Research article

**13C-Urea breath test threshold calculation and evaluation for the detection of Helicobacter pylori infection in children**

Ralf Herold and Michael Becker *

Address: Otto-Heubner-Center for Pediatric and Adolescent Medicine Charité, Virchow Hospital Humboldt University, Augustenburger Platz 1, D-13353 Berlin, Germany

E-mail: Ralf Herold - ralf.herold@charite.de; Michael Becker* - michael.becker@charite.de

*Corresponding author

Abstract

**Background:** The 13C-urea breath test (UBT) is performed in adults and children with epigastric pain for non-invasively diagnosing a suspected H. pylori infection. Criteria for UBT interpretation have not been generally agreed on and test reliability has not been established in children of different ages. This study aimed at identifying reliable UBT thresholds in children by using 251 UBTs in conjunction with reference histology and by analyzing 1232 UBTs.

**Methods:** At baseline and 30 and 60 minutes after the administration of 75 mg 13C-urea to children and adolescents (0.25 to 18 years of age), the differences (Δ) of 13CO2/12CO2 ratio in exhaled air (δ) were determined by mass spectrometry. UBT Δδ value thresholds were calculated in random subgroups and evaluated in complementary subgroups using logistic regressions on reference histology or bimodal distribution analyses of Δδ values from UBTs alone.

**Results:** Δδ values were higher (median, 15.4‰) in positive (133/251, 53 %) than in negative histology (2.4‰). At 30 minutes, the calculated cut-off was 5.3‰ (mean regression determination R² = 0.91), and sensitivity (0.95), specificity (0.97), positive (0.97) and negative predictive values (0.95) were higher than at 60 minutes (threshold 6.8‰, R² = 0.85). Similar thresholds resulted from UBTs analysis (5.8‰ and 6.2‰) when sensitivity and specificity were maximized (concordance probabilities, 0.99 and 0.99). There was no systematic age effect.

**Conclusions:** In children, 13C UBT cut-offs were obtained and specially validated, entailing high accuracy of non-invasively testing for gastric H. pylori infection.

Background

Painful chronic gastritis and duodenal ulcerations in children are mostly associated with an infection of the gastric mucosa by H. pylori (for reviews see [1,2]). Invasive tests for the detection of H. pylori are usually performed in biopsy specimens of the stomach (histological staining, bacterial culture and the rapid urease test) or in the blood (anti-HP antibodies). The use of these tests in adults and children is well established (for reviews see [3–5]). Since the introduction of the non-invasive 13C-urea breath test (UBT) by Graham et al. 1987 [6], several attempts have been made to define the minimum criteria for a positive UBT in adults [4,7–12].

For general use in children, however, criteria for the performance of the UBT and the interpretation of its results
are not yet sufficiently defined. Feasibility and promising precision of the UBT had been shown in conjunction with invasive tests, for example, in studies including 50 [13], 88 [14], and, recently, 115 children [15], in whom a 98% sensitivity and specificity at a rather low cut-off (\( \Delta \delta = 3.5\% \)) was asserted.

Another recent study in 247 Gambian children up to four years reported a cut-off of 5.47‰ for H. pylori infection of in 1532 UBTs at 30 minutes, validating with histology results in only 14 cases [16]. Their method was similar to our secondary aim’s method, which was to calculate a threshold from 1232 UBTs alone in children and adolescents from 1 to 18 years of age. Primarily, our study aimed at calculating UBT thresholds from reference histology in 251 children from 5 to 16 years of age, validating with a prospective method. We additionally demonstrate that the thresholds produced by both methods are in close agreement.

