

Reproducibility of a Rapid Human Papillomavirus Test at Different Levels of the Healthcare System in Tanzania: The AISHA Study

Armando Baena¹, Hugo De Vuyst¹, David Mesher^{1,2}, Mabula Kasubi³, Safina Yuma⁴, Julius Mwaiselage⁵, Semi Zouiouich⁶, Pendo Mlay⁷, Crispin Kahesa⁵, Sihem Landoulsi⁸, Maria de la Luz Hernandez¹, Eric Lucas⁹, Rolando Herrero¹, Maribel Almonte¹, and Nathalie Broutet¹⁰

ABSTRACT

Background: To inform policy makers in Tanzania if and how best to implement rapid HPV testing, we assessed the interobserver reproducibility of *careHPV* test at three different levels of the healthcare system in an urban and a rural region of Tanzania.

Methods: Women aged 30 to 50 years were screened by *careHPV* testing in two primary healthcare centers (PHC), two district hospitals (DiH), and two regional hospitals (ReH). Aliquots were retested at regional (ReH) and national referral laboratories (NRL). Reproducibility was evaluated using agreement and kappa index measures. Intralaboratory reproducibility was also evaluated in a set of 10 positive and 10 negative samples.

Results: Samples from 1,134 women were locally tested and retested at ReH and/or NRL. Test results from Dar es Salaam ReH and Kilimanjaro PHC showed clear quality problems

including suspicion of contamination during testing or aliquoting. After excluding these samples, 18.8% of 743 women were HPV positive at clinic level. The resulting *careHPV* reproducibility at different levels of the healthcare system was very good [agreement 95.7%, 95% confidence interval (CI), 94.0–96.9; kappa, 0.86, 95% CI, 0.81–0.91]. Intralaboratory agreement was also very good across four different experiments, with Fleiss' kappa between 0.87 (95% CI, 0.61–1.00) and 1.00 (0.75–1.00).

Conclusions: Rapid HPV testing was highly reproducible between lower and higher levels of the healthcare system in Tanzania; however, performance seems to be operator dependent.

Impact: The *careHPV* test seems to be a feasible option for cervical cancer screening in an organized, decentralized system and in limited-resource settings if quality assurance measures are in place.

Introduction

Cervical cancer is the fourth most common cancer among women in the world. Tanzania has the fifth highest incidence rate of cervical cancer worldwide (age-standardized incidence rate 59.1/100,000; ref. 1). Thus, this cancer is a major public health problem in Tanzania

and control measures are absolutely necessary and must be implemented promptly.

High-risk (hr) human papillomavirus (HPV) persistence is the primary cause of cervical neoplasia (2). HPV vaccination of young girls is ongoing in many countries around the world, including Tanzania, but cervical screening of adult women remains important. In fact, in May 2018, the World Health Organization (WHO) Cervical Cancer Elimination Initiative was launched with three main actions to reduce cervical cancer rates: (i) to increase coverage of vaccination against HPV, (ii) to increase screening coverage using HPV testing, and (iii) to offer appropriate management of women who screen positive.

Despite availability of several high sensitive HPV tests (3), the high cost of the tests and the challenges associated with their implementation (e.g., laboratory requirements, training of personnel, availability of sample collection and storage consumables, logistics for samples transport) have restrained the introduction of HPV testing in low- and middle-income countries (LMIC). Novel less-costly and easy-to-perform HPV tests may be more suitable for LMIC, although some require further field evaluation (4–6). One of them, *careHPV* (QIA-GEN) based on the Hybrid Capture 2 (HC2) signal amplification technology, has recently been prequalified by WHO. *careHPV* detects 14 oncogenic HPV types, can be performed in less than 3 hours, requires less training and laboratory equipment than HC2, and is, to our knowledge, the least expensive validated HPV test currently available (5–9).

Implementation of HPV testing in public health laboratories requires extensive planning on how to operationalize the testing (e.g., selection of the test, laboratories, and technicians to perform the test). The AISHA study aimed at investigating the robustness of *careHPV* when performed in laboratories with different capacities at different settings. Here, we present results on the intralaboratory

¹Prevention and Implementation Group, Early Detection and Prevention Section, International Agency for Research on Cancer (IARC), Lyon, France. ²Blood Safety, Hepatitis, Sexually Transmitted Infections (STI) and HIV Service, Public Health England, London, United Kingdom. ³Department of Microbiology/Immunology, Muhimbili National Hospital, Dar es Salaam, Tanzania. ⁴Reproductive and Child Health Section, Ministry of Health, Community Development, Gender, Elderly and Children, Dodoma, Tanzania. ⁵Cancer Prevention Services, Ocean Road Cancer Institute, Dar es Salaam, Tanzania. ⁶Nutritional Epidemiology Group, Nutrition and Metabolism Section, International Agency for Research on Cancer (IARC), Lyon, France. ⁷Department of Gynecology and Obstetrics, Kilimanjaro Christian Medical Centre, Kilimanjaro, Tanzania. ⁸UNDP/UNFPA/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction, World Health Organization, Geneva, Switzerland. ⁹Screening Group, Early Detection and Prevention Section, International Agency for Research on Cancer (IARC), Lyon, France. ¹⁰Department of Reproductive Health and Research, World Health Organization, Geneva, Switzerland.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (<http://cebp.aacrjournals.org/>).

