ORIGINAL ARTICLE

Deep learning-assisted digital VIA via telemedicine and HPV self-sampling for CIN2+ detection in remote archipelago women: a prospective diagnostic accuracy study

Theresia Putri Sinaga[1*], Nur Diana[2], Firman Hadi[3], Gayatri Putri[4]

[1] Department of Public Health, CMHC Research Center, Palembang, Indonesia
[2] Department of Molecular Biology, CMHC Research Center, Palembang, Indonesia
[3] Department of Obstetrics and Gynecology, Lumajang Private Family Clinic, Lumajang, Indonesia
[4] Department of Library and Informatics Science, Enigma Institute, Palembang, Indonesia

ARTICLE INFO

ARTICLE TYPE
Original Article

ARTICLE HISTORY
Received: February 15, 2026
Accepted: March 29, 2026
Published: April 30, 2026

KEYWORDS
Deep learning; early detection of cancer; papillomavirus infections; telemedicine; uterine cervical neoplasms

CORRESPONDING AUTHOR
Theresia Putri Sinaga
E-mail: theresia.ps@cattleyacenter.id
ORCID: 0009-0009-9019-3689
DOI: 10.59345/sjog.v4i1.302

ABSTRACT

Background: Cervical cancer disproportionately burdens women in low- and middle-income countries, where archipelagic geography and colposcopist shortages obstruct screening. Human papillomavirus (HPV) self-sampling is highly sensitive but poorly specific, generating colposcopy referrals that exceed remote-area capacity.
Objective: To evaluate whether deep-learning (DL)-assisted digital visual inspection with acetic acid (VIA), delivered by telemedicine, can triage HPV-positive women and detect high-grade cervical intraepithelial neoplasia (CIN2+).
Methods: In a prospective, double-blind, STARD-compliant diagnostic accuracy study, 642 women aged 30–50 years at an urban tertiary referral hospital (Center A) and five remote community health centers (Region B) in an Indonesian archipelago province underwent HPV-DNA self-sampling and smartphone-captured digital VIA analyzed by a MobileNetV2 convolutional neural network. All participants received colposcopy-directed biopsy as the reference standard, eliminating verification bias. Metrics used Wilson 95% confidence intervals (CI); multivariable logistic regression and decision-analytic triage metrics were derived.
Results: CIN2+ prevalence was 13.1% (95% CI 10.7–15.9). DL-assisted VIA achieved sensitivity 91.7% (95% CI 83.8–95.9), specificity 88.4% (85.4–90.8), and AUC 0.93, exceeding human-read VIA (sensitivity 67.9%). HPV self-sampling was most sensitive (95.2%) but least specific (81.5%). A sequential HPV→DL-VIA pathway raised specificity to 95.7% and positive predictive value to 75.5%, reducing colposcopy referrals by 46.4% and unnecessary referrals by 76.7% (number-needed-to-screen 7.6). HPV positivity dominated the multivariable model (adjusted OR 91.5, 95% CI 32.7–256.2, $p<0.001$; Nagelkerke R^2 0.50).
Conclusion: Telemedicine-delivered, DL-assisted VIA is an accurate triage for HPV-positive women that conserves scarce colposcopy capacity while preserving CIN2+ detection. This decentralized two-step pathway is a scalable strategy for advancing cervical-cancer elimination in geographically isolated populations.

All authors have reviewed and approved the final version of the manuscript.

1. Introduction

Cervical cancer remains the fourth most common malignancy in women worldwide, with an estimated 604,000 new cases and 342,000 deaths in 2020, and roughly 90% of that mortality concentrated in low- and middle-income countries (LMICs).[1-3] The burden is determined less by biology than by access: across Asian and other LMIC settings, screening coverage falls far below the level required for population impact, with the steepest deficits among rural and geographically isolated women.[4,5] In response, the World Health Organization (WHO) launched a global elimination strategy built on three targets—90% HPV vaccination, 70% twice-lifetime high-performance screening, and 90% treatment of detected disease—that reframes these disparities as measurable, actionable goals.[2,4]

Persistent infection with high-risk human papillomavirus (hrHPV) is the necessary cause of cervical carcinogenesis, which progresses through cervical intraepithelial neoplasia (CIN) grades 1 to 3 before invasion.[6,7] Because the overwhelming majority of hrHPV infections clear spontaneously, a positive HPV test is highly sensitive but only modestly specific for the actionable lesion, high-grade neoplasia (CIN2 or worse, CIN2+).[6,8] This biology creates the central operational problem of modern screening: how to triage a large HPV-positive population efficiently toward the minority who genuinely require colposcopy and treatment, without overwhelming limited specialist capacity.[8,9]

The two innovations target biologically distinct steps of carcinogenesis, and this complementarity is the conceptual core of the present study. A molecular HPV test detects the necessary cause of cervical cancer and is therefore exquisitely sensitive but cannot indicate whether infection has yet produced a lesion; an acetowhite imaging test detects the phenotypic tissue change of established neoplasia and is therefore more specific for the actionable endpoint.[6,10]

Quantifying that complementarity prospectively, in the population that stands to benefit most, is the specific knowledge gap this study addresses; prior evaluations have largely tested each modality in isolation, in referral populations, and against incompletely verified reference standards, leaving program planners without a direct, head-to-head estimate of how a combined pathway would behave in the field.[10-12]

