

HPV-BASED SCREENING AND MOLECULAR DIAGNOSTICS IN THE PREVENTION OF CERVICAL CANCER: CLINICAL PERFORMANCE, REAL-WORLD EVIDENCE, AND IMPLEMENTATION PRIORITIES

Xolboyeva Mashhura Yoqutjon qizi

Email – yoqutjonovnam@gmail.com

Termez University of Economics and Service

Faculty of Medicine

Abstract

Background: Persistent infection with carcinogenic human papillomavirus (HPV) is a necessary causal factor in most cervical cancers, but a positive molecular assay is not equivalent to precancer or invasive disease. The transition from cytology-centred prevention to validated high-risk HPV testing has changed the evidentiary basis of screening, triage, surveillance and programme evaluation. Objective: To critically integrate evidence on HPV DNA and mRNA testing, partial and extended genotyping, p16/Ki-67 dual staining, host and viral methylation, self-collection, artificial intelligence and longitudinal real-world data, with particular consideration of feasible implementation in Uzbekistan. Methods: This is a critical integrative evidence review, not a prospectively registered systematic review or a primary patient-level analysis. Official WHO, FDA, EMA and ASCCP recommendations and identifiable studies were prioritised; diagnostic accuracy and clinical-outcome evidence were distinguished from feasibility reports and hypotheses. Propensity-score methods, inverse probability weighting, target-trial emulation, longitudinal survival models and machine-learning validation are discussed as appropriate designs for future registry-based evaluations, not reported as analyses performed on a fabricated cohort. Results: WHO recommends validated HPV DNA testing as a preferred screening modality within locally implementable screen-and-treat or screen-triage-and-treat pathways. Molecular genotyping and dual-stain assessment improve risk stratification among individuals testing HPV-positive, while self-sampling can reduce participation barriers when collection devices and assays are validated and positive results trigger timely follow-up. Evidence of test accuracy does not by itself establish a reduction in mortality; linkage to colposcopy, histology, treatment and post-treatment surveillance remains decisive. Artificial intelligence may support digital cytology and quality assurance but is vulnerable to spectrum shift, verification bias and unequal

performance. Discussion: Regulatory-grade evidence requires explicit target populations, assay specifications, complete follow-up, calibration and transparency regarding clinical thresholds. Context-specific health-information systems can quantify missed follow-up and inequalities, although causal claims require stronger designs. Conclusion: Integrated HPV-based prevention should prioritise coverage, validated diagnostics, risk-adapted triage, safe treatment and verifiable surveillance rather than technology acquisition alone.

Keywords: cervical cancer; human papillomavirus; HPV DNA; HPV mRNA; molecular screening; extended genotyping; p16/Ki-67; cervical intraepithelial neoplasia; self-sampling; real-world evidence; artificial intelligence; Uzbekistan.

1. Introduction

Cervical cancer prevention is unusual among cancer-control programmes because a preventable infection occupies a central place in the causal pathway. Persistent carcinogenic HPV infection may produce transforming cellular alterations, subsequently leading to high-grade cervical intraepithelial neoplasia and, in a minority of affected individuals, invasion. The interval between infection and clinically apparent malignancy provides a window for screening. Nevertheless, most HPV infections resolve, and molecular positivity should never be conflated with a diagnosis of cervical malignancy. The clinically important question is not simply whether viral nucleic acid can be detected but whether an accessible programme can correctly identify those who need evaluation or treatment while avoiding unnecessary interventions [1–5].

The WHO elimination framework identifies a target incidence below four cases per 100,000 women-years and the complementary 90–70–90 goals for vaccination, high-performance screening and treatment by 2030 [1]. WHO's 2024 guidance reports approximately 662,000 incident cervical cancer cases and 349,000 deaths worldwide in 2022 [4]. These estimates describe the global burden, not the incidence among screened individuals, and should not be interpreted as a locally measured Uzbek rate. Programme performance depends on the full cascade from invitation through result communication, colposcopy or treatment when indicated, follow-up and quality assurance.

High-risk HPV nucleic acid amplification tests detect exposure-related viral targets and generally offer greater sensitivity for detecting high-grade precancer than cytology alone, although specificity and positive predictive value vary with age, vaccination

history, sampling and the underlying disease prevalence [2,3]. The WHO supports clinician-collected and validated self-collected samples for HPV DNA testing [2]. Genotyping, cytology, dual staining and selected molecular adjuncts are not interchangeable laboratory products: their purpose is to move a patient from a screening-positive state to a calibrated management decision. Performance claims should specify CIN2+, CIN3+ or cancer as the reference endpoint and whether disease verification was complete.

Real-world data (RWD) are routinely collected observations describing health status, services and outcomes, including electronic records, laboratory interfaces, registries, administrative claims and patient-reported information. Real-world evidence (RWE) comprises clinical inferences obtained from appropriately analysed RWD. The existence of digital records does not automatically establish a causal effect. Data completeness, treatment selection, prevalent disease and follow-up all determine whether credible RWE is possible. In cervical screening, registry-based investigations can measure coverage, failures of follow-up, time to colposcopy and interval-cancer risk, whereas properly conducted randomized trials (RCTs) are particularly important for estimating effects of invitations, modalities or management strategies [7,8].

An RCT assigns exposure prospectively and thereby reduces measured and unmeasured baseline confounding in expectation. A pragmatic trial may preserve routine practice while retaining randomization; an observational comparison needs careful design to avoid confounding by indication, immortal-time bias, differential surveillance and selection. Efficacy denotes an intervention's performance under specified trial conditions, whereas effectiveness refers to the outcome in ordinary systems with heterogeneous uptake and adherence. A highly sensitive assay may yield little population benefit when positive results are not acted upon; conversely, modest gains in participation through self-collection may be consequential even without improving intrinsic analytic sensitivity [6,9].

FDA guidance addresses whether RWD are sufficiently relevant and reliable to support decisions about medical devices; the European Medicines Agency (EMA) has developed an RWE framework and the DARWIN EU distributed-data environment for regulatory questions [7,8]. These initiatives do not imply regulatory endorsement of any specific HPV assay or AI model. WHO guidance on artificial intelligence stresses clinical benefit, transparency, accountability, inclusiveness and protection from automation bias [10]. For cervical screening, machine learning may assist image

interpretation or follow-up prediction, but prospective evaluation remains essential before clinical deployment.

