

Bayesian Hierarchical Methods to Interpret the ^{13}C -Octanoic Acid Breath Test for Gastric Emptying

Leslie J.C. Bluck^a Sarah J. Jackson^a Georgios Vlasakakis^a Adrian Mander^b

^aMRC Human Nutrition Research, and ^bCambridge Hub in Trials Methodology Research, MRC Biostatistics Unit, Institute of Public Health, University Forvie Site, Cambridge, UK

Key Words

Gastric emptying · Bayesian modelling · Breath tests

Abstract

Background: The ^{13}C -octanoic acid breath test is a convenient method for assessing gastric emptying (GE). Success depends on obtaining a well-characterized time profile of the excretion of label in breath, which may not be the case if GE is delayed. **Aims:** To use Bayesian techniques in conjunction with hierarchical modelling as a method to increase the success of the modelling process. **Methods:** Retrospective analysis of 164 individual breath tests using the WinBUGS program. The approach was tested by analysing the complete dataset simultaneously, and also as individual studies. **Results:** The time required for Bayesian modelling was comparable with that needed for the usual methods. The results obtained were almost identical to those obtained from conventional modelling for well-behaved breath tests, but much more realistic in cases where the experimental data was poor, or when GE was delayed. **Conclusions:** The use of Bayesian estimation of the parameters of the ^{13}C -octanoic acid breath test is demonstrated. By adopting a hierarchical model, realistic values for the lag phase and half-emptying time were obtained in situations when conventional param-

eter estimation failed. This is particularly relevant when GE is unexpectedly delayed. We recommend that WinBUGS become the method of choice for analysing breath test data.

Copyright © 2010 S. Karger AG, Basel

Introduction

The ^{13}C -octanoic acid breath test ($^{13}\text{COBT}$) for gastric emptying (GE) measurement [1, 2] is a useful alternative to γ -scintigraphy since it confers no radiation burden, and is therefore suitable for application in vulnerable subjects (e.g. children and pregnant women). It is particularly useful for intervention investigations in healthy volunteers. The underlying principle of the breath test is that after ingestion, short-chain fatty acids, of which octanoic acid is a convenient representative, pass unchanged through the stomach. However, on delivery to the small intestine, they are rapidly absorbed and transported to the liver where they undergo immediate oxidation [1]. On this basis the rate limiting step in their metabolism is the residence time in the stomach, and therefore the rate of metabolism is a good proxy for the rate of GE.

Unfortunately the $^{13}\text{CO}_2$ that is the detectable metabolic product of the labelled substrate is not excreted di-

rectly, but instead participates in the TCA cycle, and passes through the bicarbonate system before exhalation [3–5]. The consequence of the former is that recovery of the label in breath is incomplete, and that of the latter is that the profile of the appearance of the label in breath is a delayed, blunted and reduced representation of the GE process. It is this indirect description of GE by the breath test which has attracted critical debate of the method [2], although recently a simple method of correcting for the effects of the delay in the bicarbonate pool has been proposed [3, 6].

There is, however, a second important consequence of the delay caused by the passage through the bicarbonate pool: although GE is complete in typically 2–3 h, it is necessary to collect breath for at least twice this period in order to properly establish the kinetics of substrate oxidation for the correct inferences to be drawn. In particular the parameter $t_{1/2}$, defined as the time at which half the label has been processed, requires that the maximum amount of label recoverable in breath be well characterised. Since the fraction of the dose ever to be recovered is uncertain (due to an unknown fraction being sequestered in the TCA cycle), this maximum attainable recovery has to be taken as the asymptote of the plot of cumulative recovery versus time. The asymptote is well defined only if the breath measurements are made for a sufficiently long time. In cases of delayed or slow GE, this may well be in excess of the usual 6-hour collection time. Unless this is known before the experiment is performed so that a prolonged test can be made, the results of breath tests in these circumstances are often compromised.

A Bayesian approach to the analysis can help to overcome many of these difficulties. The Bayesian method differs from the usual nonlinear least squares (NLS) approach to the analysis of breath test output by allowing the incorporation of a priori knowledge ('prior beliefs') about the parameters of the curves used for fitting to the experimental data. There is a considerable amount of prior knowledge concerning the breath test output, e.g. the total recovery of the label cannot exceed 100%, and the limiting fractional rate constant of elimination cannot be negative. In the Bayesian approach, these restrictions are incorporated into the model by the presumptions of particular distributions for the kinetic parameters that restrict the values they are allowed to adopt.

In addition to adopting Bayesian methods to analyze the problem, we can also surmise that there will be a population distribution for the coefficients describing the actual GE parameters. This can be incorporated into our modelling strategy by drawing GE parameters for each

subject studied from a distribution of those parameters in the study population as a whole (hierarchical analysis). It should be recognized that hierarchical techniques are not solely applicable to Bayesian analysis, nor are they necessary for it. The breath tests could be analyzed on a single basis using the same prior descriptions of the fitted parameters. However, there are advantages in letting the individual parameter values be drawn from a common distribution, since by imposing some degree of consistency the individual parameter estimates can 'borrow information' from each other.

The purpose of the work reported here is to introduce the concepts of hierarchical Bayesian analysis in the study of GE, and in doing so to re-analyse a number of previously reported studies so that the advantages of the Bayesian approach can be demonstrated. In particular it will be demonstrated that adopting Bayesian philosophy and combining this with hierarchical methods can eliminate some of the modelling failures previously reported, and allow the full information to be extracted from the $^{13}\text{COBT}$ data.

Methods

A number of studies previously conducted at our laboratories have been re-analysed for this work. Since the breath test procedures were generally similar in each case, a generic protocol for GE measurement using the $^{13}\text{COBT}$ will be given, with any deviations from it noted in the specific study descriptions which follow.

Generic Protocol for the Breath Test

The studies were performed in the volunteer suite at MRC Human Nutrition Research (MRC-HNR), Cambridge, UK, and were approved by the local research ethics committee. All the reported investigations were performed on healthy adults with no history of gastrointestinal disorder. Informed written consent to the tests was obtained from each subject following written explanations of the protocol and purpose of the study.

Study Day Protocol

Subjects were asked to abstain from alcohol and strenuous physical activity, and to fast from 8.00 p.m. on the day before being studied. To prevent dehydration, a single glass of water was allowed on waking on the morning of the study. On arrival at MRC-HNR, height (without shoes) was measured to the nearest 1 cm and weight determined to the nearest 0.01 kg for the estimation of the basal metabolic rate [4]. The $^{13}\text{COBT}$ used an egg-based meal, with a standard energy content of 2 MJ. The egg was separated and a standard dose of 100 μl of ^{13}C -octanoic acid (621 μmol) was added to the yolk. The egg was then dry-fried in a non-stick pan, and served with 3 slices of toasted bread spread with a total of 10 g butter, 100 ml of orange juice and 100 ml of water. Subjects were asked to consume the meal in less than 10 min.

Basal samples of breath were collected prior to the meal being consumed and continued for 6 h. For the first 4 h, samples were taken every 15 min, and every half hour for the final 2-hour period. All breath samples were collected in duplicate. Isotope analysis was performed using isotope ratio mass spectrometry, with a reference gas traceable to an international standard V(PDB).

