Journal of
Clinical Medicine

MDPI

Article

Clinical Validation and Technological Evolution of a Biopsy-Free Endoscopic Urease Test for *Helicobacter pylori*: Two Independent Clinical Cohorts Across Two Device Generations

Anna Siwiec [1,*] and Dorota Bądziul [2]

[1] Department of Internal Medicine, Faculty of Medicine, University of Rzeszów, 35-959 Rzeszów, Poland
[2] Department of Medical Biology, University of Rzeszów, 35-959 Rzeszów, Poland; dbadziul@ur.edu.pl
* Correspondence: asiwiec@ur.edu.pl

Abstract

Background/Objectives: We assessed the sensitivity and specificity of a second-generation endoscopic urease test that samples gastric mucus without tissue excision. **Methods:** We analyzed archived diagnostic-classification data from two independent cohorts: 123 adults undergoing gastroscopy in a historical first-generation study and 200 symptomatic adults undergoing gastroscopy with independently indicated gastric biopsy in 2022 for the second-generation evaluation. Histopathology was the reference standard; Hepy/Pronto Dry was the paired comparator in the second-generation cohort. Wilson 95% confidence intervals (CIs) were calculated. **Results:** The second-generation test detected 75 of 94 histopathology-positive cases, giving a sensitivity of 79.8% (95% CI 70.6–86.7%); 19 cases were missed. No false-positive results occurred among 106 histopathology-negative participants, corresponding to an observed specificity of 100.0% (95% CI 96.5–100.0%). Comparator sensitivity and specificity were 78.7% and 90.6%; the paired specificity comparison favored the tube test ($p = 0.002$), while no statistically detectable sensitivity difference was found ($p = 1.000$). Historical first-generation sensitivity and specificity were 88.0% and 80.8%. Five minor mucosal abrasions were reported in the second-generation cohort; patients were observed for 30 min. **Conclusions:** The second-generation device showed high observed specificity but missed approximately one in five histopathology-positive infections. A negative result requires further evaluation when clinical suspicion persists. The assay avoids tissue excision for urease testing but remains an invasive endoscopic procedure. Overall clinical superiority was not established.

Keywords: *Helicobacter pylori*; rapid urease test; endoscopy; gastric mucus; biopsy-free sampling; diagnostic accuracy; medical device; patent

Check for updates

Academic Editors: Mattia Brigida and Gianluca Andrisani

Received: 21 September 2026
Revised: 7 November 2026
Accepted: 19 November 2026
Published: 26 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Helicobacter pylori infection remains a major global health problem and is causally associated with chronic gastritis, peptic ulcer disease, gastric mucosa-associated lymphoid tissue lymphoma, and gastric adenocarcinoma [1–3]. Reliable detection of active infection is therefore clinically important, and current recommendations emphasize accurate diagnosis and confirmation of eradication after treatment [1,2].

During upper gastrointestinal endoscopy, rapid urease tests (RUTs) provide a simple and inexpensive method for detecting *H. pylori*-associated urease activity. Conventional RUTs require gastric mucosal biopsy, and their performance may be influenced by bacterial density, biopsy site and size, acid-suppressive therapy, antimicrobial exposure, and

gastrointestinal bleeding [4–11]. Polish recommendations also support *H. pylori* testing in patients undergoing gastroscopy when clinically appropriate [12].

An endoscopic urease test that samples gastric mucus can avoid tissue acquisition solely for the urease assay. In this manuscript, biopsy-free refers specifically to sample acquisition without tissue excision; the procedure still requires upper gastrointestinal endoscopy and is invasive. A first-generation flexible tube prototype was developed for gastric mucus sampling and was subsequently protected as Pat.215889. Clinical experience motivated a mechanical redesign. The reagent solution impregnating the indicator paper was unchanged, whereas the second-generation construction introduced a protected, closable distal reaction chamber. The redesigned device was subsequently protected as Pat.249585.

The present study evaluates gastric mucus sampling as a potential alternative to tissue excision solely for urease testing during an already indicated upper gastrointestinal endoscopy. We assessed the second-generation device in 200 participants using histopathology as the reference standard and Hepy/Pronto Dry as the within-participant comparator; the independent 123-participant first-generation cohort provides historical, descriptive context. The clinical relevance is to characterize the diagnostic performance and limitations of this tissue-sparing sampling approach. Reductions in biopsy use, complications, procedure duration, or costs require separate evaluation.

2. Materials and Methods

2.1. Study Design and Ethics

This diagnostic accuracy study reanalyzes archived diagnostic-classification data from two independent clinical cohorts evaluating successive generations of a biopsy-free endoscopic urease test. The principal evaluation concerns the second-generation cohort; the first-generation cohort provides a separate historical description. Reporting was revised with reference to STARD 2015. Supplementary File S2 maps each item to the relevant manuscript location and specifies the information reported and any remaining gap. The timing of this reanalysis does not establish whether all aspects of the original data collection were planned before testing began. The documented prestudy procedure form and prespecified second-generation interpretation criteria are described below.

The first-generation study was approved by the Bioethics Committee of the Regional Medical Chamber in Kraków, Poland (Opinion No. 15/KBL/OIL/2011; 26 January 2011). The in vivo evaluation of the new tube-based urease test was separately approved by the same Bioethics Committee before the second-generation clinical study (Opinion No. 9/KBUOIU/2021; 18 March 2021). The 2021 opinion explicitly covered an in vivo study of a new tube-based urease test and stated that approval for the study site remained valid until expiry of the applicable insurance policy. The gastroscopy procedure form was submitted with both ethics committee applications before the respective studies began (Supplementary File S1). Written informed consent was obtained from participants.

Supplementary File S1 provides the available procedure form; it does not by itself document the complete recruitment and statistical analysis plans. Additional full protocol materials were unavailable for this analysis. Study registration status, a registry name, registration identifier, and registration date could not be established from the available materials.

2.2. First-Generation Cohort Characteristics

The first-generation prototype was evaluated in 123 adults undergoing upper gastrointestinal endoscopy. The clinical data originated from the doctoral study [13] and were reanalyzed for diagnostic accuracy. Detailed eligibility criteria, recruitment dates and setting, and the method used to identify and approach potentially eligible participants in this cohort were not available in the materials used for this analysis. The device

consisted of a flexible tube with a proximal handle, lateral openings at the distal end, and a reagent-bearing test element that was partly exposed beyond the distal tube edge (Figure 1) [14]. Histopathology was used as the reference standard and Pronto Dry as the biopsy-based comparator.