The present article therefore adopts a focused clinical hypothesis: the public-health value of HPV molecular diagnostics depends on the interaction of analytical validity, clinically meaningful triage, treatment capacity and trustworthy longitudinal information. The primary question concerns cervical cancer prevention. Examples from diabetes, cardiology, oncology, infectious disease and neurology are examined only as methodological comparators for effectiveness prediction, not as unrelated results presented as cervical-screening evidence.

2. Materials and Methods

2.1 Review design and scope

This work is a critical integrative narrative review with an explicit evidence hierarchy. It does not claim to be a PRISMA-compliant systematic review, and no invented numbers of screened records, excluded abstracts or pooled study effects are reported. The evidence base is anchored in WHO cervical screening and elimination recommendations, FDA assay and RWE guidance, EMA regulatory materials and ASCCP risk-based management publications, supplemented by identifiable trials and reviews. PubMed-indexed research and primary journal records are preferred for biomedical conclusions. Scopus, Web of Science, Embase, Cochrane Library, Nature Medicine, The Lancet, NEJM and JAMA are relevant potential indexing or publication sources, but we do not claim exhaustive interrogation of paid subscription databases without a reproducible search export.

A reproducible future search should combine (cervical cancer OR cervical intraepithelial neoplasia OR CIN3 OR CIN2) AND (HPV DNA OR HPV mRNA OR genotyping OR self-sampling OR dual stain OR methylation) AND (screening OR triage OR implementation OR real-world). Separate search blocks should cover diagnostic test accuracy, cervical cancer incidence or mortality, participation and retention, health economics and AI validation. Priority is placed on 2020–2026 guidelines and studies, with older landmark randomized evidence retained when it informs clinical outcomes. Eligible evidence includes externally validated accuracy studies, randomized screening evaluations, population registries, systematic reviews and official guidance. Exclude studies without an identifiable endpoint, assay specification

or meaningful reference standard from quantitative comparison; discuss them narratively as hypothesis-generating where appropriate.

2.3 Extraction and synthesis

For each source, the appropriate extraction fields are design, country, sample frame, age, HIV status, HPV vaccine exposure, index test, threshold, triage sequence, verification procedure, endpoint, horizon, precision and follow-up completion. Sensitivity and specificity estimate conditional test properties; positive and negative predictive values change with prevalence. For observational programme endpoints, distinguish first-round yield from incidence after repeat screening. An unpooled critical synthesis is warranted when assays, populations and reference standards are heterogeneous. Publication-level percentages are attributed to the original study or organization rather than generalized to all settings.

2.4 Propensity-score matching and inverse-probability weighting

For comparative evaluations in routine care, estimate the propensity for receiving an HPV-based versus cytology-based screening invitation from pretreatment covariates: age, area, deprivation, screening history, pregnancy, immunocompromise and service access. Matching can restrict analysis to patients with comparable scores, with balance assessed using standardized mean differences rather than a significance test. Inverse probability of treatment weighting creates a weighted pseudo-population; extreme probabilities require positivity checks and stabilized or truncated weights. These methods address only measured confounding and cannot repair absent ascertainment of CIN3+ or differential pathology verification.

2.5 Target trial emulation and causal inference

Specify time zero as eligibility and screening assignment or invitation, rather than the day of a result that already depended on attendance. Define eligibility, intervention strategies, assignment, follow-up, outcomes, estimand and analysis plan before comparison. Intention-to-screen evaluates outreach policies; per-protocol approaches require adjustment for time-varying adherence and censoring. Causal diagrams should separate mediators such as colposcopy attendance from baseline confounders, avoiding inappropriate adjustment. Instrumental-variable claims need credible exclusion restrictions; an invitation instrument may fail if outreach affects other care pathways.

2.6 Survival analysis

Kaplan–Meier estimators can summarize time from positive HPV test to completion of triage where noninformative censoring is plausible. Cox regression can estimate hazard ratios for clinically defined outcomes while testing proportional-hazards assumptions; a hazard ratio is not a risk ratio. Competing risks and interval-censored precancer detection can make simple survival models misleading. For cancer endpoints, screen-detected disease and deaths require attention to lead-time bias, length bias and differential surveillance. Here these are recommended designs, not analyses claimed to have been executed.

2.7 AI methods and validation

Machine-learning algorithms may model structured data using regularized logistic regression, random forests or boosted trees; deep learning may process digitized cytology or colposcopy images; NLP may extract referrals and pathology features; large language models may assist with document classification under clinician review. Reinforcement learning has been proposed for sequential policies but should not guide actual treatment without rigorously verified safety. Evaluate independent held-out sites, temporal drift, AUROC together with calibration, sensitivity at chosen clinical thresholds, subgroup performance, decision-curve utility and prospective impact. Split data at patient and site levels to avoid leakage between repeat samples.

3. Results: Critical Evidence Synthesis

3.1 Causal biology and screening targets

The viral life cycle and progression pathway provide the rationale for screening while revealing why viral positivity alone is an imperfect surrogate for clinically consequential disease. Carcinogenic HPV types may persist in the transformation zone, and expression of viral oncogenes E6 and E7 interferes with cellular tumor-suppressor pathways. Histological CIN2 is biologically heterogeneous and can regress, whereas CIN3 is a more robust precursor endpoint for comparative risk assessment. Molecular assays generally test for carcinogenic viral nucleic acids rather than directly proving transforming disease. Therefore, a screening programme needs a structured pathway after positive results; the assay is the entry point rather than the therapeutic endpoint [2–5].

Interpretive implication: The causal chain supports measuring progression-relevant endpoints rather than treating all HPV-positive tests as equivalent clinical events.

3.2 HPV DNA as primary screening modality

WHO recommends HPV DNA testing in primary screening beginning at age 30 years in the general population and a younger initiation age for women living with HIV, with intervals depending on population and strategy [2]. Screening intervals should not be transferred uncritically from an international guideline to a national policy without verification of local eligibility, treatment capacity and vaccination coverage. High sensitivity confers a potentially reassuring negative test but also generates referrals when transient infections predominate. Screening benefit derives from earlier precancer diagnosis followed by treatment, not merely from detecting more HPV infections [2,3].

Interpretive implication: A negative high-risk HPV test offers programmatic reassurance only when platform validation, sampling quality and rescreening mechanisms are maintained.

