Retrospective analysis of a lactose breath test in a gastrointestinal symptomatic population of Northeast Italy: use of (H₂ + 2CH₄) versus H₂ threshold

Gregorio Peron¹
Stefano Dall’Acqua¹,²
Vincenzo Sorrenti¹,³
Maria Carrara¹
Stefano Fortinguerria³
Giulia Zorzi³
Alessandro Buriani³
¹Department of Pharmaceutical and Pharmacological Sciences, University of Padova, Padova, Italy; ²Unir&d, Nutraceutical Laboratory, Padova, Italy; ³Maria Paola Belloni Center for Personalized Medicine, Data Medica Group, Synlab Limited, Padova, Italy

Background: Lactose malabsorption is normally evaluated by measuring exhaled H₂ produced by intestinal flora, from unabsorbed lactose. However, differing microbiome composition can lead to the production of CH₄ instead of H₂; hence, some authors challenge the H₂ method sensitivity and favor the evaluation of both intestinal gases.

Aim: To compare different approaches to usage of a lactose breath test for lactose malabsorption diagnosis, after medical evaluation of gastrointestinal symptoms.

Methods: In a retrospective observational study, we compared the 2 approaches in a population of 282 subjects in Northern Italy. Following oral lactose administration, exhaled samples were harvested every 30 minutes for 4 hours and prepared for H₂ and CH₄ analysis. Basal gas levels were subtracted from H₂ and CH₄ ppm and values at 4 hours and peaks were considered for analysis.

Results: Applying the standard methodology, which takes separately into consideration H₂ and CH₄ produced in the intestinal lumen, the results indicated that 11.7% of the patients were diagnosed “positive” for hypolactasia, differently from what was expected. Conversely, taking into consideration the sum of H₂ and CH₄, the percentage increased to 62.8%, closer to the expected one. No significant differences were found when comparing the 2 groups for age, gender, or symptoms. The sizable difference between the 2 approaches is likely linked to gut microbiome variability, and consequently the different production of the 2 gases, in the population.

Conclusion: The threshold normally used for lactose breath test should be reconsidered and changed, merging H₂ and CH₄ stoichiometric values to increase sensitivity.

Keywords: lactose malabsorption, intestinal malabsorption, hypolactasia, H₂ + 2CH₄ threshold

Introduction
Primary lactose intolerance is a widely diffused gastrointestinal disorder in which the organism cannot completely digest the sugar present in milk and its derivatives, leading to symptoms such as flatulence, meteorism, diarrhea, intestinal pain, and abdominal cramps.¹–³ In particular, lactose intolerance is the most common form of carbohydrate malabsorption and affects people of all ages.⁴–⁶ The cause of this malabsorption is due to the deficiency of lactase, an enzyme that is normally present in the small intestine. Under normal conditions, after ingestion, lactose is hydrolyzed by lactase in the intestine (the highest activity of this enzyme is in jejunum) to form glucose and galactose. The 2 monosaccharides are then actively absorbed by the intestinal mucosa. In case...
of lactase deficiency, lactose reaches the colon where it is fermented by the bacterial flora to produce various gases, including hydrogen (H₂) and methane (CH₄). These, in part, are absorbed and subsequently eliminated through the lungs.⁷,⁸ Given that the human organism does not produce either H₂ or CH₄ when administering lactose in controlled conditions to lactase-deficient patients, the amount of gases measured in the subject’s breath is considered a measure of the unabsorbed sugar. These “Breath Tests,” after lactose challenge, are considered a standard diagnostic approach to identify lactase deficiency.⁷–¹² Lactose malabsorption is diagnosed when breath test values are above a specific threshold value of either H₂ or CH₄ or both.⁹,¹⁰,¹³ A consensus conference back in 2009 had suggested a standardization of the analytical method to be used for lactose breath tests, with a focus mainly on H₂ intestinal production and a relatively high threshold value for positives, thus privileging specificity over sensitivity.¹⁴ During the consensus conference, the issue of intestinal variability of H₂/CH₄-producing flora was not taken into account.¹⁴ The 2017 North American consensus conference did not add much to the specific issue, confirming the guidelines.¹³ Nevertheless, a limited sensitivity can lead to missing or delaying lactate malabsorption diagnosis. Some authors have questioned the Rome guidelines and propose to increase the sensitivity of the assay, promoting analytical approaches that take into consideration equally the production of both gases and the high variability of the subject’s intestinal microbiome, with an unpredictable proportion of H₂-producing microbes, versus CH₄-producing ones.¹⁵,¹⁶ The aim of the present study was to compare different analytical approaches of the same lactose breath test in a retrospective study of patients prescribed the test after medical evaluation of gastrointestinal symptoms for a suspected lactose malabsorption condition.

