to 64 years within the previous 5.5 years (1).

The introduction of primary screening by hrHPV testing presents an opportunity for under-screened women to be offered self-sampling at home as an alternative to a speculum examination in a health-care setting (2, 3). To date self-sampling has been suggested primarily for those who are under screened or who have not been screened. However, given a preference for self-sampling by most women, accumulating data that it is as sensitive for HPV testing as clinician taken samples and the potential for major savings for health service, self-sampling has the potential to become the first choice for screening in all women (4, 5, 6).

An updated meta-analysis using a range of assays in a range of settings, found PCR based assays in screening populations had similar sensitivities of self-sampling compared to clinically taken samples (7, Table 10) with sensitivities of about 96% for CIN2+ and CIN3+ for both methods, and specificities for <CIN2 of about 79%. Similar results were seen in a large randomised trial (8). In general it has been suggested that the specificity of self-collected, vaginal samples, may be lower than for clinician taken cervical samples due to passenger hrHPV found in the vagina (9), but this is less clearly established. Increasing interest in self-sampling for HPV has led to the development of a number of kits, and dozens of devices are currently commercially available (10). These include brush-type devices and different types of swabs. Some use a liquid specimen transport medium and some are transported dry. The latter has obvious advantages in terms of safety for home use, especially when small children are present, and also fewer restrictions for posting. A number of kits for urine collection are also available (<https://www.amazon.co.uk/Medical-Specimen->

[Collection-Containers/b?ie=UTF8&node=6285950031](#)). However, none of these provide an easy way to collect initial stream urine, which has been shown to be important (11), and also do not contain a stabilization medium, which has to be added separately to avoid DNA degradation.

Cost and simplicity of use are also important factors - especially in low and middle income countries, where self-sampling may prove more cost-effective and be the only practical option. In this study we compared simple swabs against more complex and expensive devices, and also those transported dry vs wet in a liquid transport medium. As vaginal collection has an invasive element we also included urine testing, which is non-invasive and may be more acceptable to some women or cultures, but has been less extensively studied. A systematic review of 14 studies found that urine testing had good accuracy for the presence of cervical HPV, and that initial stream urine yielded more accurate results than random or midstream sampling (12). In a recent study in a referral population, we found similar performance of the Trovagene test and 5 other HPV assays in clinical cervical samples, but a lower sensitivity in a urine sample (13). Three other small studies with matched urine, self-collected vaginal and clinician taken cervical samples in a referral population have been reported. Two (14,15) found similar sensitivities for CIN2+ between all 3 samples, but the third found a lower sensitivity for urine (16). All used an Evalyn (or similar Viba) brush for the vaginal samples. Urine was collected in a sterile pot in two of these studies (14,17) before adding a preservative, and the third used the Colli-Pee device (15). Different HPV assays were used in the different studies. Here we compare the performance of four vaginal sampling devices, including wet and dry transport methods, and an initial stream urine sample in a referral population using the Becton Dickinson Onclarity™ HPV test. We also conducted additional analyses to examine the effect of adjusting for sample cellularity and different positivity thresholds.

Materials and methods

Study population and procedures

In all cases an initial stream urine sample was requested as the first sample, using the Colli-Pee® collection kit (Novosanis NV), which collects a 20ml sample of which 7ml is a prefilled Urine Conservation Medium (UCM). Two vaginal samples were provided subsequently. Women were randomised to use either the dacron based *digene*® Female Swab Specimen Collection Kit (Qiagen GmbH) placed in 1ml liquid specimen transport medium (STM) (denoted as wet dacron – **WD** below) and the Copan FLOQswab™ (Copan Diagnostics Inc) as a dry sample (dry flocced - **DF**) or the Qvintip® kit (Aprox AB) (**QT**) and the HerSwab (Eve Medical) both as dry samples (**HS**) (Supplementary Figure S1). The order of sampling was also randomised, leading to 4 equal sized groups. Samples were stored locally at 4-8°C before being transported to the laboratory within 2 weeks of collection.

Recruitment took place between April 2016 and August 2017. New referrals to the Royal London Hospital colposcopy clinic who had a positive screening result were sent information about the study and an offer to join it with their appointment letter. This was also sent to women under observation for prior abnormalities, provided that their most recent

abnormal smear test was within the last three years. Pregnancy or ablation/excision treatment to the cervix within three years were exclusion criteria. All samples were taken prior to colposcopic examination. Cervical biopsies were only taken where clinically indicated. As well as having a negative biopsy, a negative colposcopy with satisfactory evaluation of the transformation zone was considered to indicate no disease even without a biopsy. Patient outcomes were collected for one year after enrolment and this period would normally include a follow-up visit for women with low-grade disease at first visit. Histology was reported by hospital pathologists and reviewed internally according to NHS Trust quality control practices. Where cytology or histopathology results were re-graded following local review the reviewed result was used.

Women were also asked to complete a simple survey evaluating their preference for a vaginal sample vs urine, and preferred vaginal sampling device, discomfort with sample collection, confidence in sampling and ease or difficulty of each collection method.

