13CO2 breath tests: comparison of isotope ratio mass spectrometry and non-dispersive infrared spectrometry results

François MION (1), René ECOCHARD (2), Jérôme GUITTON (3), Thierry PONCHON (1)
(1) Fédération des Spécialités Digestives, Hôpital Edouard-Herriot, Lyon ; (2) Département d’Information Médicale, Hospices Civils, Lyon ; (3) Fédération de Biochimie, Hôpital Ed. Herriot, Lyon.

SUMMARY
Background and aims — Isotope ratio mass spectrometry is the standard analytical method for 13CO2 breath tests. The goal of this study was to compare the results of 13CO2 breath tests obtained by non dispersive infrared spectrometry, a new, simpler and cheaper method, with those obtained by the gold standard method.

Methods — Three hundred and eight patients were included: 150 underwent a urea breath test for Helicobacter pylori diagnosis, 140 an aminopyrine breath test to measure liver function, and 18 an octanoic acid breath test for gastric emptying evaluation. A total of 750 breath samples were obtained in duplicate for isotope ratio mass and infrared spectrometry analyses. Breath test results were compared using Bland-Altman plots.

Results — The agreement between the two methods was excellent for urea breath tests (kappa coefficient = 0.96), with only 3 discordant results. Although 13C isotopic enrichment in breath was significantly lower with infrared spectrometry (P < 0.0001), the agreement for the results of aminopyrine and octanoic acid breath tests was excellent. The clinical significance of these results was similar for both methods.

Conclusions — Infrared spectrometry results are comparable to isotope ratio mass spectrometry. Because this analytical method is simpler and less expensive, it could be used for clinical applications of 13CO2 breath tests.

Key words: Breath test. Stable isotopes. Helicobacter pylori. Liver function tests. Gastric emptying.

RéSUMÉ
Tests respiratoires au carbone 13 : comparaison des résultats obtenus par spectrométrie de masse isotopique et spectrométrie infrarouge
François MION, René ECOCHARD, Jérôme GUITTON, Thierry PONCHON
(Gastroenterol Clin Biol 2001;25:375-379)
La spectrométrie de masse isotopique est la méthode de référence pour la mesure du 13CO2 dans les gaz expirés. Le but de cette étude était de comparer les résultats de tests respiratoires au carbone 13, obtenus par spectrométrie infrarouge, une nouvelle technique plus simple et moins coûteuse, avec ceux obtenus par la méthode de référence.


Résultats — La concordance était excellente pour le test à l’urée (coefficient kappa = 0.96), avec seulement 3 résultats discordants. Bien que l’enrichissement isotopique en 13C dans les gaz expirés soit significativement plus faible en spectrométrie infrarouge (P < 0.0001), la concordance entre les 2 méthodes pour les tests à l’aminopyrine et à l’acide octanoïque était excellente. La signification clinique de ces 2 tests était identique pour les 2 méthodes d’analyse.

Conclusions — La spectrométrie infrarouge donne des résultats comparables à ceux de la spectrométrie de masse isotopique pour les tests respiratoires. Du fait de sa simplicité et de son moindre coût, cette nouvelle technique d’analyse peut être utilisée pour la pratique clinique des tests respiratoires au carbone 13.


13CO2. About 100 mL of unpurified breath samples are needed for the analysis, and the results are calculated and expressed in the same delta ‰ units as IRMS.

The goal of our study was to compare the isotopic analysis results obtained by NDIRS and IRMS in three types of 13C breath tests: a) 13C-Urea breath test for the diagnosis of H. pylori infection [2]; b) 13C-aminopyrine breath test for the quantitative evaluation of liver function [6]; and c) 13C-octanoic acid breath test for the measurement of gastric emptying [7].

Materials and methods

Between January and November 1999, 308 consecutive breath tests were performed in our GI clinic: 150 urea breath tests (UBT), 140 aminopyrine breath tests (ABT) and 18 octanoic acid breath tests (OBT) in 308 different patients (table I): a total of 750 breath samples was thus obtained. UBT were performed for H. pylori infection screening (77 cases) or to assess the efficacy of anti-H. pylori treatments (73 cases). ABT were performed to evaluate liver function before liver transplantation, and OBT to measure gastric emptying in dyspeptic or diabetic patients.