The genetic variability of intestinal lactase is responsible for lactase persistency in the adult and is extremely variable in the general population, ranging from almost 0 in some Asian Countries to high levels found in Northern Europe.⁴,⁵,⁹,¹⁷,¹⁸ The lack of lactase persistency is considered responsible for lactose malabsorption. Recently published studies took showed the incidence of lactose malabsorption among Italian population to be 72%.¹⁹ Another study recently published by Zadro et al.²⁰ investigated the frequency of lactase non-persistence genotype among Italian population. The results showed 62.3% of lactase non-persistence genotype, with no significant differences among 3 macroregions of the country (58.6%, 74.1%, and 67.1% detected in North, Center, and South, respectively).

Our study is, to date, the first lactose breath test study conducted in North East Italy on a population of patients suspected of lactose malabsorption.

Materials and methods
Subjects and sample collection
Diagnostic data were retrospectively collected, pooled, and analyzed from lactose breath tests of outpatient subjects referring to a community-based laboratory facility in the Veneto Region in Northeast Italy, Data Medica Padova, from January 2014 to June 2016. The facility was part of a group of private health care providers operating under agreement with the public health care system to deliver healthcare to the public. The center was also ISO9001 certified and accredited for excellence in health care services by Accreditation Canada International. The retrospective observational study was carried out according to Good Practice in Clinical Research, approved by the Data Medica Group Scientific Technical Committee, and, following national Italian legislation, approved by the local reference Ethics Committee for Clinical Trials of the Province of Padua (DM 2016/01). All study participants provided informed written consent prior to study enrollment.

The analysis involved 282 subjects whose breath tests results were obtained from the laboratory database and rendered anonymous before elaboration. For each patient, age, sex, self-declared reason for taking the exam (grouped into “intestinal pain,” “meteorism,” “nausea,” “diarrhea,” “no symptoms specified”), and breath test diagnostic outcome were collected. H₂ values at 4 hours and H₂ and CH₄ peak values were also considered for further analysis.

Breath test
Before the test, patients had to fast and follow a series of dietary, pharmacological, and behavioral prescriptions, as indicated by the manufacturer, and on the day of the assay they were asked to fill a questionnaire with identification data and the reasons for the examination. EXPIROLact® H₂ Breath Test Kit for the determination of lactose intolerance (Sofar SpA, Milan, Italy) was used for the assay, and a standard procedure furnished by the kit manufacturer was followed. Briefly, following the oral administration of 25 g of lactose suspended in 150 mL of tap water, exhaled samples were harvested every 30 minutes for 4 hours and prepared for H₂ and CH₄ analysis.¹⁰ Basal gas levels were subtracted from H₂ and CH₄ ppm, and values at 4 hours and peaks were considered for analysis. No sample had peaks of H₂ values above 20 ppm within 1.5 hours from lactose ingestion, which
would have suggested small intestinal bacterial overgrowth and would have been eliminated from the study.

The analysis was conducted with a Quintron microlizer BreathTracker Analyzer (QuinTron Instrument Company, Inc., Milwaukee, WI, USA), based on solid-state sensors.

Statistical analysis
Statistic analyses were performed using GraphPad Prism (v 5.01; Graphpad Inc., La Jolla, CA, USA). Statistical comparisons for subgroups were computed using 1-way analysis of variance followed by post hoc tests for selected comparisons or 2-tailed Student’s t-test.