Methods

\( \text{\^{13}} \text{C-urea breath test} \)

The \( \text{\^{13}} \text{C} \) urea breath test (UBT) was performed as described previously [17]. In short, after at least 4 hours of fasting, 75 mg \( \text{\^{13}} \text{C} \)-urea (99% \( \text{\^{13}} \text{C} \); Promochem, Wesel, Germany) was given in 50 ml water shortly after a meal, which usually consisted of half a non-sweetened roll. Breath samples were collected in duplicate before ingestion of the \( \text{\^{13}} \text{C} \)-urea (t0) and at 30 and 60 minutes afterwards (t1). If a child could not expire into the collection bag, breath samples were collected by applying a breathing mask with two unidirectional valves connected to a breathing bag. Normal free breathing activity filled the bag, from which a sample was drawn and transferred to an evacuated tube [17].

Breath samples were analyzed twice by gas isotope ratio mass spectrometry (Delta E; Finnigan MAT, Bremen, Germany). The measured \( \text{\^{13}} \text{CO}_2/\text{\^{12}} \text{CO}_2 \) ratio was compared with the known isotope ratio of a standard gas. The \( \text{\^{13}} \text{CO}_2 \) values were corrected for the international PDB-standard, then averaged and expressed as relative delta per mil difference to the standard gas (\( \delta [\% oo] \)). In the following, \( \Delta \delta [\% oo] \) is used to denote the difference between \( \delta \) at time \( t_1 \) and at baseline.

Subjects

251 children and adolescents (135 boys and 116 girls; aged 5–16 years, median 11 years) who had not received antibiotics during the previous two months were included because of recurrent epigastric pain for at least 3 months. (Children younger than four years almost never presented with symptoms indicating esophagogastrroduodenoscopy.)

Esophagogastrroduodenoscopy, UBT, and histological examinations were performed by three different investigators, each unaware of the others’ results. The histological identification of H. pylori in four biopsy specimens from the gastric antrum and corpus was used as reference. At least one biopsy specimen was required to be positive. The pathologist performed histological identification according to established staining methods and criteria [18] and was blinded to all specimens.

Additionally, a total of 1232 UBTs were performed (between January 1st 1995 and June 1st 1998) in 1232 German children with chronic abdominal pain (1 month to 18 years of age). Informed consent was given by the parents and children, where obtainable.

The study was carried out in accordance with the Helsinki Declaration for Human Studies and the protocol was approved by the local ethics committee.

Calculations

Analyses and graphics were produced with S-Plus4.5 (MathSoft, Seattle) and SAS6.12 (SAS Institute, Cary). Robust statistics included the median and the rescaled median absolute deviation, which compare to mean and standard deviation. Natural logarithms of \( \Delta \delta \) values were used throughout. In histograms, the bin width was chosen by data-based kernel smoothing [19] and the bin height was scaled to a unit area under the curve [[20], p.168].

Logistic regressions of the histological detection of H. pylori (positive or negative) on the \( \Delta \delta \) values were performed in randomly chosen 50% of the data. The resulting threshold \( \Delta \delta \) values were evaluated in the complementary 50% of the data on the basis of the observed H. pylori detection rate. Coefficients of determination (R²) were obtained according to [21].

The observed binormal distributions of logarithmic \( \Delta \delta \) values from UBTs alone were modeled with sum functions of two normal distributions using maximum likelihood estimation [[19], p.287ff]. Assuming the two normal distributions represent either absent or present H. pylori infection, threshold \( \Delta \delta \) values were calculated at the maximum of specificity plus sensitivity and of negative plus positive predictive values.

Average bias corrected estimates of the thresholds and the associated test performance characteristics were obtained by 1000-fold bootstrapping [22].

The concordance rate was the area under the receiver operating characteristic curve [23].
The Kolmogorov-Smirnov goodness-of-fit-test was used to compare distributions.

**Results**

**Feasibility and \( \Delta \delta \) values**

Children from 4 months of age on were successfully studied with the aforementioned UBT procedure. The standard error of measuring delta values (\( \Delta \delta \)) was 0.05‰. \( \Delta \delta \) values measured in breath samples taken 30 and 60 minutes after \(^{13}\)C-urea ingestion correlated neither with body weight (Spearman’s rank correlation \( R^2 = 0.03 \)) nor with age (\( R^2 = 0.04 \)).