A. Baena and H. De Vuyst are the co-first authors of this article.

Corresponding Author: Hugo De Vuyst, International Agency for Research on Cancer (IARC), Lyon Cedex 08, Lyon 69372, France. Phone: 334-7273-8421; Fax: 334-7273-8345; E-mail: devuysth@iarc.fr

Cancer Epidemiol Biomarkers Prev 2020;29:2261–8

doi: 10.1158/1055-9965.EPI-20-0422

©2020 American Association for Cancer Research.

and inter-observer reproducibility of *careHPV* when performed at laboratories from different healthcare levels in Tanzania.

Materials and Methods

Study design

This study was carried out between May 2015 and October 2017 in clinical centers from two regions in Tanzania [Dar es Salaam (DES) and Kilimanjaro] among women aged 30 to 50 years. In both regions, a recruiting center was selected from each of three different levels of healthcare (six centers in total). The first healthcare level included a primary health center (PHC) that provides basic health services. The second level corresponded to the district hospital (DiH) that provides additional services and acts as a referral hospital for the PHC. The third level corresponded to the regional/reference hospital (ReH) that provides similar services to DiH but also may have additional specialized facilities not available at the DiH level including gynecology and surgery. The laboratory services and infrastructure of the different levels reflected the respective clinical needs and we noted that urban labs were usually better equipped than their rural counterparts, except for the ReH (Supplementary Table S1). At the Kilimanjaro PHC, the lab services were very basic and run by 1 laboratory assistant (certificate holder): blood slides for malaria, sputum (TB), stools, urine, pregnancy test, and rapid HIV testing, without back-up generator. At the other end of the range, Kilimanjaro ReH employed 40 technicians including 4 lab scientists and offered a complete range of services including chemistry, hematology, pathology, cytology, histopathology, IHC, and having a back-up generator available. Asymptomatic women from the clinic catchment area were invited to attend screening by community health workers attached to the selected recruitment centers. Women who were non-sexually active, pregnant, with a history of cervical precancer or invasive cancer, with hysterectomies, or with serious preexisting conditions were not included. At the baseline visit, and after obtaining written informed consent, a

cervical sample was collected with *careBrush* and placed in *careHPV* collection medium (QIAGEN). Then, women were screened by visual inspection with acetic acid (VIA) by trained nurses. HPV testing was performed at the laboratory of the recruiting center (PHC, DiH, or ReH). Samples included in the second and later test batches from the PHC and DiH recruiting centers were aliquoted into two 200- μ L vials and retested at the corresponding ReH and the national reference laboratory (NRL) of Tanzania. Similarly, original samples from second and later batches collected at the ReH recruiting centers were aliquoted and retested at the NRL (Fig. 1). The first testing batch was considered as a “learning” experience for the clinical and laboratory staff and therefore was not retested at higher levels. Four laboratory technologists (2 from the NRL, 1 from the ReH in DES, and 1 external technician from the National Health Laboratory Quality Assurance and Training Centre) were trained by the *careHPV* manufacturer. They, in turn, trained one laboratory technician from each study center.

Women with a positive VIA result were immediately treated with cryotherapy at the recruiting center, or the corresponding ReH when cryotherapy was not locally available. Women with a VIA result suspicious for cancer were referred to the corresponding ReH for diagnosis and treatment. VIA-negative women with HPV-negative results were returned to routine screening. VIA-negative women with a positive HPV test at any testing level were recalled for a second VIA and those positive were treated with cryotherapy.

HPV testing

The *careHPV* test has been specifically designed for cervical screening of women living in low-resource regions. It is a nucleic acid hybridization assay that utilizes chemiluminescent signal detection of HPV types based on signal amplification to detect the presence of fourteen high-risk types of HPV-DNA (16/18/31/33/35/39/45/51/52/56/58/59/66/68). The *careHPV* test is a manual assay test designed to be performed by any healthcare worker. The throughput of the test is 90 specimens per 96-well kit with a running time of less than 3 hours.