Two technologies have transformed the screening landscape. HPV self-sampling allows women to collect their own specimens, and randomized and meta-analytic evidence confirms that testing of self-collected samples is non-inferior to clinician collection for CIN2+ detection while substantially increasing participation among under-screened women.[6,13,14,15] Separately, artificial intelligence (AI) has begun to address the subjectivity and inter-observer variability that have long limited visual inspection with acetic acid (VIA): deep convolutional neural networks performing automated visual evaluation of cervical images have matched or exceeded expert review, and lightweight architectures deployable on smartphones bring this within reach of frontline clinics.[10,12,16,17] Systematic reviews report high pooled accuracy for AI in cervical imaging while cautioning that most validations are retrospective, single-modality, and conducted in referral rather than true screening populations.[11,18]

Indonesia illustrates the unmet need with particular clarity. Its archipelagic geography, midwife-dependent primary care, and acute scarcity of colposcopists make centralized referral pathways logistically and economically impractical for remote island populations, and national screening coverage remains fragmentary.[3,5] For a woman on an outer island, every HPV-positive result that mandates travel to a distant tertiary center represents a formidable barrier to completing the screening cascade.[4]

Despite the maturity of each individual technology, few studies have integrated the two most scalable innovations—HPV self-sampling and AI-assisted VIA—into a single sequential pathway, evaluated prospectively against a verification-bias-free histopathologic reference, and situated in the archipelagic, colposcopist-scarce setting where decentralization matters most.[10,11,12] We therefore conducted a prospective, double-blind, STARD-compliant diagnostic accuracy study in an Indonesian archipelago province to evaluate a decentralized two-step pathway: primary HPV-DNA self-sampling followed by deep-learning-assisted digital VIA triage transmitted via telemedicine, benchmarked against colposcopy-directed biopsy for detecting CIN2+. Beyond conventional diagnostic metrics, we quantified the decision-analytic consequence most relevant to planners—how much colposcopy demand a sequential AI triage can avert without sacrificing high-grade-lesion detection.

2. Methods

Design and reporting. This was a prospective, double-blind diagnostic accuracy study conducted and reported in accordance with the Standards for Reporting of Diagnostic Accuracy Studies (STARD 2015). Index-test readers and reference-standard pathologists were mutually blinded, and all participants underwent the reference standard irrespective of index-test results to eliminate partial-verification (work-up) bias.

Setting. Recruitment took place across two anonymized geographies within a single Indonesian archipelago province: a tertiary referral hospital serving as the central telemedicine hub and histopathology laboratory (Urban Center A), and five remote community health centers (Puskesmas) equipped with mobile telemedicine kits (Archipelago Region B). No institutional identifiers are reported.

Participants. Eligible participants were non-pregnant women aged 30–50 years with an intact cervix, no history of cervical cancer, and residence within the study catchment areas. Women who were currently menstruating, had undergone hysterectomy, or had previously been treated for CIN were excluded. Written informed consent was obtained from every participant.

Sample size. Assuming an expected DL-VIA sensitivity of 90% with a 95% confidence half-width of ±8% at an anticipated CIN2+ prevalence of 12–15%, and a specificity of 85% estimated within ±4%, the specificity-precision requirement dominated and yielded a minimum of 600 evaluable participants; 642 were enrolled to allow for histopathology dropouts. All tests were two-sided with α=0.05.

Index test 1 – HPV-DNA self-sampling. Each participant received a validated vaginal self-sampling brush with pictorial instructions. Specimens were tested for 14 hrHPV genotypes using a multiplex real-time PCR assay that distinguished HPV16 and HPV18 individually from a pooled channel of twelve other high-risk types; a positive result for any hrHPV type was scored as test-positive, enabling a pre-specified genotype-stratified analysis.

Index test 2 – DL-assisted digital VIA via telemedicine. Trained midwives in Region B applied 5% acetic acid and captured standardized cervicography images at one minute post-application using a smartphone with a calibrated colposcopic attachment. Images were transmitted over an encrypted cloud platform to Center A. A convolutional neural network based on MobileNetV2, pre-trained on ImageNet and fine-tuned on 15,000 expertly annotated VIA images, classified each image as Normal/Benign, Low-grade/Suspect, or High-grade/CIN2+; the High-grade output defined a positive DL-VIA triage. The training set was strictly disjoint from the evaluation cohort, the classification threshold was fixed a priori on an independent validation partition, and an automated image-quality gate prompted recapture of unusable frames. Human-read VIA by the on-site provider was recorded independently for comparison.

Model development and reporting. The deep-learning classifier was developed and is reported in accordance with contemporary clinical-AI reporting principles (CLAIM and TRIPOD-AI). The 15,000 training images were annotated against histopathology where available and expert colposcopic consensus otherwise, and were collected separately from the present cohort; no participant image contributed to model training, precluding information leakage. Image preprocessing included illumination normalization and cropping to the cervical region, and the model output a calibrated probability for each of the three categories. The operating threshold defining a High-grade output was fixed on an independent validation partition before the evaluation cohort was analyzed, and the number of indeterminate outputs after the automated quality gate is reported in the Results. Model weights and inference code are available from the authors on reasonable request.

Reference standard. All participants, regardless of index-test results, underwent colposcopy with biopsy of any acetowhite lesion (and random quadrant biopsy where no lesion was visible), performed centrally or during scheduled specialist outreach visits. Histopathology was independently reported by two expert pathologists blinded to index results, with discordance resolved by consensus and p16 immunohistochemistry used to adjudicate equivocal CIN2. The threshold for reference-positive disease was CIN2+.

