**Tricks for interpreting and making a good report on hydrogen and $^{13}$C breath tests**


**Department of Internal Medicine, Gastroenterology Division, School of Medicine Catholic University of the Sacred Heart, Rome, Italy**

**Abstract.** Breath tests (BT) represent a valid and non-invasive diagnostic tool in many gastroenterological disorders. Their wide diffusion is due to the low cost, simplicity and reproducibility and their common indications include diagnosis of carbohydrate malabsorption, Helicobacter pylori infection, small bowel bacterial overgrowth, gastric emptying time and orocoeal transit time.

The review deals with key points on methodology, which would influence the correct interpretation of the test and on a correct report.

While a clear guideline is available for lactose and glucose breath tests, no gold standard is available for Sorbitol, Fructose or other H$_2$ BTs.

Orococcal transit time (OCTT) defined as time between assumption of 10 g lactulose and a peak >10 ppm over the baseline value, is a well-defined breath test. The possible value of lactulose as a diagnostic test for the diagnosis of small bowel bacterial overgrowth is still under debate.

Among $^{13}$C breath test, the best and well characterized is represented by the urea breath test. Well-defined protocols are available also for other $^{13}$C tests, although a reimbursement for these tests is still not available.

Critical points in breath testing include the patient preparation for test, type of substrate utilized, reading machines, time between when the test is performed and when the test is processed. Another crucial point involves clinical conclusions coming from each test. For example, even if lactulose could be utilized for diagnosing small bowel bacterial overgrowth, this indication should be only secondary to orococcal transit time, and added into notes, as clinical guidelines are still uncertain.

**Key words:**

Breath test report, H$_2$ breath test, C$_{13}$ breath test.

**Introduction**

Breath tests (BT) represent a valid and non-invasive diagnostic tool in many gastroenterological disorders. Hydrogen-breath tests are based on the fact that gases produced by colonic bacterial fermentation diffuses into the blood and are excreted in breath samples, where it can be easily quantified. There is a growing interest in breath tests but many differences persist in the methodology used.

Hydrogen Breath test (HBT) has become an important diagnostic tool in many Gastroenterological conditions, characterized by abdominal pain, discomfort, bloating etc. Its wide diffusion is due to the low cost, simplicity and reproducibility: it represents a valid and non-invasive test to diagnose carbohydrate malabsorption, small bowel bacterial overgrowth, and for the determination of orococcal transit time.

$^{13}$C-breath tests use $^{13}$C-labelled substrates in order to perform a safe non-invasive evaluation of several metabolic pathways. The turnover of the $^{13}$C-substrate throughout the specific explored pathway is assessed by measuring $^{13}$CO$_2$ in exhaled air. The availability of plenty of $^{13}$C-substrates gives these tests a wide range of potential applications, although the routine clinical use of certain substrates may be sometimes affected by their high costs. The most used $^{13}$C breath test is the one for diagnosing *Helicobacter pylori* infection.

In this paper we summarize key element on assessing and interpreting H$_2$ and major $^{13}$C breath tests, and we will provide a guide for an accurate breath test report.

**Tricks on hydrogen and methane excretion**

Hydrogen isn’t produced in healthy humans fasting and at rest, but it is only generated during
anaerobic metabolism, which occurs, for example, during bacterial fermentation. Anaerobic bacteria metabolize sugar molecules, producing CO₂, short-chain fatty acid (SCFA) and Hydrogen. Unabsorbed carbohydrates are so fermented by colonic bacteria and the hydrogen, like other gases produced within the intestine, is absorbed through the intestinal wall, passes in the blood circulation, reaches the lungs, and finally it is released and exhaled in the breath samples. The amount of hydrogen produced by carbohydrate fermentation varies according to the individual’s microbiota, as hydrogen is also used by bacteria for other metabolic processes, including conversion to methane.

