MS ID#: MEDRXIV/2023/294826

Title page

Original Research Article

Design of the HPV-Automated Visual Evaluation (PAVE) Study: Validating a Novel Cervical Screening Strategy

Silvia de Sanjosé [1,2], Rebecca B. Perkins [3], Nicole G. Campos[4], Federica Inturrisi[1], Didem Egemen [1], Brian Befano [5,6], Ana Cecilia Rodriguez[1], Jose Jerónimo[1], Li C. Cheung[1], Kanan Desai[1], Paul Han[7], Akiva P Novetsky[8], Abigail Ukwuani[1], Jenna Marcus[9], Syed Rakin Ahmed[10,11,12,13], Nicolas Wentzensen [1], Jayashree Kalpathy-Cramer [10,14], Mark Schiffman [1] for the PAVE Study Group

1. Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD, USA
2. ISGlobal, Barcelona, Spain
3. University Chobanian and Avedisian School of Medicine/Boston Medical Center, Boston, MA, USA
4. Center for Health Decision Science, Harvard T.H. Chan School of Public Health, Boston, MA, USA
5. Information Management Services Inc, 3901 Calverton Blvd Suite 200, Calverton, MD, USA
6. Department of Epidemiology, University of Washington School of Public Health, Seattle, Washington, USA
7. Division of Cancer Control and Population Sciences, National Cancer Institute, National Institutes of Health, Rockville, MD, USA
8. Westchester Medical Center/New York Medical College, Valhalla, NY, USA
9. Feinberg School of Medicine at Northwestern University, Chicago, IL, USA
10. Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Boston, MA, USA.
11. Harvard Graduate Program in Biophysics, Harvard Medical School, Harvard University, Cambridge, MA, USA
12. Massachusetts Institute of Technology, Cambridge, MA, USA
13. Geisel School of Medicine at Dartmouth, Dartmouth College, Hanover, NH, USA
14. University of Colorado Anschutz Medical Campus, Aurora, CO, USA

Corresponding Authors: de Sanjose S/ Schiffman M

Running title: HPV-Automated Visual Evaluation (PAVE) Study

Keywords: Automated visual evaluation; Cervical cancer prevention; HPV testing; Resource-limited settings; Risk-based screening

PAVE Study Group

Brazil

Ana Ribeiro - ana-ribeiro.dantas@fiocruz.br
Tainá Raiol - taina.raiol@fiocruz.br
Center for Women's Integrated Health, Oswaldo Cruz Foundation (Fiocruz), Brasília, DF, Brazil.
MARCO Clinical and Molecular Research Center, University Hospital of Brasília/EBSERH, Federal District, Brazil

Cambodia

Te Vantha, MD, Director of Takeo Provincial Hospital,Cambodia
Thay Sovannara, MD, Medical Practitioner, Raffles Medical Group, Cambodia
Judith Norman, MD, Director of Women's Health, Mercy Medical Center, Cambodia
judynorman@gmail.com
Dr. Andrew T. Goldstein, Director, Gynecologic Cancers Research Foundation.
drg.cvvd@gmail.com

Dominican Republic

Margaret M. Madeleine, MPH, PhD
Program in Epidemiology, Fred Hutchinson Cancer Center
mmadelei@fredhutch.org
Yeycy Donastorg, MD
Instituto Dermatológico y Cirugía de la Piel "Dr. Huberto Bogaert Díaz", HIV Vaccine Trials Research Unit, Santo Domingo, Dominican Republic. ydonastorg@gmail.com

El Salvador

Miriam Cremer MD; Basic Health International, Pittsburgh, PA 15205, USA. Ob/Gyn and Women's Health Institute, Cleveland Clinic, Cleveland, OH 44195, USA.
miriam.cremer@gmail.com
Karla Alfaro, MD Basic Health International, El Salvador, kalfaro@basichealth.org

Honduras

Miriam Cremer MD; Basic Health International, Pittsburgh, PA 15205, USA. Ob/Gyn and Women's Health Institute, Cleveland Clinic, Cleveland, OH 44195, USA.
miriam.cremer@gmail.com
Karla Alfaro, MD Basic Health International, El Salvador, kalfaro@basichealth.org.
Jaqueline Figueroa, MD, Programa Nacional contra el Cáncer, Tegucigalpa, Honduras.
jacqueline_figueroan@yahoo.com

Eswatini

Eyrun F. Kjetland, MD, PhD, Professor, Departments of Global Health and Infectious Diseases Ullevaal, Centre for imported and Tropical Diseases, Oslo University Hospital Ullevaal, Oslo, Norway; College of Health Sciences, Discipline of Public Health, Nelson Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa;Centre for Bilharzia and Tropical Health Research (non-profit), BRIGHT Academy, Durban, South Africa e.f.kjetland@medisin.uio.no
Teresa Norris, Founder and President, HPV Global Action, tnorris@hpvglobalaction.org

Zeev Rosberger, PhD, Department of Oncology, Psychology and Psychiatry, McGill University, Montreal, Canada, zeev.rosberger@mcgill.ca
Amelie McFadyen, MA, Chief Executive Officer, HPV Global Action, ameliemcfadyen@hpvglobalaction.org
Marc Steben, MD, Ecole de Sante Publique, Université de Montréal; International society for STD research, marc@marcsteben.com

Malawi

Amna Haider, MD, Epidemiologist, Department of Epidemiology and Training, Epicentre, Dubai, UAE, amna.haider@epicentre.msf.org
George Kassim Chilinda, MD, Médecins Sans Frontières, Operational Centre Paris, Blantyre, Malawi, gchilinda@gmail.com
Henry B.K.Phiri, MD- Sexual and reproductive health department, Ministry of Health, Malawi, henryphiri06@gmail.com

Nigeria

Ajenifuja Kayode Olusegun, MD, Obafemi Awolowo University Teaching Hospital, Ile-Ife, Osun state Nigeria, ajenifujako@yahoo.com
Adepiti Clement Akinfolarin, MD, Obafemi Awolowo University Teaching Hospital, Ile-Ife, Osun state Nigeria, akinfolarindepiti@yahoo.co.uk
Adekunbiola Banjo, MD, College of Medicine University of Lagos, Lagos, aafbanjo@cmul.edu.ng
Moharson-Bello Imran, MD, College of Medicine, University of Ibadan, Oyo state, Nigeria, imranmorhasonbello@gmail.com
Oyinloye Temitope,MD, Obafemi Awolowo University Teaching Hospitals Complex, Ile-Ife, Osun state, Nigeria, projectcoordinator.itoju@gmail.com
Bola-Oyebamiji Sekinat, MD, College of Medicine, Osun state University, Osogbo, Osun state.
Adeyemo Marydiya, MD, College of Medicine, Osun state University, Osogbo, Osun state

Tanzania

Karen Yeates-MD, MPH, Department of Medicine, Queen's University, Kingston, Ontario, Canada, yeatesk@queensu.ca
Safina Yuma, MD, Cervical Cancer Focal Person, Ministry of Health, Tanzania, sychande@yahoo.com
Bariki Mchome, MD, Head, Reproductive Health Centre, Kilimanjaro Christian Medical Centre, Kilimanjaro, Tanzania, barikimchome@gmail.com
Alex Mremi, MD, Head, Department of Pathology, Kilimanjaro Christian Medical Centre, Kilimanjaro, Tanzania, alexmremi@gmail.com

Word count (Abstract): 469

Abstract

Objective:

To describe the HPV-Automated Visual Evaluation (PAVE) Study, an international, multi-centric study designed to evaluate a novel cervical screen-triage-treat strategy for resource-limited settings as part of a global strategy to reduce cervical cancer burden. The PAVE strategy involves: 1) screening with self-sampled HPV testing; 2) triage of HPV-positive participants with a combination of extended genotyping and visual evaluation of the cervix assisted by deep-learning-based automated visual evaluation (AVE); and 3) treatment with thermal ablation or excision (Large Loop Excision of the Transformation Zone). The PAVE study has two phases: efficacy (2023-2024) and effectiveness (planned to begin in 2024-2025). The efficacy phase aims to refine and validate the screen-triage portion of the protocol. The effectiveness phase will examine implementation of the PAVE strategy into clinical practice, cost-effectiveness, and health communication.