The distribution of delta values is shown in figure 1. Median \( \Delta \delta \) values of histology validated UBTs at 30 minutes (panel A) and 60 minutes (not shown) were 2.4‰ and 3.2‰ in the absence and 15.4‰ and 17.3‰ in the presence of H. pylori, respectively. Values from the UBTs are shown in panel B (figure 1), together with the two normal functions and their sum fitted for calculate of the cut-off.

**Thresholds in histology validated UBTs**

The \( \Delta \delta \) value thresholds were 5.3 ± 0.3‰ (mean and standard deviation) at 30 minutes (figure 1, panel A) and 6.8 ± 0.5‰ at 60 minutes, respectively, in 251 histology validated UBTs (\( R^2 = 0.91 \) and \( R^2 = 0.85 \), respectively).

\( \Delta \delta \) values at 30 minutes were maximum 6.6‰ in the histological absence and minimum 3.4‰ in the histological presence of H. pylori. Only 18.3% of these 30 minutes UBT \( \Delta \delta \) values were between 3.4‰ and 6.6‰.

When positive biopsies occurred, always both antral biopsy specimens were positive. Inconsistently, the corpus was positive in addition to the positive antral specimens.

**Thresholds in UBTs**

Threshold \( \Delta \delta \) values were 5.8 ± 0.4‰ at 30 minutes and 6.2 ± 0.4‰ at 60 minutes in 1232 UBTs. In the models’ two normal distributions, i.e., in the presumed absence and presence of H. pylori, mean \( \Delta \delta \) values were 1.83 ± 1.97‰ and 16.91 ± 1.86‰ at 30 minutes. At 60 minutes, the mean \( \Delta \delta \) values were 1.96 ± 2.02‰ and 15.64 ± 1.67‰.

The modeled sum functions were not different from the observed \( \Delta \delta \) distributions (\( P = 0.45 \) and \( P = 0.08 \), respectively). The \( \Delta \delta \) values and the respective thresholds were similarly located in histology validated UBTs (figure 2, panel A) and in UBTs alone (figure 2, panel B).

The portions of presumably positive UBTs were 24.4% at 30 minutes and 23.9% at 60 minutes in UBTs alone.

**Threshold in UBTs in different age groups**

Threshold \( \Delta \delta \) values at 30 minutes varied from 9.1 ± 4.5‰ in children up to 5 years of age to 5.0 ± 1.5‰ in adolescents (14–18 years of age). This seemingly negative association in the arbitrarily chosen age groups was not evident in a plot of all UBT values by age (not shown). In the youngest age group, the bootstrap procedure’s variance indicates this model’s instability (table 1).

**Threshold evaluation**

Standard performance measures (sensitivity, specificity, negative and positive predictive values, correct classification rates, and concordance rates) were evaluated at the calculated cut-offs and almost always exceeded 0.90 (table 1).
The 251 UBT values associated with positive and negative H. pylori histology were cross evaluated using the method for analysis of the UBTs only, resulting in threshold $\Delta \delta$ values of 5.7‰ at 30 minutes and 6.7‰ at 60 minutes. The performance measures sensitivity, specificity, positive and negative predictive values were 0.95, 0.99, 0.99, and 0.94, at 30 minutes. Vice versa, binormal models of UBTs $\Delta \delta$ values were input to logistic regression, to produce figure 3, panel B.

**Discussion**

This study newly found $^{13}$C urea breath test (UBT) $\Delta \delta$ thresholds that were robustly associated with histological H. pylori detection in children. Thresholds obtained in separate groups varied from 9‰ in small children to 5‰ in children and adolescents. The non-optimistic validation and the special evaluation of the UBT demonstrated high accuracy.

