



Evaluation of commercial HPV assays in the context of post-treatment follow-up: Scottish Test of Cure Study (STOCS-H)

H A Cubie,¹ M Canham,¹ C Moore,² J Pedraza,¹ C Graham,³ K Cuschieri²

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/jclinpath-2013-202014>).

¹HPV Research Group, University of Edinburgh MRC Centre for Reproductive Health, Queen's Medical Research Institute, Edinburgh, UK
²Scottish HPV Reference Laboratory, Royal Infirmary of Edinburgh, Edinburgh, UK
³Wellcome Trust Clinical Research Facility, University of Edinburgh, Western General Hospital, Edinburgh, UK

Correspondence to

Professor Heather A Cubie, HPV Research Group, MRC Centre for Reproductive Health, Queen's Medical Research Institute, University of Edinburgh, 47 Little France Cres, Edinburgh EH16 4TJ, UK; Heather.Cubie@ed.ac.uk

Received 23 October 2013
Revised 7 December 2013
Accepted 15 December 2013
Published Online First
16 January 2014

ABSTRACT

Aims Human papillomavirus (HPV) testing is more sensitive than cytology for detection of residual/recurrent cervical disease after lesion treatment. Several HPV test comparison studies have been performed within triage and screening populations, but data on their comparative performance in a test of cure context is lacking. This study aims to address this gap.

Methods We compared the technical and clinical performance of Abbott RealTime High risk (HR)-HPV, Genprobe Aptima PV, Hologic Cervista HPV-HR, Qiagen Hybrid Capture 2 and Roche cobas HPV in the Early Implementation phase of a 'test of cure' service within the Scottish Cervical Screening Programme.

Results Valid results with all five HPV Tests from 1020 first samples taken ~6 months post-treatment showed HPV positivity ranging from 17.84% to 26.96%. There was perfect agreement in 74%, and greatest variation between assays was observed in cytologically negative samples. Clinical performance was judged on cumulative incidence of cervical intraepithelial neoplasia 2+ (CIN2+) during follow-up (mean: 13.2 months). There were 23 cases of CIN2+ of which 14 were CIN3+. All assays, including cytology, were 100% sensitive for detection of CIN3+. Of the nine cases of residual CIN2, three assays detected all, one assay missed one and one assay missed two cases. Specificity ranged from 75% to 84% according to assay.

Conclusions All assays were sensitive for detection of CIN2+ at 6 months post-treatment. The range of positivity equated to a 50% increase between assays with the lowest and highest positivity rates. The relevance of HPV positivity in the absence of cytological abnormalities requires longer follow-up to determine whether additional tools for risk stratification are required.

INTRODUCTION

Human papillomavirus (HPV) testing as a test of cure (TOC) of treatment of cervical intraepithelial neoplasia (CIN) is based on significant evidence of improved sensitivity for the detection of residual/recurrent disease (85–97%) according to the recent meta-analysis of Arbyn and colleagues.¹ HPV testing also confers a higher negative predictive value compared to cytology resulting in less protracted follow-up. According to Kocken and colleagues,² one HPV negative test at 6 months post-treatment was associated with a 10-year risk of CIN3+ of 2.1%, and if performed as a cotest with cytology, this risk reduced to 1.4% if both were negative.

It is vitally important that any HPV assay which influences patient management is validated for

clinical use. Satisfying the non-inferiority criteria of Meijer *et al*³ has been the method of clinical validation for several assays, yet these criteria were designed, a priori, to assess the performance of HPV tests for screening of women over 30 years of age. The extent of infection and associated disease will vary according to clinical context, as will the priority for performance in terms of sensitivity and specificity. Given the increasing choice of available HPV assays,⁴ the question of which HPV test is optimal for a TOC setting remains outstanding, as most multiplatform, or head-to-head studies have been performed in triage or colposcopy referral populations.^{5,6}

To address this, we undertook a performance evaluation of commercial HPV assays in the post-treatment setting. The main objectives were to compare the technical and clinical performance of five HPV consensus assays, the *digene* HC2 (Qiagen NV, Venlo, The Netherlands), Abbott RealTime High risk (HR)-HPV (Abbott Molecular, Des Plaines, Illinois, USA), cobas 4800 HPV (Roche Molecular Diagnostics, Pleasanton, California, USA) APTIMA HPV assay (formerly GenProbe, San Diego, California, USA, now Hologic) and Cervista HPV HR (Hologic, Bedford, Massachusetts, USA).