Ethical approval

The study was reviewed and approved in February 2016 by the London – City Road & Hampstead Research Ethics Committee [Reference 15/LO/2134], and was conducted in concordance with the principles sent out in the Declaration of Helsinki guidelines. Written informed consent was obtained from all participating women. The trial was registered with the ISRCTN Registry (ISRCTN68980717).

Laboratory methods

All samples were tested in the Wolfson Institute of Preventive Medicine laboratory. On receipt, the WD swab samples, which were transported in STM were frozen at -20°C until further processing. The dry DF, HS and QT samples were placed into 15ml tubes, 8ml of PreservCyt[®] solution (Hologic) was added and the tubes were vortexed thoroughly at maximum speed to re-suspend the cells into the solution. These samples were then split into 700 μl aliquots. The WD samples, collected in 1ml of STM, were split into 100 μl aliquots before freezing. All the aliquots were kept frozen until tested. Urine was refrigerated at 4°C and split into 2ml vials within seven days of receipt before being frozen at -20°C .

HPV testing

All samples were tested for hrHPV using the BD Onclarity[™] HPV Assay. This assay is based on simultaneous PCR amplification and detection of target DNA using amplification primers and fluorescent-labelled detector probes. An internal human β -globin (HBB) gene control was used to measure sample cellularity to judge sample adequacy. The assay detects 14 high-risk HPV types overall: six types are individually genotyped (16, 18, 31, 45, 51 and 52) and the other types are reported in groups (33, 58), (56, 59, 66) and (35, 39, 68). The manufacturer's predefined threshold for positivity was a Ct value 38.3 for HPV16 and 34.2 for all other genotypes and the β -globin internal control. Use of a Ct threshold of either 34.2 or 38.3 for all genotypes was also explored.

For DF, HS and QT, 500 μl of sample was added to an LBC Diluent tube. 50 μl of the WD sample was added to a tube containing the Cervical Brush Diluent (CBD) medium. For

urine, 2ml was added to empty LBC Diluent tubes. The Onclarity software produces an 'indeterminate' result when no HPV is detected and HBB internal control did not meet the amplification threshold, indicating insufficient cellularity of the specimen. Samples producing invalid results were repeat tested once. Twenty samples were indeterminate a second time and excluded from the analyses. One urine sample was of insufficient cellularity to provide a valid result.

Statistical methods

The Ct values were compared between devices using the unpaired Wilcoxon Test. Comparisons of the numbers of women who were hrHPV positive were based on binomial statistics. The Fisher's Exact Test was used when any of the four expected cell values was 10, but otherwise chi squared approximations were used.

Agreement in hrHPV positivity between devices was calculated using both an unadjusted approach and also an approach adjusted for cellularity, using the kappa statistic. The direction of discordance (+/- vs -/+) was calculated by using McNemar's odds ratio. Sensitivity for CIN2+ and CIN3+ and specificity for <CIN2 were calculated using standard approaches. Ct values adjusted for cellularity were obtained by first subtracting the internal control Ct value from the raw HPV Ct value. Due to its extensive prior use WD was selected as the reference device, and a further additive adjustment was made to give the same overall positivity rate as for the raw values for this device at the predefined genotype specific Ct cutoff (38.3 for HPV16 and 34.2 for other genotypes) for the primary analysis, and at the 34.3 and 38.3 cutoffs for all genotypes in the secondary analyses. The same additive adjustment was made for all other devices, except for urine, where cellularity was generally higher and a cellularity adjustment was not explored. Bootstrapping was used to estimate the confidence intervals for the kappa statistic and the difference between the raw and adjusted kappa statistics. A small preliminary study of DNA stability was undertaken using aliquots from 60 women (30 sampled with DF and 30 with HS) comparing baseline levels of DNA and nucleic acid using Nanodrop and RNase to those of matched aliquots stored at 25°C for one or two weeks. No significant differences were seen with either assay.

All statistical analyses were performed in R (v3.6.3).

Results

620 women consented to take part, but twenty were excluded as not eligible and the primary analysis group comprised the remaining 600 women (Figure 1). Of these 505 women also provided an initial stream urine sample of between 8-20ml, inclusive of 7ml UCM preservative.

Median age was 29y (IQR: 27-33, range 21-63), and 330 (53.2%) were aged less than 30 years, 261 (42.1%) were aged 30-44 years and 29 (4.4%) were aged 45 or older. Each woman provided two vaginal self-samples (Figure 1). Age at enrolment, referral cytology reading, time since most recent abnormal cytology and previous treatment were similar across all four randomisation groups (Supplementary Table S1).