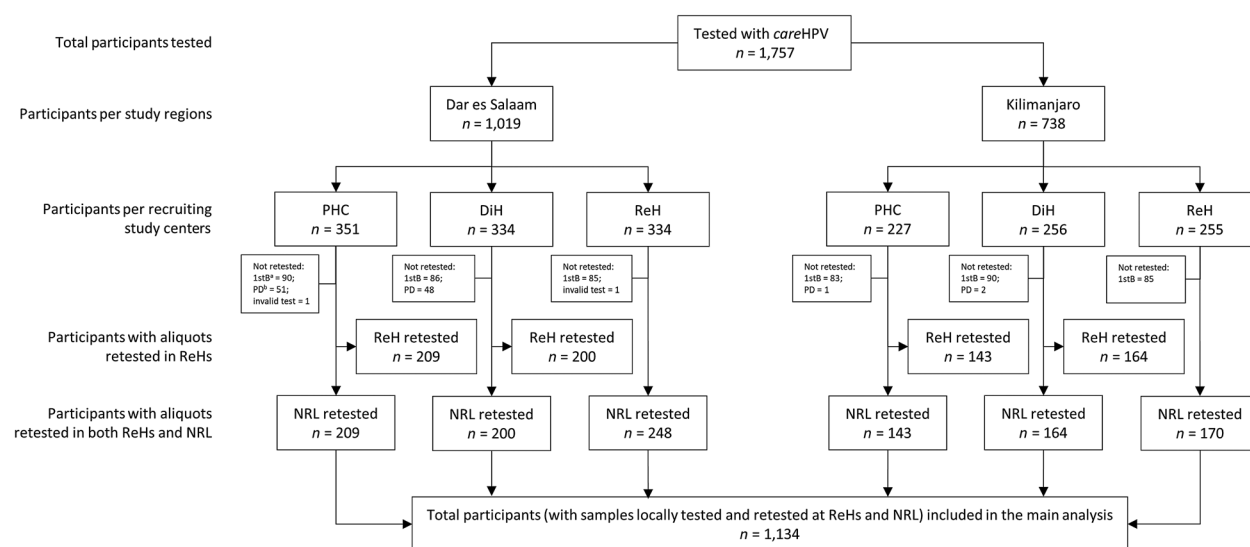


Figure 1.

Design and study population of the AISHA study. Flowchart including the number of women tested with *careHPV* at recruitment centers and their aliquots retested at higher levels of the healthcare system. ^aSamples belonging to the 1st local testing batch were not retested because they were considered a learning batch for the newly trained lab technologists. ^bSome samples were missing retests at ReH, NRL, or both due to protocol deviations. 1stB, first testing batch; PD, protocol deviation.

Six wells are used for commercial controls (3 negative and 3 positive) included in the test kits. In our study, we included 85 study samples and 5 extra negative controls per test batch. According to manufacturers' instructions, cervical specimens placed in the *careHPV* collection medium can be maintained at 15°C to 30°C for 14 days, or refrigerated at 2°C to 8°C for maximum 30 days before testing. We kept samples and aliquots refrigerated during transport and until testing. We included in the analysis 138 samples tested after 30 days of collection after confirming that testing reproducibility was not affected (Supplementary Table S2).

Statistical analysis

We estimated HPV positivity with corresponding binomial 95% confidence intervals (95% CI) for each recruiting center, for the local and retested samples separately, as the proportion of all samples with a high-risk HPV infection.

Interobserver reproducibility of the *careHPV* test was assessed between different levels of the healthcare system using percentage agreement and unweighted Cohen kappa coefficients with 95% CIs. Strength of agreement was judged based on kappa values (10): very good if >0.80; good if >0.60 and ≤0.80; moderate if >0.40 and ≤0.60, fair if >0.20 and ≤0.40 or poor if ≤0.20. Agreement was evaluated for (i) each of the six recruiting centers versus the NRL, (ii) PHC and DiH centers versus ReHs, and (iii) PHC and DiH retests by ReHs versus NRL.

Using additional samples, intralaboratory reproducibility was assessed using four laboratory schemes: In-Run (NRL), Run-to-Run (NRL), Operator-to-Operator (ReH), and Day-to-Day (NRL). In each scheme, the same batch of 20 samples (10 HPV positive and 10 HPV negative) was run three consecutive times as follows: In-Run: same batch tested three times in the same run; Run-to-Run: same batch tested in three consecutive runs on the same day; Operator-to-Operator: same batch tested in three runs by three different operators; Day-to-Day: same batch tested in one run per day on three consecutive days. The percentage agreement and Fleiss' kappa coefficient for three raters were calculated.

To investigate potential technical concerns with the testing (e.g., contamination or issues with aliquoting), we identified positive test results clustered adjacently. A cluster was defined as having at least two adjacent positive wells either horizontally or vertically in the microplate (Supplementary Fig. S1). The cluster size corresponded to the number of positive results clustered (a value of 1 corresponding to "isolated" positive result). We calculated the cluster size (median, interquartile range) by recruiting center and laboratory. Then, we calculated the proportion of samples that tested positive locally which aliquots retested negative at the NRL according to the cluster size. Using binomial regression, we estimated relative proportions (with 95% CIs and P_{trend} ; unadjusted and adjusted by recruiting center and region) to evaluate whether the number of positive-to-negative discordant results (between the local and national laboratories) increased as the size of the clusters increased.

While conducting the analysis of potential technical concerns, we identified some clear technical problems related to the cluster size that appeared to be due to contamination at one laboratory. Therefore, *a posteriori*, we decided to present the results of potential technical concerns before the results of the reproducibility as this had such a substantial impact on these results and their interpretation. We elaborate on this in detail in the Discussion.

All statistical analyses were done using R (11).

Sample size

Assuming 20% positivity of hrHPV (12), 0.05 type I error, and a null hypothesis of a kappa less than 0.4, this study was at least 90% powered to detect a kappa value greater than 0.7 between each healthcare level center (6 in total) and the NRL with 1,000 cervical samples (approximately 167 per each recruiting center). However, due to a protocol deviation in DES, which consisted of samples that were tested later than 30 days after collection, or others that were missing aliquot retests at ReH or NRL (see Fig. 1), 257 additional women were recruited in that region after corrective measures were put in place. In total, this study included 1,134 participants whose HPV tests were tested locally and aliquots retested at higher healthcare levels.