Breath test procedures

Breath samples were obtained for each time point in duplicate: one glass tube (10 mL Vacutainer®) for IRMS analysis, and one breath bag (300 mL) for NDIRS analysis. All tests were carried out after an overnight fast. For UBT, 75 mg of 13C-urea (Eurisotop, Saclay, France) dissolved in water were given orally, 5 min after the ingestion of 200 mL of a 0.1N citric acid solution to slow down gastric emptying [8]. Breath samples were obtained before and 30 min after the ingestion of urea. The results of the UBT were expressed as the isotopic difference over baseline values (DOB = T30-TB) in delta ‰, with a cut-off value for negative and positive H. pylori status of 3 delta ‰ [9]. For ABT, 140 mg of 13C-aminopyrine (Eurisotop, Saclay, France) dissolved in water were given orally, and breath samples obtained before and 60 min after the administration of the labeled substrate. The results of ABT were expressed as the DOB (T60-TB) in delta ‰ [6]. For OBT, 100 mg of 13C-octanoic acid (Eurisotop, Saclay, France) were ingested together with a test meal as described by Ghoos et al. [7]. Breath samples were obtained before and every 30 min after ingestion of the test meal until 240 min. Gastric emptying coefficient (GEC), half-emptying time (t1/2) and lag phase (tlag) were calculated as described by Ghoos et al. [7].

Analytical methods

IRMS analysis was performed with a SIRA 10 prototype (VG Isotech, Middlewich, UK). Isotopic ratio for each sample is given in delta ‰, a standard expression of 13C abundance against a defined standard. Delta ‰ values are either positive or negative from the standard according to the following formula:

\[ \Delta \% = \frac{(13C/12C) - (13C/12C)_s}{(13C/12C)_s} \times 1000 \]

with s for breath sample and r for reference gas [10].

NDIRS analysis was performed with the FANc1 2 system (Olympus-Europe, Hamburg, Germany), isotopic ratio being also expressed in delta ‰. The system also measures the amount of CO2 present in each sample (in %).

Statistical methods

A linear regression analysis was carried out to compare the DOB values obtained by both methods on the 150 urea breath tests and the 140 aminopyrine breath tests (Statview, Abacus Concept, Berkeley, California). Bland and Altman plots [11] were used to assess the concordance of breath test results. For UBT, which is the only test to give a negative/positive result, the agreement between both analytical methods was studied by calculating the kappa value [12]. Differences between DOB results obtained by IRMS and NDIRS for ABT and OBT were compared by the Student’s t test for paired data. Results were expressed as mean ± standard deviation. Statistical significance was set at P < 0.05.

Results

Urea breath test results

One hundred fifty results of UBT expressed in DOB were obtained. A direct comparison of NDIRS and IRMS DOB values is represented in figure 1. The linear regression equation obtained, y = 0.25 + 1.03 x; R2 = 0.992, yielded the following informations: a) DOB values measured by NDIRS were identical to those obtained by IRMS, as indicated by the value of the constant term of the equation (0.25, CI 95% = −0.80 to + 0.58; P > 0.05); b) differences between the two methods increased

Table I – Demographical data and main indications of the breath tests for the 308 patients included in the study. (UBT: urea breath test; ABT: aminopyrine breath test; OBT: octanoic acid breath test).

<table>
<thead>
<tr>
<th>Indications (%)</th>
<th>UBT</th>
<th>ABT</th>
<th>OBT</th>
</tr>
</thead>
<tbody>
<tr>
<td>H. pylori diagnosis</td>
<td>77 (51.3)</td>
<td>54 (38.6)</td>
<td>13 (72.2)</td>
</tr>
</tbody>
</table>
| Alcoholic cirrhosis | 40 (28.5) | 140 aminopyrine breath tests (Statview, Abacus Concept, Berkeley, California). Bland and Altman plots [11] were used to assess the concordance of breath test results. For UBT, which is the only test to give a negative/positive result, the agreement between both analytical methods was studied by calculating the kappa value [12]. Differences between DOB results obtained by IRMS and NDIRS for ABT and OBT were compared by the Student’s t test for paired data. Results were expressed as mean ± standard deviation. Statistical significance was set at P < 0.05.

