

Dietary Amelioration of *Helicobacter pylori* Infection: Design Criteria for a Clinical Trial

Jed W. Fahey,^{1,2} Alvaro Muñoz,³ Yasushi Matsuzaki,⁴ Hideo Suzuki,⁴ Paul Talalay,^{1,2} Masafumi Tauchi,⁴ Songhua Zhang,⁴ Christopher Hurt,¹ and Akinori Yanaka⁴

¹Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine and ²Center for Human Nutrition and ³Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland and ⁴Department of Gastroenterology and Hepatology, Tsukuba University, Tsukuba, Japan

Abstract

The longitudinal stability of the urea breath test (UBT), which measures urease as a biomarker for infection with *Helicobacter pylori* (a major risk factor for gastric cancer), was evaluated in the environs of Tsukuba, Japan. ¹³C-UBT measurements were monitored at four time points in 46 free-living, *H. pylori*-infected, asymptomatic volunteers over a period of 7 weeks. Subjects were asked to refrain from eating cruciferous vegetables, which might confound interpretation of results. Their compliance was monitored using both dietary records and direct biochemical testing of overnight urine. There was large between-subject UBT variation in this population (logUBT mean, 3.34; SD, 0.67). Within-subject (longitudinal) UBT values were remarkably stable in about one-quarter of the subjects (coefficients of variations for these individuals were

<21%), whereas coefficients of variations in the highest quartile of variability ranged from 40% to 80%. About half of the sequential UBTs (63 of 138 such measurement pairs) changed >10% “delta over baseline” between measurements. This study provides the elements to optimize the design of a clinical trial in this population to examine the efficacy of a dietary intervention to reduce *H. pylori* infection. The number of subjects required to detect a 30% difference in average UBT value is highly dependent on the baseline stability of UBT measurements. For the least variable quartile, as few as 12 subjects would be needed; for the most variable quartile, at least 147 subjects would be required in each arm. (Cancer Epidemiol Biomarkers Prev 2004;13(10):1610–6)

Introduction

Helicobacter pylori is recognized by the WHO as a class I carcinogen and is a causative agent for gastric carcinogenesis. Gastric cancer is the second leading cause of worldwide cancer mortality, with >875,000 cases in 2000 and ~650,000 deaths (1). More than half of the world's population is infected with *H. pylori*, and although the vast majority of infected individuals never develop gastric cancer, for those who do, the risk attributable to *H. pylori* infection is high. Recent estimates are that ~0.5 million new cases of gastric cancer yearly (about 55% overall) are directly attributable to infection with *H.*

pylori (2). Japan has a substantially higher incidence of gastric cancer than any other region of the world (3). Thus, the ability to lower the prevalence of *H. pylori* infection and the consequent risk of cancer is of compelling importance.

The recent demonstrations of both potent *in vitro* and *in vivo* antibiotic activities of sulforaphane (4-methylsulfinylbutyl isothiocyanate) against *H. pylori* (4, 5) suggest the possibility of ameliorating the effects of this organism by treatment with a sulforaphane-rich dietary supplement. Although in prior studies the efficacy of diet-based therapies was verified by invasive techniques such as gastric biopsy (e.g., ref. 6), validation in larger clinical trials requires biomarkers obtained through non-invasive testing. The urea breath test (UBT) described below is such a biomarker. Several studies that depended solely on this test to estimate the degree of *H. pylori* colonization have produced inconsistent results. Whether the variation in repeated measurements over time was due to inherent biological differences or to imprecision in the test was unclear.

The UBT is often considered the “gold standard” of noninvasive methods for diagnosis of *H. pylori* infections. It measures the activity of urease, an enzyme that is abundant in *H. pylori* but is absent from human tissues. Urease hydrolyzes urea, an end product of mammalian nitrogen metabolism, giving rise to CO₂ and NH₃. By neutralizing gastric acidity, NH₃ is believed to allow *H. pylori* to survive on the gastric mucosa. Hydrolysis

Received 2/19/04; revised 4/22/04; accepted 5/3/04.

Grant support: World Cancer Research Fund, American Institute for Cancer Research, Barbara Lubin Goldsmith Foundation, Brassica Foundation, and Lewis B. and Dorothy Cullman Foundation.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: J.W. Fahey and A. Yanaka contributed equally to this study. Two of the authors (J.W. Fahey and A. Yanaka) and Johns Hopkins University are founders of Brassica Protection Products, a company with a mission to develop chemoprotective food products and that sells broccoli sprouts. They (and Johns Hopkins University) are also stockholders as well as scientific consultants to Brassica Protection Products. Their stock is subject to certain restrictions under university policy. The terms of this arrangement are being managed by Johns Hopkins University in accordance with its conflict of interest policies.

Requests for reprints: Dr. Jed W. Fahey, Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine, 725 North Wolfe Street, 406 WBSB, Baltimore, MD 21205-2185. Phone: 410-614-2607; Fax: 410-614-3248. E-mail: jfahey@jhmi.edu

Copyright © 2004 American Association for Cancer Research.

of a dose of administered ^{13}C -urea by the infective organism gives rise to $^{13}\text{CO}_2$ in the breath and the difference in the ratio of $^{13}\text{CO}_2$ to $^{12}\text{CO}_2$ after a standard dose of ^{13}C -urea provides a measure of the urease activity and, according to some reports, the severity of infection. The measurements are usually expressed as the changes in $^{13}\text{C}/^{12}\text{C}$ ratios in the exhaled air (in parts per thousand or ‰) from the baseline levels to a specific time point after dosing. These changes are customarily expressed as "delta over baseline" (DOB). A typical dose containing 75 to 100 mg of $>99\%$ - ^{13}C urea is administered to fasting subjects in 100 mL of orange juice or citric acid, which assure gastric acidity and inhibit gastric emptying. Although some oral bacteria also exhibit urease activity, most common gastric bacteria do not (7).