Ethical considerations

The study obtained ethical approval from the Tanzania National Institute for Medical Research (DES) and the Ethical Review Committee of the WHO (Geneva, Switzerland).

Results

Study population

Figure 1 summarizes the study design and study population. In total, 1,134 30- to 50-year-old women (657 in DES, 477 in Kilimanjaro) were included in the analysis. Recruitment lasted 2 to 3 months in each center and started in DES in July 2015 followed by Kilimanjaro in June 2016. The mean age was 39.0 years (SD, 6.0) and it was similar between regions (38.9 for DES, 39.4 for Kilimanjaro) and centers (36.8, 39.1, and 40.2, respectively, for PHC, DiH, and ReH centers in DES, and 40.8, 37.7, and 39.8 for centers in Kilimanjaro). Across both regions, around 80% were married (88%, 77%, and 78%, respectively, for PHC, DiH, and ReH centers in DES, and 82%, 88%, and 78%, respectively, for centers in Kilimanjaro). A minority of women (≤10%) had not completed primary education at any of the clinics. Conversely, 18%, 21%, and 25%, respectively, for centers in DES, and 10%, 15%, and 14%, respectively for centers in Kilimanjaro had completed secondary or higher education.

Intralaboratory reproducibility

The intralaboratory reproducibility analyses (In-Run, Run-to-Run, Operator-to-Operator, and Day-to-Day) are summarized in Supplementary Table S3. Fleiss' kappa coefficients (95% CI) were "very good" (10) for all comparisons: 0.87 (0.61–1.00); 0.87 (0.61–1.00); 0.93 (0.68–1.00); and 1.00 (0.75–1.00), respectively. Only two disagreements were observed for the In-Run and Run-to-Run experiments each, and one disagreement for the Operator-to-Operator experiment.

HPV positivity

Local hrHPV positivity ranged from 13.5% in women recruited at the Kilimanjaro ReH to 49.6% in women recruited at the DES ReH (Table 1). The local HPV positivity in the remaining recruiting centers was similar and ranged between 18.2% for the DES PHC and 23.5% for the Kilimanjaro DiH. From DES, the test positivity at the NRL was between 17.2% for PHC aliquots and 22.2% for those from the ReH. From Kilimanjaro, the NRL-retest HPV positivity was between 12.9% for Kilimanjaro ReH aliquots, and a higher than expected 31.5% for the Kilimanjaro PHC aliquots.

Table 1. Positivity (% and 95% CI) of hrHPV by recruiting centers, testing laboratory, and region.

Recruiting center	DES (n = 657)			Kilimanjaro (n = 477)		
	Local result	ReH result	NRL result	Local result	ReH result	NRL result
PHC	18.2 (13.2–24.1)	39.7 (33.0–46.7)	17.2 (12.4–23.0)	23.1 (16.4–30.9)	21.0 (14.6–28.6)	31.5 (24.0–39.8)
DiH	20.5 (15.1–26.8)	35.5 (28.9–42.6)	20.5 (15.1–26.8)	23.5 (17.3–30.7)	20.7 (14.8–27.7)	22.9 (16.7–30.0)
Regional/reference hospital (ReH)	49.6 (43.2–56.0)	—	22.2 (17.2–27.9)	13.5 (8.8–19.6)	—	12.9 (8.3–18.9)

Abbreviation: —, not applicable.

Cluster size analysis

For recruiting centers that had very different HPV positivity at the different testing centers in the testing sequence, we investigated whether there was potential contamination by examining the cluster size within testing microplates.

The median cluster size for the HPV positive results was 2 or lower for the majority of testing centers (Supplementary Table S4) except for samples tested in the DES ReH (median cluster size 16, IQR 3–22) and aliquots retested in the same center (median cluster size 7, IQR 2–48). Relative proportions of a positive-to-negative discordant result (positive at the recruiting center and negative at the NRL) are presented in **Table 2**. The positive-to-negative discordant proportion was 14.1% for isolated positive results and ranged from 12.5% for cluster size 2 to 78.9% for cluster size greater or equal to 10, with a significant increasing trend ($P_{\text{trend}} < 0.001$). All except two of the clusters of 6 or more corresponded to samples tested at the DES ReH. Interestingly, if we excluded the results of possibly compromised samples from the DES ReH, the P_{trend} for clusters 1 to ≥ 4 was 0.079.

Interobserver reproducibility between recruiting centers and NRL

Agreement between recruiting centers and the NRL is shown in **Table 3**. The highest agreement was for the DES DiH (kappa 0.91, 95% CI, 0.84–0.98) and PHC (kappa 0.90, 95% CI, 0.82–0.98) followed by the Kilimanjaro ReH (0.82, 95% CI, 0.69–0.95) and DiH (kappa 0.78, 95% CI, 0.66–0.89). The agreement was significantly lower for samples collected at two sites: the DES ReH

(kappa 0.34, 95% CI, 0.24–0.43) and the Kilimanjaro PHC (kappa 0.41, 95% CI, 0.24–0.57). The overall agreement, excluding possibly compromised samples from these latter two centers was 95.7% (95% CI, 94.0–96.9) with a kappa of 0.86 (95% CI, 0.81–0.91), a “very good agreement” kappa coefficient.

