Abstract

Cervical cancer screening has been one of the most impactful human interventions in medical history, saving the lives of countless thousands of women since the introduction of organized cytology screening programs. Today, we stand at a crossroads in the fight against cervical cancer, with several countries actively engaged in introducing primary human papillomavirus (HPV) testing and vaccination as more effective means of prevention. This chapter discusses the history of organized screening and how this led to HPV test methods to detect cervical cancer. We go on to examine the technologies used to screen for high-risk HPV types and how they affect clinical performance. We examine the evidence for primary HPV screening and review recent self-collection initiatives to reach underserved women, including the use of urine as novel sample type. In addition, we critically examine the evolution of HPV test methods and make the case for the use of extended genotyping as an improved risk stratification tool for guiding clinical management. Finally, we look to the future of cervical cancer screening and consider options for future management programs.

Keywords: genotyping, HPV screening tests/strategies, Pap, risk stratification, self-sampling
1. Introduction

Cervical cancer screening has advanced considerably since the introduction of the Pap smear in the 1960s: organized cytology screening programs have successfully reduced the burden of disease associated with human papillomavirus (HPV) and sensitive new molecular methods have been developed to detect the virus. Perhaps more importantly, safe and efficacious HPV vaccines are now widely available and offer the prospect of greatly reducing the incidence of cervical cancer if sufficient numbers of the target population can be vaccinated. Vaccines offer the best hope for developing countries, which lack the resources to implement an effective screening program. Despite the introduction of vaccination, there is still a need for improved screening modalities as both screening and vaccination will necessarily coexist for some decades to come. It is against this backdrop that we briefly review the history of the Pap smear and the development of organized cytology screening programs and then go on to discuss the development of molecular methods. We discuss the pros and cons of the different molecular assay design approaches and review the case for primary HPV screening. Despite the effective tools at our disposal, reaching underserved women remains the biggest challenge to controlling cervical disease. However, new self-sampling methods offer the prospect of an effective outreach program to reach women most in need. Finally, we discuss the benefits of extended genotyping and how this might influence future screening algorithms. We conclude that the biggest hurdle to preventing and detecting cervical disease may lie in our inability to adapt to change and effectively implement new strategies. This history of cervical cancer screening suggests that the pace of change is slow but we must respond more quickly to the global threat posed by cervical cancer and make efficient use of the tools at our disposal.

2. The history of organized screening programs

George Nicholas Papanicolaou is almost single-handedly credited with developing cervical cancer screening. He was encouraged to leave his native Greece by the renowned geneticist, Thomas Hunt Morgan, who helped him gain a position in the Department of Pathology and Bacteriology at New York Hospital as an assistant, and later at the Department of Anatomy at Cornell Medical School. Dr. Papanicolaou began to study vaginal secretions in women, beginning with a case study of his wife, Mary [1]. Later work led to the observation of cancer cells while studying smears of women in the New York Hospital, which led him to propose that a systematic study of smears could lead to early cancer detection. His “New Cancer Diagnosis” theory was initially very poorly received by his peers [2]. It was not until over 10 years later that a successful collaboration with Herbert Traut (1894–1963), a renowned gynecologist and pathologist at Cornell University, resulted in a more widely received manuscript “The diagnostic value of vaginal smears in carcinoma of the uterus.” This work was presented to the New York Gynecological Society in 1941 and published in the prestigious American Journal of Obstetrics and Gynecology [3]. In 1947, Dr. Papanicolaou and the head of the Anatomy Department, Dr. Joseph Hinsey, began focusing on training other physicians on the new technique. This led to a gradual adoption of the screening method, and by 1960 it was
being adopted across the United States. However, it was not until just before Papanicolaou’s death that the impact of his discovery was becoming apparent and he finally began to get the recognition he deserved. Today, he is remembered as “the father of exfoliative cytology” and cervical cytology stands as one of the most successful clinical interventions of the twentieth century. Organized cervical cancer screening has led to an over two-fold reduction in the incidence of cervical cancer in the United States in the period 1975–2012. This success in the United States has been replicated in other countries with organized screening programs such as the United Kingdom, which has seen a similar reduction in mortality from 1985 to 2012. However, over the last decade it has become evident that cytology screening may have reached the limits of its effectiveness in terms of reducing the incidence rate, with both the US and UK trend lines reaching a plateau (Figure 1).

3. The development of molecular screening tools

The development of molecular biology techniques in the late 1970s and early 1980s paved the way for the discovery that the HPV was both necessary and responsible for cervical disease. zur Hausen played a pivotal role in establishing the link between papillomaviruses and cervical cancer [4]. As interest intensified, novel HPV types began to be isolated from genital warts (HPV6 [5–7] and laryngeal papillomas (HPV11) [8]). The detection of additional type-specific viral sequences followed: [9, 10], HPV16 [11], HPV18 [12–14], and high-risk types HPV31 and HPV45, which were first described by Lorincz et al. [15] in the United States [16]. These discoveries laid the foundation for the development of cervical cancer diagnostic tests and the development of the first HPV cancer vaccines. Harald zur Hausen went on to win the Nobel Prize in Physiology or Medicine in 2008 for this work on the association of papillomaviruses with human cancers.

While the work of Harald zur Hausen and others set the stage for modern cervical cancer diagnostics, it took another wave of development to advance HPV detection to the point that it could be widely adopted in the clinical laboratory. The early work led to rapid development of basic laboratory tests to detect various HPV types. However, the techniques employed involved the use of 32P-labeled probes that were hybridized using Southern blotting or slot blot methods, which meant they were labor intensive. These early methods (ViraPap and ViraType) were commercialized by Life Technologies Inc. and later sold to the Digene Corporation but they did not achieve widespread adoption, primarily because they did not detect all oncogenic HPV types and also lacked sensitivity. This period also saw the development of in situ hybridization (ISH) techniques by Enzo Diagnostics and Life Technologies Inc., and additional methods were later automated by Ventana Medical Systems (now a Roche Company), the Dako Corporation, and others [17]. These tests also suffered from lack of sensitivity [18] and were ultimately not adopted as screening tools. However, ISH methods may prove to be useful in assisting in the diagnosis of Cervical Intraepithelial Neoplasia cases (CIN) [19].

4. The digene hybrid capture 2 assay

The Digene Corporation (now part of QIAGEN) under the scientific leadership of another early HPV molecular pioneer, Dr. Attila Lorincz [15, 16], developed the modern market for HPV cervical cancer screening. Lorincz et al. developed a technology called hybrid capture that fused traditional immunoassay techniques with the newer nucleic acid isolation technology [20]. The core principle leveraged the fact that DNA–RNA hybrids have a distinct shape different than a DNA helix that is detectable using antibodies. By generating in vitro RNA transcripts of each of the target HPV types and pooling them into a single probe cocktail, the method was capable of detecting as many as 13 high-risk HPV types at once. The capture antibody is bound to a solid surface that binds cognate target hybrids which in turn are then detected using a secondary antibody coupled with alkaline phosphatase using a chemiluminescent substrate. The first generation hybrid capture assay utilized two mixtures of single-stranded RNA probes: one for low-risk HPV types 6, 11, 42, 43, and 44, and one for high-risk
HPV types 16, 18, 31, 33, 35, 45, 51, 52, and 56. The analytic sensitivity of the first generation test was estimated at 50,000 copies of HPV16 DNA [21, 22], and this was not found to be sensitive enough versus polymerase chain reaction (PCR) and histology [23]. Digene researchers then launched a second generation assay (Hybrid Capture 2) whose analytical sensitivity was increased to 1000 HPV DNA copies by the reformulation of hybridization reagents and by the addition of new probes for four high-risk HPV types, 39, 58, 59, and 68. This assay was launched in Europe in 1998 and was subsequently approved by the US Food and Drug Administration (FDA) in 2003, for the triage of Atypical Squamous Cells of Undetermined Significance (ASC-US) cytology patients and as an adjunctive test with cytology for cervical cancer screening [23]. The assay went on to achieve worldwide acceptance and became the de facto clinical standard for all subsequent assay development. It has the largest body of clinical evidence supporting its use, and it has been used in numerous clinical trials across the globe [24, 25].

