Isotope Selective Nondispersive Infrared Spectrometry Can Compete with Isotope Ratio Mass Spectrometry in Cumulative $^{13}$CO$_2$ Breath Tests: Assessment of Accuracy

Chleboun J., Kocna P.

1 Mathematical Institute, Academy of Sciences, Prague, and Faculty of Mechanical Engineering, Czech Technical University, Prague
2 Institute of Clinical Biochemistry and Laboratory Diagnostics, 1st Faculty of Medicine, Charles University, Prague

1. Introduction

In this paper, $^{13}$CO$_2$ breath tests evaluated by means of isotope selective nondispersive infrared spectrometry (NDIRS) are considered, their accuracy analyzed, and sources of inaccuracy identified and assessed.

The $^{13}$CO$_2$-based breath tests measure a metabolic response to a dose of an administered $^{13}$C-enriched substrate. The response results in $^{13}$CO$_2$ as the end product, which is analyzed in breath samples. The substrate for $^{13}$CO$_2$-based breath tests contains one or more functional groups labeled with $^{13}$C. In the human organism, $^{13}$C-labeled substrate is cleaved in the course of enzymatic reactions such as oxidation, decarboxylation or hydrolysis, and directly or via intermediate metabolites exhaled in the form of $^{13}$CO$_2$. The chain of metabolic reactions can be quite complex in this process. Then, the reaction responsible for the major contribution to the exhaled $^{13}$CO$_2$ can be investigated by means of a $^{13}$CO$_2$ breath test because the intensity of the $^{13}$CO$_2$ exhalation reflects the actual intensity of that metabolic reaction. See, for instance, [2, 7, 11, 12, 13, or 14] for more detailed information on $^{13}$CO$_2$ breath tests.

Generally, two ways of data acquisition are possible in $^{13}$CO$_2$ breath tests. A method can be chosen delivering the absolute abundance of $^{13}$C in a breath sample. A trusted and popular representative of such a method is isotope ratio mass spectrometry (IRMS). However, since IRMS measuring instruments are expensive and are not readily available, another approach is worth thinking about: the use of more available (inexpensive) instruments that, however, are not able to measure the absolute abundance. The NDIRS-based method discussed in this paper represents the latter approach.

Isotope abundance levels measured by NDIRS are reported as differences between a tested sample and a reference sample. The reference $^{13}$C abundance is not known exactly. This uncertainty, uncertainty in $^{13}$CO$_2$ production, and the inaccuracy of the measuring instrument contribute to the uncertainty in the results of breath tests. In this study, the particular impacts of uncertainty are estimated and expressed in a mathematical way by an uncomplicated formula illustrated by an example dealing with real-life data. It is shown that the uncertainty in the reference $^{13}$C abundance does not have severe consequences, so that the method can compete with the spectrometry methods that are able to deliver an absolute value of the $^{13}$C abundance. The inaccuracy of the measuring instrument is also manageable, though its influence is greater than in the previous case. The analysis reveals that the uncertainty in $^{13}$CO$_2$ production deserves great attention because it is difficult to estimate and its influence is rather strong. The problem of determination of a proper cut-off level is outlined.

Key words: isotope selective nondispersive infrared spectrometry, breath tests, uncertainty, accuracy.
we also have to consider the uncertainty in the total amount of CO₂ exhaled by the patient during the breath test. Since the CO₂ flow is not easy to measure, it is estimated through the body surface area (BSA) or basal metabolic rate (BMR) formulae. Again, individual deviations are present and should be taken into consideration. This sort of uncertainty is less severe in ¹³CO₂ breath test where the diagnostic conclusion is drawn not from the PDR, but simply from the increase of the ¹³C content in the patient’s breath.

Finally, the NDIRS measuring instrument itself demonstrates some uncertainty in readings, which exhibits statistical and probabilistic features.

In this paper, we rigorously analyze the sensitivity of ¹³CO₂ breath test results to uncertain input data mentioned in the previous paragraphs. We present a clear formula that shows the contributions of specific uncertain data to the overall uncertainty of results.