The UBT has been used extensively to diagnose *H. pylori* infection, identifying infected individuals but not quantifying their level of bacterial colonization. There is well-documented correspondence between UBT values and prevalence of infection on a population basis. Several studies have shown good correlation of UBT values with *H. pylori* colonization based on histology scores (8-10) and bacterial culture and serology (11, 12). Whereas inflammation did not correlate with UBT score (13), there was a significant relationship between UBT and endoscopic findings of (a) magnitude and extent of *H. pylori* colonization in both corpus and antrum as determined by histology, (b) neutrophil accumulation in the antrum, (c) atrophy of the corpus, and (d) intestinal metaplasia of the corpus ($n = 169$ subjects; ref. 14). In addition, there was excellent correlation between the number of *H. pylori* genomes per picogram of paired antrum and corpus in DNA from human biopsies ($n = 88$ subjects; ref. 15).

The reliability of the test as a biomarker of colonization density at the individual subject level remains uncertain. For studies that follow the longitudinal effects of treatment over time, it is essential to know whether measurements remain stable and reproducible. This is particularly important in determining the effectiveness of a treatment in which only a modest antibiotic effect is expected, for example, from a dietary intervention. Normal physiologic or measurement variation might mask this effect. Therefore, data collected in the present study were used to determine the variability of UBT in a Japanese population in which there is a high incidence of *H. pylori* infection to plan the most appropriate trial for treatment of this widespread risk factor. We have thus evaluated the variance components of UBT measurements within and between individuals. We have assessed the factors associated with the different components of variance and present evidence to guide the design of clinical trials by selection of individuals whose within-individuals variance falls within a predetermined range.

Materials and Methods

Subjects. The subjects in this investigation comprised the control population of a larger study of the effect of dietary intervention in subjects infected with *H. pylori*. Forty-six volunteers were studied in whom *H. pylori* colonization was verified by a urinary *H. pylori* antigen test (Rapirun, Otsuka Pharmaceutical Co., Ltd., Tokyo, Japan; ref. 16) or, for five subjects, by UBT. Subjects were followed for 7 weeks according to the protocol described

in Fig. 1, and had a total of four UBTs, on days 0 (UBT1), 7 (UBT2), 21 (UBT3), and 49 (UBT4). The subjects were not hospitalized, and compliance with the dietary restrictions of the protocol was assessed by questionnaires and by biochemical verification as described below. Demographic information obtained for each participant was supplemented by queries on smoking and alcohol use, self-reported gastrointestinal symptoms, and prior endoscopy. Data on smoking and drinking status presented in Table 1 pool respondents classifying themselves as both light (1-20 cigarettes per day) and heavy (≥ 21 cigarettes per day) smokers and pools both light (≤ 2 drinks per week) and heavy (≥ 3 drinks per week) alcohol drinkers.

Urea Breath Test. The original test developed in 1987 used mass spectrometry for measurement of $^{13}\text{C}/^{12}\text{C}$ ratios (17). Isotope-selective IR spectroscopy has made the test available more widely in doctors' offices with the use of portable IR spectrophotometers (18-20). For this study, an UBiT-IR300 was used, which had a limit of detection of 0.3‰ DOB. Fasting subjects provided baseline breath samples in 300 mL aluminized plastic collection bags (Otsuka Electronics Co., Ltd., Tokyo, Japan) and then quickly swallowed a 100 mg tablet of ^{13}C -urea (Yubitto). A second breath sample was taken 20 minutes after tablet ingestion and measurements were made in strict accordance with the manufacturer's instructions.

Test Protocol. Subjects provided a total of four UBTs (on days 0, 7, 21, and 49 of the study; see Fig. 1). They were enrolled on a rolling basis between January and March 2003, starting with a 3-day run-in period during which no cruciferous vegetables were consumed. On days 1 to 21, subjects were given 30 g of raw alfalfa sprouts three times daily as a noncruciferous vegetable supplement to be eaten in place of their normal vegetable-rich diet; no supplements were given on days 22 to 49. To assess adherence to a prescribed diet, questionnaires assessed compliance at each of the 64 meals on days 1 to 21 ($3 \times 21 = 63$, plus one extra meal), with a maximum possible score of 192 points [100% compliance was awarded 3 points, 2 points were given for consumption of between 50% and 100% of the assigned vegetable, and 1 point was given for consumption of less than half of it].

Verification of Compliance. Overnight urine samples were collected immediately before each of the UBTs and aliquots were sent to Baltimore for biochemical verification that subjects were maintaining a crucifer-free diet by using a cyclocondensation assay to detect isothiocyanates and their metabolites (21-25). Briefly, 500 μL aliquots of

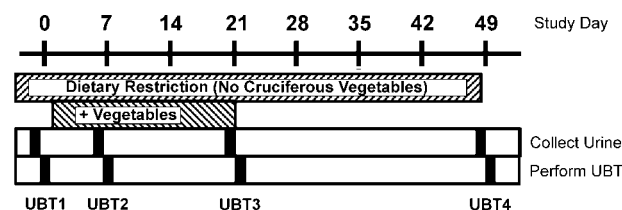


Figure 1. Experimental timeline.

