## Point by Point Response to Manuscript 81649

## "Intraprocedural gastric juice analysis as compared to rapid urease test for real-time detection of *Helicobacter pylori*"

Dear Editor,

We resubmit the revised manuscript ("Intraprocedural gastric juice analysis as compared to rapid urease test for real-time detection of *Helicobacter pylori*") on behalf of all the co-authors R. Vasapolli, F. Ailloud, S. Suerbaum, J. Neumann, N. Koch, L. Macke, J. Schirra, J. Mayerle, P. Malfertheiner, C. Schulz for consideration of publication in the *World Journal of Gastroenterology.*

We have addressed all the comments and criticisms of the reviewers as outlined below in a point by point response. We revised our manuscript, updated the table according to the suggestions and we adapted the Figure 2 according to STARDS guideline, as recommended.

A marked-up version of the revised manuscript that highlights all changes made to the original version is also attached.

We are grateful for the critical and constructive comments which helped to substantially improve our manuscript and hope that our manuscript in the revised version finds your kind consideration for publication in *WJG*.

With best regards

Sincerely

Prof. Dr. P. Malfertheiner, Dr. Christian Schulz

**First Decision by Company Editor-in-Chief**

Conclusion: Potential Acceptance

**Reviewer 1:**

**Scientific Quality:** Grade B (Very good)
**Language Quality:** Grade A (Priority publishing)
**Conclusion:** Major revision

The authors have conducted this study to evaluate the role of a gastric juice analysis technique with Endofaster versus the conventional rapid urease test used to identify H pylori Infection. This study was fascinating and well-organized throughout the whole article. However, the biggest issue of this study is the gold standard definition. Endofaster is a method designed to overcome sampling error, and histology set as the gold standard is a representative invasive method that cannot avoid sampling error. Few issues need attention of the authors. 1. Regarding biopsy and histopathological confirmation, how were the pathologists blinded regarding the sample processing from Endofaster vs. RUT; please explain. 2. The fundamental concept of this study was that it was “gastric fluid” rather than a one-point pick-up to overcome the sampling error in diagnosing H. pylori infection. Therefore, the gold standard should be the method that does not have sampling issues like a urea breath test or stool antigen test. The authors performed only histopathologic diagnosis as the gold standard definition. 3. Regarding the above question, the importance of the gold standard should be further emphasized because the study was conducted targeting a relatively low H. pylori prevalence (35.3%) cohort. The H. pylori infection in the enrolled patients in this study was lower (29.2%) than the prevalence. The study cohort may generally show a higher infection rate than the prevalence, and this is because the H. pylori infection rate is likely to be high in patients undergoing endoscopy. Therefore, the gold standard setting is not appropriate. 4. RUT is a well-known simple bedside H. pylori diagnosis method that can quickly confirm results. The authors compared Endofaster and RUT for detection time, but the time from gastric fluid collection to the final diagnosis should be provided to the reader. Authors should describe in the methods section how much additional time is required using the Endofaster. 5. In the methods section, the authors should describe whether to administer cimethicone or pronase before performing an endoscopy. 6. Although it is not an objective of the study, it would be very enriching if the authors included an economic analysis of the different techniques. It would be very interesting to know how much every diagnostic test costs. RUT is not an expensive diagnostic method for H. pylori detection. Please introduce and compare both two test methods. 7. Please provided the Figure 2 (flow chart) according to STARDS guideline 8. In the Abstract (Background & Aims, lines 68): Helicobacter pylori -> Helicobacter pylori (H. pylori).

**Reviewer 2:**

**Scientific Quality:** Grade B (Very good)
**Language Quality:** Grade B (Minor language polishing)
**Conclusion:** Minor revision

To Authors The theme is current and relevant, with adequate writing for all items: title, abstract, introduction, methods, results and conclusion. However, I recommend that the authors review the formatting of the tables and change the organization of Table 1, placing the items Overall, Patients without PPI therapy, and Patients with PPI therapy in the columns.