Interobserver reproducibility between recruiting centers and ReHs

Table 4 shows the agreement of the HPV results tested by the recruiting PHC and DiH centers and retested by the corresponding ReH. The highest disagreement was observed for DES. In line with previous results on cluster size (Supplementary Table S4) and HPV positivity (**Table 1**), samples tested at the ReH in this region seemed to be compromised due to technical issues, which led to lower agreement [kappa indexes 0.48 (95% CI, 0.37–0.60) and 0.59 (95% CI, 0.47–0.71) for PHC and DiH, respectively]. Similarly and probably also due to technical issues at the Kilimanjaro PHC, despite the similar HPV positivity between samples tested locally (23.1%) and retested at the ReH (20.7%) in this region (**Table 1**), there was low agreement (kappa 0.41, 95% CI, 0.23–0.59). Conversely, agreement between Kilimanjaro DiH and ReH was very good (kappa 0.82, 95% CI, 0.72–0.93).

Interobserver reproducibility between ReHs and NRL

Table 5 shows the interobserver reproducibility of the retested aliquot HPV results from each ReH and the NRL of the cervical samples collected at the PHC and DiH centers. As described previously, there were some technical concerns with the retesting performed

Table 2. Positive-to-negative discordance between the recruiting center and the NRL, by cluster size.

Cluster size at the recruiting center	Total local HPV positive results n	HPV retesting result NRL		RP ^a (95% CI)	RP ^b (95% CI)
		Positive n (%)	Negative n (%)		
Isolated	78	67 (85.9)	11 (14.1)	Reference	Reference
2	72	63 (87.5)	9 (12.5)	0.89 (0.38–2.02)	0.82 (0.35–1.87)
3	35	27 (77.1)	8 (22.9)	1.62 (0.68–3.66)	1.62 (0.69–3.60)
4–5	19	12 (63.2)	7 (36.8)	2.61 (1.10–5.79)	2.13 (0.88–4.89)
6–7 ^c	13	5 (38.5)	8 (61.5)	4.36 (2.09–8.92)	3.09 (1.36–6.91)
8–9 ^d	8	3 (37.5)	5 (62.5)	4.43 (1.81–9.36)	2.74 (1.07–6.36)
$\geq 10^d$	71	15 (21.1)	56 (78.9)	5.59 (3.38–10.5)	3.23 (1.75–6.85)
Total	296	192 (64.9)	104 (35.1)	<0.001 ^e	<0.001 ^e

Note: HPV by careHPV test.

Abbreviation: RP, relative proportion.

^aUnadjusted RP of positive-to-negative results.

^bRP adjusted by recruiting center and region.

^cSix out of 13 from ReH in DES.

^dAll samples (except one from PHC in Kilimanjaro with cluster size 8) belong to ReH in DES.

^e P_{trend} .

Table 3. Interobserver reproducibility of HPV results tested locally at the recruiting centers and retested at the NRL, by region.

	Number of samples	HPV retesting result NRL		Agreement (95% CI)	Kappa (95% CI)
		Positive <i>n</i>	Negative <i>n</i>		
DES					
Local HPV testing result at PHC					
Positive	209	34	4	97.1 (93.9–98.7)	0.90 (0.82–0.98)
Negative		2	169		
Local HPV testing result at DiH					
Positive	200	38	3	97.0 (93.6–98.6)	0.91 (0.84–0.98)
Negative		3	156		
Local HPV testing result at ReH					
Positive	248	48	75	66.9 (60.9–72.5)	0.34 (0.24–0.43)
Negative		7	118		
Kilimanjaro					
Local HPV testing result at PHC					
Positive	143	22	11	76.2 (68.6–82.5)	0.41 (0.24–0.57)
Negative		23	87		
Local HPV testing result at DiH					
Positive	164	31	7	92.1 (86.9–95.3)	0.78 (0.66–0.89)
Negative		6	120		
Local HPV testing result at ReH					
Positive	170	19	4	95.9 (91.7–98.0)	0.82 (0.69–0.95)
Negative		3	144		
Total ^a					
Local HPV testing result					
Positive	743	122	18	95.7 (94.0–96.9)	0.86 (0.81–0.91)
Negative		14	589		

Note: HPV by *careHPV* test.

^aExcluding compromised samples from ReH in DES and PHC in Kilimanjaro.

at the DES ReH on aliquots from PHC and DiH (0.44, 95% CI, 0.32–0.055; and 0.57, 95% CI, 0.45–0.68; respectively) and aliquots from the Kilimanjaro PHC that were retested at ReH and NRL (kappa 0.66, 95% CI, 0.53–0.80). The highest agreement for aliquots retested at ReH and NRL was found for samples collected at the Kilimanjaro DiH (kappa 0.80, 95% CI, 0.68–0.91).