The calculations performed were as follows:

$$\text{CO}_2 \text{ production rate (mol/h), } F_{\text{CO}_2} = 0.04518W^{0.5378}H^{0.3964}$$

based on the formula given by Evenepoel et al. [5], where W is the subjects weight (kg) and H the height (m).

The preferred equation taken to describe the data from a single breath test is [3]:

$$\delta(t) = \delta_b + \frac{1,000 \cdot d}{(PDB) \cdot F_{\text{CO}_2}} F_{\infty} k \beta (1 - \exp\{-kt\})^{\beta-1} \cdot \exp\{-kt\} \quad (1)$$

This can be generalized to explicitly indicate that, in this case, a number of such models were solved simultaneously by introducing an index n which specifies the particular breath test under consideration, thus:

$$\delta(n, t) = \delta_b(n) + \frac{1,000 \cdot d}{(PDB) \cdot F_{\text{CO}_2}(n)} F_{\infty}(n) k(n) \beta(n) (1 - \exp\{-k(n)t\})^{\beta(n)-1} \cdot \exp\{-k(n)t\} \quad (2)$$

Here $F_{\infty}(n)$ (the fraction of the dose recoverable in breath), $k(n)$ and $\beta(n)$ are parameters of the fit for the n -th breath test; d is the isotope dose given (which was constant for all the tests in this study); (PDB) the isotopic composition of the international standard, and $\delta(n, t)$ and $\delta_b(n)$ (‰) the measured isotopic compositions with respect to PDB for the n -th breath test at time t and in the basal state, respectively.

From the coefficients of the fit, four parameters descriptive of the GE process were calculated [7] as:

$$\begin{aligned} T_{\text{lag}}(n) &= \frac{\text{Ln}\{\beta(n)\}}{k(n)} \\ T_{1/2}(n) &= \frac{-\text{Ln}\{1 - 2^{-1/\beta(n)}\}}{k(n)} \\ T_{\text{lat}}(n) &= \frac{\text{Ln}\{\beta(n)\} + 1/\beta(n) - 1}{k(n)} \\ T_{\text{asc}}(n) &= \frac{\text{Ln}\{1 - 2^{-1/\beta(n)}\} + \text{Ln}\{\beta(n)\} + 1/\beta(n) - 1}{k(n)} \end{aligned} \quad (3)$$

$T_{1/2}$ is the time when half of all the label which will be excreted in breath has been recovered, and is analogous to the GE half time. T_{lag} is the time of maximum rate of excretion of the label, analogous to one definition of the lag phase from γ -scintigraphy, which is the time of maximum GE rate [8]. T_{lag} and T_{asc} have been described by Schommartz et al. [7] as parameters which are more discriminating in the determination of GE perturbations. T_{lat} is similar to an alternative description of the scintigraphic lag phase [9], where the breath response is approximated by a piece-wise linear model, the first phase of which is taken to represent the lag. Finally T_{asc} is suggested as a complementary parameter to T_{lat} to fully characterise the first half of the breath test.

Besides the parameters describing the observed breath curves, we also calculated the 'self-corrected' breath test profiles [3, 6]:

$$G(n, t) = \frac{\beta(n)(1 - \exp\{-k(n)t\})\beta^{(n)-1}}{\beta(n) - 1(1 - \exp\{-k(n)t\})\beta^{(n)}}$$

From this, the associated half-time, $T_{1/2(in)}(n)$, can be obtained by interpolation, and the lag time, $T_{\text{lag}(n)}(n) = \text{Ln}\{\beta(n)/2\}/k(n)$, and maximal emptying rate

$$\dot{G}_{\text{max}}(n) = \frac{4k(n)(\beta(n) - 1)}{\beta(n)} \left(1 - \frac{2}{\beta(n)}\right)^{\beta(n)-2} \quad (4)$$

are calculated.

Frequent Feeding Study

This work, designed to investigate the effect of 'snacking' on the handling of a subsequent test meal has been fully reported previously [10]. In an initial feasibility study conducted with the purpose of determining the reproducibility of the $^{13}\text{COBT}$, four subjects were studied four times, and then in the main phase of the work sixteen subjects were studied both after periods of frequent feeding (six small meals consumed hourly) or meal eating (two large meals with a 3-hour interval). In both the feasibility and the main studies the generic GE procedure was followed, except that a prolonged measurement period of 8 h rather than 6 h was used.

Obesity Study

As previously reported [11], the $^{13}\text{COBT}$ was used in two matched groups of sixteen women, one group lean ($\text{BMI} < 25$) and the other obese ($\text{BMI} > 30$). The generic protocol was used without modification in this investigation.

Meal Size Study

In the meal size study, the effect of changing the energy content of the test meal (1, 2 or 3 MJ) was investigated [12]. The only change to the generic protocol was the manipulation of meal size.

Meal Composition Study

In a parallel investigation to the meal size study, the total energy of the test meal was kept constant at 2 MJ, but the macronutrient composition was tailored so that the meals provided were either high in fat (15% protein, 60% fat and 25% carbohydrate), carbohydrate (15% protein, 25% fat and 60% carbohydrate) or protein (30% protein, 33% fat and 37% carbohydrate).

Bayesian Hierarchical Analysis of Breath Tests

Bayesian modelling was performed using the WinBUGS program [13]. A schematic overview of the hierarchical model used is shown in figure 1. The analysis requires that the form of the underlying distributions of the parameters be specified. This is fundamental to the Bayesian process since these distributions are the prior knowledge (often just termed 'priors') which is taken to the analysis. Since the Bayesian method uses the current observations to modify the priors resulting in posterior distributions, the choice of prior is very important. If a parameter's prior is specified with a high degree of precision (an informative prior), then it is possible that the new data will add little to our knowledge of the true value. Conversely, if the precision of the distribution chosen for the prior is low (vague prior) then the new data dominates the

