cancers

MDPI

Article

Risk-Based Triage of Women with Persistent HR-HPV Positivity at 12-Month Recall in Cervical Cancer Screening: A Multicenter Retrospective Evaluation of Cytology and Partial HPV Genotyping

Giampaolo Pompeo [1], Alessandra Mongia [1], Irene Paganini [1], Giuseppe Gorini [1], Donella Puliti [1], Stephanie Pacella [1], Beatrice Fuzzi [1], Annarosa Del Mistro [2], Helena Frayle [2], Giuliana Sartori [3], Chiara Castagnetti [3], Cinzia Campari [3], Andrea Amico [4], Carolina Buriani [4], Daniela Gustinucci [5], Elena Cesarini [5], Simona Dei [1], Roberto Gusinu [1], Cristina Sani [1,*], Silvia Gori [2] and ISPRO Cervical Cancer Screening Working Group [†]

[1] Institute for Cancer Research, Prevention and Clinical Network (ISPRO), 50139 Florence, Italy; g.pompeo@ispro.toscana.it (G.P.); g.gorini@ispro.toscana.it (G.G.); donella.puliti@meyer.it (D.P.)
[2] Veneto Institute of Oncology (IOV-IRCCS), 35128 Padua, Italy; annarosa.delmistro@iov.veneto.it (A.D.M.); silvia.gori@iov.veneto.it (S.G.)
[3] Azienda USL-IRCCS di Reggio Emilia, 42122 Reggio Emilia, Italy; giuliana.sartori@ausl.re.it (G.S.)
[4] University Hospital of Ferrara, 44124 Ferrara, Italy
[5] USL Umbria 1, 06127 Perugia, Italy
* Correspondence: c.sani@ispro.toscana.it
† ISPRO Cervical Cancer Screening Working Group: the names of all members are listed in the Acknowledgments.

Simple Summary

Under the current Italian HPV-based cervical screening protocol, women with results that were still positive for the HR-HPV test after the 12-month recall are referred directly to colposcopy, irrespective of cytology or partial HPV genotyping results. This approach probably maximizes precancerous lesion detection, but it results in a high number of colposcopies, most of which do not identify CIN2+. In this retrospective study, based on data from over 270,000 women screened in four Italian regions, four alternative triage strategies based on cytology and HR-HPV partial genotyping were retrospectively simulated and compared with the current Italian 12-month recall protocol. We found that the retrospective application of one of these candidate alternative strategies would have substantially reduced the immediate colposcopy referrals by more than half while still detecting most high-grade precancerous lesions (relative sensitivity for CIN3+: 85%); the seven invasive cancers reported separately in three of the four centers would also have been referred immediately, although this finding cannot be taken as evidence of prospective safety. These findings, derived from a retrospective simulation in the study cohort, suggest that this approach may be a more efficient and less invasive way to manage the 12-month recall; prospective validation is needed.

Check for updates

Academic Editor: Xavier Sastre-Garau

Received: 31 August 2026
Revised: 5 November 2026
Accepted: 6 November 2026
Published: 12 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Abstract

Background: Under the current Italian HPV-based cervical screening protocol, women with persistent high-risk human papillomavirus (HR-HPV) positivity at the 12-month recall are referred to colposcopy regardless of cytology results. Although this strategy maximizes lesion detection, it may result in a high number of colposcopies, most of which do not identify CIN2+. We evaluated whether partial HPV genotyping combined with cytology can improve risk stratification in women with persistent HPV positivity at the 12-month recall. Methods: A multicenter retrospective analysis of cervical screening data collected

from four Italian screening programs (Florence, Padua, Umbria and Emilia Romagna) was performed. Among the 270,714 women screened between 2015 and 2019 (2016–2018 in Florence), 4589 women with persistent HR-HPV positivity at 12-month recall after baseline HPV-positive/NILM (negative for intraepithelial lesion or malignancy) results were included. Four alternative triage strategies based on cytology and HPV 16/18 genotyping were retrospectively simulated and compared with the current protocol. Outcomes included CIN2 (cervical intraepithelial neoplasia grade 2) or worse (CIN2+) detection, relative sensitivity for CIN2+ and CIN3 (cervical intraepithelial neoplasia grade 3) or worse (CIN3+), positive predictive value (PPV) and reduction in colposcopy referrals. Results: Under the current protocol, 4589 colposcopies identified 327 CIN2+ lesions, including 113 CIN3+ lesions, yielding a PPV for CIN2+ of 7.1%. Overall, 36.7% of CIN2+ lesions were associated with NILM cytology at recall. HPV 16/18 genotyping identified 43.3% of these cytologically occult lesions. In this retrospective simulation, among the evaluated strategies, referral of HPV 16/18-positive women regardless of cytology and referral of women with low-grade or worse cytology regardless of genotype would have achieved the most favorable balance between sensitivity and efficiency. This approach would have reduced colposcopy referrals by 57.7% (95% CI, 56.3–59.2), nearly doubled the PPV for CIN2+ to 13.4% (95% CI, 11.9–14.9) and maintained relative sensitivities of 79.2% (95% CI, 74.5–83.3) for CIN2+ and 85.0% (95% CI, 77.2–90.4) for CIN3+. Performance varied across centers, with imprecise center-specific estimates (relative CIN3+ sensitivity of Strategy 2 from 62.5% in Emilia Romagna to 91.1% in Umbria), so generalizability should be assessed cautiously. Conclusions: Combining partial HPV genotyping with cytology would have substantially improved risk stratification in women with persistent HR-HPV positivity at 12-month recall. In this retrospective simulation, a risk-based approach would have substantially reduced immediate colposcopies while maintaining a relative sensitivity for CIN3+ of 85.0%. These findings are hypothesis-generating and support prospective evaluation of integrating molecular and cytological information into 12-month recall management.

Keywords: cervical cancer screening; HPV-DNA test; 12-month recall; triage strategies; partial HPV genotyping; cytology; risk stratification

1. Background

1.1. HPV-Based Screening and the Challenge of Persistent Infections

The pooled analysis of four European randomized clinical trials or RCTs (NTCC, POBASCAM, ARTISTIC and SWEDESCREEN) [1–4], published in 2014 [5], established that HPV-based primary screening reduces the incidence of invasive cervical cancer by approximately 70% over a 5.5-year follow-up period, marking the transition from Pap-based to HPV-based cervical cancer screening. The World Health Organization (WHO) endorsed HPV testing as the preferred primary screening method in 2021 [6] and the European Cancer Beating Plan (2021–2030) has set a target of 90% of eligible women undergoing HPV testing by 2025 [7]. More recently, the WHO issued a guideline on the use of HPV DNA genotyping in cervical screening, which provides limited and extended genotyping options for risk-based management of HPV-positive women, adaptable to the capacity of national programs [8].

A key operational challenge of HPV-based screening concerns the management of women who are HR-HPV positive with negative triage cytology (NILM, negative for intraepithelial lesion or malignancy) at baseline and remain HR-HPV positive at the 12-month recall. Persistent HR-HPV positivity could be associated with an elevated risk of underlying

high-grade disease [9]; however, a significant proportion of these infections appear to regress spontaneously without evolving into Cervical Intraepithelial Neoplasia (CIN) or cervical cancer. Therefore, HR-HPV persistence alone may lack the specificity [10] needed to efficiently guide referral decisions for further testing.

1.2. Current Italian Management of Women with Persistent HR-HPV Positivity

Pending implementation of recent multi-societal evidence-based guidelines [11], Italian cervical cancer screening is based on primary HPV testing for women aged 30–64 every five years, with cytology triage of HR-HPV-positive women at baseline. HR-HPV-positive/NILM women are recalled after 12 months to repeat the HPV test and, if HR-HPV infection persists, are referred directly to colposcopy regardless of cytology. This "all-referral" strategy maximizes sensitivity by bypassing potential false-negative cytology, but it has a low positive predictive value (PPV) for CIN2 or worse (CIN2+) and generates a high colposcopic workload with a low yield of CIN2+. In fact, in our previous publication [12] regarding the Tuscany centralized screening program, 2526 colposcopies out of approximately 300,000 screened women were performed at the 12-month recall, detecting 278 CIN2+ lesions (PPV = 11%) and suggesting that, even in high-quality settings, approximately nine out of 10 colposcopies performed at the 12-month recall may not identify a high-grade lesion.

Two mechanisms appear to support the introduction of a triage step at the 12-month recall. First, a concurrent cytological abnormality seems to be associated with a markedly higher risk of underlying disease: UK NHS pilot data [13] on 7315 women with persistent HR-HPV infection reported a CIN2+ PPV of 36% among cytology-positive women versus 13% among cytology-negative women, consistent with the Tuscan data [12]. Second, HPV 16 and HPV 18 carry a substantially higher cumulative CIN3+ risk than other oncogenic genotypes, with a 5-year CIN3+ risk at approximately 26–30% for HPV16, 14–19% for HPV18 and only 3–7% for the remaining genotypes individually [14,15]. Partial HPV genotyping (distinguishing HPV 16 and HPV 18 from other HR-HPV types) has been shown to provide a more precise risk stratification than cytology alone: HPV 16/18-positive, cytology-negative women (NILM) showing a 5-year CIN3+ risk comparable to that of cytology-positive (ASC-US, Atypical Squamous Cells of Undetermined Significance or more severe, ASC-US+) women with non-16/18 genotypes [15–17].