Results

Urea breath test results

One hundred fifty results of UBT expressed in DOB were obtained. A direct comparison of NDIRS and IRMS DOB values is represented in figure 1. The linear regression equation obtained, y = 0.25 + 1.03 x; R2 = 0.992, yielded the following informations: a) DOB values measured by NDIRS were identical to those obtained by IRMS, as indicated by the value of the constant term of the equation (0.25, CI 95% = −0.80 to + 0.58; P > 0.05); b) differences between the two methods increased

Table I – Demographical data and main indications of the breath tests for the 308 patients included in the study. (UBT: urea breath test; ABT: aminopyrine breath test; OBT: octanoic acid breath test).

<table>
<thead>
<tr>
<th>Indications (%)</th>
<th>UBT</th>
<th>ABT</th>
<th>OBT</th>
</tr>
</thead>
<tbody>
<tr>
<td>H. pylori diagnosis</td>
<td>77 (51.3)</td>
<td>54 (38.6)</td>
<td>13 (72.2)</td>
</tr>
</tbody>
</table>
| Alcoholic cirrhosis | 40 (28.5) | 140 aminopyrine breath tests (Statview, Abacus Concept, Berkeley, California). Bland and Altman plots [11] were used to assess the concordance of breath test results. For UBT, which is the only test to give a negative/positive result, the agreement between both analytical methods was studied by calculating the kappa value [12]. Differences between DOB results obtained by IRMS and NDIRS for ABT and OBT were compared by the Student’s t test for paired data. Results were expressed as mean ± standard deviation. Statistical significance was set at P < 0.05.

Results

Urea breath test results

One hundred fifty results of UBT expressed in DOB were obtained. A direct comparison of NDIRS and IRMS DOB values is represented in figure 1. The linear regression equation obtained, y = 0.25 + 1.03 x; R2 = 0.992, yielded the following informations: a) DOB values measured by NDIRS were identical to those obtained by IRMS, as indicated by the value of the constant term of the equation (0.25, CI 95% = −0.80 to + 0.58; P > 0.05); b) differences between the two methods increased
slightly along the entire range of DOB values, as shown by the value of the slope of the regression curve (1.03, CI 95% = 1.01 - 1.05); c) the amplitude of the dispersion of the values along the equality line was small, as shown by a squared correlation coefficient of 0.992. To put it differently, the value of the squared correlation coefficient indicates that 99.2% of IRMS values are known from NDIRS results.

The scatter of differences increased with larger DOB value, especially above 20 delta‰. Most of the negative UBT tests were distributed between 0 and 1 delta‰, while positive UBT tests had DOB values greater than 3.5 delta‰. The final UBT results (H. pylori positive or negative) obtained by IRMS and NDIRS are summarized in Table II. The value of the kappa coefficient is 0.96, indicating an excellent concordance between the two methods [13]. The results of the 3 discordant cases are summarized in Table III. In patient #140, IRMS DOB value was just below the cut-off value (4.9 instead of 3 delta‰), while NDIRS DOB value was above the cut-off (4.9 delta‰). No gastric biopsy specimen was available in this case to assess the real H. pylori status. For the 2 other patients with discordant results, the H. pylori status as determined by NDIRS was confirmed by the pathological analysis of gastric antrum biopsies obtained during an upper GI endoscopy performed 4 (patient #103) and 6 weeks (patient #130) after the UBT.

### Aminopyrine breath test results

ABT results are expressed as delta over baseline values (DOB = T60 - Tbaseline) in delta‰, and give a quantitative expression of the liver functional mass [6]. A direct comparison of the 140 DOB values obtained by NDIRS and IRMS is represented in Figure 2. The linear regression equation obtained, \( Y = 0.43 + 1.02X \) (\( R^2 = 0.982 \)), gave the following informations: i) DOB values obtained by NDIRS were significantly lower than those obtained by IRMS, as shown by the constant term of the equation (0.43, 95% CI: 0.27 to 0.60, \( P < 0.0001 \)); ii) however, this difference remained constant along the entire range of DOB values as shown by the value of the slope of the regression line (1.02, 95% CI = 0.99 - 1.04). Figure 3 plots the difference in DOB between IRMS and NDIRS against the mean DOB value for each aminopyrine test. When examining the data by the paired t-test, the mean difference in DOB by the 2 methods was significantly different from 0 (0.52 delta‰, \( P < 0.0001 \)), confirming that DOB values obtained by IRMS were significantly greater than those obtained by NDIRS.

