Validation of western Helicobacter pylori IgG antibody assays in Korean adults

Sun-Young Lee,1 Hee-Won Moon,2 Mina Hur2 and Yeo-Min Yun2

1Department of Internal Medicine, Konkuk University School of Medicine, Seoul, Korea
2Department of Laboratory Medicine, Konkuk University School of Medicine, Seoul, Korea

Helicobacter pylori infection is endemic in Korea, and serology testing is widely performed. The aim of this study was to validate and compare the diagnostic accuracy of Korean and Western serological assays for H. pylori detection in Korean adults. The 114 Korean adults who visited our centre over a 6-month period for the evaluation of H. pylori infection using the urea breath test (UBT) were enrolled in this prospective study. Anti-H. pylori IgG was measured using three commercially available immunoassays: Genedia H. pylori ELISA (Green Cross Medical Science), Chorus helicobacter IgG (DIESSE Diagnostica Senese) and Vidas H. pylori IgG (bioMérieux). Positive UBT findings were obtained in 40.6% of included subjects. The sensivities and the specificities of Vidas, Chorus and Genedia were 89.7%, 100% and 100% and 85.5%, 75.4% and 80.7%, respectively. We found no differences in sensitivity between the Vidas and Chorus (P = 0.125), Chorus and Genedia (P = 0.125) and Vidas and Genedia (P = 1.000) assays. There were also no differences in specificity between the Vidas and Chorus (P = 0.070), Chorus and Genedia (P = 0.508) and Vidas and Genedia (P = 0.549) assays. In Korean adults, the Genedia H. pylori ELISA, Chorus helicobacter IgG and Vidas H. pylori IgG assays exhibited a high concurrence rate with similar diagnostic accuracy. Thus, both the Korean and Western non-invasive assays are reliable for serodiagnosis of H. pylori in Korean individuals.

INTRODUCTION

Several non-invasive tests, which do not require endoscopy, have been developed to establish whether a patient is infected with Helicobacter pylori. The serum anti-H. pylori immunoglobulin G (IgG) antibody test, which is both simple and inexpensive, is one of these tests (Burucua et al., 2013; Faigel et al., 2000). Although it cannot distinguish between active and recently treated infections, serology is more accurate than biopsy in endemic areas of H. pylori infection and gastric cancer. This is due to the increased likelihood of false negative results with invasive tests in patients having severe gastric atrophy and/or intestinal metaplasia (Korstanje et al., 2008).

Various commercial enzyme-immunoassays (EIAs), including ELISA kits, are now available and are considered to be a reliable and valid method for detection of gastric colonization by H. pylori. However, there is some debate on the overall performance of H. pylori diagnostic kits from particular geographical regions. For example, a Chinese study found that Western commercial serological tests performed poorly when used in Chinese patients (Leung et al., 1999). Despite the high accuracy of these tests reported in the West, they yielded variable diagnostic performance in other regions, possibly due to strain-specific differences in serum anti-H. pylori IgG antibody responses in different populations (Marchildon et al., 2003). Conversely, a recent Chinese study showed that a Western serology test was reliable for screening H. pylori infection in Chinese populations (Ren et al., 2005).

Serology tests that were developed in Western countries are widely performed in Korea. However, it has been shown that Koreans and their Western counterparts differ in regard to the types of H. pylori they harbour, with Koreans showing a high prevalence of the East Asian CagA type (Lee, 2012, 2014; Seo et al., 2011). Therefore, serological assays should be validated in each population and a consensus report also stated that only validated commercial tests should be used (Malfertheiner et al., 2012). Despite this, most serological assays used in Korea are not validated in Asian populations and have not been directly compared to Korean assays for detection of H. pylori antigen in Korea, where the prevalence of gastric cancer is the highest in the world. Therefore, we aimed to validate and compare the diagnostic accuracy of Korean and Western serological assays for H. pylori detection in Korean adults.

METHODS

Sample size calculation. We calculated the minimal number of infected cases and non-infected cases required to ensure a 10% precision of the sensitivity and specificity estimate (α=0.1). We
estimated that the sensitivity (or specificity) of the evaluated kits is 90 % (P=0.9) with a 95 % confidence interval (Burucua et al., 2013). Using the equation \([\text{cases} = (1.96)^2 \times (1 - P) / \alpha^2]\) (Banoo et al., 2010), we required 35 infected and non-infected patients. With 50 % estimated prevalence, it was necessary to recruit a total sample of 75 patients.