**Reviewer 1 Comments:**

***The authors have conducted this study to evaluate the role of a gastric juice analysis technique with Endofaster versus the conventional rapid urease test used to identify H pylori Infection. This study was fascinating and well-organized throughout the whole article. However, the biggest issue of this study is the gold standard definition. Endofaster is a method designed to overcome sampling error, and histology set as the gold standard is a representative invasive method that cannot avoid sampling error.***

We agree that biopsy protocols may incur in sampling errors but by applying the updated Sydney system the sampling error is minimized.

We have to disagree that the histological diagnosis according to Sydney cannot be used as reference standard method for *H. pylori* detection.

***Few issues need attention of the authors.***

***1. Regarding biopsy and histopathological confirmation, how were the pathologists blinded regarding the sample processing from Endofaster vs. RUT; please explain.***

Histopathological assessment was conducted as part of our routine clinical practice using gastric biopsies taken according to the updated Sydney system and current guidelines (2 from antrum, greater and lesser curvature, 1 from angulus and 2 from corpus, greater and lesser curvature). The pathologists were not blinded to the endoscopic findings. The endoscopic standard report presented to pathologists included a) an exhaustive image documentation on the appearance of the gastric mucosa, b) a diagnosis (i.e. suspected of *H. pylori*-associated gastritis) and c) requests for pathology investigations (i.e. grading of gastritis, *H. pylori* status, presence of gastric atrophy or intestinal metaplasia). In case of positive RUT reaction detected during endoscopy this result was included in the endoscopy report. Endofaster results were not reported to the pathologists.

In each biopsy sampling set the following stainings were performed: hematoxylin and eosin, periodic acid-Schiff (PAS) and a *H. pylori* specific staining (modified Giemsa staining).

This has been clarified in material and methods of the revised version of the manuscript.

***2. The fundamental concept of this study was that it was "gastric fluid" rather than a one-point pick-up to overcome the sampling error in diagnosing H. pylori infection. Therefore, the gold standard should be the method that does not have sampling issues like a urea breath test or stool antigen test. The authors performed only histopathologic diagnosis as the gold standard definition.***

***3. Regarding the above question, the importance of the gold standard should be further emphasized because the study was conducted targeting a relatively low H. pylori prevalence (35.3%) cohort. The H. pylori infection in the enrolled patients in this study was lower (29.2%) than the prevalence. The study cohort may generally show a higher infection rate than the prevalence, and this is because the H. pylori infection rate is likely to be high in patients undergoing endoscopy. Therefore, the gold standard setting is not appropriate.***

Thank you for this comment. We agree with the reviewer that histopathological diagnosis of *H. pylori* may suffer from potential sampling error due to the patchy distribution of the bacterium. However, by using the updated Sydney system based on biopsies from 5 different sites and applying different staining methods for *H. pylori* detection the accuracy of *H. pylori*-diagnosis by histology is not inferior to any non-invasive test (13C-UBT/STA). In support for the validity of histology as gold standard for *H. pylori* detection, we found also no indirect signs of *H. pylori*-gastritis (i.e neutrophils infiltration in the gastric mucosa) in the absence of *H. pylori*.

Ad point 3.

There are no current epidemiological data on the prevalence of *H. pylori* infection in Germany. Germany is a large country with a heterogeneous population of around 83 million people. Previous studies - mostly conducted in the nineties - showed marked differences in *H. pylori* prevalence between different German regions, with values raging from 21.0%-23.4% in Baden-Württemberg (West-Germany) vs. 44.4% in Saxony-Anhalt (East-Germany). Based on those studies the German prevalence is estimated at 35.3% (31.2% - 39.4%), Hooi et al. 2017. A recent study reported a prevalence of 28.9 % in an eastern region of Germany (Caspar et al. 2017).

The local prevalence of *H. pylori* infection where this study was conducted (Bavaria region) is not precisely known. We are currently addressing this in a prospective prevalence study (HelicoPTER-Study, https://mikrobio.med.tum.de/de/helicopter, Z Gastroenterol 2022; 60(08): e541, DOI: 10.1055/s-0042-1754863).

The overall *H. pylori*-prevalence in our cohort was 29.2%. and this data corresponds to many other data on *H. pylori*-prevalence in Europe. Particularly in the group of PPI users *H. pylori*-prevalence was strongly reduced (14.9%), and we agree that the use of histology - but this is the case with any other test except serology - may have its limitations in this group of patients.