Table 1. Summary statistics of study population (n = 46)

Variable	Mean ± SD	Min	Max
UBT1 (at day 0; % DOB)	39.6 ± 30.7	6.2	161.7
UBT2 (at day 7; % DOB)	35.4 ± 22.6	4.3	102.3
UBT3 (at day 21; % DOB)	33.5 ± 26.5	5.1	137.5
UBT4 (at day 49; % DOB)	37.6 ± 24.5	4.7	98.9
UBT1 (in log scale)	3.402 ± 0.778	1.825	5.086
UBT2 (in log scale)	3.344 ± 0.729	1.459	4.628
UBT3 (in log scale)	3.240 ± 0.759	1.629	4.924
UBT4 (in log scale)	3.365 ± 0.807	1.548	4.594
Urine 1*	1.24 ± 1.37	0.04	6.16
Urine 2	1.76 ± 2.66	0.03	12.30
Urine 3	1.42 ± 2.35	0	13.38
Urine 4	0.90 ± 1.26	0	6.34

NOTE: Mean subject age was 52 years (range, 22-77 years); 59% of subjects were female, 20% of subjects were smokers, and 48% consumed alcohol (methodology for determination of smoking and alcohol consumption status in the text).

*Data are in micromolar dithiocarbamate per person per collection. Urinary dithiocarbamate excretion was calculated on entire overnight urine collection. Note that comparable expected mean values for "enthusiastic" eaters of cruciferous vegetables would be expected to be ~23 μmol per person (22).

urine were added to 500 μL of 500 mmol/L sodium borate buffer (pH 9.25) and 1.0 mL of 20 mmol/L 1,2-benzenedithiol in acetonitrile and incubated for 2 hours at 65°C. After cooling to room temperature and low-speed centrifugation to sediment any precipitates, a 200 μL aliquot of the supernatant fluid was injected automatically (Waters Autosampler, Model 717 Plus, Waters Co., Milford, MA) onto a Partisil 10 ODS-2, 4.6 × 250 mm, reverse phase HPLC column (Whatman, Clifton, NJ) and eluted with 80% methanol/20% water (v/v) at a rate of 2 mL/min. The 1,3-benzodithiole-2-thione formed in this reaction was eluted at ~5 minutes, and the peak was detected by a photodiode array detector set at 365 nm (PDA Model 996, Waters) as described by Ye et al. (23).

Statistical Analysis. Graphical depiction of the 184 (46 × 4) UBT values (Fig. 2) suggested a log-normal distribution as a reasonable probability model for the observed data. Let Y_{it} denote the logarithm of the UBT measurement taken at time t (for $t = 0, 7, 21,$ and 49 days) in the i th individual ($i = 1, 2, \dots, 46$). We denote the within-individuals mean and SD by m_i and S_i , and these were the mean and SD of the four measurements in the i th individual, respectively. To graphically depict the components of variance, we plot m_i versus S_i , so that the overall average of the m_i represents the location of the population, the variability of the m_i represents the between-individuals SD (B), and the average of the S_i represents the estimate of the within-individuals SD (W).

To characterize the components of variance of the four UBT measurements in the 46 individuals comprising the study population, we used standard methods for two-way ANOVA. As a summary of the variance components, we provided the ratio of the within SD to between SD (W/B) and its corresponding within-individuals correlation $1/(1 + (W/B)^2) = B^2/(B^2 + W^2)$. To test whether age, sex, smoking, and alcohol consumption had effects on the components of variance and the mean of the UBT measurements, we used likelihood ratio tests to

compare the two groups by sex, smoking (yes/no), alcohol consumption (yes/no), and age (older/younger than 50 years).

The estimates of W and B provided by our study are key inputs for the planning of clinical trials. In the simplest example of a treatment-plus-placebo clinical trial with premeasurements and one postmeasurement, let X_{0i} and Y_{0i} denote the pre and post values of the i th individual in the placebo group and let X_{1j} and Y_{1j} represent the corresponding values for the j th individual in the treatment group. If the analysis of the trial compares the means of the Y 's, the sample size required for each group to have 80% power to detect the difference between the mean of Y_{1j} (i.e., μ_1) and Y_{0i} (i.e., μ_0) at the 95% confidence level is given by $n = 2(B^2 + W^2)(1.96 + 0.841)^2/(\mu_1 - \mu_0)^2$. However, the strength of the design is to be able to use each individual as his/her own control by comparing the differences $D_{0i} = Y_{0i} - X_{0i}$ and $D_{1j} = Y_{1j} - X_{1j}$. In this case, the sample size formula reduces to $2(2W^2)(1.96 + 0.841)^2/(\Delta_1 - \Delta_0)^2$, where Δ_1 and Δ_0 are the means of the D_1 's and D_0 's, respectively. This formula follows from the fact that $\text{Var } D_{0i} = \text{Var } D_{1j} = (B^2 + W^2) + (B^2 + W^2) - 2\rho(B^2 + W^2) = 2W^2$, since $\rho = B^2/(B^2 + W^2)$.

To guide the planning of a clinical trial, sample size calculations were thus based on the equation: $n = 4(Z_{1-\alpha/2} + Z_{1-\beta})^2(W)^2/D^2$; W was assumed to be the same for both groups (e.g., placebo and treatment) being compared and D is defined as the detectable difference between those two groups' measures.

