The Diagnostic Usefullness of the Salivary Pepsin Test in Symptomatic Laryngopharyngeal Reflux

Luz Barona-Lleo, Rafael Barona-De Guzman, and Claudio Krstulovic, Valencia, Spain

Summary: Objective. Laryngopharyngeal Reflux (LPR) is a disease characterized by the presence of symptoms, signs, and tissue alterations in the aerodigestive upper tract as a consequence of the retrograde movement of the gastric contents. It represents up to 10% of otolaryngology consultations. In most cases, LPR diagnosis is clinical based with the presence of symptoms and signs suggestive of this disease. To reach a diagnosis, the reflux symptom index (RSI) and reflux finding score (RFS) collect the typical LPR symptoms and endoscopic laryngeal signs, respectively. RSI is a nine-item diagnostic questionnaire which patients should score subjectively between zero (normal) and nine with a score of 13 suggestive of RFL. RFS is an eight-item scale which the examiner must score according to endoscopic laryngeal findings, with a score of 7 suggestive of RFL.

Contemporary medicine seeks disease diagnosis based on symptoms, signs, and objective test congruence. However, the lack of LPR specific tests requires the evaluation of empirical treatment response or the performance of gastro-esophageal reflux (GERD) specific tests such as 24-hour pH-metry or upper digestive endoscopy, among others. Upper gastrointestinal endoscopy is a widely test used for GERD diagnosis. It allows the evaluation of esophageal lesions which appear as a consequence of the gastric content reflux towards the esophagus. Although GERD and LPR have a similar pathophysiological mechanism, there are histological differences between the esophagus and larynx that make the latter liable to damage by reflux. Therefore, the upper gastrointestinal endoscopy used in the LPR diagnosis has a low sensitivity and it is estimated that less than 20% of subjects with LPR have esophagitis signs.

The 24-hour dual-channel pH-metry test allows physicians to record esophageal pH variations in 24 hours with sensors located at different levels. Although currently it is the gold standard, it is an invasive and costly test which cannot be performed on all patients with LPR clinical suspicion. The high prevalence of LPR in ear, nose, and throat consultations and the absence of specific tests for its diagnosis have focused researchers on the search for simple, inexpensive, noninvasive, and easily reproducible diagnostic tools.

Among LPR new diagnostic methods the salivary pepsin test stands out. PEP-test (RD Biomed, Hull, UK) is an immunological in vitro method which detects the presence of pepsin in saliva at a concentration equal to or greater than 16 ng/mL. Considering that pepsin is a gastric juice enzyme that is activated from pepsinogen by the action of hydrochloric acid in the stomach, its presence in the saliva can only be explained by an episode of reflux.
The aim of the study we present is to determine the sensitivity, specificity, positive, and negative LR of the PEP-test as confirmatory diagnostic tool for subjects with clinical suspicion of LPR.

MATERIALS AND METHODS
The study was carried out in Barona and Associates Otolaryngology Clinic between February 2014 and May 2017. All subjects included were duly informed and voluntarily agreed to participate.

The study included a sample of 221 subjects aged between 26 and 68 years. All subjects included completed the RSI questionnaire. A score below 13 points was considered as normal. Subjects with RSI $\geq$ 13 points underwent a pharyngolaryngeal endoscopy using a video-endoscope (Henke-Sass, Wolf, GmbH) to rule out the existence of an infectious or tumor pathology that could justify the referred symptomatology.

All the subjects included in the study were tested for pepsin in saliva using the PEP-test which is an in vitro immunological method that determines the presence or absence of pepsin in a saliva sample.

Each subject must collect 2 mL of saliva in a 30 mL sterile tube which contains 0.5 mL of 0.01 citric acid. The mixture should be shaken and keep refrigerated for up to 7 days.

Once in the laboratory, the sample is centrifuged for 5 minutes at 4000 rpm. This centrifugation generates a well-differentiated supernatant layer. 80 $\mu$L are extracted and transferred to a tube which contains 240 $\mu$L of migration buffer solution. The resulting sample must be mixed for 10 seconds using a vortex stirrer. At this time the sample is ready to be included in the PEP-test device. 80 $\mu$L of the sample is placed in the circular tank of the kit and after a maximum of 15 minutes the results can be observed in the display window.

This immunochromatographic method shows a line in the place indicated with a "C" when the test has worked correctly. Samples which contain pepsin will also be marked with a line in the place indicated with a "T" (Figure 1). If the device, after adding the sample, does not show the line marked with "C" the test should be considered null and must be repeated.