MATERIALS AND METHODS

Sample set and algorithm

TOC is currently performed as part of the National Cervical Screening Programme in Scotland.⁷ This study used samples collected during the early implementation phase (2011–2012; figure 1), when women were sampled in primary care for HPV testing and cytology at two follow-up visits approximately 6 and 12 months after treatment. Identification of these women and their clinical outcomes was possible using the web-based national cervical screening management system, Scottish Cervical Call and Recall System (SCCRS). HPV positivity and/or cytology of mild dyskaryosis, or worse, at either follow-up visit led to colposcopic referral. For this evaluation, all five assays were performed on the 6-month sample and Hybrid Capture 2 (HC2) at a cut-off of 1 RLU/CO was the standard of care. Women who were referred back to colposcopy had biopsies taken only if clinically indicated. Histology results from these biopsies were entered into SCCRS and were made available for the study following permission from each centre.

Samples were tested retrospectively by the other four HPV assays, so non-HC2 results did not influence clinical management. Clinical performance

To cite: Cubie HA, Canham M, Moore C, *et al*. *J Clin Pathol* 2014;**67**:458–463.

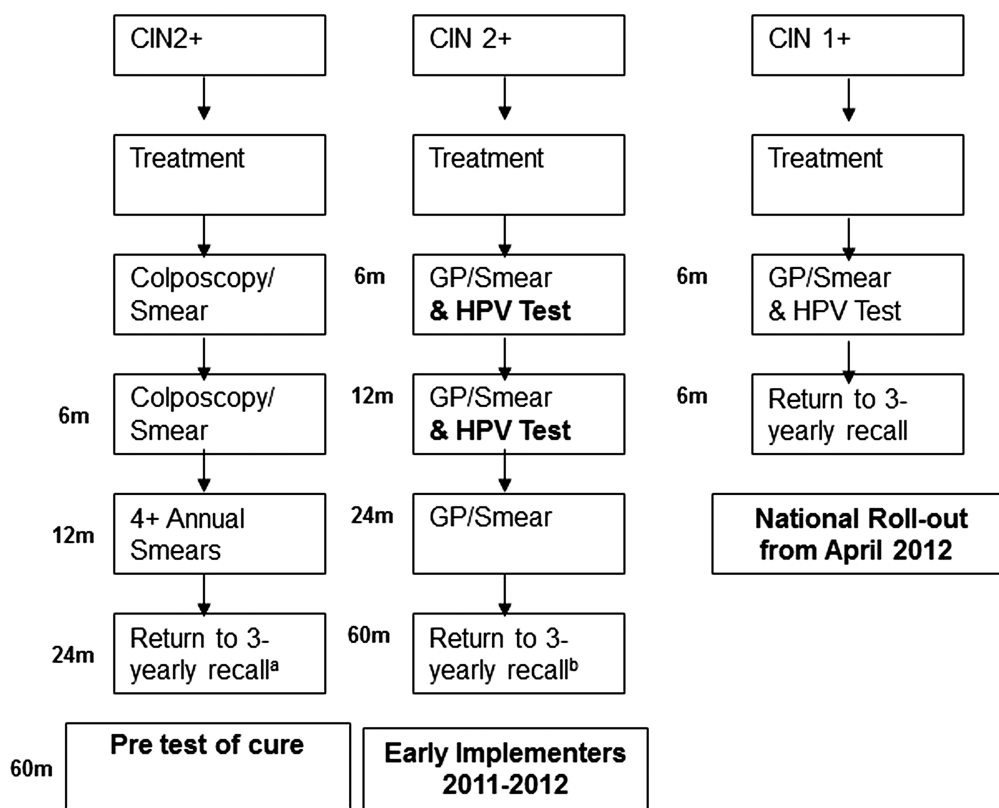


Figure 1 Schematic of test of cure development and implementation in Scotland. Left hand image shows status quo before test of cure; middle image shows early implementation phase which involved approximately 30% of the Scottish population, and the right hand image shows process in place for national roll-out. (A) relates to pre-test of cure with return to routine recall restricted to samples where no abnormality was detected on preceding smears. (B) If preceding smears HPV negative and abnormality no worse than borderline change in Squamous cells.

was judged on cumulative incidence of CIN2+ or worse during the follow-up period (on average, 13.2 months with a range of 7.2–19.8 months). A total of 1020 first (6-month) samples which gave valid results with all five HPV tests were available.

HPV testing

Automated platforms for additional HPV screening tests were available either within the Directorate of Laboratory Medicine, Royal Infirmary of Edinburgh, or the HPV Research Group, Division of Pathology, University of Edinburgh. In addition to the HC2 which was performed using the manual procedure, the other assays assessed were the RealTime HR-HPV assay (Abbott), the Aptima HPV test (Hologic), the Cervista HPV Test (Hologic) and the cobas 4800 HPV test (Roche), all used according to the manufacturers' instructions. For HC2, an RLU/cut-off of 1.0 was used for referral to colposcopy. For the purpose of this study, results were analysed at a cut-off of 1.0 and 2.0, as the latter cut-off has been suggested to enhance specificity⁸ and is used in England for HPV triage and TOC.