There were 66 cases of CIN2 and 68 cases of CIN3+ within the one-year follow-up period. This included two stage 1A1 squamous cell carcinomas (SCC), one stage 1B SCC and one stage 1A1 adenocarcinoma. A cross-tabulation of histology and cytology results is shown in Supplementary Table S2.

hrHPV positivity. The HPV assay could not be performed on 20 samples (HS= 16; QT = 4) and one urine sample due to low cellularity (Figure 1). These samples all came from different women. There were no significant differences in hrHPV positivity rates according to collection order for any of the devices (Supplementary Table S3), and results from first and second samples have been combined for all analyses. HPV positivity rates for all hrHPV types combined are shown in Table 1. At the predefined HPV test positivity cutoff of Ct 38.3 for HPV16 and 34.2 for all other genotypes, positivity of 71.7% was seen for WD and DF, with non-significantly higher values for urine (76.0%, $P = 0.17$) and non-significantly lower rates for QT (69.3%; $p=0.52$) and for HS (65.1%; $p=0.09$). Agreement on the same sample between WD and DF was high ($\kappa = 0.801$, 95% CI: 0.777, 0.826), but lower between QT and HS ($\kappa = 0.753$, 95% CI: 0.723, 0.779) and lower still between urine and the 4 vaginal samples ($\kappa = 0.568-0.646$), largely due to the higher positivity rate with urine (Supplementary Table S4). Positivity by genotype is shown in Supplementary Table S5. There were no statistically significant or obvious differences in genotype-specific positivity by device, although type 45 was lower for Herswab and Quintip, and the (35,39,68) positivity was higher for urine and Quintip using the primary cutoffs without a cellularity adjustment.

The highest sensitivity for CIN3+ was obtained with WD (91.2%). Sensitivities for CIN3+ for the other devices were as follows: urine (89.7%), DF (88.2%), QT (81.8%) and HS (77.4%), none of which were significantly lower than WD (Table 2A and Supplementary Figure S2). A similar order was seen for CIN2+, and again no differences were significant. Increasing the cutoff to 38.3 for all HPV types gave higher sensitivities for all devices (Table 2C). The order of sensitivities between the devices was similar to as before, but the differences were smaller, especially for HS and no differences were significant. However, there was a substantial loss of specificity for <CIN2 for all devices, e.g. dropping from 34.2% to 26.3% for WD in this referral population. Slightly but not significantly better agreement between the different devices in terms of overall hrHPV positivity was seen at the higher 38.3 cutoff (Supplementary Table S4).

Cellularity Adjustment

The number of cells collected in vaginal samples by the different devices and the urine sample were compared using the Ct values for the β -globin internal control as a proxy for cell number (Figure 2). The greatest number of cells was observed for DF (median Ct:24.3 (IQR:23.4, 25.5)), followed by urine (median Ct:24.6 (IQR:23.4, 26.2)), QT (median Ct:25.3 (IQR:24.0, 27.2)), WD (median Ct:25.7 (IQR:25.2, 26.8)) and HS (median Ct:26.1 (IQR:23.9, 30.0)). DF provided more cells than QT and HS ($P < 0.0001$ for both tests). A Ct value 34.2 for β -globin was used to determine sample adequacy, and samples with values below this level were deemed inadequate unless a positive value was seen for some HPV type. On average HS samples provided the smallest median number of cells and 16/300

samples (5.3%) failed to meet this sample adequacy measure. Four QT samples (1.3%) also failed to meet this requirement, as did one urine sample, but all other samples were deemed adequate (Figure 1). The largest variation in cellular yield based on β -globin Ct values was observed for samples collected with HS and the smallest with WD.

The hrHPV positivity rate was also examined after adjustment for cellularity for the vaginal devices. This was particularly relevant for HS, because of its generally low and variable cellularity. After adjustment its overall positivity rate, sensitivity for CIN2+ at the predefined genotype specific cutoff increased to 85.2%, which was comparable to WD and DF (Table 2A). Supplementary Table S4 indicated agreement between pairs of vaginal devices at the predefined type specific cutoff was non-significantly improved after cellular adjustment (WD and DF kappa 0.803 before vs 0.858 after; QT vs HS 0.739 before vs 0.850 after).

Questionnaire results

Of the 600 women, 504 provided an adequate urine sample and 454 of these replied regarding device preference (See supplement for the questionnaire). Within this group 46% preferred one of the two vaginal devices, 32% preferred urine and 22% expressed no preference (Supplementary Table S6, where further details by device are given). In general the vaginal device used first was preferred to the one used second, with the average preference for the first device being 29% vs 17% for the second sample ($P < 0.0001$). Preference for the urine sample was similar to that for the first vaginal sample. More women expressed no device preference in the WD/DF group, but Table 3 shows when restricted to women who expressed a preference between vaginal devices, there was a clear preference for DF over WD (71.4%, $P = 0.003$), but no clear preference between QT and HS (53.4% preferred HS, $P = 0.48$). Similar vaginal sampling device preferences were seen when the 95 women who did not provide a urine sample and one who had an inadequate sample were included. No correlation was seen between 'quite difficult' or 'very difficult' ease of use and cellularity ≥ 30 Ct for either QT or HS ($P > 0.5$ for both).

There was no clear preference for device according to age. Neither was time since last urination predictive of preference.

Ease of use was recorded on a five point scale: very easy, quite easy, neither easy nor difficult, quite difficult, very difficult. No effect of sample order was seen in this reply and combined results are shown in Table 4A. The HS device emerged as the device least easy to use with 12.6% indicating it was quite difficult or very difficult ($P < 0.003$ for all comparisons), with no other significant differences.