In fasting state, normal hydrogen baseline values are reported to be 7 ± 5 ppm², if it results too high (>10–16 ppm), breath test shouldn’t be performed. The ingestion of rice flour and meat is not associated with a detectable increase in breath hydrogen excretion, so they are eaten the evening before the test. In response to the test, some people produce huge amount of hydrogen in response to sugars, some produce lower amounts, and about 15% produce apparently none (designated ‘non-hydrogen producers’), as the expired hydrogen throughout the test dose not reach the limit detection rate. In some case of apparent hydrogen non production, methane determination could represent a valid alternative, although insufficient data are presented in this topic. In patients that do not produce methane neither hydrogen (a condition present in about 5% of the population), there is no indication of performing breath tests and no indication exists on whether increase test time would exert any role. This condition is mainly due to a different intestinal flora composition, which in turn is influenced by many factors including age, susceptibility to infections, dietary habits, immunological factors, intraluminal pH, interaction among components of intestinal flora and fermentable substrates, recent antibiotic therapy, use of laxatives (preparation for colonoscopy), or for the predominance of methane producing bacteria in the colon. Some groups try to overcome this problem by testing initially lactulose: this synthetic disaccharide cannot be absorbed in humans and is, therefore, always fermented producing hydrogen, except for the “non-H₂-producer”. It is, therefore, conceivable that a slow gastrointestinal transit of the substrate may be responsible for false negative results. It was shown that the prolongation of measurement from 2 to 4 hours induce a significant reduction of the prevalence of “H₂ non producer”.

**Tricks on sampling H₂ from alveolar expirates for H₂ breath testing**

We can measure hydrogen concentrations in end-expiratory breath samples using gas chromatography or electrochemical cells. Stationary dedicated gas chromatographs represent the gold standard for hydrogen determinations in breath, as they were previously validated in comparison with non-dedicated instruments in terms of linearity and reproducibility of results. Stationary dedicated gas-chromatographs could be made with solid state sensors, and they represent the gold standard for hydrogen determination, or with electrochemical sensor, which also show a good accuracy, although they will undergo to a natural “drift” over time.

Gas measurement must be performed in alveolar air, avoiding the interference represented by respiratory dead space air. The three systems are available to collect alveolar air: the modified Haldan-Priestley tube, the Y-Piece device and the two bag-system. The most efficient method to remove dead space air and guarantee complete respiratory exchanges is to inhale maximally, hold the inhalation for 15 s and expire into the bag. It could be also considered as a marker of correct sampling the concentration of CO₂, which levels in alveolar air are stable around 5–5.5%: it is possible to reduce variability due to the death space, measuring CO₂ in breath sample and correcting for the ideal value of CO₂.

Breath samples are currently stored in plastic syringes, an inexpensive method allowing the analysis of gases with no further handling. Because of the high diffusion capacity of hydrogen, samples should be analyzed as soon as possible, avoiding storage longer than 12 h; however, simple refrigeration of plastic syringes is sufficient to ensure the stability of hydrogen concentrations for a long time. At room temperature, after 5 days, the hydrogen concentration is reduced up to 30%, while at –20 °C the reduction is equal to 5% and only 7% after 15 days. Moreover, at –20 °C no hydrogen loss is detectable for 2 days.

Hydrogen concentration in breath samples must, therefore, be determined within 6 h of sampling.

**Tricks on Lactose breath test interpretation**

Lactose breath test shows good sensitivity (mean value of 77.5%) and excellent specificity (mean value of 97.6%) for lactose intolerance in patients which display or not symptoms. False negative tests, however, can be due to “hydrogen non-producers”; false positive breath tests are less frequent and are mainly secondary to of small
bacterial overgrowth. LBT was considered positive when the peak of hydrogen was 20 ppm over the baseline. In children some laboratories have used >10 ppm as cutoff value, but a rise of > 20 ppm over baseline correlates better with symptom development18.

If the value rises more than 10 ppm but less than 20 ppm above the base value, the test could be considered as “borderline”19. In this condition, it must be verified when this peak occurs: if the hydrogen value seems to increase at the end of the test, should be considered the possibility of a prolonged test, to augment the sensibility, reducing “false negative” due to a possible slow intestinal transit time.

**Tricks on other carbohydrates breath test interpretation**

Sorbitol is a sugar alcohol widespread in plants, particularly in fruits and juices, which is only partially absorbed. Sorbitol BT is effective in detecting small bowel damage, especially in celiac disease20. Tursi et al published many papers on Sorbitol Breath test (5 g) for the diagnosis of subclinical or silent celiac disease, and for the assess histological recovery after gluten-free diet21,22.