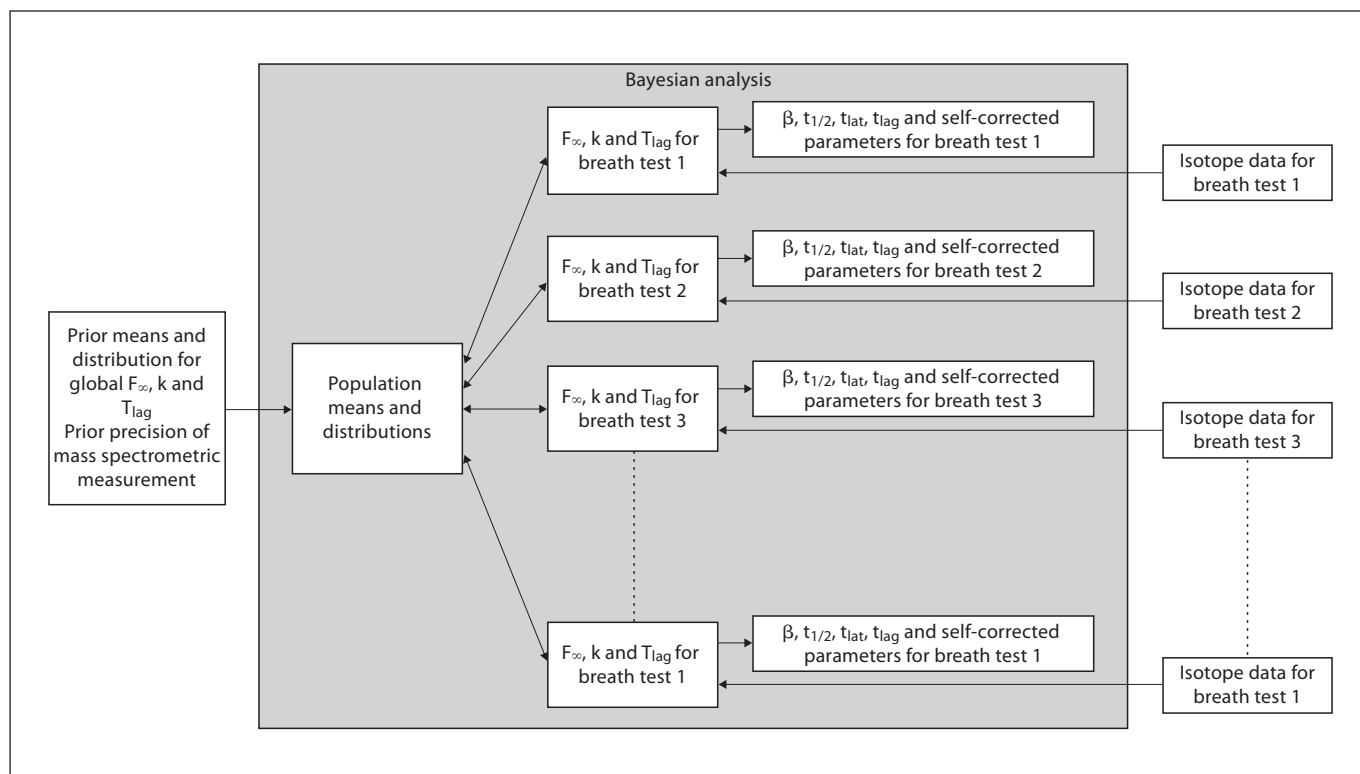


Fig. 1. Bayesian hierarchical analysis of the breath tests. Estimates of the individual parameters F_{∞} , k , and T_{lag} were made in parallel with the population characteristics (mean and distribution). The advantage of Bayesian estimation is the possibility of ‘information flow’ between individuals since the individual estimates rely on the observed mass spectral data, and also on the population distribution, which is in turn derived from the other individual estimates. This is sometimes referred to as ‘borrowing of strength’.

estimate of the posterior distribution. In this case we chose to parameterise the model in terms of three parameters F_{∞} , k and T_{lag} with the following vague prior assumptions regarding their global distributions.

Prior F_{∞} was assumed to be normally distributed with a mean of 60% and a standard deviation of 14%. This prior is relatively informative and based upon the values obtained in the many breath tests that have been reported, as well as from other estimations of the amount of ingested labelled substrate lost in the TCA cycle [14].

Prior k was assumed to be uniformly distributed in the range 10^{-4} to 10^3 h^{-1} .

Prior T_{lag} was assumed to be uniformly distributed in the range 10^{-4} to 10^3 h. Both these priors are very vague, since they allow the parameter to adopt values over seven orders of magnitude.

In all, a total of 164 sets of breath test data (16 from the frequent feeding feasibility study, 32 from the frequent feeding study, 32 from the study on obesity, 36 from the meal size study and 48 from the meal composition study) were fitted simultaneously using the hierarchical method in just under 34 min on a standard desktop workstation. For comparison of results from individual studies with their NLS counterparts, subsets of the full dataset were used.

Results from the Full Dataset

The global distribution defining the parameters F_{∞}^{BAY} , k^{BAY} and T_{lag}^{BAY} were all found to be near-normally distributed (mean and median coincident, with symmetrical confidence limits). F_{∞}^{BAY} was found to be $42.9 \pm 0.8\%$ (mean \pm standard deviation), k^{BAY} was found to be 0.511 ± 0.012 h^{-1} and T_{lag}^{BAY} was found to be 3.30 ± 0.07 h. The ranges of the individual GE time parameters were 1.66 h $< T_{lag}^{BAY} < 7.25$ h, 2.23 h $< T_{1/2}^{BAY} < 8.07$ h, 0.74 h $< T_{lat}^{BAY} < 5.20$ h, and 1.45 h $< T_{asc}^{BAY} < 4.74$ h. In general, the standard deviation of the individual value of the parameter was found to increase with its absolute value, with the average coefficient of variation (c.v.) being 2.2, 2.7, 4.3 and 6.5% for T_{lag}^{BAY} , $T_{1/2}^{BAY}$, T_{lat}^{BAY} , and T_{asc}^{BAY} , respectively.

Comparison of the estimates of (T_{lag}), and ($T_{1/2}$) from NLS methods and Bayesian analysis shows the effect of the ‘borrowing of strength’ in the hierarchical analysis (fig. 2).