Extended genotyping, which identifies individual carcinogenic HPV types or groups of types, may further refine risk stratification [18,19]. The 2026 WHO guideline on HPV DNA genotyping [8] stratifies the 12 carcinogenic types into four risk groups (group 1a: HPV 16; 1b: HPV 18 and 45; 1c: HPV 31, 33, 35, 52 and 58; 1d: HPV 39, 51, 56 and 59) and supports extended genotyping for risk-based management. This technology of genotyping is currently available through several commercial platforms, but it is not yet available in all Italian screening centers. Italian cohort data [20] identified HPV 16 as the main determinant of CIN3+ risk both at baseline (7.9%) and at the 12-month recall (4.2%), higher than HPV 18 (2% and 2.6%, respectively) or other HR-HPV types (1.3% and 0.8%), with a cumulative CIN3+ detection in HPV 16-positive women (9.8%) that substantially exceeded that of non-16/18 genotypes (1.8%).

International recommendations for managing persistent HR-HPV infection vary considerably, ranging from cytology-based triage to genotype- and biomarker-based approaches [21–23]. The p16/Ki-67 dual stain (DS), which indicates transforming or oncogenic rather than productive HR-HPV infection, showed a higher cumulative sensitivity for CIN3+ than cytology in the IMPACT trial [24] (89.2% vs. 71.3%) but a lower specificity (54.5% vs. 64.6%); its lower residual CIN3+ risk among negative women compared with cytology (3.6% vs. 7.5%) supports its potential role as a safer "rule-out" test.

These biomarkers or methodologies (partial or extended HPV genotyping, cytology and DS), now constitute the 2025 Italian multi-societal guidelines [11], which apply the GRADE methodology to the use of biomarkers in cervical screening and define explicit CIN3+ risk thresholds to guide immediate colposcopy referral, 12-month recall or return to routine screening.

1.3. Study Objectives

The study on the management of women with persistent HR-HPV positivity at the 12-month recall was conducted in 2022 on behalf of the Italian Group for Cervical Cancer Screening (GISCi), before the publication of the 2025 Italian multi-societal guidelines [11]. The study design and the alternative strategies were therefore defined independently of the guidelines, which are discussed here only to contextualize our findings. Cytology and HR-HPV partial genotyping results at the 12-month recall were already routinely available at each screening center and recorded in Information Technology (IT) systems, but are not currently used in clinical practice. This retrospective study evaluates the potential impact of using this information without additional testing, costs or infrastructure.

This retrospective evaluation builds upon previous evidence showing that partial HPV 16/18 genotyping improves risk stratification at the baseline [20], extending the same screening population and dataset to the 12-month recall protocol, with data from Florence (2016–2018) added.

The aim of this study was to evaluate whether partial HPV genotyping, combined with cytology, could improve risk stratification in women with persistent HR-HPV positivity at 12-month recall and reduce immediate colposcopy referrals, without substantially compromising the detection of CIN2+ or CIN3+ lesions.

Specifically, we characterized the performance of the current protocol, quantified the proportion of CIN2+ lesions with NILM cytology and the contribution of HPV 16/18 genotyping to their detection, evaluated four alternative triage strategies (PPV for CIN2+, relative sensitivity for CIN2+ and CIN3+, and colposcopy reduction), and identified the strategy offering the most favorable trade-off between colposcopy workload and sensitivity.

2. Materials and Methods

2.1. Study Design and Settings

This was a multicenter retrospective analysis conducted on behalf of GISCi and its outcomes were presented at the 2022 GISCi National Conference. The study involved the retrospective application of four alternative 12-month recall protocols to an existing dataset of women already screened, using data routinely collected by the participating centers. No additional clinical procedures were performed.

Four Italian cervical cancer screening centers provided the data:

- Florence (FI): 40,496 women screened in the period 2016–2018;
- Padua (PD), Emilia Romagna (ER) and Umbria (UM): 230,218 women screened in the period 2015–2019.

The total screened population considered was 270,714 women (Figure 1). The analytic cohort included 4589 women who were HR-HPV positive with negative triage cytology (NILM) at baseline and who subsequently underwent the 12-month recall, with available data on HPV partial genotype, cytology, colposcopy and histological outcomes; their characteristics are reported in the Section 3.

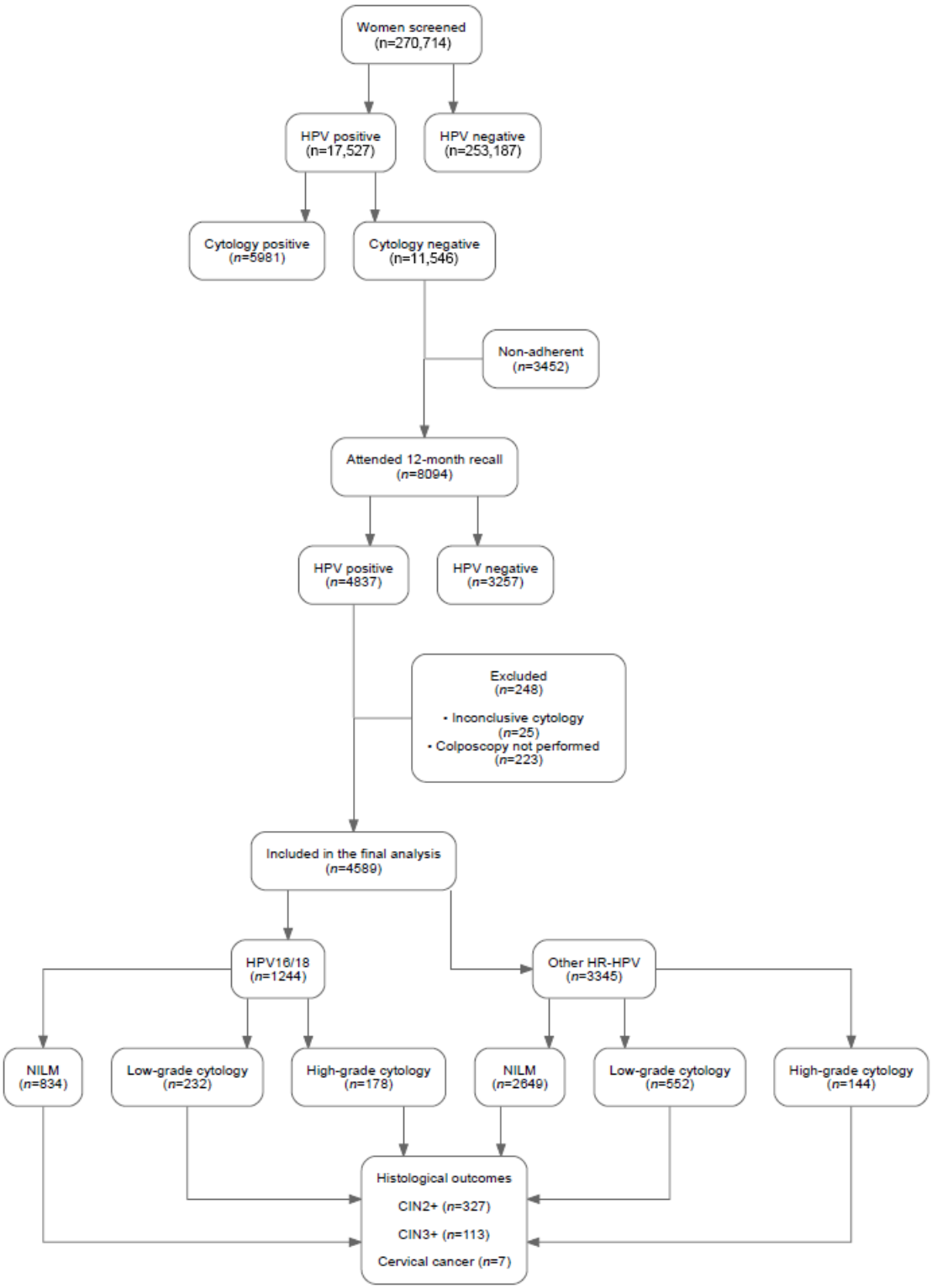


Figure 1. STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) flow diagram of the study population. Among the 270,714 screened women, 4837 women resulted with persistent HR-HPV infection at the 12-month recall and 248 were excluded (223 non-adherent to colposcopy, 25 with inconclusive cytology); 4589 women were included in the final analysis, stratified by HPV genotype and cytology result of the 12-month recall.

2.2. Laboratory Tests

All centers used a HR-HPV assay with partial genotyping capability (COBAS 4800 HPV assay, Roche Diagnostics, Basel, Switzerland), classifying HR-HPV positive results into

three categories: HPV 16-positive, HPV 18-positive, and Other HR-HPV positive (HPV 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68). Liquid-based cytology (ThinPrep, Hologic, Inc. Marlborough, MA, USA) was performed on all HR-HPV persistent women and results were reported according to the Bethesda System [25] for reporting cytology (TBS 2014); however, the cytology result at the 12-month recall was not actively used in the current protocol and all HR-HPV-persistent women were referred to colposcopy regardless of the 12-month recall cytology result. For the purposes of this analysis, cytology results were retrospectively reclassified:

- Negative (NILM: negative for intraepithelial lesion or malignancy);
- Low-grade abnormal (LG: ASC-US, LSIL, low-grade squamous intraepithelial lesion);
- High-grade abnormal (HG: ASC-H: atypical squamous cells—cannot exclude a high-grade squamous intraepithelial lesion; HSIL: high-grade squamous intraepithelial lesion; AGC: atypical glandular cells; squamous cell carcinoma; adenocarcinoma).

