Prospective study for validation of a single protocol for the 13C-urea breath test using two different devices in the diagnosis of *H. pylori* infection

Luiz Gonzaga Vaz COELHO, Osmar Reni TRINDADE, Laiane Alves LEÃO, Henrique Gomes RIBEIRO, Izabella Silva FREITAS and Maria Clara Freitas COELHO

ABSTRACT – Background – 13C-urea breath test (UBT) is the gold-standard, noninvasive method for *H. pylori* diagnosis. However, there is no uniform standardization of the test. This situation can be unpractical for laboratories running with two or more devices. Objective – To perform a prospective comparison validation study of UBT employing one validated protocol for two different devices: BreathID Hp Lab System® (Exalenz Bioscience Ltd, Israel), here called device A and IRIS-Doc2® (Wagner Analysen-Technik, Germany, now Mayoly Spindler Group, France), here called device B, in the diagnosis of *H. pylori* infection. Methods – A total of 518 consecutive patients (365 females, 153 males, mean age 53 years) referred for UBT were included. All patients received device A protocol as follow: after at least one hour fasting, patients filled two bags prior to the test, then ingested an aqueous solution containing 75 mg of 13C-urea with a 4.0 g citric acid powder and filled another two bags 15 min after ingesting the test solution. One pair of breath sample bags (before and after ingestion) was analyzed by the two different devices. A delta over baseline (DOB) ≥5‰ indicated *H. pylori* infection. Statistics: Wilcoxon test, kappa coefficient with 95% CI, Wilson’s method. Results – Considering the device A protocol as the gold standard, its comparison with device B showed a sensitivity of 99.3% (95% CI: 96.3–99.9) and a specificity of 98.9% (95% CI: 97.3–99.6). Kappa coefficient was 0.976 (95% IC: 0.956–0.997). Conclusion – Correlation between the two devices was excellent and supports a uniform standardization of UBT.


INTRODUCTION

*Helicobacter pylori* is recognized as the main etiologic agent of peptic ulcers, with a pathogenic role equally well established in gastric adenocarcinoma and MALT (mucosa-associated lymphoid tissue) lymphoma(1,2). First identified in 1983 by Marshall and Warren(3) from gastric tissue fragments, since then, different diagnostic methods have been developed for their detection, including invasive and non-invasive tests. Invasively, the bacterial presence can be identified through gastroscopy by the collection of gastric fragments for histology, culture, urease test or molecular tests. Non-invasive tests consist of serological tests, stool antigen detection, and the 13C urea breath test (UBT)(4).

UBT is the gold-standard noninvasive method for *H. pylori* diagnosis, with a sensitivity and specificity greater than 95%(5-7). Due to its high accuracy, low cost and easy performance, it is considered the first option in the control of *H. pylori* infection treatment or recurrence and it is a fundamental diagnostic tool in the “test and treat” strategy; it is also an excellent option in epidemiological studies(8-10). Such tests are based on the property of *H. pylori* in producing high amounts of urease enzyme. The principle of the test is based on the ability of *H. pylori* (if present in the gastric environment) to break down orally absorbed 13C-labelled urea. 13CO2 diffuses into the blood and is excreted via the lungs and can be easily measured in the expired air using mass or non-dispersive, isotope-selective infrared spectroscopy(11). Thus, the analysis of samples of expired air collected before (control) and after substrate ingestion will indicate the change in the 13CO2/12CO2 ratio caused by the metabolic activity induced by the administration of the labelled urea.

Since its original description in 1987 by Graham et al.(12), UBT has undergone several modifications involving the need or not of fasting before the test(13), dose of 13C-urea employed(14), concomitant administration of 13C-urea in any citrus meal(14), ideal sample collection time for exhaled air after substrate ingestion, optimal cut-off point and performance of the device used to perform UBT(13,15). Thus, there is no single worldwide standardization for testing to date, although numerous individual validation studies confirm the high accuracy of the method for the diagnosis of active *H. pylori* in adults and children over 6 years of age(4,10,16-18).