Results
A retrospective analysis was performed on a sample of 282 patients who underwent a lactose breath test between 2014 and 2015 in an outpatient laboratory in Northern Italy. Patients with a diagnosis of major gastrointestinal diseases were excluded from this study. Mean age was 35 years, and females were twice the number of males enrolled in the study. Most patients were sent for the test by a general practitioner or a specialist, with suspected hypolactasia potentially linked to 1 or more gastrointestinal symptoms as declared initially by the patients themselves (Table 1 provides a sample description). None of the samples considered for the present study had unacceptable baseline values (≥20 ppm). Two protocols were followed to determine lactose malabsorption. One is based on the maximum ppm values at 4 hours following lactose challenge of either H₂ or CH₄ (Threshold >20 ppm), and is the method ordinarily used. The second approach took into consideration the stoichiometric sum of the ppm values of both H₂ and CH₄ (H₂+2CH₄) (Threshold >18 ppm) in the same time frame following lactose oral administration. Since in some cases peak values did not correspond to those determined at 4 hours, we further collected those values for comparison.

As shown in Table 2 and Figure 1, the percentage of positive samples using the first method was 11.4%, with an average value of H₂ at 4 hours of 11.34 ppm. When using the stoichiometric sum of H₂ and CH₄, the number of positives dramatically increases up to 62.8%. Ppm values obtained using the peak values were clearly higher, although nonsignificantly (p-value >0.05). Nevertheless in the latter case the number of positive subjects increased, more so when using the second method, a difference leading to a significant diagnostic impact.

Relevant differences are thus seen between the different analytical approaches, while the same qualitative differences were found when comparing the positive samples from different subgroups of patients (Tables 3–5). When considering 5 different age ranges, all the analytical approaches show the same trends, with a higher value below 20 years, which lowers and increases again reaching top ppm values when considering the higher age range (Table 3). No differences were seen between male and females in all methods (Table 4), and the same distribution was also seen for symptomatology among methods (Tables 5 and 6), with no subgroup significantly higher than the others within each method (p-value >0.05). Hence, results suggest that the different methods have only a difference in sensitivity, and do not differ in qualitative terms.

Discussion
The present retrospective observational study was meant to give a contribution to the discussion on the diagnostic application of lactose breath test. In particular, the aim of the study was to establish the frequency of positive lactose breath tests in a group of patients affected by gastrointestinal symptoms who took the test in a local outpatient ambulatory center in Northeastern Italy. Results unexpectedly indicated...
Figure 1: Histogram reporting the percentages of patients who showed “positive” results to the breath test analysis.

Notes: Comparisons were performed between measurements of \( \text{H}_2 \) and \( \text{H}_2 + 2(\text{CH}_4) \), monitored at 4 hours after lactose challenge and at \( C_{\text{max}} \).

Abbreviation: \( C_{\text{max}} \), peak concentration.

Table 3: % of positives and mean ppm values among different age groups

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>% Positive (average ppm ± SD)</th>
<th>( \text{H}_2 ) (4 hours)*</th>
<th>( \text{H}<em>2 ) (( C</em>{\text{max}} ))*</th>
<th>( \text{H}_2 + 2\text{CH}_4 ) (4 hours)b</th>
<th>( \text{H}_2 + 2\text{CH}<em>4 ) (( C</em>{\text{max}} ))*</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;20</td>
<td>15.6 (33.1±5.0)</td>
<td>16 (32.1±5.6)</td>
<td>51.1 (37.0±14.0)</td>
<td>64.4 (34.5±13.1)</td>
<td></td>
</tr>
<tr>
<td>20–40</td>
<td>11.3 (35.5±10.6)</td>
<td>11.3 (35.5±10.7)</td>
<td>48.1 (32.4±15.9)</td>
<td>59.3 (31.1±15.0)</td>
<td></td>
</tr>
<tr>
<td>40–60</td>
<td>11.0 (34.3±4.5)</td>
<td>11.0 (34.3±4.8)</td>
<td>53.7 (31.0±12.2)</td>
<td>68.3 (30.6±11.2)</td>
<td></td>
</tr>
<tr>
<td>&gt;60</td>
<td>0 (–)</td>
<td>0 (–)</td>
<td>30.0 (23.3±5.5)</td>
<td>50.0 (24.8±3.9)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: *Threshold >20 ppm; bthreshold >18 ppm.