Statistical analysis. Sensitivity, specificity, positive and negative predictive values (PPV, NPV), and positive and negative likelihood ratios (LR+, LR-) were computed for each modality with Wilson 95% confidence intervals (CI). The area under the receiver-operating-characteristic curve (AUC) was estimated for each modality. The two primary comparisons—DL-assisted VIA versus human-read VIA, and the sequential strategy versus referring all HPV-positive women—were tested with paired McNemar tests and DeLong AUC comparisons; other comparisons were pre-specified as exploratory. Agreement with the histopathologic reference was quantified by Cohen's κ with 95% CI. Independent determinants of CIN2+ were assessed by multivariable logistic regression using Firth's penalized likelihood (adjusted odds

ratios [aOR], 95% CI, Nagelkerke R^2). A decision-analytic evaluation of the sequential pathway quantified the reduction in total and unnecessary colposcopy referrals, the number-needed-to-screen (NNS) to avert one unnecessary colposcopy with its 95% CI, and the grade distribution of additionally missed CIN2+. Subgroup performance by geography and age was exploratory. Analyses used Python 3 (NumPy); two-sided $p<0.05$ was significant and exact p-values are reported to three decimals.

Ethics. The study was approved by the Combined Medical & Health Research Ethics Committee (CMHC; approval No. CMHC/EC/2024/0461) and conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent, and data were de-identified prior to analysis.

3. Results and Discussion

Participants and baseline characteristics. Of 712 women approached, 658 consented and 642 (97.6% of consenters) completed both index tests and the reference standard; 210 were screened at Urban Center A and 432 in Archipelago Region B. The mean age was 41.2 (SD 5.4) years and did not differ between settings (p=0.120). Region B participants were more frequently of high parity (≥3 births: 50.5% vs 24.8%, p<0.010) and of primary or no formal education (37.5% vs 11.4%, p<0.010), reflecting the socio-economic gradient between urban and archipelagic communities, as detailed in Table 1. The two settings were otherwise comparable in age, and the balanced enrollment (210 urban and 432 remote participants) provided adequate numbers for the pre-specified geographic subgroup analysis.

Prevalence of high-grade disease. Histopathology confirmed CIN2+ in 84 women, an overall prevalence of 13.1% (95% CI 10.7–15.9). As shown in Table 1, prevalence was statistically indistinguishable between Center A (13.3%, 95% CI 9.4–18.6) and Region B (13.0%, 95% CI 10.1–16.5; p=0.880), and a further 8.7% had CIN1, confirming that the remote population carried a disease burden equal to the referral population it is so often unable to reach.

Table 1. Demographic and clinical characteristics of participants (N = 642).

Characteristic	Total (N=642)	Center A (n=210)	Region B (n=432)	p
Age (years), mean (SD)	41.2 (5.4)	42.1 (5.1)	40.8 (5.5)	0.120
Nulliparous, n (%)	58 (9.0)	28 (13.3)	30 (6.9)	0.040
Parity 1–2, n (%)	314 (48.9)	130 (61.9)	184 (42.6)	<0.010
Parity ≥3, n (%)	270 (42.1)	52 (24.8)	218 (50.5)	<0.010
Primary/no education, n (%)	186 (29.0)	24 (11.4)	162 (37.5)	<0.010
Secondary/tertiary education, n (%)	456 (71.0)	186 (88.6)	270 (62.5)	—
Histology: Normal/cervicitis, n (%)	502 (78.2)	160 (76.2)	342 (79.2)	0.450
Histology: CIN1, n (%)	56 (8.7)	22 (10.5)	34 (7.9)	0.380
Histology: CIN2+, n (%)	84 (13.1)	28 (13.3)	56 (13.0)	0.880

Diagnostic accuracy of individual modalities. As detailed in Table 2 and illustrated in Figure 1, DL-assisted VIA detected CIN2+ with a sensitivity of 91.7% (95% CI 83.8–95.9) and specificity of 88.4% (95% CI 85.4–90.8), corresponding to an AUC of 0.93. This substantially exceeded human-read VIA, whose sensitivity was only 67.9% (95% CI 57.3–76.9) at a comparable specificity of 82.4% (AUC 0.75). HPV-DNA self-sampling was the most sensitive single test (95.2%, 95% CI 88.4–98.1) but the least specific (81.5%, 95% CI 78.1–84.5), yielding a PPV of only 43.7%. The ROC curves for all modalities are compared in Figure 2.

Combined testing strategies. As reported in Table 2, parallel testing (positive on HPV or DL-VIA) maximized sensitivity at 98.8% (95% CI 93.6–99.8) and NPV at 99.8%, but reduced specificity to 74.2%. Conversely, the sequential strategy (positive on HPV and DL-VIA) raised specificity to 95.7% (95% CI 93.7–97.1) and PPV to 75.5% (95% CI 66.1–83.0) while retaining a sensitivity of 88.1% (95% CI 79.5–93.4) and an AUC of 0.92; this trade-off is visualized in Figure 1.

Table 2. Diagnostic performance of index tests for detecting CIN2+ (95% CI).

Modality	Sensitivity (95% CI)	Specificity (95% CI)	PPV	NPV	LR+	LR–	AUC
Human expert VIA	67.9 (57.3–76.9)	82.4 (79.1–85.4)	36.8	94.5	3.9	0.39	0.75
HPV-DNA self-sample	95.2 (88.4–98.1)	81.5 (78.1–84.5)	43.7	99.1	5.1	0.06	0.88
DL-assisted VIA	91.7 (83.8–95.9)	88.4 (85.4–90.8)	54.2	98.6	7.9	0.09	0.93
Parallel (HPV or DL)	98.8 (93.6–99.8)	74.2 (70.4–77.7)	36.6	99.8	3.8	0.02	0.86
Sequential (HPV and DL)	88.1 (79.5–93.4)	95.7 (93.7–97.1)	75.5	98.2	20.5	0.12	0.92

Wilson 95% CI; PPV/NPV at 13.1% prevalence; LR+, positive likelihood ratio; LR–, negative likelihood ratio.

