

FDA Executive Summary

New Approaches in the Evaluation for High-Risk Human Papillomavirus Nucleic Acid Detection Devices

Prepared for the March 8, 2019 meeting of the Microbiology Devices
Panel of the Medical Devices Advisory Committee

Table of Contents

| | |
|--|----|
| I. Introduction and Purpose of the Meeting:..... | 4 |
| II. Background: | 4 |
| A. HPV Biology | 4 |
| B. Regulation of HPV devices: History..... | 5 |
| C. Scope of Panel Meeting | 7 |
| III. Clinical Study:..... | 9 |
| A. Background: Clinical Studies Supporting HPV Device Approval | 9 |
| B. Panel Topic 1: Study Design | 10 |
| C. Panel Topic 2: Colposcopy Referral Protocol in Clinical Studies | 12 |
| IV. HPV Device Evaluation..... | 14 |
| A. Performance Evaluation: Background..... | 14 |
| B. Panel Topic 3: Indications for use | 20 |
| C. Panel Topic 4: Data Analysis to Support the New Indications for use..... | 22 |
| D. Panel Topic 5: Clinical Endpoint Comparator..... | 24 |
| V. Questions to Panel..... | 29 |
| Appendix | 30 |
| References | 31 |

Table of Abbreviations:

| Abbreviation | Definition |
|---------------------|---|
| ASCCP | American Society for Colposcopy and Cervical Pathology |
| ASCUS | Atypical Squamous Cells of Undetermined Significance |
| ATHENA | Addressing THE Need for Advanced HPV Diagnostics |
| CDC | Centers for Disease Control and Prevention |
| CIN | Cervical Intraepithelial Neoplasia (cervical dysplasia) |
| CRP | Centralized Pathologist Review |
| DNA | Deoxyribonucleic acid |
| ECC | Endocervical Curettage |
| FDA | Food and Drug Administration |
| HPV | Human Papillomavirus |
| HR | High-risk |
| IARC | International Agency for Research on Cancer |
| LBC | Liquid-based Cytology |
| LEEP | Loop Electrosurgical Excision Procedure |
| LSIL | Low-grade Squamous Intraepithelial Lesion |
| NHANES | National Health and Nutrition Examination Survey |
| NILM | Negative for Intraepithelial Lesion or Malignancy |
| NLR | Negative Likelihood Ratio |
| NPV | Negative Predictive Value |
| PC | PreservCyt |
| PCR | Polymerase Chain Reaction |
| PLR | Positive Likelihood Ratio |
| PMA | Premarket Approval |
| SGO | Society for Gynecologic Oncology |
| SOC | Standard of Care |
| SP | SurePath |
| STM | Specimen Transport Media |
| USPSTF | United States Preventative Services Task Force |

I. Introduction and Purpose of the Meeting:

The first high-risk human papillomavirus (HR HPV) device indicated to be used in routine cervical cancer screening was FDA approved in 2003. Over the course of the last 16 years, data from basic scientific research as well as large scale epidemiological studies firmly established the value of testing for HR HPV genotypes in ruling out the likelihood of harboring a precursor to cervical cancer (i.e., precancer). This set the foundation for widening of cervical cancer screening intervals and, in some cases, recommendations for patient observation rather than immediate colposcopy. The importance of HR HPV testing in cervical cancer screening increased substantially in 2014, when, [following advisory panel recommendations](#), the FDA approved the first device to be used for primary HPV screening. Shortly thereafter, the Society for Gynecologic Oncology (SGO) and the American Society for Colposcopy and Cervical Pathology (ASCCP) issued interim guidelines recommending primary HPV screening as an acceptable approach for women 25-65 years of age [1]. The guidance issued in September 2018 by the United States Preventative Services Task Force (USPSTF) confirms similar use, recommending primary HPV screening or HPV/cytology co-testing every 5 years for women aged 30-65 years old [2].

However, decreases in HR HPV positivity rates due to HPV vaccination and additional research have created a setting for HR HPV testing that is different from when the first generation of HR HPV devices were approved. The long term predictive values of HR HPV tests and regression rates of cervical dysplasia, as well as other important clinical predictors, are now known from over a decade of research involving HPV devices in clinical practice.

We believe the accumulated knowledge over the past 16 years, against the background of changing prevalence, could support innovation in the methods used in the development and evaluation of HR HPV devices. The purpose of the March 8, 2019, meeting is to discuss new approaches to the development and evaluation of HR HPV devices that might allow for advances and innovation in this area and reduced burden.

II. Background:

A. HPV Biology

HPV is the most common sexually transmitted infection in the US, with an approximately 80% lifetime risk of infection for any oncogenic type [3]. [According to the CDC](#), 79 million Americans are currently infected with HPV. There are over 200 HPV types, about 40 of which infect the genital tract. Of these 40 types, the International Agency for Research on Cancer (IARC) has classified the following 12 HPV types as carcinogenic (i.e., high risk): HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58 and 59. One HPV type, 68, has been classified as probably carcinogenic. Persistent infection with HR HPV is typically necessary for the development of cervical cancer. DNA from HR HPV types has been found in over 99.7% of squamous cell carcinomas of the cervix [4]. HPV types 16 and 18 account for about 70% of cervical squamous cell carcinomas, while types 16, 18, and 45 account for 94% of cervical adenocarcinomas [5]. Persistent HR HPV infection can lead to upregulation of the viral oncoproteins E6 and E7, which can interfere with cell cycle regulation and DNA repair processes. This may result in mutations that may eventually lead to cellular transformation and tumor cell growth [6].

It has been reported that 90% of HPV infections will clear on their own within 2 years [7]. Because most HPV infections clear, the goal for designing HPV devices is to optimize the assay to detect only those

infections that have a high likelihood of leading to cervical cancer. Due to both the low prevalence of cervical cancer relative to the number of women infected with HR HPV and ethical concerns of using cervical cancer as the endpoint for clinical trials, cervical precancer is used as a surrogate endpoint in clinical studies. The endpoints currently used in clinical studies supporting FDA approval are cervical intraepithelial neoplasia grades 2+ and 3+ (CIN2+ and CIN3+). CIN2+ refers to any lesion that is grade CIN2 or greater, including CIN3 and cancer. CIN3+ refers to any lesion that is grade CIN3 or greater, including cancer. To ensure that only clinically relevant infections (i.e., those associated with cervical precancer and cancer) are detected, HR HPV devices are designed to only detect potentially carcinogenic infections (i.e., “high risk”), rather than all HPV infections. However, even for HR HPV infections, about 50% will spontaneously clear within 1-2 years [8]. HR HPV viral load has been identified as a major factor associated with cervical precancer, especially for the alpha 9 genotype family, which contains types 16, 31, 33, 35, 52, and 58 [9-12]. Therefore, HR HPV devices utilize a “clinical cutoff” (also known as a “clinical cutpoint”), where only HR infections with viral nucleic acid levels resulting in signals exceeding a certain threshold are reported as positive for HR HPV, in order to minimize identification of low viral load infections that have a high probability of spontaneous clearance.

Generally, the clinical cutoffs of HR HPV devices have been optimized to detect infections associated with CIN2+ lesions in a population of women with atypical squamous cells of undetermined significance (ASC-US) cytology diagnoses (although some manufacturers have tested different populations to set the device’s cutoff). However, because there is no international reference standard for HR HPV types, there is no universally accepted viral load that is considered clinically relevant for any genotype. For this reason, the clinical cutoff is established separately by each manufacturer for each individual device, which can sometimes lead to discrepancies in the viral levels that constitute “clinically relevant infections” between devices.

B. Regulation of HPV devices: History

Assays detecting HR HPV nucleic acid are class III devices, which are regulated by the Office of In Vitro Diagnostics and Radiological Health (OIR)¹ at the FDA’s Center for Devices and Radiological Health and require the submission of a premarket approval (PMA) application. The first HPV device that was FDA-approved for marketing was the ViraType Human HPV DNA Typing Kit in 1991, a dot blot nucleic acid hybridization assay that detected HPV types 6/11, 16/18, and 31/33/35. This test was primarily indicated for the detection and differentiation of HPV nucleic acid types.

Almost a decade later, the FDA-approved indications for use for HPV tests evolved from those for use as a generic nucleic acid detection test to specifying how the test would be used to guide patient management. This was apparent in 2000 with the approval of the Digene HCII assay (now referred to as HC2), which detected the HPV types 6/11/42/43/44 (low risk HPV) and 16/18/31/33/35/45/51/52/56² (high risk HPV), with one of the indications being the following (referred to as “ASC-US Triage”):

To screen patients with ASCUS (atypical squamous cells of undetermined significance) Pap smear results to determine the need for referral to colposcopy. The results of this test are not intended to prevent women from proceeding to colposcopy.

¹ For more details on FDA regulation of in vitro diagnostic devices, please see the Appendix.

² The current High-Risk HC2 Assay detects only high risk types, 16/18/31/33/35/39/45/51/52/56/58/59/68. Please note that type 68 is probably carcinogenic.

Three years later, in 2003, the indications for use for the Digene HC2 assay further expanded upon approval of a PMA supplement to incorporate HR HPV testing in routine screening, in addition to ASC-US triage. This is referred to as the “adjunct indication” and is stated as follows:

In women 30 years and older, the HC2 High-Risk HPV DNA test can be used with Pap to adjunctively screen to assess the presence or absence of high-risk HPV types. This information, together with the physician’s assessment of cytology history, other risk factors, and professional guidelines, may be used to guide patient management.

It was not until 2009 when the second HR HPV device, Cervista HPV HR, manufactured by Third Wave Technologies (now Hologic), was approved for both the ASC-US triage and Adjunct indications. This test detects HR HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, and 68 (i.e., 14 HR HPV types) as a pooled result. At this time, HPV type 66 was added to the pool of detected genotypes. It was also during this time that the clinical utility of partial HR HPV genotyping for HPV types 16 and 18 was recognized, as these two types are the most commonly associated with cervical cancer [5, 13, 14]. In conjunction with the premarket approval of the Cervista HPV HR assay, the Cervista HPV 16/18 assay was approved. The genotyping assay was to be used adjunctively with the Cervista HPV HR assay to detect the presence or absence of HR HPV types 16 and 18 in women who were HR HPV positive. In 2011, two additional HR HPV devices were approved with the ASC-US triage and Adjunct indications: the Hologic Aptima HPV Assay, which detects the 14 HR HPV types collectively, and the Roche cobas HPV Test, which has independent readouts for HR HPV types 16, 18, and “12 other” types.³ In 2012, the Aptima HPV 16 18/45 Genotype Assay was approved to be used adjunctively with the Aptima HPV assay. This assay has readouts for type 16 individually and types 18/45 in combination.