### Table II – Comparison of UBT results obtained by IRMS and NDIRS in 150 patients, with a cut-off DOB value of +3 delta‰.

<table>
<thead>
<tr>
<th>Patient number</th>
<th>UBT indication</th>
<th>IRMS (DOB value)</th>
<th>NDIRS (DOB value)</th>
<th>Pathological analysis of gastric Biopsy</th>
</tr>
</thead>
<tbody>
<tr>
<td>#103 — Woman 70 year-old</td>
<td>control after anti-H. pylori treatment</td>
<td>negative (1.28 delta‰)</td>
<td>positive (7.32 delta‰)</td>
<td>mild H. pylori infection</td>
</tr>
<tr>
<td>#130 — Woman 63 year-old</td>
<td>control after anti-H. pylori treatment</td>
<td>positive (5.20 delta‰)</td>
<td>negative (1.96 delta‰)</td>
<td>no H. pylori infection</td>
</tr>
<tr>
<td>#140 — Woman 81 year-old</td>
<td>screening (dyspepsia)</td>
<td>negative (2.96 delta‰)</td>
<td>positive (4.90 delta‰)</td>
<td>not available</td>
</tr>
</tbody>
</table>

### Table III – Urea breath test: results in 3 patients with discordant H. pylori status as assessed by IRMS and NDIRS.

<table>
<thead>
<tr>
<th>Patient number</th>
<th>UBT indication</th>
<th>IRMS (DOB value)</th>
<th>NDIRS (DOB value)</th>
<th>Pathological analysis of gastric Biopsy</th>
</tr>
</thead>
<tbody>
<tr>
<td>#103 — Woman 70 year-old</td>
<td>control after anti-H. pylori treatment</td>
<td>negative (1.28 delta‰)</td>
<td>positive (7.32 delta‰)</td>
<td>mild H. pylori infection</td>
</tr>
<tr>
<td>#130 — Woman 63 year-old</td>
<td>control after anti-H. pylori treatment</td>
<td>positive (5.20 delta‰)</td>
<td>negative (1.96 delta‰)</td>
<td>no H. pylori infection</td>
</tr>
<tr>
<td>#140 — Woman 81 year-old</td>
<td>screening (dyspepsia)</td>
<td>negative (2.96 delta‰)</td>
<td>positive (4.90 delta‰)</td>
<td>not available</td>
</tr>
</tbody>
</table>

### Fig. 2 – Comparison of DOB values obtained by NDIRS and IRMS on 140 aminopyrine breath tests.

Comparison des valeurs en DOB obtenues par NDIRS et IRMS sur 140 tests respiratoires à l’aminopyrine-\(^{13}\)C.

### Fig. 3 – Aminopyrine breath test: Bland-Altman plot of the difference in DOB values between the 2 methods against the average DOB values in 140 patients. The thick horizontal bar represents the mean of differences (0.52 delta‰), and the dotted horizontal bars the limits of agreement (\( \pm 2 \) SD; from –0.82 to 1.86 delta‰).

Test à l’aminopyrine : graphique de Bland-Altman rapportant la différence des valeurs de DOB entre les 2 méthodes sur la moyenne des DOB obtenues par les 2 méthodes chez 140 patients. La ligne épaissie indique la moyenne des différences (0.52 delta‰) et les 2 traits pointillés les limites de l’agrément (moyenne des différences ± 2 écarts-types : de –0.82 à 1.86 delta‰).
synthetically 13C enriched substrates can be used as tracers in
natural range of variation of this stable isotope: naturally or
rather complex in terms of running and maintenance. Thus,
besides being very sensitive, IRMS systems dedicated to breath
Europe, is one such system based on NDIRS methodology. Its
also effective to measure liver function with tests such as the
octanoic acid breath test [21]. Our study compared to IRMS for urea breath test [14, 18-20], and for
non-dispersive infrared [5, 14, 15] or laser spectrometry [16,
Several methods have been made recently available, based on
tests, reliable, cheaper and easy-to-run systems are needed.
routine analysis and widespread use of stable isotope breath
tests have gained popularity during the past
decade thanks to the development of powerful methods for
isotopic analysis, such as isotope ratio mass spectrometry. These
systems allow the detection of 13C isotopic enrichment within the
natural range of variation of this stable isotope: naturally or
synthetically 13C enriched substrates can be used as tracers in
humans to study different physiological functions [10]. However,
besides being very sensitive, IRMS systems dedicated to breath
analysis are still expensive (around 400 000 French Francs) and
rather complex in terms of running and maintenance. Thus,
13CO2 analysis remains limited to large analytical centers. For
routine analysis and widespread use of stable isotope breath
tests, reliable, cheaper and easy-to-run systems are needed.
Several methods have been made recently available, based on
non-dispersive infrared [5, 14, 15] or laser spectrometry [16,
Recent studies have shown these systems to be reliable
compared to IRMS for urea breath test [14, 18-20], and for
octanoic acid breath test [21]. Our study confirms that NDIRS is
also effective to measure liver function with tests such as the
aminopyrine breath test.