Discussion

To our knowledge, this is the first study to measure the reproducibility of the *careHPV* test at three different healthcare levels in an urban (DES) and a rural (Kilimanjaro) region of a country by retesting aliquots at regional and national levels. Cervical samples from 1,134 women aged 30 to 50 were tested using *careHPV* locally by three

Table 4. Interobserver reproducibility of HPV results tested locally at the recruiting centers and retested at the ReH, by region.

	Number of samples	HPV retesting result ReH		Agreement (95% CI)	Kappa (95% CI)
		Positive <i>n</i>	Negative <i>n</i>		
DES					
Local HPV testing result at PHC					
Positive	209	37	1	77.5 (71.4–82.6)	0.48 (0.37–0.60)
Negative		46	125		
Local HPV testing result at DiH					
Positive	200	39	2	83.0 (77.2–87.6)	0.59 (0.47–0.71)
Negative		32	127		
Kilimanjaro					
Local HPV testing result at PHC					
Positive	143	17	16	79.7 (72.4–85.5)	0.41 (0.23–0.59)
Negative		13	97		
Local HPV testing result at DiH					
Positive	164	31	7	93.9 (89.1–96.7)	0.82 (0.72–0.93)
Negative		3	123		

Note: HPV by *careHPV* test.

Table 5. Interobserver reproducibility of HPV results retested at the ReH and retested at the NRL, by region.

	Number of samples	NRL HPV retesting result		Agreement (95% CI)	Kappa (95% CI)
		Positive <i>n</i>	Negative <i>n</i>		
DES					
ReH HPV retesting result collected at local PHC					
Positive	209	34	49	75.6 (69.3–80.9)	0.44 (0.32–0.55)
Negative		2	124		
ReH HPV retesting result collected at local DiH					
Positive	200	38	33	82.0 (76.1–86.7)	0.57 (0.45–0.68)
Negative		3	126		
Kilimanjaro					
ReH HPV retesting result collected at local PHC					
Positive	143	28	2	86.7 (80.2–91.3)	0.66 (0.53–0.80)
Negative		17	96		
ReH HPV retesting result collected at local DiH					
Positive	164	30	4	93.3 (88.4–96.2)	0.80 (0.69–0.91)
Negative		7	123		

Note: HPV by *careHPV* test.

recruitment centers at primary, district, and regional healthcare level, then retested by the corresponding ReH and by the NRL.

The positivity of hrHPV in this study was around 20% in most centers, which is similar to other studies in Tanzania and neighboring regions (12–15). However, there were two different technical issues that we experienced in this study. First, for samples collected by the Kilimanjaro PHC, the HPV positivity was very similar locally (23.1%) and at the ReH (21.0%) but much higher at the national laboratory (31.5%). Given that the NRL results were similar for other recruiting centers, it is most likely that a contamination issue was related to these specific PHC specimens rather than to the processes/operation at the NRL, for example, a potential issue with testing at the Kilimanjaro PHC, aliquoting or with transportation of specimens. It is worth mentioning that technical/logistical circumstances were very challenging at this PHC. The certified laboratory assistant technician had very limited experience with more complicated laboratory manipulations such as pipetting, as the lab only performed basic low tech testing. There were frequent electricity power interruptions, resulting in several spoiled test runs that needed to be repeated, and important challenges to store the samples at a constant 2°C to 8°C. As such, although part of the discordant results may have been related to our study design (re-aliquoting and transport of residual specimens for retesting), it seems likely that the failure of the testing at this center is mostly attributable to the local technical/logistical capacity of this lab. There was also a clear contamination problem during testing at the DES ReH, which was evidenced by both the far higher HPV positivity when testing was done at this laboratory and also by the extremely large clusters of positive specimens on testing microplates. This contamination issue was identified locally and improvements in the standard operating procedures were made to ensure that operators adhered to the standard laboratory protocols. As a result, there was a clear improvement in later results from this lab evidenced in the kappa increasing from 0.27 (95% CI, 0.16–0.37) in the set of 169 samples tested in 2015 to 0.55 (95% CI, 0.35–0.75) in the set of 79 samples tested in 2016. The technical issue at the DES ReH is more concerning than the former (at the Kilimanjaro PHC) as this is a regional reference laboratory with good infrastructure and the testing was performed by one of the original *careHPV* trainers hinting to an individual operator-dependent issue. This

highlights the importance of maintaining good training and laboratory protocols to ensure that the testing is performed in line with manufacturer's guidelines. In addition, it would be advisable to implement quality assurance (QA) processes to ensure that potential technical problems with the testing are identified as soon as possible. Beylerian and colleagues have previously proposed an approach to identify microplates with suspicion of contamination based on a higher than expected HPV positivity, number of clusters per microplate and/or maximum size of cluster per microplate (16). Laboratory refresher training was not part of the study protocol, as only 3 batches would be tested per clinic. In our study, contamination was apparent as some microplates were almost entirely HPV positive; however, in less clear situations, the approach proposed by Beylerian and colleagues may offer a useful tool for laboratories to use as QA. When we exclude the laboratories with technical issues, we demonstrate very good agreement between testing performed at the local level and the regional/national level in both rural and urban settings, which is in agreement with other studies (17–19). To our knowledge, there are no standardized QA guidelines for HPV testing in general. The WHO *careHPV* prequalification public report refers to the manufacturer's instructions in the *careHPV* Test Kit Handbook (20). Conversely, the Programme for Appropriate Technologies in Health (PATH) have published a detailed Guidelines for QA in programs introducing *careHPV* in low-resource settings, including operator training and proficiency, infrastructural requirements, quality control and quality assessment, and postmarketing surveillance to follow performance of the test in the intended setting (21).

