Validation of the Lyon classification for GORD diagnosis: acid exposure time assessed by prolonged wireless pH monitoring in healthy controls and patients with erosive oesophagitis

Radu-Ionut Rusu \( ^{1} \), Mark R Fox \( ^{2,3} \), Emily Tucker \( ^{3,4} \), Sebastian Zeki \( ^{1} \), Jason M Dunn \( ^{1} \), Jafar Jafari \( ^{1} \), Fiona Warburton \( ^{1,5} \), Terry Wong \( ^{1} \)

ABSTRACT

Objective Acid exposure time (AET) from ambulatory pH studies and reflux oesophagitis are independent measurements used by the Lyon classification to diagnose GORD. This study aimed to validate AET reference ranges and diagnostic thresholds by analysis of 96-hour wireless pH studies from healthy, asymptomatic controls (HCs) and patients with and without oesophagitis.

Design HC and consecutive patients referred for wireless pH studies (off acid suppressants for \( >7 \) days) underwent 96-hour pH studies at two tertiary referral centres. Erosive oesophagitis was categorised by the Los Angeles (LA) classification. Linear regression and receiver operating curve (ROC) analysis were performed to define optimal diagnostic cut-offs.

Results Prolonged, 96-hour pH studies were completed in 39 HCs (age 28 (18–53) years, 72% female) and 944 patients (age 46 (16–85) years, 65% female), of whom 136 (14.5%) had reflux oesophagitis. Median AET in HC was 1.3% (upper 95th percentile 4.6%) for any study day and 2.6% (upper 95th percentile 6.9%) for the worst day (24-hour period) during the study. ROC analysis for average AET differentiated HC from patients with moderate-to-severe oesophagitis (LA BCD; sensitivity 87%, specificity 95%, positive predictive value (PPV) 59%, negative predictive value 99% for a cut-off AET of 4.3%; area under the receiver operating curve 0.95). Specificity was higher, but PPV was substantially lower for severe oesophagitis (LA CD). ‘Worst-day’ analysis provided similar results; however, day-to-day variability was high.

Conclusion Diagnostic thresholds for average AET were identified that accurately discriminate between HCs and patients with erosive oesophagitis. The findings provide conditional support for diagnostic criteria proposed by the Lyon Consensus.

INTRODUCTION

GORD is present when the reflux of gastric content into the oesophagus causes troublesome symptoms and/or complications. Typical symptoms of GORD including heartburn and regurgitation are common, with an estimated worldwide prevalence between 4% and 14% with important variation in different regions, and up to 30% of Americans complaining of at least occasional reflux symptoms in the previous week. Empirical treatment with proton pump inhibitors (PPIs) is appropriate in patients with typical symptoms; however, the response to PPIs is neither sensitive or specific for GORD diagnosis. Endoscopic investigation is performed in patients with persistent reflux symptoms; however, less than half of patients have conclusive evidence of GORD based on the presence of severe erosive oesophagitis.

What is already known about this subject?

- Prolonged, ambulatory wireless pH studies improve diagnostic yield and reliability of reflux studies in clinical practice; however, reference ranges and diagnostic thresholds for this methodology have not been established.
- The Lyon classification defines acid exposure time (AET) \( >4\% \) as conditional and \( >6\% \) as conclusive evidence for the presence of GORD, but these thresholds have not been prospectively validated against independent markers of disease.

What are the new findings?

- Diagnostic thresholds for AET were identified that accurately discriminate between healthy controls and patients with moderate to severe erosive oesophagitis. The results provide conditional support for diagnostic criteria proposed by the Lyon Consensus.

How might it impact on clinical practice in the foreseeable future?

- The presence of moderate (Los Angeles (LA) grade B) as well as severe oesophagitis (LA grade CD) supports the diagnosis of GORD in patients referred for investigation of reflux symptoms.
- The results confirm that average AET \( <4\% \) represents normal (physiological) acid exposure, whereas AET \( >7\% \), 1% higher than the current threshold, provides conclusive evidence of GORD.

Significance of this study
Table 1 Oesophageal acid exposure time (AET) in healthy controls during each 24-hour period

<table>
<thead>
<tr>
<th></th>
<th>Day 1 (n=61)</th>
<th>Day 2 (n=60)</th>
<th>Day 3 (n=53)</th>
<th>Day 4 (n=39)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median AET % total time pH&lt;4</td>
<td>1.20 (0.01–6.15)</td>
<td>1.35 (0.10–5.59)</td>
<td>0.90 (0.00–4.85)</td>
<td>1.20 (0.00–4.50)</td>
</tr>
</tbody>
</table>

Median (5th–95th percentiles).

