A Population-Based Clinical Trial Comparing Endocervical High-Risk HPV Testing Using Hybrid Capture 2 and Cervista From the SHENCCAST II Study

Jerome L. Belinson, MD,1,5 Ruifang Wu, MD,2 Suzanne E. Belinson, PhD,1 Xinfeng Qu, MD,3 Bin Yang, MD, PhD,4 Hui Du, MD,2 Ruosong Wu, MD,3 Chun Wang, MD,2 Lijie Zhang, MD,2 Yanqiu Zhou, MD,2 Ying Liu, MD,2 and Robert G. Pretorius, MD6

Key Words: Hybrid Capture 2; Cervista; Cervical cancer screening; Human papillomavirus; HPV

Abstract

Our objective was to directly compare the accuracy of the high-risk human papillomavirus (HPV) assays, Hybrid Capture 2 (hc2; Qiagen, Gaithersburg, MD) and Cervista (Hologic, Bedford, MA), in diagnosing cervical intraepithelial neoplasia (CIN) 3 or worse (cancer).

A population-based, cross-sectional study (The Shenzhen Cervical Cancer Screening Trial II) was conducted in Guangdong Province in China. Three high-risk HPV assays, self and direct cervical sampling and cytology, were studied. Abnormal results on any of 6 study tests (33%) resulted in referral to colposcopy. At colposcopy, every patient had at least 5 cervical biopsy specimens obtained.

For 8,556 women between the ages of 25 and 59 years (mean, 38.9 years), the rate for CIN 3 or worse was 1.6% (141/8,556). The sensitivity (confidence interval) values for CIN 3 or worse were 97.9% (94.0%-99.6%) and 95.1% (90.0%-98.0%) for hc2 and Cervista, respectively (P > .05). The specificity (confidence interval) values were 87.8% (87.1%-88.5%) and 90.3% (89.6%-90.9%), respectively (P < .05).

Differences in accuracy in diagnosing CIN 3 or worse with the hc2 and Cervista tests are minor and result from the decisions made in selecting the cut points.
with data from historical hc2 and PCR-based HR-HPV assays concluded that the rate of positive results of the Cervista assay (18%) was 4 times as high as that of the composite hc2 and PCR-based assays (4%), suggesting that the Cervista assay was significantly less specific than the hc2.10

We recently completed a population-based cervical cancer screening study (the Shenzhen Cervical Cancer Screening Trial II [SHENCCAST II]) within which screened women had endocervical specimens tested with the Cervista and hc2 assays. This article describes the overall conduct of the SHENCCAST II with the focus on analyzing the subset of the data that allow direct comparison of the accuracy of Cervista and hc2 in diagnosing CIN 3 or cancer.

Materials and Methods

Participant Enrollment

Between April 2009 and April 2010, a multisite, population-based, cross-sectional study (the SHENCCAST II) was carried out in 7 sites in Guangdong Province in China to evaluate a new self-sampling device (the Preventive Oncology International/National Institutes of Health [POI/NIH] self-sampler), a new cytology imaging system (I2 Imager, Hologic, Santa Clara, CA), and 2 new HPV testing technologies (Cervista and MALDI-TOF [PCR-based mass array matrix-assisted laser desorption/ionization time-of-flight mass spectrometry system]) applied to self-sampling and direct endocervical sampling. The human subject review boards of the Cleveland Clinic (Cleveland, OH) and the Peking University Shenzhen Hospital (Shenzhen, China) approved the protocol, and all procedures followed were in accord with the ethical standards established by these institutions and are in accord with the Helsinki Declaration of 1975. The women were recruited in the urban site (Shenzhen) through a mass media campaign. In the suburban sites (Longgang and Bao’an), women were recruited through government notification or media campaign. In the rural sites (Wushi, Feng’an, Longwo, and Heping), the women were recruited through government notification or community meetings. In the rural sites (Wushi, Feng’an, Longwo, and Heping), the women were all recruited by the local health clinics under government authority.

Women were eligible if they were 25 to 59 years old, not pregnant, and had no cervical cancer screening for at least 3 years, no hysterectomy, and no pelvic radiation.

Study Visit

All willing and eligible women, after being educated on the goals and specific conduct of the study, signed an informed consent document agreeing to participate. Each woman had a one-on-one interview with a trained female interviewer for collection of demographic and medical history information.