This study has two key strengths. First, we had sufficient sample size to consider HPV test agreement at three different levels of laboratory/healthcare systems and in two disparate urban and rural regions of Tanzania. This provided us with the opportunity to demonstrate that this test could be introduced in multiple different settings. Second, we were able to evaluate both the interobserver and intralaboratory reproducibility to assess different features, which could affect the outcomes of the results. We have compared intralevel reproducibility between operators, runs and testing performed on different days, analyses that were performed following the resolution of contamination issues.

There are also some limitations to discuss. We have already discussed in detail the contamination issues in this study. We have demonstrated that contamination can be an issue even in sites with good quality equipment and training (i.e., at the urban regional laboratory), but can be avoided at any healthcare level if appropriate training and QA protocols are established, as we showed for most other sites.

We should also discuss briefly some limitations of the *careHPV* test; (i) test kits should be stored at temperatures 4°C to 25°C requiring a cold chain (including importation and transport) with a maximum shelf life of 1 year, which is challenging for many settings; (ii) the test is designed to run in a 90 samples batch mode. Few clinics in low-resource settings can achieve a sufficient screening volume which precludes a “same-day screen-and-treat” approach; (iii) samples can be kept at room temperature up to 30°C for 2 weeks, or 2°C to 8°C up to 30 days, requiring reliable refrigeration conditions in the lab for any samples held over two weeks; (iv) the sensitivity of this test has been shown to be lower than other validated HPV-DNA tests (22); (v) *careHPV* has been shown to be unsuitable for self-sampling (7). Therefore, ideally, this test should be introduced at a site that can assure adequate technical conditions to test batches of 90 samples within 2 weeks, avoiding the need for refrigeration, and also maintaining the skills of the lab technician by performing the test regularly. The number of samples can be collected either at the clinic involved, or at other surrounding health centers and transported to this lab.

In conclusion, we identified potential testing issues including contamination that could perhaps dictate the decision to implement this testing platform at centers with a set of minimal technical requirements, robust protocols, adequate training and QA procedures, within the context of an organized screening program where all steps of the screening and treatment process are adequately monitored. However, we have demonstrated that the *careHPV* test is reproducible at different levels of healthcare settings if these potential technical requirements can be met. This highlights that strengthening of laboratory systems is a key component to ensure that cervical cancer prevention and control programs are effective and of sufficient quality to achieve the goals of the cervical cancer elimination initiative.

Disclosure of Potential Conflicts of Interest

D. Meshier reports other from GlaxoSmithKline [the Blood Safety, Hepatitis, Sexually Transmitted Infections (STI) and HIV Service at Public Health England has

provided GlaxoSmithKline with postmarketing surveillance reports on HPV infections; a cost recovery charge is made for these reports] outside the submitted work. No potential conflicts of interest were disclosed by the other authors.

Disclaimer

Where authors are identified as personnel of the IARC/WHO, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the IARC/WHO.

Authors' Contributions

A. Baena: Data curation, formal analysis, investigation, methodology, writing—original draft, writing—review and editing. **H. De Vuyst:** Conceptualization, data curation, formal analysis, supervision, investigation, methodology, writing—original draft, writing—review and editing, field work. **D. Meshier:** Data curation, formal analysis, validation, investigation, methodology, writing—original draft, project administration, writing—review and editing. **M. Kasubi:** Conceptualization, supervision, validation, investigation, methodology, writing—original draft, project administration, writing—review and editing. **S. Yuma:** Conceptualization, supervision, investigation, visualization, methodology, writing—review and editing, field work. **J. Mwaiselage:** Validation, investigation, field work. **S. Zouiuich:** Data curation, formal analysis, investigation, methodology, writing—original draft. **P. Mlay:** Validation, investigation, field work. **C. Kahesa:** Validation, investigation, field work. **S. Landoulsi:** Validation, investigation. **M.L. Hernandez:** Validation, investigation, project administration, field work. **E. Lucas:** Data curation, software, investigation. **R. Herrero:** Conceptualization, resources, formal analysis, supervision, validation, investigation, methodology, writing—review and editing. **M. Almonte:** Resources, validation, investigation, methodology, writing—review and editing. **N. Broutet:** Conceptualization, resources, supervision, funding acquisition, investigation, methodology, writing—review and editing.

Acknowledgments

The HPV test kits used for this study were kindly provided by the *careHPV* manufacturer to the Tanzania Ministry of Health. For A. Baena, this article was undertaken during the tenure of a Postdoctoral Fellowship from the IARC. The study was entirely funded by the UNDP (United Nations Development Program)/UNFPA (United Nations Population Fund)/UNICEF/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction (HRP).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received March 27, 2020; revised June 23, 2020; accepted August 14, 2020; published first August 20, 2020.