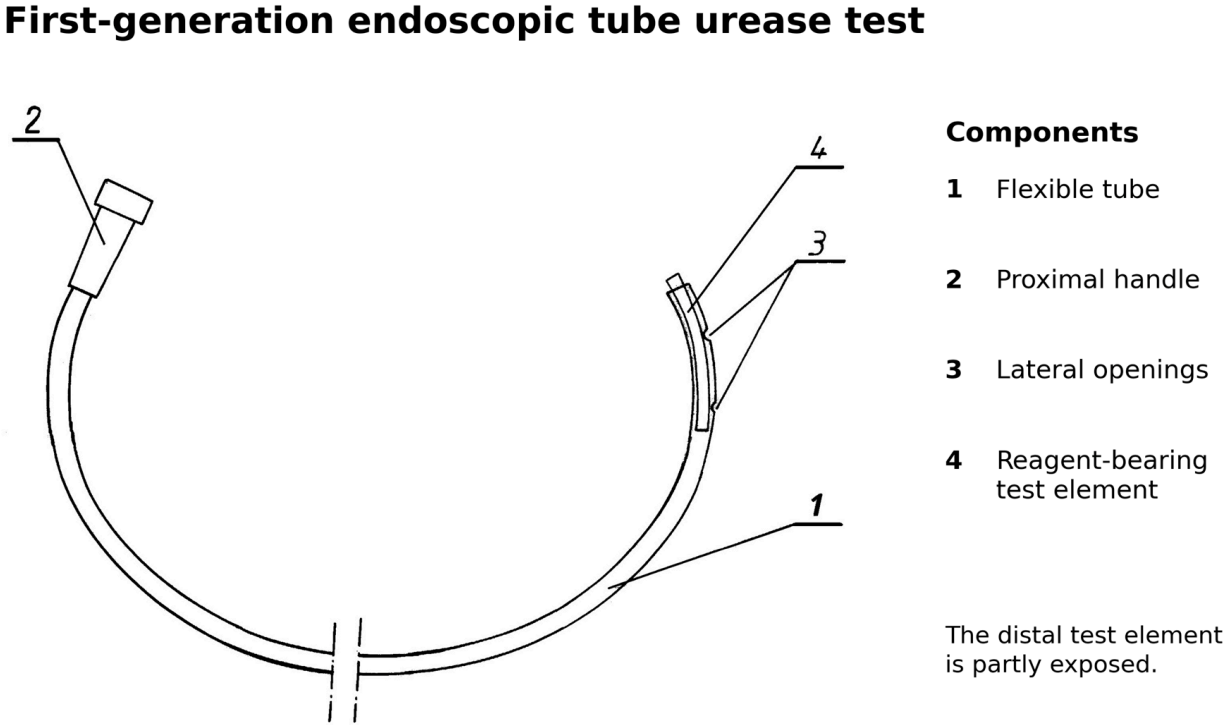

Figure 1. First-generation endoscopic tube urease test. The schematic identifies the flexible tube (1), proximal handle (2), lateral openings (3), and reagent-bearing test element (4), which is partly exposed at the distal end. Inventor-supplied patent schematic [14], with English labels added. Component numbering follows this schematic. Not to scale.

2.3. Second-Generation Device and Sampling Procedure

The second-generation device comprised a flexible plastic tube intended for contact with the human body, with a proximal handle and a fixed, domed distal end cap. A sliding cylinder inside the tube was connected to a polymer actuation line extending along the tube and beyond the handle. The reagent-bearing test element was housed in the compartment between the cylinder and the fixed end cap (Figure 2) [15]. The investigator identified the matrix as Whatman CH1 chromatography paper; the exact catalog designation was not available for verification. The paper was impregnated with urea, the pH indicator phenol red, and a buffer adjusted to a defined initial pH. The biochemical reagent system was unchanged between generations; the modification concerned the distal mechanical architecture.

During gastroscopy, the device was introduced through the working channel in the closed configuration and opened in the gastric antrum. Gastric mucus was allowed to enter the sampling compartment over a few seconds. The sliding cylinder was controlled using the internal polymer actuation line; the distal end cap remained fixed. After sampling, the device was returned to the closed configuration and withdrawn for external visual assessment. The available description does not establish the direction of cylinder movement required for reclosure. Urease-mediated urea hydrolysis generates ammonia, raising local pH and changing the phenol-red indicator color. No separate Instructions for Use were

created for this investigational tube device. The later patent documentation describes the construction but does not establish every detail of the prototype used in 2022.

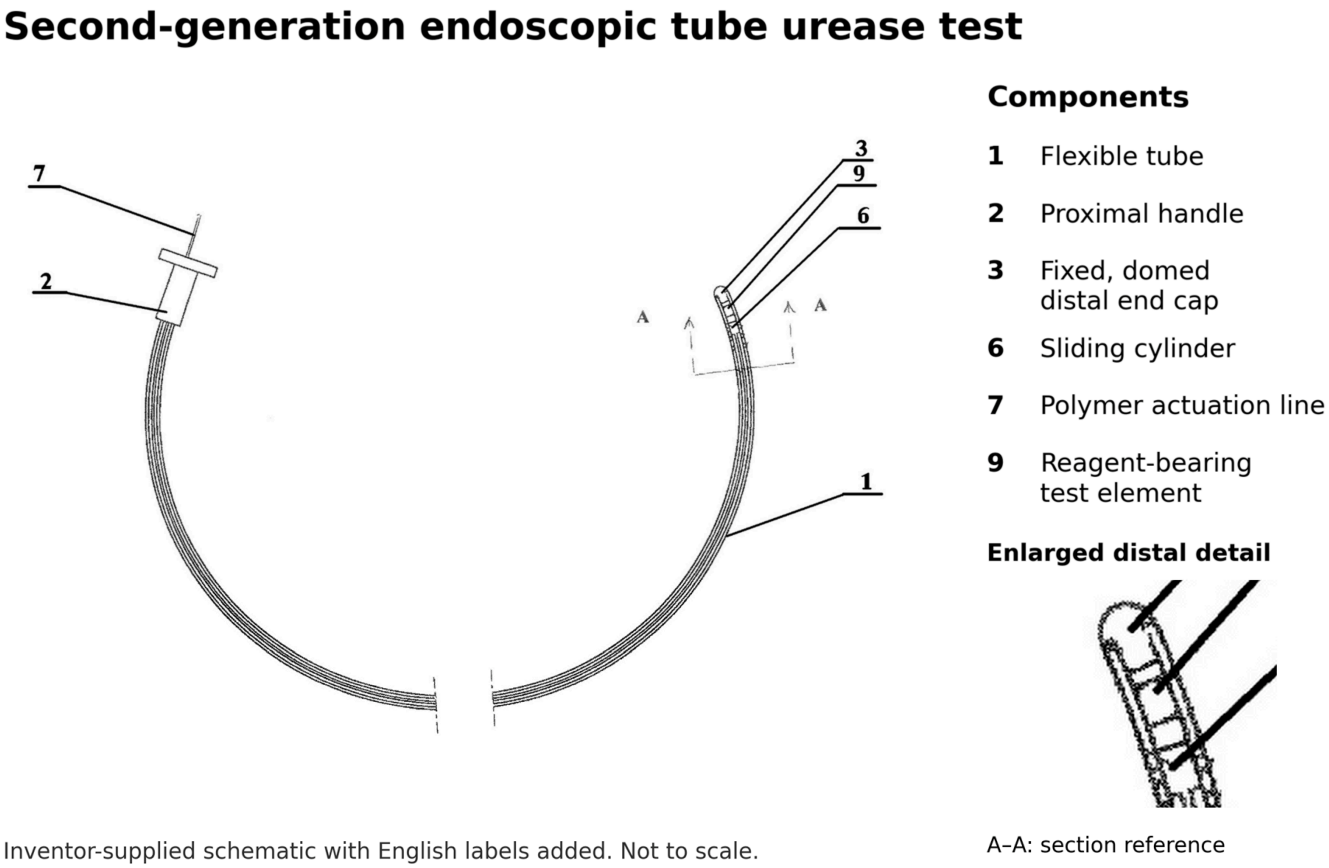

Figure 2. Second-generation endoscopic tube urease test. The schematic identifies the flexible tube (1), proximal handle (2), fixed, domed distal end cap (3), sliding cylinder (6), polymer actuation line (7), and reagent-bearing test element (9). The test element lies between the end cap and cylinder. The inset enlarges the same distal view; A–A denotes a section reference, not a direction of movement. Inventor-supplied schematic of the later patented construction [15], with English labels added. Not to scale.