**Subjects.** Korean adults who visited our centre between January and June of 2014 for the evaluation of *H. pylori* infection using the urea breath test (UBT) were included in this prospective study. The inclusion criteria were Koreans over 16 years old who had completed UBT and blood sampling on the same day at our centre. Subjects were excluded when they had a recent history of medication, including *H. pylori* eradication, within the previous 6 months. Other exclusion criteria were the presence of significant disease that required treatment, pregnancy, lactation or incomplete data in any of the performed tests.

Informed consent was obtained and the study was approved by the institutional review board (IRB) of Konkuk University School of Medicine, which confirmed that it was in accordance with the ethical guidelines of the Helsinki Declaration. After the approval, this study was registered as ClinicalTrials.gov ID: KCT0001187 (https://cris.nih.go.kr).

**Serum anti-*H. pylori* antibody test.** Blood samples were obtained from each participant after 8 h of fasting. These were subjected to centrifugation, and serum was stored at −70 °C prior to analysis. Serum levels of anti-*H. pylori* IgG antibodies were measured using three commercially available immunoassays: Genedia *H. pylori* ELISA (Green Cross Medical Science), Chorus *H. pylori* IgG (DIESSE Diagnostica Senese) and Vidas *H. pylori* IgG (bioMérieux). All three tests were performed using the same blood sample, according to the manufacturer’s instructions.

**Genedia *H. pylori* ELISA test.** The Genedia kit is an ELISA that uses peroxidase-conjugated anti-human IgG. This assay uses *H. pylori* antigen obtained from Korean *H. pylori* strains. The cut-off value for this kit was defined as mean absorbance of the negative control +0.4 OD at 450 nm, as specified in the manufacturer’s instructions. The sensitivity and specificity of this assay in Korean adults have been reported as 97.8 % and 92.0 %, respectively (Kim et al., 1998).

**Chorus *H. pylori* IgG test.** For the Chorus *H. pylori* IgG assay, *H. pylori* antigen-specific immunoglobulins are bound to solid phase and detected by peroxidase-conjugated anti-human IgG monoclonal antibodies. The titre was expressed in AU (arbitrary units); sample IgG concentrations >12 AU were considered positive, and those <8 AU were considered negative. The samples between these cut-off levels were considered equivocal. The manufacturer’s reported sensitivity and specificity are 89 % (95 % confidence interval [CI]: 77–95 %) and 100 % (95 % CI: 92–100 %), respectively. A study from Italy found that, in Italian adults, this assay displayed a sensitivity and specificity of 95 % and 100 %, respectively (Figura et al., 1994).

**Vidas *H. pylori* IgG test.** The Vidas *H. pylori* IgG assay, which utilizes the enzyme immunoassay sandwich method, is combined with a final fluorescent detection and thus is described as an enzyme-linked fluorescent assay. During the final detection step a relative fluorescent value (RFV) was calculated, and the test value (TV) was obtained by dividing the sample RFV by standard RFV. The result was considered positive when TV ≥1.00 and negative when TV <0.75. The samples between these cut-off levels were considered equivocal. The manufacturer’s claimed sensitivity and specificity for the Vidas *H. pylori* IgG test are 98.1 % (95 % CI: 93.1–99.8 %) and 90.82 % (95 % CI: 83.3–95.7 %), respectively.

**Definitive diagnosis for *H. pylori* infection.** The gold standard for *H. pylori* infection in this study was the UBT. It was conducted in the fasting state on the same day as the blood sampling for the serological assays, according to the manufacturer’s instructions. Briefly, 100 mg of C13 urea packed into a capsule was ingested by each patient, and exhaled breath samples were analysed using Heliview mass spectrometry (Medichems). Breath samples at baseline and 30 min after C13 urea administration were collected. A difference in C13CO2 concentration between the baseline and 30 min samples of over 2.4 was recorded as a positive finding for *H. pylori* infection. The results were expressed as positive or negative, with a difference in C13CO2 concentration shown as parts per thousand.

If all three serology tests and the UBT were positive, the subject was classified as having a definite *H. pylori* infection. If all of the tests were negative, the subject was classified as having no infection. In subjects who showed discrepancies between the three serology tests, those with a positive UBT were defined as likely to have an infection and those with a negative UBT were classified as unlikely to have an infection.

**Statistical analysis.** Sensitivity, specificity, positive predictive values and negative predictive values were calculated based on confirmation of *H. pylori* infection status using the UBT. Values were expressed as a percentage, with 95 % CI, and a P-value <0.05 was considered statistically significant. Statistical analysis was performed using MedCalc Statistical Software (version 11.2.1, MedCalc Software). McNemar’s test was used to calculate the statistical difference between assays. Continuous variables were summarized as mean ± SD, and categorical variables were expressed as frequency (%). Agreement between assays was assessed using Cohen’s kappa coefficient. Poor agreement was reported when the kappa value was <0.4, while excellent agreement was reported when the kappa value was >0.75. Kappa values between 0.4 and 0.75 were reported as fair to good agreement (Malletheiner et al., 2012). Equivocal results of Vidas and Chorus were considered negative for analysis.