5. PCR- and other amplification-based technologies

The development of PCR technology [26, 27] meant that small quantities of DNA or RNA could now be amplified and detected either directly in endpoint PCR [28] or through the use of accumulating fluorescence that could be monitored continuously in real-time (RT) PCR [29–31]. As with most new techniques, refinements and improvements overcame some of the early difficulties such as assay robustness and the potential for contamination/false positives [32, 33]. The use of PCR to detect HPVs mirrored that of other methods, with HPV initially being detected using gel analysis of PCR products [34] and subsequently using real-time methods [35, 36]. Type-specific primer amplification methods eventually gave way to consensus PCR approaches where multiple HPV types could be detected in one reaction [37]. Most PCR approaches targeted the L1 gene region, and included the MY09–MY11 (MY09/MY11) and GP51–GP61 (GP51/GP61) primer systems. The MY primers were later redesigned to improve performance (PGMY09/PGMY11 [38]) as were the GP primers, to GP5+/GP6+, which were subsequently widely adopted [39, 40]. PCR-based tests are intrinsically more analytically sensitive than Hybrid Capture 2 technology, detecting as little as 10–100 copies of target. However, the Hybrid Capture 2 assay established that approximately 1 pg/ml or 5000 copies of HPV16 target DNA per reaction equated to an actionable clinical result (CIN2+ histology) [41]. Therefore, assays that are clinically more sensitive than this would have poor specificity, referring low level infections unnecessarily to follow colposcopy and possible biopsy or treatment. This has been the subject of discussion in cervical cancer screening since the introduction of highly sensitive molecular methods. Key opinion leaders have cautioned against the use of analytically validated assays whose clinical performance has not been well established in large longitudinal studies, emphasizing the need to strike the proper balance between sensitivity and specificity [42]. In one case, even an FDA approved test was considered clinically unacceptable due to its high HPV positivity [43]. The issue is compounded by the fact that even large reference laboratories and smaller commercial companies cannot afford to fund large studies with a minimum of 3-year follow up of enrolled patients. This led a group
of international HPV experts to propose acceptance criteria for HPV tests that wish to be used in primary HPV screening [44]. These criteria, commonly known as “Meijer Criteria” after the lead author, were founded on the principle that “candidate high-risk HPV tests to be used for screening should reach an optimal balance between clinical sensitivity and specificity for detection of high-grade CIN and cervical cancer to minimize redundant or excessive follow-up procedures for high-risk HPV positive women without cervical lesions” and set forth the following guidelines for an acceptable HPV screening test:

1. The sensitivity of the candidate test for ≥CIN2 should be at least 90% of the sensitivity of the HC2 (i.e., relative sensitivity of at least 90%) as assessed by a noninferiority score test.

2. The specificity of the candidate test for ≥CIN2 should be at least 98% of the specificity of HC2.

3. The intralaboratory reproducibility in time and interlaboratory agreement should be determined by evaluation of at least 500 samples, 30% of which tested positive in a reference laboratory using a clinically validated assay. This should result in a percentage of agreement with a lower confidence bound not less than 87% (kappa value of at least 0.5 in this series of samples including 30% positives).

Meijer criteria have been well accepted as a critical litmus test in the absence of large-scale longitudinal trial data and have become a clinical benchmark for validation of tests, especially in the European Union. Despite the use of this leaner approach where the disease samples can be sourced retrospectively, only a small number of existing commercially available assays have actually met the criteria. A 2015 review of currently available commercial assays “identified 193 distinct commercial HPV tests” [45]. However, Arbyn et al. [46], also reported in the same year that only five commercially available tests fully met the Meijer criteria and could be considered suitable for use in primary HPV screening: PapilloCheck® HPV-Screening test; Abbott RealTime hrHPV test; cobas® 4800 HPV test; BD Onclarity™ HPV assay; HPV-Risk assay and the Aptima assay, targeting E6/E7 mRNA. The authors stated that the Cervista® assay should also be added to the list, despite one report of a lack of noninferiority for specificity (see also reference [104]). Three other assays met the sensitivity/specificity criteria but did not disclose accuracy and reproducibility data and were thus considered to partially meet the criteria (an in-house quantitative RT-PCR targeting E6/E7DNA sequences, a GP5+/GP6+ PCR with Luminex identification of high-risk types and a MALDITOF assay). The authors also concluded that the Aptima assay, while fully meeting the criteria, needed further longitudinal validation of its long-term negative predictive value (NPV) because it was an RNA-based assay and to date this had only been established for DNA assays [46].

E6 and E7 gene RNA-based assays have been extensively tested in cervical cancer screening. The impetus for using RNA- versus a DNA-based targets, is likely based on the observation that both the E6 and E7 genes encode oncogenes that are involved in the development of cancer and are upregulated as disease progresses [47]. E6 and E7 viral oncoproteins bind and modulate cellular gene products (p53 and pRb) that play a key role in cell cycle control and DNA repair. The resulting genomic instability caused by E6 and E7 oncoproteins is a necessary condition for cell transformation and immortalization [48, 49]. Thus, it is reasonable to
postulate that an E6/E7 RNA target might offer an advantage over DNA in that it should be overexpressed in high- versus low-grade disease and that it might offer both a sensitivity and specificity advantage. There are two commercially based E6/E7 RNA assays: the Proofer Assay (Norchip, Klokkarstua, Norway) is a real-time multiplex nucleic acid sequence-based amplification assay (NASBA) for isothermal amplification and detection of E6/E7 mRNA from five high-risk oncogenic types, HPV16, 18, 31, 33, and 45, using molecular beacon probes, and the Aptima® HPV Assay (Hologic—GenProbe) that is a qualitative nucleic acid amplification test that detects HPV E6/E7 mRNA from 14 high-risk HPV types [50]. The Aptima assay uses target amplification using transcription-mediated amplification (TMA) [51] and detection of the amplification products (amplicon) by the hybridization protection assay (HPA) [52]. HPV mRNA is captured on magnetic particles and then amplified using TMA that is a transcription-based nucleic acid amplification method that utilizes two enzymes, Moloney Murine Leukemia Virus (MMLV) reverse transcriptase and T7 RNA polymerase. The reverse transcriptase is used to generate a DNA copy of the target mRNA sequence containing a promoter sequence for T7 RNA polymerase. T7 RNA polymerase produces multiple copies of RNA amplicon from the DNA copy template. Detection of the amplicon is achieved by HPA using single-stranded nucleic acid probes with chemiluminescent labels that are complementary to the amplicon. The labeled nucleic acid probes hybridize specifically to the amplicon. The selection reagent differentiates between hybridized and unhybridized probes by inactivating the label on the unhybridized probes. During the detection step, light emitted from the labeled RNA–DNA hybrids is measured as photon signals called relative light units (RLU) in a luminometer. Final assay results are interpreted based on the analyte signal-to-cutoff (S/CO) [53]. The Norchip Proofer assay has been shown to be substantially less sensitive but more specific when compared to other clinically validated assays, including Roche cobas and Hybrid Capture 2 [54–56]. Arbyn et al. [57] performed a meta-analysis of the performance of the Aptima HPV assay versus Hybrid Capture 2 and found that for both triage of ASC-US and Low-Grade Squamous Intraepithelial Lesion (LSIL), Aptima is as sensitive but more specific than HC2 for detecting cervical precancer. In head to head tests, Aptima had a substantially higher sensitivity than Proofer in both a screening and a referral population, but Proofer showed improved specificity [55, 56]. The Aptima test has broader clinical application and relevance because it detects all 14 high-risk types versus just five high-risk types in the Proofer assay. However, as noted above, it needs further confirmation of its long-term NPV over at least a 5-year period, as recommended by Arbyn et al. [46]. A recent report with a 3-year longitudinal follow up found that the NPV is similar to that of Hybrid Capture 2 and again that its specificity was significantly better (96.3% compared with HC2 specificity of 94.8%; \( P < 0.001 \)) [58]. A comprehensive review of the clinical performance of the assay versus HC2 also concluded that the NPV was sufficient to justify a 3-year interval and that the specificity of the assay was consistently higher [59]. The increase in specificity is similar to that previously reported [57] and is relatively modest. It could also result from a slight clinical cut-point bias toward the specificity axis of the receiver operating characteristic (ROC) plot of sensitivity versus 1–specificity for this assay. Whatever the driver, it seems clear that there is no significant difference in sensitivity when comparing RNA- and DNA-based assays and the specificity differences are not as high as one might have predicted from the upregulation of E6/E7 RNA.
during oncogenesis. The latter may reflect that one does not see the full potential of upregulation when screening for less severe cellular abnormalities versus cancer. The answer may also lie in the observation that neither assay exclusively targets RNA: both assays have been reported to detect cognate DNA sequences present in endocervical specimens at levels where one might expect the DNA signal alone to be sufficient to record a positive clinical result [60, 61]. Thus, it is conceivable that this could render otherwise RNA-negative specimens positive, reducing the specificity of the assay.