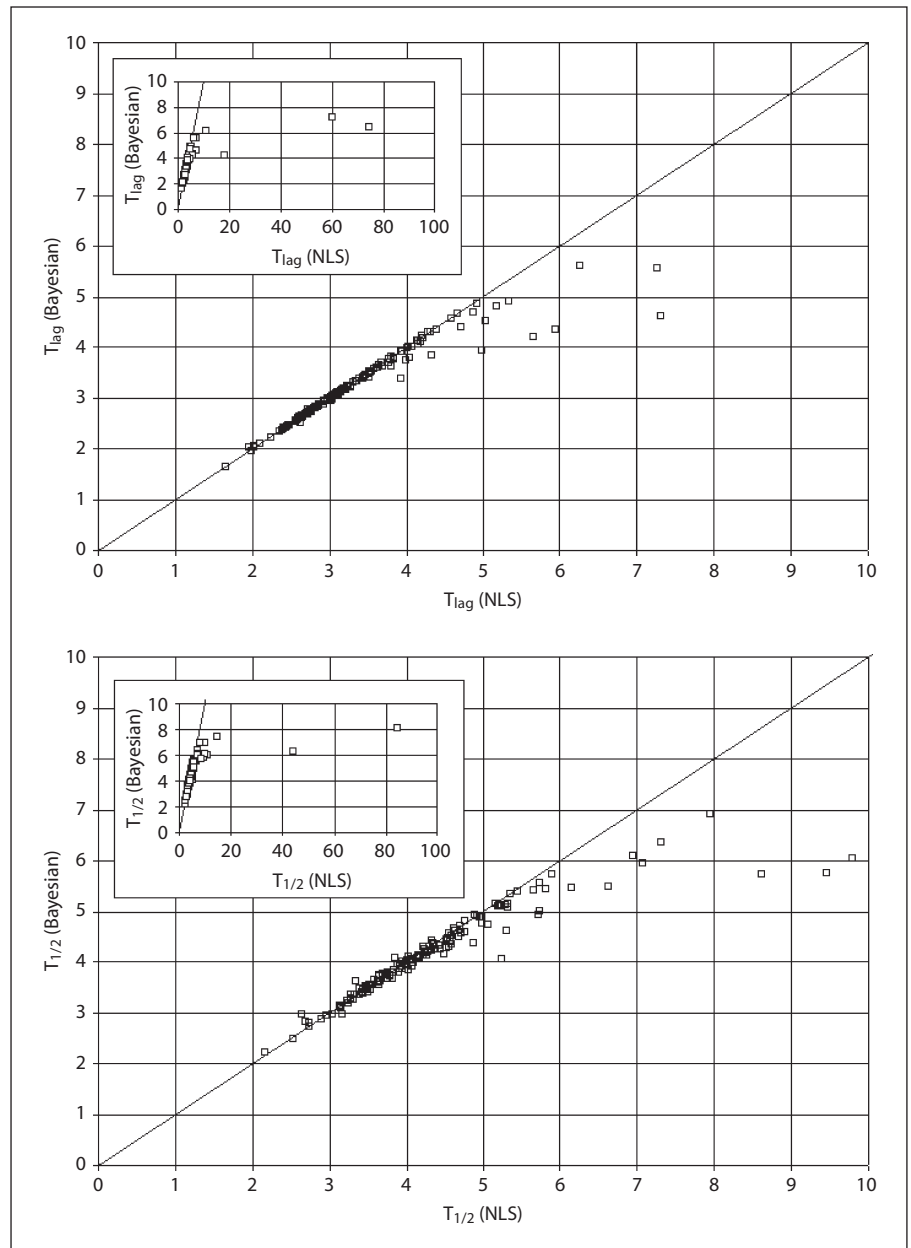


Fig. 2. Comparison of results from an hierarchical Bayesian analysis with individual least squares estimations. The unlikely high least squares estimates of T_{lag} and $T_{1/2}$ (most obvious in the inset graphs) are replaced by more realistic values by the Bayesian method.

A subset of the data (84 instances) had an independent assessment of GE made simultaneously by observing the incorporation of label from ^2H -octanoic acid into body water [15]. This methodology gives two parameters, t_1 and t_2 describing the gastric input function. Direct comparison can be made between the ‘self-corrected’ parameter $T_{lag(in)}$ and t_1 , and also between \hat{G}_{max} and $2/t_2$. The self-consistency of the WinBUGS analysis of the breath test, and the models proposed for post-absorptive processing is demonstrated by the high degrees of correlation ($r^2 = 0.80$ for the first comparison, and $r^2 = 0.65$ for the second).

Results from Individual Studies

In order to investigate the general utility of the Bayesian approach in datasets of more limited size, the WinBUGS analysis was repeated for each of the five studies individually. The results obtained were compared with the NLS estimates to investigate if the conclusions drawn are dependent on the modelling method.

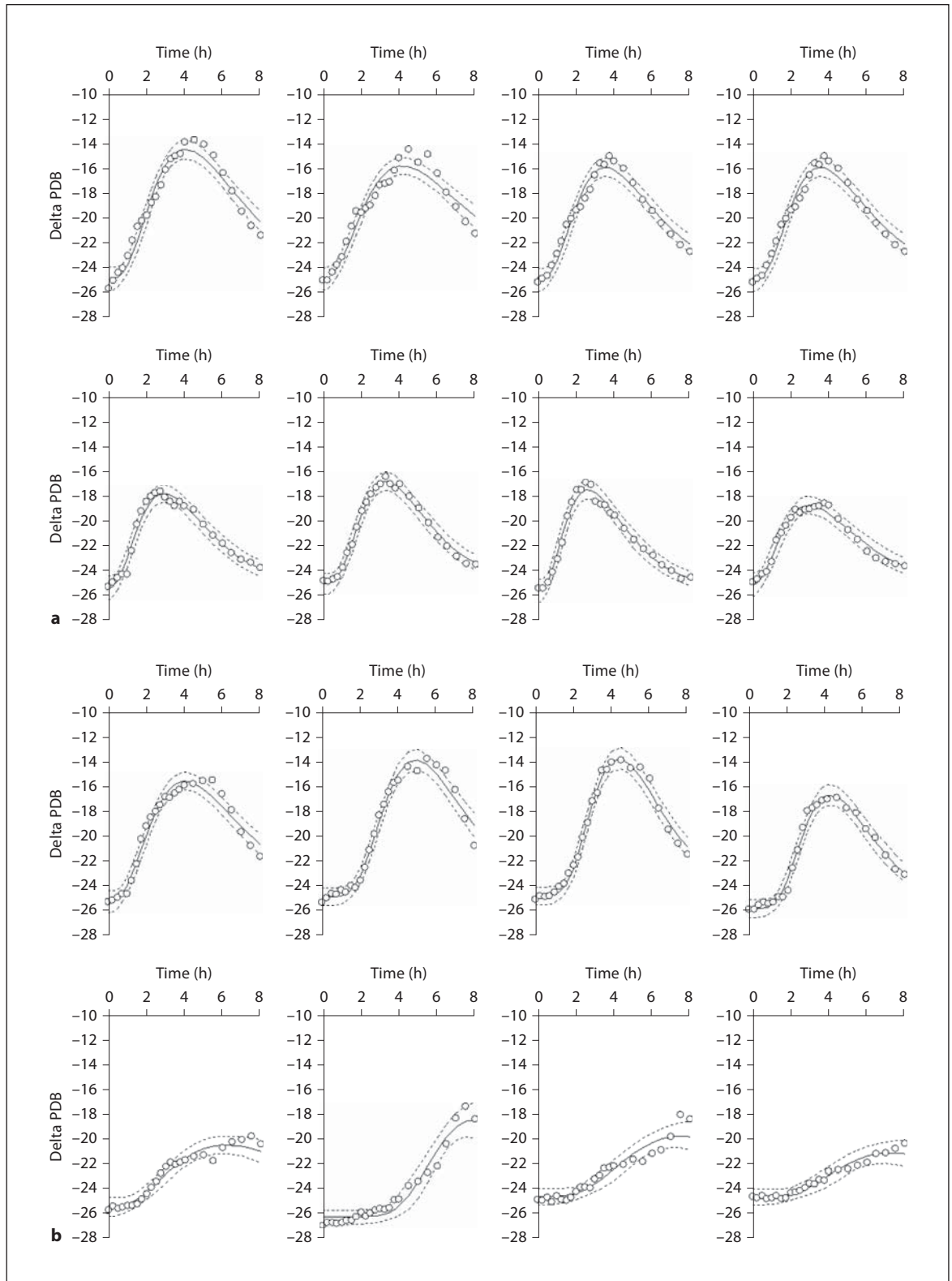


Fig. 3. Raw breath test results and the Bayesian fit (solid line median, dotted lines 95% CI) for the frequent feeding feasibility study. Note the graphs in the bottom line, which show data for a subject with slow GE, which were unable to be fitted satisfactorily using NLS methods.