2.3. Alternative Triage Strategies

The current reference strategy, used as the comparator, refers all HR-HPV persistent women to colposcopy, regardless of the 12-month recall HPV genotype or cytology result. The four evaluated alternative strategies were:

- Strategy 1: Colposcopy referral if HR-HPV+ AND cytology ≥ LG;
- Strategy 2: Colposcopy referral if HPV 16/18+ (regardless of cytology) OR cytology ≥ LG regardless of genotype (HPV Other+ and LG+);
- Strategy 3: Colposcopy referral if HR-HPV+ AND cytology ≥ HG;
- Strategy 4: Colposcopy referral if HPV 16/18+ OR cytology ≥ HG.

2.4. Outcome Ascertainment and Reference Standard

Of the 4837 women with persistent HR-HPV, 4589 (94.9%) underwent colposcopy and were managed according to the current protocol. Colposcopies were performed according to the 2011 colposcopic terminology [26] of the International Federation for Cervical Pathology and Colposcopy (IFCPC). Visibility of the squamocolumnar junction, type of transformation zone and grading of colposcopic findings were recorded (where available). Biopsy was performed in case of suspected lesions or high-grade cytology, and endocervical curettage in case of glandular atypia. Histological specimens were read locally by the pathologists of each center and the related diagnoses were reported using the cervical intraepithelial neoplasia (CIN) grading system (CIN1, CIN2, CIN3) and invasive cancer. CIN1 corresponds to low-grade and CIN2/CIN3 to high-grade squamous intraepithelial lesion in the two-tier terminology (LAST; WHO Classification of Female Genital Tumors, 5th edition). CIN3 included: CIN3 and adenocarcinoma in situ; invasive cancer included: squamous cell carcinoma and adenocarcinoma, including microinvasive. CIN2+ includes CIN2, CIN3 and invasive cancer; CIN3+ includes CIN3 and invasive cancer. The number and type (targeted vs. random) of biopsies per woman were not considered for the purposes of this study. When multiple specimens were available, the most severe histological diagnosis was used as the final outcome. The final histological outcome was the diagnosis obtained from the colposcopy performed after the positive HPV test at the 12-month recall (index colposcopy). No additional verification period was applied. In women with lesions of different grades or mixed squamous and glandular lesions, the most severe diagnosis was used for the classification. Women with normal colposcopy and no biopsy, or with negative histology or CIN1, were classified as negative/no CIN2+ detected. The all-referral strategy was used as the comparator for the simulated strategies and is not considered an infallible reference standard.

2.5. Statistical Analysis

Descriptive statistics were used to summarize the demographic and clinical characteristics of the study population. Categorical variables are presented as frequencies and percentages and were compared using Pearson's chi-square test or Fisher's exact test, as appropriate. Age was categorized into clinically relevant groups (30–39, 40–49 and 50–64 years) and differences between HPV genotype groups were assessed using the chi-square test.

Absolute risks of histologically confirmed CIN2+, CIN3+ and invasive cervical cancer were calculated for each combination of HPV genotype (HPV 16/18 versus other HR-HPV types) and cytological result at the 12-month recall. Risk estimates are presented as percentages with exact 95% confidence intervals (Clopper–Pearson method).

We calculated exact two-sided 95% binomial confidence intervals (Clopper–Pearson) for absolute CIN2+, CIN3+ and invasive cancer risks by genotype–cytology stratum. Relative sensitivities under simulated triage strategies were expressed as the proportion of lesions observed under the current all-referral pathway that would trigger immediate colposcopy; for these, as well as for the PPV and the proportion of colposcopies avoided or referred, 95% Wilson confidence intervals were calculated. Grouped-binomial logistic regression models estimated associations of HPV16/18 positivity, cytology (NILM, low-grade, high-grade) and screening center with CIN2+ and CIN3+. Genotype-by-cytology interactions were examined using likelihood-ratio tests. Between-center variation in Strategy 2 relative sensitivity was evaluated using an exact conditional Monte Carlo procedure because of sparse event counts. Center-level risks are reported in Supplementary Tables S1–S4, center-level performance of the alternative strategies in Tables S5–S8, center-specific relative sensitivity of Strategy 2 in Table S9, and between-center heterogeneity tests in Table S10. Statistical notes for the Supplementary Tables are provided in the Supplementary Materials. Age could not be included in the models because joint age-by-genotype-by-cytology-by-outcome aggregates were unavailable.

The diagnostic performance of the current management protocol was compared with four alternative risk-based management strategies. For each strategy, the number and proportion of women referred to colposcopy, the reduction in colposcopy referrals compared to the current protocol, the number and proportion of detected CIN2+, CIN3+ and cervical cancer cases as well as the number of CIN3+ lesions and invasive cancers not immediately detected were calculated. No prespecified criterion for an acceptable proportion of deferred CIN3+ sensitivity was defined; strategies were compared descriptively.

To facilitate interpretation of the clinical impact of each strategy, graphical summaries were generated, including a heatmap illustrating the absolute risk of CIN2+, CIN3+ and cervical cancer, according to HPV genotype and cytological findings, a trade-off plot displaying the relationship between colposcopy reduction and sensitivity for CIN2+ and CIN3+ detection.

To translate the observed risks into an explicit management framework, we report the absolute CIN3+ risk with exact 95% CIs for each genotype–cytology stratum alongside the corresponding management under Strategy 2, compared descriptively with the 2.5% CIN3+ risk threshold of the Italian guidelines [11]. Because invasive cancers were not reported separately from CIN3+ in Emilia Romagna, the absolute risk of invasive cancer was calculated separately from the other endpoints, using as denominator only the 3759 women from the three centers reporting cancers separately (Florence, Padua and Umbria).

All statistical analyses were performed using R software (version 4.6.0; R Foundation for Statistical Computing, Vienna, Austria). Data visualization was generated using the ggplot2, patchwork and ComplexHeatmap packages. All statistical tests were two-sided and a p-value < 0.05 was considered statistically significant.

2.6. Outcome Measures

- Detection rate (DR) for CIN2+ per 1000 screened women (‰);
- Positive predictive value (PPV) for CIN2+ (CIN2+ lesions/total colposcopies performed);
- Relative sensitivity for CIN2+ and CIN3+ versus the reference strategy;
- Number and percentage of colposcopy referrals avoided/deferred versus the current protocol;
- Proportion of CIN2+ lesions with NILM cytology at the recall;
- Contribution of HPV 16/18 genotyping to the recovery of CIN2+ lesions with NILM cytology;
- Absolute risk of CIN2+, CIN3+ and invasive cancer by genotype and cytology;
- Number of CIN2+, CIN3+ and invasive cancers deferred under each strategy.

3. Results

3.1. Overall Performance of the Current Protocol

Overall, 4837 women resulted HR-HPV persistent at the 12-month recall (Figure 1). Of these, 25 (0.5%) were excluded because of inconclusive cytology and 223 (4.6%) did not attend colposcopy. The analytic cohort therefore included 4589 women (94.9% of those eligible), all of whom underwent colposcopy. Table 1 shows the characteristics of the 4589 women according to HPV genotype group. HPV 16/18-positive women were younger than those with other HR-HPV types ($p < 0.001$) and had a higher proportion of high-grade cytology ($p < 0.001$) and more severe histological outcomes ($p = 0.003$).

Table 1. Demographic and clinical characteristics of the study population according to HPV genotype group.

Characteristics	HPV16/18 *n* (%)	Other HR-HPV *n* (%)	*p*-Value
Women	1244 (27.1)	3345 (72.9)	
Age 30–39	519 (30.3)	1194 (69.7)	<0.001
Age 40–49	501 (26.1)	1422 (73.9)	
Age 50–64	224 (23.5)	729 (76.5)	
NILM	834 (23.9)	2649 (76.1)	<0.001
Low Grade (LG)	232 (29.6)	552 (70.4)	
High Grade (HG)	178 (55.3)	144 (44.7)	
CIN2	87 (40.7)	127 (59.3)	0.003
CIN3	63 (59.4)	43 (40.6)	
Cancer	5 (71.4)	2 (28.6)	

p-values refer to comparisons between the HPV 16/18 and Other HR-HPV groups: Pearson's chi-square test for age and cytology; Fisher's exact test for histological outcome.

Among the 4589 women who underwent colposcopy under the current protocol, 327 CIN2+ lesions were detected (PPV: 7.1%), of which 113 were CIN3+ (PPV: 2.5%), with an overall DR of 1.2‰ (327/270,714). Only seven invasive cervical cancers were identified (DR: 0.032‰; 7/219,223 women screened in the three centers reporting cancers separately). Table 2 summarizes the overall performance by 12-month recall HR-HPV partial genotyping and cytology category result. The highest values for CIN2+ and CIN3+ risk were observed in HPV 16/18-positive women with HG cytology (CIN2+ PPV: 40.5%; CIN3+ PPV: 23.6%), while the lowest values were observed in other HR-HPV-positive women with NILM cytology (CIN2+ PPV: 2.6%; CIN3+ PPV: 0.6%). Substantial variability was observed among centers, with DR for CIN2+ ranging from 0.5‰ (Padua) to 2.8‰

(Umbria) and PPV for CIN2+ from 4.9% (Padua) to 10.6% (Umbria); detailed center-level results are reported in Supplementary Tables S1–S4.

Table 2. Overall 12-month recall results with detailed information on HR-HPV partial genotyping and cytology results (NILM, LG and HG).