Study design:

Phase 1 Efficacy: Nonpregnant women, aged 25-49 years, without prior hysterectomy, are being screened at nine study sites in resource-limited settings. Eligible and consenting participants perform self-collection of vaginal specimens for HPV testing using a FLOQSwab (Copan). Swabs are transported dry and undergo testing for HPV using a newly-redesigned isothermal DNA amplification HPV test (ScreenFire), which has been designed to provide HPV genotyping by hierarchical risk groups: HPV16, else HPV18/45, else HPV31/33/35/52/58, else HPV39/51/56/59/68. HPV-negative individuals are considered negative for precancer/cancer and do not undergo further testing. HPV-positive individuals undergo pelvic examination with collection of cervical images and targeted biopsies of all acetowhite areas or endocervical sampling in the absence of visible lesions. Cervical images are used to refine a

deep learning AVE algorithm that classifies images as normal, indeterminate, or precancer+. AVE classifications are validated against the histologic endpoint of high-grade precancer determined by biopsy. The combination of HPV genotype and AVE classification is used to generate a risk score that corresponds to the risk of precancer (lower, medium, high, highest). During the efficacy phase, clinicians and patients will receive HPV testing results but not AVE results or risk scores. Treatment during the efficacy phase will be performed per local standard of care: positive Visual Inspection with Acetic Acid impression, high-grade colposcopic impression or CIN2+ on colposcopic biopsy, HPV positivity, or HPV 16,18/45 positivity. The sensitivity of the PAVE strategy for detection of precancer will be compared to current SOC at a given level of specificity.

Phase 2 Effectiveness: The AVE software will be downloaded to the new dedicated image analysis and thermal ablation devices (Liger Iris) into which the HPV genotype information can be entered to provide risk HPV-AVE risk scores for precancer to clinicians in real time. The effectiveness phase will examine clinician use of the PAVE strategy in practice, including feasibility and acceptability for clinicians and patients, cost-effectiveness, and health communication.

Conclusion:

The goal of the PAVE study is to validate a screen-triage-treat protocol using novel biomarkers to provide an accurate, feasible, cost-effective strategy for cervical cancer prevention in resource-limited settings.

Funding:

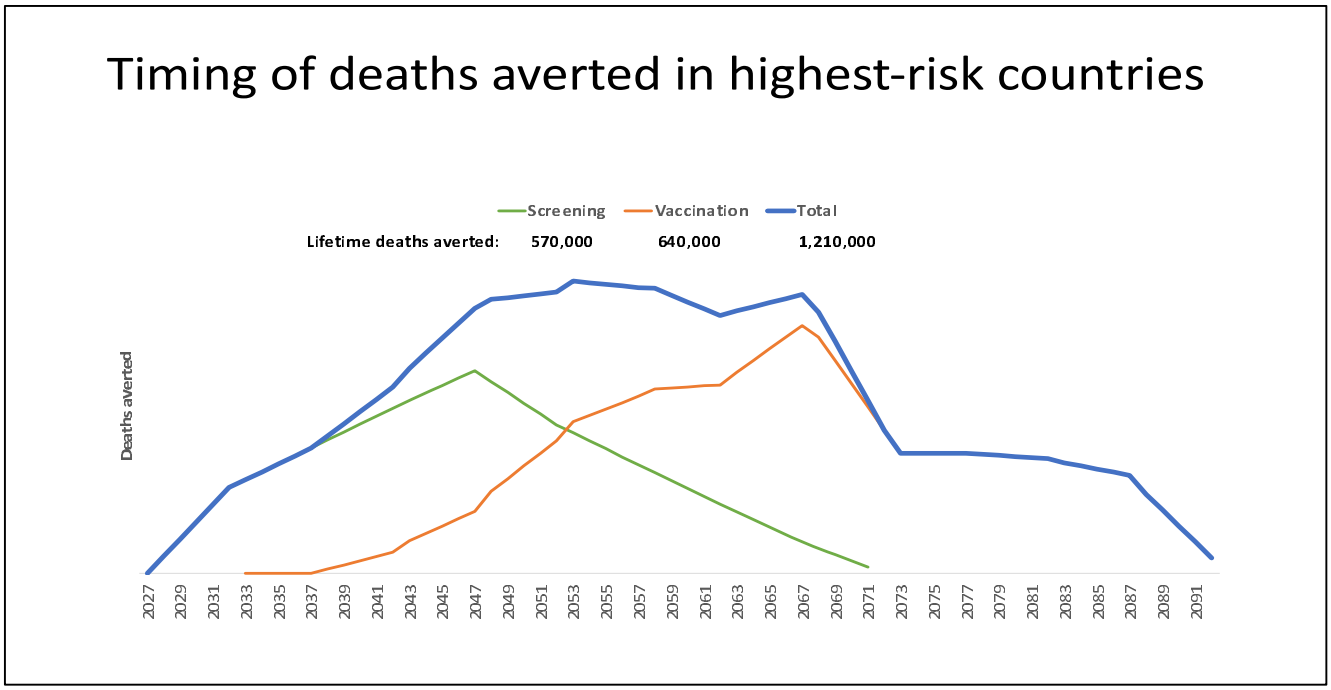
National Cancer Institute Cancer Cures Moonshot Initiative. No commercial support was obtained. Brian Befano was supported by NCI/ NIH under Grant T32CA09168.

Introduction

Global burden of cervical cancer. Cervical cancer causes substantial morbidity and mortality worldwide, with approximately 600,000 incident cases and 340,000 deaths each year.[1] Globally, cervical cancer is caused by persistent infection with one of ~13 carcinogenic human papillomavirus (HPV) types.[2] Cervical cancer rates vary greatly worldwide due to uneven access to effective preventive measures; nearly 85% of cervical cancer cases and almost 90% of cervical cancer deaths occur in low- and middle-income countries (LMIC).[3] The World Health Organization (WHO) has called for the global elimination of cervical cancer, based on an advanced understanding of the natural history of the causal carcinogenic types of cervical HPV infection and existence of effective preventive technologies, including prophylactic HPV vaccination and cervical screening.[2,4,5] However, translation of the HPV-based prevention methods has not yet occurred in many LMIC.

While prophylactic vaccination will eventually decrease cervical cancer rates[6] if high uptake can be achieved in LMIC, maximum potential health benefits of vaccinating adolescents today will not be achieved for 40 years. However, rapid implementation of a broad, effective cervical screening campaign for adult women in the highest burden areas will advance cancer control by 20 years (Figure 1). The US Cancer Moonshot initiative for Accelerated Control of Cervical Cancer has supported development of the new screening methods.[7]

Figure 1. Timing and deaths averted with one-time prevention campaigns: vaccination only, screening only, or both



Footnote Figure 1: Projection of the relative timing of health benefits, measured as deaths averted, accrued by vaccination and/or screening applied through one-time campaigns. Three scenarios were examined: 1) a one-time screening campaign providing effective management for approximately 25% of 30- to 49-year-old women in 2027 (i.e., 20 birth cohorts) (green line); 2) vaccinating 90% of 9- to 14-year-old girls in 2027 (i.e., 6 birth cohorts) with a bivalent HPV16/18 vaccination (orange line); and 3) both a screening campaign and HPV vaccination for respective birth cohorts in 2027 (blue line). We considered cervical cancer deaths averted over the lifetime of cohorts subject to the intervention, and conservatively assumed that deaths averted due to screening would only occur after age 50, to account for prevalent cancers. Projections were developed for the ~65 LMIC with age-standardized cervical cancer incidence greater than 10 per 100,000 women.[8] Even a relatively short-term intervention that combines screening and vaccination could avert over 1.2 million deaths over the lifetime of intervention cohorts, and implementation of effective screening campaigns could lead to reductions in cancer mortality almost immediately.