Fructose, widely used as a sweetener in different foods, is a monosaccharide, which is absorbed by a specific receptor for fructose (GIUT5 fructose transporter mediated facilitative diffusion), which occurs along the entire small bowel and it is slow; and by a second receptor which is rapid and efficient, but occurs only together with glucose6. The mean daily fructose intake is about 16 g per day, but eating enriched food and drinks can achieve up to 100 g per day23,24.

The clinical importance of sorbitol and fructose malabsorption remains to be evaluated, so BT should not be recommended in clinical practice in both adults and children, according to current guidelines14.

**Tricks on small intestinal bacterial overgrowth interpretation**

Small intestinal bacterial overgrowth (SIBO) is usually defined as a growth of >10⁵ colony forming units (CFU)/ml in fluid aspirated from the jejunum14, which could be associated to abdominal pain, bloating, diarrhea and/or signs of malabsorption, similar to those observed in irritable bowel syndrome (IBS)14,25,26. Glucose Breath test (GBT) is currently used for the diagnosis of SIBO. Glucose is rapidly absorbed in the proximal small bowel and usually does not reach the colon, so it is a suitable substrate to detect proximal small bowel overgrowth. A rise in H₂, after the assumption of the substrate, means that glucose meets bacteria in the small bowel, before its absorption. Because of its early absorption, GBT may not able to diagnose SIBO of the distal small intestine (ileum). Sensitivity and specificity are 62.5% and 77.8 %, respectively compared to the gold standard (jejunal culture)27.

Lactulose has also been used for diagnosis SIBO. Many criteria have been established to try to detect SIBO using Lactulose: some authors suggested that a rise in breath H₂ of 20 ppm above base levels within 90 minutes after ingestion of lactulose should be considered as diagnostic of SIBO20, but this criterion has not been validated. Actually the most diffuse criterion is the presence of a hydrogen “double peak”, the earlier representing presence of bacteria fermenting in the small bowel, the second and bigger one linked with the flora usually colonizing the colon. However, it is difficult to distinguish between an early H₂ peak caused by SIBO and a fast transit; thus, the non-standardized criteria, and lower diagnostic accuracy (55 versus 71 of Glucose14) makes lactulose breath test not indicated to diagnostic for SIBO.

**Tricks on orocaecal transit time interpretation**

Lactulose is a manufactured disaccharide, consisting of galactose and fructose, which is not metabolized by the human organism, producing always fermentation in the colon. Thus, the time from ingestion of lactulose to the hydrogen increase in breath samples reflects the orocaecal transit time. An increase >20 ppm compared to baseline is considered significant in diagnostic routine18. The orocaecal transit time in healthy subjects ranges between 40 and 170 min for a lactulose meal (192-232 min for a solid meal). A delayed orocaecal transit has been reported in constipation, alcoholism, depression, diabetes mellitus, IBS, pregnancy, obesity, cholecystectomy, cirrhosis, scleroderma, acromegaly and dyspepsia. Meanwhile, a fast transit in partial gastrectomy, post vagotomy diarrhea, hyperthyroidism, IBS, alcoholics. LHB is also useful to demonstrated drug effects on orocaecal transit.

**Tricks for sampling and measuring ¹³C/¹²C ratio**

¹³C stable isotope is used in breath testing to label specific substrates. The ¹³C-labelled tracer probe, once ingested, is cleaved in the gastroin-
testinal lumen, transported and degraded following the same metabolic pathways of its native isoform. The final result is the excretion of $^{13}$CO$_2$ through the alveolo-capillary membrane into the exhaled air, in which $^{12}$CO$_2$ amount is measured.

Since $^{13}$C stable isotope is commonly found in nature – accounting for 1.11% of all carbon atoms – and its amount may vary among individuals, it is necessary to refer results to a baseline value in order to get accurate results. Thus, total CO$_2$ production at baseline must be measured or estimated in each patient. Usually it is calculated with 300 mmol CO$_2$/m$^2$ body surface area/h$^{29}$. Moreover, factors influencing the endogenous CO$_2$ production (e.g. food ingestion, physical activity, respiratory diseases, thyroid dysfunctions, fever) should be considered.