3.3 HPV mRNA testing and translational caveats

E6/E7 mRNA tests seek transcriptional activity associated with transforming infection, yet test performance depends on covered viral types and diagnostic cutoffs. WHO published a specific 2021 recommendation set for mRNA testing and contextualized its use among nucleic-acid testing options [3]. Some investigations find increased specificity for particular mRNA-based approaches, but this cannot be assumed for every platform or every sampling medium. Clinical validation and locally implementable triage are prerequisites. Comparisons should stratify by age, HIV status and histology rather than pooling uncommensurable subgroups.

Interpretive implication: Comparative studies should use the same specimens, a prespecified reference standard and clinically relevant follow-up intervals.

3.4 Partial and extended genotyping

HPV16 and HPV18 carry particular clinical importance, and genotype-aware management can distinguish higher-risk from lower-risk screening-positive groups. The 2025 ASCCP Enduring Consensus guidance incorporated extended genotyping information from Onclarity-based cohorts into risk-threshold management; colposcopy is recommended after HPV16/18 positivity, while selected other genotypes are triaged with cytology or dual stain [5]. Such recommendations are assay-specific and jurisdiction-sensitive. A type-group result should not be treated as equivalent to complete individual genotyping. Longitudinal risk calibration after

vaccination will require repeated updates because genotype distributions and baseline risk change.

Interpretive implication: Referral thresholds should be calibrated against absolute CIN3+ risk rather than assumed solely from genotype ordering.

3.5 Dual-stain cytology and triage

The immunocytochemical p16/Ki-67 dual-stain approach identifies cells with aberrant proliferative signaling and has a role in triage of HPV-positive screening results. WHO issued guidance on dual-stain triage in 2024 [4]. Technical reproducibility relies on slide preparation, staining, reader training and quality control. A dual-stain negative result should be interpreted within the chosen risk-based algorithm and follow-up interval, not as a universal guarantee of freedom from disease. A molecular triage programme may lower unnecessary colposcopy while retaining sensitivity, but its true value also depends on referral adherence and cost.

Interpretive implication: A triage technology is valuable only when it improves the balance between cancer prevention and avoidable diagnostic procedures.

3.6 Host methylation, viral methylation and emerging biomarkers

Host and viral DNA methylation panels, p16-related pathways, cellular proliferation signatures and combined molecular classifiers seek to distinguish transient infection from lesions likely to persist or progress. Their promise is biologically plausible, yet assay thresholds, pre-analytical effects, external validity and downstream management consequences vary. Evidence should separate proof-of-concept case-control accuracy from prospective screening-population validation. A high AUROC may coexist with a poor positive predictive value at low disease prevalence. New biomarkers should not replace validated WHO-supported screening algorithms in routine services solely because a development study reports attractive discrimination.

Interpretive implication: Before clinical adoption, proposed panels need multi-site blinded validation and analysis of added net benefit beyond established triage.

3.7 Self-collected HPV testing and participation

Self-collected vaginal samples can reduce embarrassment, travel and appointment barriers, particularly for individuals with limited access to pelvic examinations. A systematic review of randomized and observational participation studies reported greater uptake with self-sampling than control strategies, pooled RR 2.13 (95% CI

1.89–2.40) across 33 studies, while noting substantial representation of high-income settings [12]. This is a relative effect on screening participation, not the sensitivity of the assay or a measured reduction in cervical cancer deaths. In May 2024 FDA expanded approved cobas HPV testing to defined self-collected vaginal specimens, subject to the corresponding indication and instructions [6].

Interpretive implication: Self-collection is an access intervention and a diagnostic pathway; the two mechanisms require separate effectiveness evaluation.

3.8 Clinical-outcome evidence from randomized trials

The landmark cluster-randomized trial in rural India enrolled 131,746 women across comparison groups and found fewer advanced cancers and cervical cancer deaths after HPV screening compared with standard care; the hazard ratio for advanced disease was 0.47 (95% CI 0.32–0.69) and for cervical cancer death 0.52 (95% CI 0.33–0.83) [11]. These trial findings demonstrate that a completed diagnostic-treatment pathway can produce measurable population benefit, while the study setting and intervention era limit direct transportability to 2026 screening programmes. Contemporary vaccination exposure, infrastructure and triage options differ markedly.

Interpretive implication: Population outcome trials and accuracy studies answer distinct questions and should not be numerically conflated.

3.9 Vaccination, screening and the interpretation of residual risk

HPV vaccination lowers infection with vaccine-targeted oncogenic types but does not eliminate the need for screening in populations for which screening remains recommended. Vaccination status and birth cohort alter prevalence and thereby the positive predictive value of HPV tests. National programmes should avoid treating vaccination registries and screening registries as independent silos: linking these data can support assessment of equitable coverage and schedule changes, provided data governance is explicit. Ecological comparisons without individual linkage cannot distinguish vaccine impact from secular changes in sexual behavior, assay mix or healthcare access.

Interpretive implication: Longitudinal analyses of vaccine and screening interactions require age-cohort stratification and repeated monitoring.

3.10 Referral, colposcopy and histopathology

After HPV positivity, risk-based decisions may include repeat testing, reflex cytology, dual stain or colposcopy according to genotype, prior history and local guideline. Histopathology remains decisive when a biopsy is indicated, and pathology quality must be integrated into screening quality assurance. Women experiencing a positive test but no completed triage may be incorrectly counted as screened successfully; programmes must instead distinguish initial coverage, proportion with valid results, referral completion, treatment completion and surveillance. Clinical communication is essential because poorly explained results can increase anxiety and disengagement.

Interpretive implication: Failure to complete referral is a measurable patient-safety event and should trigger navigation efforts rather than be treated as an administrative inconvenience.

3.11 Digital cytology and AI-assisted interpretation

Deep-learning classification of cytology and colposcopy images is sensitive to staining, image resolution, scanner differences, referral population enrichment and verification procedures. An algorithm trained primarily on highly abnormal images may show appealing internal accuracy but over-refer a low-prevalence screening cohort. AI should be judged at a clinically selected sensitivity and referral burden, not a single aggregate AUROC. Calibration, out-of-distribution monitoring and human override options are indispensable. AI-supported image assessment is not interchangeable with a validated primary HPV molecular assay.

Interpretive implication: The appropriate primary outcome for an AI-assistance trial is clinical decision quality and downstream burden, not isolated retrospective image classification.