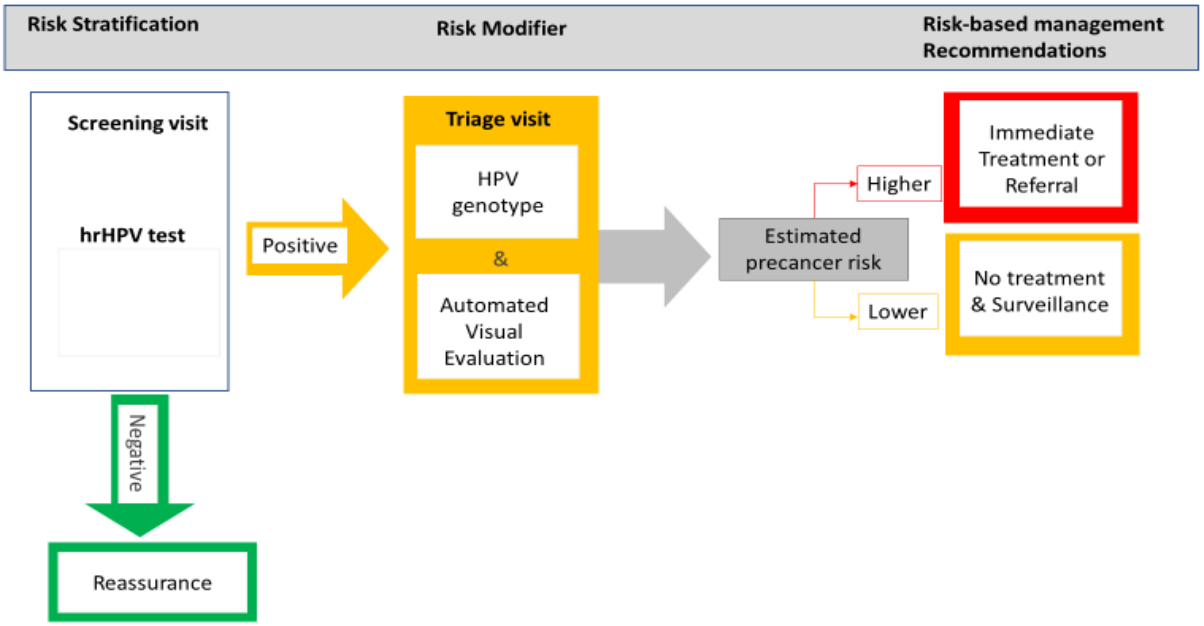
Screening using HPV testing. The detection of carcinogenic cervical/vaginal HPV DNA is currently the most sensitive screening method to distinguish those with an appreciable risk of precancer or cancer from those at low risk.[9,10] The WHO currently recommends using either screen-treat or screen-triage-treat strategies. HPV testing is preferred to using visual inspection with acetic acid (VIA) as the primary

screening method where resources permit.[4] There is a growing consensus that to achieve broad screening coverage, HPV testing of self-collected cervicovaginal specimens would be optimal for many populations.[3,4,7,11] The results from meta-analyses comparing the performance of self-collected to clinician-collected samples, using PCR-based HPV detection, showed similar sensitivity and specificity for the detection of cervical precancer. [11] As of 2022, seven LMIC were recommending HPV self-collection.[12]

There is broad consensus that HPV testing is the preferred screening method due to its high negative predictive value and reproducibility. Treatment of all HPV-positive women with thermal ablation (i.e., screen-treat strategy) is an option under current WHO guidelines[4]; however, this may substantially overtreat infections, only the minority of which would progress to cancer.[13] To make the best use of limited resources and concentrate on those at highest risk, triage strategies are critical to determine which HPV-positive women are at higher risk of cervical cancer. Triage with cytology or dual stain, as used in high resource settings, is unlikely to be a feasible solution in the majority of low-resource settings. Cervical visual examination using visual techniques, including VIA, are often used as triage methods. However, these techniques are subject to human error, have low accuracy for precancer, and require continuous training and quality control measures.[14,15] HPV genotyping is a newer, more accurate method of triage, as genotype carcinogenicity varies predictably across populations.[16,17] HPV16 is the most carcinogenic, followed by HPV18/45, followed by HPV31/33/35/52/58, followed by HPV39/51/56/59/68.[18] Automated Visual Evaluation (AVE) using an Artificial Intelligence (AI) algorithm shows promise as a relatively simple and fast triage method that could be used in conjunction with HPV genotyping to generate a highly accurate composite triage test.[19,20]

HPV-AVE (PAVE) strategy. The US National Cancer Institute (NCI) is currently undertaking a multi-centric study designed to evaluate a novel cervical screening and triage strategy for resource-limited settings, including settings with high HIV prevalence, as part of a global strategy to reduce cervical cancer burden. The PAVE strategy aims to target cervical precancer accurately and affordably by 1) self-sampled HPV screening; 2) triage among HPV-positive participants by combination of extended genotyping and visual evaluation assisted by deep-learning-based AVE; and 3) treatment using thermal ablation or excision (Large Loop Excision of the Transformation Zone (LLETZ). PAVE utilizes the concept of risk-based management, defined as patient management determined by their risk of precancer/cancer to minimize overtreatment in low-risk patients and concentrate treatment resources on high-risk patients (Figure 2). This manuscript describes the study protocol, structure, and logic of the PAVE strategy.

Figure 2. Risk-based PAVE screen-triage-treat strategy provides risk stratification to assist in the management of screening participants.



Note: hrHPV: refers to those HPV types considered as having a high potential capacity to induce cervical cancer when the infection is persistent over time. It includes HPV 16,18,45,31,33,45,52,58,39,51,56,59 and 68

Methods

The PAVE study has two phases: efficacy (2023-2024) and effectiveness (planned to begin in 2024). The efficacy phase aims to refine and validate the screen-triage portion of the protocol. The effectiveness phase will research the introduction of the PAVE strategy into clinical practice.

Phase 1: Efficacy

Setting: Study design and locations

The study aims to recruit up tens of thousands women in nine countries: Brazil, Cambodia, Dominican Republic, El Salvador, Eswatini, Honduras, Malawi, Nigeria and Tanzania (Figure 3). Criteria for study site selection included: a) existing screening programs, b) willingness to research self-sampled HPV for screening, c) capacity to run the HPV test d) availability of pathology services to process biopsies and e) access to treatment services including ablation, excision, and cancer treatment. Outreach and recruitment activities under the protocol include awareness campaigns to inform the eligible female population in the catchment area. Research protocol details including recruitment strategies, number of visits, and institutional review board approval are under the control of individual study sites. The PAVE project is integrated into the screening activities at all study sites, and at select sites is also integrated into other ongoing research studies.