Table 2 Oesophageal acid exposure time assessed during 24, 48, 72 and 96-hour study period

<table>
<thead>
<tr>
<th></th>
<th>24 hours (n=61)</th>
<th>48 hours (n=60)</th>
<th>72 hours (n=53)</th>
<th>96 hours (n=39)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average day</td>
<td>1.20 (0.01–6.15)</td>
<td>1.27 (0.10–5.04)</td>
<td>1.26 (0.07–4.94)</td>
<td>1.32 (0.10–4.65)</td>
</tr>
<tr>
<td>Worst day</td>
<td>1.20 (0.01–6.15)</td>
<td>2.20 (0.10–6.59)</td>
<td>2.50 (0.10–6.69)</td>
<td>2.60 (0.40–6.90)</td>
</tr>
</tbody>
</table>

Median (5th–95th percentiles).
removed to improve the separation of normal and abnormal oesophageal acid exposure. Reflux was defined by pH<4 and AET was documented. Data were anonymised and transferred via file encryption service to the statistician.

Statistical analysis
The upper limit of the reference interval ('normal range') in healthy controls was defined by the 95th percentile. Power calculations indicated that if 40 HCs are used to define reference intervals for 96-hour pH studies, then the CI for a diagnostic threshold for GORD of 4.4% is 3.2%–6.1% CI (2.9%). This calculation assumed a log normal distribution fitted to published intervals for 96-hour calculations indicated that if 40 HCs are used to define reference groups and incomplete data.

Median and 95th percentile values were calculated for non-normally distributed AET over 24, 48, 72 and 96 hours and compared against the single ‘worst-day’ (ie, most pathological 24-hour period). Retrospectively, the same analysis was applied to results from consecutive patients with normal endoscopy and erosive oesophagitis LA A, B, C, D that completed 96-hour wireless pH studies off acid suppressants. Patients with high-grade erosive oesophagitis (LA grades C and D) were considered a single group as small numbers in each group precluded meaningful calculations.

Linear regression analysis was performed to assess the relationship between endoscopic findings with average AET, adjusted for gender, age and duration (24, 48, 72 and 96 hours). For the worst-day AET, the analysis was adjusted for these factors and the interaction between worst day and duration. Receiver operator curve (ROC) analysis was performed on the 96-hour data to determine thresholds for discriminating between patients with definite endoscopic evidence of GORD and HCs for both average AET and worst-day AET. Youden’s index was used to determine the optimal cut-offs. Upper 95th percentile values from normal subjects and lower 5th percentile values from individuals with proven GORD defined by oesophagitis were used to assess the overlap between health and disease. Statistical analyses were performed using Stata V.15 (StataCorp 2018).

RESULTS
Healthy controls
Eighty HCs (mean age 27, 18–53 years, 65% female) were recruited: 30 in Nottingham (mean age 27, 21–50 years, 60% female) and 50 in London (mean age 27, 18–53 years, 68% female). Participants were symptom-free at screening. Per protocol, 12 HCs were excluded due to oesophagitis on endoscopy (n=10 LA grade A, 2 grade B). Early detachment or incorrect placement of the pH sensor. Thus, 39 out of 80 HCs (49%), mean age 28 (18–53) years, 72% female, mean BMI 24 kg/m² (range 19–33.9 kg/m²) had full data for inclusion in the final analysis. No demographic or physiological differences were present between HCs with prolonged pH data in the analysis and those excluded.

No complications occurred with the Bravo system (capsule retention, failure to detach or aspiration, tears in the mucosa, bleeding, perforation). The catheter-free procedures were well tolerated (satisfaction score 4.1/5) by HCs, and 84% (36/43) in education or employment attended courses or went to work during the test.