Paired comparative inference. On within-patient paired testing, DL-assisted VIA was significantly more sensitive than human-read VIA (McNemar p<0.001) at comparable specificity, and its AUC was significantly higher (0.93 vs 0.75; DeLong difference 0.18, 95% CI 0.11–0.25, p<0.001). The sensitivity difference between DL-assisted VIA and HPV self-sampling did not reach significance (91.7% vs 95.2%; McNemar p=0.219) and their AUCs were statistically indistinguishable (DeLong p=0.140), consistent with the overlapping intervals in Table 2; the two tests are therefore complementary rather than competing.

Likelihood ratios. Likelihood ratios reinforced these patterns, as summarized in Table 2: DL-assisted VIA yielded a positive likelihood ratio (LR+) of 7.9 and a negative likelihood ratio (LR-) of 0.09, the sequential strategy LR+ 20.5 and LR- 0.12, HPV self-sampling LR+ 5.1 and LR- 0.06, and human-read VIA LR+ 3.9 and LR- 0.39. The sequential strategy's LR+ above 20 indicates that a dual-positive result moves a woman convincingly into the high-probability range for CIN2+.

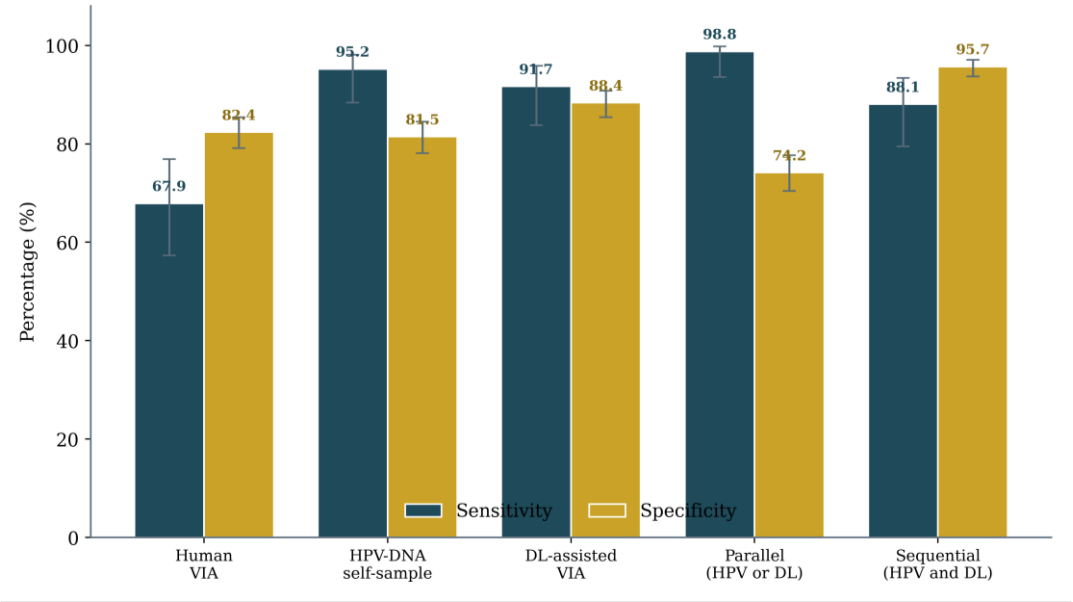

Figure 1. Sensitivity and specificity of each modality for CIN2+ detection (95% CI).

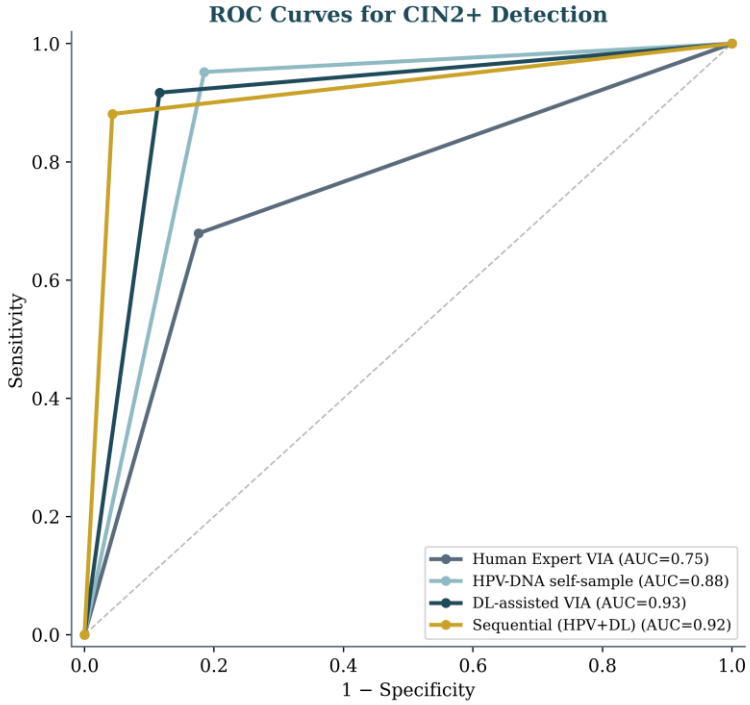

Figure 2. Receiver-operating-characteristic curves comparing modalities for CIN2+ detection (AUC annotated).