Electronic medical records capture symptoms, diagnoses and follow-up but frequently lack complete records of outside laboratory testing. E-prescriptions help document treatment and analgesic or procedural medication exposure but do not verify that precancer was cured. Laboratory databases record HPV test platform, genotype and specimen type when interfaces are structured. National cancer registries can support linkage to invasive cancer outcomes, while immunization registries define prior HPV vaccine exposure. Pharmacovigilance systems monitor suspected vaccine or treatment-related adverse events, but spontaneous reporting cannot by itself

estimate incidence. Insurance claims may measure services and utilization; wearables and mobile applications may support appointment reminders and patient-reported barriers. Each source must have provenance, denominators and purpose-specific quality assurance.

Interpretive implication: RWD diversity is valuable when linkage preserves valid denominators, dates and the distinction between observation and inference.

3.13 Data harmonization and interoperability

For a prospective national screening registry, minimum data elements include a unique pseudonymized patient key, date of birth, district, eligibility criteria, vaccine history, screening date, assay manufacturer, sample collection setting, genotype category, cytology or dual-stain results, triage recommendation, referral dates, biopsy findings, treatment method, complications and follow-up outcome. Consistent SNOMED, ICD and laboratory coding should be adopted only with local governance and validated mapping. Missing data must be classified as not measured, result pending, participant lost, or true negative; collapsing these states biases programme success estimates.

Interpretive implication: Interoperability is a scientific validity problem as much as a software engineering problem.

3.14 EHR, e-prescriptions, laboratory and registries

An electronic health record may identify longitudinal screening encounters, yet encounter-based denominators over-represent frequent service users. Electronic prescription systems can identify medication treatment surrounding procedures and prevent some duplication, although cervical neoplasia treatment is not usually captured fully through pharmaceutical dispensing. Laboratory information systems need versioned assay names and thresholds, since replacement instruments can induce apparent temporal changes in HPV positivity. National registries may clarify cancer incidence only after deduplication and stable linkage. These complementary databases create useful RWD when their sampling frames and validation errors are documented.

Interpretive implication: Historical screening comparisons must adjust for assay changes and ascertainment drift.

3.15 Pharmacovigilance, signal detection and analytics

Adverse drug reactions are clinically relevant for procedural treatments, anesthetics, supportive therapies and prophylactic medicines; vaccine safety must be interpreted using appropriate vaccine pharmacovigilance methods. Disproportionality metrics such as reporting odds ratios identify possible signals from spontaneous reporting, not causal effects or absolute event risks. Signal validation requires clinical case review, background rates, temporality and confounding assessment. Registry-based active surveillance can complement passive safety reporting but needs a defined exposed denominator. For HPV screening, excessive interventions after false-positive results and loss to follow-up after true-positive results are additional programme safety outcomes.

Interpretive implication: Pharmacovigilance supports safety monitoring but cannot replace controlled comparative effectiveness evaluation.

3.16 Machine learning, NLP, reinforcement learning and LLMs

Structured-data machine learning can estimate probability of failing follow-up and prioritize outreach; NLP can identify cytology and pathology terminology in notes; large language models can draft patient communications with safeguards and clinician review. Deep learning can segment or classify cytology images. Reinforcement learning for triage sequence optimization remains exploratory because unsafe exploration is unacceptable in clinical care. Each application requires evidence for transportability, language-specific accuracy, privacy and utility. A programme should compare any AI workflow against a simple rules-based reminder system before adopting a complex model.

Interpretive implication: AI should amplify a validated screening cascade instead of obscuring accountability for missed diagnosis.

3.17 Predicting treatment effectiveness across disease areas

Examples from diabetes, cardiovascular disease, oncology, infectious diseases and neurology show that medication effectiveness prediction depends on explicit outcomes, treatment adherence, baseline severity and indication bias. Those lessons transfer methodologically to HPV vaccination or cervical precancer management, but numerical effect sizes from unrelated diseases cannot validly estimate screening benefit. Pharmacogenomics and pharmacogenetics may explain interindividual drug response in selected diseases; however, no general pharmacogenetic test has been

established as a routine substitute for HPV DNA screening or histological triage. Precision prevention of cervical cancer is currently better grounded in viral genotype, screening history, immune status, vaccination and pathology.

Interpretive implication: Disease-spanning prediction frameworks inform study design but do not constitute new clinical results for HPV prevention.

3.18 Uzbekistan implementation priorities

A feasible Uzbek implementation strategy begins with an auditable pathway rather than unverified assertions about the existing coverage or performance of national systems. The Ministry of Health and regional programmes could establish standardized eligibility, laboratory accreditation, referral tracking, quality indicators and protected data linkage. Existing or developing electronic health, e-prescription and cancer-registry infrastructure should be evaluated for actual interoperability rather than presumed to be complete. Pilot sites should report eligibility denominators, tested proportion, positivity, follow-up within specified intervals, histologically confirmed CIN3+, overtreatment, patient costs and rural-urban inequities. Independent local evaluation is needed before nationwide AI deployment.

Interpretive implication: A phased local registry trial offers more credible evidence than launching unvalidated risk models at national scale.

Technology / strategy	Main purpose	Key limitation / action
Validated hrHPV DNA	Primary risk detection	Positive result needs protocol-based triage; [2]
HPV E6/E7 mRNA	Molecular screening option	Assay-specific validation; [3]
HPV16/18 and extended genotyping	Risk stratification	Thresholds depend on assay and history; [5]
p16/Ki-67 dual stain	Triage after hrHPV positivity	Reader quality and follow-up; [4]
Self-collected sample	Improve screening access	Validated assay/specimen combination; [6,12]
Methylation panels	Emerging stratification	Prospective external evidence required
AI cytology	Reader support / triage research	Calibration, bias and device oversight

Evidence			Documented quantitative finding	Appropriate interpretation
WHO	global	2022	~662,000 cases; ~349,000 deaths	Global burden, not Uzbek incidence
WHO	elimination	goals	90% vaccination; 70% screening; 90% treatment	Targets, not achieved performance
Self-sampling	meta-analysis	[12]	RR 2.13 (95% CI 1.89–2.40)	Screening uptake; not diagnostic accuracy
India cluster RCT	[11]		HR 0.52 (95% CI 0.33–0.83)	Cervical cancer death vs usual care, specific context

3.20 Screening pathway design in immunocompromised groups

Screening schedules derived from general populations should not automatically be applied to individuals living with HIV, transplant recipients or those experiencing prolonged immunosuppression. Earlier acquisition, persistence of carcinogenic HPV and increased burden of high-grade lesions can alter the relationship between molecular detection and future disease. A valid implementation programme must encode immunocompromise in eligibility and recall systems, with access-controlled documentation that does not disclose sensitive status. The WHO 2021 guideline distinguishes recommendations for women living with HIV from those for the general population [2,18]. Comparisons should describe immune status, antiretroviral treatment and prior screening history before attributing observed differences to HPV technology.