Figure 3. Map of PAVE study sites

Ethical and regulatory aspects

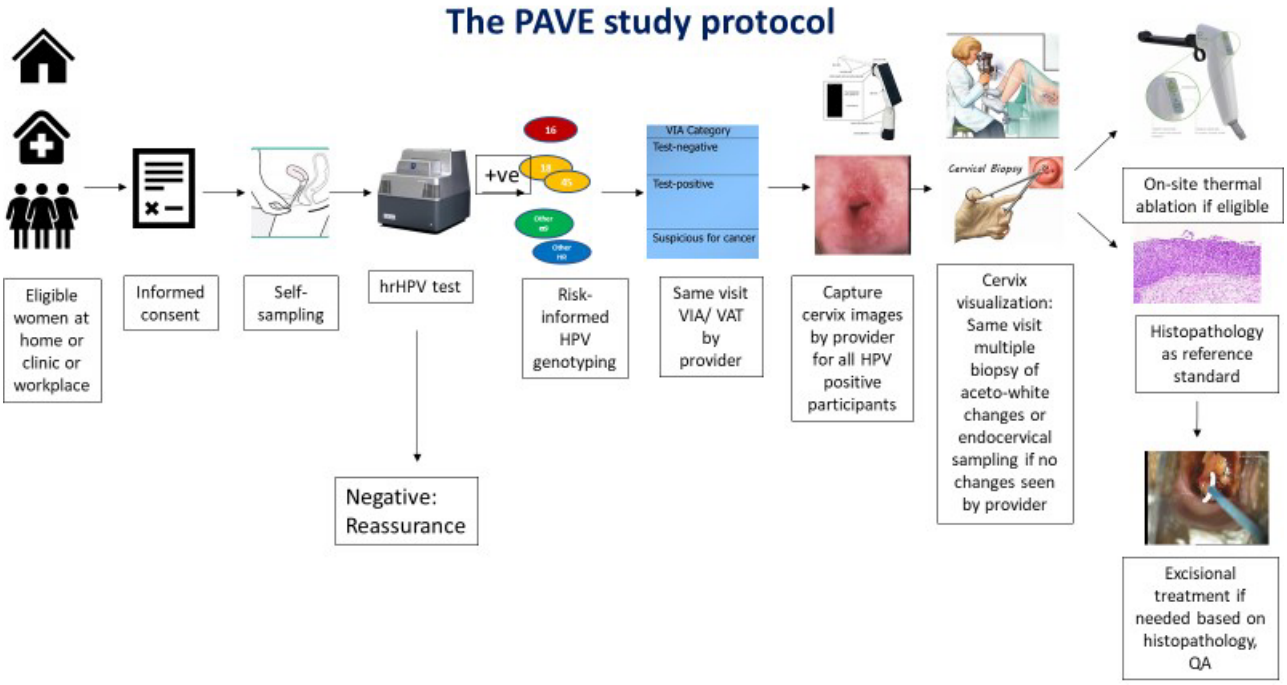
This multi-centric study is designed to function as a consortium. All ethical oversight of recruitment and clinical data collection will be done by the local sites under guidance of their own Institutional Review Boards and will follow local guidelines. All participants will have an informed consent for participation in the study and can drop their participation at any time during the process. All study documents written in languages other than English are officially translated into English for study records. Compiled analysis of de-identified data by NCI research staff for study purposes is deemed non-human subjects research by NIH.

Protocol overview

In-country elements: Patient enrollment and data collection

The steps of the PAVE protocol include 1) determination of study eligibility, 2) informed consent, 3) self-sampled HPV testing, 4) cervical image collection and biopsy collection for those testing HPV positive, and 5) treatment as indicated per local protocols.

Figure 4. Schematic of PAVE protocol elements



Enrollment. Eligibility criteria are: individuals with a cervix aged 30-49 years (general population) or 25-49 years if living with HIV (WLHIV), not currently pregnant, and able to understand study risks, benefits, and alternatives and provide informed consent in their native language. Those eligible for and interested in study participation undergo informed consent per local protocols. Those who choose to enroll provide basic demographic information (age, parity, HIV status if known).

HPV self-collection. Participants self-collect a vaginal sample for HPV testing using a FLOQSwab (Copan) following instruction by study personnel. Self-samples (FLOQSwab) are delivered dry for testing. All sites intend to use the ScreenFire HPV risk-stratification (RS) assay (ScreenFire) (US patent 11091799, Atila Biosystems Inc, Sunnyvale, CA, US).[21,22] However, other HPV tests may be acceptable alternatives if they can provide genotyping information in following groups: HPV16, HPV18/45, HPV31/33/35/52/58, and HPV39/51/56/59/68.

HPV tests are run onsite or in local laboratories in few days, with results returned to women quickly per local protocols. Women screening HPV-negative are informed of their results, reassured about their low subsequent risk of cervical cancer, and their participation in the study ends at most sites. The exception is El Salvador, at which 5% of those screening HPV negative undergo colpscopic examination. On average, approximately 80% of participants will screen HPV-negative, but this varies by study population.

Triage of HPV-positive results: image collection and biopsy. Women with HPV positive results undergo speculum exam with application of 5% acetic acid. Cervical images will be collected by a trained study provider using a dedicated device (Iris, Liger Medical LLC, Lehi, UT, US). Local clinicians also record their VIA assessment (negative, positive, suspicion of cancer) or colposcopy impression (normal, low-grade, high-grade or more severe).

Following image capture, pathology specimens are collected. Biopsies will be collected from up to four acetowhite areas for each participant. If no acetowhite areas are seen, then an endocervical sample (using curette or brush) or cytology will be collected. Sites using colposcopy will collect punch biopsies per standard practice. Sites using VIA will use Softbiopsy/SoftECC brush biopsy ™ (Histologic LLC, Anaheim, CA, US), a device that is simpler to learn and perform and is associated with lower bleeding risk. All women that have an HPV positive test are expected to have a histologic diagnosis (biopsy,

endocervical sample (ECC), and/or excisional tissue diagnosis). HPV-positive women with a negative triage evaluation initially, but who are later identified by PAVE activities to have a CIN2+, will be flagged and the clinical sites will be notified to permit "safety net" recall for adequate management.

Treatment. Treatment is provided for women meeting criteria per local protocols: VIA-positive or suspicion of cancer in sites using VIA, CIN2+ on biopsy and/or high-grade colposcopy impression in sites using colposcopy/biopsy, or HPV-positive and acetowhite changes or HPV 16,18/45 positive or HPV-positive for sites using screen-treat protocols. For sites using VIA, treatment decisions for those screening VIA-positive will be based on the standard of care, most commonly an adaptation of the WHO visual assessment for treatment (VAT) criteria for use of ablation. For sites using LLETZ, treatment decisions will follow local protocols. Table 1 describes screening, triage, biopsy, and treatment protocols for each site.

Table 1. Site-specific primary triage and treatment protocols Biopsy and Treatment Protocols

Site	Primary Screening Test	Triage method	Staff taking biopsies	Biopsy Instrument	Treatment threshold	Primary Treatment
Dominican Republic	ScreenFire HPV, Cytology	Colposcopy	Gynecologists	Biopsy forceps	CIN2+ biopsy	Ablation or LLETZ
Malawi	ScreenFire HPV	VIA	Nurses	Softbrush	VIA-positive	Ablation
Nigeria	ScreenFire HPV	Colposcopy	General Doctors, Gynecologists, GYN oncologists	Biopsy forceps	HPV-positive	Ablation or LLETZ
Brazil	ScreenFire HPV	Conventional cytology, liquid-based cytology, and colposcopy	Gynecologists	Biopsy forceps	CIN2+ biopsy or high-grade colpo impression	LLETZ
Cambodia	ScreenFire HPV	Colposcopy (mobile colposcope)	Nurse midwives, General Doctors, Gynecologists	Softbrush	VIA-positive or HPV16,18/45 positive	Ablation
Eswatini	ScreenFire HPV	VIA	Nurse midwives	Softbrush	VIA-positive	Ablation
El Salvador	ScreenFire HPV, Cytology, CareHPV	Colposcopy and VAT	General Doctors	Biopsy forceps	HPV-positive	Ablation
Tanzania	ScreenFire HPV	VIA	Gynecologists, nurses, nurse midwives	Biopsy forceps, Softbrush	VIA-positive	Ablation
Honduras	ScreenFire HPV	Colposcopy and VAT	Nurse midwives, General Doctors,	Biopsy forceps	VIA-positive	Ablation

- Malawi performed internal validation with Screenfire and Abbott HPV tests. Brazil uses cytology as an additional triage test.