We did now stress the argument of using histopathology as the gold standard as a possible limitation of the study, reporting it in the discussion section of the revised manuscript as follows:

"Using histology as the gold standard for *H. pylori*-diagnosis in a cohort with relatively low-prevalence of *H. pylori* may represent a further limitation of this study. Histopathological diagnosis of *H. pylori* may suffer from potential sampling error due to the patchy distribution of the bacterium. However, by using the updated Sydney system based on biopsies from 5 different sites and applying different staining methods for *H. pylori* detection the accuracy of *H. pylori*-diagnosis by histology is not inferior to any non-invasive test (13C-UBT/STA). In support for the validity of histology as gold standard for *H. pylori* detection, we found also no indirect signs of *H. pylori*-gastritis (i.e. neutrophils infiltration in the gastric mucosa) in the absence of *H. pylori*."

***4. RUT is a well-known simple bedside H. pylori diagnosis method that can quickly confirm results. The authors compared Endofaster and RUT for detection time, but the time from gastric fluid collection to the final diagnosis should be provided to the reader. Authors should describe in the methods section how much additional time is required using the Endofaster.***

Thank you for this important suggestion. After completion of gastric juice aspiration for Endofaster analysis the device provides a definitive diagnosis of *H. pylori* within 90 seconds. Considering that approximately 10-20 seconds (max. 30 seconds) are needed to aspirate the gastric juice through the scope a final *H. pylori* diagnosis is provided within the first 2 minutes from the beginning of the endoscopic procedure. Except for the time spent on the initial gastric juice collection no additional time is required for Endofaster use during the endoscopic procedure.

We added this information in the material and methods section of the revised manuscript.

***5. In the methods section, the authors should describe whether to administer cimethicone or pronase before performing an endoscopy.***

Thank you for this suggestion. We included the following information in the material and methods section of the revised the manuscript: "In order to avoid possible dilution of gastric juice prior to collection the administration of endoscopic premedications (i.e. dimethicone, N-acetylcysteine, pronase etc.) before endoscopy were not allowed. Furthermore, washing with water and cleaning the endoscopic lens were avoided until sampling was completed."

***6. Although it is not an objective of the study, it would be very enriching if the authors included an economic analysis of the different techniques. It would be very interesting to know how much every diagnostic test costs. RUT is not an expensive diagnostic method for H. pylori detection. Please introduce and compare both two test methods.***

At present we cannot perform an appropriate analysis of cost-effectiveness on the use of Endofaster in clinical practice compared to other tests for the following reasons:

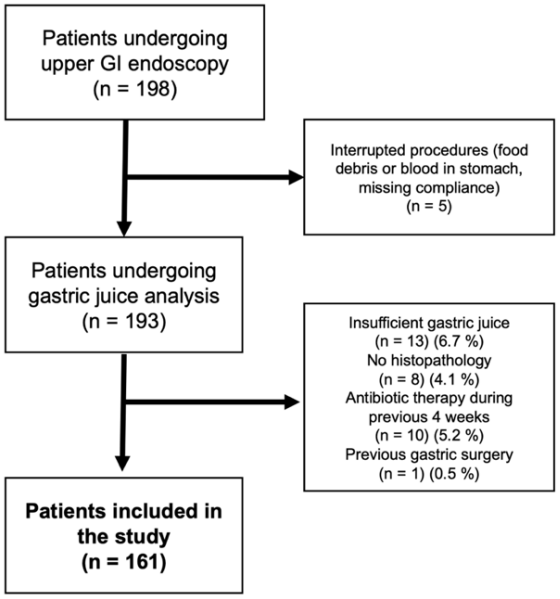
a) the present study was not designed to perform an economic analysis. For this purpose, a dedicated protocol using a larger cohort and detailed information on the costs for different histological assessments (depending on the clinical situation) are required;
b) the Endofaster device has not yet been launched on the German market, therefore reliable prices per single procedure are not available (the device was provided to our clinic for testing purpose free of charge during the recruitment period);
c) a comparison between the two methods (Endofaster vs RUT) on economic aspects would not fully exploit the diagnostic contribution of the Endofaster which allows a much more extensive analysis on the status of the gastric mucosa (i.e. pH, gastric atrophy, quantitative measurements related to *H. pylori*-presence), which is not obtained with the RUT.