The practical implication is to create stratified screening registers and risk-adapted recall pathways rather than a single undifferentiated digital reminder rule. Confirmation of sample adequacy and outcome ascertainment becomes more important when high-risk groups are underrepresented in trial datasets. Screening success should be measured by completed care, not the proportion issued an electronic invitation. Health-system audits should also identify whether privacy concerns produce differential loss to follow-up.

3.21 Pregnancy, fertility concerns and conservative management

An abnormal HPV-based result during pregnancy presents a different balance of diagnostic reassurance, procedural risk and timing than routine nonpregnant screening. HPV nucleic acid detection itself does not mandate an invasive procedure, and pregnancy-specific recommendations should determine when colposcopy or biopsy is indicated. The objective is to avoid missing suspected invasive cancer while

preventing unnecessary interventions in women with lesions that can be reassessed after delivery. Programme software should distinguish pregnancy status and anticipated follow-up windows and should never implement a generic expedited-treatment algorithm without clinician review [9].

Longitudinal studies that assess adverse pregnancy outcomes after cervical procedures must distinguish the causal effects of treatment from prior obstetric risk, infection, smoking and socioeconomic circumstances. For patients seeking future fertility, shared decisions should describe the tradeoff between sufficient treatment of high-grade disease and consequences of excessive cervical excision. RWD can support counselling when indication bias is acknowledged; an observational association between treatment and a later outcome is not by itself proof of treatment harm.

3.22 Negative results, interval cancers and quality failures

Post-screening reassurance requires a clear account of false-negative mechanisms. Sampling may miss abnormal cells, a molecular platform may fail due to inhibition, or a cancer may arise in a location difficult to sample. Administrative errors can also generate apparent negatives when rejected specimens or uncommunicated results are coded as normal. Screening registries should therefore distinguish analytical failures, specimen inadequacy, not-tested states and genuinely negative laboratory results. Interval cancers after a documented negative screen are priority sentinel events for blinded case review; they should not be automatically ascribed to assay insensitivity.

A structured case audit examines specimen collection, assay validity, interpretation, referral processes, pathology review and competing clinical diagnoses. Laboratory proficiency panels and repeat testing of selected stored specimens can identify batch effects, but retrospective retesting must respect consent and specimen storage conditions. Cancer incidence after screening must be interpreted with appropriate lead-time and surveillance adjustment. These evaluations provide stronger safety evidence than a dashboard reporting aggregate screening positivity alone.

3.23 Test accuracy metrics and clinical utility

Sensitivity is the probability of a positive test in the target disease state and specificity is the probability of a negative result without that state. These quantities are not constant across histologic thresholds, sample types, age or comparator protocols.

Positive predictive value additionally depends on prevalence and therefore declines when the background burden of CIN3+ decreases after successful vaccination, even if sensitivity is stable. Diagnostic likelihood ratios can help translate pre-test to post-test odds, but clinicians need an actionable risk threshold and an understood consequence of intervention.

Net benefit and decision-curve analysis are useful when the practical comparison is between more colposcopies and fewer missed clinically significant lesions. A model with a slightly higher AUROC may yield no clinically useful improvement if the gain lies outside the range of referral thresholds. Clinical utility studies should report absolute numbers of colposcopies, biopsies, CIN3+ detections, delayed diagnoses and unnecessary treatments per defined eligible population. It is scientifically misleading to compare products by a single vendor-reported sensitivity without specifying the validation population and reference standard.

3.24 Equity and patient-centred programme design

Participation is shaped by health literacy, transport, work obligations, stigma, linguistic barriers and trust in medical institutions. Self-collected sampling can address some obstacles, yet it can also exclude women who lack a private setting, a reliable delivery channel or understandable instructions. A programme that sends kits indiscriminately without a tested return logistics chain may record invitations rather than completed screening. Local qualitative studies should guide invitation formats, reminder schedules, specimen handoff and communication of an HPV-positive result.

Equity should be assessed with denominators by age, geography, disability-related access need and socioeconomic category where lawful and ethically appropriate. The aim is to distinguish gaps caused by insufficient outreach from gaps caused by referral bottlenecks or cost. Community engagement is not a substitute for scientific evaluation, but it can identify why an apparently technically successful service fails in practice. Patient-reported measures should include anxiety, perceived autonomy, confidentiality and understanding of subsequent management.

3.25 Economic evaluation and procurement decisions

Laboratory purchasing decisions should be framed in terms of total programme cost and outcomes rather than price per assay. Relevant cost components include personnel training, specimen collection, transportation, molecular platforms, reagent supply, quality-control panels, repeat testing, result communication, colposcopy,

histology, treatment, surveillance and the cost of navigating missed follow-up. Self-collection may reduce the clinic contact cost while adding distribution and laboratory logistics, so its incremental cost-effectiveness is conditional on uptake and downstream adherence.

Health economic models should specify time horizon, discounting, baseline incidence, HPV vaccination assumptions, test performance, transition probabilities and treatment capacity. Sensitivity analyses must explore plausible ranges rather than presenting one deterministic answer as universally applicable. For Uzbekistan, unit-cost and access data must be gathered locally before claiming savings or a favourable incremental cost-effectiveness ratio. Cost-reduction targets should not drive procurement of tests lacking demonstrated clinical validation.

3.26 Micro-level RWD architecture and data governance

The minimal dataset must preserve temporal ordering. A test ordered on one date, collected later, processed by a specific platform and communicated still later has several distinct timestamps that are essential to estimate delays. Registries should record invitation and consent status, collection mode, laboratory result, genotype or triage findings, recommended action, action actually completed, pathology, therapy and subsequent screening. For each record, source system, code version, change history and responsible laboratory should be identifiable to authorized reviewers.

Data protection is not solved by deleting patient names while preserving unconstrained linkage keys. A governance protocol should define controller responsibilities, permitted use, retention, access audits, secure transfer and breach response. Missing records may have clinical meaning: an absent histology report could imply no biopsy, biopsy elsewhere or incomplete linkage. Rather than automatically imputing an outcome, the programme should distinguish missingness causes and initiate quality investigations. Analytic readiness requires clinical oversight as well as technical interoperability.