Note: Prior to the PAVE study, El Salvador screened with primary HPV screening (clinician-collected CareHPV), Brazil and DR screened with cytology, and Cambodia, Eswatini, Honduras, Malawi, Nigeria, and Tanzania screened with VIA. All sites are introducing self-sampled HPV testing with ScreenFire as part of the PAVE protocol. In El Salvador, women are continuing to screen initially with both ScreenFire and CareHPV. Triage testing is performed in all women with HPV-positive results. In El Salvador, triage testing is also performed on 5% of those testing HPV negative.

Central elements: data management, HPV testing, AVE algorithm development, quality assurance, statistical analysis

Data management. Study data including demographic survey information, HPV test results (negative/positive, genotype for positive results), VIA or colposcopy impression, and local pathology results (cytology, biopsy and/or excisional specimen) are collected using DHIS2, Redcap or WEMA platforms. The collected data are associated with the corresponding images obtained during the triage visits. To ensure confidentiality, all personal identification information is removed from datasets before transferring outside the country of origin. De-identified datasets are securely transferred to a common server handled by the NGO specialized in country adaptation of DHIS2 information systems Enlace Hispano Americano de Salud (EHAS) and from there compiled data are transferred to Information Management Services (IMS), the NCI data management contractor, for data storage and analytic support during the course of the study. The noteworthy element in this study design element is that data rights (and residual biospecimens) will remain with the study site partners within their countries. The data on loan for PAVE analyses will be stored securely and data can be withdrawn and destroyed at any time by study site partners. This arrangement is important to many aspects of international data and biospecimen sharing.

HPV testing/typing

The ScreenFire HPV test is a new assay designed to detect the 13 high-risk HPV (hrHPV) genotypes grouped into the four risk groups described above, and specifically engineered to provide risk

stratification by HPV genotype based on carcinogenicity. ScreenFire includes an internal control for sample quality guidance. The ScreenFire HPV assay is an isothermal, multiplex nucleic acid amplification method that uses 3NT technology to reduce false positivity and increase assay performance. ScreenFire can be run on 1 to 96 samples per batch, requires only basic pipetting skills, and takes approximately 2.5 hours in total including sample preparation, pipetting, and DNA amplification with readout.

Validation: ScreenFire was compared against reference research HPV DNA assays in 2078 stored specimens. Overall concordance for both viral types was >90%, and sensitivity for CIN3+ was 94.2%, similar to Linear Array (93.1%) and TypeSeq (95.9%), indicating excellent performance.[22] Regulatory approval will require additional comparisons in screening settings. Additional clinical studies will be nested within the PAVE protocol. In El Salvador, 5% of HPV-negative women will undergo colposcopic evaluation. Comparison of ScreenFire to other WHO pre-qualified HPV tests (CareHPV and Abbott) may also be performed on subsets of patients.

Development of AVE algorithm using clinical images

Cervical images will be transferred via the digital camera to a secure server using a specially designed script. Images and non-PHI data will be shared and downloaded on a loan-basis to the PAVE AI team at NCI and the AI collaborators to train the AVE algorithm. Because images from the Iris device have not been used previously with the AI algorithm, a pilot phase will take place to retrain the AI-based algorithm. Due to the similarities of the Iris device to previously tested digital cameras, we expect to be able to retrain the algorithm successfully as done in our prior work.[23]

During the PAVE study, four AI algorithms are being developed and evaluated: (1) cervix detector, (2) image quality classifier, (3) disease classifier to identify precancer, and (4) treatability/SCJ classifier.

(1) Cervix detector. This algorithm is designed to display a bounding box on the screen, aiding healthcare providers ensure that the cervix is centralized within the image. This feature simplifies the process of locating the cervix within the digital image, enhancing efficiency and accuracy of image collection.

(2) Image quality classifier. This algorithm aims to identify images that may be unsuitable for accurate disease assessment due to factors like obstruction or inadequate visual sharpness. By flagging such images, it helps ensure that only good quality images are used when training the algorithm. If shown to be useful, the image quality classifier could provide feedback in real time to clinicians when taking images to ensure adequate image quality.

(3) Disease classifier. This algorithm aims to visually distinguish precancerous changes from lesser abnormalities, and classifies cervical images as normal, indeterminate, or precancer+. AVE results are assessed on repeatability (correct classification of replicate images of the same patient) and accuracy (correct classification of AVE based on histopathology, as well as minimization of extreme misclassification of normal as precancer or vice versa). Accuracy is defined as correct classification of participants to < precancer or precancer+. Precancer+ is rigorously defined to include HPV-positive histologic CIN3, AIS, cancer, and CIN2 diagnoses confirmed by expert pathologic review and positive for the 8 most carcinogenic HPV genotypes. CIN2 is qualified because although CIN2 is the threshold for treatment in most clinical practice worldwide, it is a less reproducible pathologic diagnosis and may regress without treatment, especially when associated with lower carcinogenicity HPV genotypes.[24,25]

The current prototype algorithm is the result of several years of development. Earlier algorithms were limited by poor repeatability, misclassification including grave errors (i.e., cases with precancer called normal or vice versa), overfitting, and lack of portability (defined as the ability

of the algorithm to accurately classify images from different image capture devices and study settings other than the dataset on which it was trained).[20] Additional techniques have been applied to develop the prototype version of the AVE algorithm that will be refined and tested in the PAVE study,[26] resulting in improved reliability and consistency of model predictions across repeat images from the same woman[27]. The disease classifier algorithm is trained using histology as the truth standard for defining the presence or absence of precancer. The outcome definition for the purpose of training a three-class ordinal classification algorithm includes normal (HPV-positive with histology normal, HPV negative, HPV negative cervicitis), indeterminate (low-grade HPV-related abnormalities, CIN1), and precancer+, as determined by histology among HPV positives (defined above). To ensure portability, the algorithm will undergo external validation using datasets distinct from the training set of images. Early data indicate that while our algorithm may function among patients across diverse geographies[28] a dedicated device may be needed because AVE fails to transfer without retraining.[23] The images collected in PAVE will be used to refine and externally validate the prototype AVE algorithm.[27] We will test repeatability, accuracy, and calibration of the model before the algorithm is tested in clinical settings during the effectiveness phase.[29]

(4) <u>Treatability/SCJ classifier.</u> This algorithm aims to classify the SCJ is as fully visible, partially visible, or not visible, using expert colposcopic impression as the truth standard. The goal is to assist providers in determining treatment eligibility. SCJ visibility is critical (i.e., necessary although not sufficient) for eligibility for thermal ablation procedures; ablation should not be performed when the SCJ is not fully visible.

Pathology quality assurance

Pathology readings are performed locally with centralized quality assurance on a subset of cases. Histotechnology adequacy via slide review from all participating laboratories includes assessment of specimen preparation, staining adequacy, and clarity/readability of scanned images by the assigned referent NCI study pathologist in collaboration with pathologists at each study site. Local pathologists involved in the PAVE project are asked to complete a performance competency review which include providing diagnoses on 20 standardized cases. Issues with either slide preparation or interpretation were addressed via videoconference between the NCI expert pathologist and local laboratories.

To assure histopathology reading standardization the following cases receive review by an expert gynecologic pathologist, making use of a Motic whole slide scan review collection and transfer: 1) all cases with histology CIN2+ or high-grade squamous intraepithelial lesion (HSIL); 2) all HPV16+ with <CIN2 pathology; 3) all cases read as precancer+ on AVE; and 4) 5% of biopsies read as normal. Results will be classified for the study purposes as normal, low-grade (CIN1), high-grade (CIN2, CIN3), adenocarcinoma in situ (AIS), or invasive cervical cancer (squamous, adenosquamous, adenocarcinoma, or other).