A major difference in confidence of correct use across devices was seen (Table 4B), with the greatest confidence seen for urine (61.6% as very confident and 4.4% not confident, $P < 0.0001$ all comparisons) and the least confidence for HS (36.4% very confident and 14.2% as not confident, $P < 0.03$ for not confident for all comparisons except DF). Lower numbers expressing high confidence were also seen for QT than for WD or DF, but these were of marginal significance.

Discussion

The focus of this study was to compare four different devices for vaginal self-sampling between themselves and also a urine sample, both for HPV positivity and user preference and confidence in collection. We have reported separately results for CIN2+ and CIN3+, because CIN2 is less likely to progress and has greater variability in diagnosis (16, 17) and CIN3 has been shown to have greater reproducibility (18).

Use of the 34.2 cutoff, as we previously have done for clinical samples, led to a clinically relevant reduction in sensitivity, but when a higher Ct cutoff was used for HPV16 (38.3), as recommended by the manufacturer and used previously (19), sensitivities and were similar to those seen in ours and other previous studies were seen (7,20–23). However as seen in Table 2, specificity was strongly dependent on the positivity threshold, and large reductions in specificity were seen when a 38.3 threshold was used for all genotypes. Sensitivity can be fairly accurately measured in a referral population, as it focuses on the diseased population, However, specificity is more problematic, as this is highly dependent on the HPV positivity rate, and passenger HPV infections not related to disease and likely to regress naturally are known to be common. While useful information on relative specificity of different tests and devices can be obtained, as was our primary goal here, a study in a screening population is required to assess absolute specificity. We are planning such a study of self vs clinician samples taken at the same visit in a screening population, which has not been previously undertaken on a large scale. In addition to evaluating the sensitivity and specificity in the intended use population, such a study will also examine preference issues in such a population, where the issues could be different than in a referral population, where women are already committed to having screening and other procedures.

High sensitivity is particularly important for self-sampling, especially for women who have declined previous invitations for a clinician taken test, as this may be a rare opportunity to screen them. Thus, technical issues, such as dilution amounts, and less restrictive thresholds for positivity that increase sensitivity, need to be explored. In general one could imagine three outcomes from a self-sample - (i) HPV negative – return to routine screening, (ii) strongly HPV positive based on higher risk genotypes, methylation profile, viral load and/or other markers identifiable in the self-sample, which would lead to immediate colposcopy and (iii) weaker positives, based on lower risk genotypes (e.g. 56,59,68), lack of methylation or other biomarkers in which case the next step could be a clinically taken sample where cytology could also be performed or a repeat self-sample at, say 3-12 months.

When not adjusted for cellularity, the WD swab transported in liquid STM and the dry transported DF gave the highest hrHPV positivity, overall and for CIN2+ and CIN3+. Women also expressed a preference for the DF swab over the WD device, but a limitation of the study design did not allow for direct comparisons of these devices with QT or HS. Urine samples gave similar sensitivities to WD and DF and were well received by the women, but specificities were lower. The HS performed least well, apparently largely due to the very large variability in cellularity of the sample between women. For 20 (13%) women no cells were detected with HS. When adjusted for cellularity the results were more similar to the other devices. However, women were least confident that they had provided a satisfactory

sample with this device. The Evalyn brush has been used in several other studies with good results, but in two previous studies in minority populations (23, 24) we found women were less confident in using it, and so in the current study we chose to put more emphasis on simpler devices, which the women had preferred. Although not included here, we note that the manufacturer of the Evalyn device also supplies a cheaper Viba brush which has the same brush head attached to a simple shaft handle and has reported equivalent performance to clinician-collected samples (25).

At a pragmatic level the DF swab appeared to be the device of choice - based on high detection rates, more favourable cost, and the important ability to transport the sample dry, which is especially important for home collection where small children may be present and transport via the post, as STM used for liquid transport contains guanidinium hydrochloride, which is known to be an irritant. We acknowledge a design weakness in that only wet dacron or dry flocked swab vs Quintip or Herswab were compared in the same woman. Ideally we would have randomised all possibilities in which each woman had an equal chance of receiving any two of the four possible vaginal samplers. However, this presented logistic challenges and a danger of errors in a busy clinic and was not implemented, although full randomisation of devices between women was achieved.

The urine sample collected in the Colli-Pee device was well received by the women and gave good results with high sensitivity, but lower specificity. In contrast, a recent study (25), where urine was collected by a different method, lower sensitivity but higher specificity was seen for urine than for self-collected vaginal samples. The liquid UCM preservative medium used here is non-toxic, which makes it suitable for home use, although urine is a body fluid and hence a biohazard.

In summary, there were clear differences in the performance and patient preferences between the 5 collection methods. Both WD and DF gave good results, and there was very good agreement in HPV positivity. Urine had a similar sensitivity to these two, but its specificity was lower. An important advantage of DF is that it can be transported dry. Sensitivity was very dependent on the positivity threshold and use of a higher Ct cutoff for HPV16 improved results. However, these results need confirmation in a routine screening population, and specificity needs evaluation in that context. The role of triage markers such as genotyping and methylation also require more study.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the staff of the gynaecology out-patients department, The Royal London Hospital, for their invaluable help. We are very grateful to all the women who took part in the study.