Through a PMA supplement in 2014, an additional indication for use for Primary HPV Screening (herein referred to as the “primary screening indication”) was FDA-approved for the cobas HPV test, consistent with the recommendation from the [Microbiology Devices advisory panel](#) on March 12, 2014. The primary screening indication states:

In women 25 years and older, the cobas HPV test can be used as a first-line primary cervical cancer screening test to detect high risk HPV, including genotyping for 16 and 18. Women who test negative for high risk HPV types by the cobas HPV test should be followed in accordance with the physician’s assessment of screening and medical history, other risk factors, and professional guidelines. Women who test positive for HPV genotypes 16 and 18 by the cobas test should be referred to colposcopy. Women who test high risk HPV positive and 16/18 negative by the cobas HPV test (12 other HPV positive) should be evaluated by cervical cytology to determine the need for referral to colposcopy.

For primary screening, the HR HPV test would be the first-line test, with patient management driven primarily by the HR HPV test results, rather than HR HPV test results serving as a triage for women with ASC-US diagnoses or as an adjunct to cytology. This approval was based on a ‘ATHENA’ prospective study conducted by Roche that compared primary HPV screening to accepted screening modalities at the time (i.e., cytology alone and co-testing). For this indication, rather than assessing performance of the *HR HPV device* itself, the performance of the *primary screening algorithm* using the investigational device was evaluated and compared to screening with cytology alone or co-testing. The benefit-risk

³ The ‘12 other’ types refers to types 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, and 68.

analysis demonstrated that primary screening was a safe and effective alternative to screening methods deemed acceptable at the time.

The most recent HR HPV device to have gained FDA premarket approval was the BD Onclarity HPV Assay, in 2018. This assay was approved for all three indications: ASC-US triage, Adjunct, and Primary Screening. The Onclarity HPV assay has independent readouts for HPV types 16, 18, 45, and “11 other” types.⁴

It should be noted that HR HPV devices are approved as a package with the collection device(s), such as the endocervical brush/spatula or broom, and collection media (i.e., PreservCyt [PC], SurePath [SP], specimen transport media [STM]), that were validated in the clinical study conducted to support the PMA. Table 1 below summarizes the FDA-approved HR HPV devices to date, for each indication and collection medium:

Table 1: FDA-approved HPV devices

| HPV assay | Manuf. | Genotyping | Indication for Use | | | | | |
|----------------------|---------|----------------------|---------------------------|-----------------|---------------------------|----|-------------------|----|
| | | | ASC-US Triage | | Adjunct | | Primary Screening | |
| Collection media | | | PC ^a | SP ^b | PC | SP | PC | SP |
| Digene* HC2 | Qiagen | No | X | | X | | | |
| Cervista HPV HR | Hologic | No | X | | X | | | |
| Cervista HPV 16/18 | Hologic | 16, 18 | Used with Cervista HPV HR | | Used with Cervista HPV HR | | | |
| Aptima HPV | Hologic | No | X | | X | | | |
| Aptima HPV 16, 18/45 | Hologic | 16, 18/45 | Used with Aptima HPV | | Used with Aptima HPV | | | |
| Cobas HPV | Roche | 16, 18, 12 other | X | X | X | X | X | X |
| Onclarity HPV | BD | 16, 18, 45, 11 other | | X | | X | | X |

*this assay is also approved to be used in STM (specimen transport media)

^a PC= PreservCyt

^b SP= SurePath

C. Scope of Panel Meeting

The FDA is seeking the panel’s input for new/updated approaches in the evaluation of premarket approval applications for HR HPV devices. The FDA believes now is the appropriate time to re-evaluate approaches to HR HPV device development and evaluation, for several reasons:

- **Broader knowledge of cervical carcinogenesis:** In the 16 years since HPV testing has been approved to be used in routine screening, a plethora of scientific data has been published on the

⁴ The ‘11 other’ types refers to types 31, 33, 35, 39, 51, 52, 56, 58, 59, 66, and 68.

clinical utility of HR HPV testing using FDA-approved devices in cervical cancer screening, as well as on the natural history of cervical carcinogenesis. FDA is looking to leverage this data for regulatory decision making so that the validation and evaluation of HR HPV devices can be less burdensome and more consistent with current clinical practice and scientific knowledge.

- **Decreased prevalence of vaccine targeted HR HPV infections due to HPV vaccination:** The uptake of HPV vaccination is resulting in a changing screening population, due to the decreased prevalence of HR HPV infections. These changes affect not only clinical study design due to decreased numbers of women who are HR HPV positive, but also procedures that are dependent on morphological characterizations (i.e., cytology, histology)[15, 16], which play pivotal roles in HR HPV device clinical evaluation. It may be necessary, therefore, to consider how vaccination will affect the intended use populations and what changes should be instituted in the evaluation of HR HPV devices.
- **Evolving screening and patient management guidelines:** Cervical cancer screening and practice guidelines are continually evolving; for example, in 2017, the first ASCCP Colposcopy Standards Recommendations were published [17]; in 2018, the USPSTF released updated cervical cancer screening recommendations [2]; and preparations for a new round of ASCCP consensus screening and management guidelines have commenced [18]. While harmonizing HR HPV device evaluation by FDA as per the clinical guidelines would be ideal, it is burdensome to continuously change regulatory recommendations for HR HPV device validation to incorporate evolving screening guidelines, as studies supporting device approvals often take several years. Therefore, approaches need to be developed for FDA's review of HR HPV devices in the context of changing screening guidelines.

The design of clinical studies and data analyses for performance evaluation of HR HPV devices will be the primary subject of discussion for this panel meeting. Specifically, the panel meeting will address the following topics:

1. Clinical Study Design- Benefits and Risks of:
 - A. Enrichment studies using specimens collected from referral populations
 - B. Archived specimens
 - C. Capping the vaccinated population
2. Colposcopy Referral Protocol in Clinical Studies
3. Indications for Use
 - A. Simplifying the indications to encompass one general screening population
 - B. Removing reference to specific triage tests and clinical actions
4. Data Analyses to Support the new indications for use
 - A. Performance evaluation **without** a clinical endpoint comparator – composite molecular comparator
 - B. Performance evaluation **with** a clinical endpoint comparator – relative device performance
5. Clinical Endpoint Comparator

The FDA is looking to use current understanding of cervical cancer screening to inform the way PMA applications for HR HPV devices are reviewed, so that the highest quality of scientific rigor will remain as the standard for evaluating HR HPV devices, while at the same time encouraging least burdensome approaches.

III. Clinical Study:

A. Background: Clinical Studies Supporting HPV Device Approval

In clinical studies supporting approval of HR HPV nucleic acid tests, the investigational device has been evaluated for its ability to detect cervical precancer (a histological diagnosis of CIN2+ and CIN3+). These prospective studies typically enrolled all comers who present for routine cervical cancer screening. The women enrolled in these studies had three screening test results: 1. Pap test, 2. HR HPV result from an FDA-approved test accepted as standard of care (SOC), and 3. HR HPV test result from the investigational device under evaluation. All women enrolled in the study who had abnormal cytology (ASC-US or greater) or a HR HPV positive test result by either the FDA-approved or investigational device (any genotype) were referred to colposcopy to have a biopsy taken. Additionally, a random subset (typically around 5%) of women who had a negative cytology diagnosis (Negative for Intraepithelial Lesion or Malignancy (NILM)) and who are HR HPV negative by both the investigational and FDA-approved HPV devices (herein referred to as NILM/HR HPV double negatives) were also referred to colposcopy in order to obtain colposcopy/biopsy data from all populations undergoing screening. Table 2 below summarizes colposcopy referrals in clinical studies that supported HPV device approval:

Table 2: Populations referred to colposcopy

| Cytology Result | HPV result (investigational) | HPV result (FDA-approved test) | Refer to colposcopy |
|------------------------|-------------------------------------|---------------------------------------|----------------------------|
| ASC-US or greater | Any | Any | Yes |
| NILM | Neg | Neg | Random subset (e.g., 5%) |
| NILM | Pos | Neg | Yes |
| NILM | Neg | Pos | Yes |
| NILM | Pos | Pos | Yes |

For details regarding age groups of these populations, please see section IV.A.

The number of CIN2+/CIN3+ in the random subset of NILM/HPV double negative women was used to extrapolate the projected number of CIN2+/CIN3+ women in the entire NILM/HR HPV double negative population. This is referred to as the verification bias adjustment.

The colposcopy visit occurred within approximately three months from the first cytology visit. The colposcopy/biopsy procedure was conducted under a standardized procedure outlined in the sponsor's clinical protocol. Histologic interpretation of the biopsies obtained during colposcopy was performed by a centralized three expert pathologist review (CPR) panel. Each slide was independently reviewed by two expert pathologists participating in the panel; if the two pathologists assigned the same diagnosis, that is the final diagnosis for that slide. If differing diagnoses were made, then a third pathologist would independently read the slide and provides a diagnosis. The majority diagnosis from the three pathologists then served as the final diagnosis for those slides. If there was no agreement between

pathologists after the third pathologist review, all three expert pathologists conferred to reach a consensus diagnosis.

B. Panel Topic 1: Study Design

Clinical studies supporting HR HPV device approvals have generally been cross-sectional prospective studies that enroll participants from screening populations in geographically diverse sites within the United States. These studies have enrolled anywhere between 11,000-45,000 women, depending on the specific indications the manufacturer is seeking and whether the study is to support a new device or an expansion of indications for an already approved device. These studies have been generally designed to enroll at least 70 women with CIN2+ in the ASC-US population. While an endpoint of CIN3+ is the more scientifically accurate surrogate endpoint for cervical cancer, the prevalence of CIN3+ is around half of that for CIN2+ lesions, which is already very low in the screening population. In a study comparing two of the largest US clinical practice datasets, the risks of CIN2+ and CIN3+ over a 3 year period in the overall population ranged from 1.12-1.39% for CIN2+ and 0.44-0.61% for CIN3+ [19]. Designing a prospective study for the detection of CIN3+ at baseline would involve enrolling an extremely large number of subjects. Therefore, the FDA accepts the use of CIN2+ as the primary endpoint to determine the size of clinical studies.

However, such study designs to support HR HPV device approvals are anticipated to become less tenable as HPV vaccination lowers the prevalence of HR HPV infection and, accordingly, HR HPV-associated disease. A 60-70% drop in HR HPV prevalence, as well as a reduction in CIN2+ prevalence, has been reported in women who were vaccinated in adolescence [20-22]. Furthermore, between 2006 and 2014, a statistically significant 34% drop in prevalence of vaccine-targeted HR HPV types was reported amongst non-vaccinated women, indicating the start of herd immunity [20]. These data were reported prior to the approval of the 9-valent vaccine, although similar reductions in prevalence for the additional genotypes in the 9-valent vaccine are anticipated. The effect of vaccination will be especially problematic for designing studies that evaluate devices offering extended genotyping capabilities beyond types 16 and 18; in these cases, not only will it be difficult to obtain a sufficient number of study endpoints, but it will also become increasingly difficult for manufacturers to obtain a statistically sound number of HR HPV infected specimens for each genotyping result output.