6. How target regions impact assay performance

6.1. Which target region to choose?

HPV is double-stranded DNA virus whose circular genome is approximately 8000 base pairs long. It encodes eight open reading frames (ORFs) that are divided into early and late genes (Figure 2) involved in replication (i.e., E1 and E2) and packaging (i.e., L1 and L2) with the remaining genes (E6, E7, E5, and E4), playing roles in driving cell cycle entry, immune evasion, and virus release (reviewed in [49]). Papillomaviruses are ancient in origin, believed to have arisen in reptiles approximately 350 million years ago. They have evolved in their various host lineages (including humans) over the millennia and they have relatively stable genomes given their (redundant) double-stranded DNA structure. The rate of nucleotide substitution for HPV18 has been estimated at ~4.5 × 10⁻⁷ subs/site/year [62]. This means that assay developers can look across the entire genome for conserved target regions in which to design gene probes.

![Figure 2. Physical map of the HPV16 genome.](image-url)
Given the aforementioned large number of commercial assays available on the market, not surprisingly most ORFs have been targeted by one or more assays, including L1 [63], E1 [64], and E6/E7 [65, 66]. However, most commercial assays to date have been developed using the L1 gene as a target region [45, 63], including the Roche cobas assay that has received FDA approval [67].

The question one must then ask is: does it matter which region of the genome you use—are all genomic regions created equal? While a detailed phylogenetic and evolutionary analysis (beyond the scope of the current work) would be required to fully address this question, the following sections discuss some important topics with respect to target analyte performance in clinical diagnostics and how this can influence clinical results.

6.2. Cross-reactivity with nontarget HPV types

There is general agreement that only high-risk types cause cervical cancer and that low-risk types should not be screened for as part of routine cervical cancer prevention [68]. Previously, there was consensus that 14 high-risk types caused the majority of cancers. However, more recently the evidence for HPV66 in cervical cancers was considered too weak to keep it in the group of 14 high-risk types and it was recommended to remove it [69]. However, all FDA tests approved to date detect HPV66 as part of their 14 high-risk panel (Roche cobas, Hologic-Genprobe Aptima, and Hologic Cervista) and the Hybrid Capture 2 assay detects it via cross-reactivity [70]. In addition, a number of possible/probably carcinogenic types (including HPV66) have been shown to have oncogenic potential and “are biologically active and affect the same cellular pathways as any of the fully recognized carcinogenic HR-HPV types” [71]. Thus, it is difficult to draw an absolute line between known carcinogenic types and those that have that potential but very infrequently result in cancers [69].

Whether you take the view that there are 14 high-risk types or 13 high-risk types (omitting HPV66), there is little doubt that screening for additional HPV types will not increase sensitivity for cancer detection and will reduce specificity. In addition, it may do harm to patients both psychologically and potentially physically if they are treated for lesions that will not result in cancer [68, 72]. Thus, it is important that clinically validated assays perform inclusivity and exclusivity studies to ensure that the assay detects only the intended HPV types for which they have claims. They Hybrid Capture 2 test has been shown to have excellent clinical sensitivity for CIN2+ endpoints and as described earlier, has been used as a clinical benchmark for the last decade. However, it does have a well-documented cross-reactivity with non-high-risk HPV types [70, 73, 74]. This results from the fact that the assay detects the entire HPV genome (Figure 2) and thus there is more potential for closely related nononcogenic sequences to be detected. One study found that “some 20% of HC2-positive samples did not contain the targeted HPV types. About two-thirds of them resulted from cross-hybridization, especially with HPV53, HPV66, and HPV70” [75]. When split sample testing is performed with HC2 and a second assay, one typically sees a larger proportion of HC2+/other assay–results which, if they are resolved by third-party genotyping or sequencing method, shows both low-risk and no HPV present results, confirming the negative result as truth with respect to high-risk HPV [66, 76]. In another recent study (n = 6172) comparing Roche cobas assay with Hybrid Capture
2, where discordants were resolved using the Roche Linear Array genotyping assay, the authors reported that “HC2+/COBAS− were less likely to contain hrHPV genotypes (12.3 versus 68.9%; \( P < 0.0001 \)) and more likely to contain only lrHPV genotypes (52.8 versus 12.1%; \( P < 0.0001 \)) than those HC2−/COBAS+” and they found “lower CIN2+ rates among women with HC2+/COBAS− results” [77].

The clinical impact of this cross-reactivity has not been widely discussed in the literature. Castle et al. used a more stringent dual reference assay method to determine the level of HC2 cross-reactivity and concluded that about 8% of all HPV positives were due to cross-reactivity with non-high-risk HPV types in the referred (ALTS Study) population. They concluded that cross-reactivity would likely be further reduced in a screening population where there is less overall infection. However, they cautioned that these non-high-risk HPV “cases of CIN2 might be treated by excisional procedures, which can cause iatrogenic morbidity and adversely affect reproductive outcomes” [70]. The previously cited study of Gillio-Tos et al. [75] concluded “If only the samples containing HC2-targeted types tested positive, the positive predictive value (PPV) would have increased from 7.0% (95% CI, 6.1–8.0%) to 8.4% (95% CI, 7.3–9.6%), although 4.9% (95% CI, 2.4–8.8%) of cervical intraepithelial neoplasia grade 2+ (CIN2+) cases would have been missed”. Regarding the latter point, “missing” CIN2+ disease caused by non-high-risk HPV types should be the goal of future assays since it has an extremely low risk of resulting in cancer. In summary, there is an 8–20% false-positivity associated with the Hybrid Capture 2 assay that unfortunately is correlated with bona fide abnormal cytology and histology (CIN2+). However, it is important to recognize that since these abnormalities are not caused by high-risk HPV types, they have a very low cancer risk and should be managed accordingly. While this phenomenon is well documented, it tends to be largely ignored and considered noise in the overall clinical performance of an assay. It does however have implications for benchmarking studies and in our view should also be taken into account in any future revisions of the Meijer criteria, since it has a direct impact on the maximum sensitivity and specificity of any assay which only detects high-risk HPV.