<table>
<thead>
<tr>
<th>Study duration</th>
<th>Total N</th>
<th>Normal n (%)</th>
<th>LA A n (%)</th>
<th>LA B n (%)</th>
<th>LA C, D n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>48 hours</td>
<td>1789</td>
<td>1496 (83.6)</td>
<td>118 (6.6)</td>
<td>127 (7.1)</td>
<td>48 (2.7)</td>
</tr>
<tr>
<td>72 hours</td>
<td>1191</td>
<td>1030 (86.5)</td>
<td>72 (6.0)</td>
<td>71 (5.9)</td>
<td>18 (1.6)</td>
</tr>
<tr>
<td>96 hours</td>
<td>944</td>
<td>808 (85.6)</td>
<td>61 (6.5)</td>
<td>60 (6.3)</td>
<td>15 (1.6)</td>
</tr>
</tbody>
</table>

The presence and severity of reflux oesophagitis did not impact on the likelihood of early detachment. LA, Los Angeles; n, number of patients.

Reference intervals
Median and 5th and 95th percentiles in HCs for AET on each day of the study and over 24, 48, 72 and 96 hours for both the average of all days recorded ('average-day analysis') and the single most pathological day ('worst-day analysis') are illustrated in tables 1 and 2. The upper limit of reference for AET for the 96-hour studies was 4.6% for any study day and 6.9% for the worst day. No difference between the study days was found (p=0.52 Kruskal-Wallis test).

Clinical study
Between March 2008 and November 2019, 2095 consecutive patients with symptoms of GORD underwent catheter-free pH monitoring. Of these, 306 patients were excluded due to studies lasting <48 hours, invalid or missing data, and non-compliance with the protocol. Endoscopy results were recorded in 1789 patients fulfilling the study criteria (tables 3 and 4). Of these, 944 patients (53%, mean age 46, 16–85 years, 65% female) completed >92-hour study. Eight hundred and eight (85%) had normal endoscopy, and 136 (15%) had erosive oesophagitis categorised by the LA classification having stopped PPI medication at least 1 week prior to investigation.

HGs had lower average and worst-day AET than patients with normal endoscopy and those with erosive oesophagitis LA ABCD (p<0.0001 Kruskal-Wallis test). Patients with normal endoscopy had lower AET than erosive oesophagitis LA ABCD (p<0.0001). Further, patients with LA A had lower AET than those with LA B or LA CD (p<0.0001). An association between median and 95th percentile values for average and worst-day AET over 96 hours and endoscopic

<table>
<thead>
<tr>
<th>Patients (n=944)</th>
<th>HC studied (n=39)</th>
<th>HC enrolled (n=80)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>27 (18–53)</td>
<td>28 (18–53)</td>
</tr>
<tr>
<td>Female</td>
<td>52 (65%)</td>
<td>28 (72%)</td>
</tr>
<tr>
<td>BMI</td>
<td>24 kg/m² (19–33.9)</td>
<td>24 kg/m² (19–33.9)</td>
</tr>
<tr>
<td>RDQ=0</td>
<td>80 (100%)</td>
<td>39 (100%)</td>
</tr>
<tr>
<td>HODQ=0</td>
<td>80 (100%)</td>
<td>39 (100%)</td>
</tr>
<tr>
<td>Normal endoscopy</td>
<td>68 (85%)</td>
<td>39 (100%)</td>
</tr>
<tr>
<td>Erosive oesophagitis</td>
<td>12 (15%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Capsule attachment &gt;92 hours</td>
<td>54 (68%)</td>
<td>39 (100%)</td>
</tr>
</tbody>
</table>

Values are mean (range) or n (%).
*Weight, height and Hospital Odynophagia Dysphagia Questionnaire (HODQ) were not routinely recorded in patients. RDQ, Reflux Disease Questionnaire.
findings was present, such that acid exposure increases with oesophagitis grade (table 5 and online supplemental table S1).

The ‘dose–response relationship’ between AET and with the presence and the severity of mucosal disease is evident; however, the box and whiskers plot highlights also the large variability in AET for each patient group (figure 1 (average AET), online supplemental figure S1 (worst-day AET)).

### Linear regression analysis

Linear regression analysis for the average-day analysis with AET as the dependent variable (table 6) shows a significant difference between HCs, symptomatic patients with normal endoscopy, patients with any grade (LA ABCD), moderate to severe (LA BCD) and severe oesophagitis (LA CD). The AET for severe oesophagitis (LA CD) is different from symptomatic patients with normal endoscopy and LA A oesophagitis (95% CIs of these groups do not overlap), but not LA B oesophagitis (95% CIs of these groups overlap).

Linear regression analysis of the worst-day data with total AET as the dependent variable reveals similar findings (online supplemental table S2).