Additionally, as more data has accumulated confirming the high regression rates of CIN2 lesions, especially for younger women [12, 23-29], it has increasingly become difficult to justify the inclusion of CIN2 lesions in the primary endpoint for cervical disease. These studies have indicated that almost half of all CIN2 lesions will regress spontaneously, with lesions that are negative for HPV types 16 and 18 having the highest rates of regression [25, 29]. For this reason, performance using the CIN3+ endpoint has also been evaluated in regulatory submissions; however, even with studies as large as 10,000 women, the incidence of CIN3+ in the study is typically too low to draw statistically meaningful conclusions regarding device performance. In a recent FDA approval, out of 8,088 women enrolled in the Adjunct intended use population, only 7 women with CIN3+ were identified. As mentioned above, with increased vaccination uptake, it is expected that the prevalence of CIN3 will decrease, resulting in even fewer samples from women with CIN3+ to evaluate. The proportion of CIN2 lesions that are associated with HR HPV infections will also decrease, leaving a higher proportion of benign, regressive lesions (those that are caused by low risk, noncarcinogenic HPV infections) in the CIN2+ endpoint. For these reasons, finding a way to enroll a sufficient number of women with CIN3+ histology will be critical in accurately establishing HR HPV device performance.

Therefore, in order for new HPV devices to be evaluated efficiently and rigorously, the FDA is seeking input for alternative clinical study design strategies that can enroll a sufficient number of HR HPV positive and CIN3+ women in clinical studies to support approval. Below are several proposals that the FDA is requesting that the panel consider, in addition to open discussion of different options not outlined here.

Proposal 1: Enrichment Studies using Specimens Collected from Referral Populations.

One potential approach could be to enrich a prospective study with patients from referral populations, (e.g., colposcopy clinics), which consist of women with abnormal screening results who were referred for colposcopy. The prospective study could include screening sites that also perform colposcopy. Because referral populations consist of women with abnormal cytology and/or repeat HR HPV positivity, obtaining additional HR HPV positive specimens and enrolling women with CIN3+ lesions will be more likely than if solely recruiting from screening populations. However, one risk in adopting this approach is that this referral population is not completely representative of the entire screening population. Namely, this enriched population would exclude: 1. women who are negative by the standard of care (SOC) HPV device that was responsible for referral but would have been positive by the investigational device, 2. women who are negative by both the SOC and investigational HPV devices but may have underlying disease, and 3. Women with NILM cytology who are HR HPV positive but have not yet been rescreened at the 1 year mark. This could lead to performance estimates that are not entirely representative of the intended use population. Another risk associated with this approach is that colposcopy referrals are not conducted under a standardized clinical protocol. Referrals would instead be made at the physician's discretion based on SOC HPV and cytology results.

Proposal 2: Archived Specimens.

Another approach could be the use of archived liquid-based cytology (LBC) specimens from large, well-characterized patient populations (e.g., biobanks, academic/hospital institutions, state run public health laboratories, etc.). These could be retrospective specimens or specimens that were archived after prospective collection. For this approach, testing of archived specimens associated with patient medical information, including SOC screening results, histology diagnoses, follow-up information and test results, and, potentially, additional medical history, could be conducted. The HR HPV device results, as well as the patient data in medical records, could be used to evaluate clinical performance. One major benefit to this approach is the ability to obtain a sufficient number of HR HPV positive specimens and include women with CIN3+ lesions without the burden and time of an extremely large prospective study. Additionally, if, during the course of FDA's review, it is determined that additional longitudinal data would be useful to address any issues, this could be accessed through linked patient medical records.

However, as mentioned in the previous section, risks associated with a lack of a standardized colposcopy protocol, as well as the colposcopy referral population potentially not being completely representative of the entire intended use population, would be present. Another risk is that there are likely limited resources for archived specimens and having these specimens as the sole source of clinical data may be problematic.

Proposal 3: Capping the vaccinated population

Another option would be to cap the number of vaccinated women in the clinical study such that the majority of enrolled subjects will be nonvaccinated. The benefit to this approach is a higher probability of recruiting subjects with CIN3+ in a nonvaccinated population. However, one risk associated with such a protocol is that HPV device performance against a histological endpoint will be primarily driven by the

genotype distribution found in non-vaccinated communities, which may be different than in vaccinated communities. This risk could be mitigated by determining the HPV type as part of the clinical study in order to demonstrate that there is adequate representation of all genotypes and ensure that the device can adequately detect both vaccine-targeted and non-targeted infections (please see section IV.C. for a more detailed discussion regarding this approach).

The FDA is requesting that the panel discuss the benefits and risks of these proposals, possible approaches to mitigate the associated risks, and whether these proposals are acceptable paths forward for new devices. The panel will also be asked to propose any additional approaches not described above, along with their risks and benefits.

C. Panel Topic 2: Colposcopy Referral Protocol in Clinical Studies

Clinical studies supporting HR HPV device approvals are different from those supporting other microbiological devices in that the comparator has historically been histological diagnosis from a biopsy taken during the colposcopy visit, rather than detection of a pathogen. Because of the need for a biopsy to assess the presence of cervical dysplasia, the clinical study design for HR HPV devices has needed to simultaneously consider two factors. First and foremost of these is patient safety, in terms of which subjects are asked to undergo the colposcopy/biopsy procedure. The cervical cancer screening recommendations dictate the preferred and acceptable management strategies for women with cervical abnormalities (e.g., colposcopy, follow-up, etc.) and guide who should undergo these procedures. The second factor is keeping the study unbiased by ensuring that all patient populations with different HPV/cytology result combinations have colposcopy/biopsy data for clinical evaluation. Using the clinical study design described above in Section III.A, it is possible to adjust for verification bias and calculate clinical performance, since all combinations of cytology/HPV test results would have colposcopy data.

The first prospective clinical study conducted to support HR HPV device approval for routine screening took place when the 2006 consensus screening/management guidelines were in effect [30]. However, since then, guidelines have evolved, with the trend moving towards less aggressive management (i.e., increased intervals between visits, follow-up rather than immediate colposcopy or treatment). This is for two reasons. First, several large-scale studies conducted in the US and overseas reported long-term risks of CIN2+, CIN3+, and cervical cancer for women with different HPV and/or cytology screening result combinations [31-36]. These studies demonstrated the high negative predictive value of HR HPV testing and indicated that incorporating HR HPV testing in screening and ruling out an HR HPV infection could lead to less frequent screening visits without putting patients at additional risk for cervical cancer. Second, studies where women with CIN2 (and even CIN2/3) lesions were observed over time (rather than immediately treated) reported high spontaneous regression rates of these lesions over 1-2 years, especially for younger women [25], further supporting lengthening intervals. Because CIN2 is the current threshold for treatment, screening too frequently before these lesions have the time to clear on their own would lead to potential overtreatment of otherwise healthy women using procedures, such as LEEP (loop electrosurgical excision procedure), that have been associated with poor reproductive outcomes (e.g., pre-term birth) [37].

Because of these changes in screening and patient management practices, the FDA is re-assessing clinical study protocols supporting HR HPV device approval to ensure that colposcopy referral protocols of enrolled subjects do not subject women to unnecessary risk. Studies have been referring ASC-US/HR

HPV double negative⁵ and a subset of NILM/HR HPV double negative women to colposcopy, where a biopsy (either directed or random, if no lesions are present) and/or, at times, an ECC was obtained. However, current screening and management guidelines [2, 33, 34, 38] indicate that these populations should not be referred to colposcopy, but rather, be re-screened in 3-5 years due to: 1. The established high negative predictive value of HPV testing in ruling out disease over this time period, as evidenced by the extremely low 3 and 5 year risks for CIN2+ and CIN3+ [35, 36, 39-42], and 2. The need to allow sufficient time to pass for benign infections/lesions to clear/regress. Referring these women to colposcopy could potentially expose these low risk populations to harms they would not normally be exposed to under standard of care (e.g., due to the colposcopy procedure itself or potential overtreatment of transient lesions).

Below is a discussion of the benefits and risks of referring these populations to colposcopy for purposes of a clinical study supporting HR HPV device approval. The FDA is requesting that the panel review these benefits and risks, as well as any other benefits and risks not described below, and discuss if having these populations undergo a procedure they would not normally undergo under standard of care is imperative for the unbiased review of an HR HPV device.

Benefits of colposcopy referral:

Referring ASC-US/HR HPV double negative and a subset of NILM/HR HPV double negative women to colposcopy permits an unbiased clinical study design by ensuring that all populations with different HPV/cytology result combinations have colposcopy/biopsy data. The risk of CIN2+ and CIN3+ in the subset of women with NILM/HR HPV double negative results is extrapolated to the entire population, where the projected number of CIN2+ and CIN3+ cases is imputed. This allows for the unbiased calculation of performance metrics (sensitivity, specificity, predictive values, likelihood ratios) of the investigational HR HPV device. If certain populations of women are not referred to colposcopy, then assumptions would have to be made regarding the disease status of those women; namely, they would automatically be considered as “non-diseased.” However, it is possible that NILM/ASC-US women who are negative by both the FDA-approved and investigational HR HPV devices can still have precancer. Cervical cancers that have been negative by a HR HPV device have been reported [5, 43]. This may lead to a potential bias in the performance estimates, since negative device results in women with precancer would not be identified without referral to colposcopy. The presence of colposcopy/biopsy data for all populations of women ensures that the HR HPV device is optimized to detect precancers in all women in the screening population, and a better understanding of an HPV device’s ability to do so can be gained.

Risks of referring these populations to colposcopy:

There are two types of risks associated with referring NILM/HR HPV double negative and ASC-US/HR HPV double negative women to colposcopy: physical risks to the patient and the underestimation of test sensitivity.

Physical Risks to patient: As mentioned above, these populations of women would not be referred to colposcopy under standard of care. One of the main risks in referring these women to colposcopy is the identification of benign lesions destined to regress, which could lead to unnecessary treatment, such as LEEP, which has been associated with poor reproductive outcomes, including pre-term births and low birth weights [37]. Other associated risks with the colposcopy/biopsy procedure include pain, discomfort, vaginal bleeding/discharge, and emotional distress for the patient [44].

⁵ HR HPV double negative refers to being negative by *both* the FDA-approved and investigational devices.