Considering the current moment, where there are several devices in the world market, including hospitals and clinics working with more than one device from different manufacturers in the daily routine, the search for a single UBT preparation and reading protocol should be investigated to standardize processes and improve efficiency.
The aim of this study is to perform a prospective comparison validation study of $^{13}$C-urea breath test for the diagnosis of \textit{H. pylori} infection employing one validated protocol for two different devices.

**METHODS**

The study was performed at the Breath Tests Laboratory of the Alfa Institute of Gastroenterology at Clinics Hospital of Federal University of Minas Gerais, Belo Horizonte, MG, Brazil.

**Patients**

We recruited individuals consecutively referred for UBT for the purpose of the initial diagnosis or control of \textit{H. pylori} infection treatment from November 2017 to September 2018. Informed consent was obtained from all patients to participate in the study after being duly informed about the purpose of the study. The study inclusion criterion was the non-use of proton pump inhibitors or antibiotics in the last 14 or 30 days, respectively, prior to UBT.

$^{13}$C-urea Breath Test (UBT)

All study participants underwent UBT, which was processed and analysed simultaneously by two different devices. Our laboratory has two devices: the BreathID HP Lab System® (Exalenz Bioscience Ltd. Israel), here called device A, and IRIS-Doc 2® (Wagner Analysen-Technik, Bremen, Germany, now Mayoly Bioscience Ltd, Israel), here called device B. Although these two devices have independent protocols previously validated and recommended by the manufacturers\(^{(17,19)}\), we chose a single protocol as recommended by the manufacturer of device A, cleared by FDA in November 2016, due its particularities and practicality\(^{(19)}\).

The protocol employed can be summarized as follows: after at least 1h fasting, exhaled air samples were initially collected from the participants in two small collection bags (120 mL), which corresponded to time zero (sample-1, control). Next, the patients ingested, within 2 min, an aqueous solution (200 mL) containing 75 mg of $^{13}$C-urea and 4.0 g of citric acid powder, with added edulcorant. A second exhaled air collection was performed 15min after the ingestion of the substrate in two new small collection bags, which corresponded to sample-2. Each pair of collected material (sample-1 and sample-2) was analysed and processed by one of the two infrared analyser devices of the study. According to the manufacturer’s instructions, patients were considered positive for \textit{H. pylori} when they had a delta over baseline (DOB) equal to or greater than 5‰\(^{(19)}\). This parameter indicates the change in the $^{13}$CO/$^{12}$CO ratio in metabolic activity induced by the administration of the labelled urea.

For statistical analysis, the data were expressed as percentages, means (standard deviation), median and minimum and maximum values. Continuous variables were compared using t-test or the Mann-Whitney test (non-parametric data) and Wilcoxon test for paired samples. The coefficient of concordance (kappa) of the tests between the two devices was calculated. Considering device A, whose protocol was used in the study as the gold standard, the sensitivity, specificity and 95% confidence interval were calculated for the results obtained with device B. Wilson’s method was used to calculate the confidence intervals\(^{(20)}\). All statistical analyses were performed using the MINITAB statistical package (Minitab Inc., State College, PA, USA) version 16 and Excel (Office 10).

**RESULTS**

Five hundred eighteen patients were consecutively included in the study: 365 (70.5%) patients were women and 153 were men (29.5%), with a mean age of 53 years (10-89 years) and a standard deviation (SD) of 15.3 years. Among the 518 patients, 161 had never received anti-\textit{H. pylori} treatment and 357 patients underwent the test to evaluate the anti-\textit{H. pylori} treatment result. TABLE 1 shows the observed values in DOB‰ in the UBT obtained using the single protocol (device A) in the two different devices.