Financial support

Funding was provided by Cancer Research UK programme grant C569/A16891 to J Cuzick. Assay kits and an additional financial contribution was provided by Becton Dickson. Colli-Pee urine collection kits and HerSwab kits

were donated by Novosanis and Eve Medical respectively. Commercial companies had no role in study design, data collection, analyses or writing of this paper but were allowed to see the paper prior to submission for publication.

References

1. Public Health England. Cervical screening standards data report 1 April 2018 to 31 March 2019. PHE publications gateway number GW-968; 2020.
2. Freeman M, Waller J, Sasieni P, Lim AW, Marlow LA. Acceptability of non-speculum clinician sampling for cervical screening in older women: A qualitative study. *J Med Screen*. 2018; 25:205–10. [PubMed: 29439604]
3. Haguenoer K, Sengchanh S, Gaudy-Graffin C, Boyard J, Fontenay R, Marret H, et al. Vaginal self-sampling is a cost-effective way to increase participation in a cervical cancer screening programme: a randomised trial. *Br J Cancer*. 2014; 111:2187–96. [PubMed: 25247320]
4. Forrest S, McCaffery K, Waller J, Desai M, Szarewski A, Cadman L, et al. Attitudes to self-sampling for HPV among Indian, Pakistani, African-Caribbean and white British women in Manchester, UK. *J Med Screen*. 2004; 11:85–8. [PubMed: 15153323]
5. Waller J, McCaffery K, Forrest S, Szarewski A, Cadman L, Austin J, et al. Acceptability of unsupervised HPV self-sampling using written instructions. *J Med Screen*. 2006; 13:208–13. [PubMed: 17217611]
6. Darlin L, Borgfeldt C, Forslund O, Hénic E, Hortlund M, Dillner J, Kannisto P. Comparison of use of vaginal HPV self-sampling and offering flexible appointments as strategies to reach long-term non-attending women in organized cervical screening. *J Clin Virol*. 2013 Sep; 58(1):155–60. [PubMed: 23867008]
7. Arbyn M, Smith SB, Temin S, Sultana F, Castle P, Collaboration on S-S. et al. Detecting cervical precancer and reaching underscreened women by using HPV testing on self samples: updated meta-analyses. *BMJ*. 2018; 363:k4823 [PubMed: 30518635]
8. 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 Feb; 20(2):229–238. [PubMed: 30658933]
9. Belinson JL, Hu S, Niyazi M, Pretorius RG, Wang H, Wen C, et al. Prevalence of type-specific human papillomavirus in endocervical, upper and lower vaginal, perineal and vaginal self-collected specimens: Implications for vaginal self-collection. *Int J Cancer*. 2010; 127:1151–7. [PubMed: 20039323]
10. Othman NH, Zaki FHM. Self-Collection Tools for Routine Cervical Cancer Screening: A Review. *Asian Pac J Cancer Prev*. 2014; 15(20):8563–8569. [PubMed: 25374168]
11. Vorsters A, Van den Bergh J, Micalessi I, Biesmans S, Bogers J, Hens A, De Coster I, Ieven M, Van Damme P. Optimization of HPV DNA detection in urine by improving collection, storage, and extraction. *Eur J Clin Microbiol Infect Dis*. 2014 Nov; 33(11):2005–14. [PubMed: 24916950]
12. 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: 25232064]
13. Cuzick JM, 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 Jul; 26(7):1053–1059. [PubMed: 28223432]
14. Sargent A, Fletcher S, Bray K, Kitchener HC, Crosbie EJ. 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
15. Leeman A, Del Pino M, Molijn A, Rodriguez A, Torné A, de Koning M, Ordi J, van Kemenade F, Jenkins D, Quint W. 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 Aug; 124(9):1356–1363. [PubMed: 28391609]