Statistical analyses

Prevalence of HPV in the 6-month post-treatment sample was assessed for all assays with 95% CIs and compared to the prevalence of cytological abnormality at a threshold of mild dyskaryosis or worse. The number of discordant samples according to HPV assay was also quantified and compared, as was the number of infections identified by a single assay only.

Clinical performance was measured as sensitivity, specificity, negative predictive value (NPV) and positive predictive value (PPV) for the cumulative detection of CIN2+ over the

follow-up period. Given that not all women were referred to colposcopy, clinical performance was also measured according to concurrent cytology at 6 months with disease, considered to be a cytological result of moderate dyskaryosis or worse.

Proportional agreement and McNemar test were performed for each test relative to histology (CIN2+) and cytology separately. Additionally, HPV 16 and/or 18 status (cobas and *rt*HPV assays only) was correlated with histology outcomes.

RESULTS

Prevalence of HPV at 6 months according to assay and discordance

One thousand and twenty samples were available for analysis. HPV positivity ranged from 18% to 27% depending on assay (figure 2). Using HC2 at a cut-off of one RLU/CO as the analytical standard for comparison, statistically significant differences between the assays were observed, with the Aptima HPV (AHPV) having fewer positives ($p < 0.0001$) and the COBAS and Cervista assays having a greater number of positives ($p = 0.0001$ and $p = 0.0009$, respectively). No significant difference in positivity between the *rt*HPV and the HC2 was observed at a cut-off of 1 ($p = 0.2008$). At a cut-off of 2, HC2 had significantly fewer positives ($p < 0.0001$) compared to all assays, with the exception of AHPV ($p = 0.00304$).

Perfect agreement (concordance of all 5 assays) was found in 753 (74%) samples, with 136 and 617 samples positive and negative for all five tests, respectively. The number of singleton positives (ie, where a positive test was generated by 1 assay only) ranged from 4 for AHPV; 11 for *rt*HPV; 11 for HC2; 28 for cobas and 80 for Cervista. At a cut-off of two RLU/CO,

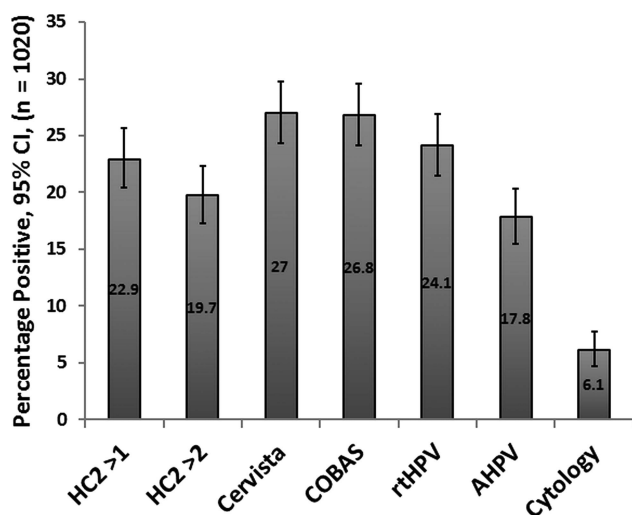


Figure 2 HPV positivity according to assay at first visit (~6 months) post-treatment. Cytology positive is set at mild dyskaryosis or worse.

HC2 was associated with only five singleton positives (data not shown). A table which shows further details of concordance where samples are stratified into those positive for ≥ 1 test at manufacturers' cut-off is available as online supplementary data.

HPV positivity relative to underlying cytology

Of the 1020 samples, 902 (88.4%) were cytologically negative, 53 (5.2%) were classed as having borderline squamous abnormalities, 32 (3.1%) showed mild dyskaryosis, and 33 (3.2%) showed moderate dyskaryosis or worse. As expected, the number of cytological abnormalities was lower than the number of HPV infections. Follow-up histology revealed that of the 33 with high-grade cytological abnormalities, 16 were associated with CIN2+, 10 were associated with CIN1, three had a negative biopsy and four had no biopsy/treatment. Table 1 shows HPV prevalence according to cytology grade stratified into the following categories (Negative or borderline; Mild; Moderate/Severe). Positivity according to assay, within the negative or borderline category, was wide (13.6%–23.4%). More consistency was observed in the smaller mild and moderate/severe groups, (78.8% and 84.8%, respectively) reflecting a difference between 26 and 28 infections only.

Clinical performance relative to cytology

During follow-up, 33 women had cytology which was moderate or worse (21 cases of moderate dyskaryosis, 11 cases of severe dyskaryosis and 1 case of glandular abnormality). Sensitivity for detecting moderate dyskaryosis or worse ranged from 79% (61, 91) for two of the assays (HC2 at a cut-off of 2 and Cervista) to 85% for COBAS, Aptima, rtHPV and HC2 at a cut-off of 1, reflecting seven and five 'missed' cases, respectively. Table 2 shows clinical performance parameters according to high-grade cytology. For all HPV assays, McNemars test confirmed differences in the distribution of discordant results between HPV and cytology.