Statistical analysis

The primary objective of the PAVE protocol is to compare the sensitivity at a given specificity level of the PAVE approach for triaging HPV-positive women to current SOC (HPV testing without genotyping triaged by VIA or colposcopic impression). The PAVE screen-triage-treat protocol combines the three-class classification label (normal, indeterminate, precancer+) provided by the AVE algorithm with the four hrHPV risk group strata to create 12 strata of risk of precancer (Table 2).

Table 2. HPV-AVE risk strata

	AVE Risk Classification		
HPV risk group	Precancer+	Indeterminate	Normal
HPV16	Highest	High	High
HPV18/45	High	High	High
HPV31/33/35/52/58	High	Medium	Medium
HPV39/51/56/59/68	High	Medium	Low
Negative	Lowest		

The PAVE protocol will be compared to SOC VIA or colposcopy screen-triage protocols using the visual impressions recorded during the study. In SOC scenarios, participants are classified as positive or negative in the HPV test and as normal or abnormal in visual evaluations (e.g., VIA negative or positive, colposcopic impression less than high-grade or high-grade+) (Table 3).

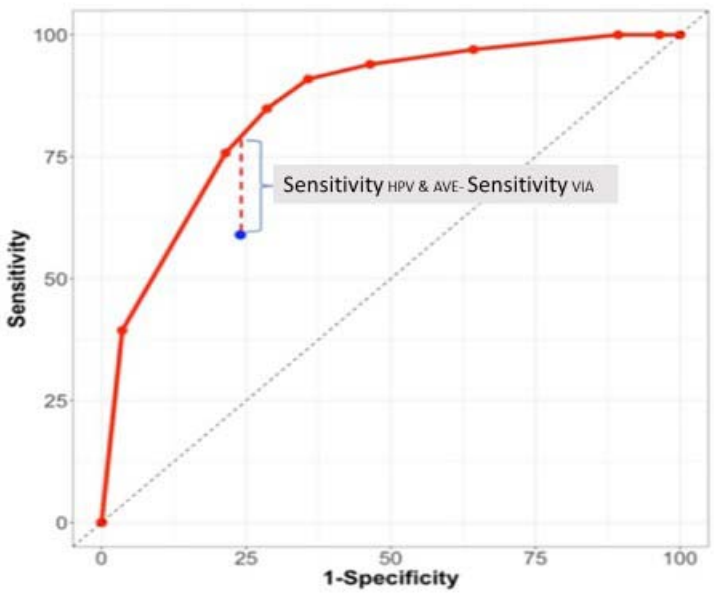
Table 3. Risk strata for participants with an HPV positive women and visual Standard of Care (SOC) as the triage (i.e. VIA, colposcopy) test

	SOC Classification*	
HPV test result	Positive/High grade	Normal
Positive	High	Low
Negative	Lowest	

Note: Participants with a negative test for HPV will not have a VIA nor colposcopy assessment. see Table 1 for SOC at individual PAVE sites.

The strata from Table 2 will be plotted as an ROC-like curve of sensitivity versus (1-specificity). The ROC curve for each study site will yield an area under the curve (AUC) for both the PAVE and the SOC strategies.

Figure 5. Theoretical approach to compare the HPV AVE strategy and the standard of care (SOC) screening and triage outcome



Footnote: Figure 5 is a hypothetical example showing how PAVE and SOC will be compared. Under a specific specificity value (which will be determined by the SOC), we will compare the difference between the sensitivities of PAVE and SOC. In this example, we had three categories for HPV genotype groups and three categories for AVE *(normal-indeterminate-precancer/cancer), and in total nine PAVE categories.*

The 12-stratum AVE risk score will be compared to visual assessment of SOC: VIA (negative or positive) or colposcopic impression (less than high grade vs. high grade+).[29] At each site, we will compare the sensitivity of the two approaches, at the (1-specificity) value produced by SOC visual evaluation. We will

compile these values for the consortium and calculate the average (weighted by study size). We hypothesize that the difference of 2 sensitivities (PAVE minus SOC) conducted on the overall consortium data will be significantly greater than zero (the null hypothesis of no difference in sensitivity), indicating that PAVE detects more precancer than SOC triage at the same level of specificity. Where available, the analysis will be stratified by HIV status. Throughout the study we will be checking for the consistency of results at the different steps across sites (e.g., performance of HPV results, quality of the images, histopathology reports). We will assess the reproducibility of the PAVE strategy across different sites by measuring the variability of the AUC values.

Additional analyses that may be performed during the efficacy phase include:

- ***Development of treatability algorithm:*** In addition to defining SCJ visibility, an AI algorithm to assess whether a lesion fulfills WHO ablation criteria is in development. If AI algorithm development is successful, output will be compared to the VIA or colposcopy assessment of eligibility for ablation or referral for surgical management. Accuracy of the three-class classification label provided by the AVE treatability algorithm (treatable, uncertain, not treatable) will be compared to the against a truth label based on the evaluation of three experts.
- ***Impact of HIV status on PAVE algorithm:*** We will evaluate the impact of HIV in the PAVE performance for the accuracy in detection precancer+ by comparing the AUC in women with and without HIV, with some consideration of role of effective ART.[23]

Phase 2: Effectiveness

During the efficacy phase, the PAVE algorithm is undergoing evaluation and development, and clinicians will not be provided with HPV genotyping, AI algorithm outputs, or risk strata. When the efficacy phase is complete, if the PAVE algorithm outperforms the SOC, we will begin the effectiveness phase. Sites that chose to participate in the effectiveness phase will test the following screen-treat-triage protocol.

Screening: self-sampled HPV testing with genotyping (ScreenFire or equivalent test). Triage of HPV positive individuals: image collection using the Iris device, upon which the AI algorithms have been installed. The AI algorithms will guide the clinician in taking high-quality images (cervix identifier, Image quality classifier) and then provide a disease classification score of normal, indeterminate, or precancer+ and an SCJ visibility assessment of fully visible, partially visible, or not visible. The clinician will then enter the HPV genotyping test result and the device will output a risk category using the strata in Table 2 (lower, medium, high, highest). The risk category and SCJ visibility assessment are designed as clinical management tools to aid clinicians in determining which patients are most likely to benefit from treatment, and among those needing treatment, whether ablation can be considered. This phase will assess the feasibility and acceptability of the PAVE strategy in clinical practice.

Cost-effectiveness analysis

To inform decisionmakers designing cervical cancer prevention programs in resource-limited settings, we will analyze the cost-effectiveness (i.e., cost per precancer detected) and affordability (the impact on a payer's budget) of the PAVE strategy at several sites. Micro-costing efforts to estimate the cost of all resource ingredients used for a screening episode are underway with technical assistance provided to study sites. A microsimulation model of genotype-specific HPV natural history and cervical carcinogenesis is being developed specifically for evaluation of novel biomarker triage tests, including AVE.[30] By adapting this natural history model to setting-specific HPV prevalence patterns (by genotype and age); overlaying screening, triage, and treatment strategies; and using setting-specific healthcare delivery data inputs for uptake, adherence to management, and costs, we will evaluate the cost per precancer detected by the PAVE strategy relative to SOC. We will explore the health and economic implications of applying different management approaches (e.g., triage, treatment, and follow up) to different risk strata, depending on health system capacity.

