The water load test: observations from healthy controls and patients with functional dyspepsia

Michael P. Jones, Seth Hoffman, Dhiren Shah, Ketan Patel, and Christine C. Ebert

Division of Gastroenterology and Hepatology, Northwestern University Medical School, Chicago, Illinois 60611-2908

Submitted 26 August 2002; accepted in final form 13 January 2003

The water load test: observations from healthy controls and patients with functional dyspepsia. Am J Physiol Gastrointest Liver Physiol 284: G896–G904, 2003. First published 15 January 2003; 10.1152/ajpgi.00361.2002.—Gastric sensation and accommodation are studied by barostat, but this is invasive. The drink test is noninvasive and may provide similar information. We evaluated relationships between drink test, gastric function, symptoms, and psychiatric distress. Controls (73) and functional dyspeptics (FD) (92) were studied using a 5-min water load test (WL5), gastric emptying, and electrogastrography (EGG). Symptoms, quality of life, and psychiatric distress were measured using standardized measures. Controls underwent test-retest of WL5 and comparison of WL5 with 100 ml/min water-based drink test (WL100) or nutrient drink. Controls, FD, and gastroparetics estimated drinking capacity before WL5 using a visual analog scale. WL5 correlated with WL100 (r = 0.7929) but not nutrient drink test (r = 0.1995). WL5 was significantly less in FD than controls, and abnormal WL5 was seen in 46%. In FD, volume to fullness inversely correlated with symptom severity (r = −0.29; P = 0.0154) and WL5 produced more symptoms, particularly nausea. Gastric function was not different between FD with normal or abnormal WL5. Symptoms and psychiatric distress were similar between normal and abnormal WL5 groups, but the abnormal group had significantly poorer quality of life. Controls and gastroparetics had good correlation of estimated and ingested volumes, but FD did not. Versus FD with normal WL5 capacity, FD with impaired drinking capacity have normal gastric function and similar symptoms but poorer quality of life. FD are less able to predict drinking capacity. These data suggest that WL5 identifies FD with intact gastric function but abnormal visceral perception.

FUNCTIONAL DYSPEPSIA (FD) is a common, often difficult problem in clinical practice. The pathophysiology of the disorder is likely heterogeneous with a number of mechanisms proposed. These include gastric and small intestinal motor abnormalities, visceral hypersensitivity, and psychological factors. Recent investigations have focused on impaired meal-induced fundic relaxation as well as visceral hypersensitivity (14, 21, 25, 29, 33, 35). These studies have generally been performed using a barostat with an intragastric balloon. Although the utility of the barostat to study mechanical properties of the proximal stomach is well established, its utility as an investigational or clinically applicable device in the investigation of FD is questionable. The technique is cumbersome, not widely available, and unpleasant.

The investigation of patients with functional or neuromuscular digestive disorders, whether from a clinical or research point of view, will benefit from minimally invasive, patient-friendly techniques that are sensitive, reproducible, and can be applied across large study populations. The water load or drink test is such a tool that was developed as a noninvasive method to assess gastric sensation. The test is easily performed and well tolerated, and preliminary studies have demonstrated reproducibility in healthy adults (11, 15, 34). FD patients have been shown to have impaired drinking capacity, although there is substantial overlap with normal subjects (13, 15, 34).

Initial studies of the drink test suggested that it correlated well with barostat studies (1). Boeckxstaens et al. (3), however, recently showed that a drink test using either water or a caloric drink was a less sensitive predictor of impaired accommodation or visceral hypersensitivity than the barostat. It thus appears that, although the results of drink tests are due at least in part to gastric sensation and motor function, the test may also be influenced by sensory and psychological factors. The present study was undertaken to further evaluate the performance characteristics of the drink test in both healthy controls and patients with FD.