### 6.3. L1 consensus-based versus gene-specific PCR approaches

All molecular assays require some sort of amplification technology in order for them to have the required sensitivity to detect small quantities of target nucleic acid. This can take the form of signal amplification such as that described for the Hybrid Capture 2 assay (where the signal rather than the target is amplified using a chemiluminescent substrate), or in the case of the Hologic Cervista assay, where signal amplification is achieved through the use of a unique isothermal Invader® chemistry that leverages a reusable universal flap and fluorescence energy transfer (FRET) to amplify the signal [78, 79] or it can use RNA amplification (Proofer and Aptima assays described earlier) or PCR that accounts for the majority of HPV assays on the market [45]. Traditional PCR methods where the amplicon is detected in a secondary process such as on an agarose gel or via hybridization on strips have gradually been replaced by real-time detection methods for clinical use, although the Roche Linear Array and InnoLIPA line blot assays have remained popular where researchers are interested in determining which genotypes are present in clinical samples or for discordant result anal-
ysis [76, 80, 81]. With PCR one can choose to selectively amplify each viral gene target individually (gene-specific PCR) or utilize a consensus (broad spectrum) primer approach. The latter typically uses degenerate primer sequences to detect conserved protein coding regions of multiple HPV types [82]. This has the advantage that multiple (e.g., all high-risk) HPV types can be detected in a single reaction but it means that you can only record a pooled type result, unless you use a secondary method such as a line blot or bead approach to subsequently identify the specific genotypes present [83–85]. Gene-specific detection approaches offer the possibility of direct individual genotype detection but require the use of multiple reactions [86] or the combining of one or more genotypes in different fluorescent channels using real-time PCR technology [66, 87]. Table 1 lists the major approaches to HPV DNA detection and details some of their advantages and disadvantages. Head to head testing suggests that the top performing assays have broadly similar clinical sensitivity and specificity [55, 56, 88, 89] so clinical laboratories have a choice in the assay they use. The decision on which assay to adopt may be made based on the needs of a particular laboratory. However, it is instructive to consider the following assay characteristics when choosing an HPV test:

<table>
<thead>
<tr>
<th>Assay name</th>
<th>HPV target region</th>
<th>Nucleic acid type</th>
<th>Consensus or gene-specific coverage</th>
<th>High-risk type coverage</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>QIAGEN Hybrid Capture 2</td>
<td>Entire genome</td>
<td>DNA (signal amplification)</td>
<td>Not applicable</td>
<td>13</td>
<td>Single reaction, Clinically validated CE marked/FDA approved</td>
<td>No genotyping or internal control. High level of nontarget cross-reactivity</td>
<td>[41, 70, 75]</td>
</tr>
<tr>
<td>PreTect HPV-Proofer E6/E7</td>
<td>RNA</td>
<td>Gene-specific</td>
<td>5</td>
<td></td>
<td>High specificity CE marked</td>
<td>Low sensitivity, Does not detect all cancer-causing types [55, 56]</td>
<td></td>
</tr>
<tr>
<td>Hologic Aptima (screening) E6/E7</td>
<td>RNA</td>
<td>Gene-specific</td>
<td>14</td>
<td></td>
<td>Single reaction, Clinically validated CE marked/FDA approved</td>
<td>No genotyping, Sample processing control only [53, 90]</td>
<td></td>
</tr>
<tr>
<td>Hologic Aptima (genotyping) E6/E7</td>
<td>RNA</td>
<td>Gene-specific</td>
<td>3—HPV16, HPV18+45</td>
<td></td>
<td>Single reaction, Clinically validated CE marked/FDA approved</td>
<td>Additional test required for limited genotyping Sample processing control only [91]</td>
<td></td>
</tr>
<tr>
<td>Assay name</td>
<td>HPV target region</td>
<td>Nucleic acid type</td>
<td>Consensus or gene-specific coverage</td>
<td>Advantages</td>
<td>Disadvantages</td>
<td>References</td>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>-------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>Hologic Cervista (screening)</td>
<td>E6/E7/L DNA</td>
<td>Gene-specific</td>
<td>14</td>
<td>Single reaction</td>
<td>Clinically validated.</td>
<td>[43, 78]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Internal control</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CE marked/FDA approved</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No genotyping</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>High positivity rate (clinical specificity has been questioned)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hologic Cervista (genotyping)</td>
<td>E6/E7/L DNA</td>
<td>Gene-specific</td>
<td>2—HPV16, HPV18</td>
<td>Single reaction</td>
<td>Clinically validated.</td>
<td>[79]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Internal control</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CE marked/FDA approved</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Additional test required for limited genotyping</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Limited genotyping</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abbott Real-time high-risk HPV</td>
<td>L1 DNA</td>
<td>Consensus</td>
<td>14</td>
<td>Single reaction that includes HPV16/18 genotyping. Internal control</td>
<td>Literat-positive</td>
<td>[92, 93]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Semiautomated, CE marked</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Additional test required</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Limited genotyping</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greiner Papillocheck</td>
<td>E1 DNA</td>
<td>Consensus</td>
<td>24— includes 14 CE marked</td>
<td>Full genotyping</td>
<td>Literat-intensive workflow requiring three separate rooms</td>
<td>[64, 94]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>includes 14 CE marked</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Or 14 high-risk only</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BYC Laboratories (formerly RIATOL) E6/E7 real-time PCR test</td>
<td>E6/E7 DNA</td>
<td>Gene-specific</td>
<td>14 + 3 moderate risk types</td>
<td>Full genotyping</td>
<td>Partially validated per Meijer criteria</td>
<td>[86]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Semiautomated</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17 individual PCR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>reactions plus controls required</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Laboratory developed test</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Three-well assay design reduces throughput</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BD Onclarity HPV Assay</td>
<td>E6/E7 DNA</td>
<td>Gene-specific</td>
<td>14</td>
<td>Extended genotyping</td>
<td>HPV16, 18, 45, 51, 52</td>
<td>[66, 95, 96]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Clinically validated</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Fully automated</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CE marked</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Limited genotyping capability, Meijer criteria not yet established</td>
<td>[87]</td>
<td></td>
</tr>
<tr>
<td>Cepheid Xpert® HPV</td>
<td>E6/E7 DNA</td>
<td>Gene-specific</td>
<td>14</td>
<td>Single reaction that includes HPV16/18 genotyping. Internal control</td>
<td>Literat-positive</td>
<td>[87]</td>
<td></td>
</tr>
<tr>
<td>Assay name</td>
<td>HPV target region</td>
<td>Nucleic acid type</td>
<td>Consensus or gene-specific coverage</td>
<td>High-risk type coverage</td>
<td>Advantages</td>
<td>Disadvantages</td>
<td>References</td>
</tr>
<tr>
<td>------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------------------------</td>
<td>------------------------</td>
<td>------------</td>
<td>---------------</td>
<td>------------</td>
</tr>
<tr>
<td>Roche cobas assay</td>
<td>L1</td>
<td>DNA</td>
<td>Consensus</td>
<td>14</td>
<td>Single reaction that includes HPV16/HPV18 genotyping, Internal control, Clinically validated, CE marked and FDA approved (cotesting and primary screening)</td>
<td>Limited genotyping capability</td>
<td>[72, 97]</td>
</tr>
</tbody>
</table>

Table 1. HPV detection methods.

1. Is the assay and required equipment cost competitive?
2. What level of automation is involved—manual processing, semiautomated, or a fully integrated workflow?
3. What is the time to result and the hands on time of laboratory personnel to run the assay?
4. Is the assay part of a complete offering or does the user have to procure third party equipment or reagents to complement the workflow?
5. Does the assay target only high-risk HPV types?
6. Does the assay provide the required level of genotyping information to assist with informed patient management and risk stratification?
7. Has the assay demonstrated the required ability to accurately diagnose both single and mixed infections?
8. Has the clinical performance of the assay been demonstrated in the intended target population using the specified collection devices and preservative media (sensitivity/specificity/longitudinal NPV and PPV)?
9. Have the required regulatory approvals and quality requirements for the product been fully met?
10. Has the clinical cutoff of the assay been tested in several different patient populations with different risk profiles to ensure that patients are neither being over- or underreferred for biopsy and treatment?