2. PDR-Oriented ¹³CO₂ Breath Test

The test is arranged as follows:

1. At time \( t_0 = 0 \), the patient fills bag I (the reference bag) by his/her breath containing a basal level of ¹³C in the form of ¹³CO₂, and receives a dose \( D \) of ¹³C in the form of a ¹³C-enriched substrate.

2. At time \( t_i = i \), \( i = 0,1,…,6 \) (hours), the patient fills bag II, and the ¹³C concentration in bag II is compared to the ¹³C concentration in bag I; the measuring instrument produces a value \( \delta(t) \) (delta over base, DOB).

3. The instantaneous substrate-origin-¹³C content in the patient’s breath at time \( t \), i.e., \( 1/r_{13}(t) \) is the CO₂/¹³CO₂ ratio. Since we assume that CO₂ comprises only ¹²CO₂ and ¹³CO₂, it holds

\[
\frac{1}{r_{13}(t)} = 1 + \frac{1}{R(t)}
\]

Thus,

\[
r_{13}(t) = \frac{R(t)}{1 + R(t)}
\]

It is assumed that the patient’s CO₂ production rate \( P \) is constant during the test. Under this assumption, the ¹³CO₂ production rate \( v_{13}(t) \) is simply

\[
v_{13}(t) = Pr_{13}(t) = P \frac{R(t)}{1 + R(t)}
\]

We are interested in the part of \( v_{13}(t) \) that comes from the ¹³C-enriched substrate; this part will be denoted by \( \alpha(P, R_0; \delta; t) \). Thus, we have to take away the background (basal, dose independent) part. It is supposed that this background part of ¹¹CO₂ is constant, and, consequently, can be determined from the breath stored in the reference bag I. Then (see (2))

\[
\alpha(P, R_0, \delta; t) = v_{13}(t) - v_{13}(0) = P \left( \frac{R(t)}{1 + R(t)} - \frac{R_0}{1 + R_0} \right)
\]

and, after substituting (1) for \( R(t) \),

\[
\alpha(P, R_0, \delta; t) = \frac{PR_0 \delta(t)}{1000 \left( \frac{\delta(t)}{1000} + 1 \right) R_0 + 1} \left( R_0 + 1 \right)
\]

Since (4) can be viewed as a \( P \) multiple of a fraction, it is convenient to introduce \( A(R_0, \delta; t) \) and write

\[
\alpha(P, R_0, \delta; t) = PA(R_0, \delta; t).
\]

The key quantity that we need to evaluate is denoted by \( C(P, R_0, \delta; T) \) and stands for the total substrate-origin-¹³C production from time \( t = 0 \) till time \( t = T \).

\[
C(P, R_0, \delta; T) = \int_0^T \alpha(P, R_0, \delta; t) \, dt = P \int_0^T A(R_0, \delta; t) \, dt.
\]

As \( A(R_0, \delta; t) \) is known only at \( t_i, i = 0,1,…,6 \), the integral (5) is approximated by the trapezoidal rule:

\[
C(P, R_0, \delta; T) = \sum_{i=0}^6 \frac{A(R_0, \delta; t_i) + A(R_0, \delta; t_{i+1})}{2}.
\]

Remark 1: The patient’s PDR is calculated (in %) as 100\( C_{\text{app}}(P, R_0; \delta_i)/D, j = 1,…,6 \). Since the ratio equals (6) divided by the known constant \( D \), we can limit our analysis to (6), i.e., to the CDR.

Remark 2: If the investigation concentrates only on the increase of the content of ¹³C in the patient’s breath (i.e., \( r_{13}(T) - r_{13}(0) \)), then \( \alpha(P, R_0, \delta; t)/P = A(R_0, \delta; t) \) (see (4)) is the target quantity.

Remark 3: The integral (5) can be approximated by Simpson’s rule. Since the trapezoidal rule has been used in calculating the results of clinical breath tests, we confine ourselves to this rule in the uncertainty analysis too.
3. Uncertainty Sources and Their Influence

The exact values of P, R, and δ(t) are not known, i.e., P, R, and δ(t) are uncertain. As a consequence, C_{app} is also uncertain. However, if we put realistic upper and lower bounds on P and R, and if we determine the parameters of the probabilistic behavior of δ, then we will be able to assess the range of possible values C_{app}(P,R,δ;t). Let us investigate the particular sources of uncertainty.