The primary aim of this study was to determine the utility and clinical correlates of the water load test in FD. To achieve this, we first established normal values and test characteristics for the 5-min water load test (WL5) and compared it with other methods of “stomach loading.” We also administered WL5 to patients with FD and correlated the volume required to produce fullness with gastric emptying, electrogastrography (EGG), symptom severity, quality of life, and psychopathological variables.

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

Subjects. Healthy volunteers were recruited by advertisement. Exclusion criteria included prior gastrointestinal surgery (except appendectomy), presence of any digestive complaints, or the use of any medications to treat digestive disorders or medications that might alter digestive motility.

Patients with FD were consecutively enrolled from the Gastrointestinal Physiology Laboratory after being referred from the outpatient gastroenterology clinic at Northwestern University. Although enrollment from the Physiology Laboratory was consecutive, it was not possible to know the total pool of dyspeptics from which these subjects were selected. Patients were eligible if they met Rome II criteria for FD (31). Patients were excluded if they had dominant complaints of heartburn, prior digestive surgery other than cholecystectomy or appendectomy, or were taking medications known or suspected of altering digestive motility. All subjects were required to have a normal upper gastrointestinal endoscopy, with further evaluation performed as indicated by the evaluating gastroenterologist. Patients with FD were further categorized using Rome II guidelines as ulcer-like dyspepsia or dysmotility-like dyspepsia. Patients were considered to have ulcer-like dyspepsia if the predominant symptom was pain centered in the upper abdomen. Patients were considered to have dysmotility-like dyspepsia if their predominant symptom was a nonpainful sensation characterized or associated with upper abdominal fullness, early satiety, bloating, or nausea. Subjects were allowed to remain on acid suppressive therapy because proton pump inhibitors do not alter gastric emptying, EGG, or the volume of water required to produce satiety (12).

Subjects were also excluded if they were infected with Helicobacter pylori. H. pylori status was determined by either biopsy for histology or rapid urease testing, whole blood

EGG. EGG was performed after a 12-h overnight fast using a Cycle3 EGG machine (3CPM, Crystal Bay, NV). After skin preparation, three standard Ag-AgCl electrodes (Sensor-

Medics, Anaheim, CA) were positioned on the anterior abdominal surface. One electrode was positioned midway between the xiphoid and umbilicus. A second electrode was placed just below the costal margin in the right midclavicular line. Electrodes were connected to a BMA-931 Bioamplifier with frequency cutoffs set at 0.016 Hz and 0.25 Hz (0.95 and 15 cpm). Output from the bioamplifier was simultaneously sent to a strip chart recorder as well as being digitized at a sampling frequency of 4,267 Hz and stored in a PC for further analysis. A pneumo-

belt was placed around the patient’s upper abdomen and connected to a pressure transducer to record respiratory rate. All recordings were performed in a quiet room with the subject instructed not to talk and to remain as still as possible during recording to minimize motion artifacts.

After signal stabilization, baseline tracing was performed for 15 min. This was followed by the drink test. After completion of the drink test, recording was continued for an additional 30 min.

The raw EGG tracing was inspected, and segments with artifact were excluded. The digitized signal was Fourier transformed, and a running spectral analysis (RSA) was generated using a 4-min epoch with 75% overlap as previously described by Koch et al. (16). EGGS were classified as normal, abnormal, or technically inadequate based on the raw signal, the RSA, and the percentage of EGG signal power in various bandwidths. An EGG was considered normal if there was a dominant frequency between 2.5 and 4 cpm for >70% of the study period. The study was considered abnormal if 2- to 4-cpm activity was seen for >70% of the time. Activity of 0.95–2.4 cpm was considered bradygastria, and activity of 4–15 cpm was considered tachygastria. Tracings were considered technically inadequate if there was excessive artifact, persistently low amplitude signal, or an RSA demonstrating numerous, chaotic peaks.