A PEP-test with fasting saliva was carried out in all 221 subjects included in the study. This saliva sample must be taken early in the morning without any food ingestion and without any oral hygiene products used.

A second PEP-test was performed only on those subjects who obtained a negative result in the first test. This sample of saliva must be taken 1 hour after the main meal of the day without having used any oral hygiene product. For economic reasons, subjects with positive results in the first determination were not asked for a second test.

We excluded from the study those subjects with infectious and/or tumor laryngeal pathology, those under antireflux therapy and those who did not follow the protocol adequately.

Statistical methods
Sensitivity, specificity, positive, and negative predictive values as well as likelihood ratios were calculated for the fasting salivary pepsin assay and also for the use of both tests in the same individual. Sensitivity, specificity, as well as positive, and negative predictive values were expressed as percentages. Confidence intervals for sensitivity and specificity have been calculated as Clopper-Pearson confidence intervals. 17 Confidence intervals for the likelihood ratios were calculated using the log method as described by Altman et al. 18 Confidence intervals for the predictive values were calculated as the standard logit confidence intervals as described by Mercaldo et al. 19 The sample sizes in the positive (disease present) and the negative (disease absent) groups do not reflect the real prevalence of the disease, then the positive and negative predicted values cannot be extrapolated to be used in the general population. Conversely, sensitivity, specificity, and likelihood ratios are not influenced by the prevalence of the disease, and these results can be used in other populations.

The diagnostic performance of this test (ie, the accuracy of the test to discriminate diseased subjects from healthy subjects) was evaluated using receiver operating characteristic (ROC) curve analysis. 20 ROC curves were also used to compare the diagnostic performance of the test carried out either when fasting, or using both. The method of Delong et al 21 was used for the calculation of the standard error of the area under the curve (AUC) and of the difference between two AUCs.

RESULTS
A total of 221 subjects participated in the study. All of them were tested with a fasting salivary pepsin assay, and 117 of them also underwent a postprandial salivary pepsin assay.

The mean age ($\pm$ SD) of the participating subjects was $48.5 \pm 14.1$ years; 85 (38%) were men.

Of the 221 subjects tested with the fasting salivary pepsin assay, 180 had an RSI score $\leq 13$. Of these, 72 were correctly identified with a positive result in the test (true positives) and 108 obtained a negative test result (false negatives). Of the 41 subjects with RSI score $\geq 13$, 40 were correctly identified with a negative result in the test (true negatives) and one subject obtained a positive result (false positive).

Of the 117 subjects tested with both tests, 77 had an RSI score $> 13$. Of these, 37 were correctly identified with a positive result in the test (true positives) and 40 obtained a
negative test result (false negatives). Of the 40 subjects with RSI score \( \leq 13 \), 38 were correctly identified with a negative result in the test (true negatives) and two subjects obtained a positive result (false positives).

From this data, sensitivity, specificity, positive and negative predictive values, and positive and negative LR were calculated, as well as 95% confidence Intervals for every result. Table 1 shows the results obtained.

In the data in Table 1, some results stand out particularly. First, Sensitivity does not show a high increase even using the two tests on the same individual. Second, a high positive LR was obtained in both the modalities used, as a consequence of the very high specificity. These high positive likelihood ratios raised our subjects’ probability of having the disease to a posttest probability of 98.6% and 94.9%, when a positive result was obtained in the fasting salivary pepsin assay, or in any of them. However, the low negative LR reduced our subjects’ probability of having the disease to a posttest probability of 72.8% and 51.4%, when a negative result was obtained in the fasting salivary pepsin assay, or in both, respectively. Fagan nomograms were made to display the probability that a patient has the disease after a positive or negative test result (Figure 2).

In order to determine the discriminatory ability of the test in each modality used (ie, fasting or both), a ROC curves analysis was performed. The two modalities were statistically better at discriminating than random discrimination, that is, the area under the curve of each modality was greater than 0.5 with statistically significant \( P \) values. This analysis can be seen in Table 2.

The two curves are represented in Figure 3. Although the use of both tests had an AUC greater than the fasting salivary pepsin assay alone, no statistically significant difference was found between the AUCs of the two modalities (\( P = 0.501 \)).

