Use of Breath Hydrogen and Methane as Markers of Colonic Fermentation in Epidemiological Studies: Variability in Excretion

Lynne R. Wilkens, Loic Le Marchand, Patricia Harwood, and Robert V. Cooney

Abstract
Breath hydrogen and methane are specific end products of colonic fermentation, a process which may play a protective role against colon cancer. To assess the possibility of using these markers in epidemiological studies, we characterized the intra- and intersubject variability of breath hydrogen and methane excretion over 15 consecutive days among 32 men and women of various ethnic backgrounds (16 Asians, 8 Caucasians, 8 Hawaiians). Participants were asked to collect four end-expiratory samples each day, which we had shown previously would optimally characterize daily hydrogen excretion. There was substantial within-subject variation in breath hydrogen over the study, although breath methane levels were more constant over time. We found that about 4 days of measurement for breath hydrogen and 1 day of measurement for breath methane are required to correctly characterize individuals according to their long-term excretion of these gases. This was true for Asians and non-Asians. Although breath methane appears to be more practical to measure, it is a less sensitive marker of colonic fermentation than breath hydrogen. Whereas all subjects excreted hydrogen, only 28% of the subjects excreted methane, and methane excretor status of a few participants varied during the study. Because the breath test is noninvasive and reliable, we tested the multiple day collection protocol among colon cancer patients and controls and found it to be well accepted. We conclude that it is practical to measure breath hydrogen and methane in large epidemiological studies conducted at the individual level. The potential use for these markers is discussed.

Introduction
Fermentation of dietary and other substrates in the large bowel by anaerobic bacteria has been postulated to play a protective role against colon cancer (1). Possible mechanisms include a bulking and diluting effect, a lowering of the colonic pH, and the production of butyrate, which has been shown to act as a differentiating agent and a tumor growth inhibitor (1).

Colonic fermentation is controlled to a great extent by the availability in the colon of dietary carbohydrates (resistant starch and fiber), which depends not only on the composition of the diet, but also on factors that cannot be assessed by questionnaire, such as ripeness, coarseness, temperature of the food, extent of processing and cooking, etc. (1, 2). Although these factors suggest many possibilities for increasing the fermentability of an individual’s diet, they make the fermentation-colon cancer association difficult to test in epidemiological studies.

Measuring hydrogen and methane in the breath may provide useful and practical biomarkers of colonic fermentation. Hydrogen and methane are end products of fermentation that are absorbed into the portal bloodstream and excreted via expired air. Levels of hydrogen in expired breath have been shown to correlate very well (r = 0.9) with concentrations produced in the large intestine (3). These breath gases have been used mainly as measures of colonic fermentation to assess carbohydrate malabsorption in clinical practice (4) and to study populations at different risks for colon cancer (5).

Although data for free living subjects are scant, breath hydrogen excretion appears to exhibit day-to-day variation and a circadian pattern related to carbohydrate ingestion (4, 6, 7). However, breath methane levels appear to be relatively constant over time (4, 7). To assess the possibility of using these markers in epidemiological studies conducted at the individual level, we characterized the inter- and intra-subject variability of breath hydrogen and methane excretion over 15 consecutive days among 32 individuals. Because the 5 main ethnic groups in Hawaii differ considerably in their carbohydrate food sources, we included each of these groups as a way to investigate the generalizability of our results.

Materials and Methods
Participants were identified among the population controls from one of our ongoing case-control studies of diet and cancer (8). None of the subjects had taken laxatives or antibiotics in the previous 4 weeks.

The subjects were asked to collect four end-expiratory breath samples every day for 15 consecutive days. The collection times were upon rising in the morning, between noon and 1:00 p.m., at 6:00 p.m., and at 10:00 p.m. We had determined in an earlier investigation that collecting samples...
at these times would optimally characterize the daily excretion levels (7). Samples collected at three of these times were also adequate (7).

End-expiratory breath samples were collected into sealed, aluminum-lined bags (GaSampoers, Quintron Instrument Company, Milwaukee, WI). The collection procedure was demonstrated to the subjects who then personally collected the samples at the specified times. Smokers were instructed not to smoke during the 15 minutes before sample collection. All participants were asked to brush their teeth before collecting the first sample of the day. Breath samples collected in this manner have been found to be stable for 14 days when stored at room temperature in the collection bags (7).