Underestimation of test sensitivity: Because the prevalence of precancer at baseline in NILM/HR HPV double negative and ASC-US/HR HPV double negative women is extremely low, as evidenced by long-term studies following these women over time, it can be presumed that the majority of colposcopy procedures will have no clinical findings and those that do reveal lesions have a high likelihood of being a transient lesion due to benign abnormalities caused by low-risk HPV infections destined to clear (since these women would be HR HPV negative by an FDA-approved test, which all have NPVs above 99% for CIN3+, and HR HPV negative lesions almost always regress [25]). If women with transient lesions are considered as being positive for precancer in the evaluation of an HR HPV device, this may lead to an underestimation of device sensitivity for true precancer.⁶ Evaluating the clinical data in this way may also result in manufacturers seeking to optimize devices to detect infections that are causing lesions destined to regress, in an effort to achieve better sensitivity estimates (e.g., designing primers to cross react with low-risk types, choosing a cutoff that is more likely to pick up non-clinically relevant infections). In clinical practice, this could lead to otherwise healthy women being referred unnecessarily to colposcopy and potentially experiencing overtreatment.

No screening result/result combinations can ensure a woman has a 0% risk of harboring precancer/cancer, mainly because it is impossible to know at a given timepoint whether a lesion will regress or progress, even with the best pathological evaluations. Therefore, the decision to either refer or not refer these women to colposcopy has both benefits and risks that will affect the enrolled subjects and HR HPV device performance estimates in different ways. Weighing these benefits and risks, the FDA seeks the panel's input in assessing whether it is necessary to have ASC-US/HR HPV double negative and a subset of NILM/HR HPV double negative women undergo the colposcopy/biopsy procedure for clinical studies supporting HPV device approval. Additionally, please discuss whether possible alternative proposals that do not involve referring these populations to colposcopy exist that would maintain an unbiased study design. Please discuss these issues separately for the NILM/HR HPV double negative and ASC-US/HR HPV double negative populations.

IV. HPV Device Evaluation

A. Performance Evaluation: Background

As mentioned in section II.B, there are three main indications for which sponsors have successfully sought a claim for HR HPV devices: (i) ASC-US triage for women 21 years and older, (ii) adjunctive testing for women 30 years and older, and (iii) primary screening for women 25 years and older. The introduction of each indication was based on cervical cancer screening practices at the time of the device approval, namely according to the populations (defined by age and sometimes cytology result) that would be tested for HR HPV. HPV triage of women aged 21 and older with ASC-US cytology was the first main approved indication for use for HPV testing. In 2003, it was determined by a Microbiology Devices Advisory Panel that the benefits of incorporating HR HPV testing in routine screening (i.e., in asymptomatic women) outweighed the risks, so an additional adjunctive indication was approved for women 30 years and older. This resulted in two populations in which subsequent HR HPV tests were evaluated by FDA during premarket review of additional applications: ASC-US women aged 21 years and older and NILM women aged 30 years and older. The evaluation of the HR HPV device performance has

⁶ Previously sponsors have submitted typing for low-risk HPV in their PMAs as additional information to help clarify CIN2+/CIN3+ lesions missed by the device. However, this has not recently been a part of the main analysis. Please see section IV.C.4 for discussion on HPV typing in the comparator.

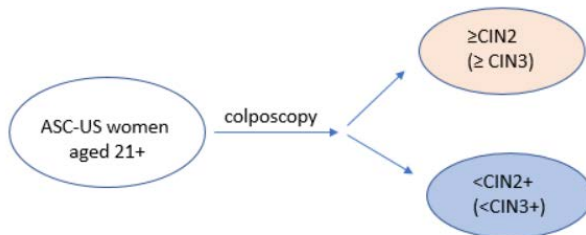
been slightly different in these two populations due to the different indications: HR HPV results in the ASC-US triage indication aid in colposcopy referral, while HR HPV results in routine screening are used adjunctively with cytology results to assess risk of cervical disease. Several years later, it was observed through large scale clinical studies/trials that a negative HR HPV result not only has predictive value for ruling out cervical precancer, but triage of HR HPV positive test results by genotyping and cytology helps further decrease unnecessary colposcopy referrals. This led to the approval of the most recent indication for primary HPV screening. This has resulted in three different indications for which HR HPV devices have been reviewed for FDA approval, with each having had device performance evaluated using different analyses.

Below are descriptions and flow charts of the performance evaluation analysis that has been conducted for approval of each indication:

1. ASC-US triage indication

Clinical Study Population: This population consists of all women with ASC-US cytology who are aged 21 years and older. Because every woman with ASC-US cytology is referred to colposcopy in the clinical study regardless of HR HPV result, there is histology data for all women (Figure 1):

Figure 1: ASC-US triage indication flow diagram



Performance Evaluation: All women who have CIN2+ (or CIN3+, depending on the analysis) were defined as “comparator positives” (shaded in peach in Figure 1 and Table 3) and the device sensitivity calculations were performed in these populations. Women whose biopsy diagnoses were <CIN2+ (or <CIN3+) were “comparator negatives” (shaded in blue) and device specificity calculations were performed in these populations. The 2X2 table below was then generated to evaluate clinical performance parameters, including sensitivity, specificity, risk of CIN2+/CIN3+ associated with different HPV result outputs (including positive and negative predictive values), as well as positive and negative likelihood ratios⁷ (PLR and NLR, respectively):

⁷ Positive Likelihood Ratio, or PLR, indicates how many times more likely the subjects with disease are to have a positive result than subjects without disease. Negative Likelihood Ratio, or NLR, indicates how many times less likely the subjects with disease are to have a negative result than subjects without disease. For more information, please see <http://www.thennt.com/diagnostics-and-likelihood-ratios-explained/>.

Table 3: 2X2 table for ASC-US triage performance evaluation

| | CIN2+ | <CIN2 |
|----------------|-------|-------|
| HPV Device POS | a | b |
| HPV Device NEG | c | d |

Example performance calculations:

Sensitivity: $[a/(a+c)]*100$

Specificity: $[d/(b+d)]*100$

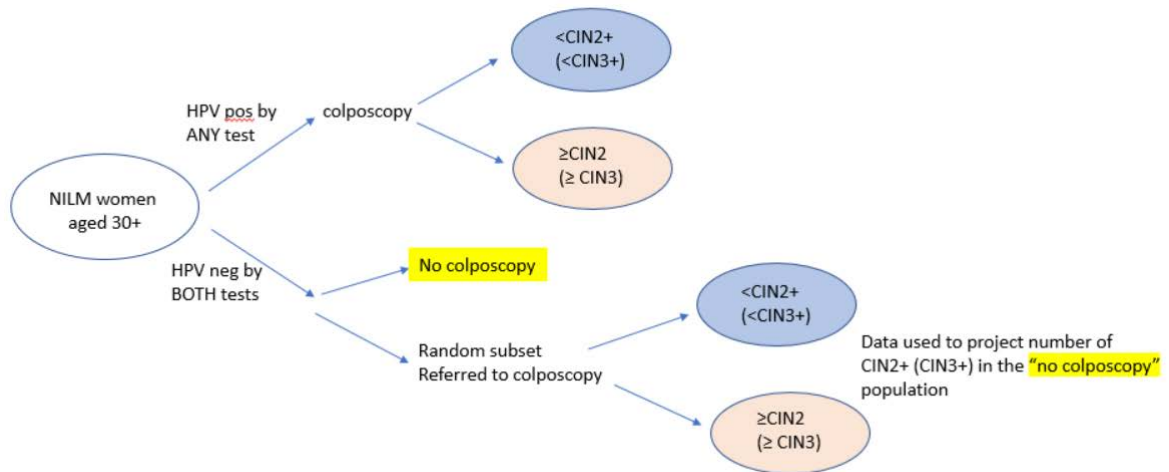
The above table was also generated for CIN3+ and <CIN3+ during device performance evaluation.

2. *Adjunct indication*

Clinical Study Population: The clinical study population that was evaluated for this indication consisted of women aged 30 years and older with NILM cytology. This is because at the time this indication was first approved, the screening guidelines dictated that the only women 30 years and older whose management was affected by a HR HPV test result were ones with ASC-US and NILM cytology; all women with >ASC-US cytology were immediately referred to colposcopy. Since ASC-US women 30 years and older were already included in the ASC-US triage intended use population, that left NILM women 30 years and older as the evaluable population for the Adjunct indication. However, consensus screening guidelines changed in 2012 to include HR HPV testing of women with LSIL cytology results as an acceptable strategy to rule out immediate colposcopy [34]. For this reason, all subjects aged 30 and above (rather than just women who are NILM 30 years and older) were enrolled in studies to monitor for safety signals, although performance was typically evaluated and presented only in the NILM population.

As stated in section IIIA, women with NILM cytology who were HR HPV positive (any genotype) by either the FDA-approved or investigational device were referred to colposcopy, as well as a subset of NILM/HR HPV double negative women, for purposes of the verification bias adjustment. The number of CIN2+/CIN3+ in this subset population was used to project the number of CIN2+/CIN3+ women in the entire NILM/HR HPV double negative population. This is summarized in Figure 2:

Figure 2: Adjunct indication flow diagram



Performance Evaluation: Similar to the ASC-US triage performance evaluation, women with CIN2+ (or CIN3+) were considered “comparator positives” and women with <CIN2+ (or <CIN3+) were considered “comparator negatives”. For the adjunct indication, the 2X2 Table 4 below was generated, and the main performance metrics evaluated included the absolute risks for CIN2+ and CIN3+ that were associated with each HPV result output (the proportion of women with a particular test result that had CIN2+/CIN3+), as well as the relative risks for different result outputs (ratio of absolute risks for different result outputs). Other parameters may also be evaluated depending on the available data.

Table 4: 2X2 table for Adjunct performance evaluation

| | CIN2+ | <CIN2 |
|----------------|-------|-------|
| HPV Device POS | a | b |
| HPV Device NEG | c | d |

Example performance calculations:

Absolute risk if Device POS result: $[a/(a+b)]*100$

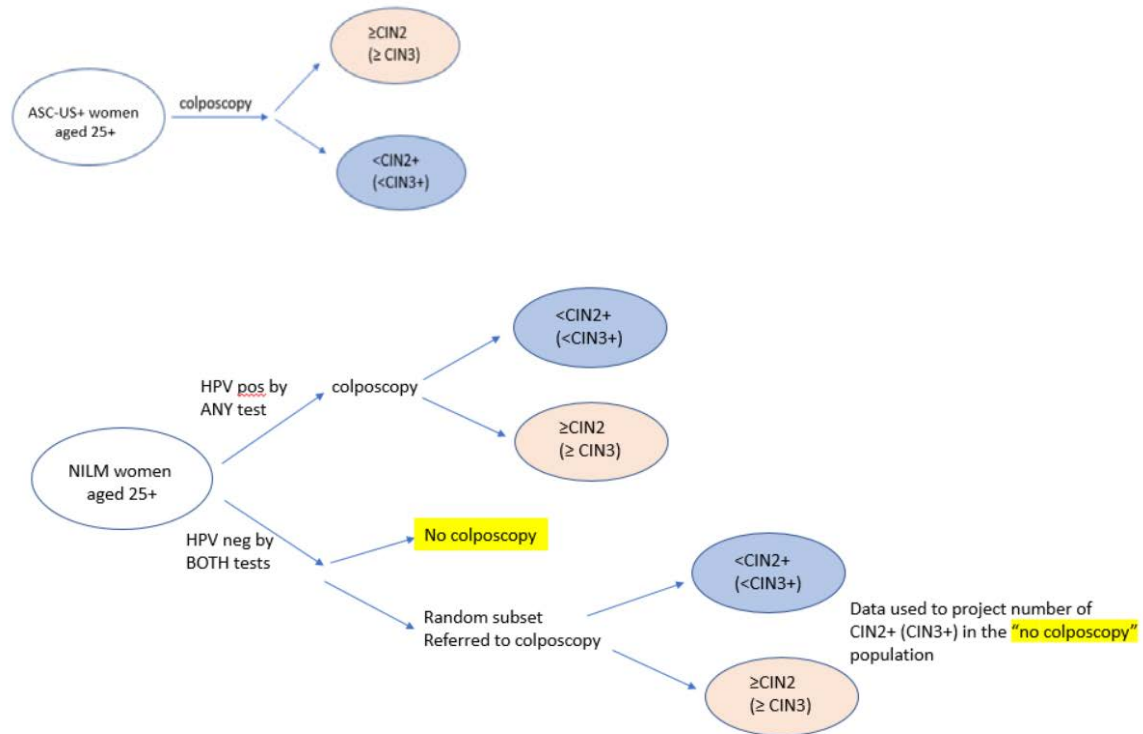
Absolute risk if Device NEG result: $[c/(c+d)]*100$

The above table was also generated for CIN3+ and <CIN3+ during device performance evaluation.