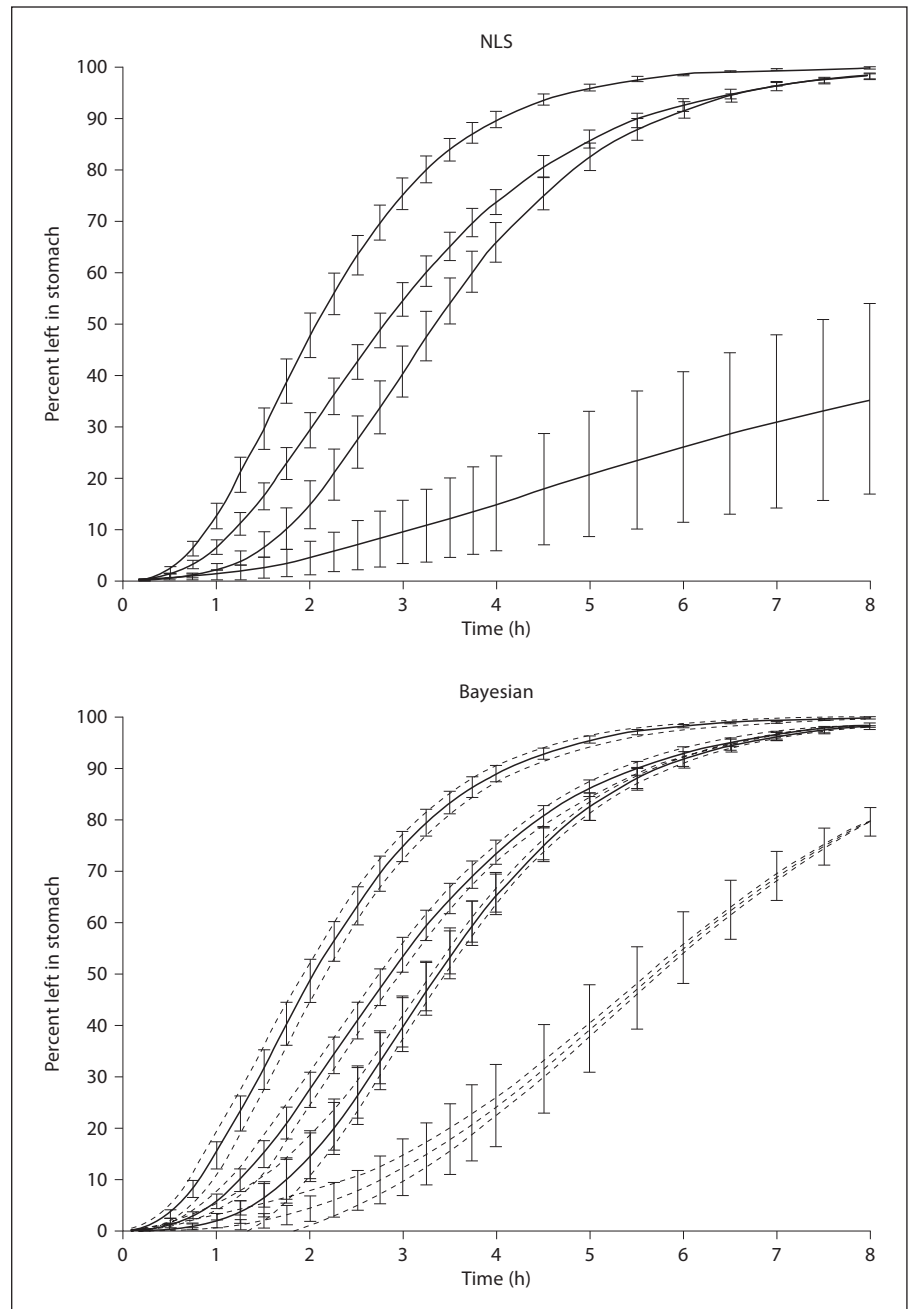


Fig. 4. ‘Self-corrected’ GE curves for the four subjects in the feasibility study, derived from NLS and Bayesian fitting. The error bars represent the standard error obtained from four breath tests in the same subject. The dotted lines in the lower graph indicate the mean uncertainty (standard deviation) at each time point derived from the Bayesian analysis.

Table 1. Population parameters obtained from hierarchical Bayesian analysis for the frequent feeding feasibility study

Population characteristic	Mean	Median (95% interval)
$F_{\bullet\bullet}^{BAY}$	0.489	0.488 (0.398–0.585)
k^{BAY}, h^{-1}	0.417	0.412 (0.323–0.500)
T_{lag}^{BAY}, h	4.66	4.64 (3.71–5.61)

Frequent Feeding (Snacking) Study

Hierarchical analysis of the feasibility study data gave estimates of $F_{\bullet\bullet}^{BAY}$, k^{BAY} and T_{lag}^{BAY} with good precision (c.v. of 7.9–28, 7.6–27 and 1.9–11.3%, respectively). The fits to the experimental data are shown in figure 3, and population characteristics are indicated in table 1. Using the conventional NLS method, it was not possible to get a feasible model for one subject [10] due to excessively slow GE. With the Bayesian approach, realistic interpretations

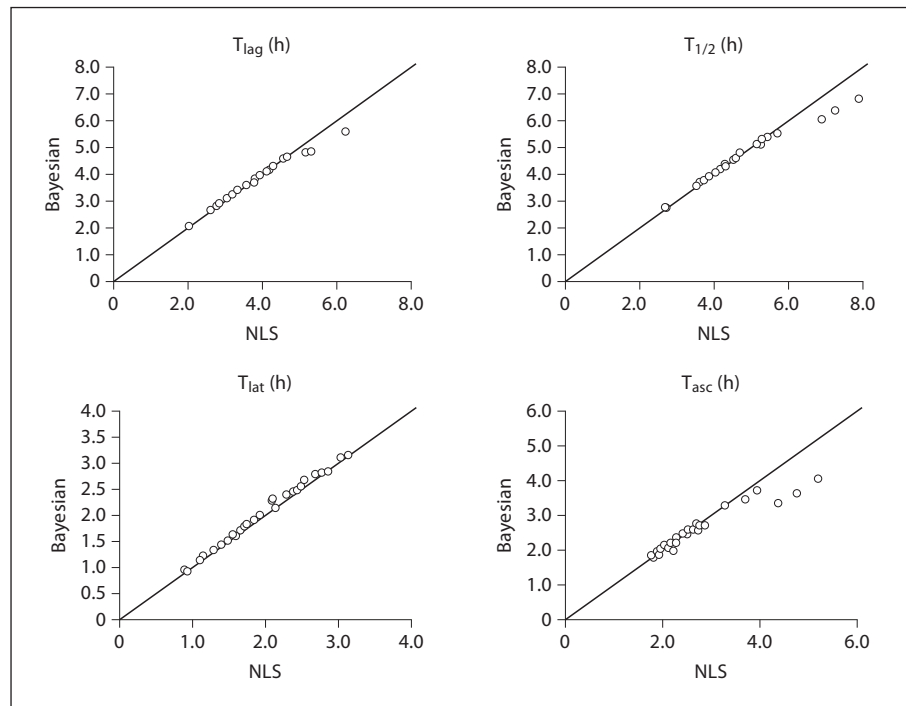


Fig. 5. Scatter plots of the GE parameters obtained from the frequent feeding study from the NLS and Bayesian analyses.