The first decision one may want to make is whether to choose an assay that targets RNA or DNA. Clinical performance is very similar so it may come down to a practical decision on workflow and laboratory suitability. At the current time, the long-term NPV of DNA-based assays is well established and enshrined in consensus guidelines whereas the evidence for
RNA assays is still accumulating [58, 59, 98]. DNA assays also have the advantage of improved target stability which imposes stricter adherence to laboratory cleaning methods to avoid degrading the more labile RNA targets [53]. As mentioned above, real-time PCR approaches now predominate in the market so one has the choice of either a consensus-based primer design or a gene-specific primer approach (Table 1). Consensus approaches yield a pooled high-risk result with HPV16/HPV18 genotype identification from a single well whereas gene-specific approaches can offer more extended genotyping information which can limit throughput. The clinical benefits of extended genotyping will be discussed later in this chapter but here we will focus on the analytical performance of the two assay design approaches.

6.4. The link between analytical and clinical assay performance

Good analytical assay performance is a requirement but not a guarantee of clinical assay performance. The viral load that correlates well with prediction of CIN2+ risk is approximately 5000 copies per reaction as established by the Hybrid Capture 2 test [41]. Thus, it is important to determine the clinical cutoff (a level of positivity in the assay that best correlates with histologically confirmed disease) for each individual assay using standard receiver operating characteristic plots of sensitivity versus 1-specificity so as to provide the highest possible sensitivity and an optimal specificity [55, 56, 88]. Assays that are analytically too sensitive will refer an unacceptably high number of women to colposcopy for possible biopsy and treatment [43]. Table 2 describes some of the unique diagnostic challenges associated with cervical cancer screening that make this task a complicated one. This is reflected in the reports in the literature on viral load measurements, which are mixed at best. While HPV16 viral load has been shown to correlate with disease progression [99, 100], non-HPV16 types have been reported to show less correlation [101]. A recent well controlled study that focused exclusively on single infections found that HPV16/HPV18/HPV31/HPV45 viral load was correlated with abnormal cytology. This study is likely to have benefited by excluding the complication of mixed infections and the use of a high quality E6 gene-specific real-time PCR assay [102]. A positive correlation for these types (and HPV33) with disease severity was also confirmed in an independent study with histological endpoints [103]. These apparent discrepancies may be explained by analytical performance differences in the methods used to measure viral load. As mentioned above, most commercial assays use consensus L1 primers because of the ease of use offered by a single reaction that can detect 14 high-risk viruses. However, consensus primer designs exhibit poor detection of mixed infections due to HPV-type suppression or restriction, where the overabundance of one HPV type in a mixed infection can lead to failure to detect lower levels of a coinfecting virus. Van Doorn et al. were one of the first to demonstrate this phenomenon using spiking experiments where detectable levels of HPV18 were eliminated using an increasing concentration of a competing HPV16 target [104]. This result was later confirmed in cervical smear and biopsy specimens where the L1 SPF10 consensus primer assay was compared to a gene-specific E6 assay where significantly more genotypes ($P < 0.0001$) were identified by the E6 assay, especially for HPV types 16, 35, 39, 45, 58, and 59 and the authors concluded “that broad-spectrum PCRs are hampered by type competition when multiple HPV genotypes are present in the same sample” [105]. Similar results have been
reported by Mori et al. who found that “three consensus primers frequently caused incorrect genotyping in the selected clinical specimens containing HPV16 and one or two of HPV18, 31, 51, 52, and 58” and went on to conclude “that PCR with consensus primers is not suitable for genotyping HPV in specimens containing multiple HPV types” [106]. Given that approximately one-third or more of clinical specimens can harbor more than one HPV type [107], this should be considered carefully when designing vaccine monitoring or other studies requiring genotyping. This has recently been underlined by a post-hoc analysis of the PATRICIA vaccine trial where Struyf et al. found that an E6-based multiplex type-specific PCR and reverse hybridization assay showed improved sensitivity versus the L1-based SPF10 PCR-DNA enzyme immunoassay (DEIA)/line probe assay (LiPA25) used in the original trial, resulting in higher vaccine efficacy estimates for nonvaccine oncogenic HPV types [108]. The E6-assay was developed by Van Doorn et al. and had previously been shown to increase genotype detection by 14.3% [105]. Another negative impact of HPV-type suppression was reported by Cornall et al. who hypothesized that the sensitivity of consensus-based PCR approaches could be altered in highly vaccinated populations such as Australia. They confirmed their hypothesis using the Roche HPV linear array genotype assay by simulating samples containing 1000 copies of one or two high-risk HPV DNA genomes in the presence and the absence of 10,000 copies of the HPV16 genome. HPV16 alone did not affect detection of other high-risk genotypes; however, when HPV16 and an additional genotype were present, detection of HPV31, 33, 51, or 59 was impeded, indicating potential for misrepresentation of population-based prevalence of these genotypes and false evidence for type replacement following vaccination [109]. A next-generation sequencing (NGS) method also found that consensus MY09/MY11 primers had “lower sensitivity for some HPV types than LiPA, conceivably due to the poor sensitivity of the MY09/MY11-based primers” [110]. The results from the WHO LabNet genotyping panel broadly support these findings. The panel consists of approximately 43 DNA standardized DNA samples with single and mixed infections of the 14 high-risk HPV types (and low-risk types HPV6 and HPV11) and three cell line extraction controls. Candidate assays are considered proficient if they can detect 50 international units (IU) of HPV type 16 (HPV16) and HPV18 DNA and 500 genome equivalents (GE) for the other 14 HPV types. The 2010 panel results reported the data from 98 laboratories who submitted 132 datasets, only ~20% of which were deemed proficient for all HPV types. In addition, approximately 35% of the test panels had multiple false positive results and were considered nonproficient. Virtually all of the assays that submitted test results were L1 consensus-primer based [63]. The results from the 2014 panel from 119 participating global laboratories (146 datasets) were recently reported and the overall results showed an improvement, with 59% of the test results deemed 100% proficient and 20% nonproficient [111]. Nevertheless, there is little room for complacency when one considers that this implies that 40% of current assays do not accurately detect HPV types. In summary, assay design has a direct impact on clinical performance and the ability to accurately genotype both single and mixed infections will play an ever-increasing role, especially in a postvaccination era.
<table>
<thead>
<tr>
<th>Diagnostic challenge</th>
<th>Impact</th>
<th>Mitigation</th>
<th>Current status</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most infections do no result in cancer</td>
<td>Cancer incidence can be as low as 5–8 per 100,000 women screened [112]</td>
<td>Organized screening programs need to reach as many women as possible</td>
<td>50% of cancers still occur in women who have not been screened</td>
<td>Need to continue to expand programs to reach underserved women (self-sampling)</td>
</tr>
<tr>
<td>Younger women are more likely to be infected with HPV</td>
<td>Screening of women &lt; 25–30 may lead to overdiagnosis [113]</td>
<td>Current guidelines recommend not screening women &lt;25 or &lt;30</td>
<td>Primary screening approved in US for &gt;25, with &gt;30 more common for cotesting with cytology</td>
<td>Ongoing countrywide programs will shed more light on the correct age to start screening (which could be population specific)</td>
</tr>
<tr>
<td>Older women are more likely to develop cancer</td>
<td>Overtesting may cause unnecessary anxiety in younger patients</td>
<td>Adhere to expert guidelines and screen appropriately</td>
<td>Meijer guidelines and FDA approval provide reassurance but in-country real-world validation is also informative</td>
<td>Need more research on the impact of cross-reactivity and molecular mimics as well as mixed infections</td>
</tr>
<tr>
<td>Analytical positive ≠ clinical positive [42]</td>
<td>Over- and underreferral possible</td>
<td>Establish appropriate clinical cutoff in target population</td>
<td>Performance of different collection devices should be considered [114, 115]</td>
<td>Physicians and laboratory personnel need to adhere to recommended procedures for collection and processing of specimens</td>
</tr>
<tr>
<td>Results impacted by quality of specimen collection</td>
<td>Poorly collected or expressed samples increases unsatisfactory rate and can lead to false negative results if no cellular control present</td>
<td>Include cellular versus PCR processing control in the assay</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large number of normal versus abnormal cells present in sample leads to sampling variation</td>
<td>Sampling errors can occur and lead to false negative results [116]</td>
<td>Assay needs to be robust to differences in pipetting volumes around recommended test volume</td>
<td>Most of the newer assays have a cellular processing control which helps mitigate this risk</td>
<td>Most manufacturers have adapted their methods for the two common liquid-based-cytology media on the market</td>
</tr>
<tr>
<td>Most endocervical samples are collected in preservative which lead to clumping of exfoliated cells</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Human Papillomavirus - Research in a Global Perspective
### Table 2. Diagnostic challenges associated with cervical cancer screening.