Statistical analyses were done using both SAS (version 8, 1999; SAS Institute, Inc., Cary, NC) and Stata (version 7, 2001; Stata Corp., College Station, TX). Specific comparisons are described in Results.

The conduct of this trial was approved by the appropriate institutional review boards at both Tsukuba University and Johns Hopkins University.

Results

We assessed the within-subject and between-subject variance in 46 individuals not taking antibiotics, in a target population for a dietary intervention. Four

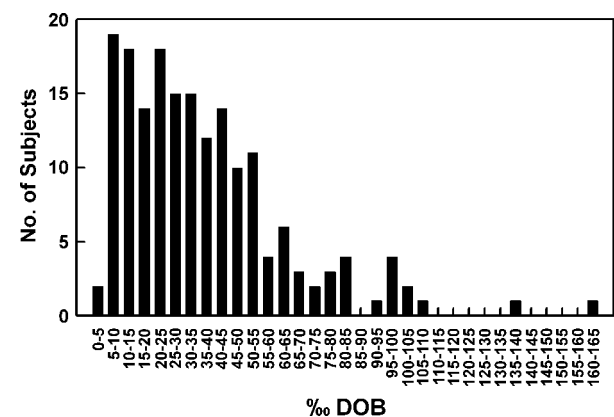


Figure 2. Distribution of 184 UBT values (UBT1-UBT4 for 46 subjects).

measurements of UBT were made over a period of 7 weeks (called UBT1, UBT2, UBT3, and UBT4, respectively). Subjects were administered supplementary green vegetables (alfalfa sprouts) which they consumed for 21 days, beginning after UBT1 and ending at the time of UBT3 (Fig. 1). Average age of the 46 subjects was 52 years (range, 22-77 years) and 27 subjects (59%) were female. Descriptive statistics of their UBT and urine measurements (an index of dietary compliance), are given in Table 1. Raw UBT scores ranged from 4.3‰ to 161.7‰, with an overall mean of 36.5‰. Mean scores for UBTs 1 to 4 were 39.6‰, 35.4‰, 33.5‰, and 37.6‰ DOB, with coefficients of variation of 77.6%, 63.7%, 79.1%, and 65.2%, respectively. In addition, Table 1 presents the descriptive statistics of the UBT measurements in the log scale. This transformation was suggested by the shape of the distribution of scores for the pooled data from UBT1 to UBT4, which was consistent with a log-normal distribution (Fig. 2); thus, the statistical analysis was conducted on log-transformed data.

Table 1 also shows the compliance data. Compliance of subjects with the study protocol (abstention from cruciferous vegetables) was excellent: levels of dithiocarbamates measured in overnight urine collections (biochemical markers of consumption of cruciferous vegetables) averaged between 1 and 2 μmol per person. These values were entirely consistent with self-reported consumption in which scores ranged between 165 and 192 out of 192 possible points (reflecting complete adherence to the dietary protocol over all 64 meals). There were no significant differences overall between any of the four UBT scores when evaluated by paired t tests on log-transformed data, direct comparisons, quartile-parsed data, the sums of differences in UBT values, the sum of the absolute values of the differences, or the sum of absolute values of differences as a fraction of the mean.

Figure 3 depicts the within-individuals means and SDs. The overall mean (depicted with a *large arrow* on the x axis) was 3.34 and the SD of the individual means (i.e., between-individuals SD) was 0.67, which is depicted by two arrows located at +1 SD and -1 SD from the overall mean. Notably, there was a striking difference among subjects in the stability of their UBT measurements as indicated by the location of the *open circles* along the y axis. About half (63 of 138) of the pairs of sequential UBT measurements across all subjects and all times changed by $>10\%$ between measurements. Partitioning the study population according to the quartiles of the within-individuals SDs (i.e., quartiles of the S_i for $1 \leq i \leq 46$), Fig. 4 illustrates the trajectories of UBT (log scale). Figure 4A provides the trajectories of the 11 individuals in the first quartile who had a remarkably stable UBT profile over four measurements spaced across 7 weeks. Coefficients of variations for these subjects were $<21\%$ (Fig. 4A) compared with coefficients of variations between 40% and 80% for the highest quartile (Fig. 4D).

Table 2 shows the means and variance components for the overall group as well as for two strata according to age, sex, smoking, and alcohol consumption. Calculation of omnibus P 's showed that, although there was no difference in means between male and female subjects (P for homogeneity of means = 0.240), the variance

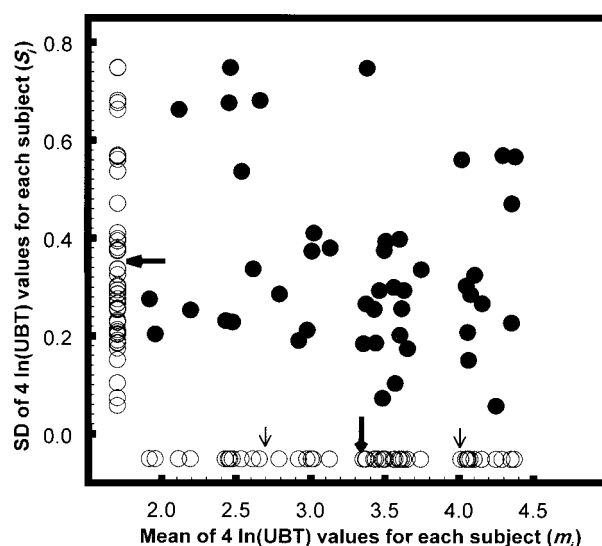


Figure 3. Mean and SD of $\ln(\text{UBT1-UBT4})$ for each of the 46 subjects. The x and y coordinates for each of the 46 means are projected onto their respective axes. *Large arrows*, overall means; *smaller arrows* (x axis), points at -1 SD and $+1$ SD from the overall mean. The overall SD of the $\ln(\text{UBT})$ (0.380) is also equivalent to the coefficient of variation for all (untransformed) UBTs. *Points* mapped to the x -axis reflect $\text{SD}_{\text{between}}$ and those mapped to the y axis reflect $\text{SD}_{\text{within}}$ (see Table 3).