1. Uncertain P

Let μ be a real parameter and P = (1 + μ)P. Then

\[ C_{app}(P, R, δ;t) = (1 + μ)C_{app}(P, R, δ;t) \] (7)

because C_{app} is linear in P.

2. Uncertain R

A drawback of the used isotope selective nondispensive infrared spectroscopy method lies in the unknown value R. As a rule of thumb, the \(^{13}\text{CO}_2/^{12}\text{CO}_2\) ratio of the Pee Dee Belemnite (PDB) standard, is often chosen for R, i.e., \(R = R_{\text{PDB}} = 0.0112372\).

However, measurements indicate that the PDB value overestimates the individual basal \(^{13}\text{CO}_2/^{12}\text{CO}_2\) abundance in the European population. This observation is in conformity with the generally accepted fact that the usual Western European diet is low in \(^{13}\text{C}\). A European group on a habitual and a \(^{13}\text{C}\)-enriched diet showed basal \(^{13}\text{CO}_2/^{12}\text{CO}_2\) values well between 1.080 and 1.084 atom percent in [17]. Similar results are presented in [5]. By taking values from [5], [17] and by considering intervals of confidence characterized by ±3SD (SD = standard deviation), we infer that the interval

\[ I_1 = [0.0109180, 0.0109588] \]

covers more than 99% of the range of the basal \(^{13}\text{CO}_2/^{12}\text{CO}_2\) abundance in the breath of the group on a low \(^{13}\text{C}\)-enriched diet. If a high \(^{13}\text{CO}_2\)-enriched diet is to be taken into consideration, then

\[ I_2 = [0.010918, 0.011053] \]

encompasses the basal \(^{13}\text{CO}_2/^{12}\text{CO}_2\) abundance values as can be calculated from data published in [17]. To allow even larger fluctuations, let us consider

\[ I_{R_{\text{mean}}} = [0.01091, 0.01106], \]

i.e., we have \(R_{\text{mean}} = 0.010985\) perturbed by ± 0.7%.

The following analysis focuses on the sensitivity of C_{app} to a change of R, i.e., to the uncertainty in R. Let C_{app}(P,R,δ;t) be calculated. We are interested in C_{app}(P,R,δ;t), where \(R = (1 + \rho)R_0\) for a small parameter \(\rho\). Although it is easy to calculate C_{app}(P,R,δ;t) – C_{app}(P,R,δ;t) exactly, the following approximation is more convenient for its simplicity:

\[ C_{app}(P,R_0,δ;t) = (1 + \rho)C_{app}(P,R_0,δ;t), \] (8)

where (8) and the estimate of the relative error

\[ \left| \frac{C_{app}(P,R_0,δ;t) - (1 + \rho)C_{app}(P,R_0,δ;t)}{C_{app}(P,R_0,δ;t)} \right| \leq \sum_{i=0}^{1} 2.2\rho R_0 \] (9)

are valid for physiological values, say \(\delta_0 \leq 100\) and \(0 \leq R_0 \leq 0.02\), i.e., (8) and (9) hold even if a value different from \(R_{\text{PDB}}\) or \(R_{\text{mean}}\) is taken as the baseline R. Note that (9) does not depend on P because the fraction is reduced by P.

By referring to the above-mentioned observation concerning the range of individual basal \(^{13}\text{CO}_2/^{12}\text{CO}_2\) ratios, we can expect that \(\rho \in [-0.007,0.007]\) (if \(R_0 = R_{\text{PDB}}\) or \(R_{\text{mean}}\) will cover the majority of cases. Then, the relative error (9) is less than 0.11% at time \(t_6\) (at the end of the breath test) and, consequently, the error of the approximate formula (8) is acceptable in practice.

3. Uncertain δ

Values \(\delta = \delta(t), i = 0, 1, \ldots, 6\), come from measurements therefore their accuracy is limited by the accuracy of the measuring instrument. The determination of the ISOMAX 4000 accuracy is necessary for an insight into the sources of uncertainty.