Cancer (%) *	Cancer (*n*) *	PPV CIN3+ (%)	CIN3+ (*n*)	PPV CIN2+ (%)	CIN2+ (*n*)	*n*	Triage Cytology	HR-HPV Type
0.00% (0.00–0.52%)	0	2.40% (1.47–3.68%)	20	6.24% (4.69–8.10%)	52	834	Negative (NILM)	16/18
0.00% (0.00–1.77%)	0	2.59% (0.95–5.54%)	6	13.36% (9.26–18.43%)	31	232	Low-grade (LG)	16/18
5.15% (1.69–11.62%)	5	23.60% (17.57–30.53%)	42	40.45% (33.17–48.05%)	72	178	High-grade (HG)	16/18
0.00% (0.00–0.17%)	0	0.64% (0.37–1.03%)	17	2.57% (2.00–3.24%)	68	2649	Negative (NILM)	Non 16/18
0.21% (0.01–1.15%)	1	1.45% (0.63–2.84%)	8	9.24% (6.96–11.97%)	51	552	Low-grade (LG)	Non 16/18
0.79% (0.02–4.31%)	1	13.89% (8.69–20.63%)	20	36.81% (28.93–45.24%)	53	144	High-grade (HG)	Non 16/18
0.19% (0.07–0.38%)	7	2.46% (2.03–2.95%)	113	7.13% (6.40–7.91%)	327	4589		Total

* Invasive cancers were reported separately in Florence, Padua and Umbria; in Emilia Romagna they were aggregated with CIN3+ and could not be reported separately. Cancer percentages are therefore calculated on the 3759 women from the three centers reporting cancers separately.

Figure 2 graphically presents the detailed risk heatmap for CIN2+, CIN3+ and cervical cancer across all six HPV genotype–cytology strata at the 12-month recall (HPV 16/18 or other HR-HPV–NILM or LG or HG cytology result); exact 95% CIs are reported in Table 2. The color gradient reveals a gradual and consistent risk escalation along both axes: risk increases progressively from NILM to HG cytology within each genotype group and from other HR-HPV to HPV 16/18 within each cytology category. The highest risk combination is HPV 16/18 with HG cytology, reaching a CIN2+ PPV of 40.5%, a CIN3+ PPV of 23.6% and a cancer risk of 5.2%. Conversely, the lowest-risk stratum, other HR-HPV with NILM cytology, shows a CIN2+ risk of only 2.6%, a CIN3+ risk of 0.6% and zero cancer cases, identifying this group as the primary candidate for follow-up deferral rather than colposcopy. Notably, HPV 16/18 types with NILM cytology (CIN3+ PPV 2.4%) carries a higher risk than other HR-HPV/LG cytology (CIN3+ PPV 1.4%). The absolute risk of invasive cancer in Table 2 and Figure 2 was calculated on the 3759 women from the three centers reporting cancers separately (Florence, Padua, Umbria), since in Emilia Romagna cancers were aggregated with CIN3+.

3.2. CIN2+ Lesions with NILM Cytology at 12-Month Recall

A finding with direct implications for triage design is the proportion of CIN2+ lesions detected under the current protocol that were associated with NILM cytology, the so-called "cytologically silent" lesions. As shown in Table 3, 120 out of 327 CIN2+ lesions (36.7%; range across centers: 24.7–45.7%) had NILM cytology at the 12-month recall, underscoring a limitation of cytology as a stand-alone triage tool.

Our data showed that HPV16/18 genotyping could recover 43.3% of CIN2+ lesions (52/120) undetected by cytology triage alone. No cytology-based triage strategy, however sensitive, appears able to detect almost all CIN2+ lesions on its own.

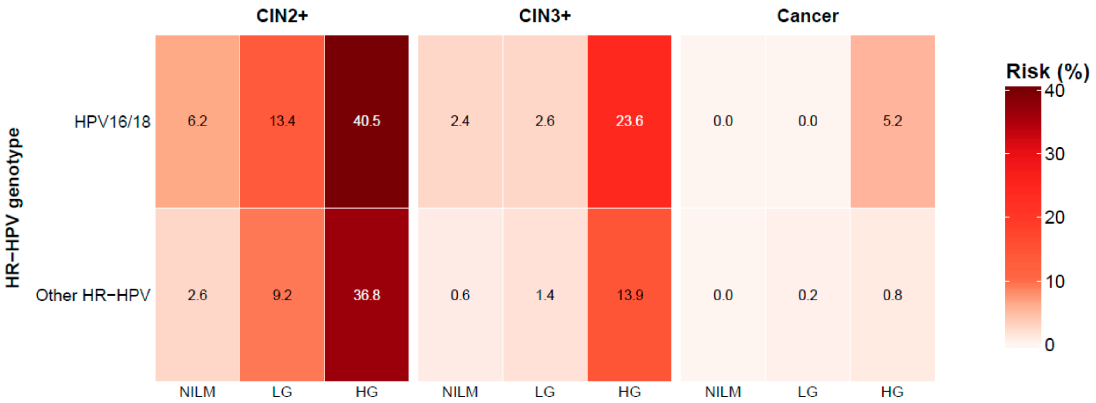


Figure 2. Risk heatmap for CIN2+, CIN3+ and cervical cancer by HPV genotype and cytology at the 12-month recall. Each cell displays the absolute risk (%) for the corresponding histological endpoint within each genotype/cytology stratum. Color intensity of the box encodes absolute risk on a continuous scale from white (lowest) to deep red (highest). Denominators and exact 95% confidence intervals for each stratum are reported in Table 2. HPV 16/18 = positive to HPV 16 and/or 18; other HR-HPV = positive for at least one of the subsequent HPV types: HPV 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68; NILM = negative for intraepithelial lesion or malignancy; LG = low-grade cytology; HG = high-grade cytology; CIN = cervical intraepithelial neoplasia. Cancer risks refer to the three centers reporting invasive cancers separately (Florence, Padua, Umbria; $n = 3759$); in Emilia Romagna cancers were aggregated with CIN3+.

Table 3. CIN2+ lesions with NILM cytology at the 12-month recall and contribution of HPV16/18 genotyping to their identification (n, of total CIN2+ lesions = 327).

HPV16/18 "Recovery" of CIN2+ Lesions (%)	*n*. of HPV16/18+ Among NILM/CIN2+ (%)	% NILM CIN2+ of All CIN2+	NILM CIN2+/Total CIN2+	Centre
40.00%	16/40	39.60%	40/101	Florence
47.40%	9/19	24.70%	19/77	Padua
38.10%	8/21	45.70%	21/46	Emilia Romagna
47.50%	19/40	38.80%	40/103	Umbria
43.30%	52/120	36.70%	120/327	Total

3.3. Performance of Alternative 12-Month Recall Strategies

Center-level retrospective application of the four alternative strategies (see also Supplementary Tables S5–S8) showed that Strategy 2 consistently offered the most favorable balance between lesion detection and colposcopy reduction in all centers, with some quantitative variation reflecting differences in local population. The relative CIN3+ sensitivity of Strategy 2 (Table S9) ranged from 62.5% in Emilia Romagna to 91.1% in Umbria. The difference was not statistically significant (heterogeneity test $p = 0.067$; Table S10), but center-specific estimates are imprecise because of the small number of CIN3+ cases (16 in Emilia Romagna). In Florence, Strategy 2 would have deferred 59.3% of immediate colposcopies while maintaining a relative sensitivity for CIN3+ of 84.6% (PPV 15.4%). In Padua, colposcopy immediate deferred reached 54.7%, with a relative sensitivity for CIN3+ of 88.5% (PPV 9.5%). In Umbria, Strategy 2 would have deferred 57.0% of immediate colposcopies and achieved the highest CIN3+ relative sensitivity among the four centers (91.1%; PPV 19.6%). Emilia Romagna showed the lowest performance of Strategy 2 (relative sensitivity 62.5% for CIN3+, 10/16, 95% CI 38.6–81.5; six CIN3+ lesions deferred), although estimates were imprecise and heterogeneity between centers was borderline (Pearson χ^2 $p = 0.048$; exact Monte Carlo $p = 0.067$; Table S10).

3.4. Pooled Performance of Alternative Strategies: Identification of the Protocol Showing the Most Favorable Trade-Off

Table 4 and Figure 3 present the pooled performance retrospective application of all four strategies across all centers, comparing each strategy to the current protocol in terms of relative sensitivity for CIN3+ and CIN2+, colposcopy deferred and PPV.