FIGURE 1 shows the absolute values of DOB‰ observed in all 518 study participants. Considering the cut-off point of DOB ≥5‰ for the presence of \textit{H. pylori}, only 5/518 participants (FIGURE 2) presented discordant results between the two devices, a positive participant on device A and a negative one on device B, while four subjects were negative on device A and positive on device B. Four in five participants with discordant results underwent the test for post-

**TABLE 1. UBT results employing two different devices with a unique protocol (n=518).**

<table>
<thead>
<tr>
<th>DOB ‰</th>
<th>Device A</th>
<th>Device B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Negative</td>
</tr>
<tr>
<td>Mean value</td>
<td>13.1 (24.7)</td>
<td>0.4 (1.0)</td>
</tr>
<tr>
<td>Median value</td>
<td>0.9</td>
<td>0.4</td>
</tr>
<tr>
<td>Min. value</td>
<td>-1.0</td>
<td>-1.0</td>
</tr>
<tr>
<td>Max. value</td>
<td>119.1</td>
<td>4.9</td>
</tr>
</tbody>
</table>

UBT: $^{13}$C-urea breath test; DOB: delta over baseline; SD: standard deviation.
In its initial description, UBT was performed using isotope ratio mass spectrometry technology, which has a high cost and complexity regarding use and maintenance\(^{(12)}\). More recently, UBT has been performed using non-dispersive, isotope-selective infrared spectroscopes, which has a lower cost, smaller size, easy maintenance and operation, and which results are available a few minutes after the procedure; this allows UBT to be carried out in doctors’ offices or small-to-medium sized laboratories. Its excellent performance has made this methodology attractive to the industry and physicians, since there are numerous devices being commercialized in the international market\(^{(11,23,24)}\). Despite the existence of small variations in the methods employed by each device, regarding fasting time before UBT, ideal test meal, best \(^{13}\)C-urea dose, optimal breath sampling after ingestion of the substrate and best cut-off point to discriminate infected from non-infected individuals, several local validation studies have shown high diagnostic accuracy to detect \(H.\) pylori infection\(^{(6,9)}\). However, the absence of definitive standardization of the test makes it impractical to manage different devices at the same place and to compare the results from different studies.

In our study, although the protocols for the two devices were previously validated, we chose to use the protocol suggested by device A for its practicity, especially in relation to fasting time (1h instead of 8h), use of citric acid instead of orange juice and 15 min instead of 30 min for the optimal breath sampling after ingestion of the substrate. The fasting time of only one hour makes it more convenient to perform the test at different times, with several studies demonstrating that the differences between DOB‰ values fasting or not fasting are minimal or non-existent\(^{(25-28)}\). The incorporation of citric acid as a test meal instead of orange juice administered together with \(^{13}\)C-urea is currently well established\(^{(14,29)}\). Its administration delays gastric emptying, allowing the labelled substrate to distribute throughout the stomach and maximizes the reaction with the preformed bacterial urease. Additionally, it increases the hydrolysis of urea, both by increasing the availability of intracytoplasmic urease and increasing the activity of intragastric urease, providing higher \(^{13}\)CO\(_2\) recovery values\(^{(14,29)}\). Its effect is dose dependent, and the 4.0 g dose used here seems to be palatable and efficient to maximally enhance urease activity\(^{(29,30)}\). Recently, a new test meal using a high-dose mixture of citric, malic and tartaric acid has been proposed to be used as a test meal in patients taking proton pump inhibitors with good accuracy, although 7.2% of the patients complained of dyspeptic symptoms with the test meal\(^{(31)}\). The dose of \(^{13}\)C-urea administered is between 50 mg and 100 mg, and 75 mg is the dose most commonly used currently\(^{(11)}\). Studies have shown that doses above 125 mg are unnecessary\(^{(32)}\), and a dose of 25 mg is inefficient to provide good accuracy\(^{(33)}\). The interval of 15 min instead of 30 min for optimal breath sampling after ingestion of the substrate has been considered sufficient to avoid interference of the pharyngeal flora with the possible presence of urease-producing bacteria and sufficient to allow hydrolysis \(^{13}\)C-urea by contact with the preformed urease by \(H.\) pylori. The use of citric acid, might also contribute to obtain greater accuracy of UBT with the collection performed 15 min after substrate ingestion by increasing the saliva production and reducing possible interaction with the urease that might be present in the oral cavity\(^{(11,27,34)}\).