By testing statistical hypotheses and by applying Monte Carlo simulation, we concluded that, for a fixed pair of bags I and II, the measured values are distributed normally with the standard deviation \(\sigma = 0.3\). Within the physiological \(^{13}\text{C}\) range (i.e., \(-5 \leq \delta(t) \leq 20\) in this study), we consider \(\sigma\) independent of the \(^{13}\text{CO}_2/^{12}\text{CO}_2\) ratio.

In practice, four measurements of \(\delta\) are performed and their mean \(\delta_i\) is used in (4) and (6) to improve the accuracy of C_{app}. Since C_{app} is a nonlinear function of \(\delta\) (see (4) – (6)), statistical properties of C_{app}(P,R,δ;t) were inferred via Monte Carlo simulation. We write \(\delta = (\delta_0, \ldots, \delta_6)\) and use \(\delta\) in C_{app} though it is formally incorrect; it simply means that \(\delta\) is used in the trapezoidal rule (see (6)) to calculate C_{app}(P,R,δ;t).

Let P and R be fixed and let \(\delta_i\) be the true delta over base (DOB) at time \(t_i\). In the simulation, if four instances of DOB are generated with the normal distribution (the mean is set to \(\delta_i\) and the standard deviation to 0.3), \(C_{app}(P,R_0,δ_i,t_i)\) were inferred via Monte Carlo simulation. We write \(\delta = (\delta_0, \ldots, \delta_6)\) and use \(\delta\) in C_{app}, though it is formally incorrect; it simply means that \(\delta\) is used in the trapezoidal rule (see (6)) to calculate C_{app}(P,R,δ;t).

Let P and R be fixed and let \(\delta_i\) be the true delta over base (DOB) at time \(t_i\). In the simulation, if four instances of DOB are generated with the normal distribution (the mean is set to \(\delta_i\) and the standard deviation to 0.3), \(C_{app}(P,R_0,δ_i,t_i)\) were inferred via Monte Carlo simulation. We write \(\delta = (\delta_0, \ldots, \delta_6)\) and use \(\delta\) in C_{app}, though it is formally incorrect; it simply means that \(\delta\) is used in the trapezoidal rule (see (6)) to calculate C_{app}(P,R,δ;t).

Let \(\delta_i\) be the true \(^{13}\text{CO}_2/^{12}\text{CO}_2\) ratios in the respective bags, and let \(\delta_i\) be the corresponding true DOB; see (1). Since inaccessible \(\delta_i\) is approached
through four measurements of $\delta$, we use their mean $\bar{\delta}$ to calculate $C_{app}$. Then

\[
C_{app}(P,R_0,\bar{\delta};t) = 7 \times 10^{-4} PR_0 \leq C_{app}(P,R_0,\delta^*;t) \\
\leq C_{app}(P,R_0,\delta^*;t) + 7 \times 10^{-4} PR_0
\]

with the probability level 0.95 (95%). The constant $7 \times 10^{-4}$ originates from $1.96\sigma_m/R_0 \approx 7 \times 10^{-4}$.

The 95% confidence interval for $\delta$ expressed in terms of $\bar{\delta}$ (the mean of four measurements at time $t$) is

\[
\bar{\delta} - 0.3 \leq \delta \leq \bar{\delta} + 0.3,
\]

where $0.3 = 1.96\sigma_m$, $\sigma_m = 0.151$.

4. Combined uncertainty

If both $P$ and $R_0$ are uncertain, then (7), (8), and (10) can be combined.

Let $\mu$, $\rho \geq 0$ be parameters, let $P,R_0 > 0$ be fixed, and let

\[
(1 - \rho)\bar{R}_0 \leq R_0 \leq (1 + \rho)\bar{R}_0, \quad (1 - \mu)\bar{P} \leq P \leq (1 + \mu)\bar{P}
\]

be the respective bounds for an uncertain $^{13}$CO$_2$/$^{12}$CO$_2$ basal ratio $R_0$ and an uncertain physiological CO$_2$ production $P$. Then (10) is valid for any $P$ and $R_0$ fulfilling (12). After applying (7), (8) and realizing that $C_{app} = PK$, we arrive at

\[
(1 - \mu)(1 - \rho)\bar{P}[K(\bar{R}_0,\bar{\delta};t) - 7 \times 10^{-4}\bar{R}_0] \leq C_{app}(P,R_0,\delta^*;t) \leq (1 + \mu)(1 + \rho)\bar{P}[K(\bar{R}_0,\bar{\delta};t) + 7 \times 10^{-4}\bar{R}_0].
\]