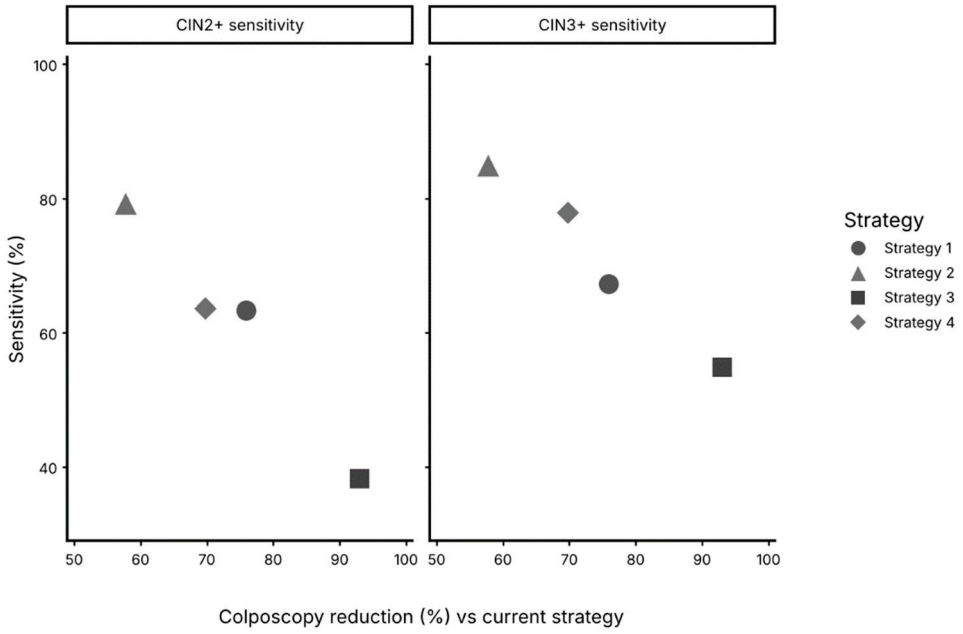


Figure 3. Trade-off plot: colposcopy reduction versus relative sensitivity for CIN2+ and CIN3+ for the four alternative triage strategies. Each point represents one strategy applied to the pooled cohort ($n = 4589$). The x-axis shows the percentage of immediate colposcopy referrals deferred compared to the current strategy; the y-axis shows the relative sensitivity for CIN2+ (**left** panel) and CIN3+ (**right** panel) versus the same reference. The figure is descriptive: no clinical utility threshold or prespecified criterion for acceptable sensitivity or colposcopy reduction was defined.

Strategy 2 showed the most favorable trade-off among the four strategies evaluated, offering the best balance between CIN2+ and CIN3+ sensitivity and colposcopy reduction. Strategies 1 and 3 achieved the highest PPV for CIN2+ (18.7% and 38.8%, respectively), but at the cost of markedly lower CIN3+ sensitivity. Strategy 2 was the only strategy to simultaneously achieve: (a) an 85% relative sensitivity for CIN3+, (b) a 57.7% reduction in colposcopy referrals and (c) an approximate doubling of the PPV for CIN2+ (13.4%). Under the simulated Strategy 2, 17/113 CIN3+ lesions (15.0%) would not have triggered immediate colposcopy and would have been deferred. Their subsequent detection cannot be assumed from the present data, as information on follow-up, persistence or regression of these lesions, and adherence to surveillance was not available (see Limitations). All seven invasive cancers reported separately would have been referred immediately under Strategy 2; the cohort total may be higher, because cancers in Emilia Romagna were aggregated with CIN3+.

In multivariable logistic regression models including HPV genotype, cytology and screening center (Table 5), HPV 16/18 positivity (CIN2+: aOR 1.81, 95% CI 1.40–2.33; CIN3+: aOR 2.83, 95% CI 1.87–4.30) and high-grade cytology (CIN2+: aOR 17.65; CIN3+: aOR 19.63) were independently associated with higher risk of lesions. Low-grade cytology was associated with CIN2+ (aOR 3.42) but not clearly with CIN3+ (aOR 1.67, 95% CI 0.89–3.12). There was no evidence of genotype × cytology interaction ($p = 0.131$ and $p = 0.550$), and risk differed across centers after adjustment. Age could not be included in the models.

Table 4. Pooled performance of all triage strategies versus the current protocol; all centers combined (FI + PD + UM + ER).

Strategy	Women Referred to Colposcopy *n* (%)	Immediate Colposcopy Reduction vs. Current (%)	CIN2+ Detected *n*	PPV CIN2+ (95% CI)	CIN3+ Detected *n*	PPV CIN3+ (95% CI)	Rel. Sens. for CIN2+ (%)	Rel. Sens. for CIN3+ (%)	Cervical Cancers Detected *n* (%; 95% CI) *	Deferred CIN2+ *n*	Deferred CIN3+ *n*	Deferred Cancers * *n*
Current protocol	4589 (100)	-	327	7.13 (6.42–7.91)	113	2.46 (2.05–2.95)	100.00	100.00	7 (0.19; 0.09–0.38)	-	-	-
Strategy 1	1106 (24.1)	75.90 (74.64–77.11)	207	18.72 (16.53–21.12)	76	6.87 (5.53–8.52)	63.30 (57.95–68.34)	67.26 (58.16–75.22)	7 (0.77; 0.37–1.58)	120	37	0
Strategy 2	1940 (42.3)	57.72 (56.29–59.15)	259	13.35 (11.91–14.94)	96	4.95 (4.07–6.01)	79.20 (74.48–83.25)	84.96 (77.22–90.39)	7 (0.43; 0.21–0.89)	68	17	0
Strategy 3	322 (7)	92.98 (92.21–93.69)	125	38.82 (33.66–44.24)	62	19.25 (15.32–23.91)	38.23 (33.12–43.60)	54.87 (45.68–63.73)	6 (2.68; 1.23–5.72)	202	51	1
Strategy 4	1388 (30.2)	69.75 (68.41–71.07)	208	14.99 (13.20–16.96)	88	6.34 (5.17–7.75)	63.61 (58.26–68.64)	77.88 (69.38–84.54)	6 (0.53; 0.24–1.14)	119	25	1

Rel. Sens. = relative sensitivity compared to the current protocol. Deferred lesions: lesions that would not have triggered immediate colposcopy under the simulated strategy. * Cancer percentages refer to colposcopies performed in the three centers reporting invasive cancers separately (n = 3759, 911, 1623, 224 and 1142 for the current protocol and Strategies 1–4, respectively); in Emilia Romagna cancers were aggregated with CIN3+ and could not be reported separately. PPV = positive predictive value (detected lesions/colposcopies performed). 95% confidence intervals for relative sensitivity, PPV, colposcopy reduction and the percentage of cancers detected: Wilson method.

Table 5. Multivariable logistic regression analyses for CIN2+ and CIN3+.

Predictor	CIN2+ Adjusted OR (95% CI)	*p*-Value	CIN3+ Adjusted OR (95% CI)	*p*-Value
HPV genotype				
Other high-risk HPV	Reference	—	Reference	—
HPV16/18	1.81 (1.40–2.33)	<0.001	2.83 (1.87–4.30)	<0.001
Cytology				
NILM	Reference	—	Reference	—
Low-grade (LG)	3.42 (2.54–4.61)	<0.001	1.67 (0.89–3.12)	0.110
High-grade (HG)	17.65 (12.98–24.01)	<0.001	19.63 (12.54–30.73)	<0.001
Screening center				
Florence	Reference	—	Reference	—
Padua	0.51 (0.37–0.71)	<0.001	0.84 (0.47–1.50)	0.559
Umbria	1.25 (0.91–1.71)	0.167	2.22 (1.31–3.77)	0.003
Emilia Romagna	0.39 (0.26–0.59)	<0.001	0.46 (0.23–0.90)	0.024

Models include center fixed effects. Genotype × cytology interaction LR $p = 0.131$ for CIN2+ and $p = 0.550$ for CIN3+. No age adjustment available.

Table 6 shows the observed CIN3+ risk for each genotype–cytology stratum and the corresponding management under Strategy 2. Only the other HR-HPV/NILM stratum, the one deferred under Strategy 2, had a risk clearly below the 2.5% threshold (upper 95% CI limit 1.03%), whereas the HPV 16/18 and other HR-HPV strata with high-grade cytology were clearly above it. HPV 16/18/NILM, HPV 16/18/LG and other HR-HPV/LG had confidence intervals including the threshold. This comparison is exploratory: the 2.5% benchmark [11] was not prespecified in this study and is not presented as a validated decision rule for this cohort.

Table 6. Exploratory risk matrix: observed CIN3+ risk by HPV genotype and cytology at the 12-month recall ($n = 4589$), with exact 95% confidence intervals, and corresponding management under Strategy 2.

HPV Genotype	Cytology	Women, *n*	CIN3+, *n*	CIN3+ Risk, % (95% CI)	95% CI vs. Reference Level †	Management Under Strategy 2 ‡
16/18	NILM	834	20	2.40 (1.47–3.68)	Includes	Immediate colposcopy
16/18	LG	232	6	2.59 (0.95–5.54)	Includes	Immediate colposcopy
16/18	HG	178	42	23.60 (17.57–30.53)	Above	Immediate colposcopy
Other HR-HPV	NILM	2649	17	0.64 (0.37–1.03)	Below	Deferral
Other HR-HPV	LG	552	8	1.45 (0.63–2.84)	Includes	Immediate colposcopy
Other HR-HPV	HG	144	20	13.89 (8.69–20.63)	Above	Immediate colposcopy
Total		4589	113	2.46 (2.03–2.95)	–	–

95% confidence intervals: exact Clopper–Pearson method. † Position of the 95% CI relative to an illustrative reference level of 2.5% CIN3+ risk, as proposed by the Italian multi-societal guidelines [11] for direct colposcopy referral: "Below" = upper CI limit < 2.5%; "Above" = lower CI limit ≥ 2.5%; "Includes" = CI includes 2.5%. This reference level was not prespecified in the present study and is not presented as a validated decision rule for this cohort; the comparison is descriptive. ‡ Deferral = no immediate colposcopy referral; follow-up of deferred women was not assessed in this study.

4. Discussion

The present multicenter retrospective analysis suggests that the current management of women with persistent HR-HPV infection at the 12-month recall may offer opportunities for the application of a risk stratification strategy. Although the current "all-referral" strategy maximizes lesion detection, our findings indicate that it is associated with a relatively low positive predictive value for CIN2+ (7.1%), reflecting the limited specificity of HPV

persistence alone as a referral criterion. These observations support the hypothesis that integrating information already routinely available in screening practice, cytology and partial HPV genotyping, may improve the balance between diagnostic sensitivity for CIN3+ and colposcopy workload.