We determined that 100 samples with estimated prevalence of 50 % were required to achieve 80 % power, in order to detect a difference of 0.1000 between a diagnostic test with an area under the ROC curve (AUC) of 0.8000 and another diagnostic test with an AUC of 0.9000 using a two-sided z-test at a significance level of 0.0500. The data were discrete (rating scale) responses. The AUC was computed between false-positive rates of 0.000 and 1.000. The ratio of the SD of the responses in the negative group to the SD of the responses in the positive group for diagnostic test 1 was 1.000 and for diagnostic test 2 was 1.000. The correlation between the two diagnostic tests was assumed to be 0.600 for both positive and negative groups (Obuchowski & McClish, 1997). All analyses were conducted using PASS (http://www.ncss.com/software/pass).

**RESULTS**

**Subject characteristics.** Of the 114 subjects who visited our centre for evaluation, 18 were excluded due to a recent history of *H. pylori* eradication or other medication. Positive UBT findings were obtained for 39 (40.6 %) of the 96 included subjects (Table 1). Thirty-five subjects (36.5 %) showed positive findings in all three serological tests and UBT, whereas 39 subjects (40.6 %) showed negative findings in all three serological tests and UBT. There were four subjects (34-, 36-, 42- and 61-year-old males) who tested negative by UBT but gave positive findings in all three serological tests.

**Comparison of serological assays using UBT as the predefined gold standard**

The sensitivities and the specificities of Vidas, Chorus and Genedia were 89.7 %, 100 % and 100 % and 85.5 %, 75.4 %, respectively.
and 80.7 %, respectively (Table 2). There were no differences in sensitivity between the Vidas and Chorus (P = 0.125), Chorus and Genedia (P = 0.125) and Vidas and Genedia (P = 1.000) assays. In addition, there were no differences in specificity between the Vidas and Chorus (P = 0.070), Chorus and Genedia (P = 0.508) and Vidas and Genedia (P = 0.549) assays. Weighted kappa values of Vidas, Chorus and Genedia for concordance with UBT were 0.741, 0.714 and 0.773, respectively, and did not differ significantly (Table 2).

**Concordance rates among the three serological tests**

Weighted kappa values for concordance rates between the Vidas and Chorus, Chorus and Genedia and Vidas assays were 0.747 (0.617–0.878), 0.812 (0.695–0.929) and 0.683 (0.537–0.828), respectively. Pairwise comparison of ROC curves showed that there were no statistically significant differences between Vidas and Chorus (P = 0.659), Genedia and Vidas (P = 0.810) and Chorus and Genedia (P = 0.483), as shown in Fig. 1.

**Subjects showing discrepancy between the diagnostic tests**

Only 18 of the 96 subjects included in this study showed discrepancy in the results of the three serology tests (Table 3).

Of the three tests, only Vidas gave a false-negative finding in four subjects who tested positive by UBT. The titres measured by Vidas were between 0.7 and 0.99 (one negative and three equivocal) in these four subjects.

Of the 14 samples that tested negative by UBT, false-positives were obtained by Chorus in ten cases, by Genedia in seven cases and by Vidas in four cases. All four Vidas false-positives had low titres, measuring between 1.02 and 1.15 TV. Similarly, the ten samples that gave false-positive results in the Chorus assay exhibited low titres, giving values of between 18.0 and 108 AU, and the seven Genedia positives showed variable OD values.

**DISCUSSION**

In this study, we found that the diagnostic accuracy of the three serological assays for detection of *H. pylori* did not differ relative to the UBT findings. Notably, there were no differences in sensitivity and specificity among Genedia, Chorus and Vidas assays. In most of the subjects the three serology tests showed results consistent with UBT, suggesting that all of three of these non-invasive assays are reliable for serodiagnosis in Korean individuals.

Here we measured only IgG antibodies, as it has been shown that IgM and IgA antibodies do not perform as well in the detection of *H. pylori* (Laheij et al., 1998; Talley et al., 1991). We recently showed that serum anti-*H. pylori* IgG antibodies are also useful in assessing the risk of gastric neoplasia and predicting pathology type in Korean adults (Choi et al., 2014). Of the 3328 participants included in this previous study, the incidence of gastric neoplasia was highest in the gastric atrophy (+)/*H. pylori* (−) group (4.17 %; odds ratio 25.8, P = 0.009), but the neoplasm exhibited the least advanced histology using the IgG serology test. Importantly, in this previous study we could not confirm whether there was a difference between Western and Korean serology kits (Choi et al., 2014) because the Genedia *H. pylori* ELISA test, which utilizes *H. pylori* antigen obtained from Korean strains, was not tested.