Gastric emptying. Determination of solid-phase gastric emptying was performed using the 13C-Spirulina platensis gastric emptying breath test (Meretek Diagnostics, Houston, TX). We employed a similar method to that reported by Lee et al. (19). The S. platensis used in this study was grown in a closed hydroponic chamber purged with pure 13CO2 and was uniformly labeled with 13C. As a result, the level of 13CO2 in the breath was increased from normal background levels of ~1% to ~99%. Gastric emptying was measured after an overnight fast. Subjects ingested a standard test meal that consisted of a 60-g rye roll (160 kcal), 30 g cream cheese (90 kcal), and 120 ml white grape juice (80 kcal). The rye roll contained 200 mg of 13C-S. platensis. Breath samples were collected at baseline and at 45, 90, 105, and 150 min after ingestion of the test meal.

full (LM). The LM was administered using an identical method to WL5. WL100 has been reported by Boeckxstaens and colleagues in several studies on proximal gastric function (2, 3). LM is a method similar to that used by both Boeckx-

staens et al. (2, 3) and Tack et al. (29) but did differ, however, in that we used Boost for the test meal (Mead Johnson Nutritionalis, Evansville, IN). Boost contains 1.1 kcal/ml and is 75% carbohydrate, 15% fat, and 16% protein, compared with Nutridrink (N.V. Nutricia, Zoetermeer, Netherlands), which contains 1.5 kcal/ml and is 39% fat. Symptoms of nausea, fullness, and bloating were rated using a four-point Likert scale (none, mild, moderate, and severe). Symptoms were assessed at baseline and at 10, 20, and 30 min after completion of the drink test.

EGG. EGG was performed after a 12-h overnight fast using a Cycle3 EGG machine (3CPM, Crystal Bay, NV). After skin preparation, three standard Ag-AgCl electrodes (Sensor-

Medics, Anaheim, CA) were positioned on the anterior abdominal surface. One electrode was positioned midway between the xiphoid and umbilicus. A second electrode was placed just below the costal margin in the right midclavicular line. Electrodes were connected to a BMA-931 Bioamplifier with frequency cutoffs set at 0.016 Hz and 0.25 Hz (0.95 and 15 cpm). Output from the bioamplifier was simultaneously sent to a strip chart recorder as well as being digitized at a sampling frequency of 4,267 Hz and stored in a PC for further analysis. A pneumo-

belt was placed around the patient’s upper abdomen and connected to a pressure transducer to record respiratory rate. All recordings were performed in a quiet room with the subject instructed not to talk and to remain as still as possible during recording to minimize motion artifacts.

After signal stabilization, baseline tracing was performed for 15 min. This was followed by the drink test. After completion of the drink test, recording was continued for an additional 30 min.

The raw EGG tracing was inspected, and segments with artifact were excluded. The digitized signal was Fourier transformed, and a running spectral analysis (RSA) was generated using a 4-min epoch with 75% overlap as previously described by Koch et al. (16). EGGS were classified as normal, abnormal, or technically inadequate based on the raw signal, the RSA, and the percentage of EGG signal power in various bandwidths. An EGG was considered normal if there was a dominant frequency between 2.5 and 4 cpm for >70% of the study period. The study was considered abnormal if 2- to 4-cpm activity was seen for >70% of the time. Activity of 0.95–2.4 cpm was considered bradygastria, and activity of 4–15 cpm was considered tachygastria. Tracings were considered technically inadequate if there was excessive artifact, persistently low amplitude signal, or an RSA demonstrating numerous, chaotic peaks.