Exact tube dimensions, reagent concentrations, and the numerical initial pH are not disclosed in this publication because they are retained as confidential proprietary know-how associated with the patented device. This restriction limits full independent reproduction of the device and reaction formulation. Other technical details, including compartment dimensions, the exact polymer identity, minimum compatible working-channel diameter, buffer composition, preparation and storage conditions, and intended single-use or reprocessing procedure, were unavailable in the materials used for this analysis. The precise time-zero definition, the actual sampling order relative to biopsy, operator experience, invalid- or equivocal-result criteria, and the full comparator sampling and interpretation procedures remained incompletely documented, limiting procedural reproducibility.

2.4. Second-Generation Clinical Cohort

The redesigned tube test was evaluated during 2022 in an independent cohort of 200 adult patients at a specialist medical practice and the Specialist Hospital in Brzozów, Poland. The source population comprised patients with upper gastrointestinal symptoms referred for clinically indicated upper gastrointestinal endoscopy. Inclusion in the diagnostic-accuracy analysis required an independent clinical indication for gastric mucosal biopsy and

histopathological examination. Histopathological sampling was therefore part of routine diagnostic work-up rather than tissue acquisition required for performance of the index tube test. At the initial eligibility assessment, patients were selected on the basis that they had not taken proton pump inhibitors (PPIs), H2-receptor antagonists, bismuth preparations, or antibiotics during the four weeks before testing. The available materials did not specify the method used to verify medication history. Data on previous *H. pylori* eradication therapy were also unavailable.

The 200 analyzed participants had final histopathology, tube-test, and comparator classifications. The process used to identify and approach potentially eligible participants was not documented. The numbers screened, potentially eligible, declining participation, or excluded under each criterion, and whether recruitment was consecutive, random, or based on another sampling strategy were unavailable. Consequently, the screening-to-inclusion proportion could not be calculated. The available participant disposition is shown in Figure 3.

In the second-generation cohort, five gastric mucosal biopsies were obtained for clinically indicated histopathological examination: two from the antrum, two from the corpus, and one from the incisura angularis. Specimens from different anatomical sites were placed in separate containers and immediately fixed in 10% formalin. Sections were examined using hematoxylin and eosin (H&E) and Giemsa staining. Three pathologists participated in assessment. They received biopsy specimens accompanied by the clinical referral diagnosis (e.g., gastritis) and interpreted the specimens without knowledge of the tube-test or conventional biopsy-based RUT results. Immunohistochemistry was not performed. *H. pylori* density was graded semiquantitatively according to the Updated Sydney System: 0, absent; 1, mild (+); 2, moderate (++); and 3, marked (+++) [16]. The grades represented scattered organisms or small groups (+), increased numbers and distinct clusters (++), and dense, extensive colonization (+++). These categories describe bacterial density separately from inflammatory activity.

For the present second-generation analysis, histopathology was selected as the reference because it was available from independently indicated gastric biopsies in all 200 participants. This provided the same patient-level reference for evaluation of both urease tests. Recorded patient-level grade 0 was classified as *H. pylori*-negative and grades 1–3 as positive. Histopathology was not treated as error-free, and no composite adjudication of discordant results was available.

The available documentation did not identify the histopathology laboratory or establish gastrointestinal subspecialty expertise, allocation of cases among pathologists, procedures for repeat reading or resolution of disagreement, or the rule for combining grades from different biopsy sites into the recorded patient-level grade. The original timing of specification of that aggregation rule was also unavailable. Independent assessment of every case by all three pathologists was not assumed.

Mucus for the tube test and gastric biopsies for histopathology and the biopsy-based Hepy/Pronto Dry RUT were obtained during the same endoscopic examination. Histopathological interpretation followed specimen processing and was not performed during endoscopy. The exact sampling order, intervals between sampling steps, and any interventions between those steps were not documented. The available clinical variables included sex, age, and the final results of histopathology and both urease tests.

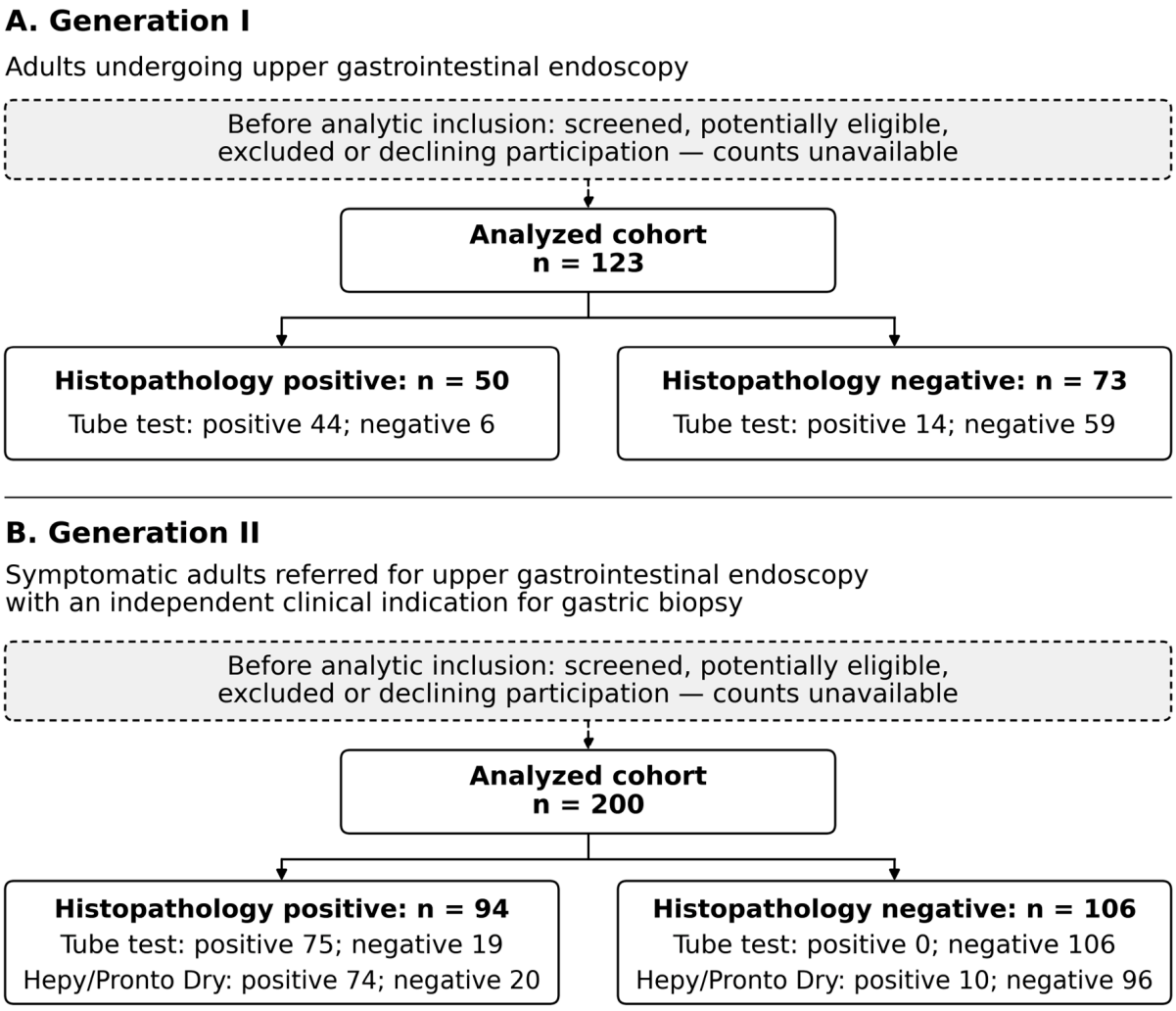

Figure 3. Available participant disposition and final diagnostic classifications. The two cohorts are independent. Dashed boxes indicate unavailable counts before inclusion in the present analysis. In the second-generation cohort, both urease tests were evaluated in the same 200 participants; the displayed results are marginal counts for each test against histopathology and do not reconstruct the full joint distribution of the two urease tests.