Nevertheless, we agree on the importance of a study focusing on the economic impact of using Endofaster routinely and see this as a task for a future work.

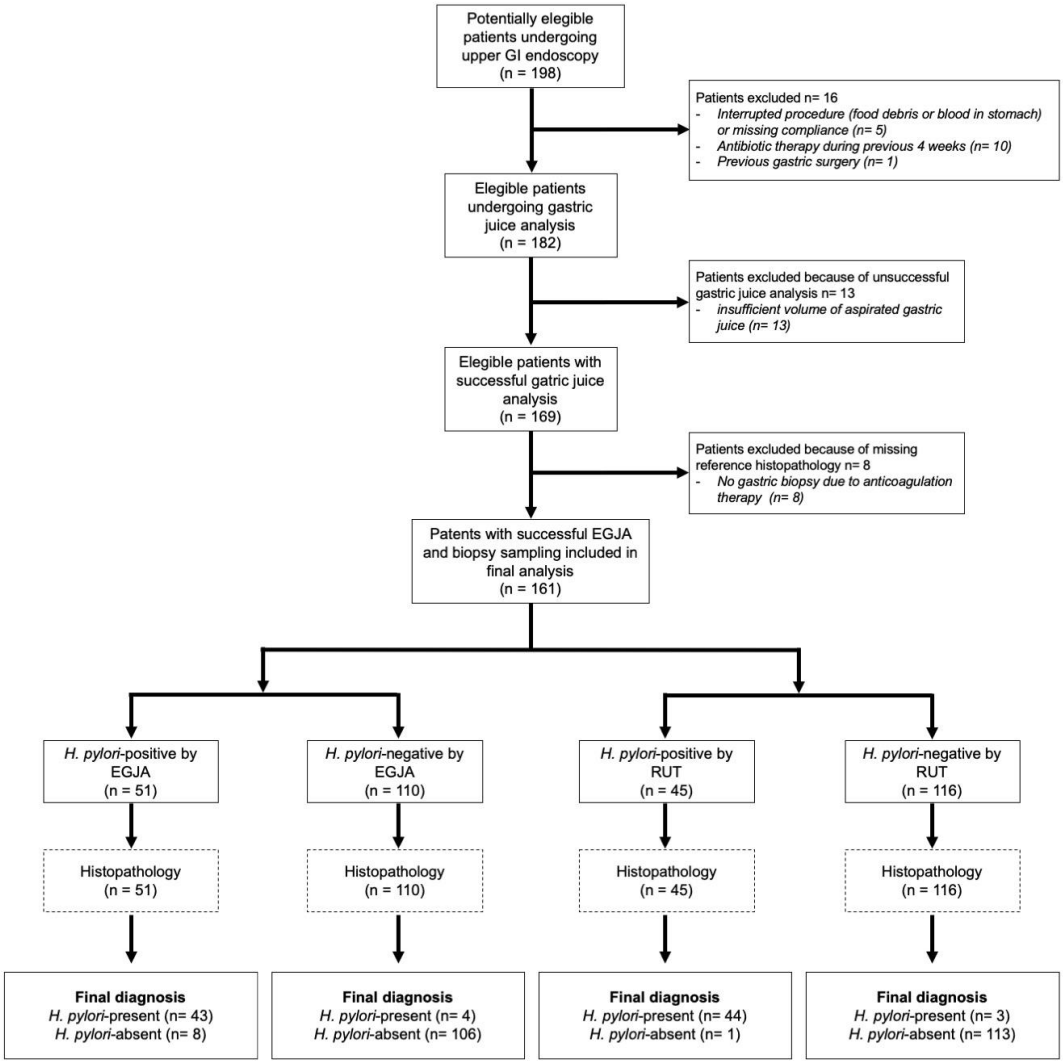
***7. Please provided the Figure 2 (flow chart) according to STARDS guideline.***

As suggested, we adapted Figure 2 according to the STARD guidelines.

Old Figure 2



New Figure 2




***8. In the Abstract (Background & Aims, lines 68): Helicobacter pylori -> Helicobacter pylori (H. pylori)..***

Thank you for this. We added the missing abbreviation as suggested.