Our findings suggest that cytology alone would not be sufficient as a triage strategy in women with persistent HR-HPV infection. More than one-third of CIN2+ lesions occurred in women with NILM cytology, consistent with the intrinsic limitations of cytology in this specific clinical setting rather than questioning its overall value within cervical screening. Three principal explanations could be considered for false-negative cytology (so-called "cytologically silent" lesions) in HPV-positive women: (i) inadequate exfoliation of dysplastic cells from the lesion surface (particularly for small, endocervical or immature transformation zone lesions), (ii) sampling errors at the time of specimen collection and (iii) interpretive variability in cytological reading. Notably, a recent blinded review study [27] of 384 liquid-based cytology specimens from HPV-positive women at the 12-month recall, including 128 cases with histologically confirmed CIN2+, demonstrated that the original NILM diagnosis was confirmed, also after the blinded revision, in 92.5% of cases applying the Bethesda System 2014 classification. These findings suggest that a majority of cytologically silent CIN2+ lesions at 12-month recall could be considered a "genuine" morphological negativity at the time of sampling (i.e., inadequate lesion exfoliation) or a sampling error, rather than an interpretive error, reinforcing the argument that cytology, however expertly performed, has intrinsic sensitivity limitations in the HR-HPV-positive triage context and could not be used as "stand-alone" criterion for deferring colposcopy at the 12-month recall.

Conversely, partial HPV 16/18 genotyping provided information complementary to cytology by identifying a substantial proportion of CIN2+ lesions that would otherwise have remained undetected among women with NILM cytology. These findings are consistent with the concept that, in the 12-month recall cohort, combining molecular and cytological information may improve risk stratification compared with either test alone.

The heatmap in Figure 2 provides visually intuitive representation of the CIN2+, CIN3+ and cancer risk: high-grade cytology elevates risk within both partial genotype groups, while HPV 16/18 positivity confers a clinically significant risk even in the absence of cytological abnormality. In fact, as we previously argued, HPV 16/18 genotyping appeared complementary to cytology, recovering 43.3% of the "cytologically silent" CIN2+ lesions (Table 3). Furthermore, HPV 16/18-positive women with NILM cytology showed a CIN2+ risk (6.2%) that already approaches or exceeds thresholds typically used to justify direct colposcopy referral in other settings, and their observed CIN3+ risk (2.4%) is very close to the "≥2.5% threshold" recently proposed by the Italian multi-societal guidelines [11] for direct referral, which is consistent with that risk-tier assignment, although we would caution that a single retrospective dataset cannot substitute for prospective validation.

Given the complementary roles of cytology and HPV 16/18 genotyping, we then compared four strategies combining them, in terms of colposcopy reduction and lesion detection. Among the four retrospectively simulated strategies, Strategy 2 (colposcopy referral for HPV 16/18-positive women regardless of cytology and any HR-HPV-positive woman with low-grade or worse cytology) showed the most favorable trade-off between colposcopy reduction and lesion detection within the present dataset. Although no strategy completely preserved the sensitivity of the current protocol, Strategy 2 maintained a substantial reduction in colposcopy volume (57.7%) with a relatively high sensitivity for CIN3+ (85.0%). These findings suggest that this approach appears to offer the most favorable trade-off among the strategies evaluated, based on our retrospective study, but deserves prospective evaluation before any implementation in routine clinical practice.

Our data indicate that, using laboratory tests currently available (HPV partial genotyping and cytology) a perfect solution does not exist; therefore, improving the efficiency of the clinical protocol necessarily implies accepting a marginal loss in sensitivity, in our case a reduction in relative CIN3+ sensitivity to 85% (17/113 lesions not immediately referred), to achieve greater operational efficiency (a significant reduction of immediate colposcopies). Strategies 1 and 3 achieved greater colposcopy reduction, but at what appears to be a less acceptable cost in CIN3+ detection (relative sensitivity 67.3% and 54.9%, respectively); Strategy 3, despite its high PPV (38.8%), would have deferred the detection of 51 CIN3+ lesions. Strategy 4 offered intermediate performance without a clear advantage over Strategy 2.

Overall, Strategy 2 seems to represent a reasonable balance between diagnostic yield and resource use, although its adoption should ideally be followed by close monitoring of outcomes in the deferred population.

These findings are in line with evidence published after the study period and with the recent Italian guidelines [11], all of which are in line with a progressively more risk-based management of HPV-positive women. Our analysis was conducted retrospectively on historical data, before the publication of the 2025 Italian guidelines [11]; its findings are therefore hypothesis-generating, consistent with (but independent of) the more recent recommendations, and should be confirmed in prospective studies. The Danish HCP (Health Care Policy) Trial [20] showed that combining genotyping with cytology outperforms cytology alone, with sensitivity gains broadly in line with our retrospective results. The NTCC2 (New Technology in Cervical Cancer 2) study [18], using extended genotyping, identified a low-oncogenic-genotype/cytology-negative subgroup with a CIN3+ risk below 0.4% at 24 months; the analogous deferred group in our historical data (non-16/18 HR-HPV, NILM cytology) showed a residual CIN3+ risk of 0.64% (95% CI 0.37–1.03%), in the same low range. This is, however, an aggregate estimate: since the Cobas 4800 platform did not allow the heterogeneous "other HR-HPV" group to be sub-stratified, this low risk should not be assumed to apply uniformly to all non-16/18 types. Even the 2025 Italian multi-societal guidelines [11], explicitly endorse complete or at least partial genotype- and cytology-based risk stratification and an "equal management for equal risk" principle, providing a normative framework broadly consistent with the rationale underlying Strategy 2, even though the limitations of the instrument used by the participating centers prevented a more precise risk stratification in relation to the specific HPV type.

Despite this consistency with the external evidence, the performance of Strategy 2 was not uniform across the participating centers, and this heterogeneity must be considered before generalizing the pooled results. The six-fold variation in DR (0.5‰ in Padua vs. 2.8‰ in Umbria) and the more than two-fold variation in PPV for CIN2+ (4.9% in Padua vs. 10.6% in Umbria) across the four centers carry implications for national-scale implementation of triage protocols. These differences may reflect differences in local populations and clinical practice, as well as in the screening periods, which differed for Florence (2016–2018) compared with the other centers (2015–2019). While Strategy 2 consistently appeared to deliver the most favorable balance across centers, the magnitude of its benefits varied by local context. Centers with lower baseline PPV for CIN2+ or CIN3+ (e.g., Padua, Emilia Romagna) may gain proportionally more from a triage step, since the absolute potential reduction in colposcopies not identifying CIN2+ is larger; conversely, the relative sensitivity of Strategy 2 appeared more robust in higher-prevalence centers (Umbria: 91.1% for CIN3+) than in lower-prevalence settings (Emilia Romagna: 62.5%). In Emilia Romagna, Strategy 2 would have deferred six of 16 CIN3+ lesions: although imprecise, this suggests that the safety margin of the strategy may be narrower in some local contexts. This heterogeneity may argue against a single and inflexible national threshold, in favor of a possible strategy based on locally calibrated implementation, consistent with the risk-based, threshold-driven

framework of the 2025 Italian guidelines. Screening programs with lower HPV 16/18 prevalence or lower CIN3+ rates may need extended genotyping, or identification of individual genotypes beyond 16/18 versus other, to safely customize a 12-month recall triage.

Nevertheless, the observed variability between-center also highlights that the performance of any triage strategy is likely to depend on the local epidemiological context, suggesting that external validation across additional screening programs would be desirable.

Beyond center-level differences, the safety of any deferred-management strategy (i.e., Strategy 2) ultimately depends on the ability or capacity of screening programs to ensure adequate follow-up adherence among women not immediately referred to colposcopy, although the risk of CIN3+ is relatively low (i.e., 0.64% in our dataset). Therefore, the findings of the present study should be interpreted within the context of organized screening programs capable of ensuring high compliance with repeat testing and appropriate surveillance of deferred women. The clinical harm–benefit balance of reducing immediate colposcopies depends not only on the burden of colposcopy but also on the consequences of delayed diagnosis in deferred women, which could not be assessed in this study. Prospective studies and formal decision-analytic or net-benefit analyses are needed to confirm these findings and to define an acceptable threshold of CIN3+ sensitivity loss.

5. Limitations

Limitations of this study must be acknowledged.

The 248 women (5.1%) excluded because of non-adherence to colposcopy ($n = 223$) or inconclusive cytology ($n = 25$) could not be assessed, and verification bias cannot be excluded. No additional verification period was applied after the index colposcopy (12-month recall), so lesions diagnosed at later follow-up in women with negative colposcopy were not captured. Details on the number and type of biopsies per woman were not analyzed, so we could not assess variability in sampling practices across centers. The retrospective design precludes direct assessment of the outcomes of women who, under alternative strategies, would have been deferred from colposcopy. Adherence to subsequent follow-up of women not immediately referred to colposcopy could not be assessed, as no data on their subsequent examinations, regression or cancer incidence were available; the safety of deferred management cannot therefore be inferred from the present data and depends on the capacity of screening programs to ensure follow-up. Adherence to follow-up and colposcopy in Italian organized screening programs is generally high (about 90%; GISCI- National Cervical cancer screening survey, 2025), but this was not verified in the present cohort.