Our study also demonstrated that mean UBT values, in both devices, were significantly higher in females (device A: 47.5‰, SD:27.8, \(P=0.036\) and device B:47.4‰, SD:27.8, \(P=0.027\) than

### DISCUSSION

Our study shows that the UBT employing the same protocol for two different devices from two different manufacturers presented very similar results. This allows their standardization in daily practice. Only 5/518 (0.97%) participants (four of them performed for treatment control) showed discordant results, and the DOB values were close to the cut-off point in three of them. Characteristically, in the UBT, DOB‰ values in individuals infected or not infected by \(H.\) pylori are situated far from the established cut-off point, and our study shows very similar absolute values, both positive and negative, between the two devices analysed. Typically, borderline cut-off UBT observed values (grey or indeterminate zone) should be interpreted with caution, and it is suggested that the test should be repeated or the diagnosis should be confirmed by another method.

The prevalence of UBT results in the grey area has been estimated at 1% to 2%, similar to the one observed here\(^{(21,22)}\).

### TABLE 2. DOB‰ absolute values in \(H.\) pylori-infected patients (n=147), regarding sex.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Device A</th>
<th></th>
<th>Device B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>56.4 (21.9)</td>
<td>47.5 (27.8)*</td>
<td>35.8 (21.2)</td>
<td>47.4 (27.8)**</td>
</tr>
<tr>
<td>Median</td>
<td>33.5</td>
<td>41.4</td>
<td>30.3</td>
<td>43.1</td>
</tr>
<tr>
<td>Min</td>
<td>6.1</td>
<td>5.6</td>
<td>5.6</td>
<td>5.1</td>
</tr>
<tr>
<td>Max</td>
<td>87.4</td>
<td>119.1</td>
<td>84.4</td>
<td>123.8</td>
</tr>
</tbody>
</table>

DOB‰: delta over baseline; *\(P<0.036\); **\(P<0.027\).
in males (device A: 36.4%, SD: 1.9, and device B: 35.8%, SD: 21.2). These results demonstrate that, for the first time, using two distinct devices simultaneously corroborate findings already described in studies using only one device. This demonstrates that UBT has absolute values significantly higher in adult females than in adult males, with a significant increase varying from 4.5% to 11% in the median UBT, and similar to our results from 11.1% in device A and 11.6% in device B\(^{29,30}\). The reasons for these findings are still largely unknown. Variables such as higher bacterial density in women, hormonal changes, body surface area and sex differences in intragastric pH are being investigated, but further investigations are clearly needed\(^{29,30}\).

In conclusion, our study showed that UBT performed by two different devices employing a single protocol presents excellent agreement between them. This harmonization, while improving and simplifying the operational procedures, represents an important contribution in the search for a single standardization for UBT.

REFERENCES


ACKNOWLEDGEMENTS

The authors thank to Maurilio M. Fernandes and Maria de Lourdes M. Fernandes for technical and statistical assistance.

Authors’ contribution

Coelho LG and Trindade OR contributed to the conception and design of the study; Leão LA and Trindade OR contributed to collection of the samples; Ribeiro HG and Freitas IS interpreted and analysed the data and Coelho LG and Coelho MC wrote the paper.


Orcid

Luiz Gonzaga Vaz Coelho. Orcid: 0000-0002-8721-7696.
Laiane Alves Leão. Orcid: 0000-0001-6547-3798.
Izabella Silva Freitas. Orcid: 0000-0002-5500-8881.
Maria Clara Freitas Coelho. Orcid: 0000-0001-8028-6114.