2.5. Test Interpretation

In the second-generation cohort, tube-test results were assessed visually against a color scale at 1, 5, 10, 30, and 60 min. The qualitative positivity criterion was a visible change from yellow to a color ranging from pink to magenta (amaranth); absence of this change at the final 60 min assessment was interpreted as negative. The positivity criterion and final 60 min negative-result cutoff were prespecified before the first participant was tested. Interpretation was qualitative; color intensity was not used to estimate bacterial density.

The written gastroscopy procedure form was submitted with the ethics committee application for each study before that study began; its English translation is provided in Supplementary File S1. In the second-generation cohort, endoscopy nurses and physicians received training. The conventional biopsy-based RUT was read by a different physician from the physician reading the tube test. The physician interpreting the tube test was blinded to the conventional biopsy-based RUT and histopathology results. Blinding of the comparator RUT reader and the clinical information available to either urease-test reader were not documented. Detailed training content and competency assessment, total reader numbers and case allocation, procedures for equivocal results, inter-reader agreement, and

the precise event defining time zero were also unavailable. These second-generation reader details were not assumed to apply to the historical first-generation cohort.

Individual reaction times and verified per-time-point numerators and denominators were unavailable. An archived percentage-only kinetic summary could not be reconciled uniquely with the final diagnostic counts and was excluded from the quantitative results. The final diagnostic classifications were retained; no time-specific sensitivity, cumulative-positivity estimates, or formal time-to-positivity analysis was performed.

2.6. Patent History

The first-generation design was filed as P.392212 on 24 August 2010 and subsequently covered by Polish Patent PL215889B1 [14]. The redesigned second-generation construction was filed as P.445161 on 6 June 2023 and is associated with Polish Patent Pat.249585; the application announcement was published on 9 December 2024 [15]. The first design had a partly exposed distal test element. The later construction housed the element inside the tube between a fixed distal end cap and a sliding cylinder controlled by an actuation line. The application postdated the 2022 clinical evaluation and is cited as technical documentation. The distinction between generations concerns the mechanical construction rather than a change in the urease indicator chemistry.

2.7. Statistical Analysis

The participant was the unit of analysis. The main diagnostic outcomes were the sensitivity and specificity of the second-generation tube test relative to histopathology. No a priori sample-size calculation was performed for either cohort. The present analyses used the available diagnostic-classification data for 123 first-generation and 200 s-generation participants; confidence intervals describe the resulting precision. The original recruitment targets and the reasons for ending recruitment could not be established from the available materials.

Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy were calculated from the final diagnostic-classification counts using histopathology as the reference. Sensitivity used the histopathology-positive denominator, specificity the histopathology-negative denominator, PPV, the test-positive denominator, and NPV, the test-negative denominator; accuracy used all analyzed participants in the relevant cohort. Wilson 95% confidence intervals (CIs) were used for proportions. Positive and negative likelihood ratios (LR+ and LR−, respectively) were calculated where estimable. Zero observed false-positive results were reported with the corresponding specificity CI. No equivalence or noninferiority margin was specified; a nonsignificant comparison was not interpreted as equivalence.

The present calculations used the reported final binary diagnostic-classification counts. No missing results were imputed in this reanalysis. Final tube-test and histopathology counts were available for all 123 analyzed first-generation participants, and final counts for both urease tests and histopathology were available for all 200 analyzed second-generation participants. The available materials did not establish the frequency or original handling of invalid or equivocal results, or the number of potentially eligible participants excluded because of missing test results before the final analysis sets were assembled. Availability of complete final counts was not interpreted as evidence that no earlier missing or indeterminate results occurred.

The cross-generation comparison in this reanalysis was descriptive and exploratory because the cohorts were independent and studied in different periods; no confirmatory between-generation superiority analysis was performed. The available aggregate data did not permit adjustment for differences in participant selection, case mix, testing conditions,

operator experience, or recording procedures across the study periods. Histological-grade-specific sensitivity could not be estimated without joint counts linking tube-test results to the grades. The original prespecification status of analyses of accuracy variation was not established.

Within the second-generation cohort, two-sided exact McNemar comparisons were derived where identifiable from the available marginal counts. Among histopathology-negative participants, all tube results were negative, fixing the discordant counts at 10 Hepy-positive/tube-negative pairs and zero pairs in the opposite direction. Among histopathology-positive participants, 75 tube-positive versus 74 Hepy-positive results imply discordant counts differing by one; every feasible pairing therefore gives a two-sided exact p value of 1.000. The complete two-by-two counts for each test against histopathology were available. However, the full joint tube-test versus Hepy/Pronto Dry cross-tabulation within the histopathology-positive stratum was unavailable. Individual agreement patterns and the identities of jointly missed cases were therefore not reconstructed.

2.8. Assessment of Adverse Events

In the second-generation cohort, patients were observed for 30 min after endoscopy. This patient-observation period was distinct from the 60 min tube-test reading window. The supplied procedure form contains a field for procedural complications and instructions to monitor pain, bleeding, and nausea after endoscopy. The safety information available for this analysis concerned minor mucosal abrasions, damage to the endoscope working channel, device retention within the endoscope, and technical failures. The completeness and timing of event recording in individual participants and information on follow-up beyond the first 30 min remained unavailable; this analysis therefore provides a limited descriptive safety assessment.

3. Results

Figure 3 summarizes the available participant disposition and final diagnostic classifications in the two independent cohorts. The analyzed samples comprised 123 first-generation and 200 s-generation participants. Counts for screening, eligibility assessment, exclusion, or refusal before analytical inclusion were unavailable and were not reconstructed.

3.1. First-Generation Cohort

In the 123-patient first-generation cohort, histopathology was positive in 50 and negative in 73 participants. The tube test yielded 44 true-positive, 6 false-negative, 59 true-negative, and 14 false-positive results. Sensitivity was 88.0% (95% CI 76.2–94.4%), specificity 80.8% (70.3–88.2%), PPV 75.9% (63.5–85.0%), NPV 90.8% (81.3–95.7%), and accuracy 83.7% (76.2–89.2%).