The women were then block randomized by the day of screening to self-sample with the POI/NIH sampler or the conical-shaped brush (Qiagen). Education in self-sampling was by video and personal instruction. Following self-sampling, a physician placed a vaginal speculum and the direct endocervical sample was obtained using the “broom” sampler (Rovers Medical Devices, Oss, the Netherlands).

Specimen Preparation

Self-sampling specimens were placed in 20-mL vials of PreservCyt (Hologic, Marlborough, MA) for testing for HR types of the HPV by the Cervista assay and by MALDI-TOF.

Physician-collected endocervical specimens were placed in 20-mL vials of PreservCyt Liquid. The 20 mL was then processed for cytology and the HPV assays in the following order: Preparation of ThinPrep slide (Hologic, Marlborough, MA; number of milliliters variable), hc2 (4 mL), Cervista (2 mL), and MALDI-TOF (1 mL).

The ThinPrep slides were prepared and stained by research staff trained by Hologic, according to the Hologic protocol designed for the I2 imager. Slides were read using the imager for computer-assisted diagnosis per the Hologic protocol.

The hc2 is an in vitro nucleic acid hybridization assay with signal amplification using microplate chemiluminescence for the qualitative detection of 13 HR types of HPV DNA, in aggregate, in cervical specimens and was performed per the manufacturer’s product insert.

The Cervista HPV HR assay is an in vitro diagnostic test for the qualitative detection of 14 HR-HPV types (the same 13 types as hc2 plus type 66). It uses Invader chemistry (Hologic, Madison, WI), a signal-amplification method for the detection of nucleic acid sequences. There are 2 types of isothermal reactions that occur simultaneously, a primary on the targeted DNA sequence, and a secondary, which is a reaction that produces fluorescence. The assay was performed according to the manufacturer’s product insert by the research laboratory staff trained in the procedure by Hologic.

The PCR-based MALDI-TOF assay is a mass spectrometry method that uses a multiplex primary PCR, with several HPV primers (GP5+/6+) that target type-specific base pairs in the L1 region of the HPV genome for 14 HPV types, followed by a mass extension reaction with a single primer of distinct mass that is also specific for each genotype. If present in the specimen, each specific type is reported by the software.11

The computer algorithm and hardware were from Sequenom (San Diego, CA) with modification by the scientists at the Beijing Genomics Institute (Shenzhen, China).

All specimens were received, triaged, and processed for cytology, hc2, and Cervista at the Royal Ladies Clinic, the POI research center in Shenzhen, China. The MALDI-TOF assay was done by the staff of the Beijing Genomics Institute, Shenzhen.
Women with positive results on any 1 of the 5 HPV assays (Cervista self, MALDI-TOF self, hc2 direct, Cervista direct, or MALDI-TOF direct) or with ASCUS or worse by cytology were contacted and asked to return for colposcopy. All women who returned for colposcopy were evaluated using the POI micro biopsy protocol of directed and random biopsies.2

At colposcopy, the cervix was visually divided into 4 quadrants by lines drawn from 12 to 6 o’clock and from 9 to 3 o’clock. Each quadrant of the cervix was examined independently. All abnormal areas were biopsied by quadrant. Quadrants with normal colposcopic impressions had 1 “random” biopsy sample obtained at the squamocolumnar junction at the 2, 4, 8, or 10 o’clock position depending on the quadrant. Endocervical curettage was then performed. All cervical biopsy specimens were collected with our standard 2-mm POI biopsy instrument, which allows rapid healing of the biopsy sites and minimizes patient discomfort. Histologic slides were interpreted by a gynecologic pathologist from the Peking University Shenzhen Hospital (C.W.) and adjudicated by a gynecologic pathologist from the Cleveland Clinic (B.Y.). The adjudication consisted of reviewing all high-grade abnormal slides (CIN 2 or worse) and 31% of the CIN 1 slides.

A team of 5 data managers controlled the original source data, which were independently entered into 2 specially designed Access databases (Microsoft, Seattle, WA) for the SHENCAST II project. On completion of the study and data entry, the 2 independent databases were transferred to the POI Epidemiology and Biostatistical Center, Chicago, IL, for adjudication, cleaning, and statistical analysis.

This report focuses on a paired sample comparison of endocervical hc2 and Cervista assays for HR-HPV.