The breath samples were analyzed for hydrogen and methane using a Model DP Quintron Microlyzer and for carbon dioxide with a Model 24 Quintron Alveolyzen. End alveolar CO2 partial pressure was assumed to be constant, and the breath hydrogen and methane concentrations were corrected for possible atmospheric contamination by normalizing values to the measured CO2 concentration (2, 6, 9). Instruments were calibrated with a standard gas mixture from Quintron, containing 93 ppm hydrogen, 22 ppm methane, and 4.62% CO2. To test the reproducibility of our measurement process, we had volunteers collect more than one sample at a time, resulting in 49 multiple samples from 18 subjects. The average coefficient of variance across replicates was 8.7% for log H2 and 0.6% for log CH4, indicating excellent reproducibility. Samples where the CO2 measurements were very low (<1%) were excluded from this analysis. Methane values less than 2 ppm were set to zero because the instrument has a 2-ppm detection limit for CH4.

Thirty-two days of measurement were excluded from the analysis because, for these days, there were fewer than the three usable samples required to adequately characterize hydrogen excretion during a 24-h period (7). This resulted in 442 days being available for analysis, with a mean number of days/subject of 13.8 (range of 9–15) and a mean number of samples/day of 3.8 (range of 3–4). Methane production was defined as a concentration of CH4 in breath over 2 ppm, and a “true methane producer” was defined as a person excreting methane on 90% or more of the observation days.

In the analysis of the data, all corrected breath measurements were log transformed as log(x + 1) so that the distributions would approximate normality. For each day, the mean breath hydrogen and methane levels were computed. The daily means were then modeled with a one-way random effects ANOVA to obtain estimates of the inter- and intrasubject variability. Age and sex were adjusted for by including them in the ANOVA models as fixed covariates (10). ICCs (11) were computed as

\[ ICC = \frac{s_g^2}{s_g^2 + s_w^2} \]

where \( s_g^2 \) and \( s_w^2 \) are the observed within- and between-subject variances computed by the ANOVA model.

The number of days of measurement \( D \) necessary to rank individuals correctly according to their long-term hydrogen and methane excretion was computed using the formula proposed by Nelson et al. (12):

\[ D = \frac{r^2}{1 - r^2} \times \frac{s_w^2}{\frac{k}{k} F_{0.025, a,v}} \]

\[ D = \frac{r^2}{1 - r^2} \times \frac{k}{k} F_{0.025, a,v} \]

where \( r \) is the unobservable correlation between the observed and the true mean excretion values for an individual over a day. A value of \( r = 0.90 \) was chosen because at least 80% of the subjects in the extreme thirds of the distribution would be correctly classified, and less than 1% would be grossly misclassified in the incorrect extreme thirds (12). 95% confidence intervals were computed about \( D \) as:

\[ r^2 \times \frac{s_g^2 + s_w^2}{k} F_{0.025, a,v} \]

\[ r^2 \times \frac{k}{k} F_{0.025, a,v} \]

where \( F_{a,v} \) is the critical value of the \( F \) distribution, \( u = n - 1 \), where \( n \) is the number of subjects, \( v = \sum n_i - 1 \), where \( m_i \) is the number of days of observation for subject \( i \), and \( k \) is \( 1 - \frac{1}{n} \) (11).

**Results**

Table 1 presents characteristics of the participants. Among the 13 men and 19 women studied, the age range was 47–85, with a mean of 69 years. Ten major ethnic groups in Hawaii were represented: Japanese (n = 11), Caucasian (n = 8), part Hawaiian (n = 8), Filipino (n = 3), and Chinese (n = 2). There were seven current smokers in the study. Asians were much less likely to drink alcohol than were Caucasians and Hawaiians. Nutrients were adjusted for caloric intake by the method of residuals (13) so that we could compare the diet composition between the ethnic groups. Non-Asians consumed more of their starch from breads and tubers than did Asians, who consumed most of their starch from rice. Caucasians consumed more cereal than the other ethnic groups. The participants in our 15-day study were older and, by design, included proportionally more Hawaiians; however, the smoking and eating patterns were similar between the subsample and all of the controls from the parent colorectal cancer study. All 32 subjects excreted breath hydrogen during the totality of the study period, whereas only 8–10 (an average of 9, or 28.4%) excreted breath methane on any particular day.

Table 2 presents the within- and between-subject variances for the daily hydrogen means and the number of days required to correctly characterize individuals, separately for all subjects, Asians (n = 16 Japanese, Filipinos, and Chinese) and non-Asians (n = 16 Caucasians and Hawaiians). As expected, the intrasubject variability was substantial, resulting in ratios of within- to between-subject variation close to 1, and 4 days of measurement (95% confidence interval: 2–7) being required to correctly rank subjects. The ICCs around 0.5 indicate that half of the total variability is a result of within-subject variability. The variances were somewhat smaller in Asians than in Caucasians and Hawaiians. However, the ratio of intra- to intersubject variation and thus the required number of days of measurement were similar between the two groups. The results were little changed by adjusting for age and sex (data not shown).