3.2. Second-Generation Cohort Characteristics

The second-generation cohort included 200 patients: 94 women (47.0%) and 106 men (53.0%). Mean age was 51.4 ± 18.1 years, median age 51.5 years (interquartile range 33–64), and range 21–90 years. Histopathology was positive for *H. pylori* in 94 patients (47.0%) and negative in 106 (53.0%). Among the 94 histopathology-positive participants, bacterial density was mild (+) in 45 (47.9%), moderate (++) in 33 (35.1%), and marked (+++) in 16 (17.0%). These grades describe the spectrum of bacterial colonization, not clinical disease severity. Cohort characteristics are summarized in Table 1.

Table 1. Baseline characteristics of the second-generation clinical cohort. Sex and overall histopathology percentages use the full cohort denominator of 200. Bacterial-density grade percentages use the 94 histopathology-positive participants; the corresponding $n/94$ denominators are shown explicitly. SD, standard deviation; IQR, interquartile range. +++ amaranth; ++ dark pink; + light pink.

Characteristic	Second-Generation Cohort ($n = 200$)
Female sex	94 (47.0%)
Male sex	106 (53.0%)
Age, mean ± SD	51.4 ± 18.1 years
Age, median (IQR)	51.5 (33–64) years
Age range	21–90 years
Histopathology *H. pylori*-positive	94 (47.0%)
Histopathology-negative	106 (53.0%)
H. pylori density grade +	45/94 (47.9%)
H. pylori density grade ++	33/94 (35.1%)
H. pylori density grade +++	16/94 (17.0%)

The 106 reference-negative participants came from the same symptomatic endoscopy population. Their alternative endoscopic or histopathological diagnoses were unavailable in the data used for this analysis; *H. pylori*-negative classification should not be interpreted as absence of other gastric disease. Detailed symptom distributions and comparable demographic data for the historical cohort were also unavailable.

3.3. Diagnostic Accuracy in the Second-Generation Cohort

The second-generation tube test yielded 75 true-positive, 19 false-negative, 106 true-negative, and no false-positive results relative to histopathology. Sensitivity was 79.8% (75/94; 95% CI 70.6–86.7%); 19/94 reference-positive participants (20.2%) tested negative. No tube-test-positive result occurred among the 106 histopathology-negative participants, giving an observed specificity of 100.0% (95% CI 96.5–100.0%). Of the 125 participants with a negative tube test, 19 (15.2%) were histopathology-positive; the corresponding NPV was 84.8% (95% CI 77.5–90.0%). Hepy/Pronto Dry yielded 74 true-positive, 20 false-negative, 96 true-negative, and 10 false-positive results. Diagnostic performance is summarized in Table 2.

The data available for this analysis did not link individual tube-test results to histological bacterial-density grades. Therefore, the distribution of the 19 false-negative results across the +, ++, and +++ categories and sensitivity within each category could not be determined.

Among histopathology-positive participants, observed sensitivity was 79.8% for the tube test and 78.7% for Hepy/Pronto Dry, with no statistically detectable paired difference (two-sided exact McNemar $p = 1.000$). This does not establish equivalence or identical patient-level results. Among histopathology-negative participants, all 10 discordant pairs favored the tube test: Hepy/Pronto Dry was positive while the tube test was negative; no discordant pair favored Hepy/Pronto Dry. The observed specificity difference was 9.4 percentage points (two-sided exact McNemar $p = 0.002$). These comparisons concern agreement with the histopathological reference.

Table 2. Diagnostic performance relative to histopathology in the second-generation cohort ($n = 200$). Proportions are reported as n/N (percentage), followed by Wilson 95% confidence intervals. Likelihood ratios are reported separately.

Parameter	Second-Generation Tube Test	Hepy/Pronto Dry
TP/FN/TN/FP	75/19/106/0	74/20/96/10
Sensitivity	75/94 (79.8%) 95% CI 70.6–86.7	74/94 (78.7%) 95% CI 69.4–85.8
Specificity	106/106 (100.0%) 95% CI 96.5–100.0	96/106 (90.6%) 95% CI 83.5–94.8
PPV	75/75 (100.0%) 95% CI 95.1–100.0	74/84 (88.1%) 95% CI 79.5–93.4
NPV	106/125 (84.8%) 95% CI 77.5–90.0	96/116 (82.8%) 95% CI 74.9–88.6
Accuracy	181/200 (90.5%) 95% CI 85.6–93.8	170/200 (85.0%) 95% CI 79.4–89.3
LR+	No finite estimate (0 observed FP)	8.34
LR−	0.20	0.23

Abbreviations: TP, true-positive; FN, false-negative; TN, true-negative; FP, false-positive; PPV, positive predictive value; NPV, negative predictive value; CI, confidence interval; LR+, positive liklihood ratio; LR−, negative likelihood ratio.

3.4. Performance Across Device Generations

The mechanical construction differed between generations, while the reagent solution was unchanged. In the historical first-generation cohort, observed specificity was 80.8% (95% CI 70.3–88.2%), with 14 false-positive results relative to histopathology. In the independent second-generation cohort, observed specificity was 100.0% (95% CI 96.5–100.0%), with no false-positive results among 106 reference-negative participants. Observed accuracy was 83.7% and 90.5%, respectively. These cross-generation differences are descriptive and cannot establish superiority of the redesigned device or a causal effect of its mechanical construction, because participant selection, case mix, and study circumstances may also differ. Table 3 summarizes the construction differences alongside the separate cohort results.

Table 3. Reported construction and observed diagnostic performance of the two device generations. Values in parentheses are Wilson 95% confidence intervals; the clinical cohorts were independent and not paired.

Characteristic	Generation I ($n = 123$)	Generation II ($n = 200$)
Distal construction and test element	Lateral distal openings; reagent-bearing element partly exposed beyond the tube edge	Fixed domed end cap and internal sliding cylinder; reagent-bearing element housed between them
Actuation and sampling	A controlled opening and reclosure sequence was not specified in the available description	Cylinder controlled by an internal polymer actuation line; closed insertion, opening for mucus entry, and reclosure before withdrawal
Biochemical reagent solution	Urea, phenol red, and buffer	Same reagent solution as Generation I
TP/FN/TN/FP	44/6/59/14	75/19/106/0
Sensitivity	88.0% (76.2–94.4)	79.8% (70.6–86.7)
Specificity	80.8% (70.3–88.2)	100.0% (96.5–100.0)
PPV	75.9% (63.5–85.0)	100.0% (95.1–100.0)

Table 3. *Cont.*

Characteristic	Generation I (n = 123)	Generation II (n = 200)
NPV	90.8% (81.3–95.7)	84.8% (77.5–90.0)
Accuracy	83.7% (76.2–89.2)	90.5% (85.6–93.8)
Patent	Pat.215889	Pat.249585

Abbreviations: TP, true-positive; FN, false-negative; TN, true-negative; FP, false-positive; PPV, positive predictive value; NPV, negative predictive value. Construction descriptions are based on the reported device information and patent schematics [14,15]; the later patent does not establish every detail of the 2022 prototype. Differences in clinical performance cannot be attributed causally to the mechanical redesign.

3.5. Adverse Events and Device Performance

Five cases of minor mucosal abrasion were reported in the 200-participant second-generation cohort (5/200, 2.5%). No damage to the endoscope working channel, device retention within the endoscope, or technical failure was reported. Patients were observed for 30 min after the procedure. The locations and clinical course of the abrasions, associated bleeding, and any required intervention were not available for this analysis. Information was insufficient to assign event counts for aspiration, device breakage, retention outside the endoscope, or delayed device-related events.