Clinical performance of HPV tests and cytology relative to histology

In the total cohort of 1020 women, 23 had residual CIN2+ at any time during the follow-up period, of which 14 were CIN3+ including one cancer. At manufacturers' cut-off, all tests were over 90% sensitive for the detection of CIN2+, with three

Table 1 HPV positivity in 1020 samples stratified by cytology grade and HPV assay

Cytology	HC2: c/o 1.0 n pos/total % (95% CI)	HC2: c/o 2.0 n pos/total % (95% CI)	rtHPV n pos/total % (95% CI)	AHPV n pos/total % (95% CI)	Cervista n pos/total % (95% CI)	COBAS n pos/total % (95% CI)	Positive by any test % (95% CI)	Number of CIN2+
Negative or borderline	181/955 19.0 (16.5 to 21.6)	150/955 15.7 (13.5 to 18.2)	194/955 20.3 (17.8 to 23.0)	130/955 13.6 (11.5 to 16.0)	223/955 23.4 (20.7 to 26.2)	220/955 23.0 (20.4 to 25.8)	345/955 36.1 (33.1 to 39.3)	4
Mild	25/32 78.1 (60.0 to 90.7)	25/32 78.1 (60.0 to 90.7)	24/32 75.0 (56.6 to 88.5)	24/32 75.0 (56.6 to 88.5)	26/32 81.2 (63.6 to 92.8)	25/32 78.1 (60.0 to 90.7)	28/32 87.5 (71.0 to 96.5)	3
Moderate or severe	28/33 84.8 (68.1 to 94.9)	26/33 78.8 (61.1 to 91.0)	28/33 84.8 (68.1 to 94.9)	28/33 84.8 (68.1 to 94.9)	26/33 78.8 (61.1 to 91.0)	28/33 84.8 (68.1 to 94.9)	30/33 90.1 (75.7 to 98.1)	16
Total	234/1020 22.9 (20.4 to 25.6)	201/1020 19.7 (17.3 to 22.3)	246/1020 24.1 (21.5 to 26.9)	182/1020 17.8 (15.5 to 20.3)	275/1020 27.0 (24.3 to 29.8)	273/1020 26.8 (24.1 to 29.6)	403/1020 39.5 (36.5 to 42.6)	23

AHPV, Aptima HPV Test; CIN, cervical intraepithelial neoplasia; c/o, cut off; HC2, hybrid capture 2 test; rtHPV, real time HPV test.

Table 2 Clinical performance of assays according to detection of high grade cytology (defined as moderate dyskaryosis or worse)

	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Proportional Agreement* (95% CI)
HC2 at <i>c/o</i> 1	85 (68 to 95)	79 (76 to 82)	12 (8 to 17)	99 (99 to 100)	79 (77 to 82)
HC2 at <i>c/o</i> 2	79 (61 to 91)	82 (80 to 85)	13 (9 to 18)	99 (98 to 100)	82 (80 to 84)
Cervista	79 (61 to 91)	75 (72 to 77)	9 (6 to 14)	99 (98 to 100)	75 (72 to 78)
AHPV	85 (68 to 95)	84 (82 to 87)	15 (10 to 21)	99 (99 to 100)	84 (82 to 87)
<i>rt</i> HPV	85 (68 to 95)	78 (75 to 80)	11 (8 to 16)	99 (98 to 100)	78 (75 to 81)
COBAS	85 (68 to 95)	75 (72 to 78)	10 (7 to 14)	99 (98 to 100)	75 (73 to 78)

*At a threshold of moderate dyskaryosis or worse.

AHPV, aptima HPV test; *c/o*, cut off; HC2, hybrid capture 2 test; NPV, negative predictive value; PPV, positive predictive value; *rt*HPV, real time HPV test.

HPV assays detecting all CIN2+ (HC2, *rt*HPV and COBAS; table 3). The Cervista and AHPV assays were negative for one and two cases of CIN2, respectively. Assay performance relative to CIN2 and CIN3 was not performed separately given the relatively small numbers of disease cases. Using a threshold of mild dyskaryosis which would have triggered a referral to colposcopy independently of the HPV result, cytology did not detect four cases of CIN2, whereas with a threshold of borderline, this reduced to three cases. No HPV assay, nor cytology (at \geq mild) missed any case of CIN3. Although sensitivity was high, the PPV for CIN2+ of all HPV assays was relatively low and ranged from 8% to 12% depending on assay.

Detailed assessment of residual/recurrent disease (CIN2+)

Table 4 provides details on all 23 cases of CIN2+ detected in this population, annotated according to HPV status and cytology. Of the 14 cases of CIN3+, 11 were positive for HPV 16 by either or both the *rt*HPV and the COBAS. The two cases of CIN2 that were negative for at least one HPV test were associated with 'other' HPV types, and had negative cytology.