components for female subjects were significantly higher than those of males (P for homogeneity of variances = 0.018; Table 2). Likewise, there was no difference in means between older and younger subjects (P for homogeneity of means = 0.660) but a marginally increased between-individuals variance in older subjects (P for homogeneity of variance = 0.086). Neither smoking nor alcohol consumption had any impact on mean values or variance components.

Outcomes obtained in the present trial (Tsukuba) are contrasted in Table 3 with those in three previously published trials. Using the data from the 46 subjects in the present trial as the basis for projection, we calculate that, to have 80% power to detect a difference (D) of 30% ($D = \log 1.30 = 0.26$) from the mean between treated and control groups, a sample size of $n = 66$ in each arm would be needed to conclude that there was a difference at the 95% confidence level (Table 3).

This sample size falls within the range that one can calculate based on some of the best test/retest data available in the literature (Table 3), but our data are unique in that they permit assessment of the variability for UBT measurements within individuals. Published test/retest data have used two or at best three repeated measurements on an individual. Computation of required sample size based on the lower compared with the upper quartile of variability in our subjects suggests that as few as 12 or as many as 147 subjects, respectively, per treatment arm, would be required for similar power to detect a difference (Table 3). Thus, prescreening of subjects for stable UBT baseline has great bearing on the sample size required.

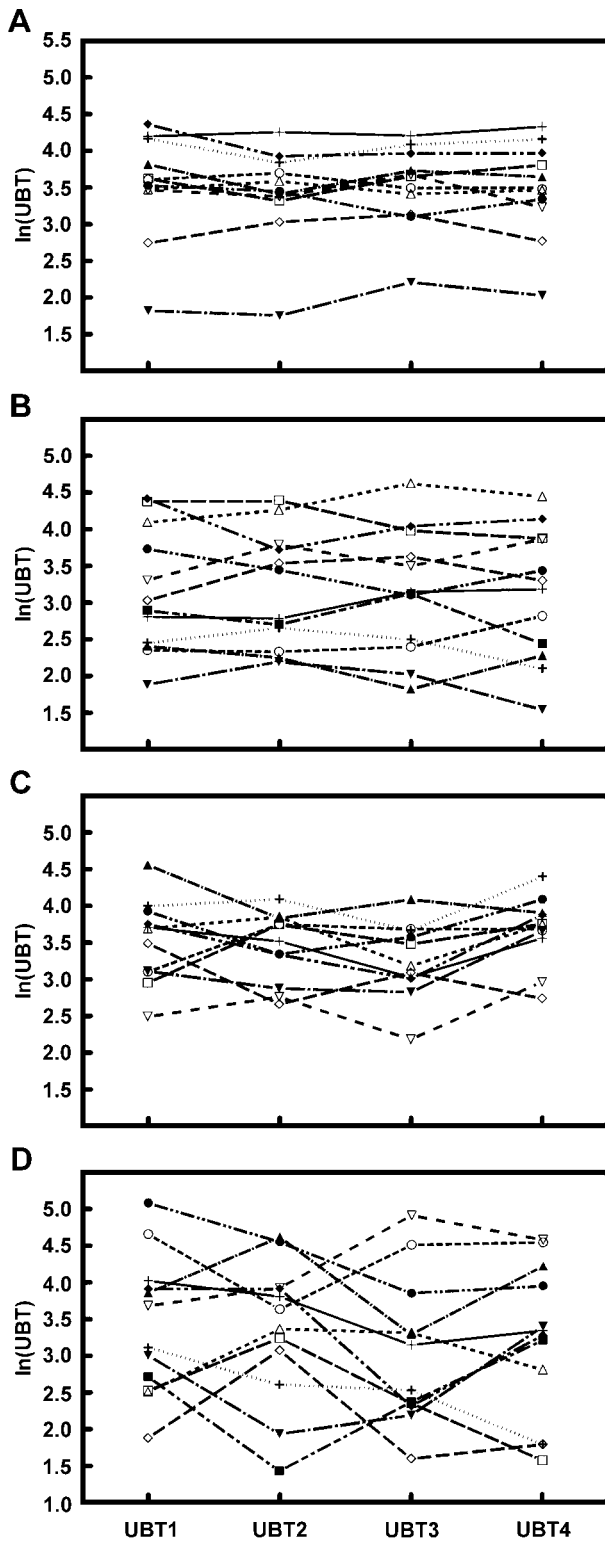


Figure 4. Repeat UBTs (log scale) of the quartiles of subjects with the lowest (A) to highest (D) coefficients of variation for UBT measurement. Mean SD_{within} for each quartile of variation is 0.159 for the lowest quartile (A), 0.254 for the second quartile (B), 0.338 for the third quartile (C), and 0.586 for the fourth quartile (D).