3. Primary screening indication

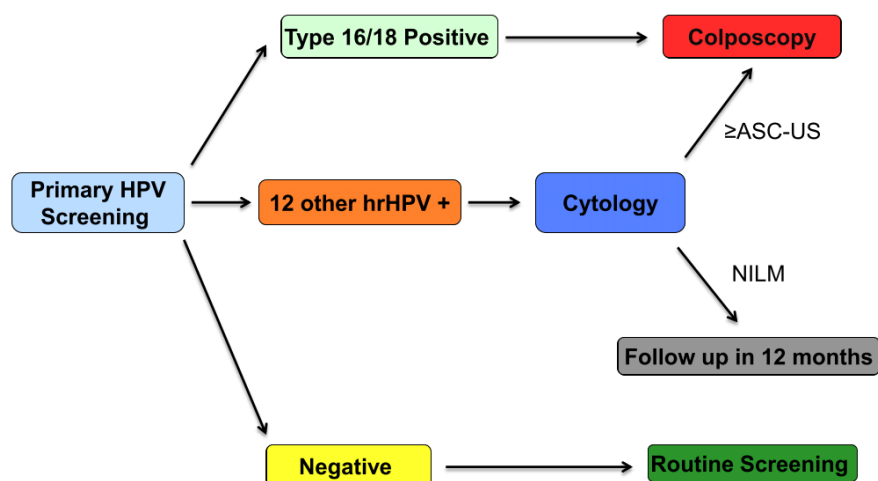
Clinical Study Population: The clinical study population for the primary screening indication included all women aged 25 years and older presenting for routine screening. All women with abnormal cytology results and/or positive HPV results by either test were referred to colposcopy, as well as a random subset of NILM/HPV double negative women (Figure 3):

Figure 3: Primary screening indication flow diagram



Performance Evaluation: Rather than assessing performance of the *HR HPV device* itself, the performance of the *primary screening algorithm* using the investigational device was evaluated and compared to screening with cytology alone or with co-testing. This was to determine the risks and benefits of screening women using primary HPV screening versus cytology alone or with co-testing. In the primary screening algorithm, women are first screened with a HR HPV device. Women who are HR HPV negative return to routine screening. Women who are HPV 16/18 positive are referred directly to colposcopy. Women who are HR HPV positive but 16/18 negative (i.e., 12 other HPV positive) are triaged by cytology, where those who have \geq ASC-US cytology are referred directly to colposcopy, and those who have NILM cytology are followed up according to the physician's discretion (current guidelines state 1-year follow-up [1]). Figure 4 is a flow chart from the 2015 interim clinical guidelines [1] depicting the primary screening algorithm:

Figure 4: Primary screening indication algorithm



Unlike the ASC-US and adjunct indications where the “device positive/negative” results were used to evaluate clinical performance, the *algorithm positives and negatives* were used to calculate performance of the HR HPV device used in the primary screening algorithm. An *algorithm positive* was defined as a woman who was referred directly to colposcopy according to the particular screening algorithm being assessed. In contrast, *algorithm negatives* were women who are not referred directly to colposcopy. Table 5 below depicts the populations considered “positive” (sent immediately to colposcopy) and “negative” according to the primary screening algorithm:

| Table 5: Definition of Positive and Negative Results According to the Primary Screening algorithm | Cytology | | | |
|---|----------|---------|-------|------|
| | ASC-US | >ASC-US | NILM | |
| | | | 25-29 | > 30 |
| HPV 16/18 Pos | | | | |
| 12 Other HR HPV Pos | | | | |
| HR HPV Neg | | | | |

Green denotes positive and gray denotes negative results. Positive results are defined as women sent immediately to colposcopy according to the primary screening algorithm. Negative results are defined as women who are not sent immediately to colposcopy according to the primary screening algorithm.

The 2X2 table below (Table 6) was generated, and performance metrics of the *algorithm* were calculated, including sensitivity for CIN2+ /CIN3+ (the proportion of women with CIN2+ /CIN3+ who were referred to colposcopy), specificity (the proportion of women <CIN2+ /<CIN3+ who were not referred to colposcopy), positive/negative predictive values, positive/negative likelihood ratios, and absolute risks for different result outputs (including certain HPV and cytology result combinations):

Table 6: 2X2 table for Primary screening performance evaluation

| | CIN2+ | <CIN2 |
|---|-------|-------|
| Primary Screening Algorithm positives (green boxes from Table 5) | a | b |
| Primary Screening Algorithm negatives (gray boxes from Table 5) | c | d |

Example performance calculations:

Sensitivity: $[a/(a+c)]*100$

Specificity: $[d/(b+d)]*100$

The above table was also generated for CIN3+ and <CIN3+ during device performance evaluation.

For a detailed summary regarding the parameters assessed in the primary screening indication, please refer to the [executive summary in the 2014 panel pack](#).

B. Panel Topic 3: Indications for use

The approaches to evaluating HR HPV devices described above were originally developed due to the importance of establishing HR HPV device performance in the context of then-current clinical practice. However, this additive approach in evaluating HR HPV devices, with successive indications reflecting new ways HR HPV results were used in screening, has at times resulted in redundant analyses being conducted (by sponsors) and reviewed (by the FDA), since there are some populations of women that overlap between the indications (i.e., NILM women 30+ fall into both the adjunct and primary screening indications, ASC-US women 25+ fall into both the ASC-US triage and primary screening indications). Furthermore, clinical guidelines are continually changing, and device indications that are directly tied to clinical guidelines at the time of approval present a challenge in maintaining consistency and relevance as newer devices are submitted for approval. Below are examples of some of the challenges FDA has faced for each of the indications:

1. **ASC-US Triage – triage to colposcopy not recommended for women aged 21-24:** The ASC-US triage indication has necessitated a clinical study population that included ASC-US women aged 21 and older who would be triaged to colposcopy. However, consensus cervical cancer screening guidelines no longer recommend HPV triage of ASC-US women aged 21-24 to guide colposcopy referral [34]. Assessing HR HPV device performance in this population may be inappropriate and extra burdensome because performance in this population may no longer be clinically relevant.
2. **Adjunct – HR HPV testing now acceptable for women aged 30+ with LSIL⁸ cytology:** The adjunct indication was evaluated in a population of NILM women aged 30 years and older because at the time this indication was approved, the only women whose management was affected by HR HPV results were those with ASC-US and NILM cytology (ASC-US women were included in the ASC-US triage population). However, in 2012, consensus guidelines recommended HR HPV testing in women aged 30+ with LSIL cytology to rule out the need for

⁸ LSIL (low-grade squamous intraepithelial lesion) in this context refers to a cytology diagnosis according to the Bethesda system [45], not to be confused with LSIL histology according to the LAST terminology [46].

immediate colposcopy as an acceptable approach. Therefore, the evaluation in the NILM population leaves out an additional population that can currently be co-tested for HPV.

3. **Primary Screening – potential future changes in clinical management and triage strategies:**
While the primary screening indication allows for more flexibility in the intended use population than the ASC-US triage and adjunct claims (i.e., all patients aged 25 and older), this indication dictates a very specific clinical action depending on the combination of HPV and cytology results (i.e., who should be referred to colposcopy). While this strategy is currently deemed acceptable by the clinical community [1], it is possible that this may change in the future as vaccination will decrease the prevalence of the targeted genotypes (and CIN3+ lesions caused by these genotypes), leading to potential changes in management strategies. Furthermore, it is anticipated that new triage methodologies will soon be developed [47]. Specifically stating the follow-on triage test (e.g., cytology) in the indications for the HR HPV device would likely require manufacturers of approved HPV devices to submit a new PMA supplement any time a new triage method is approved, in order to update their intended use/indications (provided the HPV device manufacturer would *want* to update the indications). This would create a large administrative and regulatory burden on both industry and the FDA, considering that the HR HPV device itself would not have changed.

One approach to address these challenges could be to continually modify the indications for use⁹ and required data analyses as clinical guidelines change. However, this approach may be problematic for several reasons. First, the amount of time between initial FDA/industry discussions regarding the studies/data needed to support approval of a particular indication and the actual approval of the device is often several years. If clinical guidelines change during that time, it may not be feasible for FDA to request that industry change their clinical plan to support a modified indication, especially if the clinical study has already commenced. Second, continually changing HR HPV device indications to be in line with current clinical guidelines could create a lack of consistency in the FDA’s expectations for approval, which could create confusion for manufacturers of new devices. Third, the indications for use remain in the device’s package insert as long as the device is on the market, even if clinical guidelines change. Having indications tied to specific guidelines and management strategies runs the risk of a marketed device being indicated for a use that is no longer recommended by the clinical community.

The FDA is therefore soliciting feedback on a proposal to simplify the recommended indications for use statements and make them independent of potential changes in clinical guidelines. This proposal includes the following two major changes:

1. Simplifying the indications to encompass one general screening population

With primary HPV screening approved for all women aged 25 years and older, there is now an opportunity to support the use of a screening population that consists of one intended use population that encompasses “all women presenting for routine screening” rather than specifying cytology groups (i.e., ASC-US, NILM, etc.) or age groups (i.e., 21+, 25+, 30+). A benefit to this approach is that the enrolled population would be relevant to current screening practices. If guidelines in the future dictate changes in which populations should be tested for HPV (e.g., based on age, cytology result, etc.), then this generalized intended use wording would encompass these changes. Please note that although the intended use population would be generalized to all patients, HPV device performance in specific sub-

⁹ Modifications to a device’s indications for use may require a new PMA or PMA supplement

populations (e.g., different age groups, cytology populations, etc.) could also still be evaluated and presented in the package insert. Furthermore, explicit inclusion/exclusion criteria (regarding age groups, etc.) for the clinical studies could still be maintained with such a generalized intended use statement.