The threshold calculation of this study comprised logistic regressions and iterative modeling analyses. The analytical principle is superior to threshold selection procedures, even if the often used single value dichotomy is replaced by receiver operator characteristics including all test values (for example, [11]). For comparison, we used iterative analyses in a large number of UBTs from symptomatic children in whom gastric biopsies were not performed. These analyses returned threshold values, which were similar to the aforementioned results, apart from our youngest age group, which had a rather large error. However, a recent study of 3 months to 4 years old Gambian children agreeably complements our findings by obtaining a 5.5‰ threshold with a very similar method [16].

This study's design was based on others' findings that factors suspected to influence the UBT, such as the quantity and formulation of $^{13}$C-urea, the fasting period, the test meal composition, and the timings of breath sample col-
lection were generally not important [3,7–12,14,23]. Here, ∆δ values were higher at 60 than at 30 minutes, which is in agreement with the 13C-urea elimination kinetics. According to our results, breath test samples collected at 30 minutes were slightly more useful when analyzed for correspondence with the histological detection of H. pylori in gastric biopsy specimen. On the other hand, the two peaks of UBT ∆δ value distribution were slightly more separated in samples obtained 60 minutes after the test meal. In face of the small differences, either time point seems appropriate.

The many approaches for performing the UBT in adults have resulted in various threshold proposals ranging from about 3.5‰ to 6‰. In our study, only 16.3% of UBT samples at 30 minutes were in this comparatively wide range. In order to clarify the impact of H. pylori in gastric biopsy specimen. On the other hand, the two peaks of UBT ∆δ value distribution were slightly more separated in samples obtained 60 minutes after the test meal. In face of the small differences, either time point seems appropriate.

With regard to the effect of age, there were only few studies on rather small groups of children, not including infants [3,14,23]. Recently, a correlation of UBT values with the children's age was proposed, while the portion of explained variance was not given and the significance might be explicable by the large numbers (1016, of which 149 were histologically validated [25]). We could not demonstrate a systematic influence of age, notwithstanding the higher, but greatly variable thresholds in children up to 5 years. A partial shortcoming rests in (not only) our assumption that the modes of the UBT results represent negative or positive H. pylori detection. In contrast to other studies, we demonstrated that the model distributions based on this assumption are not different from the observed distributions. However, clarification of an age effect would require endoscopy in all age groups and groups of different H. pylori prevalence.

In our study, endoscopy with biopsies was prompted only by patients' symptoms, and not by UBT results. Resampling by bootstrapping was used to additionally reduce patient bias. Furthermore, thresholds were cross-validated, that is, thresholds obtained in one portion of the tests were evaluated prospectively in the complementary portion. Thus, the great accuracy reported in this study is not optimistically biased. Although these methods are very useful to validate conclusions and cut-offs in data analysis, they have not been reported with the UBT previously.

Additionally, thresholds other than the proposed ones could be chosen on the basis of the provided figures, which indicate the associated H. pylori detection probabilities.

Table 1: Results of the 13C-UBT in children and adolescents*

| Patients and Methods | Sample Collection Time [Minutes] | 30 | 60 | 30 | 60 | 30 | 30 | 30 | 30 | 30 | 30 | 60 |
|---------------------|---------------------------------|----|----|----|----|----|----|----|----|----|----|----|
| Age Limits [Years]  |                                 | 0.3–5 | 5–8 | 8–11 | 11–14 | 14–18 | |
| Number of Tests     |                                 | 251 | 251 | 251 | 138 | 289 | 383 | 278 | 137 | 1232 | 1232 | |
| Method of Analysis  |                                 | LR | LR | BM | BM | BM | BM | BM | BM | BM | BM | BM |
| Histology used      |                                 | Yes | Yes | Yes | Yes | No | No | No | No | No | No | No |
| Threshold Results   |                                 | 5.29 | 6.76 | |
| Threshold from Logistic Regression (LR) [%δ] | |
| LR                  | 0.3 | 0.5 | |
| R² LR               | 0.91 | 0.85 | |
| Threshold from Specificity [%δ] | 5.66 | 6.72 | 9.10 | 5.57 | 6.83 | 6.46 | 4.97 | 5.78 | 6.19 | |
| SD [%δ]             | 4.46 | 1.22 | 0.86 | 1.05 | 1.46 | 0.45 | 0.42 | 0.54 | 0.53 | |
| Positive plus Negative Predictive Values [%δ] | 5.54 | 6.53 | 12.25 | 7.78 | 8.11 | 7.24 | 5.56 | 7.19 | 7.55 | |
| SD [%δ]             | 5.10 | 1.66 | 1.00 | 1.19 | 1.76 | 0.56 | 0.53 | 0.99 | 0.99 | |