Abbreviation: \( C_{\text{max}} \), peak concentration.

Table 4: % of positives and mean ppm values between the sexes

<table>
<thead>
<tr>
<th>Sex</th>
<th>% Positive (average ppm ± SD)</th>
<th>( \text{H}_2 ) (4 hours)*</th>
<th>( \text{H}<em>2 ) (( C</em>{\text{max}} ))*</th>
<th>( \text{H}_2 + 2\text{CH}_4 ) (4 hours)b</th>
<th>( \text{H}_2 + 2\text{CH}<em>4 ) (( C</em>{\text{max}} ))*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>11.9 (11.2±9.6)</td>
<td>11.9 (13.3±8.9)</td>
<td>51.1 (21.8±14.9)</td>
<td>59.8 (24.5±14.1)</td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td>10.6 (11.4±10.2)</td>
<td>11.2 (13.5±9.4)</td>
<td>50.0 (21.1±15.2)</td>
<td>64.4 (24.0±14.2)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: *Threshold >20 ppm; bthreshold >18 ppm.

Abbreviation: \( C_{\text{max}} \), peak concentration.

Table 5: % of positives and mean ppm values among patients grouped by common symptoms (H\(_2\) test results)

<table>
<thead>
<tr>
<th>Symptoms</th>
<th>Total</th>
<th>( \text{H}_2 ) (4 hours) N</th>
<th>% Positive</th>
<th>ppm</th>
<th>( \text{H}<em>2 ) (( C</em>{\text{max}} )) N</th>
<th>% Positive</th>
<th>ppm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intestinal pain</td>
<td>102</td>
<td>15</td>
<td>15.2</td>
<td>36.6±6.4</td>
<td>15</td>
<td>15.2</td>
<td>36.6±6.3</td>
</tr>
<tr>
<td>Meteorism</td>
<td>90</td>
<td>11</td>
<td>12.3</td>
<td>33.2±5.1</td>
<td>11</td>
<td>12.3</td>
<td>33.1±5.0</td>
</tr>
<tr>
<td>Nausea</td>
<td>15</td>
<td>2</td>
<td>7.1</td>
<td>30.7±6.6</td>
<td>2</td>
<td>7.2</td>
<td>31.2±6.7</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>72</td>
<td>10</td>
<td>14.1</td>
<td>38.4±10.2</td>
<td>10</td>
<td>14.2</td>
<td>38.2±11.5</td>
</tr>
<tr>
<td>No symptoms</td>
<td>84</td>
<td>8</td>
<td>10.6</td>
<td>34.3±6.3</td>
<td>9</td>
<td>11.1</td>
<td>32.7±6.7</td>
</tr>
<tr>
<td>Whole sample</td>
<td>32</td>
<td>3</td>
<td>1.4</td>
<td></td>
<td>3</td>
<td>11.7</td>
<td></td>
</tr>
</tbody>
</table>

Note: Data presented as mean ± SD unless otherwise stated.

Abbreviation: \( C_{\text{max}} \), peak concentration.
a rather low frequency of positive tests when the standard analytical methodology was used.9 The same results were thus reevaluated using a different approach used by some other authors,15 which was characterized by a higher sensitivity and the same level of precision.15 This approach led to a much higher proportion of positive samples, as one would have expected in a population of gastrointestinal patients suspected of lactose intolerance. Lactase deficiency is considered a rather frequent condition, though it is known that it is highly variable in the general population and dependent on race and latitude,4,18 making it difficult to establish fixed rules for diagnostic appropriateness without a precise figure of what to expect in a given geographical area. Recent studies have been conducted on the Italian local population, showing no significant variations in the frequency of lactase nonpersistence genotype among 3 macro-areas of the country.20 The values of the frequency of this genetic condition affecting lactose absorption are expected to fall in the range of 60%–70%, according to the studies of Storhaug et al19 and Zadro et al.20 When we applied the standard analytical methodology, which takes into consideration H2 and CH4 produced in the intestinal lumen separately, we identified only 11.7% of the patients as positive for the lactose breath test, a figure rather different from what was expected. When we changed the analytical approach, taking into consideration the sum of the 2 bacterial gases produced from lactose by the intestinal flora, the result changed and approached closer to the expected one (62.8%).