2. Removing reference to specific triage tests and clinical actions:

The second change would be to provide more generalized language to describe how positive results should be followed-up/triaged, rather than stating the specific triage test and patient management action based on the combination of results. This would prevent the need for each HPV device manufacturer to submit a new regulatory application to the FDA to update their indications any time a new triage method is introduced in clinical practice. Furthermore, since this wording would not include specific patient management strategies regarding which population should be referred to colposcopy, clinical practice recommendations would remain up to the clinical community and guidelines committees rather than designated by an FDA-approved indication.

Taking these two proposals together, the following is an example of such a proposed intended use:

[Description of technological characteristics of test and trade name] is a qualitative in vitro test for the detection of Human Papillomavirus in cervical specimens collected by a clinician using [collection device/media]. This assay should be used to test women presenting for routine cervical cancer screening to assess the risk for cervical dysplasia and cancer. Women who test positive or negative for the HR HPV types [list types detected from test] should be triaged/followed-up in accordance with professional guidelines, the physician's assessment of screening and medical history, and other risk factors.

The FDA is requesting the panel's input regarding the risks and benefits of supporting simplifying the indications for use to encompass one general screening population and removing reference to specific triage tests and clinical actions, as well as whether this could be an appropriate path forward. In addition, please discuss any potential risk mitigation measures for this new generalized intended use (e.g., populations that should be included/excluded for clinical studies, warnings/limitations in labeling, sub-analyses to be performed, etc.). The FDA also welcomes alternative proposals for intended use labeling that would address the issues outlined above.

C. Panel Topic 4: Data Analysis to Support the New Indications for use

Data analyses that the FDA has recommended manufacturers submit for premarket approval of their device has been dependent on the intended use/indications for which the manufacturer was seeking a claim. If the indications for use were changed as described above, then the data analyses needed to support this more generalized intended use would have to be modified to be more generalizable across all screening populations. To do this, FDA proposes to change the paradigm of HR HPV device evaluation and instead adopt an approach that focuses on comparing HR HPV assays, since to date, five HR HPV assays have been approved. There are two potential approaches for conducting these evaluations: 1. without a clinical endpoint comparator (i.e., no histological diagnoses), and 2. with a clinical endpoint comparator.

Proposal 1: Performance evaluation without a clinical endpoint comparator

This approach would involve comparison of the investigational device to other methods detecting HPV nucleic acid without a clinical endpoint comparator (as is often done with other microbiological assays).

One potential way to do this is to use three FDA-approved devices to create a “composite comparator” and defining the majority result (e.g., two out of three) as the comparator result. Positive and negative agreement between the investigational device and the composite comparator would then be assessed. In this approach, colposcopy referral and histological adjudication would not be a part of the clinical protocol (although standard of care for patient management would have to be followed for women enrolled in the study). Rather, the study would be designed such that multiple FDA-approved devices could be used to test the collected specimens.

This approach has several benefits. First, this type of analysis would likely result in substantially smaller clinical studies (potentially several thousand vs. tens of thousands of subjects as has been needed in the past), which would ensure more timely access to approved devices. Second, a greater understanding of how results from multiple FDA-approved devices from a single patient compare will be gained. This could provide the clinical community with valuable information when deciding which devices to use in clinical practice.

However, one risk associated with this approach is that it may be difficult to ensure that the enrolled population includes representation of women who are at highest risk of harboring an early precancer since the vast majority of women in the intended use population would be healthy. If this approach was adopted, it would be important to clearly outline the specific populations that should be included in the evaluation to ensure that the smaller sample size did not mean that important populations were missing from the analysis. Another risk is that, without histology, it would be difficult to assess the clinical relevance of any discrepancies between the investigational device and comparator (although utilizing three devices in the comparator and using the majority result may mitigate cutoff-related variability between individual devices). Lastly, devices have differing result outputs. It is expected that in the future, devices offering genotyping result outputs that are not present in currently approved HR HPV devices will be submitted for regulatory approval. Therefore, this type of analysis could only be conducted for outputs that all devices (the investigational device and comparator devices) have in common. For genotyping outputs not present in any FDA-approved HPV device, a validated HPV typing technology would have to be used to assess detection, such as DNA amplification followed by sequencing.

It is important to note that if this approach is utilized, then the indications for use for the device would not include mention of assessing for risk of cervical dysplasia or cancer. The indications for use of a device are supported by the data generated in the study, and the data in this case would not include histology to make that assessment.

As assessing performance against a composite molecular comparator would represent a paradigm shift in the way HR HPV devices are evaluated, the FDA would like the panel to discuss the benefits and risks of using a composite molecular comparator without histology to evaluate the safety and effectiveness of an investigational device and whether this approach would be an acceptable path forward. If so, please discuss which study populations should be required and whether they should be prospectively enrolled or analyzed as part of the overall dataset; the performance parameters to be evaluated as well as required acceptance criteria; and any other pertinent considerations relevant to this topic that have not been mentioned.

Proposal 2: Performance assessment with a clinical endpoint comparator

This alternative approach would assess the relative clinical performance between an investigational HR HPV device and an FDA-approved device against a clinical endpoint comparator that includes histology

(please see section IV.D for further discussion on the clinical endpoint comparator). In this comparison, performance parameters such as sensitivity/specificity/etc., would be calculated for both the new and FDA-approved devices using the clinical endpoint comparator. Then, the ratios of these estimates would be calculated to determine how performance of the new device compares to performance against an already approved device. (For an example of these “relative performance” calculations, please see [48]). The most recent FDA approval for a device of this type utilized this analysis for an approved test seeking to expand claims to include an additional collection media (Table 7):

Table 7: Ratio (Adjusted) of Sensitivities and (1-Specificity) Between the new test in SurePath and an FDA-approved test using PreservCyt

| Disease Endpoint | Sensitivity Ratio (95% CI) | (1-specificity) Ratio (95% CI) |
|---|----------------------------|--------------------------------|
| ≥CIN2 | 1.01 times (0.86, 1.22)* | 1.19 times (1.13, 1.25) |
| ≥CIN3 | 1.00 times (n/a)* | 1.18 times (1.13, 1.24) |
| *ratio of sensitivities indicates that sensitivities of two tests are similar | | |

The table is reproduced from the package insert for the cobas HPV test

This approach has several benefits. First, evaluating relative performance against a clinical endpoint (CIN2+, CIN3+, etc., as discussed in section IV.D) provides information regarding the clinical relevance of any differences between an investigational and FDA-approved device (i.e., if differences occur in women with or without precancer). Second, since safety and effectiveness of an investigational device would be evaluated relative to an FDA-approved, clinically validated molecular assay, this approach also adds a layer of objectivity to the evaluation, in which potential variances in primarily morphological assessments (i.e., cytology, histology) could be normalized.

However, one risk is that not all devices share the same result outputs. The relative performance described above could only be conducted for overall HR HPV positivity and for device result outputs that are common amongst the investigational and FDA-approved devices. For devices that do not share the same result outputs, FDA proposes to incorporate a validated PCR/sequencing technology to assess agreement in women with and without cervical dysplasia (see section IV.D).

The FDA is seeking the panel’s input on the benefits and risks of this approach, as well as several topics that may affect how this new approach may be implemented. The FDA is requesting that the panel discuss the performance parameters that should be evaluated as well as required minimum performance criteria (e.g., What should be the allowable margin of non-inferior performance when comparing assays? To what extent can analytical data contribute to these evaluations?). The FDA is also requesting that the panel discuss any potential sub-populations that should be evaluated in addition to the overall screening population (e.g., age groups, cytology groups, etc.), and whether these populations should be prospectively enrolled or analyzed as part of the overall data. Lastly, please discuss any other pertinent factors regarding this topic that have not been mentioned.

D. Panel Topic 5: Clinical Endpoint Comparator

If the panel determines that the comparator to assess device performance should include histology (in other words, a clinical endpoint comparator), then the next question is what clinical comparator should be used. While CIN2+/CIN3+ are the most commonly accepted clinical endpoints for cervical precancer and have been used as the reference comparator for all previous approvals, the FDA is seeking to determine if, for the purposes of evaluating the safety and effectiveness of a HR HPV device, additional

molecular information, such as HPV typing, should be incorporated into this histologic assessment to generate a mixed histological/molecular clinical endpoint comparator. This comparator would then be used to evaluate performance of the new and FDA-approved devices for the “relative performance” assessment (it should be noted that the HPV typing assay used in the mixed histologic/molecular comparator would not be the same FDA-approved device that is being used to assess the relative performance). As stated above in Section IV.C, not all devices share the same genotyping outputs, so evaluating performance relative to an FDA-approved device may not suffice for assessing detection of genotyping results that are unique to the new device (however, this may change in the future as additional devices are approved). Furthermore, even for devices that do share the same result outputs, characterizing the HPV genotype associated with lesions through HPV typing may still be important. This is based on two reasons:

First, current molecular technologies have enabled more sensitive detection of HPV nucleic acid in cervical specimens than earlier technologies, especially regarding mixed infections. This has allowed sponsors to provide more data to characterize the lesions from women in their studies. These data have shown that a proportion of CIN2+/CIN3+ lesions that are negative for HR HPV nucleic acid by multiple devices are positive for low risk, non-carcinogenic HPV [49]. This phenomenon has also been observed in unpublished data reviewed by the FDA. Including these lesions in the clinical comparator and penalizing devices for missing them (i.e., classifying negative results as “false negatives” even though by design these devices should not be detecting these infections) may encourage sponsors to optimize their devices to detect these infections in an effort to report better performance than an already FDA-approved device. This could lead to overtreatment of women with lesions caused by non-carcinogenic HPV types that are likely to regress spontaneously. This also indicates that even amongst CIN2+/CIN3+ lesions, discordant results between an investigational and FDA-approved device may have varying degrees of clinical relevance depending on the type of HPV infection (i.e., high-risk vs. low-risk) that is responsible for the lesion.¹⁰

Second, the uptake of HPV vaccination, especially the 9-valent vaccine, will lead to changing screening populations with regards to HR HPV prevalence that can affect device performance for both new and previously approved HPV devices. This is for several reasons:

1. *Proportion of remaining lesions that are HR HPV negative may increase:* The uptake in vaccination with the 9-valent vaccine will lead to a decrease in the lesions that are caused by the 7 HR HPV types that are responsible for 90% of cervical cancers worldwide [51]. Subsequently, not only will the absolute number of lesions decrease, but lesions that are HR HPV negative, which can be caused by non-carcinogenic HPV types and are likely to regress, may constitute a higher proportion of the remaining lesions. Studies have found that the likelihood of lesion regression is highly dependent on the HPV genotype associated with that lesion, with the vast majority of CIN2 lesions that are HR HPV negative reported to regress within 1 year [25]. This will have a direct effect on HPV device performance estimates. Continuing to evaluate HR HPV devices against CIN2+/CIN3+ diagnoses may make it appear that devices in the post-vaccine era have lower sensitivity than in the pre-vaccine era, but in actuality, a higher proportion of the remaining lesions will be HR HPV negative and caused by factors that are not on the causal pathway to cervical cancer (i.e., low risk HPV infections, which are not what HR HPV devices are designed to detect). This will also have an effect when assessing relative performance between an investigational and FDA-approved assay. If a previously approved device has poor sensitivity

¹⁰ For more information regarding IARC classification of low vs. high risk HPV genotypes, please see [50].

using a purely CIN2+/CIN3+ endpoint in the post-vaccine era due to the reasons described above, newer devices that cross react with low risk HPV infections may appear to have better sensitivity, which as noted could encourage manufacturers to optimize their devices to detect these non-clinically relevant infections.