Table 2. Comparison of the breath test parameters derived in the frequent feeding feasibility study

Subject	F_{∞}		k, h^{-1}		T_{lag}, h	
	NLS	Bayesian	NLS	Bayesian	NLS	Bayesian
1	54 ± 8%	51 ± 6%	0.377 ± 0.054	0.396 ± 0.054	3.99 ± 0.31	4.00 ± 0.30
2	37 ± 3%	40 ± 3%	0.532 ± 0.056	0.500 ± 0.047	2.91 ± 0.28	2.93 ± 0.28
3	54 ± 10%	52 ± 9%	0.452 ± 0.077	0.461 ± 0.073	4.38 ± 0.38	4.39 ± 0.37
4	N/A	53 ± 8%	N/A	0.291 ± 0.061	N/A	7.28 ± 0.75

Note that realistic estimates are obtained in all cases using Bayesian methods, whereas this is not the case for NLS.

of the breath test could be made in all cases. This is due to the hierarchical nature of the approach (i.e. borrowing of strength between tests). For the subjects where satisfactory analysis was obtained in both cases, their results were comparable (table 2). Particularly striking is the improvement the Bayesian method gives for the ‘self-corrected’ GE curves (fig. 4).

Turning now to the data obtained comparing frequent feeding with meal eating, there was a strong correlation obtained between the GE parameters deduced from the NLS and Bayesian methods (fig. 5). The correlation coefficients found were 0.992 (T_{lag}), 0.985 ($T_{1/2}$), 0.997 (T_{lat}) and 0.961 (T_{asc}).

In general, the Bayesian method was better at discriminating between frequent feeding and meal eating behaviour. Whilst NLS indicated significant differences in T_{lag} ($p = 0.028$) and T_{lat} ($p = 0.036$) only, WinBUGS indicated greater significance in both parameters T_{lag} ($p = 0.015$) and T_{lat} ($p = 0.028$), but also showed that $T_{1/2}$ was also greater after frequent feeding ($p = 0.043$). The mean GE curves deduced from the Bayesian ‘self-corrected’ profiles are illustrated in figure 6. The Bayesian analysis was also more discriminatory in the self-corrected GE parameters, the NLS estimate of the increase in $t_{1/2(in)}$ on the frequent feeding pattern being 0.32 h ($p = 0.014$) compared with 0.30 h ($p = 0.010$) from the

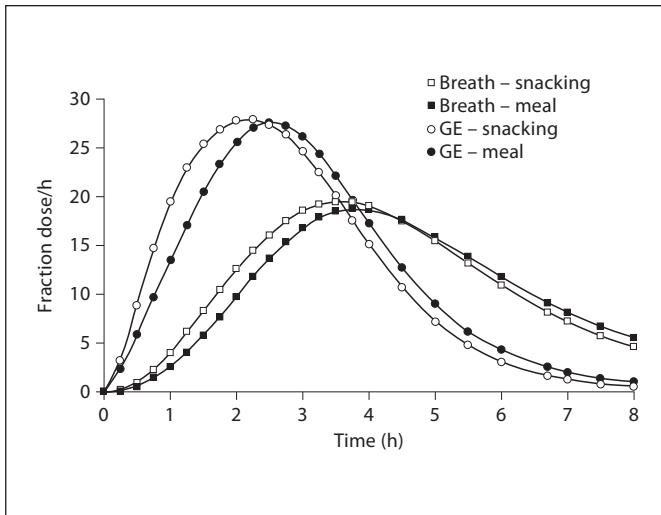


Fig. 6. The mean breath test data and derived rates of GE obtained by 'self-correction' of the Bayesian fits to the breath test. The y-axis represents the amount of the meal which has passed out of the stomach.

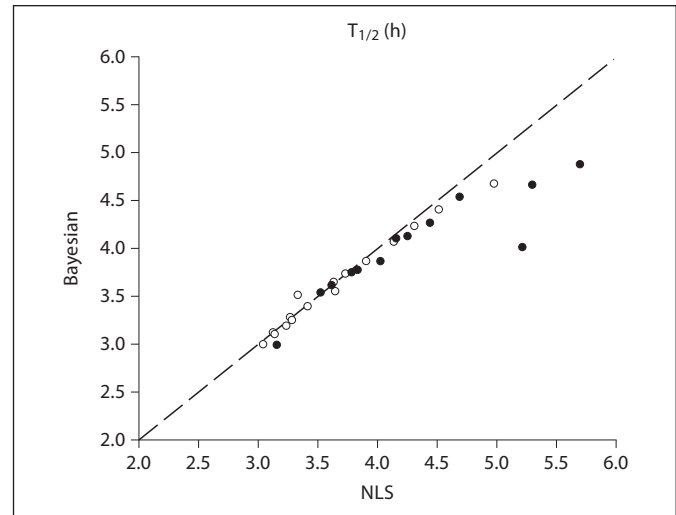


Fig. 7. Comparison of $T_{1/2}$ for the obese (solid circles) and lean (open circles) subjects.

Table 3. $^{13}\text{COBT}$ parameters obtained by NLS methods and Bayesian hierarchical techniques in a study comparing lean and obese subjects

Parameter	Method	Lean	Obese	p
F_{∞}	NLS	0.425 ± 0.023	0.535 ± 0.028	0.006
	Bayesian	0.416 ± 0.020	0.472 ± 0.018	0.04
k	NLS	0.576 ± 0.028	0.529 ± 0.035	N/S
	Bayesian	0.606 ± 0.027	0.608 ± 0.031	N/S
β	NLS	5.24 ± 0.31	5.95 ± 0.54	N/S
	Bayesian	5.79 ± 0.33	7.54 ± 0.61	0.016
T_{lag}	NLS	2.88 ± 0.11	3.34 ± 0.11	0.005
	Bayesian	2.88 ± 0.10	3.26 ± 0.09	0.008
$T_{1/2}$	NLS	3.67 ± 0.14	4.23 ± 0.18	0.019
	Bayesian	3.63 ± 0.13	3.99 ± 0.12	0.04
T_{lat}	NLS	1.454 ± 0.062	1.716 ± 0.058	0.004
	Bayesian	1.497 ± 0.065	1.812 ± 0.076	0.001
T_{asc}	NLS	2.22 ± 0.11	2.51 ± 0.17	N/S
	Bayesian	2.13 ± 0.09	2.18 ± 0.10	N/S
$t_{lag(in)}$	NLS	1.632 ± 0.077	1.920 ± 0.076	0.013
	Bayesian	1.724 ± 0.078	2.101 ± 0.070	0.001
$t_{1/2(in)}$	NLS	2.076 ± 0.081	2.422 ± 0.068	0.003
	Bayesian	2.110 ± 0.008	2.466 ± 0.067	0.002

Bayesian method. The Bayesian analysis also gave a significant difference in $t_{lag(in)}$ of 0.28 h ($p = 0.02$). Although a change of approximately the same magnitude was observed in the NLS parameters, it failed to reach a significant level.