16. Silver MI, Gage JC, Schiffman M, Fetterman B, Poitras NE, Lorey T, et al. Clinical Outcomes after Conservative Management of Cervical Intraepithelial Neoplasia Grade 2 (CIN2) in Women Ages 21-39 Years. *Cancer Prev Res (Phila)*. 2018 Mar; 11(3):165–170. [PubMed: 29437696]
17. 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 Feb 24; 58(3):e01443–19. [PubMed: 31896666]
18. Darragh TM, Colgan TJ, Cox JT, Heller DS, Henry MR, Luff RD, et al. Members of the LAST Project Work Groups. The Lower Anogenital Squamous Terminology Standardization Project for HPV-Associated Lesions: background and consensus recommendations from the College of American Pathologists and the American Society for Colposcopy and Cervical Pathology. *Arch Pathol Lab Med*. 2012; 136(10):1266–129716. [PubMed: 22742517]
19. Stoler MH, Wright TC Jr, Ferenczy A, Ranger-Moore J, Fang Q, Kapadia M, Ridder R. Routine Use of Adjunctive p16 Immunohistochemistry Improves Diagnostic Agreement of Cervical Biopsy Interpretation: Results From the CERTAIN Study. *Am J Surg Pathol*. 2018 Aug; 42(8):1001–1009. [PubMed: 29697437]
20. Szarewski A, Mesher D, Cadman L, Austin J, Ashdown-Barr L, Ho L, et al. A comparison of seven tests for high grade cervical intraepithelial neoplasia in women with abnormal smears: the Predictors 2 study. *J Clin Microbiol*. 2012 Jun; 50(6):1867–73. [PubMed: 22422852]
21. Cuzick J, Cadman L, Mesher D, Austin J, Ashdown-Barr L, Ho L, et al. Comparing the performance of six human papillomavirus tests in a screening population. *Br J Cancer*. 2013 Mar 5; 108(4):908–13. [PubMed: 23370211]
22. Cuzick J, Ahmad AS, Austin J, Cadman L, Ho L, Terry G, et al. A comparison of different human papillomavirus tests in PreservCyt versus SurePath in a referral population-PREDICTORS 4. *J Clin Virol*. 2016 Sep; 82:145–51. [PubMed: 27498250]
23. Szarewski A, Cadman L, Ashdown-Barr L, Waller J. Exploring the acceptability of two self-sampling devices for human papillomavirus testing in the cervical screening context: a qualitative study of Muslim women in London. *J Med Screen*. 2009; 16(4):193–8. [PubMed: 20054094]
24. Cadman L, Ashdown-Barr L, Waller J, Szarewski A. Attitudes Towards Cytology and Human Papillomavirus Self-Sample Collection for Cervical Screening Among Hindu Women in London, UK: A Mixed Methods Study. *J Fam Plann Reprod Health Care*. 2015 Jan; 41(1):38–47. [PubMed: 24521934]
25. Szarewski A, Ambroisine L, Cadman L, Austin J, Ho L, Terry G, et al. Comparison of Predictors for High-Grade Cervical Intraepithelial Neoplasia in Women with Abnormal Smears. *Cancer Epidemiol Biomarkers Prev*. 2008; 17(11):3033–42. [PubMed: 18957520]