2. *Possible performance bias:* For lesions that are caused by carcinogenic HPV infections, the proportion of the lesions caused by vaccine- and non-vaccine targeted types will be different in studies with varying rates of vaccinated subjects. Using a CIN2+/CIN3+ endpoint could potentially mask important characteristics in the device's ability to detect cancer-causing infections. For example, a device that tends to miss infections due to vaccine-targeted HPV types but detects non-vaccine-targeted types very well may still have acceptable clinical performance if the majority of the lesions in the study population are caused by non-vaccine types (if the study includes a large proportion of vaccinated individuals). While this device may be beneficial in highly vaccinated populations, it may miss critically important infections in settings where vaccination may not be as prevalent. The converse is also true: if a device has a reduced ability to detect non-vaccine types, the performance may look acceptable if the majority of the lesions in the clinical study population are caused by vaccine targeted types (if there is a very low proportion of vaccinated individuals enrolled in the study).

One way to mitigate this risk is to evaluate the HR HPV device performance separately for vaccinated and non-vaccinated populations in the study, which the FDA has started doing. However, once herd immunity plays a greater role, this problem still may exist, since herd immunity will cause the HPV type distribution in non-vaccinated populations to look similar to the distribution in vaccinated populations. In a study by the CDC analyzing the NHANES data, there was a statistically significant 34% drop in vaccine-targeted HPV infections in non-vaccinated populations in the post-vaccine era (2011-2014) compared to the pre-vaccine era (2003-2006), indicating herd immunity is already taking effect [20]. Furthermore, the extent of herd immunity, and subsequently, the distribution of HPV types causing lesions amongst non-vaccinated populations, may be different in different studies depending on where the subjects are recruited. It has been reported that vaccination rates vary among different states in the US; Rhode Island, for example, has 70.8% coverage, while in Mississippi and South Carolina, rates are as low as 29.1% [52]. This suggests that in states with vaccination rates exceeding the threshold for herd immunity, the distribution of HPV types in non-vaccinated populations will appear closer in makeup to vaccinated populations. In states with very low vaccination coverage, on the other hand, the distribution of genotypes leading to CIN2+/CIN3+ lesions will overlap less between vaccinated and non-vaccinated populations.

FDA Proposal: Mixed histologic/molecular comparator [CIN2+/CIN3+/HPV genotyping]:

One possible way to mitigate the above-mentioned risks is to assess relative HPV assay performance (as discussed in section IV.C) against a comparator that utilizes both histology and the presence of HR HPV nucleic acid (histology (CIN2+/CIN3+) with HPV typing). It should be noted that this proposal is different from the "analytical comparator" that has been used, which has consisted of either two FDA-approved HPV tests, or one FDA-approved test and a laboratory developed PCR followed by sequencing of the amplicon. These analyses are generally only conducted on a subset of the specimens, and only the analytical characteristics of the test are evaluated (i.e., does a test result agree with PCR/sequencing results). Furthermore, in these analyses, any specimens that yield discordant results between the two comparator assays are deemed "indeterminate" and excluded from analyses. Up until recently, this information has generally not been analyzed in conjunction with histological diagnoses for clinical

performance evaluation. Using the proposed mixed histologic/molecular comparator would combine “histologic” and “molecular” data into one comparator algorithm against which clinical performance would be assessed.

There are several benefits to a mixed histologic/molecular comparator:

1. Determination of which lesions contain high risk vs. low risk HPV types (or even no types at all). For result outputs that are shared between the investigational and FDA-approved devices, this could be especially informative if differing results are obtained, as HPV typing data would provide additional information regarding the clinical relevance of these differences.
2. Ability to assess clinical performance for genotyping outputs that are unique to the investigational device.
3. More accurate assessment of the types of “false negative” device results that could occur: i.e., whether a negative device result occurs in a CIN2+/CIN3+ lesion that is positive for HR HPV nucleic acid vs. a negative device result that occurs in a CIN2+/CIN3+ lesion that is also negative for HR HPV nucleic acid. Knowing which type of “false negative” is occurring is important in assessing whether a device is safe and effective. If a HR HPV device yields a negative result in a CIN2+/CIN3+ lesion that has been shown to be positive for high-risk HPV nucleic acid, this could indicate issues with the sensitivity of the device in detecting nucleic acid of a particular HPV subtype. However, if a device yields a negative result in a CIN2+/CIN3+ lesion that has shown to be negative for high-risk nucleic acid (negative for any HPV type or HR HPV negative but positive for low risk HPV by a reference typing test) this could indicate issues that are independent of the device (e.g., sampling issue, anomalies in the patients, etc.).
4. A more detailed assessment of an investigational device’s ability to detect lesions caused by vaccine- and non-vaccine targeted HPV type infections would be obtained. This would enable the FDA to ensure that the device will work in populations with varying vaccination rates/herd immunity effects, and subsequently, differing distributions of circulating HPV types. Additionally, since relative performance against an FDA-approved HR HPV device could be assessed, and most FDA-approved HR HPV devices were evaluated in primarily non-vaccinated populations, this will provide additional information regarding previously approved device performance in detecting vaccine- and non-vaccine-targeted types.

The FDA is requesting that the panel discuss whether utilizing a mixed histologic/molecular comparator against which relative HPV device performance is assessed is an appropriate approach to assess the safety and effectiveness of new HR HPV devices. In this discussion, please consider how the association of HR HPV types with CIN2+/CIN3+ lesions should factor in when assigning “comparator positive” (for sensitivity analyses) and “comparator negative” (for specificity analyses) results (please see Table 8 below, where some cells have already been filled in - the FDA requests that the panel discuss the comparator results in blank cells); the considerations for choosing an appropriate HPV typing molecular comparator; and risks and benefits of this approach. Additionally, please discuss any alternative proposals that could mitigate the risks associated with using a purely histologic endpoint in the vaccine era.

Table 8: Determination of reference status using mixed histologic/molecular comparator

| Histology Diagnosis | HPV typing result using molecular comparator | | | | |
|------------------------|--|---------------------------|-----|-----------------|---------|
| | High Risk HPV | | | Low Risk HPV | HPV Neg |
| | 16, 18, 45, 31, 33, 52, 58 | 35, 39, 51, 56, 59, 68 | 66* | | |
| NEG | | | | NEG | NEG |
| CIN1 | | | | NEG | NEG |
| CIN2 | | | | | |
| CIN3 | POS | POS | | | |
| CIN3+ | POS | POS | | | |

This table outlines the different combinations of HPV genotypes and histological diagnoses with the majority of the HR HPV types divided up according to the 9-valent vaccine-targeted and non-targeted types although, if the panel would like to discuss all of the types separately, that would be acceptable as well.

*HPV66 is considered separately because in 2009, the International Agency for Research on Cancer (IARC) no longer classified this type as “carcinogenic” [50]. However, most HR HPV devices still include this genotype

V. Questions to Panel

1. Does the panel agree with the three proposals (enrichment from referral clinics, archived specimens, capping the vaccinated population) to increase the number of women positive for CIN2+/CIN3+ and/or HR HPV in clinical studies supporting HPV device approval?
2. Do the benefits of referring women to colposcopy who are ASC-US/HR HPV double negative and a subset of women who are NILM/HR HPV double negative outweigh the risks that the procedure and potential overtreatment pose to these women enrolled in the clinical studies? Please answer separately for ASC-US/HPV double negative and NILM/HPV double negative.
3. Do the benefits outweigh the risks for:
 - A. Simplifying the indications for use to encompass one general screening population, and
 - B. Removing reference to specific triage tests and clinical actions based on the results of the device

Please state any potential risk mitigation measures that the FDA should take if this new indications for use is used in the future (e.g., appropriate inclusion/exclusion criteria for clinical studies, warnings/limitations in labeling, sub-analyses to be performed, etc.).

4. Which of the two comparator methods does the panel recommend for the assessment of safety and effectiveness of new HR HPV devices:
 - A. Agreement against a composite molecular comparator consisting of 3 FDA-approved molecular assays, or
 - B. Relative performance between a new and FDA-approved assay using a clinical endpoint comparator

What should be the minimum performance criteria?

5. If the panel recommends assessing HR HPV device performance against a clinical endpoint comparator:
 - A. Is utilizing a mixed histologic/molecular comparator acceptable?
 - B. If so, how should the association of HR HPV types with CIN2+/CIN3+ lesions factor in when assigning “comparator positive” and “comparator negative” results?