### ROC analysis

Receiver operating curve (ROC) analysis estimating the diagnostic accuracy of endoscopic findings for GORD diagnosis with reference to average AET and worst-day AET from 96 hours studies is presented in table 7 and online supplemental table S3 respectively. The optimal thresholds for discriminating between HCs and patients with any grade of oesophagitis (LA ABCD), between HCs and patients with moderate to severe endoscopic evidence of GORD (LA BCD) and between HCs and patients with severe erosive oesophagitis (LA CD) for average AET and worst-day AET are shown in table 7, highlighted in bold text on the ROC curves in figure 2 and online supplemental figure S2, respectively. This analysis shows that 87% of patients with moderate to severe reflux oesophagitis (LA BCD) had average AET>4.3%, and 85% had worst-day AET>6.9%. By comparison, 93% of patients with severe oesophagitis (LA CD) had average AET>4.7%, and 100% had worst-day AET>7.1%.

### Use of AET to discriminate between health and disease

The 5th and 95th percentiles for average and worst-day AET in controls and patients with erosive oesophagitis are shown in table 8 and online supplemental table S4. Histograms of the data, for controls and patients with the 95th centile for controls and the 5th centile for cases marked, reveal a clear division (ie, essentially no overlap) between the ‘upper limit of normal’ AET in HCs and results from individuals with LA CD reflux

---

**Table 5** Comparison of average AET in healthy controls and patients with reflux symptoms referred for ambulatory pH studies

<table>
<thead>
<tr>
<th>Group</th>
<th>96 hours average AET</th>
<th>IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy controls (n=39)</td>
<td>1.32 (0.10–4.65)</td>
<td>0.60, 2.57</td>
</tr>
<tr>
<td>Patients normal endoscopy (n=808)</td>
<td>2.95 (0.05–9.94)</td>
<td>1.05, 5.71</td>
</tr>
<tr>
<td>LA A erosive oesophagitis (n=61)</td>
<td>6.10 (0.80–15.50)</td>
<td>3.5, 7.73</td>
</tr>
<tr>
<td>LA B erosive oesophagitis (n=60)</td>
<td>8.23 (1.61–18.41)</td>
<td>5.12, 12.47</td>
</tr>
<tr>
<td>LA C, D erosive oesophagitis (n=15)</td>
<td>9.95 (4.07–24.90)</td>
<td>9.05, 15.58</td>
</tr>
</tbody>
</table>

Median (5th–95th percentiles). IQR, inter-quartile range.

---

**Table 6** Results of linear regression analysis of the 96-hour average AET and endoscopic findings

<table>
<thead>
<tr>
<th>Group</th>
<th>Coefficient</th>
<th>95% CI</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Reference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symptoms normal endoscopy</td>
<td>1.92</td>
<td>0.70 to 3.13</td>
<td>0.002</td>
</tr>
<tr>
<td>B</td>
<td>5.09</td>
<td>3.64 to 6.55</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>C</td>
<td>7.82</td>
<td>6.37 to 9.27</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>D</td>
<td>11.02</td>
<td>9.23 to 12.82</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>A+B+C, D</td>
<td>8.6</td>
<td>6.30 to 10.91</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>A+B+C</td>
<td>7.14</td>
<td>5.81 to 8.48</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>All patients</td>
<td>2.54</td>
<td>1.21 to 3.87</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Healthy controls provide reference values.

R² represents the amount of variation in AET (dependent variable) that is explained by the grade of oesophagitis.

The total % time for patients with symptoms and normal endoscopy, LA A, LA B, LA C, D, LA B+C, D and LA A-D is on average 1.92, 5.09, 7.82, 11.02, 8.60 and 7.14 higher than controls, respectively. The total % time for all patients is 2.54 higher than controls.

AET, acid exposure time; CI, confidence interval.

---

**Table 7** Summary of results from receiver operating curve (ROC) analysis for the diagnostic accuracy of endoscopic findings for GORD diagnosis based on 96-hour average AET

<table>
<thead>
<tr>
<th>Group</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
<th>Cut-off AET %</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control vs LA A, B, C, D</td>
<td>80</td>
<td>95</td>
<td>72</td>
<td>97</td>
<td>4.3</td>
<td>0.92</td>
</tr>
<tr>
<td>Control vs LA B, C, D</td>
<td>87</td>
<td>95</td>
<td>59</td>
<td>99</td>
<td>4.3</td>
<td>0.95</td>
</tr>
<tr>
<td>Control vs LA C, D</td>
<td>94</td>
<td>98</td>
<td>42</td>
<td>100</td>
<td>4.7</td>
<td>0.99</td>
</tr>
</tbody>
</table>

AET, acid exposure time; AUC, area under the ROC curve; PPV, positive predictive value; NPV, negative predictive value.
Oesophagus

Reflux oesophagitis (figure 3 and online supplemental figure S3). There is limited overlap between HCs and individuals with LA BCD but considerable overlap if any grade of oesophagitis is considered. Overall, the average day data display less overlap than worst-day data (table 8 and figure 3, online supplemental table S4 and figure S3).