Table 2 also presents the within- and between-subject variances for the daily methane means and the number of days required, separately for all subjects, Asians and non-Asians. For methane, the intrasubject variability is small, re-
Table 1  Characteristics of participants in 15-day breath hydrogen and methane study in Hawaii

<table>
<thead>
<tr>
<th>Variables</th>
<th>Total (n = 32)</th>
<th>Asians (n = 16)</th>
<th>Non-Asians (n = 16)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage male</td>
<td></td>
<td></td>
<td></td>
<td>0.28</td>
</tr>
<tr>
<td>Mean age (years)</td>
<td>68.6</td>
<td>68.5</td>
<td>68.8</td>
<td>0.94</td>
</tr>
<tr>
<td>Mean years of school</td>
<td>10.9</td>
<td>10.7</td>
<td>11.2</td>
<td>0.71</td>
</tr>
<tr>
<td>Percentage ever smoked cigarettes</td>
<td>62.5</td>
<td>68.8</td>
<td>56.2</td>
<td>0.53</td>
</tr>
<tr>
<td>Percentage currently smoking</td>
<td>21.9</td>
<td>25.0</td>
<td>18.8</td>
<td>0.33</td>
</tr>
<tr>
<td>Mean lifetime pack-years</td>
<td>13.1</td>
<td>12.7</td>
<td>11.4</td>
<td>0.92</td>
</tr>
<tr>
<td>Mean ethanol intake (g/day)</td>
<td>7.8</td>
<td>0.1</td>
<td>15.4</td>
<td>0.01</td>
</tr>
<tr>
<td>Mean calories (Kcal/day)</td>
<td>1886</td>
<td>1668</td>
<td>2104</td>
<td>0.07</td>
</tr>
<tr>
<td>Mean daily nutrient intake, adjusted for calories</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carbohydrate (g)</td>
<td>247</td>
<td>250</td>
<td>244</td>
<td>0.69</td>
</tr>
<tr>
<td>Starch (g)</td>
<td>106</td>
<td>113</td>
<td>99</td>
<td>0.14</td>
</tr>
<tr>
<td>Starch (g) from</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice</td>
<td>38</td>
<td>55</td>
<td>22</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Breads</td>
<td>32</td>
<td>30</td>
<td>34</td>
<td>0.56</td>
</tr>
<tr>
<td>Tubers (potatoes &amp; taro)</td>
<td>9</td>
<td>4</td>
<td>14</td>
<td>0.07</td>
</tr>
<tr>
<td>Cereals</td>
<td>8</td>
<td>7</td>
<td>10</td>
<td>0.34</td>
</tr>
<tr>
<td>Pasta</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>0.74</td>
</tr>
<tr>
<td>Dietary fiber (g) from</td>
<td>16</td>
<td>16</td>
<td>17</td>
<td>0.76</td>
</tr>
<tr>
<td>Dietary fiber (g) from</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cereal</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>0.45</td>
</tr>
<tr>
<td>Fruit</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>0.40</td>
</tr>
<tr>
<td>Vegetables</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.33</td>
</tr>
</tbody>
</table>

a  Eleven Japanese, three Filipinos, and two Chinese.
b Eight Caucasians and eight Hawaiians.
c P-values from t tests for continuous variables and x2 tests for categorical variables, comparing Asians and non-Asians.
d Adjusted for caloric intake by the method of residuals (13).

Table 2  Within- and between-subject variation over 15 days for breath hydrogen and methane, and the number of days necessary to correctly characterize individuals on their long-term excretion