<table>
<thead>
<tr>
<th>Diagnostic challenge</th>
<th>Impact</th>
<th>Mitigation</th>
<th>Current status</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samples can be inhibited with a variety of exogenous and endogenous interfering substances [118]</td>
<td>False negative or indeterminate results obtained</td>
<td>Product needs to undergo rigorous analytical validation with a variety of interfering substances</td>
<td>Most assays are robust to interfering substances</td>
<td>Less of an issue than sampling variation</td>
</tr>
<tr>
<td>Infections are dynamic and can produce a large dynamic range of viral titers</td>
<td>Infections may be regressing or proliferating leading to false results</td>
<td>Correlate with cytology or other biomarker results and repeat test as needed</td>
<td>Clinically validated assays have good correlation with disease endpoints. One assay reported predictive ability of viral load measurements taken sequentially [119]</td>
<td>Viral load measurements are difficult to perform and may not be clinically practical if sequential. Genotyping provides information on persistence which may aid in patient management</td>
</tr>
<tr>
<td>The same pathology can be caused by non-high-risk HPV or other “mimics”</td>
<td>Non-high-risk HPV pathology can lead to overtreatment</td>
<td>Assays should not detect nontarget HPV types. Histology results should be considered in the context of high-risk HPV results</td>
<td>Newer assays have no cross-reactivity to nontarget types. Pathology review and biomarker qualification improve diagnostic cytopathology accuracy</td>
<td>The belief that “gold standard” pathology is truth needs to be critically challenged [120]</td>
</tr>
<tr>
<td>Viral clearance may precede resolution of cellular abnormalities</td>
<td>Lingering pathology complicates patient management</td>
<td>Adhere to screening guidelines and retest as needed</td>
<td>p16 biomarker is increasingly being used to augment diagnosis and guidelines have been proposed [121]</td>
<td>Extended genotyping offers improved risk stratification which may aid in patient management [96]</td>
</tr>
</tbody>
</table>

#### 6.5. HPV integration-induced deletions: impact on screening programs?

It is now well established that the HPV viral genome can integrate into the host genome during disease progression. The ability to integrate varies by high-risk HPV type with approximately 70% of HPV16-associated cervical cancers containing integrated HPV16 sequences, rising to almost 100% for HPV18 [122, 123]. Next-generation sequencing methods are now shedding new light on this phenomenon and the emerging picture suggests the following:
1. Viral integration appears to be a random (nontargeted) event with disruption of a wide variety of host genes across the human genome [124–126].

2. Integration is associated with fragile sites in human DNA and is likely opportunistic, with the HPV virus taking advantage of exposed DNA to integrate into the host genome. Regions of microhomology between the viral and human sequences are enriched near integration sites suggesting that integration may be driven by the host DNA repair machinery [124, 125].

3. Viral sequences are frequently deleted during the integration process [124, 125].

4. Contrary to what was previously believed E6 and E7 oncogenes can also be deleted during integration. However, this occurs less frequently than in other viral genes such as L1, L2, E1, and E2 and there appears to be significant enrichment of (intact) viral gene E7 > E4 > E5 > E6 reads among the cervical tumor samples [124, 125, 127]. (It is possible that this may simply due to the random nature of the integration process and reflect the fact that L1, L2, E1, and E2 are the largest ORFs in the HPV genome, Figure 2).

5. Viral integration appears to occur at a higher frequency than perhaps previously understood and can also be detected in Pap smears, one study reporting integration in 53% of cytology samples with histology grades CIN1–3 [124]. Thus, it appears that viral integration HPV integration is an unintended consequence of HPV replication and that following integration, the virus life cycle is aborted. Nevertheless, virus integration is a routine consequence of infection and the observation that it can occur earlier in the disease progression prompts the question whether it has a measureable effect on the ability of HPV assays to detect the virus. Several authors have expressed concern that L1-based assays can fail to detect late-stage cancers due to target deletion [128–130] and there is no doubt that L1-deleted cancers occur as evidenced by case studies [129, 131]. Several studies have reported that the E6/E7 assays have increased sensitivity for cancer versus L1-based assays [132–134]. Some have dismissed this as due to earlier methodological differences due to the difficulty of detecting longer L1 amplicons in formalin fixed paraffin embedded (FFPE) tissue. However, this is not supported by more recent literature where even using gold standard FFPE processing methods, one group reported only a 91% detection rate in cancer specimens with two different L1 primer assays [135]. Another recent study of cervical adenocarcinomas reported that PCR using a very sensitive L1-based SPF10-PCR resulted in 482 cases (67%) positive for HPV DNA, but that testing using type-specific E6 PCR added 53 HPV-positive cases (a 7% increase in detection). The study design also accounted for DNA adequacy in the samples [136]. Adenocarcinomas represent about 25% of all cervical cancers [137, 138] and are known to be prone to HPV DNA integration, so this increase in sensitivity for detection of adenocarcinomas could potentially translate to the detection of another 1–2% of total cancer cases (assuming an increase in just adenocarcinoma cases). Large-scale studies will be needed to address the question of whether or not the above reported difference in assay performance (due to increased maintenance of intact target regions postintegration) has a measureable impact on clinical sensitivity in cervical cancer screening. This warrants further study, especially given the reported ability of the virus to integrate in approximately 50% of cervical
smears from patients with CIN [124]. Finally, the Roche cobas assay has recently been reported to have a 94% detection rate in the smears from women with histologically confirmed invasive cervical cancer [139].