3.27 Designing a prospective pragmatic evaluation in Uzbekistan

A pragmatic stepped-wedge or cluster-randomized implementation design could compare standard invitation pathways with a validated HPV self-sampling and navigation programme across participating districts, provided there is genuine equipoise and sufficient capacity for treatment of screen-detected disease. Alternatively, a carefully emulated target trial could compare staggered rollouts, using

a prespecified eligibility definition and common time zero. The primary endpoint could be completion of indicated follow-up within an agreed interval, with secondary endpoints including CIN3+ yield, patient experience and resource use. Sample-size calculation must follow locally observed baseline rates and prespecified effect size; no artificial enrollment target should be inserted into the protocol.

A data safety monitoring mechanism should review unexpected delays or excessive referrals. Analyses ought to account for clustering, baseline differences and potential contamination between facilities. Before scaling, investigators must test whether data systems correctly link repeat HPV tests, biopsies and referrals. Publication should report CONSORT extension requirements for any randomized implementation trial or STROBE principles for the observational version. A pragmatic assessment would allow decision-makers to judge real programme effectiveness while avoiding false equivalence between imported efficacy estimates and local clinical outcomes.

3.28 Regulatory validation of algorithms and diagnostics

Regulatory science begins with a declared intended use: primary screening, adjunctive triage, referral prioritization or research-only decision support. Evidence for one use is not proof for another. Device documentation should include analyte targets, specimen collection instructions, reproducibility, analytical interference, defined clinical populations and laboratory performance checks. For algorithmic triage, locked-model testing should precede prospective clinical evaluation, and version updates should trigger review of calibration and safety. FDA good machine-learning-practice principles and RWE guidance provide useful methodological benchmarks rather than locally automatic approvals [7,13].

Monitoring after implementation is essential because vaccination can change the class balance and referral yield over time. Calibration and subgroup sensitivity may drift even if the software code is unchanged. External audits should be empowered to review false-negative pathways, algorithm errors and inappropriate follow-up recommendations. The programme should retain simple fallback workflows when an AI component fails, since digital sophistication should never become a single point of failure in a preventive service.

3.29 Interpretation of evidence and future research agenda

Future work should prioritize external prospective comparisons of validated HPV platforms; triage strategies that minimize overtreatment; specimen logistics for self-

collection; pathway designs that reduce loss to follow-up; and risk models robust to changing vaccine-era prevalence. Research teams need to distinguish a biomarker-discovery question from an implementation question. In the former, the primary concern may be independent discrimination of CIN3+; in the latter, an intervention may be valuable primarily because more eligible women complete a previously inaccessible screening pathway.

Reporting should include preplanned subgroup evaluations, uncertainty estimates and explicit handling of incomplete verification. Equally important are studies comparing basic reminder systems with elaborate AI systems, since complexity is not equivalent to clinical benefit. Linking a validated molecular diagnostic, a transparent triage algorithm and an auditable treatment pathway presents the most defensible route to cervical cancer prevention that is both effective and equitable.

3.30 Clinical integration and audit-ready implementation scenarios

Scenario A concerns a woman eligible for first-round HPV screening who has never undergone cytology. The clinician first confirms the nationally applicable age and risk criteria, explains that HPV is common and does not automatically indicate cancer, and offers a validated assay with an approved specimen type. If the result is negative and the sample is valid, the system schedules the guideline-consistent return interval. If positive, the report should provide enough genotype or reflex-test information to determine subsequent care under the selected local algorithm. A laboratory message without a completed patient notification cannot be counted as clinical closure. Quality indicators include result turnaround time, missing or rejected specimen rates and the proportion of screening-positive women completing the recommended next step.

Scenario B concerns a woman who has an HPV16-positive result. In the setting of applicable ASCCP guidance, this can trigger colposcopic assessment rather than simply repeating an HPV test [5]. The programme records the actual appointment completion and histological endpoint instead of inferring disease severity from viral positivity. If colposcopy is delayed, a patient-navigation mechanism should identify the barrier and assist with access. The process also preserves individual autonomy and confidential communication. A management algorithm embedded in software must clearly disclose the jurisdiction, publication version and exclusion populations for which its recommendations do not apply. Clinical documentation should retain the rationale for exceptions.

Scenario C involves an individual declining speculum-based sampling because of pain, disability-related access barriers or prior distress. Where a specific validated self-collection assay is available and allowed by local rules, this alternative may enable participation. Staff must explain appropriate specimen handling and what a positive result could entail. Self-collection should not be represented as equivalent to a complete pelvic examination for symptomatic patients or as suitable for every molecular platform. The performance audit should evaluate not just the number of kits distributed but completed testing, invalid sample rate and attendance after a positive result. Patient satisfaction outcomes should be measured alongside clinical endpoints.

Scenario D concerns a previously treated CIN2+ patient entering surveillance. The patient's risk differs from that of an individual at first screening, and recall logic must incorporate previous pathology and treatment rather than relying on the most recent HPV result alone. ASCCP updates through 2023 clarify surveillance following treatment [9]. Reliable longitudinal records are particularly important when patients move between districts. Fragmented laboratory and pathology systems can cause unnecessary repeat investigations or missed intensified surveillance. Such errors are not evidence of the failure of molecular biology; they are preventable health-system weaknesses requiring targeted interoperability solutions.

These scenarios highlight a shared design principle: automated workflows should implement rules transparently and recognize when cases require human review. Programme leaders should review adverse process outcomes including loss of specimens, incorrect result routing, referral delays, incomplete treatment and incorrect surveillance intervals. Clinical quality committees could sample cases for manual verification, identify error mechanisms and publish de-identified findings. Any deployment of predictive AI should demonstrate incremental improvement over deterministic reminders. A system that labels patients high risk without providing access to timely diagnosis may increase anxiety without achieving prevention.

A nationally tailored monitoring framework could distinguish structure indicators (accredited laboratory capacity, trained readers, secure registries), process indicators (invitation reach, valid test rate, turnaround time, triage completion) and outcome indicators (CIN3+ detection among those appropriately verified, interval cancer, invasive stage distribution and treatment-related complications). Composite indices should not conceal poor performance on an essential step. Stratification is required

to show whether national improvements are equitably distributed. A well-designed baseline evaluation is essential, because projected benefits cannot honestly be reported as already observed in the absence of local measurements.