**Reviewer 2 Comments:**

***To Authors The theme is current and relevant, with adequate writing for all items: title, abstract, introduction, methods, results and conclusion. However, I recommend that the authors review the formatting of the tables and change the organization of Table 1, placing the items Overall, Patients without PPI therapy, and Patients with PPI therapy in the columns.***

Thank you for your comment. As suggested, we reviewed the formatting of the tables and merged together table 2 and table 3.

Old Table 2:

| | Endofaster | | Rapid Urease Test | |
|---|---|---|---|---|
| | Value | 95% CI | Value | 95% CI |
| **Sensitivity** | 91.5% | 79.6% - 97.6% | 93.6% | 82.5% - 98.7% |
| **Specificity** | 93.0% | 84.6% - 96.9% | 99.1% | 95.2% - 100% |
| **PPV** | 84.3% | 73.3% - 91.3% | 97.8% | 86.2% - 99.7% |
| **NPV** | 96.4% | 91.2% - 98.6% | 97.4% | 92.7% - 99.1% |
| **Accuracy** | 92.6% | 87.3% - 96.1% | 97.5% | 93.8% - 99.3% |

Old Table 3:

| | Endofaster | | | | Rapid Urease Test | | | |
|---|---|---|---|---|---|---|---|---|
| | no PPI | | ongoing PPI Therapy | | no PPI | | ongoing PPI Therapy | |
| | Value | 95% CI | Value | 95% CI | Value | 95% CI | Value | 95% CI |
| **Sensitivity** | 97.3% | 85.8% - 99.9% | 70.0% | 34.8% - 93.3% | 97.3% | 85.8% - 99.9% | 80.0% | 44.4% - 97.5% |
| **Specificity** | 96.5% | 87.9% - 99.6% | 89.5% | 78.5% - 96.0% | 100% | 93.7% - 100% | 98.3% | 90.6% - 100% |
| **PPV** | 94.7% | 82.2% - 98.6% | 53.9% | 33.1% - 73.4% | 100% | - | 88.9% | 52.8% - 98.3% |
| **NPV** | 98.2% | 88.8% - 99.7% | 94.4% | 86.8% - 97.8% | 98.3% | 89.2% - 99.8% | 96.6% | 89.0% - 99.0% |
| **Accuracy** | 96.8% | 91.0% - 99.3% | 86.6% | 76.0% - 93.7% | 98.9% | 94.2% - 100% | 95.5% | 87.5% - 99.1% |

New Table 2:

| | Endofaster | | | | | | Rapid Urease Test | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Overall | | No PPI | | Ongoing PPI therapy | | Overall | | No PPI | | Ongoing PPI therapy | |
| | Value | 95% CI | Value | 95% CI | Value | 95% CI | Value | 95% CI | Value | 95% CI | Value | 95% CI |
| **Sensitivity** | 91.5% | 79.6% - 97.6% | 97.3% | 85.8% - 99.9% | 70.0% | 34.8% - 93.3% | 93.6% | 82.5% - 98.7% | 97.3% | 85.8% - 99.9% | 80.0% | 44.4% - 97.5% |
| **Specificity** | 93.0% | 84.6% - 96.9% | 96.5% | 87.9% - 99.6% | 89.5% | 78.5% - 96.0% | 99.1% | 95.2% - 100% | 100% | 93.7% - 100% | 98.3% | 90.6% - 100% |
| **PPV** | 84.3% | 73.3% - 91.3% | 94.7% | 82.2% - 98.6% | 53.9% | 33.1% - 73.4% | 97.8% | 86.2% - 99.7% | 100% | - | 88.9% | 52.8% - 98.3% |
| **NPV** | 96.4% | 91.2% - 98.6% | 98.2% | 88.8% - 99.7% | 94.4% | 86.8% - 97.8% | 97.4% | 92.7% - 99.1% | 98.3% | 89.2% - 99.8% | 96.6% | 89.0% - 99.0% |
| **Accuracy** | 92.6% | 87.3% - 96.1% | 96.8% | 91.0% - 99.3% | 86.6% | 76.0% - 93.7% | 97.5% | 93.8% - 99.3% | 98.9% | 94.2% - 100% | 95.5% | 87.5% - 99.1% |