4. Discussion

The independent 200-participant cohort provides evidence of high observed specificity together with a clinically relevant sensitivity limitation. The second-generation tube test detected 75 of 94 histopathology-positive cases, giving a sensitivity of 79.8% (95% CI 70.6–86.7%); 19 cases (20.2%), approximately one in five, were missed. No false-positive tube-test results were observed among 106 histopathology-negative participants. The resulting specificity estimate of 100.0% (95% CI 96.5–100.0%) describes this sample and does not establish an intrinsic property of invariably perfect specificity. Hepy/Pronto Dry produced 10 positive results among histopathology-negative participants, and the paired specificity comparison favored the tube test.

The redesign changed the distal test chamber without changing the reagent solution. The protected configuration is intended to limit unintended exposure during insertion and to permit contact with gastric mucus at the target site. The integrated sampling and reaction compartment may also help retain reaction products near the pH indicator. These are mechanistic hypotheses: contamination, ammonia retention, and their effects on diagnostic performance were not directly measured. The independent cohorts do not establish that the mechanical modification caused the difference in observed specificity. Comparable information on patient characteristics, testing conditions, operator experience, and recording procedures was insufficient to quantify their contributions to the observed differences.

The second-generation tube test and Hepy/Pronto Dry had similar observed sensitivities (79.8% and 78.7%), but the nonsignificant paired comparison does not demonstrate equivalence. The NPV of 84.8% means that 19 of 125 tube-test-negative participants (15.2%) had positive histopathology. A negative result therefore has limited value for excluding infection in this cohort. PPV and NPV also depend on infection prevalence, which was 47.0% against the reference standard in this selected population.

Low organism density or sampling variability could plausibly contribute to false-negative urease results [5,11]. These explanations were not investigated in the 19 missed cases, and the available data do not establish whether false-negative results were concentrated in any histological density category.

When clinical suspicion persists after a negative tube test, infection should be assessed with a validated test for active *H. pylori*, such as a urea breath test or a laboratory-based monoclonal stool antigen assay; biopsy-based testing remains appropriate when tissue sampling is independently indicated [1,2,12]. Follow-up testing should respect the relevant preparation and medication-withdrawal guidance. In particular, PPIs should usually be withheld for 2 weeks and antibiotics and bismuth for 4 weeks; P-CAB exposure should also be considered when planning testing [2,12]. This is a proposed clinical pathway informed by guidance, not a follow-up strategy whose outcomes were measured in the present study.

The reported reading window extended to 60 min. The scheduled assessments at 1, 5, 10, 30, and 60 min describe when the tube test was assessed; they do not establish the observed distribution of time to positivity. The archived aggregate timing percentages could not be reconciled with verified time-point counts and were excluded from the quantitative results. Consequently, the present analysis does not establish the proportion of positive tests detectable at any earlier time point, a comparative advantage in time to detection, or a reduction in endoscopy duration.

The tube test obtains material for the urease assay without excising gastric tissue. All participants in the validation cohort nevertheless underwent biopsy for independently indicated histopathology. The study therefore does not demonstrate a reduction in the total number of biopsies obtained in clinical practice. Avoiding an additional biopsy solely for RUT is a possible application, but the present population does not directly establish performance or benefit in patients declining biopsy or without an indication for tissue sampling. Diagnostic mucosal biopsy is generally classified as a low-bleeding-risk procedure [17,18]. Five minor mucosal abrasions were reported in the second-generation cohort, demonstrating that sampling without tissue excision does not eliminate the possibility of mucosal injury. The absence of reported endoscope-channel damage, retention within the endoscope, and technical failure does not establish an incremental safety benefit. Histopathology should still be performed whenever clinically indicated.

The described tube-sampling sequence involves one insertion and withdrawal through the working channel. Comparative procedure duration, numbers of instrument passages, patient comfort, costs, training time and resources, and downstream diagnostic or treatment outcomes were not evaluated. Technical failures were reported descriptively for the tube test rather than evaluated as a comparative endpoint. A conventional RUT workflow also depends on its sampling protocol and accessories. The present results therefore support a diagnostic-accuracy comparison with the selected comparator, but do not establish overall procedural, clinical, or economic superiority.

The intended setting is upper gastrointestinal endoscopy that is already clinically indicated. The procedure remains invasive despite avoiding tissue excision for the urease assay. It does not create an indication for endoscopy, replace recommended non-invasive testing when endoscopy is unnecessary, or replace histopathology when mucosal assessment is required [1,2,12].

Gastric surface and mucus sampling without tissue excision is an established area of investigation. Soh et al. compared swab-based and biopsy-based RUT in 138 participants [19]. Yoshikawa et al. studied gastric mucus swabbing across four institutions in 94 participants and reported sensitivity of 93.3% and specificity of 92.2% against a urea breath test [20]. In 216 previously treated patients, Noh et al. reported sensitivity of 81.2% and specificity of 91.2% for a sweeping method [21]. Bang et al. subsequently reported positive brush-RUT results in all 20 patients with confirmed infection in a pilot study; the absence of uninfected controls prevented estimation of specificity [22]. Differences in populations, sampling sites, reading windows, and reference standards preclude direct comparisons with the present cohort.

The contribution of the present study is the clinical evaluation of a particular integrated tube architecture with a protected, closable distal sampling and reaction compartment. The reagent-bearing element remains within the device, whereas swab and brush methods may transfer collected material to a separate RUT [20,22]. Protection during withdrawal is not unique to the tube concept: the brush in the study by Bang et al. was retracted into a sheath [22]. The specific integrated construction is therefore the focus of the technological contribution; superiority in contamination control, diagnostic performance, safety, or clinical utility over these other approaches has not been demonstrated.

The two cohorts document an initial prototype and a subsequent mechanical redesign while preserving the biochemical reaction system. The second-generation clinical evaluation took place in 2022, before the later patent application. The analytical performance of a separate dry biopsy-based Helibac RUT has been reported by Anna Siwiec and Agnieszka Siwiec [23]; that publication concerns a different test format under controlled laboratory conditions and does not establish clinical performance or reading criteria for the investigational tube device.

Strengths and Limitations

Strengths include two independent clinical cohorts and paired evaluation of the second-generation tube test and Hepy/Pronto Dry in the same 200 participants. In the second-generation cohort, the histopathological reference was based on five gastric biopsies from the antrum, corpus, and incisura angularis with H&E and Giemsa staining. Histopathological interpretation in this cohort was performed without knowledge of either urease-test result. The generations are analyzed separately. Several limitations restrict interpretation and generalizability.

First, the cohorts were independent and evaluated different generations in different periods. Their differences cannot be attributed causally to device construction, and the available data did not permit adjustment for differences in case mix, testing conditions, or operator experience. The second-generation cohort comprised symptomatic adults already selected for endoscopy and independently indicated biopsy. This biopsy requirement may select a different clinical spectrum from patients needing only assessment of infection. The eligibility requirement of no use of PPIs, H2-receptor antagonists, bismuth preparations, or antibiotics during the four weeks before testing further restricts generalizability. Detailed symptom categories, screening and exclusion counts, recruitment sequence, and the method of verifying medication history were unavailable. Figure 3 presents the known analyzed cohorts and final classifications; it does not resolve the missing earlier recruitment counts. The absence of a screening denominator and documented recruitment sequence limits assessment of selection bias and representativeness; the direction and magnitude of any resulting bias cannot be quantified. Performance in patients without an indication for biopsy or declining biopsy was not directly evaluated. The comparison was not designed to establish equivalence or noninferiority. No a priori sample-size calculation was performed for either cohort, and the reported confidence intervals describe the precision achieved with the available samples.