A final methodological concern is the difference between screening test verification and causal effects on cancer outcomes. CIN3+ ascertainment requires diagnostic procedures, so disease status will often be unknown among screen-negative women. Estimates of sensitivity from fully verified referral cohorts therefore cannot automatically describe asymptomatic population screening. Longitudinal linkage to cancer registries helps identify interval cancers, but does not fully remove verification bias. Analysts should report reference standard completeness and sensitivity analyses for plausible undetected disease. Intervention comparisons should also separate invitation, attendance, testing, interpretation and treatment as distinct processes. Their failure mechanisms differ and require different remedies.

A clinically mature programme should treat interpretability, follow-up completeness and fairness as co-primary dimensions of quality. Software teams may optimise a classifier using an internally held-out validation set, yet a district-wide programme may have different infection prevalence, staining practices, laboratory instruments and patient case mix. Transportability therefore needs evaluation both before adoption and continuously after rollout. Transparent procurement, laboratory accreditation and accessible clinical escalation can deliver meaningful safety benefits even before AI is considered. Every proposed system change should specify its measurable outcome, responsible professional and predefined review date.

Finally, scientific reporting must separate achieved empirical measurements from planned service improvements. Estimates derived from WHO or randomized studies may establish an evidence-informed rationale, but cannot be relabelled as observed national coverage or local effectiveness. Formal implementation research should register its protocol, predefine a primary outcome, describe missingness and provide audit trails for all derived variables. Public-health impact should be evaluated over sufficient follow-up to distinguish temporary increases in precancer detection from subsequent reductions in invasive cancer. This distinction is essential for honest interpretation of cervical cancer elimination progress.

4. Discussion

The strongest interpretation of the available evidence is that a validated HPV-based programme offers a biologically coherent and clinically effective strategy, but tests and programmes operate at different levels. The central translational challenge is integration: an assay with excellent technical performance can coexist with weak overall effectiveness if specimen transport, result notification, colposcopy scheduling or definitive treatment is unreliable. Thus programme-level evaluations should measure a linked screening cascade rather than a single cross-sectional diagnostic indicator.

Risk-based management is a response to the heterogeneity of HPV infection. A positive HPV16 result, an HPV-positive result accompanied by transforming cytology, and a low-risk genotype constellation do not imply identical probabilities of CIN3+. ASCCP extended-genotyping recommendations and WHO triage guidance illustrate the movement from binary molecular screening toward stratified follow-up [4,5]. However, a published algorithm should not be transplanted without checking test regulatory status, population risk, local colposcopy access and the consequences of over-referral.

Methodological bias warrants particular attention. Verification bias emerges when biopsies are performed mainly after positive index tests; spectrum bias occurs when case-control datasets overrepresent advanced lesions; selection bias follows clinic-based recruitment; and incorporation bias arises when the index test affects the reference diagnosis. In observational comparisons, confounding by screening history, vaccine status and service access may remain after adjustment. Missing follow-up may not be random: people who live farther from facilities may both miss colposcopy and have worse late outcomes. Multiple imputation should be sensitivity-tested against plausible missing-not-at-random scenarios rather than treated as a universal solution.

A clinically interpretable AI model should report absolute predicted risk, uncertainty and the reasons a case was referred for review. Explainable AI techniques such as SHAP or saliency maps can aid debugging but do not prove causal explanations or patient safety. Ethical AI implies consent-appropriate data use, privacy-preserving linkage, performance audit across rural and urban service areas, clinician accountability and the ability to contest an automated recommendation. An AI tool that improves average sensitivity while disadvantaging a poorly represented language group or laboratory platform would not meet acceptable deployment standards [10].

FDA's evolving RWE guidance and AI-enabled device approaches underscore the need for intended-use clarity, robust dataset quality and lifecycle validation [7,13]. EMA's RWE policy and DARWIN EU are useful examples of multi-source regulatory science, but they do not validate an individual HPV assay in Uzbekistan [8]. WHO's 2021 AI guidance emphasizes human rights and responsible governance [10]. A trial or registry protocol should specify the oversight body, software change procedures and predefined escalation when sensitivity or calibration deteriorates.

The requested cross-disciplinary drug effectiveness examples provide methodological analogies but limited clinical relevance to a cervical-screening question. Antidiabetic therapy response prediction, cardiovascular survival modelling, oncology treatment heterogeneity, antimicrobial effectiveness and neurological disease management illustrate the importance of treatment selection and longitudinal follow-up. Yet these should not displace the primary prevention endpoints of CIN3+, interval cancer, adherence to evaluation and eventual mortality. Pharmacogenomics is promising across medicine but not a current replacement for genotype-informed viral screening.

For Uzbekistan, priority should be given to low-cost, measurable improvements: a standard laboratory dictionary, a unique referral identifier, routine follow-up alerts, quality audit of pathology and publicly reported performance metrics with appropriate privacy protection. Procurement decisions should distinguish assay availability from demonstrated service readiness. An incremental implementation study across diverse facilities can assess participant acceptability, out-of-pocket burden, result turnaround, clinician workload and retention without inventing a baseline national estimate. Integration of immunization and screening registries should proceed only with appropriate legal authorization and transparent governance.

This review has important limitations. It is not a registered systematic review; no exhaustive database counts or quantitative meta-analysis are claimed. The underlying materials include official recommendations, selected randomized data and observational syntheses with differing outcomes and years. Molecular adjuncts have uneven maturity; the absence of a listed pooled estimate is deliberate where comparable verified evidence is lacking. Clinical and policy recommendations must therefore be distinguished from hypothesized benefits or proposed future research.

5. Conclusion

HPV-based primary screening is a cornerstone of cervical cancer prevention, but sustainable benefit requires validated assays, risk-adapted triage, rapid linkage to colposcopy and precancer treatment, and longitudinal follow-up. HPV16/18 and extended genotyping, p16/Ki-67 dual staining and appropriately validated self-collection can refine pathway performance when supported by local resources and quality controls. RWE can reveal participation disparities and implementation failures; causal claims demand appropriately specified designs and carefully documented data. AI may support interpretation and patient navigation, but should undergo external, prospective and subgroup validation. Uzbekistan would benefit from a phased, interoperable cervical-screening registry with transparent indicators and independent evaluation. No fabricated patient-level outcomes or original statistical analyses are presented in this review.