<table>
<thead>
<tr>
<th></th>
<th>Meana</th>
<th>Within-Subject Variation (W)</th>
<th>Between-Subject Variation (B)</th>
<th>Ratio W/B</th>
<th>ICCb</th>
<th>Number of Daysc</th>
<th>95% Confidence Interval for D</th>
</tr>
</thead>
<tbody>
<tr>
<td>All participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log (H2CO2 + 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (n = 32)</td>
<td>1.1857</td>
<td>0.0613</td>
<td>0.0649</td>
<td>0.94</td>
<td>0.51</td>
<td>4.02</td>
<td>2.18-6.68</td>
</tr>
<tr>
<td>Asians (n = 16)</td>
<td>1.2171</td>
<td>0.0544</td>
<td>0.0584</td>
<td>0.93</td>
<td>0.52</td>
<td>3.97</td>
<td>1.57-7.99</td>
</tr>
<tr>
<td>Non-Asians (n = 16)</td>
<td>1.1516</td>
<td>0.0676</td>
<td>0.0757</td>
<td>0.89</td>
<td>0.53</td>
<td>3.81</td>
<td>1.50-7.70</td>
</tr>
<tr>
<td>Log (CH4CO2 + 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (n = 32)</td>
<td>0.4323</td>
<td>0.0495</td>
<td>0.4978</td>
<td>0.10</td>
<td>0.91</td>
<td>0.42</td>
<td>0.24-0.68</td>
</tr>
<tr>
<td>Asians (n = 16)</td>
<td>0.0920</td>
<td>0.0034</td>
<td>0.0864</td>
<td>0.04</td>
<td>0.96</td>
<td>0.17</td>
<td>0.07-0.32</td>
</tr>
<tr>
<td>Non-Asians (n = 16)</td>
<td>0.8014</td>
<td>0.0801</td>
<td>0.7205</td>
<td>0.11</td>
<td>0.90</td>
<td>0.47</td>
<td>0.19-0.91</td>
</tr>
<tr>
<td>Among methane producers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log (CH4CO2 + 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (n = 8)</td>
<td>1.5468</td>
<td>0.0683</td>
<td>0.3017</td>
<td>0.23</td>
<td>0.82</td>
<td>0.96</td>
<td>0.23-2.39</td>
</tr>
<tr>
<td>Non-Asians (n = 7)</td>
<td>1.1451</td>
<td>0.0733</td>
<td>0.3237</td>
<td>0.23</td>
<td>0.82</td>
<td>0.96</td>
<td>0.19-2.54</td>
</tr>
</tbody>
</table>

a  The actual mean levels can be obtained by exponentiating the value, subtracting 1, and multiplying by 5, the commonly accepted percentage concentration for the partial pressure of CO2 in alveolar air.
b  ICC = B/(W+B).
c  The number of days required to characterize individuals (see “Materials and Methods” for formula).
d A methane producer excretes methane over 2 ppm on 90% or more of the observation days.

resulting in ratios of within- to between-subject variation close to zero, and only 1 day of measurement required to correctly rank subjects. The ICCs were around 0.9. The intra- and intersubject variability was markedly smaller for Asians than non-Asians, as was the ratio of the variances. However, only one day of measurement was required for both groups. Adjustment for age and sex had little effect (data not shown).

Because the distribution of daily breath methane, even after log transformation, was skewed to the left, we investigated variability of methane levels among producers only, and of methane production status. Table 2 shows that among methane producers the within-subject variability was small for methane compared with between-subject variability, and that one day of measurement was sufficient to correctly characterize the levels among producers.

Classification tables of true methane excretor status and daily status were produced for each of the 15 days. The percentage of subjects correctly classified based on a single day of measurement ranged from 93.8% to 100.0%, with an average of 96.5%. The percent of false positives ranged from 0.0% to 12.5%, with an average of 1.7%, while the percent of false negatives ranged from 0.0% to 8.3%, with an average...
of 4.2%. Thus, the agreement using one day of observation was very good and did not improve greatly by observing two days.

Discussion

In addition to SCFAs (acetic, propionic, and butyric acids), lactic acid, and energy, the end products of colonic fermentation include methane, hydrogen, and carbon dioxide. Proportions of the former two gases are absorbed into the portal blood stream and excreted unchanged by the lungs. The formation of hydrogen and methane is unique to anaerobic bacteria, and no mammalian cell is known to produce these gases. Thus, their excretion in the breath should be a valid and specific indicator of bacterial fermentation in the colon, as suggested by the marked reduction in breath hydrogen levels brought about by oral antibiotic administration. Fecal SCFA concentrations may not reflect SCFA production because these acids are rapidly absorbed. Butyrate and propionate are metabolized by the colonic mucosa and liver, respectively. Acetate and lactic acid can be measured in peripheral blood, but because of their endogenous production from other sources, they are not specific. Fermentation also affects pH in the large bowel, especially in the cecum, because SCFAs are the principal anions in the colon. However, the use of fecal pH in large epidemiological studies appears problematic.

In determining the feasibility of using breath hydrogen and methane as markers of colonic fermentation in epidemiological studies, we have already demonstrated that at least four samples for breath hydrogen and two samples for breath methane are needed daily to optimally characterize daily excretion. The present data show that because of the variation in carbohydrate intake in the usual diet, about 4 days of measurement for breath hydrogen and 1 day of measurement for breath methane are required to correctly characterize individuals according to their long-term excretion of these gases. Thus, in epidemiological studies where assessment of an individual’s long-term breath hydrogen excretion is desired, 16 measurements would be required, 4 measurements/day for 4 days. In contrast, only two samples on a single day would be needed to characterize an individual on breath methane excretion. In studies where these gases are assessed at the group level, 4 measurements of breath hydrogen and 2 measurements of breath methane during a single day would be sufficient because random within-subject variability does not affect group means. However, the substantial within-subject variability will increase the total variation, reducing the power of studies at the group level.