Agreement and multivariable determinants. Against the histopathologic reference, DL-assisted VIA showed substantial agreement (Cohen's κ 0.62, 95% CI 0.54–0.70), nearly double that of human-read VIA (κ 0.37, 95% CI 0.27–0.47), as reported in Table 3. In multivariable logistic regression, also shown in Table 3 and visualized in Figure 3, HPV positivity was the overwhelming independent determinant of CIN2+ (aOR 91.5, 95% CI 32.7–256.2, p<0.001), whereas geography (aOR 1.33, p=0.377), low education (aOR 0.67, p=0.224), high parity (aOR 0.90, p=0.732) and age (aOR 0.94 per SD, p=0.669) were not independently associated; the model explained half of the outcome variance (Nagelkerke R^2 0.50).

Conditional dependence and missed lesions. Within diseased women, HPV and DL-VIA results were positively but incompletely correlated (74 of 84 CIN2+ positive on both; 6 HPV-positive/DL-negative and 3 DL-positive/HPV-negative), so the sequential strategy's sensitivity loss arose chiefly from a small set of DL-missed but HPV-detected lesions. Of the ten CIN2+ lesions missed by the sequential strategy, seven were CIN2 and three were CIN3, with no invasive cancers missed; of the four lesions missed by HPV self-sampling alone, all were CIN2 (Table 3).

Genotype-stratified yield. In the pre-specified genotype analysis presented in Table 3, HPV16/18 positivity carried a markedly higher CIN2+ yield than other high-risk types: among HPV16/18-positive women the PPV for CIN2+ was 71.4% (50 of 70), versus 26.5% (30 of 113) for women positive only for other high-risk genotypes. This supports a genotype-informed refinement in which HPV16/18-positive women might proceed directly to colposcopy while DL-VIA triage is reserved for the larger, lower-risk pool.

Table 3. Multivariable determinants of CIN2+ and decision-analytic triage metrics.

Parameter	Estimate (95% CI)	p
HPV positive — adjusted OR (Firth)	91.5 (32.7–256.2)	<0.001
Region B (remote) — adjusted OR	1.33 (0.71–2.50)	0.377
Low education — adjusted OR	0.67 (0.35–1.28)	0.224
Parity ≥3 — adjusted OR	0.90 (0.51–1.61)	0.732
Age (+1 SD) — adjusted OR	0.94 (0.72–1.24)	0.669
Model fit — Nagelkerke R^2	0.50	—
DL-VIA agreement — Cohen's κ	0.62 (0.54–0.70)	<0.001
HPV16/18 PPV vs other hrHPV PPV	71.4% vs 26.5%	—
Missed CIN2+ (seq): CIN2/CIN3/invasive	7 / 3 / 0	—
Referral reduction (sequential)	−46.4% (183→98)	—
Unnecessary-referral reduction	−76.7% (103→24)	—
NNS to avert 1 unnecessary colposcopy	7.6 (5.9–10.6)	—

Triage efficiency. Among the 183 HPV-positive women, applying DL-assisted VIA as a second-line triage reduced colposcopy referrals from 183 to 98—a 46.4% reduction in total referrals and a 76.7% reduction in unnecessary (false-positive) referrals (from 103 to 24). As shown in Table 3, the number-needed-to-screen to avert one unnecessary colposcopy was 7.6 (95% CI 5.9–10.6), achieved at the cost of six additional missed CIN2+ relative to referring all HPV-positive women (sensitivity 88.1% vs 95.2%).

Subgroup performance. DL-assisted VIA performance was consistent across geography, as shown in Figure 3, with a sensitivity of 89.3% and specificity of 90.1% at Urban Center A and 92.9% and 87.5%, respectively, in Archipelago Region B, indicating that the model retained accuracy under the variable lighting and field conditions of remote clinics.

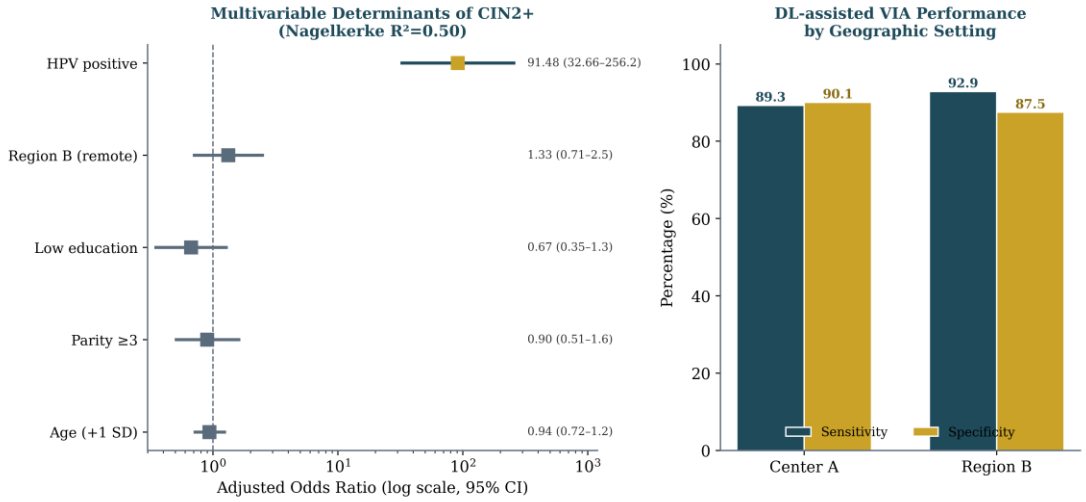

Figure 3. Multivariable determinants of CIN2+ (left) and DL-assisted VIA performance by geographic setting (right).