Declarations

Study type: critical integrative evidence review. Funding: not specified. Conflict of interest: authors to declare. Ethics: no patient-level data were collected or analysed for this manuscript. Author contributions: to be completed by the authors. Date of literature synthesis: 16 October 2026.

References

1. World Health Organization. Global strategy to accelerate the elimination of cervical cancer as a public health problem. Geneva: WHO; 2020. <https://www.who.int/publications/i/item/9789240014107>
2. World Health Organization. WHO guideline for screening and treatment of cervical pre-cancer lesions for cervical cancer prevention. 2nd ed. Geneva: WHO; 2021. <https://www.who.int/publications/i/item/9789240030824>
3. World Health Organization. WHO guideline for screening and treatment of cervical pre-cancer lesions for cervical cancer prevention: use of mRNA tests for HPV. Geneva: WHO; 2021. <https://www.who.int/publications/i/item/9789240040434>
4. World Health Organization. WHO guideline for screening and treatment of cervical pre-cancer lesions: use of dual-stain cytology after a positive HPV test. Geneva: WHO; 2024. <https://www.who.int/publications/i/item/9789240091658>

5. Massad LS, Clarke MA, Perkins RB, et al. Applying results of extended genotyping to management of positive cervicovaginal HPV test results: enduring guidelines. *J Low Genit Tract Dis.* 2025;29(2):134-143. doi:10.1097/LGT.0000000000000865
6. US Food and Drug Administration. cobas HPV for 5800/6800/8800 systems: FDA approval, May 14 2024. <https://www.fda.gov/medical-devices/recently-approved-devices/cobas-hpv-580068008800-systems-p190028s009>
7. US Food and Drug Administration. Use of real-world evidence to support regulatory decision-making for medical devices. Guidance. 2025. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-real-world-evidence-support-regulatory-decision-making-medical-devices>
8. European Medicines Agency. Real-world evidence framework to support EU regulatory decision-making: third report, February 2024–February 2025. Amsterdam: EMA; 2025. https://www.ema.europa.eu/en/documents/report/real-world-evidence-framework-support-eu-regulatory-decision-making-3rd-report-experience-gained-regulator-led-studies-february-2024-february-2025_en.pdf
9. American Society for Colposcopy and Cervical Pathology. Management guidelines and enduring guidelines process. <https://www.asccp.org/management-guidelines/>
10. World Health Organization. Ethics and governance of artificial intelligence for health. Geneva: WHO; 2021. <https://www.who.int/publications/i/item/9789240029200>
11. Sankaranarayanan R, Nene BM, Shastri SS, et al. HPV screening for cervical cancer in rural India. *N Engl J Med.* 2009;360:1385-1394. doi:10.1056/NEJMoa0808516
12. Yeh PT, Kennedy CE, de Vuyst H, Narasimhan M. Self-sampling for human papillomavirus testing: a systematic review and meta-analysis. *BMJ Glob Health.* 2019. PMID:31179035. <https://pubmed.ncbi.nlm.nih.gov/31179035/>
13. US Food and Drug Administration. Good Machine Learning Practice for Medical Device Development: Guiding Principles. <https://www.fda.gov/medical-devices/artificial-intelligence-enabled-medical-devices/good-machine-learning-practice-medical-device-development-guiding-principles>

14. World Health Organization. Cervical Cancer Elimination Initiative: achieving elimination and 90–70–90 targets. <https://www.who.int/initiatives/cervical-cancer-elimination-initiative>
15. American Society for Colposcopy and Cervical Pathology. Enduring guidelines process and published documents. <https://www.asccp.org/guidelines/management-guidelines-enduring-guidelines-process/enduring-guidelines-process/>
16. World Health Organization. Artificial Intelligence for Health. Geneva: WHO; 2024. <https://www.who.int/publications/m/item/artificial-intelligence-for-health>
17. World Health Organization. Ethics and governance of artificial intelligence for health: large multi-modal models. Geneva: WHO; 2024. <https://www.who.int/publications/b/70584>
18. World Health Organization. Recommendations and good practice statements on cervical screening and treatment. NCBI Bookshelf. <https://www.ncbi.nlm.nih.gov/books/NBK572318/>
19. US Food and Drug Administration. cobas HPV test for 4800 system: approved self-collected specimens. May 2024. <https://www.fda.gov/medical-devices/recently-approved-devices/cobas-hpv-test-4800-system-p100020s055>
20. Egemen D, Cheung LC, Chen X, et al. Risk estimates supporting the 2019 ASCCP risk-based management consensus guidelines. J Low Genit Tract Dis. 2020;24:132-143. <https://pubmed.ncbi.nlm.nih.gov/32243308/>
21. World Health Organization. Global strategy on cervical cancer elimination, publication record. <https://www.who.int/publications/b/55745>
22. World Health Organization. Regulatory considerations on artificial intelligence for health. WHO AI programme resources. <https://www.who.int/teams/digital-health-and-innovation/harnessing-artificial-intelligence-for-health>
23. World Health Organization. Health ethics and governance: big data and artificial intelligence. <https://www.who.int/teams/health-ethics-governance/emerging-technologies/big-data-and-artificial-intelligence>
24. World Health Organization. WHO consultation towards the development of

26. World Health Organization. WHO issues first global report on AI in health and six guiding principles. 2021. <https://www.who.int/news/item/28-06-2021-who-issues-first-global-report-on-ai-in-health-and-six-guiding-principles-for-its-design-and-use/>
27. Cheung LC, Egemen D, Chen X, et al. 2019 ASCCP risk-based management consensus guidelines: methods for risk estimation, recommended management, and validation. J Low Genit Tract Dis. 2020;24(2):90-101. doi:10.1097/LGT.0000000000000528
28. American Society for Colposcopy and Cervical Pathology. Self-collected vaginal specimens for HPV testing: enduring consensus recommendations. ASCCP; 2025. <https://www.asccp.org/guidelines/management-guidelines-enduring-guidelines-process/enduring-guidelines-process/>
29. Perkins RB, Guido RS, Castle PE, et al. 2019 ASCCP risk-based management consensus guidelines for abnormal cervical cancer screening tests and cancer precursors. J Low Genit Tract Dis. 2020;24(2):102-131. doi:10.1097/LGT.0000000000000525
30. US Food and Drug Administration. Clinical Decision Support Software; guidance listings. 2026. <https://www.fda.gov/vaccines-blood-biologics/general-biologics-guidances/devices-guidances>