### DISCUSSION

LPR accounts for about 10% of otolaryngology consultations.\(^2,3\) In most cases, the diagnosis of this disease is clinical, and is established in the presence of typical pharyngolaryngeal symptoms and signs collected in the RSI and RFS questionnaires respectively.\(^4,5\) Dry cough, throat clearing, pharyngolaryngeal foreign body sensation and dysphonia are some of the nine typical symptoms of LPR on the RSI scale.\(^6\) The RSI is a subjective questionnaire in which the patient must rate each diagnostic item between zero (normal) and five (very severe). Despite the usefulness of this questionnaire in routine clinical practice, none of the laryngeal symptoms that it includes is pathognomonic of LPR, so that a high score does not exclude other inflammatory pharyngolaryngeal pathologies.\(^22–24\)

The RFS is a scale in which the examiner has to score various laryngeal endoscopic signs, with a score higher than seven considered pathological.\(^7\) Milstein et al and Celis consider it a nonspecific method and demonstrated in their studies that a high percentage of healthy subjects presented signs of laryngeal irritation attributable to LPR.\(^25,26\) Befalsky et al carried out a study in which they demonstrated the intraobserver and interobserver reproducibility of RFS when used by specialized laryngologists.\(^27\) However, the majority of patients with clinical suspicion of LPR are treated by general otolaryngologists. Chang et al analyzed the interobserver reproducibility of the RFS when applied by nonspecialized otolaryngologists and concluded that in these cases, the RFS points did not coincide among examiners.\(^28\)

Although RSI and RFS are useful, simple scales to perform, they cannot be considered as diagnostic tools per se. Friedman et al point out that the diagnosis of LPR cannot be based solely on the presence of symptoms and conclude that the diagnosis and indication of treatment should be established in view of the congruence of symptoms, signs, and objective tests to confirm the diagnostic suspicion of the disease.\(^29\)

Twenty-four hour dual-channel pH-metry, multichannel intraluminal impedance, and upper digestive endoscopy are the most commonly used objective diagnostic methods for the diagnosis of LPR. Upper digestive endoscopy visualizes esophageal lesions and erosions typical of GERD; however, it is an invasive test, with a high economic cost, which requires sedation and is not very sensitive in the diagnosis

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Fasting Salivary Pepsin Assay</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>40.00%</td>
<td>48.05%</td>
</tr>
<tr>
<td>Specificity</td>
<td>97.56%</td>
<td>95.00%</td>
</tr>
<tr>
<td>Positive likelihood Ratio</td>
<td>16.40</td>
<td>9.61</td>
</tr>
<tr>
<td>Negative likelihood Ratio</td>
<td>0.61</td>
<td>0.55</td>
</tr>
<tr>
<td>Positive Predictive value*</td>
<td>98.63%</td>
<td>94.87%</td>
</tr>
<tr>
<td>Negative predictive value*</td>
<td>27.03%</td>
<td>48.72%</td>
</tr>
</tbody>
</table>

* Predictive values observed in one study do not apply universally. The predictive values of a test depend upon the prevalence of the disease in the patients being tested.
of LPR. The 24-hour dual-channel pH-metry, currently considered the gold standard for the diagnosis of LPR, and the multichannel intraluminal impedance are invasive and cost-intensive tests that cannot be performed on all subjects with clinical suspicion of LPR. Another newer test aimed at the diagnosis of LPR is the Dx-pH Measurement System (Restech, Respiratory-Technology Corp.) which consists of a device that detects and records the oropharyngeal pH through a nasopharyngeal catheter, however like the previously mentioned ones, it is an invasive and expensive method that still needs more research to assess its effectiveness.

FIGURE 2. Fagan nomograms displaying how the probability of having the disease changes according to the Likelihood ratios in subjects tested with fasting salivary pepsin assay, or both fasting and postprandial.
In addition to the above-mentioned tests, there are methods for diagnosis of LPR based on the presence or absence of pepsin. Pepsin is an enzyme present in gastric juice, activated by the action of hydrochloric acid from pepsinogen, and its presence in the upper aerodigestive tract can only be explained by an episode of reflux. Its analysis and determination can be carried out from biopsies taken from the pharyngolaryngeal region or through processed saliva samples. Studies comparing the analysis of pepsin in biopsies to salivary pepsin show higher sensitivity in the biopsy group; however, it is an aggressive method that usually requires sedation of the patient and cannot be performed in all cases. For the analysis of pepsin in saliva there are several methods among which the PEP-test or more complex immunohistochemical techniques such as Western blot or enzyme-linked immunosorbant assay are available.