As (13) couples probabilistic and nonprobabilistic estimates, it is not fully justified to say that (13) is the 95% confidence interval for $C_{app}(P,R_0,\delta^*;t)$ under fixed $\mu$, $\rho$. Instead of that, we can say that $C_{app}(P,R_0,\delta^*;t)$ will break the bounds (13) with a low probability, which hardly exceeds 0.05.

Remark 4: It should be $\delta = 0$ because the breath in bag I is, in fact, equal to the breath in bag II at $t = 0$. However, measurements at $t = 0$ suggest that a systematic error $e_0 = -0.4$ could be present, see Table 1. The error was neglected but, in the case of necessity, relevant corrections can be made in the above analysis.

4. Uncertainty Analysis of Real-Life Data

We illustrate (13) by a numerical example originating from a pancreas-oriented $^{13}$CO$_2$ breath test [9, 10]. Table 1 shows the means of four measured DOB values respective to tested individuals $\alpha$ and $\beta$.

Table 1: Means of measured data

\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
 & $t_0$ & $t_1$ & $t_2$ & $t_3$ & $t_4$ & $t_5$ & $t_6$ \\
\hline
$\delta$ & $-0.8175$ & $5.4737$ & $5.9897$ & $4.7690$ & $3.5598$ & $2.5590$ & $1.9628$ \\
$\bar{\delta}$ & $-0.4475$ & $0.2655$ & $1.2540$ & $1.7575$ & $1.5075$ & $1.1163$ & $0.6777$ \\
\hline
\end{tabular}

The BSA-based $\bar{P}$ equals 707 and $\bar{P} = 673$. The Haycock formula [4] is used to calculate the body surface area, then the area is multiplied by a constant (= 300) to obtain the CO$_2$ production rate. Let $\mu = 0.2$, $\rho = 0.007$, see (12), and $\bar{R}_0 = R_{mean}$. Then

\[
K(\bar{R}_0,\bar{\delta};t) = 2.46 \times 10^{-4} \quad \text{and} \quad K(\bar{R}_0,\bar{\delta};t) = 6.45 \times 10^{-4}.
\]

and (13) reads

\[
0.1338 \leq C_{app}(P,R_0,\bar{\delta};t) \leq 0.2167, \quad (14)
\]

\[
0.0304 \leq C_{app}(P,R_0,\bar{\delta};t) \leq 0.0587. \quad (15)
\]

By comparing (14) and (15), we observe that the ranges of $C_{app}$ do not intersect even if a rather considerable amount of uncertainty is taken into account. This indicates that these two individuals (almost) certainly do not belong to one and the same metabolic response class.

We can also express (14)–(15) in terms of the CDR, that is,

\[
40.3 \% \leq \text{CDR}_\alpha \leq 65.2 \%, \quad (16)
\]

\[
9.2 \% \leq \text{CDR}_\beta \leq 17.7 \%; \quad (17)
\]

see Remark 1; it is $D = 0.3224$.

According to our experience supported by clinical data and information available in literature, the BSA-based cut-off level is equal to 23% for this test, where $^{13}$C-mixed triglyceride is administered. As a consequence, (17) clearly indicates a disorder.

5. Carbon Dioxide Production

Although the uncertainty in $P$ is easily tractable, see (7), this sort of uncertainty certainly has a large share in the uncertainty of breath tests results. It seems reasonable to use the BMR-based estimates of $P$ instead of the BSA-based estimates. The BMR-based formulae take into consideration not only body weight and height (as the BSA formulae do), but also age and sex; see [8], [16].

In general, $P$ calculated from BMR ($P_{BMR}$) differs from $P$ calculated from BSA ($P_{BSA}$), but the difference is not uniform. Figure 1 depicts the shifts in our test results if $P_{BMR}$ is used instead of $P_{BSA}$ inferred by multiplying the body surface area, see [4], by 300 mmol/hour. To obtain $P_{BMR}$, the formula published in [8] was modified. We set the patient energy expenditure equal to 1.1 BMR, which, in our opinion, corresponds better to the conditions of our tests than 1.4 BMR used in [8]. The PDR values are calculated from functional breath tests measurements after $^{13}$C-mixed triglyceride or $^{13}$C-xylose administration in a group of 181 subjects.