7. The evolution of HPV screening

7.1. Cotesting with cytology versus primary HPV screening

It has been known for 10 years that HPV testing is substantially more sensitive in detecting CIN2+ than cytology (96.1 versus 53.0%) but less specific (90.7 versus 96.3%) [140]. In the intervening period, large randomized controlled trials (RCTs) have continued to reaffirm this point. Pileggi et al. [141] recently performed a meta-analysis of global RCTs which included trials from Italy (NTCC trials I+II [142, 143]), the UK (ARTISTIC Trial [144]), Finland [145], India [146], Canada [147], and the Netherlands (POBASCAM [148]). The analysis from this larger more comprehensive dataset showed a significantly higher detection of both CIN2+ and CIN3+ by HPV testing versus cytology and that the relative specificity of cytology was higher. However, in women greater than 30 years of age, the specificity was not statistically different (“almost overlapping”). In addition, the pooled relative PPV was not significantly lower for HPV compared with cytology. These results suggest that the difference in specificity is actually less than previously thought, at least in women over 30. A recent US post-hoc analysis of a large group of women from multiple practices reported that positive Pap + HPV (cotest) was more sensitive than either a positive HPV alone or Pap alone for detection of CIN3+ and suggested that HPV primary screening might miss cancers if not combined with the Pap [149]. However, the study design has come under criticism from an independent expert, because the short follow up time of 1 year biases results in favor of cytology and all women were not followed equally [150]. Thus, large global studies support the use of primary HPV testing as a means to improve cervical cancer detection and it explains why many countries are now in the process of implementing either pilot or countrywide programs to replace cytology and/or cotesting [151, 152]. In the United States, one HPV test has received FDA approval for primary screening in women >25 years of age [153] and interim guidelines for its use have been issued [113]. The reaction to this decision has been mixed with some arguing that cotesting detects more disease [154]. However, data continue to accumulate both in the United States and in Europe that supports the long-term NPV of a primary HPV negative result, thus permitting safe interval extension [155, 156]. A study comparing primary HPV versus cotesting versus primary cytology concluded that “primary HPV testing every 3 years might provide as much, if not more, reassurance against precancer and cancer, compared to primary Pap testing every 3 years and cotesting every 5 years” [156]. Thus, the choice of which screening paradigm to adopt will likely be influenced by the resources available together with the desired screening interval and risk tolerance of the medical community and the patients they serve. Despite the slow pace of change, it seems clear that HPV primary screening will gradually be adopted worldwide, especially in countries which do not have current cytology infrastructure. HPV testing is considerably more reproducible than cytology [157] and with the advent of automation, highly skilled personnel are no longer required to implement IVD-qualified tests. It should be noted,
however, that primary HPV screening cannot be used in isolation to refer women to colposcopy. Biomarkers such as p16 [121] or cytology or HPV16/HPV18 genotyping need to be considered to increase specificity and reduce the number of unnecessary colposcopies [72, 153, 158].

7.2. Reaching underserved women: self-sampling methods

It is well established that the effectiveness of cervical cancer screening programs is limited by the number of women who do not participate—with 50% or more of disease being detected in women who have not been screened [159–162]. While there may be many reasons (cultural, socioeconomic, and religious) why women choose not to participate in screening programs available to them, it is clear that self-collection offers an effective outreach tool with which to increase participation. For example, in Finland the participation rate was <70% but was increased to 72.6–79.9% by sending reminder letters to nonattendees in 22 municipalities, and to 83.4% by sending self-collection kits to those who did not attend after receiving the initial invitation letter [163]. Similar finding on outreach were reported in Sweden where telephone invitations to long-term nonattendees increased the participation rate within the following 12 months to 18.0% versus 10.6% in a control group [164]. In Canada, reminder letters were compared head to head with sending a self-sampling kit and standard screening and women receiving the self-collected HPV kit were 3.7 times more likely to undergo screening compared with the standard of care [165]. Finally, a recent Danish outreach program where 5000 women were invited to “opt-in” and receive a self-collection kit resulted in the detection of nine cancers. This translates to yield of 1.8 per 1000, a dramatic enrichment for disease detection versus the Danish population rate, estimated at 12.9 per 100,000 women [166, 167]. Thus, self-sampling is emerging as a very important tool in national screening programs and one that continues to generate a high degree of interest [168]. Two principal self-collection methods that have been described in the literature: cervico-vaginal collection and urine collection.

7.2.1. Cervico-vaginal collection methods

A number of different cervico-vaginal methods have been described in the literature (reviewed in [169, 170]) and range from simple brush devices currently used by physicians to custom designed self-collection devices such as the Rovers® Evalyn® brush (Rovers Medical Devices BV, the Netherlands) and HerSwab™ (Eve Medical, Canada). Earlier literature reported mixed results versus physician collected endocervical samples in terms of sensitivity and specificity. However, more recent studies have confirmed that self-collected specimens can have the same sensitivity for CIN2+ as physician-collected specimens [171]. A recent meta-analysis concluded that signal amplification methods were not sensitive enough for self-collection use but that certain PCR methods had similar sensitivity to physician-collected samples [172]. This has been confirmed in a number of small studies using clinically validated HPV tests where sensitivities for CIN2+ were similar to physician-collected samples [173–176]. It is likely that both the quality of the instructions used and the sample workflow/assay play a role in the ability to accurately detect infections. Further large-scale studies are needed to further demonstrate the utility of these methods in screening populations.
7.2.2. HPV testing from urine

HPV testing from urine has been performed for over 20 years and similar to cervico-vaginal methods, early studies were somewhat discouraging with decreased sensitivity versus physician-collected endocervical samples being reported [177]. This was compounded by the demonstration that male urine samples were a poor sample type for HPV detection [178], which has been confirmed in more recent studies and likely reflects a true biological difference between men and women [179, 180]. The general interest in self-sampling and the introduction of HPV vaccines, which can benefit from noninvasive population-based monitoring, has resulted in a renewed interest in urine as an alternative sample type for cervical cancer screening [181]. There is now a growing body of evidence that urine, in particular first void urine, has a lot of potential for detecting HPV infections [182, 183]. The basic premise of HPV urine testing is founded on the hypothesis that “at the start of the void, urine gets contaminated by debris and impurities lining the urethra opening, including mucus and debris of exfoliated cells from the vagina, cervix and uterus” [181]. Recent research has also demonstrated that optimal sample workflows and sensitive detection methods improve clinical performance:

1. Urine contains substantial amounts of free-DNA and virus, so processing methods that utilize recovered cell pellets may underestimate the amount of virus present. Processing of neat urine is preferable [184, 185]. A recent study found similar performance using 0.5 ml of neat urine versus up to 80-fold concentrated cell pellets [186]. Another study found HPV DNA sequences in both urinary tract infection positive and negative patient urine using cesium chloride density gradient-purified virus particles, suggesting that intact virions are present [187]. (Urine also contains transrenal DNA [short DNA fragments from the blood circulation that has passed the kidney barrier [188], but this is not believed to be the main source of HPV target DNA as evidenced by the observation that first void urine collected during menses is contaminated with blood and cells.)

2. Urine from HPV positive women appears to be consistently positive even when sampled multiple times from the same patient. However, first void urine samples are analytically more positive, especially if the initial catchment is in a smaller volume [181, 182, 186, 189].

3. Virus appears to be shed continuously from infected women and is simply flushed out of the vaginal canal during urination. Therefore, it is not advisable to clean the local area with alcohol wipes prior to collection in order to maximize recovery of cells, virus, and free-DNA [181].

4. Urine samples are labile at room temperature (due to the action of proteases and nucleases) so urine samples should be mixed promptly with an appropriate preservative such as EDTA to ensure sample integrity. Addition of the preservative to the urine cup or collection tube is optimal since degradation can occur immediately post-collection prior to the arrival of the sample in the testing laboratory [184, 188, 190–192].

5. An assay cellularity control targeting a human gene is recommended to ensure sample adequacy. Routine urine samples should be positive for human DNA and confirms that an adequate catchment has been obtained. Analytically sensitive molecular methods such
as real-time PCR should be used to ensure optimal results and clinical cutoff adjustments may be required versus endocervical specimens [193].