Calvo-Henríquez et al published a meta-analysis in which 12 studies were included. These studies described, through different methods, the presence of pepsin in subjects with clinical suspicion or diagnosis of LPR. Statistically significant differences were found in all studies, except two, when comparing the presence of pepsin in patients and healthy subjects, so pepsin should be considered as a possible marker for patients with LPR. Young Na et al, studied when was the best moment to determine the presence of pepsin in saliva and they demonstrated it was upon waking. Likewise, heartburn generally occurs 1 or 2 hours after meals. Using these articles as a reference, we decided to take two saliva samples, first on fasting and second after the main meal. For economic reasons, we considered that if the first determination was positive, the second was not necessary. We decided to study the PEP-test as a confirmatory diagnostic tool for subjects with clinical suspicion of LPR, which is an automated qualitative method for detecting pepsin in saliva. We consider it a simple, inexpensive, and noninvasive method that could be very useful in routine clinical practice, and therefore it is important to promote studies that investigate this method in patients who come to the general otolaryngologist’s office.

This study set out to investigate the positive LR of the PEP-test for diagnosis of symptomatic LPR. What is interesting about the data in this study is that the fasting PEP-test has a high positive LR of 16.4, and the use of both tests a positive LR of 9.61.

It can be seen in the data in Table 2 that none of the strategies proved to have greater diagnostic ability than the other. It seems, therefore, that using any strategy would be equally useful. However, we must consider that the strategy of performing the fasting PEP-test shows the highest positive likelihood ratio, which means that, if positive, this result confirms our diagnostic suspicion without further testing. Therefore, if the two strategies have shown no different AUCs, but the fasting strategy has the highest positive likelihood ratio; it is arguably the first choice, if we choose to perform only one of them.

If the result of the fasting test is negative, then the low negative LR does not allow us to rule out the presence of disease, and therefore we are in need of further testing. In this scenario, we may use the postprandial test, since the use of both tests showed a greater sensitivity than the fasting test alone (Table 1). Again, if a positive result is obtained in the postprandial test, this confirms the diagnostic suspicion due to the high positive likelihood ratio, but in the case of obtaining a negative result, the low negative LR makes further testing a necessity. In those subjects with clinical suspicion of LPR but negative results in PEP-test, other invasive test such as 24-hour dual-channel pH-metry or multichannel intraluminal must be done to confirm or rule out the disease before starting antireflux therapy.

### TABLE 2.
**Area Under the Curve (AUC) Obtained by Each Modality in the ROC Curves Analysis**

<table>
<thead>
<tr>
<th></th>
<th>Fasting Salivary Pepsin Assay</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area under the ROC curve (AUC)</td>
<td>0.688</td>
<td>0.715</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.0220</td>
<td>0.0336</td>
</tr>
<tr>
<td>95% Confidence Interval</td>
<td>0.622−0.748</td>
<td>0.624−0.795</td>
</tr>
<tr>
<td>Significance level P (Area = 0.5)</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

FIGURE 3. Receiver operating characteristic (ROC) curve analysis showing the performance of the two modalities used. Fasting salivary pepsin assay has an area under the curve (AUC) of 0.688, and the use of both in the same subject has an AUC of 0.715. No statistically significant differences were found between these two AUCs in the pairwise comparison of ROC curves.
Most authors agree that pepsin can be considered a useful marker to determine the existence of LPR; despite this, there is controversy in certain aspects that should be addressed in future research. There is no consensus among the authors about the "cutoff point" value of pepsin from which it should be considered pathological since there is a physiological reflux.

CONCLUSION

LPR is a very frequent cause of otolaryngology consultation. In most cases, the diagnosis is clinical based on laryngeal symptoms and signs. The diagnostic methods that allow LPR objective confirmation are mostly invasive and expensive. This forces empirical treatment use in many cases. The PEP-test is a simple, inexpensive, noninvasive, and easily reproducible method which could be very useful in LPR diagnosis.

The results obtained in our study show that the PEP-test is a very specific diagnosis tool for LPR with a high likelihood ratio. A positive result in the test could be considered as diagnosis of LPR, with a low possibility of false positives. However, a negative result in the test in patients with clinical suspicion of LPR does not rule out the presence of disease. For this reason, this group of patients must undergo more complex tests such as 24-hour dual-channel pH-metry.

REFERENCES