Obesity Study

Comparison of the $^{13}\text{COBT}$ data from NLS and Bayesian methods (table 3) indicates much the same characteristics as was found for the frequent feeding study. The highest values of F_{∞} obtained (always from obese subjects) were reduced by the Bayesian methodology so that the apparent significant difference observed between the two groups by NLS is much reduced with the Bayesian method. The corollary to this is that the estimates of $T_{1/2}^{BAY}$ also become more uniform between the groups, and the statistical significance of this parameter is also lost.

The effects of borrowing influence from other subjects are clearly illustrated in figure 7. For subjects where the GE curve is well-defined in the sampling interval, there is close agreement between NLS and Bayesian methods. However, for subjects with prolonged GE (i.e. when $T_{1/2}$ approaches or exceeds the sampling interval), the agreement becomes much poorer as can be seen by the deviations from the line of unity in the top right of the graph.

A somewhat surprising finding is that Bayesian analysis indicates that the macroparameter k^{BAY} is identical for lean and obese subjects, but that β^{BAY} appears to be greater in the obese. Associated with this is the much increased significance in the difference in $t_{lag(in)}$.

Meal Size Study

As stated in the original report, 5 of the 36 breath tests could not be satisfactorily modelled by NLS methods [12],

Table 4. ¹³COBT parameters obtained for the meal size study

	NLS			Bayesian		
	1 MJ (n = 12)	2 MJ (n = 10)	3 MJ (n = 9)	1 MJ (n = 12)	2 MJ (n = 12)	3 MJ (n = 12)
$F_{\bullet\bullet}$ ¹	0.500 ± 0.032	0.499 ± 0.028	0.515 ± 0.038	0.471 ± 0.024	0.496 ± 0.023	0.495 ± 0.021
k	0.585 ± 0.046*	0.437 ± 0.025	0.380 ± 0.046	0.588 ± 0.043*	0.418 ± 0.028	0.361 ± 0.033
β	4.35 ± 0.28	4.08 ± 0.037	4.44 ± 1.42	4.49 ± 0.32	4.30 ± 0.42	4.29 ± 1.01
T_{lag}	2.56 ± 0.12*	3.17 ± 0.14	3.39 ± 0.29	2.58 ± 0.12*	3.40 ± 0.18	3.62 ± 0.24
$T_{1/2}$	3.39 ± 0.18*	4.25 ± 0.17	4.82 ± 0.40	3.40 ± 0.17*	4.61 ± 0.26	5.09 ± 0.31
$T_{lag(in)}$	1.30 ± 0.07	1.53 ± 0.15	1.33 ± 0.32	1.35 ± 0.07	1.65 ± 0.18	1.57 ± 0.25
$T_{1/2(in)}$	1.77 ± 0.07*	2.16 ± 0.12	2.19 ± 0.25	1.81 ± 0.08*	2.34 ± 0.16	2.43 ± 0.22

* Indicates significant difference ($p < 0.05$) from the 2-MJ meal according to paired t test.

¹ Values differ from those previously quoted [12] due to a different method of calculating CO₂ production being used.

but this is apparently resolved using the Bayesian methods. A summary of the results is given in table 4. Although the conclusions are largely unchanged, Bayesian methodology consistently reports increased significance in observed differences (for example $p = 0.001$ vs. 0.006 for T_{lag} , and $p = 0.001$ vs. 0.004 for $T_{1/2}$).

Additionally, the Bayesian method produced estimates of T_{lag} ($p = 0.04$), $t_{lag(in)}$ ($p = 0.04$) and $t_{1/2(in)}$ ($p = 0.02$) which were significantly correlated with body mass index for the 3-MJ meal only. These associations were not observed with NLS estimates.

Meal Composition Study

In common with the NLS analysis, WinBUGS indicated that the 'standard' meal was retained in the stomach longest. The only material difference in the results from the two different analyses was that the Bayesian method indicated a significant difference for T_{lag} using ANOVA ($p < 0.05$), which was not observed by NLS. Post-ANOVA Bonferroni-corrected t tests indicated that WinBUGS detected a difference between the high protein meal and the standard meal. The meals did not give significant differences in $T_{1/2}$ either by Bayesian or NLS methods, but in both cases the input parameters $t_{lag(in)}$ and $t_{1/2(in)}$ showed that the standard meal emptied at a lower rate than the other three.

Discussion

We have demonstrated that Bayesian hierarchical methods form the basis of a convenient method of interpreting breath test data in GE studies. There are several

reasons that make the method appealing, which can be summarised as:

- (1) a 100% success rate in obtaining meaningful estimates of the model parameters, which is due to the internal consistency of the underpinning theoretical foundations of the method;
- (2) the automatic generation of credible intervals (confidence limits) for all parameters;
- (3) an apparent greater discrimination between populations that might be expected to display altered GE; and
- (4) the analysis can be performed with the freely available WinBUGS package (<http://www.mrc-bsu.cam.ac.uk/bugs/>); the code used to perform the modelling is relatively short (around two thirds of a standard notebook page).

Criticism of the method might include that it is computationally intensive, and that since all the model fits required for the study are performed simultaneously, a change in a single datum (for example to correct a typing error) requires that the whole set of breath-test curves be refitted. This is the case for hierarchical analyses, whether Bayesian or not. However, with relatively small studies such as those described (with around 50 breath tests in total), the fitting can be performed with WinBUGS on an average computer workstation in less than 10 min. In fact, this compares favourably with using the non-linear fitting routines in Microsoft Excel, for example, with several seconds being required for each individual fit to be performed. The advantage of the Bayesian approach is that it provides theoretically self-consistent and plausible estimates of individual and population precisions, albeit with extra care being required during data preparation.

Although a hierarchical approach was adopted in this illustration, Bayesian methods are equally applicable to the analysis of a single breath test. In this case the fitting took less than 30 s. Prior knowledge is still incorporated into the analysis, but when used in this way the model loses the advantages of shared information. However, this approach might be considered more relevant in a clinical setting where the result of a single breath test for diagnostic purposes is required.

No independent estimates (or true values) of the breath test parameters are available to confirm the results generated by the Bayesian methodology. However, from the fact that the results obtained are in almost perfect agreement with NLS methods when the breath test output curve is adequately defined within the timeframe of the measurement period, we can infer that the Bayesian estimates are reliable.

Indeed, in cases where GE proceeds at normal rates, and the sampling timeframe of the experiment is sufficiently long, there is no difference between the results obtained from NLS or Bayesian methods (for example in the frequent feeding study; fig. 5). The advantages of the Bayesian method are demonstrable only when the dataset contains subjects illustrating delayed substrate processing, for which the data would have to be discarded if NLS methods are deployed.

A further advantage of the Bayesian method over NLS techniques is that the former can accommodate errors in the basal (pre-test) isotopic composition in breath. Obtaining a good estimate of the basal value is vital when using NLS methods since this is subtracted from all of the other time points. Under the Bayesian scheme the measured value can be associated with a suitable prior (in this work a normal distribution with a standard deviation of 1% was used), and then treated as a stochastic node in the analysis. As an example, the prior and post distributions of a typical basal value are shown in figure 8.