Discussion. This prospective, double-blind, fully verified diagnostic accuracy study shows that deep-learning-assisted digital VIA, delivered by telemedicine, is an accurate and operationally valuable triage for HPV-positive women in a remote archipelago setting. As summarized in Table 2, the DL model detected CIN2+ with a sensitivity of 91.7% and specificity of 88.4% (AUC 0.93), and when sequenced after HPV self-sampling it raised specificity to 95.7% and PPV to 75.5%, cutting colposcopy referrals among HPV-positive women by nearly half and unnecessary referrals by more than three quarters (Table 3). To our knowledge this is among the first prospective, fully verified evaluations to integrate HPV self-sampling with an on-island, telemedicine-delivered AI triage and to report not only diagnostic accuracy but the downstream colposcopy-conservation that determines real-world feasibility.

These findings extend a rapidly maturing evidence base. Work on automated visual evaluation has demonstrated that deep learning can equal or surpass expert image review, and a prospective Zambian study reported an AUC of 0.91 for smartphone-captured AVE.[10,12] Systematic reviews report high pooled accuracy for AI in cervical imaging, with CIN2+ sensitivities ranging from roughly 72% to 98%.[11,17] Our results are concordant with these estimates but add three features much of the prior literature lacks: prospective recruitment in a genuine screening population, complete verification of every participant by the histopathologic reference, and deployment on frontline smartphones rather than research-grade colposcopes.[10,12]

The contrast with human-read VIA is instructive. Conventional VIA is endorsed by WHO as a low-cost screen-and-treat tool but is constrained by subjectivity and inter-observer variability, and telehealth review of VIA images has been shown to improve concordance in low-resource programs.[16,19] In our cohort, human-read VIA missed nearly one third of high-grade lesions (sensitivity 67.9%, Table 2), whereas the DL model missed fewer than one in ten. The near-doubling of agreement with histopathology (κ 0.62 vs 0.37, Table 3)

quantifies the reproducibility gain that automation confers.[11,12]

Our HPV self-sampling results reaffirm a now-robust consensus. Sensitivity of 95.2% with a high NPV mirrors meta-analyses showing that self-collected testing is non-inferior to clinician collection and markedly improves participation among under-screened women.[6,13,14,15] The limitation of self-sampling is specificity: a PPV of only 43.7% (Table 2) means that referring every HPV-positive woman for colposcopy would direct most travel, cost, and specialist time toward women without high-grade disease—precisely the inefficiency that overwhelms remote-area capacity.[5,9]

Mechanistically, the complementary performance of the two tests reflects their distinct biological targets, as the conditional-dependence analysis confirmed: HPV testing detects the molecular cause of carcinogenesis and is therefore exquisitely sensitive but agnostic to whether infection has produced a lesion, whereas acetowhite imaging detects the phenotypic tissue change of established neoplasia and is therefore more specific for the actionable endpoint.[6,8] Sequencing a sensitive molecular screen with a specific morphological triage exploits this complementarity, which is why the sequential pathway in Table 2 approached the sensitivity of HPV alone while nearly eliminating false-positive referrals.

The safety of any specificity-gaining triage must be weighed explicitly. As reported in Table 3, the sequential pathway missed six more CIN2+ lesions than universal referral, but these were predominantly CIN2—the grade most likely to regress—and no invasive cancers were missed; the three missed CIN3 lesions nonetheless mandate that HPV-positive, DL-negative women enter a structured recall pathway rather than being discharged. We therefore do not present the sequential strategy as universally superior. Where colposcopy capacity exists, the maximally sensitive parallel strategy (Table 2) remains the safer default; the sequential strategy is best reserved for settings where the realistic alternative is no triage and high default rates. A genotype-informed hybrid—direct referral for HPV16/18 and DL-VIA triage for other high-risk types—offers a principled way to retain sensitivity for the highest-risk lesions, and our genotype data (Table 3) and external evidence suggest it is feasible.[8,20]

These pathways map naturally onto established risk-based frameworks. WHO's screen-triage-treat model designates HPV as the primary screen and a second test to triage screen-positives; our DL-assisted VIA occupies that triage position, and a High-grade output aligns with the threshold for colposcopy or, in a screen-and-treat configuration, same-visit thermal ablation—an appropriate option in Region B where excisional services are remote.[4,16,21] Positioning the modality within these algorithms clarifies exactly where in a national program it would slot and how triage-positive women would proceed to treatment.

An equity lens sharpens the public-health case. As shown in Table 1, the remote cohort carried a CIN2+ burden equal to the urban referral population yet faces far greater barriers—distance, cost, and competing demands—to completing a conventional referral. A decentralized pathway that returns an actionable triage decision during the same visit preferentially benefits exactly those women, flattening rather than reproducing the access gradient.[4,5]

The approach should generalize to other geographically fragmented populations facing the same structural barriers. Archipelagic and mountainous regions across South-East Asia and the Pacific share the defining features that make centralized colposcopy impractical, and the robustness of DL performance across our urban and remote subgroups (Figure 3) suggests the model's accuracy does not depend on idealized imaging conditions.[3,4] Because diagnostic-AI performance is sensitive to distribution shift, however, any transfer should be accompanied by local validation and, where necessary, recalibration under a governance framework.[10,11]