*Mean results from bootstrapping. Empty fields: not applicable. BM: Binormal maximum likelihood modeling. LR: Logistic regression. SD: Standard Deviation. UBT: Urea Breath Test. AUROC: Area under the receiver operator characteristic curve.
Conclusions
In conclusion, this study calculated and validated urea breath test thresholds in a large pediatric patient population, and showed that this test reliably and non-invasively detects a H. pylori infection. In children, cut-offs such as 5.3‰ and 6.8‰ in UBT samples obtained at 30 and 60 minutes after ingestion of 75 mg 13C-urea entail high accuracy (0.94 and 0.98, respectively) for gastric H. pylori detection.

List of abbreviations
AUROC Area under the receiver operator characteristic curve
BM Binormal maximum likelihood modeling
LR Logistic regression
SD Standard Deviation
UBT Urea Breath Test

Competing interests
None declared.

Acknowledgements
The authors thankfully acknowledge the contributions of Prof Dr H. Helge, Dr M. Unger, Dr H. Brosicke, and A. Pfundstein (Berlin, Germany).

References
1. Drumm BB: Helicobacter pylori in the pediatric patient. Gastroenterol Clin North Am 1993, 22:169-82
2. Prieu G, Polanco I, Larrauri J, Rota L, Lama R, Carrasco S: Helicobacter pylori infection in children: clinical, endoscopic, and histologic correlations. J Pediatr Gastroenterol Nutr 1992, 14:420-5
3. Vandenbergas Y, Blecker U, Druzeke T, Keppens E, Nijs J, Cadranel S, Mathias C, Durnin M, Drumm B: Carbon 13-labeled urea breath test for the detection of Helicobacter pylori gastritis in children. Pediatr 1991, 90:608-11
4. Thijs JC, van Zwer AA, Thijs WJ, Oey HB, Karrenbeld A, Stellaard F, Vandenplas Y, Blecker U, Druzeke T, Keppens E, Nijs J, Cadranel S, Mathias C, Durnin M, Drumm B: Diagnostic tests for Helicobacter pylori: a prospective evaluation of their accuracy, without selecting a single test as the gold standard. Am J Gastroenterol 1996, 91:2125-2129
5. Logan RP: Urea breath tests in the management of Helicobacter pylori infection. Gut 1998, 43:547-550
6. Graham DY, Klein PD, Evans DJ, Evans DG, Alpert LC, Opekun AR, Bouton TW: Campylobacter pylori detected noninvasively by the 13 C-urea breath test. Lancet 1987, 1:1174-1177
7. Klein PD, Graham DY: Minimum analysis requirements for the detection of Helicobacter pylori infection by the 13 C-urea breath test. Am J Gastroenterol 1993, 88:1865-1869
8. Klein P, Malaty H, Martin R, Graham K, Genta R, Graham D: Noninvasive detection of Helicobacter pylori infection in clinical practice: The 13 C urea breath test. Am J Gastroenterol 1996, 91:690-694
9. Megraud F: Advantages and disadvantages of different current diagnostic tests for the detection of Helicobacter pylori. Scand J Gastroenterol Suppl 1996, 215:57-62
10. Mion F, Delecluse HJ, Rousseau M, Berger F, Brazier J, Minaire Y: 13C-urea breath test for the diagnosis of Helicobacter pylori infection. Comparison with histology. Gut 1994, 31:1106-1111
11. Mion F, Rosner G, Rousseau M, Minaire Y: 13C-urea breath test for Helicobacter pylori: cut-off point determination by cluster analysis. Clin Sci 1997, 93:3-6