References

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424.
- de Martel C, Georges D, Bray F, Ferlay J, Clifford GM. Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis. *Lancet Glob Health* 2020;8:e180–90.
- Arbyn M, Ronco G, Anttila A, Meijer CJ, Poljak M, Ogilvie G, et al. Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. *Vaccine* 2012;30:F88–99.
- Xue P, Gao LL, Yin J, Han LL, Zhao J, Li L, et al. A direct comparison of four high-risk human papillomavirus tests versus the cobas test: detecting CIN2+ in low-resource settings. *J Med Virol* 2019;91:1342–50.
- Labani S, Asthana S, Sodhani P, Gupta S, Bhambhani S, Pooja B, et al. CareHPV cervical cancer screening demonstration in a rural population of north India. *Eur J Obstet Gynecol Reprod Biol* 2014;176:75–9.
- Zhao FH, Jeronimo J, Qiao YL, Schweizer J, Chen W, Valdez M, et al. An evaluation of novel, lower-cost molecular screening tests for human papillomavirus in rural China. *Cancer Prev Res* 2013;6:938–48.
- Jeronimo J, Bansil P, Lim J, Peck R, Paul P, Amador JJ, et al. A multicountry evaluation of *careHPV* testing, visual inspection with acetic acid, and papanicolaou testing for the detection of cervical cancer. *Int J Gynecol Cancer* 2014;24:576–85.
- Tuerxun G, Yukeisaier A, Lu L, Aierken K, Mijiti P, Jiang Y, et al. Evaluation of *careHPV*, cervista human papillomavirus, and hybrid capture 2 methods in diagnosing cervical intraepithelial neoplasia grade 2+ in Xinjiang Uyghur women. *Oncologist* 2016;21:825–31.
- Kelly H, Mayaud P, Segondy M, Pant Pai N, Peeling RW. A systematic review and meta-analysis of studies evaluating the performance of point-of-care tests for human papillomavirus screening. *Sex Transm Infect* 2017;93:S36–S45.
- Altman DG. *Practical statistics for medical research*. London: Chapman and Hall; 1991.
- R Core Team. *R: a language and environment for statistical computing*. Vienna (Austria): R Foundation for Statistical Computing; 2020. Available from: <https://www.R-project.org/>.
- Dartell M, Rasch V, Kahesa C, Mwaiselage J, Ngoma T, Junge J, et al. Human papillomavirus prevalence and type distribution in 3603 HIV-positive and

- HIV-negative women in the general population of Tanzania: the PROTECT study. *Sex Transm Dis* 2012;39:201–8.
13. Womack SD, Chirenje ZM, Gaffikin L, Blumenthal PD, McGrath JA, Chipato T, et al. HPV-based cervical cancer screening in a population at high risk for HIV infection. *Int J Cancer* 2000;85:206–10.
 14. Kahesa C, Kjaer SK, Ngoma T, Mwaiselage J, Dartell M, Iftner T, et al. Risk factors for VIA positivity and determinants of screening attendances in Dar es Salaam, Tanzania. *BMC Public Health* 2012;12:1055.
 15. De Vuyst H, Parisi MR, Karani A, Mandaliya K, Muchiri L, Vaccarella S, et al. The prevalence of human papillomavirus infection in Mombasa, Kenya. *Cancer Causes Control* 2010;21:2309–13.
 16. Beylerian EN, Slavkovsky RC, Holme FM, Jeronimo JA. Statistical modeling for quality assurance of human papillomavirus DNA batch testing. *J Low Genit Tract Dis* 2018;22:219–24.
 17. Obiri-Yeboah D, Adu-Sarkodie Y, Djigma F, Akakpo K, Aniakwa-Bonsu E, Amoako-Sakyi D, et al. Options in human papillomavirus (HPV) detection for cervical cancer screening: comparison between full genotyping and a rapid qualitative HPV-DNA assay in Ghana. *Gynecol Oncol Res Pract* 2017;4:5.
 18. Lin CQ, Chen F, Liu B, Zhang YZ, Cui XL, Li AM, et al. A parallel study of careHPV and Hybrid Capture 2 human papillomavirus DNA testing for cervical cancer screening in rural China. *J Virol Methods* 2014;202:73–8.
 19. Ying H, Jing F, Fanghui Z, Youlin Q, Yali H. High-risk HPV nucleic acid detection kit—the careHPV test—a new detection method for screening. *Sci Rep* 2014;4:4704.
 20. World Health Organization. WHO prequalification of in vitro diagnostics: public report. Product: careHPV™ Test. Geneva (Switzerland): WHO; 2018 Jul. WHO reference No.: PQDx 0085-028-00. Available from: https://www.who.int/diagnostics_laboratory/evaluations/pq-list/180713_pqpr_pqdx_0085_028_00_carehpv_with_labelling.pdf?ua=1.
 21. PATH. Guidelines for organizing a quality assurance program for introduction and routine use of careHPV in low-resource settings. Seattle (WA): PATH; 2017.
 22. Gage JC, Ajenifuja KO, Wentzensen N, Adepiti AC, Stoler M, Eder PS, et al. Effectiveness of a simple rapid human papillomavirus DNA test in rural Nigeria. *Int J Cancer* 2012;131:2903–9.