Statistical Methods

The performance characteristics of the screening tests were evaluated by calculating the sensitivity, specificity, and positive and negative predictive values and areas under the receiver operating characteristic curve (ROC) according to the standard definitions for CIN 2 or worse and CIN 3 or cancer. The McNemar test was used to compare the sensitivities and specificities of the tests.

All women in the study received the reference standard (cervical biopsy) or had negative results of HPV testing and cytologic examination, which has been shown to predict histologically negative results.2 Because of this, the aforementioned estimates were calculated directly in this study. Biopsy data were collapsed across participants to generate variables using the highest grade biopsy result in any quadrant as the final histologic diagnosis. All confidence intervals (CIs) are exact binomial confidence intervals. The ROC curves12 were drawn using the published cutoffs for positivity for each test, for CIN 2 or worse and CIN 3 or cancer. The partial areas under the curves with specificity between 85% and 100% were compared with the Z-test. In addition, an analysis exploring test characteristics at different cut points for the Cervista and hc2 was undertaken. These results were then plotted. All data analyses were performed using STATA 10.0 (StataCorp, College Station, TX).

Results

For the study, 10,000 women were screened and, therefore, entered into the SHENCAST II. There were 8,556 women who had all prescribed screening and diagnostic procedures with no missing data, and their data are included in this analysis. The data for 1,444 women were dropped from the analysis because of missing data as follows (some missing more than 1 test): 196, cytology data; 3, hc2 direct test data; 264, Cervista direct test data; 537, MALDI-TOF direct test data; 106, Cervista self-test data; and 159, MALDI-TOF self-test data; in addition, 625 did not return for colposcopy. Table 1 shows the rate of return for colposcopy for each abnormal screening test.

The mean age was 38.9 years, 2.2% of the women were smokers, and 94.3% reported they were married. The cytologic result was ASCUS or worse in 12.1% (1,031/8,556) of the women. The overall HPV positivity rates were 13.6% (CI, 12.9%-14.3%) and 11.1% (CI, 10.4%-11.8%) for endocervical hc2 and Cervista tests, respectively (P < .000001). The HPV positivity rates in women with normal cytologic results were 7.9% (594/7,525) and 6.0% (453/7,525) for the hc2 and Cervista tests, respectively (P < .000001). Similarly, the respective values for ASCUS cytologic results were 41.4% (256/618) and 34.0% (210/618) for the hc2 and Cervista tests, respectively (P < .000001). The overall rates for CIN 2 or worse and CIN 3 or worse were 2.7% (233/8,556) and 1.6% (141/8,556) respectively.

The sensitivity rates for CIN 3 or cancer were 97.9% (CI, 94.0%-99.6%) and 95.1% (CI, 90.0%-98.0%) for the hc2 and Cervista tests, respectively (P > .05; not statistically significant). The specificity rates were 87.8% (CI, 87.1%-88.5%) and 90.3% (CI, 89.6%-90.9%), respectively (P < .05).
The ROC curves with their respective cut points for these data are shown in Figure 1.

Cut-point analysis revealed the following results: (1) If the sensitivity of the hc2 test for CIN 3 or cancer were adjusted to be equal to that of the Cervista test at 95%, the specificity of the hc2 would calculate to be approximately 90.5% (range, 90.46%-90.51%). There are actually a few cut points in this range that will give approximately 95.0% sensitivity. This value (90.5%) is not different from the 90.3% specificity of the Cervista; therefore, the Cervista cut point is on the hc2 ROC curve. (2) If the sensitivity of the Cervista for CIN 3 or cancer were adjusted to be equal to that of the hc2 at 97.9%, the specificity of Cervista would calculate to be approximately 35% (range, 30.9%-39.8%, with the variation due to the range of viable cut points). This value for the Cervista (35%) is clearly different from the specificity of hc2 of 87.8%. Therefore, the ROC curves for the Cervista and hc2 diverge beyond the established Cervista cut point. (3) The hc2 at a cut point of about 3.2 pg (range, 3.21-3.29 pg) would equal the Cervista cut point. We compared HR-HPV testing by the Cervista with testing by the hc2 using CIN 2 or worse as the end point and found results similar to those reported for CIN 3 or cancer (data not shown).