All malignant lesions (Cancers) are defined as "detected" basing on the observed historical cohort. Invasive cancers were reported separately in three of the four centers (Florence, Padua, Umbria); in Emilia Romagna they were aggregated with CIN3+ and could not be distinguished. Therefore, the total number of invasive cancers in the cohort may exceed seven, and cancer risks refer to the three centers reporting them separately. The small number of cancers does not allow conclusions on the safety of Strategy 2 for invasive cancer. No further data are available. Regression and heterogeneity: age could not be included in the models because joint age-by-genotype-by-cytology-by-outcome aggregates were unavailable.

Screening periods differed across centers (2016–2018 for Florence and 2015–2019 for the other centers), and this may have contributed to between-center variability.

The number of CIN3+ cases was small in some centers (16, Emilia Romagna), limiting the precision of center-specific estimates and the power to detect heterogeneity. No prespecified criterion for acceptable loss of sensitivity was defined, so the identification of Strategy 2 as the most favorable trade-off is descriptive. This strategy should therefore be regarded

as a candidate strategy requiring prospective validation, and future studies should apply formal approaches such as decision-analytic modelling or net-benefit analysis.

The HPV genotyping strategy used in this retrospective evaluation was the "partial" one (HPV 16, HPV 18 and other HR-HPV), precluding discrimination of the oncogenically heterogeneous "other" group into its specific genotypes; as Strategy 2 defers mainly other HR-HPV/NILM women, a uniformly low risk across these types could not be verified.

This evaluation was conducted in areas with a well-established HPV screening programs and, after the transition from Pap-based to HPV-based, characterized by a strong and consolidated Pap-based screening program; consequently, regions that adopted HPV-based screening more recently or who lacked a previous Pap-based screening program may present a different HR-HPV distribution or CIN2+ burden, and probably may need to modulate 12-month recall triage strategies.

All data predate widespread HPV vaccination uptake in the screened cohorts; as vaccinated cohorts enter the screening program, HPV 16/18 prevalence is expected to decline, and the risk gradient between genotype strata may shift, potentially altering the performance of a partial HPV genotype-driven triage strategy over time.

The alternative management strategies were evaluated through retrospective simulation rather than prospective clinical implementation. Consequently, the observed diagnostic performance should be interpreted as an estimate of the potential impact of each strategy rather than evidence of real-world effectiveness. Prospective validation remains extremely important.

6. Conclusions

This multicenter retrospective study suggests that integrating partial HPV genotyping with cytology at the 12-month recall may improve risk stratification among women with persistent HR-HPV infection compared with an "all-referral" strategy based solely on HR-HPV persistence. In multivariable analyses adjusted for center, HPV 16/18 genotype and high-grade cytology were independently associated with CIN2+ and CIN3+, supporting the combined use of genotype and cytology. Within this dataset, the combined approach would have substantially reduced colposcopy referrals while maintaining relatively high sensitivity for CIN3+ detection. Furthermore, this improvement could be implemented without requiring women to be recalled for additional testing, but instead actively exploiting laboratory results already routinely available at screening centers (partial HPV genotyping and cytology) and recorded in IT systems.

Among the four alternative triage strategies evaluated, Strategy 2 (immediate colposcopy referral for HPV 16/18-positive women or for HPV-other-positive women with low-grade or worse cytology, with a deferred follow-up for HPV-other-positive/NILM-cytology women) appears to offer the most favorable trade-off between diagnostic performance and colposcopy workload. Applied across the 270,714 screened women from four Italian centers, it would reduce colposcopy referrals at 12-month recall by 57.7%, nearly double the current PPV for CIN2+ lesions (13.4%), and maintain an 85% relative sensitivity for CIN3+. The 17 CIN3+ lesions deferred applying Strategy 2 are not necessarily biologically or clinically "missed"; rather, they are lesions that would not have triggered immediate colposcopy under the simulated algorithm. In addition, invasive cancers were reported separately in only three of the four centers, and their small number does not allow conclusions on the safety of Strategy 2 for invasive cancer. However, these findings derive from a retrospective simulation and should therefore be regarded as hypothesis-generating rather than practice-changing.

Strengthening recall infrastructure therefore appears to be a prerequisite for safely implementing risk-based triage; screening programs should consider investing in multi-

modal active recall systems (SMS, e-mail, telephone contact, social media instant messaging, push-notification on a dedicated app), dedicated colposcopy pathways for deferred women in follow-up and robust tracking mechanisms to flag non-responders for priority outreach.

Future prospective studies should evaluate the safety and effectiveness of Strategy 2 in routine clinical practice, particularly with respect to long-term CIN3+ incidence among women whose colposcopy would be deferred. Further research should also assess the incremental value of extended HPV genotyping, p16/Ki-67 dual staining [28] and other emerging biomarkers and define risk thresholds appropriate for increasingly HPV-vaccinated populations.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/cancers18223621/s1. Table S1. Detailed 12-month recall results—Florence. Table S2. Detailed 12-month recall results—Padua. Table S3. Detailed 12-month recall results—Umbria. Table S4. Detailed 12-month recall results—Emilia Romagna. Table S5. Performance of alternative triage strategies at the 12-month recall—Florence. Table S6. Performance of alternative triage strategies at the 12-month recall—Padua. Table S7. Performance of alternative triage strategies at the 12-month recall—Umbria. Table S8. Performance of alternative triage strategies at the 12-month recall—Emilia Romagna. Table S9. Performance of Strategy 2: variation between centers in the relative sensitivity for CIN2+ and CIN3+. Table S10. Tests of between-center heterogeneity.

Author Contributions: Conceptualization, G.P. and S.G.; methodology, G.P., G.G., D.P. and S.G.; formal analysis, G.P., G.G., D.P. and S.G.; investigation, G.P., A.M., I.P., S.P., B.F., C.S., ISPRO Cervical Cancer Screening Working Group and S.G.; data curation, G.P., G.G., D.P. and S.G.; writing—original draft preparation, G.P., A.M., I.P. and S.G.; writing—review and editing, G.P., A.M., I.P., G.G., D.P., S.P., B.F., A.D.M., H.F., G.S., C.C. (Chiara Castagnetti), C.C. (Cinzia Campari), A.A., C.B., D.G., E.C., S.D., R.G., C.S. and S.G.; supervision, C.S. and S.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study is a retrospective analysis of anonymous aggregated data routinely collected within the organized cervical cancer screening programs of Florence, Padua, Umbria and Emilia Romagna; no additional procedures were performed and clinical management followed standard of care. Ethical review and approval were waived for this study because it involved only the secondary analysis of anonymous aggregated routine data. Data were processed in accordance with the GDPR (Regulation (EU) 2016/679) and Italian Legislative Decree 196/2003 as amended by Legislative Decree 101/2018.

Informed Consent Statement: Informed consent for the use of data for scientific and epidemiological research was obtained from all women at the time of sample collection, when they participated in the respective screening programs.

Data Availability Statement: The individual-level data are held by the regional screening programs; the analyses were performed on aggregated data. These data cannot be made publicly available because of privacy regulations (GDPR and Italian Legislative Decree 196/2003). The aggregated, non-identifiable counts underlying all analyses (by center, HPV genotype, cytology and histological outcome) are reported in Table 2 and Supplementary Tables S1–S4.