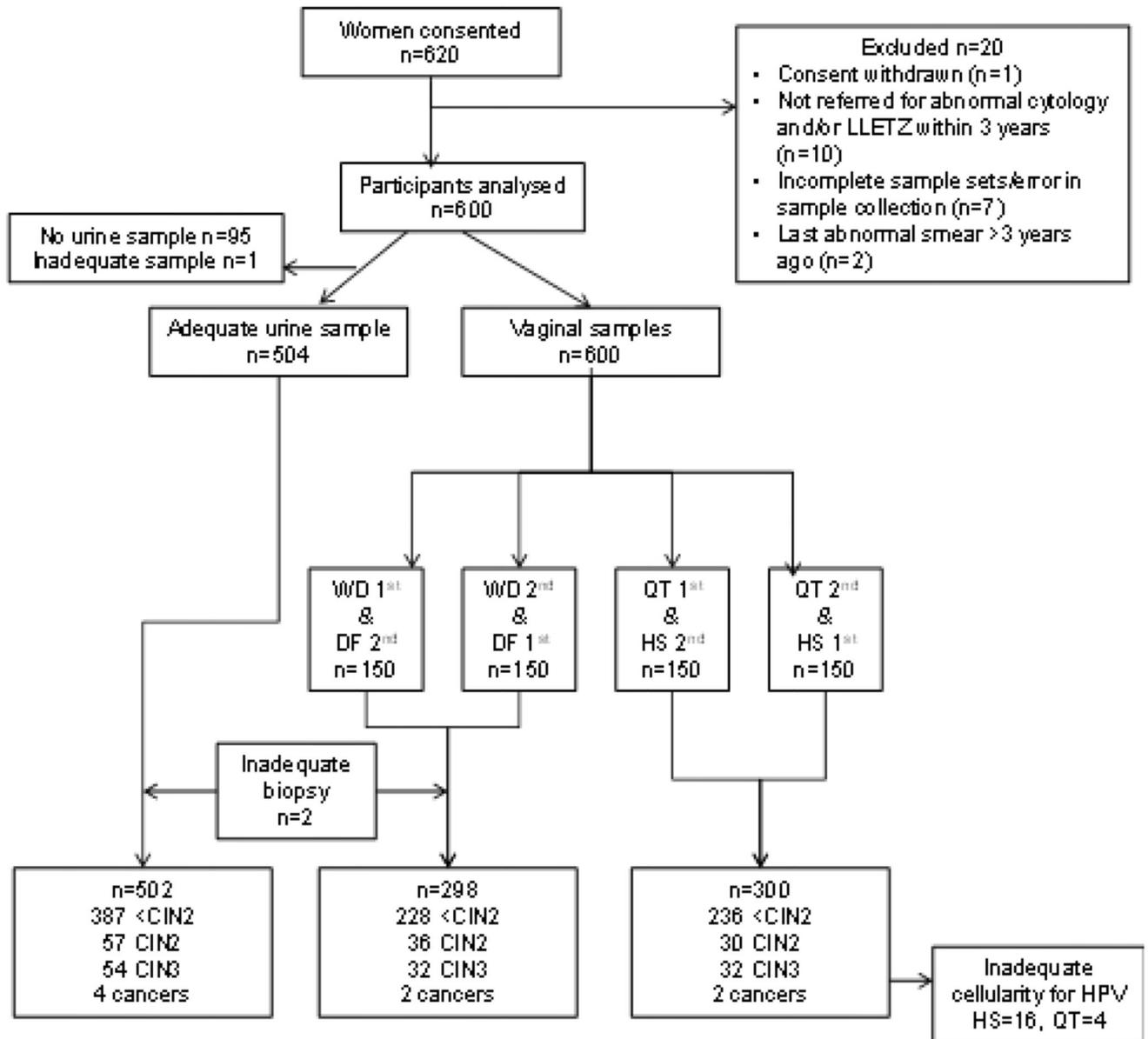


Figure 1. Study flowchart for Predictors 5.1 trial. WD: digene® Female Swab Specimen Collection Kit; DF: FLOQswab; QT: Qvintip; HS: HerSwab; 1st: first sample collected, 2nd: second sample collected

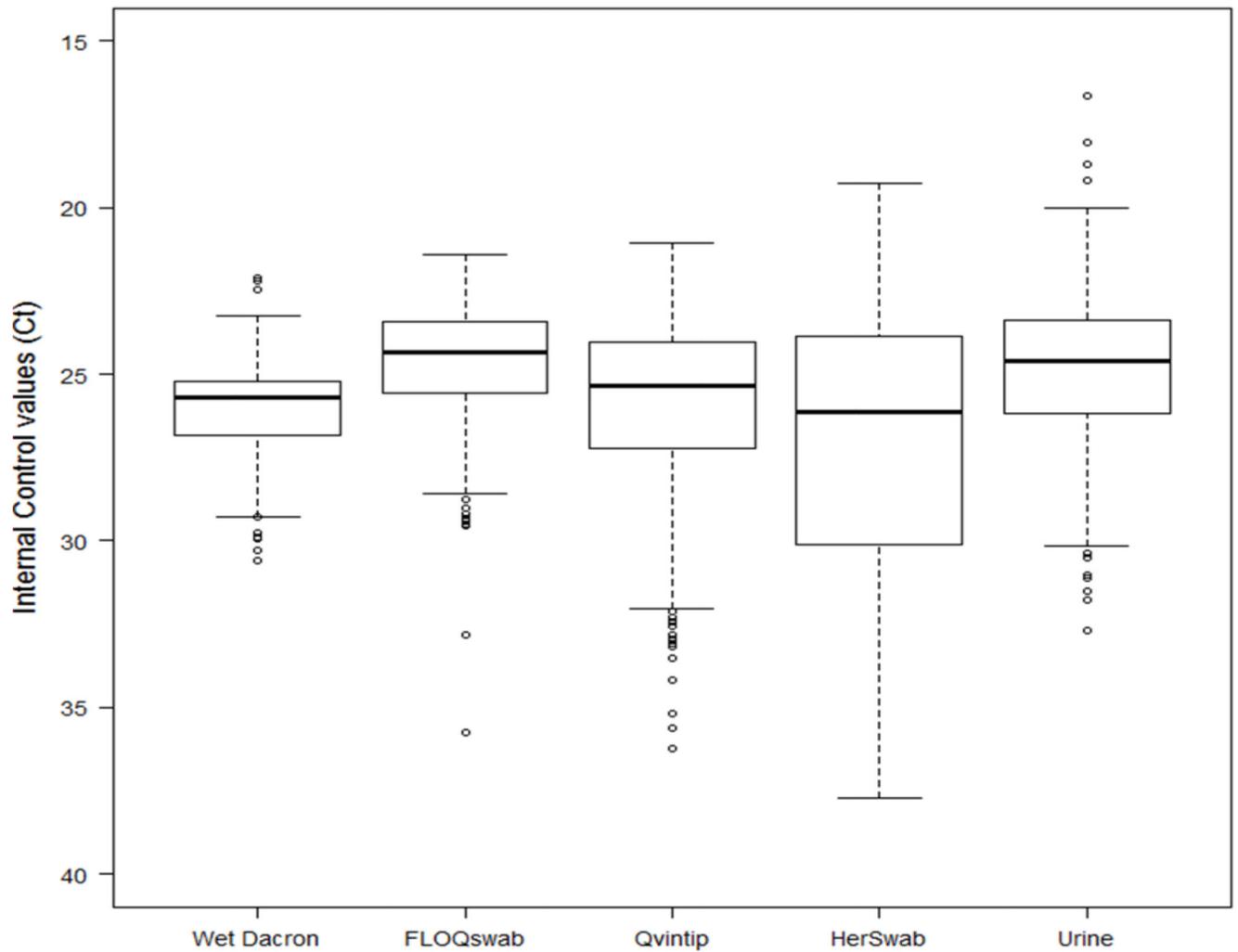


Figure 2. Cellularity of samples by device type. The cellularity is represented by the Ct values of the internal control (β -globin). A low Ct value indicates a higher cellularity and the vertical scale is ordered by decreasing Ct values.