The FANc2 analyser, distributed in Europe by Olympus
Europe, is one such system based on NDIRS methodology. Its
price is roughly half of standard IRMS systems (around 200 000
French Francs). This user-friendly system runs on Windows 98,
functions at a constant temperature (50 °C) and pressure (1 000
hPa), which have been shown to be important parameters to
avoid spurious results [22]. Breath samples obtained in 300 mL
bags are then directly connected to the analyser without any further
sample preparation. The system uses about 100 mL of breath for
analysis, thus allowing a second measurement on the same
breath sample if needed. Because of the volumes of the bags,
NDIRS is adapted to test samples where they are collected. This
represents a limitation of the system compared to IRMS: because
of the small volumes of breath gas needed for IRMS (less than 100
µL), breath samples are collected in 10 mL glass tubes that can
then be sent by mail to an analytical center running IRMS.

The comparison of the results of the tests obtained by NDIRS
and IRMS showed a good agreement between these two different
analytical methods. For UBT, only 3 of 150 cases were found
discordant between the 2 analytical methods: interestingly,
NDIRS results were confirmed in 2 cases by pathological analysis
of gastric biopsy specimens. The reason for the false positive and
negative results by IRMS remains unknown. In the third case, the
IRMS value was very close to the cut-off (2.9 delta %), a situation
where the predictive value of the UBT is quite low (“grey zone” of
the test) [9]. We observed that the scattering of the differences
between the two methods increased for DOB values over 20 delta
%. Since these positive results are far away from the cut-off value,
the differences do not change the final result of the UBT.
Our data thus confirm previous studies showing a good agree-
ment between NDIRS and IRMS results [14, 19, 23].

For ABT and OBT, the results are only quantitative. Our
findings show a good agreement between results obtained by
IRMS and NDIRS. We observed a significant trend towards
higher DOB values with IRMS. Because this difference was
constant all over the DOB values spectrum, the clinical signi-
ficance of the 2 tests (normal or decreased liver function mass
for ABT; rapid, normal or slow gastric emptying for OBT) remained
identical whatever the method used for isotopic analysis. The sole
consequence of this observation would be the need to determine
the normal range for ABT and OBT for both analytical procedu-
res. Furthermore, it has been shown that for gastric emptying
studies, the calculated variables GEC, half-emptying time and lag
phase depend mainly on the kinetics of 13CO2 excretion curves
[24], rather than on the individual DOB values. For this reason,
the lower DOB values obtained at each time point with NDIRS
compared to IRMS do not affect significantly the final results of the
test.

In conclusion, our findings confirm that NDIRS gives similar
results as IRMS in all types of breath tests used to measure gut
function. This analytical method is simpler (it does not requires
vacuum and helium to function, as it is the case for IRMS), and its
fixed costs are significantly less. The two disadvantages related to
the use of NDIRS are: i) the price of the breath bags, slightly more
expensive than the glass tubes used for IRMS, and ii) the necessity
to have the NDIRS analyser near-by the place where the breath
tests are performed. NDIRS analysers could thus be used in
centers or laboratory facilities involved in breath 13C measure-
ments in relation with the clinical use of 13CO2 breath tests.

**REMERCIEMENTS - Ce travail a été réalisé grâce à une bourse de recherche Olympus-Europe (H. Drenel).
Une partie de ce travail a été présentée aux Journées Francophones de Pathologie Digestive, Nice, 20-22 mars 2000.**

**RÉFÉRENCES**


8. Graham DY, Runke D, Anderson SY, Malaty HM, Klein PD. Citric acid as the test meal for the $^{13}$C-urea breath test. Am J Gastroenterol 1999;94:1214-7.