Table 2. Comparison of UBT (log scale) according to age, sex, smoking, and alcohol consumption

	n	Mean	SD_{within}	$SD_{between}$	P for homogeneity	
					Variation	Means
Overall	46	3.338	0.380	0.669		
Age (y)					0.086	0.660
<50	21	3.388	0.323	0.648		
>50	25	3.296	0.422	0.698		
Sex					0.018	0.240
Male	19	3.194	0.319	0.496		
Female	27	3.439	0.417	0.762		
Smoking					0.999	0.675
No	37	3.316	0.380	0.671		
Yes	9	3.426	0.379	0.697		
Alcohol					0.287	0.367
No	24	3.249	0.404	0.742		
Yes	22	3.435	0.351	0.582		

Discussion

The present study provided data to determine the reproducibility of the UBT over time in a population known to be infected with *H. pylori*, with a view to testing the effects of dietary intervention on reducing colonization with these organisms. Although Blaser and Berg (26) have highlighted reasons why universal eradication may not be indicated, it is clear that *H. pylori* eradication in certain high-risk groups will reduce the risk of gastric cancer. Very few dietary interventions have monitored the level of *H. pylori* by the UBT. Notably, those studies comprised very small numbers of subjects and, perhaps not surprisingly, reported lack of effect. For example, the following numbers of subjects were studied: (a) 15 treated subjects and 8 untreated controls with single before and

Table 3. Characterization of the distribution of UBT (log scale) in the present (Tsukuba) and previously published studies

Descriptor	Tsukuba	Logan (38)	Logan (37)	Wang (39)
No. subjects in study	46	20	39	19
Measurements per subject	4	2	2	2
Mean ln(UBT)	3.338	3.234	2.300	2.155
W	0.380	0.279	0.440	0.395
B	0.669	0.321	1.132	1.360
W/B*	0.567	0.870	0.389	0.290
Corr _{within} †	0.76	0.57	0.87	0.92
Projected no. subjects to detect 30% difference				
Based on overall W‡	66	35	88	71
Lowest quartile W§	12			
Highest quartile W	147			

NOTE: UBT values for Logan (37), Logan (38), and Wang (39) were estimated from their published graphs.

*W/B ratio = ratio of SD_{within} to $SD_{between}$ (inter/intra).
 †Corr_{within} = intraindividual correlation $(B^2 / (B^2 + W^2)) = 1 / [1 + (W/B)^2]$.
 ‡Calculation employing 80% power to detect a difference of 30% (i.e., $D = \log 1.30 = 0.2624$) at the 95% confidence level, inferred SDs from published studies, and actual SD from the data presented herein. $n = 4 (Z_{1-\alpha/2} + Z_{1-\beta})^2 (W)^2 / D^2$.
 §Based on mean SD_{within} for the lowest quartile of variability (0.159).
 ||Based on mean SD_{within} for the highest quartile of variability (0.586).

after treatment UBTs (27); (b) 5 treated subjects with two baseline and one or two follow-up UBTs (28); and (c) 5 subjects with only single before and after treatment UBTs (29). Examination of the data in light of the variability we present here leads to the conclusion that their experiments did not have the statistical power to identify an effect. Such underpowering has been highlighted as a critical inadequacy in the conduct of randomized trials (30, 31), even when pilot samples are used in sample size determination (32). Indeed, one of the developers of the UBT assay (P.D. Klein) recently cautioned (19):

The simplicity of the breath test method is seductive if one could distinguish between two groups or two populations on the basis of a difference in the recovery of labeled CO₂, valuable diagnostic information should have been generated. Unfortunately, this belief is often unsupported by external comparisons. . . Good statistical design considerations point to a study size of at least 60 individuals in each category or a total of 120 to 150 comparisons to the standard.

This problem was also recently highlighted by Marshall et al., who questioned whether the UBT was able to assess the intragastric bacterial load even semiquantitatively. His group concluded no effect based on the individual data of nine subjects given UBTs before, during, and after lactoferrin therapy (33).

Our studies showed that there is variability in longitudinal measurements in some subjects. With only a small change in colonization, even in the subjects with the most stable baseline measurements, the number of subjects required would be large. Indeed, in studies of the most commonly examined dietary approach to *H. pylori* treatment, formulations containing *Lactobacillus* sp., demonstration of a significant effect of intervention required a much larger sample size than has been used in most such trials. With a sample size of 326 asymptomatic children, Cruchet et al. (34) have shown a significant effect of daily ingestion of *Lactobacillus johnsonii* LA1 for 4 weeks. Two other studies serve to illustrate the point that development of robust baseline information on the subject pool is critical to conducting a conclusive intervention. In the first study, Sakamoto et al. (35) followed 29 subjects with three UBTs spread over 19 weeks—the first two designed to establish baseline values and the third designed to reveal the effects of *Lactobacillus gasseri* intervention. Although their mean values were virtually identical, close examination of individual subjects' values reveals dramatic fluctuations in many of them during the baseline phase of the study. In the second study, *Lactobacillus casei* treatment of 14 *H. pylori*-colonized subjects (and six controls) was followed with single before and after treatment UBTs (36). Neither of these studies showed convincing results due to low subject number and high interindividual variability.