Breath samples are collected at baseline and for a variable number of times and at different time intervals depending on the type of $^{13}$C breath test performed.

The tracer could be detected and expressed as a static variable, e.g. the delta over baseline [DOB: $\delta_{\text{sample}} - \delta_{\text{basal value before tracer application}}$] in urea breath test, or as a dynamic one, in tests where $^{13}$C/$^{12}$C is expressed as recovery rate per hour or percent dose recovery and more than 2 dosages are necessary and special algorithms are used. Moreover, in order to get accurate results, measurement or estimation of basal $^{13}$C value is recommended. $^{13}$C recovery is never complete as a substantial amount of tracer is retained in the carbon pool of the body. Thus, $^{13}$C breath test analysis is a semi-quantitative diagnostic tool.

Breath samples are analyzed by high-resolution mass spectrometers, which are able to measure the slight mass difference of one neutron between $^{13}$C-labelled carbon dioxide and the naturally most common carbon dioxide with the carbon isotope $^{12}$C, on the base of different light absorbance between the two isotopes. The precision of high-resolution mass spectrometers is very high, so very low isotopic quantity can be detected, permitting the use of small samples of exhaled air. On the other hand, mass spectrometers are highly expensive, they have relatively long analysis time and they require well-trained personnel for their use. The introduction of infrared technologies (non-dispersive isotope selected infrared spectrometers, NDIRS), based on the use of photoacoustic detectors, made this goal cheaper with adequate accuracy. Operating and handling of NDIRS is easy, even for non-experienced users. This enables not only the performance, but also the diffusion of $^{13}$C-breath tests in primary care settings. Generally, there is a good agreement between mass spectroscopy and infrared based technology, for $^{13}$C/$^{12}$C ratio estimate, particularly for Urea breath test$^{30,31}$. The agreement of infrared technologies on dynamic $^{13}$C breath tests is not well addressed and some concerns remains on reproducibility and accuracy.

**Tricks on Helicobacter pylori breath test**

_Helicobacter pylori (H. pylori)_ is a common Gram-negative pathogen bacterium colonizing the gastric mucosal environment in about 50% individuals. $^{13}$C-urea breath test [UBT] is considered the gold standard for non-invasive diagnosis of _H. pylori_ infection, because of its accuracy and its easy execution$^{32}$. Urea hydrolysis can be generally entirely accounted to the presence of _H. pylori_, since it is the most common gastric pathogen expressing urease$^{33}$. Several studies reported the high diagnostic reliability of UBT, that reaches values >95% both for sensitivity and specificity prior and after eradication therapy$^{34-36}$. $^{13}$C-UBT protocol underwent multiple modifications through years. Nowadays, 75 mg $^{13}$C-urea in 200 ml water with the adjunction of 1-4 g citric acid$^{37}$ represents the most commonly accepted protocol. Adding citric acid enhances _H. pylori_ urea hydrolysis by both increasing intra-citoplasmic urea availability and enhancing urease activity$^{38}$. Hence it allows to spare urea and to contain costs. Breath samples are collected through a straw in a glass phial prior to the substrate administration and after 20 or 30 minutes.

The coexistence of some factors might alter UBT results affecting its accuracy. In particular, the assumption of proton pump inhibitors reduces UBT sensitivity in 31%$^{39}$ since it reduces the bacterial load especially in the antrum$^{32,40,41}$. False-positive tests may be provided by urease expressed by oropharyngeal flora. Mouth washing or administering the substrate with a straw or in pills may be helpful strategies to avoid this disadvantage.

Although performing the test in a non-fasting condition has been associated with a higher rate of both false-negative and false-positive results$^{42}$, several other studies found only small differences between DOB values in fasting and non-fasting status$^{43,44}$ or even no differences$^{45}$. However, it seems prudent to perform the test in a fasting status until this issue is better clarified.
The choice of a reliable cut-off to define the positivity of the test is another incompletely addressed question. A unique DOB cut-off is not definable because it has to be adapted to different factors, such as the test meal, doses and types of urea, or the pre-/post-treatment setting in which the test is employed. Although, because positive and negative UBT results tend to cluster outside of the DOB range between 1.8 and 5‰, a change in cut-off value within this range would be expected to have little effect on clinical accuracy of the test.