In Figure 2, the ROC curves comparing the cut points for the hc2 test of 1, 3, 5, and 10 pg of HPV DNA are shown in blue to gray, and the cut point for the Cervista test is shown in red. As the prior analysis suggested, the cut point of Cervista lies on the hc2 ROC curve at a point similar to hc2 with a cut point of 3.2 pg of HPV DNA. The ROC curves for hc2 and Cervista are virtually identical until the cut point of 3.2 pg of HPV DNA is reached. For cut points lower than 3.2 pg of HPV DNA, the hc2 ROC curve is higher than that of Cervista. This difference in ROC curves is shown by a significantly higher partial area under the curve: hc2, 0.1164; Cervista, 0.1101; \( P = .03 \).

Discussion

In the development of any diagnostic assay, there is always a determination to be made as to where to place the positive/negative cut point. This decision is based on finding...
Belinson et al / Accuracy of HC-II vs Cervista

the favorable balance between clinical sensitivity and specificity. This clinical trial of paired specimens demonstrated a nonsignificant difference between the sensitivities of the hc2 and Cervista HR-HPV assays for the detection of CIN 3 or cancer. It also demonstrated a significant difference (favoring Cervista) in specificity. This difference may be due in part to the documented lack of cross-reactivity with low-risk or non-oncogenic HPV types with the Cervista assay that is known to occur with hc2.12 However, we believe it is primarily because the cut point for the hc2 (1 pg of HPV DNA) is lower than that for the Cervista (~3.2 pg of HPV DNA). As noted in the analysis, when the cut point of the hc2 is increased to about 3.2 pg of HPV DNA, the sensitivity and specificity of the hc2 are the same as those of the Cervista. On the other hand, if the cut point were moved so the sensitivity of the Cervista is 97.9% and equal to the hc2, the specificity falls significantly to about 35%. This means that at a cut point less than 3.2 pg of HPV DNA, the 2 assays perform quite differently. This conclusion is shown again in Figure 1, in which the cut point of Cervista seems to be on the hc2 ROC curve at a cut point of about 3 pg of HPV DNA. At that point, the curves begin their major diversion, creating the difference noted in the comparative calculation of the areas under the curves.

The clinical question is whether an assay (Cervista, in this case) that acts like the hc2 with a cut point of about 3.2 pg of HPV DNA is significantly different from the hc2 with a cut point of 1 pg of HPV DNA. It is likely that the 2.8% lower sensitivity for CIN 3 or cancer and the 2.5% higher specificity for CIN 3 or cancer of the Cervista are not different enough to matter in clinical practice. This thesis is not new, and one can refer to the work of Cuzick et al,13 who reached the conclusion in a population 34 years or older that the sensitivity of the hc2 for CIN 2 or worse with a cut point of 1 pg of HPV DNA did not significantly differ from that with a cut point of 4 pg of HPV DNA, whereas the specificity for CIN 2 or worse of the hc2 with a cut point of 1 pg of HPV DNA was significantly lower than that with a cut point of 4 pg of HPV DNA. Our group reached a similar conclusion in 2003 for a cut point of 3 pg HPV DNA.14

The strengths of this study are that it is a population-based, paired study of the 2 assays being evaluated using the same sample. In addition, CIN 3 or cancer was used as the study end point, which is a more reproducible diagnosis than CIN 2. The study also benefited from a very high referral rate to return, or they could not be located.

These data can certainly be applied to other populations. Although there are differences in the distribution of HR-HPV types and differences in the proportion of cervical cancer associated with a particular HPV type among different populations, there are far more similarities. In addition, as noted, only HPV-66 separates the type coverage for the 2 assays. Worldwide HPV types 16 and 18 are found in 68% of the invasive cervical cancers, HPV-66 is present in 0.4% of cancers and 1.9% of high-grade lesions.15 In the study by Chen et al16 of 630 invasive cervical cancers in China, HPV-16 and HPV-18 accounted for 84.5% of the cervical cancers, and HPV-66 was present in 1% of cases. In the SHENCAST II data, HPV-66 was present in 1.8% of high-grade lesions. These data from our large, paired-sample, population-based study are very different from the findings reported by Kinney et al10 of a study that compared the Cervista package insert with historical hc2 and PCR-based HR-HPV assays. It seems to us that the differences in accuracy in diagnosing CIN 3 or cancer between HR-HPV testing with the Cervista and with the hc2 are minor and result from the decisions made in selecting the cut points. The choice as to which test a laboratory should choose will come down to price, ease of use, and rate for specimen throughput.