Second, histopathology was the sole reference standard. In the second-generation cohort, pathologists received the clinical referral diagnosis and interpreted specimens without knowledge of either urease-test result. Immunohistochemistry was not performed. Sampling variability and low organism density may affect histopathology as well as urease testing, and no composite assessment to adjudicate discordant cases was available. Apparent false-positive RUT results should therefore be interpreted relative to histopathology. Only 106 reference-negative participants contributed to the specificity estimate, and zero observed

false positives does not exclude a nonzero false-positive rate in the target population. The 19 false negatives also limit the ability to rule out infection.

Third, previous *H. pylori* eradication data were unavailable, and diagnostic performance could not be stratified by linked histological bacterial-density grades. The aggregate grade counts do not show whether the 19 false-negative tube results were concentrated in the +, ++, or +++ category. Previous treatment and medication exposure can influence urease-test performance [2,5,21], but they cannot be assigned as explanations for missed cases in this dataset. The present analysis does not validate color intensity as a measure of bacterial density or infection severity. Missing details of specimen interpretation, pathology case allocation, blinding of the comparator RUT reader, comparator sampling, and operational procedures limit assessment of bias. The gastroscopy procedure form was submitted with both ethics committee applications before the respective studies began. Staff training, use of a color scale, and different physician readers were confirmed. In the second-generation cohort, tube-test interpretation was blinded to the comparator RUT and histopathology results. The positivity criterion and final 60 min negative-result cutoff were prespecified before the first participant was tested. Details of training and competency assessment remained undocumented, limiting assessment of interpretation bias and reproducibility. The component architecture and qualitative reagent composition are described. Exact tube dimensions, reagent concentrations, and the numerical initial pH are intentionally withheld as confidential proprietary know-how associated with the patented device, limiting full independent reproduction of the device and reaction formulation. Other unavailable technical details include exact material specifications and preparation and storage conditions. These reporting limitations are relevant to STARD principles [24,25].

Finally, the kinetic percentages were excluded because their underlying counts and denominators could not be verified; individual reaction times were also unavailable. A reduction in biopsy use, comparative procedure duration, patient comfort, costs, training time and resources, comparative technical failure rates, and downstream diagnostic or treatment outcomes were not evaluated. Five minor mucosal abrasions were reported, and observation lasted only 30 min. The limited safety information does not establish the absence of other or delayed adverse events. No overall safety advantage or reduction in biopsy-related harm was demonstrated. Protection against contamination or loss of ammonia remains a mechanistic hypothesis. Independent prospective validation with controlled device specifications and a prespecified clinical and reading protocol is needed.

5. Conclusions

In this 200-participant cohort, the second-generation endoscopic tube test detected 75 of 94 histopathology-positive cases, giving a sensitivity of 79.8% (95% CI 70.6–86.7%). No false-positive results were observed among 106 histopathology-negative participants, corresponding to an observed specificity of 100.0% (95% CI 96.5–100.0%). Observed sensitivity was similar to Hepy/Pronto Dry, while the paired specificity comparison favored the tube test. A negative result does not reliably exclude infection and warrants further testing when clinical suspicion persists. The device permits gastric mucus sampling without tissue excision for the urease assay but remains an invasive endoscopic procedure; five minor mucosal abrasions were reported. Independent prospective evaluation is needed to establish its clinical utility, safety, and practical advantages.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/jcm15239277/s1, Supplementary File S1 contains an English translation of the investigator-supplied gastroscopy procedure form. Supplementary File S2 provides an item-by-item STARD 2015 reporting map with manuscript locations, the information reported, and specific remaining gaps.

Author Contributions: Conceptualization, A.S.; methodology, A.S.; validation, A.S.; formal analysis, A.S.; investigation, A.S.; resources, A.S.; data curation, A.S.; writing—original draft preparation, A.S.; writing—review and editing, A.S. and D.B.; visualization, A.S.; supervision, A.S.; project administration, A.S. All authors have read and agreed to the published version of the manuscript.

Funding: The historical clinical program was conducted within the author's academic research. The archived materials available for this analysis do not identify an external funding source. A complete account of financial or in-kind support and the role of any funder was not available; absence of a named source in the archive is not treated as confirmation that no support was received.

Institutional Review Board Statement: The study program was conducted in accordance with applicable ethical requirements and the principles of the Declaration of Helsinki. The first-generation clinical study was approved by the Bioethics Committee of the Regional Medical Chamber in Kraków, Poland (Opinion No. 15/KBL/OIL/2011; 26 January 2011). The second-generation in vivo study of the new tube-based urease test was separately approved by the same Committee (Opinion No. 9/KBUOIU/2021; 18 March 2021), before recruitment of the 2022 cohort. The 2021 opinion stated that approval for the study site remained valid until expiry of the applicable insurance policy.

Informed Consent Statement: Written informed consent was obtained from study participants.

Data Availability Statement: The aggregate diagnostic-classification counts used in this analysis are reported in Tables 2 and 3 and Figure 3. The clinical data are retained in the authors' archived study materials. A de-identified dataset may be made available by the corresponding author on reasonable request, subject to applicable ethical and data-protection requirements. Individual time-to-positivity records and verified time-point counts were not available for the present analysis.

Conflicts of Interest: A.S. is the inventor and rights holder named in Polish Patent Pat.249585 concerning the current-generation urease test and was also the inventor and rights holder of the earlier Pat.215889. These intellectual-property interests represent a potential conflict of interest and are disclosed transparently. No patent-related interest of D.B. is identified.