To be meaningful from a public health perspective, the ability to discriminate between quartiles or quintiles of *H. pylori* density in an infected population is a reasonable objective. The large differences in variance that we observed would make detecting such a difference in the sample sizes used in most dietary eradication studies problematic. There are several potential explanations for these large differences, including the underlying pathology of *H. pylori* infection in which there may be physiologic perturbations that are detected by the biomarker

(UBT) yet not reported as a symptom by the patient (e.g., mild gastritis) or detected on examination. Colonization with multiple *H. pylori* strains may result in competition for niches within the stomach or in evolution to colonize those niches more efficiently. None of the additional data collected in this study (e.g., smoking status, age, and sex) could account for the variation we saw.

We have thus presented a longitudinal series of four UBT measurements made over a 7-week period in 46 free-living, asymptomatic, *H. pylori*-infected individuals and have generated statistical descriptors for that population. We are not aware that such test/retest data are publicly available. This information now permits the design of dietary interventions in this population in which absolute eradication of *H. pylori* is not a required end point. Although evidence indicates that it would be feasible to preselect individuals with stable baseline UBT measurements by performing three or four temporally spaced UBTs and thus reduce the required sample size, much caution is required. Such preselection may well carry with it a bias toward tolerance to *H. pylori* colonization, so that the expected change due to an intervention in these individuals may be of diminished magnitude, reducing the perceived advantage of such an approach.

Acknowledgments

We thank Dr. Masayuki Yamamoto for facilitating this study and for providing seed funds to assist in its startup; Murakami Noen (Hiroshima, Japan) for producing and delivering all of the sprouts prescribed for ingestion in this study, for donating two UBIT-IR300 machines to the Tsukuba University Medical School, and for providing all of the stable isotope (¹³C-urea) that was required for this study; the AICR and the Brassica Foundation for funding; Naoko Itagaki and Mamoru Amagai for helping with the collection of urine samples and for assisting with UBT and data collection; Katherine Stephenson for assisting with laboratory analyses; Shintaro Kamo for growing and distributing the sprouts; Lisa Jacobson for providing initial statistical consultation; Pamela Talalay for providing editorial assistance; the members of Citizen Oriented Health and Medicine Network, many of whom volunteered to participate in this study; and the numerous volunteers whose interest in helping with *H. pylori* research motivated them to participate in this study.

References

1. IARC. Globocan 2000: cancer incidence, mortality and prevalence worldwide. Version 1.0. Cancer Base No. 5. Lyon: IARC Press; 2001.
2. Pisani P, Parkin DM, Muñoz N, Ferlay J. Cancer and infection: estimates of the attributable fraction in 1990. *Cancer Epidemiol Biomarkers Prev* 1997;6:387–400.
3. Parkin DM, Pisani P, Ferlay J. Global cancer statistics. *CA Cancer J Clin* 1999;49:33–64.
4. Fahey JW, Haristoy X, Dolan PM, et al. Sulforaphane inhibits extracellular, intracellular, and antibiotic-resistant strains of *Helicobacter pylori* and prevents benzo[a]pyrene-induced stomach tumors. *Proc Natl Acad Sci U S A* 2002;99:7610–5.
5. Haristoy X, Angoi-Duprez K, Duprez A, Lozniewski A. Efficacy of sulforaphane in eradicating *Helicobacter pylori* in human gastric xenografts implanted in nude mice. *Antimicrob Agents Chemother* 2003;47:3982–4.
6. Correa P, Fontham ET, Bravo JC, et al. Chemoprevention of gastric dysplasia: randomized trial of antioxidant supplements and anti-*Helicobacter pylori* therapy. *J Natl Cancer Inst* 2000;92:1881–8.
7. Turner NJ, Thomson BM, Shaw IC. Bioactive isoflavones in functional foods: the importance of gut microflora on bioavailability. *Nutr Rev* 2003;61:204–13.
8. Labenz J, Stolte M, Aygen S, Hennemann O, Bertrams J, Borsch G. Qualitative and semiquantitative invasive and noninvasive diagnosis