The development of the microsimulation model and the micro-costing tools for the PAVE consortium will serve as the basis for estimating the real-world costs and health benefits of implementing novel screening and management strategies. These tools can be adapted to different settings, with refinement of management algorithms, health care delivery variables, and cost estimates as implementation and scale-up occur. An early exercise to approximate the potential costs and benefits of a highly effective screening campaign delivered to women aged 30- 49 years in the ~65 highest burden LMIC (Figure 1; Supplementary Materials) and an HPV vaccination program delivered to girls ages 9-14 years found that the number of screening or adolescent HPV vaccinations needed to avert one cervical cancer death was similar for each intervention (i.e., 293 for screening; 278 for vaccination). Assuming a bundled cost of US$15 per woman screened and managed appropriately, a onetime screening campaign that achieves 40% uptake and ~60% adherence to recommended treatment for screen-positive women yielded a financial cost of ~US$2.5 billion to avert ~570,000 deaths, or US$4,400 per death averted. On a similar order of magnitude, a onetime single-dose bivalent HPV vaccination campaign achieving 90% coverage of girls aged 9- 14 years in the same countries (US$4.50 vaccine cost; US$7 delivery cost) would cost ~US$2.0 billion and avert ~640,000 deaths, or US$3,200 per death averted. Of note, these ballpark estimates are undiscounted and do not account for cancer treatment cost offsets. While data are not yet available on the costs of implementing novel highly effective screening strategies for adult women and single-dose HPV vaccination for female adolescents, these data are forthcoming from the PAVE consortium and single-dose vaccination studies. Refining these cost and effectiveness estimates, and obtaining country-specific data, is a high priority and a critical component of the PAVE consortium objectives.

Stakeholder knowledge and attitudes regarding cervical cancer prevention and screening interventions in the PAVE Study

Effective dissemination and implementation of the PAVE strategy in the future will require clear and consistent strategies for communicating information to healthcare providers and patients. The field of HPV and screening is rapidly evolving and constantly being enriched with new scientific information. However, this influx of information can sometimes lead to unclear or conflicting messages, which in turn may diminish the effectiveness of interventions aimed at improving screening rates and optimizing management strategies. Furthermore, both healthcare providers and patients can differ in their knowledge and perceptions of cervical cancer risk, tolerance of risk, and their personal values and attitudes regarding cervical cancer prevention and screening, all of which can influence uptake of prevention strategies such as the PAVE strategy. To address these challenges, it will be necessary to provide healthcare providers with accurate, consistent information about the latest scientific advances and guidelines, and as well as training and tools that can help them effectively inform and engage patients in cervical cancer prevention and screening.

To address these needs and prepare for broader dissemination and implementation of the PAVE strategy, the Communication and Retention Workgroup is conducting a mixed-methods pilot study utilizing both qualitative interviews and survey questionnaires administered to scientific experts as well as key stakeholders—patients and healthcare providers—at four participating sites (Brazil, El Salvador, Nigeria, Tanzania). The specific objectives of this pilot study are to explore stakeholders' knowledge, perceptions, and attitudes regarding cervical cancer prevention and screening, and their preferences for information and participation in decision making. The study will generate evidence to enable the future development of effective, ethical strategies for engaging eligible members of LMIC communities in cervical cancer prevention and screening, including using the PAVE strategy.

Discussion

Timeline and Next Steps

At the time of this writing, the study has been initiated in five countries. It is expected that a preliminary interim analysis on the efficacy of the PAVE strategy will be completed by the end of 2023, and field recruitment will be completed by July 2024. The efficacy phase is designed to assess the validity of the PAVE protocol by refining the AVE protocol using histopathology specimens and demonstrate the superiority of PAVE for detecting precancer and minimizing unnecessary referrals compared to the SOC. Following the initial efficacy phase, the effectiveness phase is planned to examine additional factors including program feasibility, acceptability, and cost-effectiveness.

Regulatory considerations

The rapid advancement of AI in healthcare has prompted significant ethical and regulatory discussions. Global ethical principles specific to AI in healthcare are in development and are rapidly evolving, regulatory authorities such as the WHO and U.S. Food and Drug Administration (FDA) are responsible for developing guidelines that ensure the safe, effective, and appropriate use of AI technologies in healthcare and therapeutic development. In addition to complying with existing ethical principles in medicine, AI solutions must demonstrate scientific validity, ensuring their effectiveness and reliability. It is imperative that AI technologies do not perpetuate discrimination or bias, or exclude specific population segments. One major concern is the potential creation of a permanent digital identity, linked to individuals' health and personal data, without obtaining proper consent.[31,32]

The Iris device, used in PAVE, is an example of an AI-based system in healthcare. It will incorporate AI algorithms that aim to generate scores evaluating image quality, presence of precancerous lesions, and visibility of the SCJ based on digital cervix images. These scores provide additional information to assist healthcare providers in making informed decisions. To arrive at a comprehensive decision, the provider will need to integrate the HPV information, gynecological examination findings, scoring system outputs

generated from the AI algorithms, and the patient's medical history. Considering its functionality, the AI algorithm used in the PAVE system, based on FDA specifications, can be considered an assistant to medical management offering the joint evaluation of the HPV data and the AVE outcome. As WHO is actively evaluating screening and triage approaches, it is expected that clear regulatory guidance on the use of the algorithms will be available before their implementation .[31]

Competition and commercial aspects

The PAVE strategy utilizes certain products of significant commercial importance. The selection of each device or assay is based on factors such as accuracy to achieve its purpose (i.e. high sensitivity of the ScreenFire HPV test to detect precancer, high-quality image capture by the Iris device) and affordability. It is important to note that the PAVE project does not preclude future commercial developments. The presence of competition with similar assays is desirable as long as they can provide similar qualities. If the PAVE strategy is proven to be more accurate than the SOC for screening and triage, there will likely be an increased demand for devices and assays. Procurement of these resources may present challenges and require effective communication with the respective Ministries of Health to ensure a smooth introduction in countries where the need is evident.

Dissemination

After the demonstration of efficacy and effectiveness, we expect to have a screen-treat-triage protocol fulfilling WHO principles that can be used at the discretion of local health authorities in LMIC for cervical cancer prevention programs. If the PAVE strategy is proven to outperform the local SOC, and be feasible, acceptable, and affordable for resource-limited settings, then countries may switch their current SOC to the PAVE strategy. De-implementation of existing strategies, such as cytology, colposcopy, and VIA, as well as implementation of self-sampled HPV with AVE triage will require buy-in from stakeholders and policymakers as well as substantial investment in educating and retraining the laboratory and clinical

workforces. Cytology/colposcopy programs, though effective, are limited in scope and are costly to maintain, therefore switching may be attractive to health ministers. However, laboratories will need funding for the purchase of equipment for running HPV testing and materials for self-sampling, and the cytotechnologist and pathologist workforce will be reduced. Countries with existing VIA programs will require significant introduction costs such as laboratory machinery and training of healthcare personnel, and recurrent costs including reagents and self-sampling materials. In all settings, program continuation beyond the initial study will require local governments and programs to address issues of procurement and implementation, as well as de-implementation of existing strategies.

In conclusion, the PAVE project will develop and validate a strategy using self-sampled HPV with genotyping and AVE to identify precancer in a large group of women from many different settings. The PAVE objective is to create an accurate, feasible, cost-effective screening and triage protocol for cervical cancer prevention in resource-limited settings. If proven effective, cost-effective, feasible, and acceptable, the strategy can have a major impact in reducing cervical cancer among non-vaccinated adult women.

Contributors: All authors affirm that this commentary is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted, and that any discrepancies from the study as planned (and, if relevant, registered) have been explained. This research was conducted with the appr approval of the institutional review board of the National Cancer Institute.

Declaration of interests: No authors have a conflict of interest to declare related to this work

References

1. Singh D, Vignat J, Lorenzoni V, et al. Global estimates of incidence and mortality of cervical cancer in 2020: a baseline analysis of the WHO Global Cervical Cancer Elimination Initiative. *Lancet Glob Health*. 2023;11(2):e197-e206. doi:10.1016/S2214-109X(22)00501-0

2. Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. *The Lancet*. 2007;370(9590):890-907. doi:10.1016/S0140-6736(07)61416-0

3. IARC. Cervical Cancer Screening IARC Handbooks of Cancer Prevention Volume 18. Published online 2022. Accessed December 23, 2022. https://publications.iarc.fr/Book-And-Report-Series/Iarc-Handbooks-Of-Cancer-Prevention/Cervical-Cancer-Screening-2022

4. World Health Organization. WHO guideline for screening and treatment of cervical pre-cancer lesions for cervical cancer prevention, second edition. Published online 2021.