Validation of Lyon classification
Reflux oesophagitis

Linear regression analysis shows a significant difference between patients with severe oesophagitis (LA CD), moderate to severe erosive oesophagitis (LA BCD) and any grade of oesophagitis (LA ABCD), respectively, and the HC group for average AET and worst-day AET. There is also a highly significant difference between LA grade A, LA grade B and the control group for average AET and worst-day AET. Severe oesophagitis (LA CD) is significantly different from LA A for average AET and worst-day AET.

Table 8  Summary of results showing 5th and 95th percentiles in healthy controls and patients with erosive oesophagitis for average day AET from the 96-hour studies

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Control LA C+D</th>
<th>LA B+CD</th>
<th>LA A+B+CD+D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average day</td>
<td>5.08</td>
<td>4.08</td>
<td>1.80</td>
</tr>
<tr>
<td></td>
<td>4.65</td>
<td>24.9</td>
<td>22.3</td>
</tr>
<tr>
<td></td>
<td>17.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AET, acid exposure time; LA, Los Angeles.

Figure 2  ROC analysis with tables presenting results for the full range of potential diagnostic thresholds for discriminating between healthy controls and patients with erosive oesophagitis based on average AET from the 96-hour studies. Upper panel: any oesophagitis grade. Middle panel: moderate to severe reflux oesophagitis (LA grades B, C, D). Lower panel: severe oesophagitis (LA grades C, D). In each case, the optimal diagnostic threshold is highlighted in bold text.

Figure 3  Histogram showing the distribution of 96-hour average AET in healthy controls and patients with erosive oesophagitis. Upper panel: any oesophagitis grade. Middle panel: moderate to severe reflux oesophagitis (LA grade B, C, D). Lower panel: severe oesophagitis (LA grade C, D). The overlap between health and disease is highlighted (upper 95th percentile limit for healthy controls, 5th percentiles for cases).
Oesophagus

AET; however, this did not reach statistical significance for LA B (online supplemental table S2).

Ambulatory pH studies
The upper limit of reference (‘normal’) of AET for the 96-hour studies calculated using the median and 95th percentile was 4.6% for the average day and 6.9% for the worst-day analysis.

In HCs and each patient group, there is much less variability of the average-day AET data compared with worst-day AET data, whether this is reported as median (5th–95th percentiles), IQR or displayed in box and whiskers plots. Further, histograms show less overlap between normal subjects and individuals with LA BCD and LA ABCD oesophagitis for average-day data when compared with the worst-day data (figure S3).

ROC analysis shows that the optimal diagnostic thresholds that discriminate between healthy controls and patients with severe oesophagitis (LA CD), moderate to severe oesophagitis (LA BCD) and any oesophagitis grade (LA ABCD) was 4.7%, 4.3% and 4.3%, respectively, based on average AET, and 7.1%, 6.9% and 6.9%, respectively, based on worst-day AET.

DISCUSSION
Ambulatory measurement of oesophageal acid exposure provides an objective basis for the diagnosis of GORD. This analysis of endoscopic and physiological data from a large group of HCs and patients referred for investigation of reflux symptoms provides an important validation of diagnostic thresholds for GORD from the Lyon classification. This research also provides insight into the optimal analysis of prolonged pH studies.

The average AET is almost identical on all four test days; however, there is important day-to-day variation in AET (table 1). As a result, the reproducibility of GORD diagnosis based on any 24-hour period is suboptimal. Increasing the duration of pH measurement from 24 to 96 hours progressively improves the reproducibility of results. What remains uncertain is the most appropriate diagnostic threshold and the best method for analysis of pH data for GORD diagnosis. The Lyon Consensus considers AET<4% to be physiological and >6% to be pathological; however, these limits that are based on expert opinion have not been tested prospectively. To validate these thresholds, it is necessary to compare AET results with an independent assessment of GORD severity. The presence of reflux oesophagitis has been used for this purpose since the introduction of ambulatory pH measurement by Johnson and DeMeester.