Table 1

hrHPV positivity rates by device (combined samples). HPV positive is defined as HPV16 with Ct \leq 38.3 and/or \leq 34.2 for all other genotypes. A total of 20 samples could not be evaluated (HS = 16; QT= 4)

All samples			
Device	N	Positive (%) (95% CI)	P-value vs WD
WD	300	71.7 (66.2, 76.7)	-
DF	300	71.7 (66.2, 76.7)	1.00
QT	296	69.3 (63.7, 74.5)	0.52
HS	284	65.1 (59.3, 70.7)	0.090
Urine	504	76.0 (72.0, 79.7)	0.17

Table 2
Sensitivity and specificity analysis for hrHPV by device at various cut-offs.

A) 38.3 cut-off for HPV16 with the other HPV types at 34, B) All HPV types at 34.2 cut-off, 2 and C) All HPV types at 38.3 cut-off. There are two sets of sensitivity and specificity results given – those for the raw values (and those adjusted for cellularity in brackets except urine). Here N denotes number of samples and n the number with CIN2+ or CIN3+

A) 38.3 HPV16, 34.2Aall other genotypes	N	n	CIN2+ Sensitivity (adjusted)	n	CIN3+ Sensitivity (adjusted)	< CIN2 Specificity (adjusted)
WD	298	70	90.0 (85.7)	34	91.2 (88.2)	34.2 (32.9)
DF	298	70	88.6 (84.3)	34	88.2 (88.2)	33.8 (37.7)
QT ¹	296	63	82.5 (82.5)	33	81.8 (84.8)	34.3 (35.3)
HS ¹	284	61	82.0 (85.2)	31	77.4 (87.1)	39.5 (36.9)
Urine	502	115	87.0	58	89.7	27.4
B) 34.2 All	N	n	CIN2+ Sensitivity(adjusted)	n	CIN3+ Sensitivity(adjusted)	< CIN2 Specificity (adjusted)
WD	298	70	85.7 (84.3)	34	88.2 (85.3)	34.6 (34.2)
DF	298	70	84.3 (82.9)	34	85.3 (85.3)	34.6 (38.6)
QT ¹	296	63	77.8 (82.5)	33	72.7 (84.8)	38.2 (37.9)
HS ¹	284	61	68.9 (75.4)	31	61.3 (77.4)	41.3 (38.3)
Urine	502	115	77.4	58	84.5	29.5
C) 38.3 All	N	n	CIN2+ Sensitivity (adjusted)	n	CIN3+ Sensitivity (adjusted)	< CIN2 Specificity (adjusted)
WD	298	70	92.9 (92.9)	34	94.1 (94.1)	26.3 (26.3)
DF	298	70	92.9 (92.9)	34	94.1 (94.1)	26.8 (28.9)
QT ¹	296	63	92.1 (90.5)	33	90.9 (93.9)	24.5 (25.9)
HS ¹	284	61	90.2 (91.8)	31	90.3 (93.5)	30.5 (32.9)
Urine	502	115	93.9 (58	93.1	19.6 (

¹Note that for the adjusted sensitivity and specificity for CIN2+ and <CIN2, there is one less record for both Qvintip and Herswab (N=295 and 283, respectively). This is due to missing internal control values.

Table 3

Device preference by randomisation group (row percentages), amongst 454/504 women with adequate urine sample who provided preference information, ignoring sample order. Results are split according to if woman was in the WD/DF group or QT/HS group

Randomisation Group	N	Device preference			
		WD	DF	Urine	No preference
WD/DF	222	22 (9.9%)	55 (24.8%)	80 (36.0%)	65 (29.3%)
		QT	HS	Urine	No Preference
QT/HS	232	61 (26.3%)	70 (30.2%)	65 (28.0)	36 (15.5)
		Either vaginal device			
Total	454	208 (45.8)		145 (31.9)	101 (22.2)

Table 4
A) Women's reported ease of use of five collection devices, and B) their confidence of correctly using devices, amongst 504 women with adequate urine sample, ignoring sample order of vaginal device use.

A Ease of using device								
Device	N	Very easy	Quite easy	Neither easy nor difficult	Quite difficult	Very difficult	Excluded	Overall Total
WD	247	128 (51.8%)	91 (36.8%)	19 (7.7%)	9 (3.6%)	0(0.0%)	2	249
DF	247	155 (62.8%)	71 (28.7%)	12 (4.9%)	8 (3.2%)	1 (0.4%)	2	249
QT	254	119 (46.9%)	87 (34.3%)	35 (13.8%)	13 (5.1%)	0 (0.0%)	1	255
HS	254	108 (42.5%)	89 (35.0%)	25 (9.8%)	29 (11.4%)	3 (1.2%)	1	255
Urine	501	304 (60.7%)	146 (29.1%)	29 (5.8%)	17 (3.4%)	5 (1.0%)	3	504

B) Confidence in correctly using devices						
	N	Very confident	Fairly confident	Not confident	Excluded	Overall Total
WD	247	97 (39.3%)	132 (53.4%)	18 (7.3%)	2	249
DF	247	109 (44.1%)	113 (45.7%)	25 (10.1%)	2	249
QT	254	88 (34.6%)	146 (57.5%)	20 (7.9%)	1	255
HS	253	92 (36.4%)	125 (49.4%)	36 (14.2%)	2	255
Urine	500	308 (61.6%)	170 (34.0%)	22 (4.4%)	4	504