When lactose reaches the intestinal lumen, the first bacterial population encountered produces H2 as an end product. If lactose escapes this first level, it can be metabolized by a different microbial population, producing CH4 as metabolic end product.15 As a consequence, these 2 molecules together, and not separately, account for the total amount of lactose that is not absorbed. When considering them singularly, 2 problems could arise: first of all, variability in the results linked to variability of microbiome in different patients. A second issue is linked to low sensitivity, potentially leading to the identification of a lower amount of lactose malabsorption cases, with important consequences on patients’ health and therapeutic options. Conversely, the sum of the 2 molecules, stoichiometrically corrected (since the hydrogen produced from lactose is twice as much as the amount of CH4), is not dependent on the microbiome variability, since it includes both types of microbial populations. The first consensus meeting on the subject in 2009 recommended the use of either H2 or CH4 to determine positivity to the lactose breath test,14 but since then others have challenged the decision, while the 2017 consensus did not touch upon the issue.15 In particular, our results indicate that the 2 analytical approaches yield 2 different diagnostic outcomes. While the quantitative difference between the 2 tests is very high, the 2 methods do not seem to carry qualitative differences, at least in the parameters considered in the present retrospective study. Considering different age groups, the frequency trends are similar, and there are no significant differences in the sex stratification of the positive samples. Similarly, we could not find significant differences in the stratification of the groups according to the symptoms declared by the patients.

**Conclusion**

Our results suggest that the analytical methodology used and the threshold for lactose breath tests should be reconsidered, especially in view of the extreme variability known to affect intestinal microbiome, a condition that can greatly affect the diagnostic outcome. Taking into account previous studies, indicating a significant underestimation of lactase genetic nonpersistence when using H2/CH4 standard breath tests,11 our results can contribute to the discussion on the subject that has relevant consequences on lactose malabsorption diagnosis that might be missed.

### Table 6 % of positives and mean ppm values among patients grouped by common symptoms (H2+2CH4 test results)

<table>
<thead>
<tr>
<th>Symptoms</th>
<th>Total</th>
<th>H2+2CH4 (4 hours)</th>
<th></th>
<th>H2+2CH4 (Cmax)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>% Positive</td>
<td>ppm</td>
<td>N</td>
<td>% Positive</td>
</tr>
<tr>
<td>Intestinal pain</td>
<td>102</td>
<td>53</td>
<td>52.1</td>
<td>33.2±15.0</td>
<td>66</td>
</tr>
<tr>
<td>Meteorism</td>
<td>90</td>
<td>41</td>
<td>46.4</td>
<td>33.2±13.7</td>
<td>50</td>
</tr>
<tr>
<td>Nausea</td>
<td>15</td>
<td>8</td>
<td>53.4</td>
<td>32.4±9.5</td>
<td>10</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>72</td>
<td>32</td>
<td>44.2</td>
<td>36.1±19.0</td>
<td>45</td>
</tr>
<tr>
<td>No symptoms</td>
<td>84</td>
<td>48</td>
<td>57.0</td>
<td>30.7±13.2</td>
<td>59</td>
</tr>
<tr>
<td>Whole sample</td>
<td>139</td>
<td>49.3</td>
<td>177.0</td>
<td>62.8</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** Data presented as mean ± SD unless otherwise stated.

**Abbreviation:** Cmax, peak concentration.
Availability of data and materials
The data that support the findings of this study are available from Data Medica Group, Padova, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Data Medica Group, Padova.

Acknowledgment
We thank the laboratory staff in Data Medica for helping with data collection, in particular Dr Giovanni Lenzo and Dr Alessia Caltabiano.

Disclosure
The authors report no conflicts of interest in this work.

References