The consequences of the sampling regime on the results of GE experiments have been investigated by others [16]. The conclusion from this work was that the breath test should be conducted over a 4-hour period with a 30-min period between samples. Clearly, this protocol would not be satisfactory for some of the studies considered in this work. Even with the test prolonged to 6 h there were an appreciable number of failures when NLS methods were used. The Bayesian approach can be used to eliminate these failures, but caution must be exercised. For example, by limiting the period over which fitting is performed, the 4-hour protocol can be investigated. A marked decrease in precision of the estimated parameters

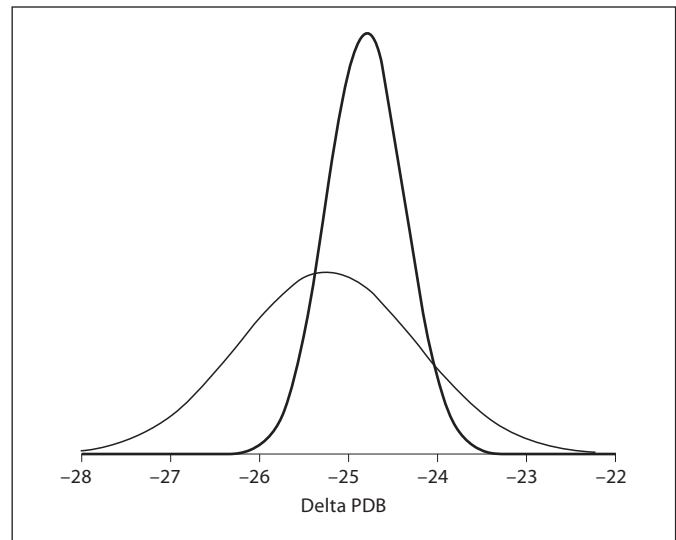


Fig. 8. Prior (light line) and posterior (heavy line) distributions for a typical basal delta value.

occurs with the reduced sampling time, e.g. the fraction of cases where $T_{1/2}$ is estimated with a coefficient of variation of less than 5% drops from 66% to only 17%, and whilst 99% of all breath tests returned estimates of this parameter with a c.v. of less than 10% using the full 6-hour data, the fraction decreased to 78%.

A difficulty with the $^{13}\text{COBT}$ is that it has no well-defined physiological basis to the model used for its interpretation: the form of the equation used is purely based on the empirical observation that the output has the form of a bell-shaped curve. This means that the two parameters k and β cannot be unambiguously associated with a particular physiological process. Because of this, the significance of observations such as the apparent change in β in obesity revealed by the Bayesian method are unclear. For any given k , increasing β not only represents a delay in substrate processing, but also an increase in the maximum rate at which the substrate is processed [10].

Studies of kinetics at the population level are widely used in drug pharmacokinetics since knowledge of intersubject variability ensures the safety of dosing regimes. The first introduction of these methods into physiological modelling was their application to the minimal model of glucose kinetics [17–19]. As is the case for GE, using NLS techniques to fit the minimal model is subject to failure in some instances, producing implausible (i.e. negative) values for insulin sensitivity or glucose effectiveness [20]. Bayesian analysis with suit-

ably selected priors has been shown to eliminate this problem [17].

In summary, Bayesian hierarchical methods provide a robust and reliable way to analyse the GE breath test. The methods avoid parameter estimation failure in cases

where data quality is sub-optimal, yet reproduce the results of standard NLS methods when the data is well-behaved. We would recommend that the WinBUGS package be adopted routinely for the estimation of GE by $^{13}\text{COBT}$.

References

- 1 Guillot E, et al: Intestinal-absorption and liver uptake of medium-chain fatty-acids in non-anesthetized pigs. *Br J Nutr* 1993;69: 431–442.
- 2 Sanaka M, Yamamoto T, Kuyama Y: Theoretical flaws in the gastric emptying breath test: why is it dubious? *Dig Dis Sci* 2005;50: 15–17.
- 3 Bluck LJ, Coward WA: Measurement of gastric emptying by the C13-octanoate breath test – rationalization with scintigraphy. *Physiol Meas* 2006;27:279–289.
- 4 Schofield W: Predicting basal metabolic rate, new standards and review of previous work. *Hum Nutr Clin Nutr* 1985;39(Suppl 1):5–41.
- 5 Evenepoel P, et al: Digestibility of cooked and raw egg protein in humans as assessed by stable isotope techniques. *J Nutr* 1998;128: 1716–1722.
- 6 Jackson SJ, Bluck LJ: Measurement of gastric emptying by octanoate metabolism. *Curr Opin Clin Metab Care* 2005;8:538–544.
- 7 Schommartz B, Ziegler D, Schadewaldt P: Significance of diagnostic parameters in [C-13]octanoic acid gastric emptying breath tests. *Isotopes Environ Health Stud* 1998;34: 135–143.
- 8 Siegel JA, et al: Biphasic nature of gastric emptying. *Gut* 1988;29:85–89.
- 9 Couturier O, et al: Gastric emptying of solids: estimates of lag phase and constant emptying times. *Nucl Med Commun* 2000;21: 665–675.
- 10 Jackson SJ, et al: Frequent feeding delays the gastric emptying of a subsequent meal. *Appetite* 2007;48:199–205.
- 11 Jackson SJ, et al: Delayed gastric emptying in the obese: an assessment using the non-invasive (13)C-octanoic acid breath test. *Diabetes Obes Metab* 2004;6:264–270.
- 12 Jackson SJ, Bluck LJ, Coward WA: Use of isotopically labelled octanoic acid to assess the effect of meal size on gastric emptying. *Rapid Commun Mass Spectrom* 2004;18:1003–1007.
- 13 Lunn DJ, et al: WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput* 2000;10:325–337.
- 14 Sidossis LS, et al: A new correction factor for use in tracer estimations of plasma fatty-acid oxidation. *Am J Physiol* 1995;32:E649–E656.
- 15 Bluck LJ, et al: Measurement of gastric emptying in man using deuterated octanoic acid. *Rapid Commun Mass Spectrom* 2002;16: 127–133.
- 16 Yamamoto T, et al: Modified C-13-octanoate breath test and impact of sampling points. *J Clin Gastroenterol* 2004;38:669–670.
- 17 Agbaje OF, et al: Bayesian hierarchical approach to estimate insulin sensitivity by minimal model. *Clin Sci* 2003;105:551–560.
- 18 Andersen KE, Hojbjerg M: A population-based Bayesian approach to the minimal model of glucose and insulin homeostasis. *Stat Med* 2005;24:2381–2400.
- 19 Anderson KE, Hojbjerg M: A Bayesian Approach to Bergman's Minimal Model; in: Ninth International Workshop on Artificial Intelligence and Statistics. Key West, Microsoft, 2003.
- 20 Godsland IF, Walton C: Maximising the success rate of minimal model insulin sensitivity measurement in humans: the importance of basal glucose levels. *Clin Sci* 2001;101:1–9.