12. Cutler AF, Havstad S, Ma CK, Blaser MJ, Perez Perez GI, Schubert TT: Accuracy of invasive and noninvasive tests to diagnose Helicobacter pylori infection. Gastroenterology 1995, 109:136-141
13. Etieurne M, Brueton MJ, Fraun S, De Brands M, De Jonge C: Diagnosis of Helicobacter pylori gastritis in children using the 13 C urea breath test. J Clin Gastroenterol 1999, 28:238-240
14. Rowland M, Lambert I, Gormally S, Daly LE, Thomas JE, Hetherington C, Durnin M, Drumm B: Carbon 13-labeled urea breath test for the diagnosis of Helicobacter pylori infection in children. J Pediatr 1997, 131:815-820
15. Bazzoli F, Cecchini L, Covaglia L, Dell’Antonia M, De Giacomo C, Fossi S, Casalli LG, Gullini S, Lazzuri R, Leggeri G, Lerro P, Valdarni V, Mandrioli G, Marani M, Martelli P, Minato A, Nicolini G, Oderda G, Pazzi P, Pozzato A, Ricciardiello L, Roda E, Simonini P, Sottilli S, Zagari RM: Validation of the 13 C-urea breath test for the diagnosis of Helicobacter pylori infection in children: a multicenter study. Am J Gastroenterol 2000, 95:646-650
16. Thomas JE, Dale A, Harding M, Coward WA, Cole TJ, Sullivan PB, Campbell DI, Warner BF, Weaver LT: Interpreting the 13 C-urea breath test among a large population of young children from a developing country. Pediatr Res 1999, 46:147-151
17. Kohrt R, Becker M, Brosicke H, Kruger N, Helge H: Prevalence of Helicobacter pylori infection in Nicaraguan children with persistent diarrhea, diagnosed by the 13 C-urea breath test. J Pediatr Gastroenterol Nutr 1997, 25:84-88
18. Caspary WF, Arnold R, Bayerdorffer E, Behrens R, Birkner B, Brand W, Domshke W, Labenz J, Koletzko S, Halfereirher B, Menge H, Ro- schke Y, Schepp W, Strauch M, Stolte M: Diagnosis and therapy of Helicobacter pylori infection. Guidelines of the German Society of Digestive and Metabolic Diseases. Z Gastroenterol 1996, 34:392-401
19. Wand M: Data-based choice of histogram bin width. The American Statistician 1997, 51:59-64
20. Venables W, Ripley B: Modern applied statistics with S-Plus. New York, Springer 1997
21. Nagelkerke NJ, Fidler V, Buwalda M: Instrumental variables in the evaluation of diagnostic test procedures when the true disease state is unknown. Stat Med 1988, 7:739-744
22. Shao J, Dongsheng T: The jackknife and bootstrap. New York, Springer 1995
23. Hanley JA, McNeil BJ: The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982, 143:29-36
24. Kalach N, Briet F, Raymond J, Benhamou PH, Barbier P, Bergeret M, Senocci L, Maurel M, Flourie B, Dupont C: The 13 carbon urea breath test for the noninvasive detection of Helicobacter pylori in children: comparison with culture and determination of minimum analysis requirements. J Pediatr Gastroenterol Nutr 1998, 26:291-296
25. Kinnermann A, Demmlmair H, Koletzko B, Krauss-Etschmann S, Wiebecke B, Koletzko S: Influence of age on 13 C-urea breath test results in children. J Pediatr Gastroenterol Nutr 2000, 30:85-91

Pre-publication history
The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1471-230X/2/12/pre-pub