**Table 2. Performance of three different serological assays based on UBT findings**

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
<th>Efficiency (overall accuracy)</th>
<th>Weighted kappa value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vidas</td>
<td>89.7 %</td>
<td>85.5 %</td>
<td>81.4 %</td>
<td>92.2 %</td>
<td>87.2 %</td>
<td>0.741</td>
</tr>
<tr>
<td><em>H. pylori</em> IgG</td>
<td>(75.8–97.1 %)</td>
<td>(73.3–93.5 %)</td>
<td>(66.6–91.6 %)</td>
<td>(81.1–97.8 %)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chorus</td>
<td>100 %</td>
<td>75.4 %</td>
<td>73.6 %</td>
<td>100 %</td>
<td>85.4 %</td>
<td>0.714</td>
</tr>
<tr>
<td><em>Helicobacter</em> IgG</td>
<td>(91.0–100.0 %)</td>
<td>(62.2–85.9 %)</td>
<td>(59.7–84.7 %)</td>
<td>(91.8 %–100.0 %)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genedia</td>
<td>100.0 %</td>
<td>80.7 %</td>
<td>78.0 %</td>
<td>100.0 %</td>
<td>88.5 %</td>
<td>0.773</td>
</tr>
<tr>
<td><em>H. pylori</em> ELISA</td>
<td>(91.0–100.0 %)</td>
<td>(68.1–90.0 %)</td>
<td>(64.0–88.5 %)</td>
<td>(92.3–100.0 %)</td>
<td></td>
<td>(0.649–0.896)</td>
</tr>
</tbody>
</table>

PPV, positive predictive value; NPV, negative predictive value.

All values except the overall accuracy are expressed as a percentage with 95 % confidence interval. Equivocal results were considered negative in the analysis.
In our current study, we found a high concurrence rate among the three serological tests with each displaying similar diagnostic accuracies. Our findings are consistent with another recent report which demonstrated that Western serology testing can provide a reliable method for screening of *H. pylori* infection in Chinese populations (Ren et al., 2005). Although a meta-analysis also found that a similar overall accuracy of common commercial serological kits (Loy et al., 1996), studies from China and Japan have shown poor performance of Western commercial serological tests when used in Asian subjects (Leung et al., 1999; Marchildon et al., 2003; Miwa et al., 2000). These results may due to the use in Western kits of *H. pylori* antigens that are less prevalent in East Asians (Miwa et al., 2000), as antigen composition will vary according to the different strains utilized in the generation of a particular serological assay (Marchildon et al., 2003). Therefore, Western serological assays should be validated in Asians where the prevalence of *H. pylori* infection and gastric cancer is high. Importantly, prior to this study, the Vidas and Chorus assays had not been validated in Asians while various Western kits were evaluated in previous studies. Also, these assays had not been directly compared to Korean serological assay in Korean individuals. The Vidas *H. pylori* IgG test uses a pool of native purified antigens derived from five different strains (ATCC 53721, 53722, 53725, 53726, and 53727), and the Chorus *H. pylori* IgG test also uses *H. pylori* G21 and G39 cells as antigens. Antigens in these Western kits are prepared from diverse strains and represent various antigens, including both common and strain-specific. Moreover, the prevalence of the East Asian CagA subtype was found to be 94.4% in our population according to a previous report from our institution, suggesting that the majority of Koreans are infected with the East Asian CagA subtype of *H. pylori* (Seo et al., 2011). Less diversity in our population may partly explain this discrepancy in results among studies.

It is notable that the three serological assays showed different diagnostic trends as compared with UBT. For example, Vidas had a tendency to give more false-negative findings, whereas Chorus had a tendency to give more false-positives. Adjustment of the cut-off level could enhance the performance of each assay according to test purposes. For example, all four

![Fig. 1. Pairwise comparison of ROC curves. Area under the curves of Vidas, Chorus and Genedia were 0.952, 0.964, and 0.945, respectively. There was no significant difference between the groups.](image)