Although there are differences between individual $P_{BMR}$ and $P_{BSA}$ values, the mean of $P_{BMR}/P_{BSA}$ is close to 1. This indicates that the widely accepted cut-off value 23% inferred on the basis of the BSA-based estimate...
of CO₂ production is valid even if the BMR-based estimate is used. This cut-off level has also been confirmed by the breath tests of about 40 individuals without a pancreas disorder. In detail, the mean of their PDR reduced by 2 SD equals 24%.

However, it must be stressed that the estimation of the individual CO₂ production rate is a delicate matter and that the assessment of its accuracy asks for further investigation.

6. Discussion

The uncertainty in the results of ¹³CO₂ breath tests has three sources: (a) uncertain baseline ¹³C abundance R₀, (b) inaccuracy of the spectrometer, and (c) uncertainty in CO₂ production P.

Regarding (a), it is fairly justified to consider the uncertainty in R₀ very limited. Its impact is almost negligible if compared with (b) and (c). Although the basal ¹³C abundance is estimated on the basis of general facts, the bounds are rather tight. It means that the accuracy gain in a cumulative breath test using IRMS, where the reference basal ¹³C is known, would not be significant.

The accuracy of the spectrometer includes two aspects: (b1) the accuracy of repetitive measurements, i.e., a deviation when one gas sample is measured repeatedly, and (b2) the absolute accuracy, i.e., the relation to a true δ value.

In [1], subjects (b1) and (b2) are treated and an isotope selective nondispersive infrared spectrometer is tested on gases with different δ values measured also by isotope ratio mass spectrometry (IRMS). The measured samples have the true δ values (approximately) 118, 98, 75, -20, -24, and -45. These sample δ values are irregularly distributed and do not cover the range of our prime interest in this study (−1 up to 12 per mille). Nevertheless, the standard deviation value is significantly less than 3% of the true δ value in [1], except for the 75 per mille sample, where it is more than 5%.

The standard deviation of NDIRS measurements reported in [1] is in conformity with our findings. Let us recall that we consider the standard deviation of measurements equal to 0.3, when samples with δ between −1 and 12 are measured; see Section 3.

The isotope selective nondispersive infrared spectrometer used in our tests (ISOMAX 4000) has been regularly tested on samples supplied by the manufacturer. Although the accuracy tolerance has not been exceeded, this sort of “black-box” testing has not contributed much to our knowledge of the absolute accuracy of the device (the inaccuracy is less than 10%). However, according to [1, Table 1], the difference between the IRMS and NDIRS readings of δ is about 1% or less, except for the 75 per mille sample, where it is 6%. The 1% figure indicates that a properly maintained and calibrated infrared spectrometer is a satisfactorily accurate device. Similar or even better agreement between IRMS and NDIRS readings is reported in [6].

7. Conclusion

By virtue of the sensitivity and uncertainty analysis, we conclude that the uncertainty in the reference ¹³C abundance is not particularly harmful. It does not prevent the NDIRS method from being successfully applied in the ¹³CO₂ breath tests (cf. [15], where different reasons lead to the same conclusion). However, the analysis also reveals that a more serious danger for the credibility of ¹³CO₂ breath tests could be the uncertainty in the amount of CO₂ exhaled by the patient. We consider the uncertainty in CO₂ production the most important source of uncertainty in breath tests results. The IRMS-based cumulative breath tests also suffer from this sort of uncertainty, so that to substitute IRMS for NDIRS is not a remedy.

Acknowledgements: The research pursued by Jan Chleboun was supported by the Grant Agency of the Czech Republic under grant number 201/02/1058.

The research was also supported by the Academy of Sciences of the Czech Republic, Institutional Research Plan No. AV0Z10190503.

References

2. Ghoos, Y. ¹³CO₂-Breath Tests at the Laboratory Digestion Absorption, University of Leuven, May 1996.