From our data, breath methane would appear to be a more practical marker than breath hydrogen for use in large epidemiological studies, especially conducted at the individual level. However, breath methane is not a very sensitive marker of colonic fermentation, as only 33–48% of whites who are secretors of blood group substances (ABH antigen) are methane producers. Other substrates could be more rate limiting.

The higher rate of methane production in colon cancer patients may also result from delayed transit. Partial colonic obstruction by the tumor may enhance anaerobic conditions favoring growth of methanogenic bacteria. Indeed, Melcher et al. have shown that a high cereal fiber intake concomitantly decreases transit time and reduces breath methane excretion in methane producers. Thus, although the relationship between colonic fermentation and breath methane production is complex, breath methane still constitutes an interesting bowel function parameter that can be measured in epidemiological studies.

Although somewhat less practical, breath hydrogen appears to be a promising marker for assessing colonic fermentation. Because of its high sensitivity and specificity, breath hydrogen has been used widely by nutritionists to investigate carbohydrate digestibility. For example, change in the level of breath hydrogen after a test meal was used by Levitt et al. to estimate that 20% of the carbohydrate from baked beans and 7–10% of that from wheat, oats, potatoes, and corn reached the large intestine undigested, but only 0.9% of that from rice. These findings are similar to more direct measurements obtained in gastrointestinal intubation studies and studies of ileostomy patients. This comparison not only supports the validity of the breath test for epidemiological studies, but suggests that breath hydrogen also could be used as a compliance marker in interventions attempting to increase colonic fermentation.

Information on the variability of an exposure, both between and within subjects, is required to design an epidemiological study. Substantial, random, within-person variation can severely attenuate risk estimates in epidemiological studies. The β-coefficient for the exposure has been shown to be attenuated by a factor of 0.91, where n_k is the number of replicates. For studies collecting only one replicate, the attenuation factor is the ICC. Therefore, in studies collecting 1 day of information, the β-coefficient for breath hydrogen would be reduced by a factor of 0.51, and that for breath methane would be reduced by a factor of 0.91. In a study analyzing the mean for breath hydrogen and methane across 4 days of information, the attenuation factors would be 0.81 and 0.98, respectively. In the presence of considerable intrasubject variability, the investigator can either collect repeated measurements of the exposure, or execute a reliability study and use the information on within- and between-subject variability to correct the results from the main study.
For breath hydrogen, it seems that the relatively high number of repeated samples required for the correct measurement is not a major limitation to its use in large studies; the test is non-invasive, the participants can easily collect the samples themselves, the samples can be stored for up to 2 weeks at room temperature, and the biochemical analysis is simple, reliable, and inexpensive. We already have demonstrated the high acceptability of this test among population controls interviewed in an ongoing case-control study conducted by our group (30). Among the 331 consecutive individuals offered the breath test (four samples at specified times on 1 day), only 10% refused. To test the feasibility of the four-sample collection protocol over multiple days in a population-based sample, we conducted a pilot study among 31 consecutive colon cancer patients and their matched population controls interviewed as part of the same case-control study. The participation rate for the parent study was 70%. The participation rate for the additional breath collection protocol was 87%, indicating a good acceptability for collecting samples on multiple days.

Although we studied breath hydrogen and methane excretion over a relatively long period (15 days), we may have missed variation that can exist over a longer time period, such as that resulting from seasonal variation in diet. We think this is unlikely because there are few seasonal changes in food availability in Hawaii. Our findings on breath H2 and CH4 variability did not differ greatly across the ethnic groups studied, who, in Hawaii, demonstrate marked differences in dietary habits. This suggests that our findings may be generalizable to other populations.

In conclusion, we have demonstrated that it is practical to use breath hydrogen and methane as markers of colonic fermentation in large epidemiological studies. However, these studies need to be limited to subjects with a normal bowel environment. For example, these markers could be used in cross-sectional and prospective studies of healthy individuals and in case-control studies of adenomatous polyps.

Acknowledgments
The authors thank the subjects and interviewers for their participation. The technical assistance of Terry Shimamoto is also gratefully acknowledged.

References
Use of breath hydrogen and methane as markers of colonic fermentation in epidemiological studies: variability in excretion.

L R Wilkens, L Le Marchand, P Harwood, et al.


Updated version
Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/3/2/149

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.