**Table 3.** Detailed results of subjects showing discrepancies among the three serological tests

<table>
<thead>
<tr>
<th>UBT</th>
<th>Subject</th>
<th>Genedia (OD)</th>
<th>Chorus (AU)</th>
<th>Vidas (TV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probably infected (n=4)</td>
<td>Positive (n=4)</td>
<td>M/29</td>
<td>P, 1.0930</td>
<td>P, 83.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M/72</td>
<td>P, 0.4630</td>
<td>P, 109</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M/41</td>
<td>P, 0.7570</td>
<td>P, 132</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/48</td>
<td>P, 3.0660</td>
<td>P, 178</td>
</tr>
<tr>
<td>Probably not infected (n=14)</td>
<td>Negative (n=4)</td>
<td>M/36</td>
<td>P, 0.5830</td>
<td>P, 18.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M/57</td>
<td>P, 0.7990</td>
<td>P, 43.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/61</td>
<td>P, 0.4980</td>
<td>P, 99.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/51</td>
<td>P, 1.7200</td>
<td>P, 108</td>
</tr>
<tr>
<td></td>
<td>Negative (n=3)</td>
<td>M/58</td>
<td>P, 0.4200</td>
<td>N, 6.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M/29</td>
<td>P, 0.8680</td>
<td>N, 7.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M/32</td>
<td>P, 0.5560</td>
<td>E, 9.6</td>
</tr>
<tr>
<td></td>
<td>Negative (n=3)</td>
<td>M/57</td>
<td>N, 0.2790</td>
<td>P, 18.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M/39</td>
<td>N, 0.2210</td>
<td>P, 24.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M/64</td>
<td>N, 0.0380</td>
<td>P, 40.4</td>
</tr>
<tr>
<td></td>
<td>Negative (n=3)</td>
<td>F/48</td>
<td>N, 0.2720</td>
<td>P, 14.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M/36</td>
<td>N, 0.0330</td>
<td>P, 33.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M/59</td>
<td>N, 0.3930</td>
<td>P, 61.0</td>
</tr>
<tr>
<td></td>
<td>Negative (n=1)</td>
<td>M/63</td>
<td>N, 0.2840</td>
<td>E, 10.2</td>
</tr>
</tbody>
</table>

UBT, Urea breath test; M, male; F, female; AU, arbitrary unit; TV, test value; P, positive; N, negative; E, equivocal.
subjects who showed false-negative (negative or equivocal) findings with the Vidas test had antibody titres between 0.7 and 0.99, values close to the positive cut-off (≥1.00 TV). Additionally, while there is a wide equivocal zone for the Chorus and Vidas tests, retesting is recommended for samples within 10% of the cut-off values for the Genedia test due to the absence of an equivocal zone. For the Chorus and Vidas tests, using a higher cut-off value would decrease the overall H. pylori seroprevalence as compared with that obtained with the lower cut-off values for these tests. In the present study, equivocal results were considered negative when we compared the results to UBT. This approach is preferable for Chorus assay, but use of a low limit for the equivocal zone (0.75 TV) can increase sensitivity in the Vidas assay, which shows a sensitivity of 97.4% and specificity of 78.2%.

The differential sensitivities and specificities obtained in various studies with the same kit indicate that it is more important to test a kit in a particular study population than to choose any specific kit (Jensen et al., 1993). Of the 14 UBT negative subjects, ten tested positive with the Chorus assay, indicating that among the three serology tests, the false positive rate was highest in Chorus. Of these ten subjects, three showed false-positive results with the Vidas test, albeit with relatively low titres between 1.11 and 1.15. Since these three subjects also showed low titres in the Chorus assay, between 18.2 and 40.4 AU, past infection or spontaneous regression of H. pylori might be the reason for false-positives in these cases. Conversely, serum from the four subjects who showed false-positives in both the Chorus and Genedia tests showed a wide range of titres, from 0.498 to 1.720 OD for Genedia, and from 18.0 to 108 AU for Chorus. Taken together, these data suggest that recommendations for the optimal use of different interpretation criteria should be made for each serology test.

13C-UBT was used as the gold standard in this study because there is no single test that suffices as a criterion standard to detect H. pylori infection. Although UBT is considered to be one of the most reliable tests, it is not always accurate due to the various optimal cut-off values of 13C-UBT. The presence of four subjects who tested negative by UBT but gave positive findings in all three serological tests in our study suggests a high probability of false-negative findings by UBT under cryptic conditions. Our findings with the Vidas test had antibody titres between 0.7 and 0.99, values close to the positive cut-off (¢1.00 TV).

In conclusion, the Genedia H. pylori ELISA, Chorus helicobacter IgG and Vidas H. pylori IgG assays exhibited a high concurrence rate, with similar diagnostic accuracy, in Korean adults. Therefore, all three of these non-invasive assays are reliable for serodiagnosis in the Korean population.

ACKNOWLEDGEMENTS

This paper was supported by Konkuk University in 2015. The authors declare no competing interest. References