Appendix

FDA regulations applicable to *in vitro* diagnostic devices are based on the FDA classification of the device. The current approach to classification is the result of several laws, most prominently the 1976 Medical Device Amendments to the original Food, Drug and Cosmetic Act (the FD&C Act) (<http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Overview/ClassifyYourDevice/>). Medical devices, including *in vitro* diagnostic devices, are classified on the basis of risk. The three regulatory classes for device categorization are based on the level of regulatory control necessary to provide reasonable assurance of the safety and effectiveness of a device:

- Class I: Devices of low risk for which general controls are sufficient to provide a reasonable assurance of safety and effectiveness of the device.
- Class II: Devices for which there is sufficient information to establish special controls that are necessary, in combination with the general controls, to provide reasonable assurance of the safety and effectiveness of the device.
- Class III: Devices intended to be used in supporting or sustaining human life or preventing impairment of human health, or that may present a potential unreasonable risk of illness or injury for which insufficient information exists to determine special controls, that in combination with the general controls, would be sufficient to provide reasonable assurance of the safety and effectiveness of a device

HR HPV devices are class III devices. Class III devices require 'pre-market approval' (PMA) applications for which additional materials are necessary at the time of regulatory filing by the sponsor/manufacturer. In addition, class III devices may have additional post-approval requirements. For a more detailed overview of PMA applications:

<https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/HowtoMarketYourDevice/PremarketSubmissions/PremarketApprovalPMA/default.htm>

References

1. Huh, W.K., et al., *Use of primary high-risk human papillomavirus testing for cervical cancer screening: interim clinical guidance*. *Gynecol Oncol*, 2015. **136**(2): p. 178-82.
2. Force, U.S.P.S.T., et al., *Screening for Cervical Cancer: US Preventive Services Task Force Recommendation Statement*. *JAMA*, 2018. **320**(7): p. 674-686.
3. Chesson, H.W., et al., *The estimated lifetime probability of acquiring human papillomavirus in the United States*. *Sex Transm Dis*, 2014. **41**(11): p. 660-4.
4. Walbloomers JMM, J.M., Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJF, Peto J, Meijer CJLM, Munoz N, *Human Papillomavirus is a necessary cause of invasive cervical cancer worldwide*. *Journal of Pathology*, 1999.
5. de Sanjose, S., et al., *Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study*. *The Lancet Oncology*, 2010. **11**(11): p. 1048-1056.
6. Doorbar, J., et al., *Human papillomavirus molecular biology and disease association*. *Rev Med Virol*, 2015. **25 Suppl 1**: p. 2-23.
7. Ho, G., et al, *NATURAL HISTORY OF CERVICOVAGINAL PAPILOMAVIRUS INFECTION IN YOUNG WOMEN*. *NEJM*, 1998. **338**(7): p. 5.
8. Castle, P.E., et al., *Short term persistence of human papillomavirus and risk of cervical precancer and cancer: population based cohort study*. *BMJ*, 2009. **339**: p. b2569.
9. Gravitt, P.E., et al., *High load for most high risk human papillomavirus genotypes is associated with prevalent cervical cancer precursors but only HPV16 load predicts the development of incident disease*. *Int J Cancer*, 2007. **121**(12): p. 2787-93.
10. Kovacic, M.B., et al., *Relationships of human papillomavirus type, qualitative viral load, and age with cytologic abnormality*. *Cancer Res*, 2006. **66**(20): p. 10112-9.
11. Schmitt, M., et al., *Viral load of high-risk human papillomaviruses as reliable clinical predictor for the presence of cervical lesions*. *Cancer Epidemiol Biomarkers Prev*, 2013. **22**(3): p. 406-14.
12. Wang, S.M., et al., *Six-year regression and progression of cervical lesions of different human papillomavirus viral loads in varied histological diagnoses*. *Int J Gynecol Cancer*, 2013. **23**(4): p. 716-23.
13. Guan, P., et al., *Human papillomavirus types in 115,789 HPV-positive women: a meta-analysis from cervical infection to cancer*. *Int J Cancer*, 2012. **131**(10): p. 2349-59.
14. Li, N., et al., *Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: Variation by geographical region, histological type and year of publication*. *Int J Cancer*, 2011. **128**(4): p. 927-35.
15. Tota, J.E., et al., *Introduction of molecular HPV testing as the primary technology in cervical cancer screening: Acting on evidence to change the current paradigm*. *Prev Med*, 2017. **98**: p. 5-14.
16. Munro, A., et al., *The impact of human papillomavirus type on colposcopy performance in women offered HPV immunisation in a catch-up vaccine programme: a two-centre observational study*. *BJOG*, 2017. **124**(9): p. 1394-1401.
17. Wentzensen, N., et al., *Evidence-Based Consensus Recommendations for Colposcopy Practice for Cervical Cancer Prevention in the United States*. *J Low Genit Tract Dis*, 2017. **21**(4): p. 216-222.

18. Schiffman, M., et al., *Preparing for the Next Round of ASCCP-Sponsored Cervical Screening and Management Guidelines*. J Low Genit Tract Dis, 2017. **21**(2): p. 87-90.
19. Gage, J.C., et al., *Similar Risk Patterns After Cervical Screening in Two Large U.S. Populations: Implications for Clinical Guidelines*. Obstet Gynecol, 2016. **128**(6): p. 1248-1257.
20. Oliver, S.E., et al., *Prevalence of Human Papillomavirus Among Females After Vaccine Introduction-National Health and Nutrition Examination Survey, United States, 2003-2014*. J Infect Dis, 2017. **216**(5): p. 594-603.
21. Benard, V.B., et al., *Population-Based Incidence Rates of Cervical Intraepithelial Neoplasia in the Human Papillomavirus Vaccine Era*. JAMA Oncol, 2017. **3**(6): p. 833-837.
22. Hariri, S., et al., *Population-based trends in high-grade cervical lesions in the early human papillomavirus vaccine era in the United States*. Cancer, 2015. **121**(16): p. 2775-81.
23. Moscicki, A.B., et al., *Rate of and risks for regression of cervical intraepithelial neoplasia 2 in adolescents and young women*. Obstet Gynecol, 2010. **116**(6): p. 1373-80.
24. Godfrey, M.A.L., et al., *Conservative management of cervical intraepithelial neoplasia grade 2 (CIN2) in women under 30 years of age: A cohort study*. Eur J Obstet Gynecol Reprod Biol, 2018. **228**: p. 267-273.
25. Tainio, K., et al., *Clinical course of untreated cervical intraepithelial neoplasia grade 2 under active surveillance: systematic review and meta-analysis*. BMJ, 2018. **360**: p. k499.
26. Munro, A., et al., *Spontaneous regression of CIN2 in women aged 18-24 years: a retrospective study of a state-wide population in Western Australia*. Acta Obstet Gynecol Scand, 2016. **95**(3): p. 291-8.
27. Discacciati, M.G., et al., *Outcome of expectant management of cervical intraepithelial neoplasia grade 2 in women followed for 12 months*. Eur J Obstet Gynecol Reprod Biol, 2011. **155**(2): p. 204-8.
28. Castle, P.E., et al., *Evidence for frequent regression of cervical intraepithelial neoplasia-grade 2*. Obstet Gynecol, 2009. **113**(1): p. 18-25.
29. Trimble, C.L., et al., *Spontaneous regression of high-grade cervical dysplasia: effects of human papillomavirus type and HLA phenotype*. Clin Cancer Res, 2005. **11**(13): p. 4717-23.
30. Wright, T., Massad LS, Dunton CJ, Spitzer M, Wilkinson EJ, Solomon D, 2006 *Consensus Guidelines for the Management of Women With Abnormal Screening Tests*. Journal of Lower Genital Tract Disease, 2007. **11**(4): p. 201-222.
31. Dillner, J., et al., *Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study*. BMJ, 2008. **337**: p. a1754.
32. Katki, H.A., et al., *Cervical cancer risk for women undergoing concurrent testing for human papillomavirus and cervical cytology: a population-based study in routine clinical practice*. The Lancet Oncology, 2011. **12**(7): p. 663-672.
33. Saslow, D., et al., *American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer*. Am J Clin Pathol, 2012. **137**(4): p. 516-42.

34. Massad, L.S., et al., *2012 updated consensus guidelines for the management of abnormal cervical cancer screening tests and cancer precursors*. *Obstet Gynecol*, 2013. **121**(4): p. 829-46.
35. Katki, H.A., et al., *Five-year risks of CIN 3+ and cervical cancer among women with HPV testing of ASC-US Pap results*. *J Low Genit Tract Dis*, 2013. **17**(5 Suppl 1): p. S36-42.
36. Ronco, G., et al., *Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial*. *The Lancet Oncology*, 2010. **11**(3): p. 249-257.
37. Kyrgiou, M., et al., *Adverse obstetric outcomes after local treatment for cervical preinvasive and early invasive disease according to cone depth: systematic review and meta-analysis*. *BMJ*, 2016. **354**: p. i3633.
38. ACOG, *Practice Bulletin: Cervical Cancer Screening and Prevention*. *Obstet Gynecol*, 2016.
39. Gage, J.C., et al., *The low risk of precancer after a screening result of human papillomavirus-negative/atypical squamous cells of undetermined significance papanicolaou and implications for clinical management*. *Cancer Cytopathol*, 2014. **122**(11): p. 842-50.
40. Katki, H.A., et al., *Benchmarking CIN 3+ risk as the basis for incorporating HPV and Pap cotesting into cervical screening and management guidelines*. *J Low Genit Tract Dis*, 2013. **17**(5 Suppl 1): p. S28-35.
41. Schiffman, M., et al., *A long-term prospective study of type-specific human papillomavirus infection and risk of cervical neoplasia among 20,000 women in the Portland Kaiser Cohort Study*. *Cancer Epidemiol Biomarkers Prev*, 2011. **20**(7): p. 1398-409.
42. Cuzick, J., et al., *Long-term follow-up of cervical abnormalities among women screened by HPV testing and cytology-Results from the Hammersmith study*. *Int J Cancer*, 2008. **122**(10): p. 2294-300.
43. Schiffman, M., et al., *Relative Performance of HPV and Cytology Components of Cotesting in Cervical Screening*. *J Natl Cancer Inst*, 2017.
44. Khan, M.J., et al., *ASCCP Colposcopy Standards: Role of Colposcopy, Benefits, Potential Harms, and Terminology for Colposcopic Practice*. *J Low Genit Tract Dis*, 2017. **21**(4): p. 223-229.
45. Solomon, D., et al., *The 2001 Bethesda System: terminology for reporting results of cervical cytology*. *JAMA*, 2002. **287**(16): p. 2114-9.
46. Darragh, T.M., et al., *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*. *J Low Genit Tract Dis*, 2012. **16**(3): p. 205-42.
47. Wentzensen, N., et al., *Triage of HPV positive women in cervical cancer screening*. *J Clin Virol*, 2016. **76 Suppl 1**: p. S49-S55.
48. Polman, N.J., et al., *Evaluation of the Clinical Performance of the HPV-Risk Assay Using the VALGENT-3 Panel*. *J Clin Microbiol*, 2017. **55**(12): p. 3544-3551.
49. Petry, K.U., et al., *Evaluating HPV-negative CIN2+ in the ATHENA trial*. *Int J Cancer*, 2016. **138**(12): p. 2932-9.

50. Schiffman, M., G. Clifford, and F.M. Buonaguro, *Classification of weakly carcinogenic human papillomavirus types: addressing the limits of epidemiology at the borderline*. Infect Agent Cancer, 2009. **4**: p. 8.
51. Vesikari, T., et al., *A Randomized, Double-Blind, Phase III Study of the Immunogenicity and Safety of a 9-Valent Human Papillomavirus L1 Virus-Like Particle Vaccine (V503) Versus Gardasil(R) in 9-15-Year-Old Girls*. Pediatr Infect Dis J, 2015. **34**(9): p. 992-8.
52. Dilley, S., K. Miller, and W. Huh, *HPV vaccination*. Gynecol Oncol, 2018. **148**(1): p. 3-4.