Realistic implementation will hinge on more than accuracy. Reliable connectivity for image transmission, robust data-governance and consent for cloud handling of intimate images, device calibration and maintenance, clear medico-legal responsibility for an AI-assisted decision, and the recurring cost of the telemedicine hub and specialist outreach all condition real-world impact; telehealth-VIA programs elsewhere have shown such systems can be operationalized.[19,22] Critically, a screening pathway prevents cancer only if triage-positive women actually receive treatment, so programs must guarantee an ablation or excision pathway—ideally same-visit thermal ablation on the island.[21]

From a health-systems perspective, the decision-analytic results are arguably more consequential than the headline accuracy figures. In an archipelago province, the binding constraint is not the sensitivity of any single test but the number of women who can physically reach a colposcopy service; a triage that conserves nearly half of colposcopy demand, as quantified in Table 3, therefore translates directly into more women screened per unit of scarce specialist capacity. Economic evaluations from Indonesia have shown that the affordability of HPV-based pathways is highly sensitive to downstream test and procedure volumes, so a triage that compresses unnecessary referrals by more than three quarters could materially improve the cost-effectiveness of a national program.[4,5] Modeling the full budgetary impact of the two-step pathway, incorporating local colposcopy and ablation costs, is an important next step that our diagnostic data can parameterize but cannot by themselves resolve.

Our findings also speak to the live debate over how HPV-positive women on self-collected samples should be managed. Current guidance in many programs requires self-sample-positive women to return for a clinician visit before any triage, a step that is precisely where remote populations are lost.[20] The present pathway collapses that gap by performing the triage—an AI-read image—at the same encounter, on the same island, within minutes. Combined with the genotype data in Table 3, which show that HPV16/18-positive women carry a far higher CIN2+ yield, the architecture supports a pragmatic, risk-tiered algorithm: immediate referral or same-visit treatment for HPV16/18, AI-VIA triage for other high-risk types, and structured recall for triage-negative women. Such tiering mirrors the direction of extended-genotyping and molecular-triage research and could recover much of the sensitivity that simple sequential triage sacrifices.[8,9]

Several practical caveats temper enthusiasm. The three missed CIN3 lesions, although few, are a reminder that no triage is free; programs adopting a sequential strategy must fund and audit the recall of HPV-positive, triage-negative women, or the efficiency gain will be paid for in delayed diagnoses. The dependence of the decision-analytic benefit on local colposcopy access means the 46.4% referral reduction observed here should be re-estimated wherever the pathway is deployed, since a setting with abundant colposcopy would value the sensitivity of the parallel strategy more highly than the specificity of the sequential one. And the durability of model performance over time, as cameras, software, and the screened population evolve, will require post-deployment monitoring rather than a one-time validation.[10,11]

This study has strengths and limitations. Its strengths are the prospective design, complete verification of all participants by an independent dual-pathologist reference standard (eliminating work-up bias), the head-to-head comparison of four modalities within one cohort, real-world deployment in remote clinics, and an upgraded inferential analysis with Wilson intervals, multivariable adjustment, agreement statistics, and explicit decision analysis. Limitations include

the single-province scope, which is ideal for the archipelago use-case but limits generalizability; the fixed, single-device DL model, whose performance may vary with hardware and operator technique and which requires multicenter external validation; reference-standard imperfection, since CIN2 has limited histologic reproducibility despite dual reading and p16 adjudication;[8,11] the powering of the study to estimate rather than to compare accuracy, so non-significant comparisons should not be read as equivalence; and the cross-sectional design, which cannot by itself establish cancer prevention—an outcome requiring linkage to treatment completion in future implementation research.[12]

Future work should proceed on three fronts. First, prospective external validation across multiple provinces, devices, and operator-training regimes is needed to confirm that the model's accuracy is robust to distribution shift before any recommendation for national scale-up. Second, an implementation-effectiveness study should follow HPV-positive women—especially the triage-negative group—through to documented treatment and, ideally, to incident-cancer outcomes, converting the present diagnostic evidence into the effectiveness evidence a program requires. Third, a formal budget-impact and cost-effectiveness analysis, parameterized with local colposcopy and ablation costs, should quantify the economic value of the conserved referral capacity demonstrated here. Together these studies would establish whether a decentralized, AI-assisted, telemedicine pathway can move from a promising single-province result to a dependable instrument of cervical-cancer elimination in archipelagic and other geographically isolated settings.

4. Conclusion

In a remote Indonesian archipelago population, deep-learning-assisted digital VIA delivered by telemedicine detected CIN2+ with high accuracy (sensitivity 91.7%, specificity 88.4%, AUC 0.93) and, when sequenced after HPV self-sampling, raised specificity to 95.7% while reducing colposcopy referrals by 46.4% and unnecessary referrals by 76.7% (number-needed-to-screen 7.6). This decentralized two-step pathway offers a scalable, resource-conserving strategy to advance cervical-cancer elimination among geographically isolated women, and warrants prospective multicenter validation and implementation research linking it to treatment completion.

Declarations

Acknowledgements. The authors thank the participating community health centers and the women who took part in this study.

Author Contributions. Conceptualization: TPS and ND. Methodology and data curation: TPS and ND. Image annotation and model development: TPS and ND. Histopathology: FH and GP. Formal analysis: TPS and ND. Writing—original draft: TPS. Writing—review and editing: TPS, ND, FH, and GP. Supervision: ND. All authors approved the final version.

Conflict of Interest. The authors declare no conflict of interest.

Funding. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethics and Consent. The study was approved by the Combined Medical & Health Research Ethics Committee (CMHC; approval No. CMHC/EC/2024/0461) and conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants. No identifiable patient material is included; the SJOG Patient Consent for Publication form is therefore not applicable.