**Tricks on $^{13}$C octanoic acid breath for studying gastric emptying time**

Gastric emptying is a complex process requiring variable durations, depending on quantity and quality of food ingested, and on the presence of certain pathologies. In a clinical setting, a gastric emptying study might be particularly useful to assess how gastric physiology varies depending on age and diet, to study functional dyspepsia, diabetic gastroparesis and to evaluate the effect of prokinetic therapy. Radioscintigraphy represents the gold standard for gastric emptying studies.

Breath testing for gastric emptying assessing variable durations, depending on quantity and quality of food ingested, and on the presence of certain pathologies. In a clinical setting, a gastric emptying study might be particularly useful to assess how gastric physiology varies depending on age and diet, to study functional dyspepsia, diabetic gastroparesis and to evaluate the effect of prokinetic therapy. Radioscintigraphy represents the gold standard for gastric emptying studies.

Breath testing for gastric emptying assessing variable durations, depending on quantity and quality of food ingested, and on the presence of certain pathologies. In a clinical setting, a gastric emptying study might be particularly useful to assess how gastric physiology varies depending on age and diet, to study functional dyspepsia, diabetic gastroparesis and to evaluate the effect of prokinetic therapy. Radioscintigraphy represents the gold standard for gastric emptying studies.

In the original description by Ghoos et al., $^{13}$C-octanoic acid is mixed together with egg yolk and baked. The frying process anchors the molecule in the solid phase of the egg yolk. After its administration, when the meal has passed through the pylorus, the $^{13}$C-octanoic acid is rapidly absorbed in the duodenum and jejunum and is oxidised to $^{13}$C-labelled carbon dioxide, then eliminated by lungs. Collecting breath samples for more than 4 h gives reproducible $^{13}$CO$_2$-kinetics in breath, which reflects gastric emptying; thus the appearance of $^{13}$CO$_2$ in breathed air reflects the transit of the test meal from the stomach through the duodenum, where this is digested and absorbed with the final production of $^{13}$C-labelled carbon dioxide.

Several parameters are calculated to evaluate gastric emptying and this issue, requiring mathematical formulae, makes breath testing an indirect technique for the study of gastric emptying. Half-emptying gastric time ($t_{0.5}$) indicates the time required to dismiss half of the gastric content; gastric emptying coefficient [GEC] is an index of the gastric emptying rate; lag phase [tlag] is the time occurring for the solid meal to be triturated into <2 mm$^3$ fragments.

Test protocols vary depending on the test meal used. The main variations regard the formulation of test meals, caloric burden, time intervals between sample collections and the total duration of sampling. There is a reasonable agreement on a 4-6 hours-long sampling at time intervals of 15 or 30 minutes. Regarding the test meal, the standardisation of the test meals is required but to date it is still lacking. To this purpose, Perri et al. proposed a new standardised test meal in form of a gluten-, lactose-, glucose-free muffin.

$^{13}$C-acetate is used as marker in liquid phase studies because of the hydrophilic character of the molecule and has showed an excellent correlation toward scintigraphy. Liquid phase studies are particularly useful for research purposes, whilst in a clinical setting only certain neurological pathologies affect gastric emptying for the liquids.

Breath tests can be easily performed by children, pregnant women and critically ill patients. Their analysis does not require expert personnel operating in the same place where they are performed, thanks to the easy transportability of the samples. Although, breath-test and scintigraphy results are not identical, different results originate from different approaches in describing the gastric emptying. Radioactive tests study the progressive tracer disappearing from the stomach and are affected by tissue attenuation, angle of detection, radiation decay. $^{13}$C-Breath-tests study absorption of $^{13}$C-labelled molecules in the small bowel and are delayed pulmonary excretion and partial retention of the tracer amount in the body.

On the other hand, simultaneous labelling of solid and liquid phases, although could be done in nuclear medicine, is possible in breath-testing only with the addition of $^{14}$C, whose clinical use is obsolete because of its radioactivity.