This study analysed data from HCs and patients that successfully completed prolonged, 96-hour wireless pH studies with only minimal loss of data. About half the study participants met these stringent criteria. Those included had similar characteristics as those that did not complete 96-hour studies and are likely representative of the full cohort. Patients were instructed to stop PPI therapy at least 7 days before the study. The effects of these medications on gastric acid secretion are no longer present after 1 week; however, recurrence of oesophagitis requires more time with the extension of a PPI-free period from 1 week to indefinitely result in a deterioration of erosive oesophagitis by one grade.24 25 This is one reason why only a minority of patients had mucosal disease on endoscopy. Nevertheless, the large size of the cohort ensured that sufficient patients with reflux oesophagitis (n=136) contributed 96-hour pH data to the analysis.

There is a significant difference between AET measured in HCs and patients without mucosal disease compared with patients with reflux oesophagitis (table 6, online supplemental table S2). A positive, ‘dose-dependent’ relationship between AET with the severity of oesophagitis is present (figure 1, online supplemental figure S1). This confirms that such data can be used for purposes of validation. A key issue concerns the oesophagitis grade that should be considered conclusive for GORD diagnosis. In a large, population-based study, 16% of healthy, asymptomatic subjects had reflux oesophagitis (10% grade A, 4% grade B, 2% severe oesophagitis or Barrett’s oesophagitis). Similarly, in the current series, 12 (15%) HCs were excluded because oesophagitis was present on endoscopy (n=10 LA grade A, 2 grade B). These findings support the view that grade A oesophagitis is not pathological; however, the average AET in patients with LA grade B is in the conclusively pathological range (>8% AET) and higher than that measured in those with mild, grade A oesophagitis or non-erosive disease (table 5). Only 9 out of 60 (15%) patients with LA B had average day AET<4%. Additionally, the ‘overlap’ of AET results from HCs and from patients with moderate to severe reflux oesophagitis (LA BCD) is small (7.8%), just above the traditional 5% threshold used to define the pathological range in clinical diagnostics (figure 3). By contrast, discrimination between health and disease is almost complete for severe, LA grade CD oesophagitis (table 8, figure 3, online supplemental table S4 and figure S3). Taken together, these results highlight that LA grade B oesophagitis provides at least conditional evidence of objective, pathological GORD. This assertion is supported by results from the longitudinal ProGERD study that observed frequent, spontaneous resolution of mild oesophagitis on repeated endoscopy but a significant risk of disease progression in patients with moderate to severe oesophagitis. Similarly, clinical studies report good outcomes for medical and surgical antireflux treatment in patients with this form of erosive GORD.

The upper limit of oesophageal AET in HCs as determined by the 95th percentile of the normal range is 4.6% for the average day and 6.9% for the worst-day AET analysis (table 2). Thus, although most HCs have a low AET on most days (96 hours average 1.32% AET), it is ‘normal’ for some healthy individuals to have occasional days with 6%–7% AET. These findings are in remarkably close agreement with the results of an ROC analysis that indicates the optimal diagnostic threshold that discriminates between HCs and patients with moderate to severe oesophagitis (LA grade BCD) is 4.3% based on average AET and 6.9% based on the worst-day AET (table 7, online supplemental table S3). The optimal threshold to discriminate HCs and patients with severe oesophagitis (LA grade CD) was somewhat higher, 4.7% based on average AET and 7.1% based on the worst-day AET (table 7, online supplemental table S3). These results provide conditional support for the Lyon classification which defines AET<4% as normal (physiological). By contrast, based on analysis of worst-day AET data, the AET >7% threshold provides more conclusive evidence for GORD diagnosis than the current AET >6% threshold. These results suggest that intermediate values between 4% and 7% AET, a slightly wider range than in version 1.0 of the classification, should be considered inconclusive and require additional supportive evidence to establish GORD diagnosis. Thresholds derived from physiological measurement require support from clinical studies. A recent double-blinded clinical trial in patients who underwent 96-hour pH studies demonstrated that the optimal predictor of successful withdrawal of PPIs was an average day AET <4.4% and a worst-day AET <7.4% on reflux monitoring. The close agreement with this study confirms that the normal and pathological ranges defined by the present analysis are clinically relevant.
Independent of diagnostic thresholds, it is important to identify the most appropriate analysis of prolonged pH measurements in clinical practice. Worst-day analysis maximises the sensitivity and yield for GORD diagnosis in patients with reflux symptoms (ie, low risk of false-negative result). However, this method increases the risk of a false-positive diagnosis due to high day-to-day variation in AET. Graphical illustration of results demonstrate a greater overlap between results from HCs and patients with oesophagitis for the worst day than the average AET analysis. Average AET analysis has a high specificity for GORD diagnosis because inconclusive (borderline) or discordant data from any one 24-hour measurement period will increasingly become ‘more representative’ of the true result as more information is collected. (Note that diagnosis based on average day AET is closely correlated with that based on ‘predominant AET pattern’ analysis proposed by Hasak et al. The latter method tallies the number of days with negative, inconclusive and positive AET diagnoses with the final result determined by the most frequent (predominant) pattern. This simple approach may increase the proportion of patients with a conclusive diagnosis; however, it cannot be used for measurements <3 days, discards incomplete data (<24 hours) and is inconclusive if the number of positive and negative days are identical). Thus, consistent with recent guidelines, current findings indicate that the result of prolonged pH studies should be based on all the data available and not focused on the ‘worst day’ because this is a more stable and statistically robust approach when applied to individual cases.