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229–263. doi: 10.3322/caac.21834
2. 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
3. Ueda Y. Epidemiology of cervical cancer and HPV infection in Asia and Oceania. J Obstet Gynaecol Res. 2024;50(Suppl 1):31–41. doi: 10.1111/jog.15943
4. Ong SK, Abe SK, Thilagaratnam S, et al. Towards elimination of cervical cancer - human papillomavirus (HPV) vaccination and cervical cancer screening in Asian National Cancer Centers Alliance (ANCCA) member countries. Lancet Reg Health West Pac. 2023;39:100860. doi: 10.1016/j.lanwpc.2023.100860
5. Hafidz F, Icanervilia AV, Rizal MF, et al. Economic evaluation of cervical cancer screening by HPV DNA, VIA, and Pap smear methods in Indonesia. Asian Pac J Cancer Prev. 2024;25(9):3015–3022. doi: 10.31557/APJCP.2024.25.9.3015
6. Arbyn M, Simon M, de Sanjose S, et al. Accuracy and effectiveness of HPV mRNA testing in cervical cancer screening: a systematic review and meta-analysis. Lancet Oncol. 2022;23(7):950–960. doi: 10.1016/S1470–2045(22)00294–7
7. Magdi R, Elshafeey F, Elshebiny M, et al. A systematic review and meta-analysis of diagnostic accuracy of HPV tests for the screening of cervical cancer in low-resource settings. Int J Gynaecol Obstet. 2021;152(1):12–18. doi: 10.1002/ijgo.13455
8. Stoler MH, Parvu V, Yanson K, et al. Risk stratification of HPV-positive results using extended genotyping and cytology: data from the baseline phase of the Onclarity trial. Gynecol Oncol. 2023;174:68–75. doi: 10.1016/j.ygyno.2023.04.022
9. Zhang L, Zhao X, Hu S, et al. Triage performance and predictive value of the human gene methylation panel among women positive on self-collected HPV test: results from a prospective cohort study. Int J Cancer. 2022;151(6):878–887. doi: 10.1002/ijc.34041
10. 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. Int J Cancer. 2022;150(5):741–752. doi: 10.1002/ijc.33879
11. Allahqoli L, Lagana AS, Mazidimoradi A, et al. Diagnosis of cervical cancer and pre-cancerous lesions by artificial intelligence: a systematic review. Diagnostics (Basel). 2022;12(11):2771. doi: 10.3390/diagnostics12112771
12. Hu L, Mwanahamuntu MH, Sahasrabuddhe VV, et al. Internal validation of automated visual evaluation (AVE) on smartphone images for cervical cancer screening in a prospective study in Zambia. Cancer Med. 2024;13(11):e7355. doi: 10.1002/cam4.7355
13. Li DM, Liu QY, Xue SL, et al. Accuracy analysis of cervical cancer screening using urine and vaginal self-sampling versus clinician-collected samples: a systematic review and meta-analysis. Int J Gynaecol Obstet. 2025;171(1):207–216. doi: 10.1002/ijgo.70207
14. Costa S, Verberckmoes B, Castle PE, et al. Offering HPV self-sampling kits: an updated meta-analysis of the effectiveness of strategies to increase participation in cervical cancer screening. Br J Cancer. 2022;128(5):805–813. doi: 10.1038/s41416–022–02094-w
15. Di Gennaro G, Licata F, Trovato A, et al. Does self-sampling for human papilloma virus testing have the potential to increase cervical cancer screening? An updated meta-analysis of observational studies and randomized clinical trials. Front Public Health. 2022;10:1003461. doi: 10.3389/fpubh.2022.1003461
16. Poli UR, Gudlavalleti AG, Bharadwaj J, et al. Development and clinical validation of visual inspection with acetic acid application-artificial intelligence tool using cervical images in screen-and-treat visual screening for cervical cancer in South India: a pilot study. JCO Glob Oncol. 2024;10:e2400146. doi: 10.1200/GO.24.00146
17. Fu L, Xia W, Shi W, et al. Deep learning based cervical screening by the cross-modal integration of colposcopy, cytology, and HPV test. Int J Med Inform. 2022;159:104675. doi: 10.1016/j.ijmedinf.2021.104675
18. Liu J, Liang T, Peng Y, et al. Segmentation of acetowhite region in uterine cervical image based on deep learning. Technol Health Care. 2022;30(2):469–482. doi: 10.3233/THC-212890
19. Ogunsola OO, Olawepo JO, Ajayi O, et al. AVIVA: a telehealth tool to improve cervical cancer screening in resource-constrained settings. BMJ Glob Health. 2023;8(7):e012311. doi: 10.1136/bmjgh-2023–012311
20. Lei J, Cuschieri K, Patel H, et al. Human papillomavirus genotype and cycle threshold value from self-samples and risk of high-grade cervical lesions: a post hoc analysis of a modified stepped-wedge implementation feasibility trial. PLoS Med. 2024;21(12):e1004494. doi: 10.1371/journal.pmed.1004494
21. Denny L, Saidu R, Boa R, et al. Point-of-care testing with Xpert HPV for single-visit, screen-and-treat for cervical cancer prevention: a demonstration study in South Africa. Sci Rep. 2023;13(1):16182. doi: 10.1038/s41598–023–43467–2
22. Okunade KS, Adejimi AA, Adekanye TV, et al. Impact of mobile health technologies on human papillomavirus vaccination uptake among mothers of unvaccinated girls aged 9–14 years in Lagos, Nigeria (mHealth-HPVac): study protocol of a randomised controlled trial. BMC Cancer. 2024;24(1):751. doi: 10.1186/s12885–024–12538–6