5. WHO. Global strategy to accelerate the elimination of cervical cancer as a public health problem. Published online November 17, 2020. Accessed June 20, 2023. https://www.who.int/publications/i/item/9789240014107

6. Lei J, Ploner A, Elfström KM, et al. HPV Vaccination and the Risk of Invasive Cervical Cancer. *N Engl J Med*. 2020;383(14):1340-1348. doi:10.1056/NEJMoa1917338

7. Perkins RB, Smith DL, Jeronimo J, et al. Use of risk-based cervical screening programs in resource-limited settings. *Cancer Epidemiol*. 2023;84:102369. doi:10.1016/j.canep.2023.102369

8. Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I, Bray F. Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer. Published online 2020. Accessed March 8, 2023. https://gco.iarc.fr/today,

9. Demarco M, Egemen D, Hyun N, et al. Contribution of Etiologic Cofactors to CIN3+ Risk Among Women With Human Papillomavirus-Positive Screening Test Results. *J Low Genit Tract Dis*. 2022;26(2):127-134. doi:10.1097/LGT.0000000000000667

10. Castle PE, Glass AG, Rush BB, et al. Clinical Human Papillomavirus Detection Forecasts Cervical Cancer Risk in Women Over 18 Years of Follow-Up. *Journal of Clinical Oncology*. 2012;30(25):3044-3050. doi:10.1200/JCO.2011.38.8389

11. Arbyn M, Smith SB, Temin S, Sultana F, Castle P, Collaboration on Self-Sampling and HPV Testing. Detecting cervical precancer and reaching underscreened women by using HPV testing on self samples: updated meta-analyses. *BMJ*. 2018;363:k4823. doi:10.1136/bmj.k4823

12. Serrano B, Ibáñez R, Robles C, Peremiquel-Trillas P, de Sanjosé S, Bruni L. Worldwide use of HPV self-sampling for cervical cancer screening. *Preventive Medicine*. 2022;154:106900. doi:10.1016/j.ypmed.2021.106900

13. Demarco M, Hyun N, Carter-Pokras O, et al. A study of type-specific HPV natural history and implications for contemporary cervical cancer screening programs. *EClinicalMedicine*. 2020;22:100293. doi:10.1016/j.eclinm.2020.100293

14. Sankaranarayanan R, Nene BM, Dinshaw K, et al. Early detection of cervical cancer with visual inspection methods: a summary of completed and on-going studies in India. *Salud Publica Mex*. 2003;45 Suppl 3:S399-407.

15. Catarino R, Schäfer S, Vassilakos P, Petignat P, Arbyn M. Accuracy of combinations of visual inspection using acetic acid or lugol iodine to detect cervical precancer: a meta-analysis. *BJOG*. 2018;125(5):545-553. doi:10.1111/1471-0528.14783

16. Sankaranarayanan R, Nene BM, Shastri SS, et al. HPV Screening for Cervical Cancer in Rural India. *N Engl J Med*. 2009;360(14):1385-1394. doi:10.1056/NEJMoa0808516

17. Qiao YL, Sellors JW, Eder PS, et al. A new HPV-DNA test for cervical-cancer screening in developing regions: a cross-sectional study of clinical accuracy in rural China. *Lancet Oncol*. 2008;9(10):929-936. doi:10.1016/S1470-2045(08)70210-9

18. Guan P, Howell-Jones R, Li N, 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):2349-2359. doi:10.1002/ijc.27485

19. Desai KT, Ajenifuja KO, Banjo A, et al. Design and feasibility of a novel program of cervical screening in Nigeria: self-sampled HPV testing paired with visual triage. *Infect Agent Cancer*. 2020;15:60. doi:10.1186/s13027-020-00324-5

20. Desai KT, Befano B, Xue Z, et al. The development of "automated visual evaluation" for cervical cancer screening: The promise and challenges in adapting deep-learning for clinical testing: Interdisciplinary principles of automated visual evaluation in cervical screening. *Int J Cancer*. 2022;150(5):741-752. doi:10.1002/ijc.33879

21. Desai KT, Adepiti CA, Schiffman M, et al. Redesign of a rapid, low-cost HPV typing assay to support risk-based cervical screening and management. *Int J Cancer*. 2022;151(7):1142-1149. doi:10.1002/ijc.34151

22. Inturrisi F, de Sanjose S, Desai KT, Dagnal C, Egemen D, Befano B, Rodrigue AC, Jerónimo J, Zuna RE, Hoffman A, Nozzari SF, Walker JL, Perkins RB, Wentzensen N, Palefsky JM, Schiffman M. A rapid HPV typing assay to support cervical cancer screening and risk-based management: a cross-sectional validation study. *Under Review*.

23. Parham G, Egemen D, Befano B, Mwanahamuntu M, Rodriguez AC, Antani S, Chisele S, Munalula MK, Kaunga F, de Sanjose S, Schiffman M, Sahasrabuddhe V. Validation in Zambia of a cervical screening strategy including HPV genotyping and artificial intelligence (AI)-based automated visual evaluation. *Under Review*.

24. Lee MH, Finlayson SJ, Gukova K, Hanley G, Miller D, Sadownik LA. Outcomes of Conservative Management of High Grade Squamous Intraepithelial Lesions in Young Women: *Journal of Lower Genital Tract Disease*. 2018;22(3):212-218. doi:10.1097/LGT.0000000000000399

25. Kylebäck K, Ekeryd-Andalen A, Greppe C, Björkenfeldt Havel C, Zhang C, Strander B. Active expectancy as alternative to treatment for cervical intraepithelial neoplasia grade 2 in women aged 25 to 30 years: ExCIN2-a prospective clinical multicenter cohort study. *Am J Obstet Gynecol*. Published online June 29, 2022:S0002-9378(22)00527-0. doi:10.1016/j.ajog.2022.06.051

26. Ahmed SR, Befano B, Lemay A, et al. Reproducible and Clinically Translatable Deep Neural Networks for Cancer Screening. *Res Sq*. Published online March 3, 2023:rs.3.rs-2526701. doi:10.21203/rs.3.rs-2526701/v1

27. Lemay A, Hoebel K, Bridge CP, et al. Improving the repeatability of deep learning models with Monte Carlo dropout. *NPJ Digit Med*. 2022;5(1):174. doi:10.1038/s41746-022-00709-3

28. Ahmed SR, Egemen D, Befano B, Rodriguez AC, Jeronimo J, de Sanjose S, Kalpathy-Cramer J, Schiffman M. Assessing generalizability of an ai-based visual test for cervical cancer screening. *Under Review*.

29. Egemen D, Perkins R, Cheung L, Befano B, Rodriguez AC, Desai K, Lemay A, Ahmed SR, Antani S, Jeronimo J, Wentzensen N, Kalpathy-Cramer J, de Sanjose S, Schiffman M. AI-based image analysis in clinical testing: lessons from cervical cancer screening. *Under Review*.

30. Campos NG, Demarco M, Bruni L, et al. A proposed new generation of evidence-based microsimulation models to inform global control of cervical cancer. *Prev Med*. 2021;144:106438. doi:10.1016/j.ypmed.2021.106438

31. WHO. Ethics and governance of artificial intelligence for health. Published online June 28, 2021. Accessed June 29, 2023. https://www.who.int/publications/i/item/9789240029200

32. FDA. Artificial Intelligence and Machine Learning in Software as a Medical Device. Accessed July 13, 2023. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device