Clinical performance of HPV 16/18 testing

HPV 16 and/or 18 infection could be assessed with the *rt*HPV and COBAS assays. COBAS and *rt*HPV detected 98 and 85 infections, respectively, in the whole population of 1020 samples. The sensitivity of HPV 16 and/or 18 detection (excluding other types) for identification of CIN2+ was 61% (39 to 80) and 65% (43 to 84) for the *rt*HPV and COBAS, respectively. The specificity was over 90% for both assays (93% (91 to 94) for *rt*HPV and 92% (90 to 93) for COBAS) with PPVs of 16.5 (9.6 to 26.4) and 15.3 (9.1 to 24.3), respectively.

DISCUSSION

The meta analyses of Arbyn and colleagues^{1 9} provided evidence to show that after treatment of CIN, HPV testing detected residual and recurrent high-grade CIN more quickly, with higher sensitivity than follow-up cytology. Additionally, the systematic review by Chan *et al*¹⁰ showed that in a post-treatment population, more than 90% of residual CIN2 was detected by HPV testing compared with 76% for cytology. Although these analyses have incorporated large datasets, the conclusions are drawn largely on the performance of the HC2 assay and this is one of the first studies to assess the comparative performance of several HPV assays within the context of a TOC service.

We acknowledge certain limitations to the study. First, HC2 is advantaged, in that this was the only assay used for colposcopic referral, and it is feasible that some of the 'HC2 negative/other test positive' were associated with underlying CIN2+. Second, we assumed that there was no disease in cases that were HPV negative/cytology negative, and we did not account for verification bias nor take blind biopsies. Furthermore, the number of residual disease cases (23) which were used to judge clinical performance was relatively low. Finally, follow-up relates to a relatively short period (average 13.2 months) and it will be important to review the data at a later time point. Conversely, an advantage of the study is that it represents a real-life population undergoing TOC as part of a routine programme, and provides data on comparative test performance in an under-researched setting.

In Scotland, all women are tested with HPV and cytology at 6 months concurrently, and the cytology status is not known at the time of the HPV test. This enables us to assess how HPV testing would perform as a stand-alone test. Prevalence of HPV at 6 months post-treatment varied according to assay, and it is

Table 3 Clinical performance of HPV assays for detection of CIN2+

	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Proportional agreement* (95% CI)
HC2 at <i>c/o</i> 1	100 (lower bound 85)	79 (76 to 81)	10 (6 to 14)	100 (lower bound 100)	79 (77 to 82)
HC2 at <i>c/o</i> 2	96 (78 to 100)	82 (80 to 84)	11 (7 to 16)	100 (99 to 100)	82 (80 to 85)
Cervista	96 (78 to 100)	75 (72 to 77)	8 (5 to 12)	100 (99 to 100)	75 (72 to 78)
AHPV	91 (72 to 99)	84 (81 to 86)	12 (7 to 17)	100 (99 to 100)	84 (81 to 86)
<i>rt</i> HPV	100 (lower bound 85)	78 (75 to 80)	9 (6 to 14)	100 (lower bound 100)	78 (75 to 81)
COBAS	100 (lower bound 85)	75 (72 to 78)	8 (5 to 12)	100 (lower bound 100)	75 (73 to 78)
Cytology†	83 (61 to 95)	95 (94 to 97)	29 (19 to 42)	100 (99 to 100)	95 (94 to 96)

*Relative to histology (CIN2+).

†At a threshold of mild dyskaryosis or worse.

Detailed clinical information relating to cytology referral categories is available as online supplementary tables.

AHPV, aptima HPV test; CIN, cervical intraepithelial neoplasia; *c/o*, cut off; HC2, hybrid capture 2 test; HPV, human papillomavirus; NPV, negative predictive value; PPV, positive predictive value; *rt*HPV, real time HPV test.

Table 4 Details of all cases of CIN2+ detected during follow-up of 1020 patients who had received treatment for cervical disease