**Tricks on $^{13}$C mixed triglyceride acid breath exploring exocrine pancreatic function**
Except for secretin-pancreozymin test and secretin-enhanced magnetic resonance pancreatography [s-MRP] that are both able to reach good levels of sensitivity and specificity, functional studying for the diagnosis of chronic pancreatitis plays a minor role today. Nevertheless, their high costs and invasiveness, makes useful to develop cheaper tests with similar accuracy.

Several substrates have been studied for assessing exocrine pancreatic disorders, e.g. $^{13}$C-starch, $^{13}$C-egg white and $^{13}$C-peptides demonstrated low sensitivity because of the non-exclusivity of exocrine pancreas to digestion processes of carbohydrates and proteins. Thus, the most suitable breath tests for the study of exocrine pancreatic disorders seem to be those exploring fat malabsorption. This kind of studies reach sufficiently good accuracy even in the study of early or mild disease mainly because lipase is the first enzyme reduced during pancreatic insufficiency; gastric lipase participates only for 15% to intestinal lipolysis; moreover, gastric lipase is not able to adequately compensate the reduced pancreatic lipase activity.

The best investigated triglyceride breath test uses 1,3-distearyl, 2-[13$^\text{C}$-carboxyl] octanoyl glycerol and is known as $^{13}$C-mixed triglyceride breath test administered at the doses of 200-300 mg. Breath samples are collected before the test meal and for the subsequent 5 hours at time intervals of 30 minutes. Labeled triglycerides are metabolized in the duodenum under physiological conditions; they are easily absorbed by mucosal wall and metabolized in the liver cells with production of $^{13}$C labeled carbon dioxide, excreted by lungs. Low $^{13}$C/$^{12}$C ratio in expired air indicate a not completed digestion of marked substrates due to impaired pancreatic secretion.

Sensitivity and specificity of the mixed triglyceride breath test compared to the stimulated lipase output in the pancreozymin test are 89% and 81%, respectively.

It is important that the patient follows certain indications prior to the test: food naturally rich in $^{13}$C (i.e. corn, pineapple, cane sugar) should be avoided at least from 2 days before the test; a 12 hours fasting is recommended, with the exception of natural water. Moreover, smoking, sleeping or doing physical exercise should be avoided from 30 minutes before the test through the entire duration of the test.

Even if breath tests are valid methods for evaluate pancreatic function, the costs of the substrates, the high time expenditure and the lack of standardization limit the utilization and the diffusion in common clinical settings.

Drawing a good and complete report

The medical report is necessary to reduce variability between tests performed in different centers and to give the opportunity to experienced gastroenterologist to fully interpret results, particularly in the case of tests not fully standardized.

It must include all the methodologies and clinical characteristics to let the test interpretable also by non-specialists.

In addition to patient data, it should be specified the substrate administrated, in term of dosage and concentration, the methods of sample collecting (intervals and duration of the test) and machine used for the measurement.

It is good to indicate some clinical notes, like for instance the clinical indications for requiring the test (i.e suspect lactose intolerance, diarrhea, constipation, bloating, etc.).

If necessary, on the basis of the results, gastroenterologist or family doctors could require further exams, to complete the diagnostic process.

In Figure 1 we summarized in a virtual format all informations, which should be clearly written in a good report.

We think that, dealing with $H_2$ breath testing, the key points for the validity of the test would be the distance between the day of the breath sampling and the day of breath assessment. Type of machines utilized, status of the machines, being the “age” of the machine, a limiting step particularly for gas-chromatography with electro-chemical sensor.

Duration of the test, contemporary measurement of $CO_2$ and/or methane is also another important parameter. Even for $^{13}$C breath test, a clear indication of protocol used for the test is crucially important, particularly for non conventional breath tests. Despite $H_2$ breath testing, samples for $^{13}$C breath testing, when kept in proper tubes, are more stable, being $^{13}$C a stable isotope. The indication of the machine utilized could be of interest, as mass spectroscopy has the highest accuracy and, usually, highest precision.

Conflict of interest

The Authors declare that they have no conflict of interests.
References


37) **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; 5: 1214-1217.