8. “Safety in numbers”: time for extended genotyping in cervical cancer screening

The majority of (clinically validated) HPV screening tests have the ability to provide genotyping information on only two types, HPV16 and HPV18 [46]. There is now a growing body of evidence that extended genotyping (beyond types 16 and 18) better stratifies a woman’s risk for subsequent disease: A recent Danish study looked at the long-term absolute risk for CIN3+ in women following a baseline HPV genotyping result and found that HPV31 and HPV33 had the same or greater longitudinal risk as HPV18 over a 10-year period [194]; Schiffman et al. [195] confirmed this finding for HPV31 in a US cohort over a 15-year follow up period; Cuzick et al. [196] looked at the positive predictive value of genotyping in a referred UK population and found that HPV33 had a higher PPV than HPV16 (59.8 versus 57.8) and that HPV31 was higher than HPV18 (39.5 versus 29.3). Similar trends were reported in a much larger New Mexico study of 21,297 women of whom 77% had biopsies. Among women with CIN3+ (n = 1880), 14.9% were attributable to HPV31, 5.2% to HPV33 and 4.9% to HPV18, with HPV16 responsible for 54.1% [197]. Roche Linear Array genotyping analysis of patients from the ATHENA trial also support these findings: among 40,901 women aged ≥25 years HPV16 conferred the greatest absolute risk of ≥CIN3 both in women aged 25–29 and ≥30 years (14.2% and 15.1%, respectively) followed by HPV31 (8.0% and 7.9%), HPV52 (6.7% and 4.4%), and HPV18 (2.7% and 9.0%) [198].

While the majority of cancers are caused by HPV16 and HPV18, presumably due to the increased ability of these genotypes to persist and induce oncogenic changes, one must equally concede that HPV31, 33, 45, 52, and 58 are the next most potent carcinogens of the high-risk HPV types [199]. This is supported by the introduction of the nine-valent HPV vaccine, where the addition of these types is estimated to provide protection against an additional ~20% of cancers [200]. The clinical paradigm of equal management of equal risk [201] strongly suggests that knowledge of HPV31 and HPV33 infections has clinical value and can aid in patient management. In addition, the widespread introduction of the new nine-valent vaccine is expected to have a profound effect on the incidence of CIN3+ disease, further underlining the need for a more comprehensive genotype profile of pre- and postvaccinated women [202]. Finally, the use of gene-specific versus consensus PCR genotyping approaches can alleviate the issue of apparent type replacement in vaccinated women where HPV types such as HPV31 and HPV33 may appear to be increasing in prevalence but are simply being unmasked due to the reduced incidence of HPV16 and HPV18 in the postvaccination era [109].

Extended genotyping also offers a potential solution to key emerging issues in cervical cancer screening:

I. Accurate risk stratification—absolute risk calculations a priori are underestimated when two or more HPV types with different CIN3 risks are pooled together (e.g., the 10-year longitudinal study of Kahn et al. attributes a very low risk to pooled non-16/18
types [203] which has been consistently found not to be the case in subsequent studies where individual genotypes have been detected [194, 196–198]). Precise risk stratification is essential for consistent and informed patient management. This is especially true in cytology-negative HPV-positive women where genotyping can provide clinicians with a much more informed assessment of future risk versus a pooled HPV positive result.

II. On the other hand, extended genotyping allows one to group HPV types with similar risk, thereby simplifying patient management algorithms. Cuzick et al. [196] proposed a three-tiered risk group approach to patient stratification based on the positive predictive value of different genotypes for CIN3+ disease in a referred population. Schiffman et al. recently published the largest longitudinal ASC-US population analysis to date (Persistence and Progression cohort with 13,890 women aged 21+ with HC2 (QIAGEN)-positive ASC-US at enrollment and median follow-up of 3.0 years). The authors used the concept of equal management of equal risk approach and calculated the 3-year CIN3+ risk for all HC2-positive women with ASC-US (5.2%), using this as the “benchmark” risk for colposcopic referral. They concluded that the 3-year risk for developing CIN3+ associated with high-risk HPV types 35, 39, 51, 56, 59, 66, and 68 (2.7% for HPV51, 1.6% for HPV39/HPV68/HPV35, and 1.3% for HPV59/HPV56/HPV66) “might be low enough to recommend 1-year retesting, permitting viral clearance. This strategy would defer colposcopy for 40% of women with HPV-positive ASC-US, half of whom would be cotest-negative at 1-year return” [96].

Thus, extended genotyping offers the potential of improved risk stratification and simpler patient management, helping to improve patient outcomes with reduced intervention.

9. Future perspectives

While we have made great strides in both detecting and preventing cervical cancer over the last 20 years, it still represents a significant challenge in both developed and developing settings. In countries with a high disease burden such as Mexico, Ecuador, Samoa, and Colombia, it is the number one cancer in young women and the number two cancer in women of all ages, whereas in several areas in Africa and in Cambodia, it is primary cause of cancer and death [112]. With 50% of cancers occurring in underserved women, it is clear that the biggest remaining challenges are in implementation, not technology development. Self-collection methods offer new hope of reaching women who are currently not captured in traditional screening programs. Ongoing studies should shed further light on the ability of these methods to equal those used in the physician’s office or at least provide a means to target these women for follow up in the medical system (analogous to current home pregnancy test kits). HPV vaccines have been shown to be safe and efficacious but here again the biggest hurdle is achieving the required vaccine coverage in the target population. It is hoped that the successful implementation
models such as those in the UK and Australia will be adopted elsewhere to boost global vaccine coverage. It should also not go unmentioned that the other barrier to improved disease prevention and detection is resistance to change. While medicine can be understandably conservative, waiting for substantial bodies of evidence to accumulate prior to changing clinical management, we should also constantly challenge ourselves on this point, given the severity of the disease burden imposed by HPV. With the benefit of hindsight, one can justifiably argue that even the Pap smear itself took 30 years too long to be implemented, HPV primary screening could have been adopted 10 years ago based on the available evidence, and vaccine coverage is suffering today due to the emergence of vaccine opposition which is not supported by scientific evidence [204]. We should also continue to challenge the status quo on cervical cytology and histopathology—while they have an important role to play, their future role should be evaluated in the context of comprehensive screening strategies [205, 206]. There is an urgent need to eradicate cervical cancer for the sake of both the current generation of women and the men and women who follow them. At the same time, there is an ever increasing pressure on global healthcare systems to do more with less. Ironically, this will likely serve as a catalyst for change as health economic studies identify which screening and vaccination strategies are the most cost-effective and offer the highest adoption rate. Ongoing and future research should focus on answering key questions in cervical cancer program implementation such as the effect of vaccination on HPV prevalence, and the ability of extended genotyping to better stratify a patient’s risk. Optimal triage strategies should also be investigated together with new studies on the long-term effects on the physical and psychological health of patients who are overtested and overtreated. As we settle into the twenty-first century of cervical cancer screening, one thing is clear, we have effective means to detect and prevent disease. The real question is: do we have the resolve to deploy these tools in an effective manner? We join others in the hope that we will learn from the past and quicken the pace of change in response to the global threat posed by cervical disease [207].

Trademarks are the property of their respective owners.

© 2016 BD. BD, the BD Logo and BD Onclarity are trademarks of Becton, Dickinson and Company.

**Author details**

Laurence M. Vaughan†, Brian R. Faherty, Erin C. Gutierrez, James M. Harris, William A. Nussbaumer and Ryan J. Schwab

*Address all correspondence to: laurence.vaughan@bd.com*

Diagnostics Systems, BD Life Sciences, Sparks, MD, USA
References


[25] Patanwala IY, Bauer HM, Miyamoto J, Park IU, Huchko MJ, Smith-McCune KK. A systematic review of randomized trials assessing human papillomavirus testing in


[41] QIAGEN. Hybrid Capture 2 High-Risk HPV DNA Test Package Insert. 2007;L00665 Rev. 2.


[67] Roche. cobas 4800 HPV Test, Instructions For Use. 2013;Reference 05641225001-10EN(Doc. Rev, 8.0).


[159] van Ballegooijen M, van den Akker-van Marle E, Patrick J, Lynge E, Arbyn M, Anttila A, Ronco G, Dik J, Habbema F. Overview of important cervical cancer screening process...


[193] Thomsen LT, Frederiksen K, Munk C, Junge J, Iftner T, Kjaer SK. Long-term risk of cervical intraepithelial neoplasia grade 3 or worse according to high-risk human...