Case number	HC2 RLU/CO	COBAS (with typing)	rtHPV (with typing)	Cervista (HPV species)	Aptima (with value)	Cytology	Histology
1	777.61	16	16	A9	10.79	Severe dyskaryosis	CIN 3
2	242.09	16	16	A9	10.22	Mild dyskaryosis	CIN 3
3	1095.42	Other HR-HPV	Other HR-HPV	A5/6	7.41	Moderate dyskaryosis	CIN 3
4	6.04	Other HR-HPV	Other HR-HPV	A5/6, A9	4.85	Borderline squamous changes	CIN 2
5	2777.21	16,Other HR-HPV	16, Other HR-HPV	A7, A9	19.02	Severe dyskaryosis	CIN 2
6	327.67	Other HR-HPV	Other HR-HPV	A9	16.87	Moderate dyskaryosis	CIN 2
7	273.43	16	16	A9	10.16	Severe dyskaryosis	CIN 3
8	61.89	16	16	A9	10.07	Moderate dyskaryosis	CIN 3
9	247.1	16	16	A9	10.46	Severe dyskaryosis	CIN 3
10	2.37	Other HR-HPV	Other HR-HPV	Negative	Negative	Negative	CIN 2
11	894.48	Other HR-HPV	Other HR-HPV	A9	9.6	Mild dyskaryosis	CIN 3
12	180.4	Other HR-HPV	Other HR-HPV	A9	17.49	Moderate dyskaryosis	CIN 3
13	2209.85	16	16	A5/6, A9	9.62	Mild dyskaryosis	CIN 2
14	2762.86	16	16	A9	12.4	Moderate dyskaryosis	CIN 3
15	2224.68	16,Other HR-HPV	Other HR-HPV	A5/6, A9	20.06	Moderate dyskaryosis	CIN 2
16	132.39	16	16	A9	11.8	Moderate dyskaryosis	CIN 3
17	13.79	16	16	A9	11.73	Severe dyskaryosis	CIN 2
18	964.15	16	16	A5/6, A9	10.53	Severe dyskaryosis	CIN 3
19	445.86	Other HR-HPV	Other HR-HPV	A5/6	7.98	Negative	CIN 2
20	4.39	16,Other HR-HPV	16, Other HR-HPV	A7	18.74	Severe dyskaryosis	CIN 3
21	539.15	16,Other HR-HPV	16, Other HR-HPV	A9	11.24	Severe dyskaryosis	CIN 3
22	1.85	Other HR-HPV	Other HR-HPV	A5/6	Negative	Negative	CIN 2
23	5.87	16	16	A9	10.27	Severe dyskaryosis	Invasive SCC

CIN, cervical intraepithelial neoplasia; HC2, hybrid capture 2 test; HPV, human papillomavirus; HR-HPV, high risk HPV; rtHPV, real time HPV test; RLU/CO, relative light unit/cut off.

notable that when considering the assays associated with the lowest and the highest prevalence, colposcopy referrals would differ by ~50%. The AHPV showed a significantly lower prevalence compared to all other assays at manufacturer's cut-off, other than when HC2 cut-off of two was assessed. The observation that the AHPV assay showed the highest specificity reconciles with other multitest comparisons in triage and primary screening contexts such as the Predictors 2 and 3 studies.^{1 5 6} The particularly high number of singleton positives with the Cervista assay gives cause for concern. These were largely attributable to samples with normal cytology where each of the three HPV channels gives a reading just above cut-off. We believe a reassessment of cut-off is warranted, and a manuscript to this effect is in preparation.

As women treated for CIN2+ are at greater risk of developing high-grade lesions for at least 10 years and possibly up to 20 years after treatment,^{11 12} the priority for clinical performance is sensitivity, and it is reassuring that all assays were 100% sensitive for the detection of residual CIN3+ at the 6-month follow-up visit. Small differences were observed at the level of CIN2+, with the Cervista assay missing one case of CIN2+ and the AHPV assay missing two cases. The HC2 at a cut-off of two RLU/CO, as used in England,^{13 14} also missed one of the cases of CIN2 which was also negative by AHPV. These differences did not translate into a statistically significant difference in sensitivity between assays. Post-treatment cytology at a threshold of mild had a higher PPV than any of the HPV tests, but would have missed the greatest number of CIN2 cases in this series (n=4).

We acknowledge that in other TOC settings, cytology is performed in advance of a HPV test, and women with ≥moderate dyskaryosis would be referred to colposcopy directly. We therefore stratified the data to assess the performance of HPV assays by cytology category. The range of assay positivity was largest in

the normal and borderline group, and ranged from 13.6% to 23.4%, compared to 75–81.2% and 78.8–84.5% in the mild and moderate or severe categories, respectively (see online supplementary data showing detailed results stratified for normal and borderline; normal, borderline and mild; and moderate and severe categories). As described above, HPV testing detected four additional cases of CIN2 (3 associated with negative cytology, 1 with borderline cytology). Although no case of CIN3 was missed by cytology at mild dyskaryosis or worse, a missed CIN2 in the TOC population may arguably have more clinical significance than a CIN2 missed in a primary screening context.

Based on these data, it is therefore challenging to determine the added value of the cytology cotest at 6 months post-treatment, and this aligns with the conclusions of Arbyn *et al.*¹ However, although the sensitivity of HPV testing is confirmed by this project, the specificity and PPV for CIN2+ was low for all tests, and consistently lower than cytology. Striking a balance between optimal sensitivity and unnecessary referrals is challenging and enhanced tools which aid risk stratification of clinically significant HPV positives would be welcome. While typing for HPV 16/18 appeared to confer greater specificity compared to consensus detection, the sensitivity of this approach at 61–65% would be unacceptable.