Strengths and limitations of this study
A group of 39 well-characterised HCs with no evidence of GORD and complete 96-hour pH studies provided 156 days reference data. The numbers of patients with reflux oesophagitis in the clinical cohort (136/944) allowed for a statistically well-powered validation study to be performed. This addresses limitations of previous studies that analysed more heterogeneous groups and/or did not have sufficient numbers to provide conclusive results. The mean age of HCs was lower than that of patients (table 4); however, the increase in AET that can be attributed to the 15–20 year age gap is estimated to be <2%, based on results from the same population. The study was performed in a tertiary referral unit; however, the Oesophageal Physiology Laboratory in Guy’s Hospital (GSTT) is the main provider for this service in the South-East Thames region. The relatively short, 7-day time between stopping PPI medication and endoscopy can be criticised. Stopping acid suppressant medication for prolonged periods of time in highly symptomatic patients is difficult and not acceptable to the ethics committee. In patients with conclusive GORD diagnosis, based on the presence of moderate to severe erosive oesophagitis. The results support the use of prolonged, ideally 4-day, ambulatory pH studies to counter the large day-to-day variation in AET that is seen in patients with reflux symptoms. The study will also inform the decisions of the committee when the next version of the Lyon Consensus is drafted. First, it adds to the evidence that moderate, LA grade B oesophagitis provides at least conditional evidence of GORD. Second, it indicates that the threshold for a conclusive diagnosis of GORD by ambulatory pH measurement should be refined upwards from 6% to 7% AET, with average day AET analysis preferred to worst-day AET analysis when prolonged pH measurements are available. These findings will improve the diagnosis of GORD and facilitate the selection of patients for medical and surgical treatment.

Conclusions and implications for clinical practice
This study defines the normal reference range for 96-hour ambulatory wireless pH monitoring and identifies pathological thresholds that discriminate between healthy controls and patients with conclusive GORD diagnosis, based on the presence of moderate to severe erosive oesophagitis. The results support the use of prolonged, ideally 4-day, ambulatory pH studies to counter the large day-to-day variation in AET that is seen in patients with reflux symptoms. The study will also inform the decisions of the committee when the next version of the Lyon Consensus is drafted. First, it adds to the evidence that moderate, LA grade B oesophagitis provides at least conditional evidence of GORD. Second, it indicates that the threshold for a conclusive diagnosis of GORD by ambulatory pH measurement should be refined upwards from 6% to 7% AET, with average day AET analysis preferred to worst-day AET analysis when prolonged pH measurements are available. These findings will improve the diagnosis of GORD and facilitate the selection of patients for medical and surgical treatment.
Oesophagus

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information. Available on reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iDs
Radu-Ionut Rusu http://orcid.org/0000-0002-1355-9300
Mark R Fox http://orcid.org/0000-0003-4294-5584
Emily Tucker http://orcid.org/0000-0002-7518-1147
Sebastian Zeki http://orcid.org/0000-0003-1673-2663
Jason M Dunn http://orcid.org/0000-0002-1973-6072
Jafar Jafari http://orcid.org/0000-0003-3453-6324
Fiona Warburton http://orcid.org/0000-0001-3954-7758
Terry Wong http://orcid.org/0000-0002-5236-9139

REFERENCES