Acknowledgments: The authors would like to thank all the healthcare professionals, operators, and colleagues involved in the cervical cancer screening program for their invaluable support and dedication. The ISPRO Cervical Cancer Screening Working Group includes: Simonetta Bisanzi, Angela Cappadona Sciametta, Giulia Fantacci, Jessica Viti, Serena Giunti, Sara Galastri, Chiara Di Stefano, Stefania Cannistrà, Claudia Giachini and Lisa Lelli.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ronco, G.; Giorgi-Rossi, P.; Carozzi, F.; Confortini, M.; Palma, P.D.; Del Mistro, A.; Ghiringhello, B.; Girlando, S.; Gillio-Tos, A.; De Marco, L.; et al. Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: A randomised controlled trial. *Lancet Oncol.* **2010**, *11*, 249–257. [CrossRef] [PubMed]
2. Rijkaart, D.C.; Berkhof, J.; Rozendaal, L.; van Kemenade, F.J.; Bulkmans, N.W.; Heideman, D.A.; Kenter, G.G.; Cuzick, J.; Snijders, P.J.; Meijer, C.J. Human papillomavirus testing for the detection of high-grade cervical intraepithelial neoplasia and cancer: Final results of the POBASCAM randomised controlled trial. *Lancet Oncol.* **2012**, *13*, 78–88. Correction in *Lancet Oncol.* **2015**, *16*, E480. [CrossRef] [PubMed]
3. Kitchener, H.C.; Almonte, M.; Thomson, C.; Wheeler, P.; Sargent, A.; Stoykova, B.; Gilham, C.; Baysson, H.; Roberts, C.; Dowie, R.; et al. HPV testing in combination with liquid-based cytology in primary cervical screening (ARTISTIC): A randomised controlled trial. *Lancet Oncol.* **2009**, *10*, 672–682. Erratum in *Lancet Oncol.* **2009**, *10*, 748. [CrossRef] [PubMed]
4. Naucler, P.; Ryd, W.; Törnberg, S.; Strand, A.; Wadell, G.; Elfgren, K.; Rådberg, T.; Strander, B.; Johansson, B.; Forslund, O.; et al. Human Papillomavirus and Papanicolaou Tests to Screen for Cervical Cancer. *N. Engl. J. Med.* **2007**, *357*, 1589–1597. Erratum in *N. Engl. J. Med.* **2008**, *359*, 1637. [CrossRef] [PubMed]
5. Ronco, G.; Dillner, J.; Elfström, K.M.; Tunesi, S.; Snijders, P.J.F.; Arbyn, M.; Kitchener, H.; Segnan, N.; Gilham, C.; Giorgi-Rossi, P.; et al. Efficacy of HPV-based screening for prevention of invasive cervical cancer: Follow-up of four European randomised controlled trials. *Lancet* **2014**, *383*, 524–532. Erratum in *Lancet* **2015**, *386*, 1446. [CrossRef] [PubMed]
6. World Health Organization. *WHO Guideline for Screening and Treatment of Cervical Pre-Cancer Lesions for Cervical Cancer Prevention*, 2nd ed.; WHO: Geneva, Switzerland, 2021.
7. European Cancer Beating Plan (2021–2030). Available online: https://health.ec.europa.eu/document/download/26fc415a-1f28-4f5b-9bfa-54ea8bc32a3a_en (accessed on 3 August 2026).
8. World Health Organization. *WHO Guideline for Screening and Treatment of Cervical Pre-Cancer Lesions for Cervical Cancer Prevention: Use of Human Papillomavirus (HPV) DNA Genotyping*; WHO: Geneva, Switzerland, 2026. Available online: https://www.who.int/publications/i/item/9789240121744 (accessed on 26 October 2026).
9. Kjær, S.K.; Frederiksen, K.; Munk, C.; Iftner, T. Long-term absolute risk of cervical intraepithelial neoplasia grade 3 or worse following human papillomavirus infection: Role of persistence. *J. Natl. Cancer Inst.* **2010**, *102*, 1478–1488. [CrossRef] [PubMed]
10. Polman, N.J.; Veldhuijzen, N.J.; Heideman, D.A.M.; Snijders, P.J.F.; Meijer, C.J.L.M.; Berkhof, J. HPV-positive women with normal cytology remain at increased risk of CIN3 after a negative repeat HPV test. *Br. J. Cancer* **2017**, *117*, 1557–1561. [CrossRef] [PubMed]
11. Gori, S.; Venturelli, F.; Carozzi, F.; Giorgi Rossi, P.; Del Mistro, A.; on behalf of the Multisocietal Contributor Group. Italian multisocietal guidelines for cervical cancer screening: Biomarker use in an HPV-based screening programme with a risk-based approach and GRADE methodology. *Br. J. Cancer* **2025**, *133*, 1076–1084. Erratum in *Br. J. Cancer* **2025**, *133*, 1927. [PubMed]
12. Carozzi, F.; Burroni, E.; Confortini, M.; Pompeo, G.; Bisanzi, S.; Cellai, F.; Paganini, I.; Mantellini, P.; Iossa, A.; Lelli, L.; et al. Implementation of a centralized HPV-based cervical cancer screening programme in Tuscany: First round results and comparison with the foregoing Pap-based screening programme. *J. Med. Screen.* **2022**, *29*, 110–122. [CrossRef] [PubMed]
13. Rebolj, M.; Rimmer, J.; Denton, K.; Tidy, J.; Mathews, C.; Ellis, K.; Smith, J.; Evans, C.; Giles, T.; Frew, V.; et al. Primary cervical screening with high risk human papillomavirus testing: Observational study. *BMJ* **2019**, *364*, l240. [CrossRef] [PubMed]
14. Castle, P.E.; Schiffman, M.; Wheeler, C.M.; Solomon, D. Evidence for Frequent Regression of Cervical Intraepithelial Neoplasia–Grade 2. *Obstet. Gynecol.* **2009**, *113*, 18–25. [CrossRef] [PubMed]
15. Schiffman, M.; Hyun, N.; Raine-Bennett, T.R.; Katki, H.; Fetterman, B.; Gage, J.C.; Cheung, L.C.; Befano, B.; Poitras, N.; Lorey, T.; et al. A cohort study of cervical screening using partial HPV typing and cytology triage. *Int. J. Cancer* **2016**, *139*, 2606–2615. [CrossRef] [PubMed]
16. Castle, P.E.; Stoler, M.H.; Wright, T.C.; Sharma, A., Jr.; Wright, T.L.; Behrens, C.M. Performance of carcinogenic human papillomavirus (HPV) testing and HPV16 or HPV18 genotyping for cervical cancer screening of women aged 25 years and older: A subanalysis of the ATHENA study. *Lancet Oncol.* **2011**, *12*, 880–890. [CrossRef] [PubMed]
17. Hashim, D.; Engesæter, B.; Baadstrand Skare, G.; Castle, P.E.; Bjørge, T.; Tropé, A.; Nygård, M. Real-world data on cervical cancer risk stratification by cytology and HPV genotype to inform the management of HPV-positive women in routine cervical screening. *Br. J. Cancer* **2020**, *122*, 1715–1723. [CrossRef] [PubMed]
18. De Marco, L.; Bisanzi, S.; Ronco, G.; Mancuso, P.; Carozzi, F.; Allia, E.; Rizzolo, R.; Gustinucci, D.; Frayle, H.; Viti, J.; et al. Extended HPV genotyping by the BD Onclarity assay: Concordance with screening HPV-DNA assays, triage biomarkers, and histopathology in women from the NTCC2 study. *Microbiol. Spectr.* **2025**, *13*, e0089724. [CrossRef] [PubMed]
19. Stoler, M.H.; Baker, E.; Boyle, S.; Aslam, S.; Ridder, R.; Huh, W.K.; Wright, T.C., Jr. Approaches to triage optimization in HPV primary screening: Extended genotyping and p16/Ki-67 dual-stained cytology-Retrospective insights from ATHENA. *Int. J. Cancer* **2020**, *146*, 2599–2607. [CrossRef] [PubMed]

20. Gori, S.; Battagello, J.; Gustinucci, D.; Campari, C.; Zorzi, M.; Frayle, H.; Passamonti, B.; Sartori, G.; Bulletti, S.; Fodero, C.; et al. Clinical relevance of partial HPV16/18 genotyping in stratifying HPV-positive women attending routine cervical cancer screening: A population-based cohort study. *BJOG Int. J. Obstet. Gynaecol.* **2020**, *128*, 1353–1362. [CrossRef] [PubMed]
21. Schroll, J.B.; Bonde, J.; Lynge, E.; Waldstrøm, M.; Viborg, P.H.; Frandsen, A.; Andersen, R.H.; Nielsen, S.M.; Kristensen, B.K.; Schledermann, D.; et al. Health care policy trial of primary human papillomavirus–based cervical screening in Denmark: Comparison of three triage algorithms. *Int. J. Cancer* **2026**, *158*, 3197–3207. [CrossRef] [PubMed]
22. Canfell, K.; Saville, M.; Caruana, M.; Gebski, V.; Darlington-Brown, J.; Brotherton, J.; Heley, S.; Castle, P.E. Protocol for Compass: A randomised controlled trial of primary HPV testing versus cytology screening for cervical cancer in HPV-unvaccinated and vaccinated women aged 25–69 years living in Australia. *BMJ Open* **2018**, *8*, e016700. Correction in *BMJ Open* **2018**, *8*, e016700corr1. [CrossRef] [PubMed]
23. Massad, L.S.; Clarke, M.A.; Perkins, R.B.; Garcia, F.; Chelmow, D.; Cheung, L.C.; Darragh, T.M.; Egemen, D.; Lorey, T.S.; Nayar, R.; et al. Applying Results of Extended Genotyping to Management of Positive Cervicovaginal Human Papillomavirus Test Results: Enduring Guidelines. *J. Low. Genit. Tract Dis.* **2025**, *29*, 134–143. [CrossRef] [PubMed]
24. Wright, T.C.; Stoler, M.H.; Ranger-Moore, J.; Fang, Q.; Volkir, P.; Safaeian, M.; Ridder, R. Clinical validation of p16/Ki-67 dual-stained cytology triage of HPV-positive women: Results from the IMPACT trial. *Int. J. Cancer* **2021**, *150*, 461–471. [CrossRef] [PubMed]
25. Nayar, R.; Wilbur, D.C. The Pap test and Bethesda 2014. *Cancer Cytopathol.* **2015**, *123*, 271–281. [CrossRef] [PubMed]
26. Bornstein, J.; Bentley, J.; Bösze, P.; Girardi, F.; Haefner, H.; Menton, M.; Perrotta, M.; Prendiville, W.; Russell, P.; Sideri, M.; et al. 2011 colposcopic terminology of the International Federation for Cervical Pathology and Colposcopy. *Obstet. Gynecol.* **2012**, *120*, 166–172. [CrossRef] [PubMed]
27. Cannistrà, S.; Carozzi, F.; Di Stefano, C.; Matucci, M.; Pompeo, G.; Gorini, G.; Puliti, D.; Zappa, M.; Sani, C.; Confortini, M. Accuracy and Reproducibility of Cytology Triage in a HPV-Based Primary Screening Setting: A Revision of 384 Pap Tests. *Acta Cytol.* **2023**, *67*, 618–628. [CrossRef] [PubMed]
28. Benevolo, M.; Ronco, G.; Mancuso, P.; Carozzi, F.; De Marco, L.; Allia, E.; Bisanzi, S.; Rizzolo, R.; Gustinucci, D.; Del Mistro, A.; et al. Comparison of HPV-positive triage strategies combining extended genotyping with cytology or p16/ki67 dual staining in the Italian NTCC2 study. *EBioMedicine* **2024**, *104*, 105149. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.