References

1. Malfertheiner, P.; Megraud, F.; Rokkas, T.; Gisbert, J.P.; Liou, J.-M.; Schulz, C.; Gasbarrini, A.; Hunt, R.H.; Leja, M.; O'Morain, C.; et al. Management of *Helicobacter pylori* infection: The Maastricht VI/Florence consensus report. *Gut* **2022**, *71*, 1724–1762. [CrossRef] [PubMed]
2. Chey, W.D.; Howden, C.W.; Moss, S.F.; Morgan, D.R.; Greer, K.B.; Grover, S.; Shah, S.C. ACG Clinical Guideline: Treatment of *Helicobacter pylori* Infection. *Am. J. Gastroenterol.* **2024**, *119*, 1730–1753. [CrossRef] [PubMed]
3. Li, Y.; Choi, H.; Leung, K.; Jiang, F.; Graham, D.Y.; Leung, W.K. Global prevalence of *Helicobacter pylori* infection between 1980 and 2022: A systematic review and meta-analysis. *Lancet Gastroenterol. Hepatol.* **2023**, *8*, 553–564. [CrossRef] [PubMed]
4. Costa, L.C.M.C.; Carvalho, M.G.; Pereira, A.C.L.C.; Pereira, A.C.L.G.C.; Neto, R.G.T.; Figueiredo, L.C.A.; Barros-Pinheiro, M. Diagnostic Methods for *Helicobacter pylori*. *Med. Princ. Pract.* **2024**, *33*, 173–184. [CrossRef] [PubMed]
5. Uotani, T.; Graham, D.Y. Diagnosis of *Helicobacter pylori* using the rapid urease test. *Ann. Transl. Med.* **2015**, *3*, 9. [CrossRef] [PubMed]
6. Ricci, C.; Holton, J.; Vaira, D. Diagnosis of *Helicobacter pylori*: Invasive and non-invasive tests. *Best Pract. Res. Clin. Gastroenterol.* **2007**, *21*, 299–313. [CrossRef] [PubMed]
7. Gisbert, J.P.; Abraira, V. Accuracy of *Helicobacter pylori* diagnostic tests in patients with bleeding peptic ulcer: A systematic review and meta-analysis. *Am. J. Gastroenterol.* **2006**, *101*, 848–863. [CrossRef] [PubMed]
8. Dechant, F.X.; Dechant, R.; Kandulski, A.; Selgrad, M.; Weber, F.; Reischl, U.; Wilczek, W.; Mueller, M.; Weigand, K. Accuracy of different rapid urease tests in comparison with histopathology in patients with endoscopic signs of gastritis. *Digestion* **2020**, *101*, 184–190. [CrossRef] [PubMed]
9. Laine, L.; Lewin, D.; Naritoku, W.; Estrada, R.; Cohen, H. Prospective comparison of commercially available rapid urease tests for the diagnosis of *Helicobacter pylori*. *Gastrointest. Endosc.* **1996**, *44*, 523–526. [CrossRef] [PubMed]
10. Yousfi, M.M.; El-Zimaity, H.M.T.; Cole, R.A.; Genta, R.M.; Graham, D.Y. Detection of *Helicobacter pylori* by rapid urease tests: Is biopsy size a critical variable? *Gastrointest. Endosc.* **1996**, *43*, 222–224. [CrossRef] [PubMed]
11. Ansari, S.; Yamaoka, Y. *Helicobacter pylori* infection, its laboratory diagnosis, and antimicrobial resistance: A perspective of clinical relevance. *Clin. Microbiol. Rev.* **2022**, *35*, e00258-21. [CrossRef] [PubMed]

12. Jaroń, K.; Pietrzak, A.; Daniluk, J.; Adrych, K.; Gąsiorowska, A.; Skrzydło-Radomańska, B.; Małecka-Wojciesko, E.; Zwolińska-Wcisło, M.; Waluga, M.; Reguła, J.; et al. Diagnostic and therapeutic recommendations on *Helicobacter pylori* infection. Recommendations of the Working Group of the Polish Society of Gastroenterology. *Prz. Gastroenterol.* **2023**, *18*, 225–248. [CrossRef] [PubMed]
13. Siwiec, A. Non-invasive urease test for rapid diagnosis of *Helicobacter pylori*. Doctoral Dissertation, Medical University of Warsaw, Warsaw, Poland, 2013. (In Polish)
14. Siwiec, A. Urease Test for Detection of *Helicobacter pylori*. Polish Patent PL215889B1, Published 28 February 2014; Application P.392212, Granted on 23 April 2013. Available online: https://patents.google.com/patent/PL215889B1/pl (accessed on 20 September 2026).
15. Siwiec, A. Urease Test for Detection of *Helicobacter pylori*. Polish Patent Application P.445161, Filed 6 June 2023; Publication PL445161A1, Granted on 6 June 2023; Associated Patent Pat.249585. *Biuletyn Urzędu Patentowego* **2024**, *50*, 7. Available online: https://uprp.gov.pl/sites/default/files/bup/2024/Wynalazki%20i%20wzory%20u%C5%BCytkowe/12_49-53/50/bup50_2024.pdf (accessed on 9 December 2024).
16. Dixon, M.F.; Genta, R.M.; Yardley, J.H.; Correa, P. Classification and grading of gastritis. The updated Sydney System. International Workshop on the Histopathology of Gastritis, Houston 1994. *Am. J. Surg. Pathol.* **1996**, *20*, 1161–1181. [CrossRef] [PubMed]
17. Veitch, A.M.; Radaelli, F.; Alikhan, R.; Dumonceau, J.M.; Eaton, D.; Jerrome, J.; Lester, W.; Nylander, D.; Thoufeeq, M.; Vanbiervliet, G.; et al. Endoscopy in patients on antiplatelet or anticoagulant therapy: British Society of Gastroenterology (BSG) and European Society of Gastrointestinal Endoscopy (ESGE) guideline update. *Gut* **2021**, *70*, 1611–1628. [CrossRef] [PubMed]
18. Yuki, T.; Ishihara, S.; Yashima, K.; Kawaguchi, K.; Fujishiro, H.; Miyaoka, Y.; Yuki, M.; Kushiyama, Y.; Yasugi, A.; Shabana, M.; et al. Bleeding Risk Related to Upper Gastrointestinal Endoscopic Biopsy in Patients Receiving Antithrombotic Therapy: A Multicenter Prospective Observational Study. *Curr. Ther. Res. Clin. Exp.* **2017**, *84*, 32–36. [CrossRef] [PubMed]
19. Soh, H.; Kim, N.; Moon, J.; Lee, H.M.; Kim, B.; Cho, S.-J.; Kim, S.G.; Chung, H. Gastric mucosal swab as an alternative to biopsy for *Helicobacter pylori* detection. *Helicobacter* **2023**, *28*, e12980. [CrossRef] [PubMed]
20. Yoshikawa, T.; Yamauchi, A.; Kou, T.; Murao, T.; Kamada, T.; Suehiro, M.; Kawano, K.; Haruma, K.; Yazumi, S. Validity of rapid urease test using swab of gastric mucus to mucosal forceps and 13 C-urease breath test: A multicenter prospective observational study. *BMC Gastroenterol.* **2024**, *24*, 258. [CrossRef] [PubMed]
21. Noh, C.-K.; Lee, G.H.; Lee, E.; Park, B.; Lim, S.G.; Shin, S.J.; Lee, K.M. Comparative diagnostic performance of rapid urease test with the sweeping method versus tissue sampling method after *Helicobacter pylori* eradication (with video). *Gastrointest. Endosc.* **2024**, *100*, 660–669.e3. [CrossRef] [PubMed]
22. Bang, C.S.; Choi, A.I.; Choi, J.-G.; Lee, J.J.; Gong, E.J. Mucosal Brush Sampling Increases the *Helicobacter pylori* Detection Rate of the Rapid Urease Test: A Pilot Study. *Gut Liver* **2026**, *20*, 77–85. [CrossRef] [PubMed]
23. Siwiec, A.; Siwiec, A. Analytical performance evaluation of the Helibac rapid urease test for detection of *Helicobacter pylori* urease activity under controlled laboratory conditions. *Front. Microbiol.* **2026**, *17*, 1889913. [CrossRef] [PubMed]
24. Bossuyt, P.M.; Reitsma, J.B.; Bruns, D.E.; Gatsonis, C.A.; Glasziou, P.P.; Irwig, L.; Lijmer, J.G.; Moher, D.; Rennie, D.; de Vet, H.C.W.; et al. STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *BMJ* **2015**, *351*, h5527. [CrossRef] [PubMed]
25. Cohen, J.F.; Korevaar, D.A.; Altman, D.G.; Bruns, D.E.; Gatsonis, C.A.; Hooft, L.; Irwig, L.; Levine, D.; Reitsma, J.B.; de Vet, H.C.W.; et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: Explanation and elaboration. *BMJ Open* **2016**, *6*, e012799. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.