- of *Helicobacter pylori* colonization of gastric mucosa. *Z Gastroenterol* 1993;31:437–43.
9. Labenz J, Barsch G, Peitz U, et al. Validity of a novel biopsy urease test (HUT) and a simplified ¹³C-urea breath test for diagnosis of *Helicobacter pylori* infection and estimation of the severity of gastritis. *Digestion* 1996;57:391–7.
 10. Bazzoli F, Zagari RM, Pozzato P, et al. ¹³C-Urea breath test to quantify *H. pylori* colonization of gastric mucosa an association with severity of inflammation. *Gastroenterology* 1994;106:A48.
 11. Ellenrieder V, Glasbrenner B, Stoffels C, et al. Qualitative and semi-quantitative value of a modified ¹³C-urea breath test for identification of *Helicobacter pylori* infection. *Eur J Gastroenterol Hepatol* 1997; 9:1085–9.
 12. Perri F, Clemente R, Pastore M, et al. The ¹³C-urea breath test as a predictor of intragastric bacterial load and severity of *Helicobacter pylori* gastritis. *Scand J Clin Lab Invest* 1998;58:19–27.
 13. Hilker E, Domschke W, Stoll R. ¹³C-urea breath test detection of *Helicobacter pylori* and its correlation with endoscopic and histologic findings. *J Physiol Pharmacol* 1996;47:79–90.
 14. Chen X, Haruma K, Kamada T, et al. Factors that affect results of the ¹³C urea breath test in Japanese patients. *Helicobacter* 2000;5:98–103.
 15. Kobayashi D, Eishi Y, Ohkusa T, et al. Gastric mucosal density of *Helicobacter pylori* estimated by real-time PCR compared with results of urea breath test and histological grading. *J Med Microbiol* 2002;51: 305–11.
 16. Leodolter A, Dominguez-Munoz JE, Von Arnim U, Malfertheiner P. Citric acid or orange juice for the ¹³C-urea breath test: the impact of pH and gastric emptying. *Aliment Pharmacol Ther* 1999;13:1057–62.
 17. Graham DY, Klein PD, Evans DJ Jr, et al. *Campylobacter pylori* detected noninvasively by the ¹³C-urea breath test. *Lancet* 1987;1: 1174–7.
 18. Kato C, Yamaguchi T, Yusa K, et al. ¹³C-urea breath test. *Nippon Rinsho* 2001;59:267–71.
 19. Klein PD. ¹³C breath tests: visions and realities. *J Nutr* 2001;131: S1637–42.
 20. Savarino V, Vigneri S, Celle G. The ¹³C urea breath test in the diagnosis of *Helicobacter pylori* infection. *Gut* 1999;45 Suppl 1:118–22.
 21. Shapiro TA, Fahey JW, Wade KL, Stephenson KK, Talalay P. Human metabolism and excretion of cancer chemoprotective glucosinolates and isothiocyanates of cruciferous vegetables. *Cancer Epidemiol Biomarkers Prev* 1998;7:1091–100.
 22. Shapiro TA, Fahey JW, Wade KL, Stephenson KK, Talalay P. Chemoprotective glucosinolates and isothiocyanates of broccoli sprouts: metabolism and excretion in humans. *Cancer Epidemiol Biomarkers Prev* 2001;10:501–8.
 23. Ye L, Dinkova-Kostova AT, Wade KL, Zhang Y, Shapiro TA, Talalay P. Quantitative determination of dithiocarbamates in human plasma, serum, erythrocytes and urine: pharmacokinetics of broccoli sprout isothiocyanates in humans. *Clin Chim Acta* 2002;316:43–53.
 24. Zhang Y, Cho CG, Posner GH, Talalay P. Spectroscopic quantitation of organic isothiocyanates by cyclocondensation with vicinal dithiols. *Anal Biochem* 1992;205:100–7.
 25. Zhang Y, Wade KL, Prestera T, Talalay P. Quantitative determination of isothiocyanates, dithiocarbamates, carbon disulfide, and related thiocarbonyl compounds by cyclocondensation with 1,2-benzene-dithiol. *Anal Biochem* 1996;239:160–7.
 26. Blaser MJ, Berg DE. *Helicobacter pylori* genetic diversity and risk of human disease. *J Clin Invest* 2001;107:767–73.
 27. Nir Y, Potasman I, Stermer E, Tabak M, Neeman I. Controlled trial of the effect of cinnamon extract on *Helicobacter pylori*. *Helicobacter* 2000;5:94–7.
 28. McNulty CA, Wilson MP, Havinga W, Johnston B, O'Gara EA, Maslin DJ. A pilot study to determine the effectiveness of garlic oil capsules in the treatment of dyspeptic patients with *Helicobacter pylori*. *Helicobacter* 2001;6:249–53.
 29. Opekun AR, Opekun JL, Graham DY. Broccoli (sulforaphane glucosinolate) for the treatment of *H. pylori* infection in humans [abstract]. *Am J Gastroenterol* 2003;M1502.
 30. Moher D, Dulberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. *JAMA* 1994;272: 122–4.
 31. Vickers AJ. Underpowering in randomized trials reporting a sample size calculation. *J Clin Epidemiol* 2003;56:717–20.
 32. Browne RH. On the use of a pilot sample for sample size determination. *Stat Med* 1995;14:1933–40.
 33. Guttner Y, Windsor HM, Viiala CH, Marshall BJ. Human recombinant lactoferrin is ineffective in the treatment of human *Helicobacter pylori* infection. *Aliment Pharmacol Ther* 2003;17:125–9.
 34. Cruchet S, Obregon MC, Salazar G, Diaz E, Gotteland M. Effect of the ingestion of a dietary product containing *Lactobacillus johnsonii* La1 on *Helicobacter pylori* colonization in children. *Nutrition* 2003;19: 716–21.
 35. Sakamoto I, Igarashi M, Kimura K, et al. Suppressive effect of *Lactobacillus gasseri* OLL 2716 (LG21) on *Helicobacter pylori* infection in humans. *J Antimicrob Chemother* 2001;47:709–10.
 36. Cats A, Kuipers EJ, Bosschaert MA, Pot RG, Vandenbroucke-Grauls CM, Kusters JG. Effect of frequent consumption of a *Lactobacillus casei*-containing milk drink in *Helicobacter pylori*-colonized subjects. *Aliment Pharmacol Ther* 2003;17:429–35.
 37. Logan RP, Polson RJ, Misiewicz JJ, et al. Simplified single sample ¹³Carbon urea breath test for *Helicobacter pylori*: comparison with histology, culture, and ELISA serology. *Gut* 1991;32:1461–4.
 38. Logan RPH, Dill S, Bauer FE, et al. The European ¹³C-urea breath test for the detection of *Helicobacter pylori*. *Eur J Gastroenterol Hepatol* 1991;3:915–21.
 39. Wang WM, Lee SC, Ding HJ, et al. Quantification of *Helicobacter pylori* infection: Simple and rapid ¹³C-urea breath test in Taiwan. *J Gastroenterol* 1998;33:330–5.