In addition to clinical performance, choice of test may also be influenced by cost, physical space requirements, liquid handling capability, existence of managed equipment contracts and throughput in addition to differences in conventional comparators such as sensitivity and specificity. Long-term follow up and associated linkage to pathology records will demonstrate over time where the balance in terms of colposcopy referrals should rest. The relevance of HPV positivity in the absence of cytological abnormalities also requires longer follow-up to

determine whether additional tools for risk stratification are required in a post-treatment context.

Take home messages

- ▶ Positivity ranged between 18% and 27% in five widely used commercial human papillomavirus (HPV) tests in the context of post-treatment follow-up.
- ▶ Extrapolation of these results in a national programme would result in considerable variation in referrals to colposcopy, depending on assay used.
- ▶ All cases of CIN3+ within 13 months post-treatment were detected at 6 months by all five HPV tests.
- ▶ A sensitive HPV assay could be considered as a stand-alone test in the post-treatment setting to detect residual disease.

Acknowledgements We would like to acknowledge the considerable support from our cytopathology colleagues: Drs Ted Duvall, Louise Smart and Tim Palmer with their respective cytology managers: Fiona McQueen, Aileen Milne and Dennis Walker, technical assistance from Alison Fleming and Chris Ward and administrative support from Evelyn Wallace. The contribution of HPV testing kits from Abbott, GenProbe, Hologic and Roche to allow this study to take place is gratefully acknowledged. This is, however, an investigator-led study, and the companies have had no say in the production of this manuscript.

Contributors HAC conceived and directed the study. HAC and KC were responsible for the manuscript although all authors contributed to the writing/editing process; KC and CM coordinated transfer of residual samples and were responsible for some of the HPV testing and clinical data collection; MC and JP were responsible for archiving samples, HPV data collection and some of the HPV testing; CG was responsible for statistical analysis.

Competing interests HAC and KC have periodically received funding from each of the companies whose products were involved in the study, to attend and present at scientific meetings, and occasionally at expert panel discussions for these companies. The Institutions have received free or discounted test kits and associated equipment on loan for a number of studies.

Ethics approval Samples were HPV tested for clinical management and linked to clinical data prior to the study. The South-East Scotland Research Ethics Committee (SESREC) operates a proportionate ethical review service and considered the project to have no material ethical issues as samples only had additional HPV tests. In addition, these residual LBC samples are stored in the Scottish HPV Archive and come under the generic approval of the archive as a Research Tissue Bank (ref 11/AL/0174) for HPV related research on anonymised archive samples.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The companies that provided kits have requested that the manuscript is shared with them, but they do not have control of content. They have also asked for their individual results and data to be released to them and have accepted that this would happen after the article is accepted.

REFERENCES

- 1 Arbyn M, Ronco G, Anttila A, *et al.* Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. *Vaccine* 2012;30S:F88–99.
- 2 Kocken M, Helmerhorst TJM, Berkhof J, *et al.* Risk of recurrent high-grade cervical intraepithelial neoplasia after successful treatment: a long-term multi-cohort study. *Lancet Oncol* 2011;12:441–5.
- 3 Meijer C, Berkhof H, Heideman DAM, *et al.* Validation of high-risk HPV tests for primary cervical screening. *J Clin Virol* 2009;53:S1–4.
- 4 Cubie HA, Cuschieri K. Understanding HPV tests and their appropriate applications. *Cytopathology* 2013.doi:10.1111/cyt.12083. [Epub ahead of print].
- 5 Szarewski A, Ambroisine L, Cadman L, *et al.* Comparison of predictors for high-grade cervical intraepithelial neoplasia in women with abnormal smears. *Cancer Epidemiol Biomarkers Prev* 2008;17:3033–42.
- 6 Szarewski A, Mesher D, Cadman L, *et al.* Comparison of seven tests for high-grade cervical intraepithelial neoplasia in women with abnormal smears: the Predictors 2 study. *J Clin Microbiol* 2012;50:1867–73.
- 7 Scottish Cervical Screening Programme. Cervical Screening-changes to screening test result categories. <http://www.nsd.scot.nhs.uk/services/screening/cervicalscreening/> (accessed 4 Oct 2013).
- 8 Sargent A, Bailey A, Turner A, *et al.* Optimal threshold for a positive hybrid capture 2 test for detection of human papillomavirus: data from the ARTISTIC trial. *J Clin Microbiol* 2010;48:554–8.
- 9 Arbyn M, Paraskevidis E, Martin-Hirsch P, *et al.* Clinical utility of HPV–DNA detection: triage of minor cervical lesions, follow-up of women treated for high-grade CIN: an update of pooled evidence. *Gynecologic Oncology* 2005;99: S7–11.
- 10 Chan BKS, Melnikow J, Slee CA, *et al.* Posttreatment human papillomavirus testing for recurrent cervical intraepithelial neoplasia: a systematic review. *Am J Obstet Gynecol* 2009;200:422.e1–422.e9.
- 11 Kyrgiou M, Koliopoulos G, Martin-Hirsch P, *et al.* Management of minor cervical cytological abnormalities: a systematic review and a meta-analysis of the literature. *Cancer Treat Rev* 2007;33:514–20.
- 12 Soutter WP, Sasieni P, Panoskaltsis T. Long-term risk of invasive cervical cancer after treatment of squamous cervical intraepithelial neoplasia. *Int J Cancer* 2005;118:2048–55.
- 13 NHS Cancer Screening programmes. HPV triage and test of cure in the cervical screening programme in England. <http://www.cancerscreening.nhs.uk/cervical/hpv-triage-test-of-cure.html> (accessed 4 Oct 2013).
- 14 Legood R, Smith M, Lew J-B, *et al.* Cost effectiveness of human papillomavirus test of cure after treatment for cervical intraepithelial neoplasia in England: economic analysis from NHS Sentinel Sites Study. *BMJ* 2012;345:e7086–96.

Supplementary Tables

Supplementary Table 1. Concordance of HPV positivity according to 5 separate HPV tests at manufacturers' cut-off. Shaded cells indicate a positive result.

	HC2	Cervista	COBAS	rHPV	AHPV	Frequency	Percent
						617	60.49
						4	0.39
						11	1.08
						2	0.2
						28	2.75
						23	2.25
						4	0.39
						80	7.84
						1	0.1
						1	0.1
						4	0.39
						2	0.2
						3	0.29
						6	0.59
						11	1.08
						1	0.1
						2	0.2
						1	0.1
						4	0.39
						1	0.1
						24	2.35
						12	1.18
						9	0.88
						6	0.59
						1	0.1
						1	0.1
						5	0.49
						20	1.96
						136	13.33

Supplementary Tables 2 a, b and c

Clinical performance of HPV assays for detection of CIN2+, by cytology referral category

A. Normal and borderline (n= 952 samples)

	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)	Proportional Agreement % (95% CI)
HC2 at c/o 1	100 (lower bound 3)	81 (76,81)	1 (0,3)	100 (lower bound 100)	81 (79, 84)
HC2 at c/o 2	75 (19,99)	85 (82,87)	2 (0,6)	100 (99,100)	85 (82,87)
Cervista	75 (19,99)	77 (74,80)	1 (0,4)	100 (99,100)	77 (74,79)
AHPV	50 (7,93)	87(84,89)	2 (0,5)	100 (99,100)	86 (84,88)
rHPV	100 (lower bound 40)	80 (77,83)	2 (1,5)	100 (lower bound 100)	80 (77,83)
COBAS	100 (lower bound 40)	77 (74,80)	2 (0,5)	100 (99,100)	77 (75,80)

B. Normal/borderline and mild (n=987)

	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)	Proportional Agreement % (95% CI)
HC2 at c/o 1	100 (lower bound 59)	80 (77,82)	3 (1,7)	100 (lower bound 100)	80 (77, 82)
HC2 at c/o 2	86 (42,100)	83 (80,85)	3 (1,7)	100 (99,100)	83 (80,85)
Cervista	86 (42,100)	75 (72,78)	2 (1,5)	100 (99,100)	75 (72,78)
AHPV	71 (29,96)	85(82,87)	3 (1,7)	100 (99,100)	85(82,87)
rHPV	100 (lower bound 59)	78 (76,81)	3 (1,7)	100 (lower bound 100)	79 (76,81)
COBAS	100 (lower bound 59)	76 (73,78)	3 (1,6)	100 (lower bound 100)	77 (75,80)

C. Moderate/severe (n= 33)

	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)	Proportional Agreement % (95% CI)
HC2 at c/o 1	100 (lower bound 79)	29 (10,56)	57 (37,76)	100 (lower bound 48)	64 (45,80)

HC2 at c/o 2	100 (lower bound 79)	41 (18,67)	62 (41,80)	100 (lower bound 59)	70 (51,84)
Cervista	100 (lower bound 79)	41 (18,67)	62 (41,80)	100 (lower bound 59)	70 (51,84)
AHPV	100 (lower bound 79)	29 (10,56)	57 (37,76)	100 (lower bound 48)	64 (45,80)
rtHPV	100 (lower bound 79)	29 (10,56)	57 (37,76)	100 (lower bound 48)	64 (45,80)
COBAS	100 (lower bound 79)	29 (10,56)	57 (37,76)	100 (lower bound 48)	64 (45,80)