Gastric emptying. Determination of solid-phase gastric emptying was performed using the 13C-Spirulina platensis gastric emptying breath test (Meretek Diagnostics, Houston, TX). We employed a similar method to that reported by Lee et al. (19). The S. platensis used in this study was grown in a closed hydroponic chamber purged with pure 13CO2 and was uniformly labeled with 13C. As a result, the level of 13CO2 in the breath was increased from normal background levels of ~1% to ~99%. Gastric emptying was measured after an overnight fast. Subjects ingested a standard test meal that consisted of a 60-g rye roll (160 kcal), 30 g cream cheese (90 kcal), and 120 ml white grape juice (80 kcal). The rye roll contained 200 mg of 13C-S. platensis. Breath samples were collected at baseline and at 45, 90, 105, and 150 min after ingestion of the test meal.
$^{13}$C breath enrichment was determined using isotope ratio mass spectrometry performed using an automated breath $^{13}$C analyzer (PDZ Europa; Cheshire, UK). $^{13}$CO$_2$ in a breath sample was calculated as the percentage of $^{13}$CO$_2$ in total CO$_2$ above background $^{13}$CO$_2$ concentrations in nature. Results were expressed as both the percentage of $^{13}$CO$_2$ recovery and the cumulative recovery over the study period. CO$_2$ production was calculated as the percentage of $^{13}$CO$_2$ in total body surface area as has been previously determined (8). Body surface area was calculated from existing height-weight nomograms.

Determination of the quantity of $^{13}$CO$_2$ in breath per unit time was performed using delta over baseline (DOB): $^{13}$C$_{90} =$ DOB $\times$ 0.0112372 $\times$ CO$_2$ production. The constant in this equation is the isotopic abundance of Pee Dee Belemnite, which is a limestone standard. CO$_2$ production was corrected for age and gender as determined by Schofield (26).

Calculation of half-time for solid emptying ($T_{1/2}$) and the lag phase ($T_{lag}$) were made using the mathematical model derived by Lee et al. (19). $T_{1/2} = 1/LP_{lag}$, where $LP_{lag} = 0.0097 + 0.0021(13C_{105}) - 0.0012(13C_{150})$. Determination of $T_{lag}$ was made using $T_{lag} = 1/LP_{lag}$, where $LP_{lag} = 0.250 + 0.0063(13C_{90}) - 0.0032(13C_{150})$. $LP_{lag}$ and $LP_{lag}$ are the linear predictors, and $13C_{90}$, $13C_{105}$, and $13C_{150}$ are the excreted concentrations of $^{13}$C ($\mu$mol/min) at 90, 105, and 150 min after ingestion of the meal test.

Questionnaires. Subjects were asked to complete several self-report measures pertaining to symptomatology, quality of life, and psychopathologic traits. FD severity was evaluated using the Nepean Dyspepsia Index (NDI). This is a validated, disease-specific measure that evaluates symptoms related to FD over the 2-wk before administration (30, 32). Raw scores were converted into percentage scores for comparative purposes. Ability to predict the volume of water required to produce fullness was determined in a subset of subjects. Before the water load test, subjects estimated the volume they could ingest on a 100-mm visual analog scale with the endpoints “almost none” and “a very large amount.”

Quality of life was assessed using two general measures, the SF-36 and the Psychological General Well Being Index (PGWB). The SF-36 is a widely used and well-studied measure that assesses symptom impact on eight subscales that are broadly grouped into mental and physical well being (7, 24, 36). Although a nonspecific measure, the SF-36 has been used in a variety of digestive disorders including FD (20). The PGWB is an index to measure a person’s subjective well-being. The person self-reports on 22 items that are indicators of the affective states of anxiety, depressed mood, sense of positive well-being, self-control, general health, and vitality (6).

Psychiatric distress was measured using the SCL-90. The SCL-90 is a self-reporting, clinical symptom rating scale consisting of 90 questions. Responses indicate symptoms associated with nine psychiatric constructs. These constructs are somatization, obsessive-compulsive behavior, feelings of inadequacy or inferiority (interpersonal sensitivity), depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism (5).

Study protocol. All studies were performed after a 12-h overnight fast. Informed consent was obtained before evaluation for all subjects. Healthy subjects and patients meeting Rome II criteria for FD completed the drink testing using WL5. A subset of healthy subjects underwent WL5 on a second occasion at least 2 wk after the initial testing session. A subset of these subjects also underwent WL5 on a third occasion at least 2 mo after the initial testing session. A second subset of controls underwent, in addition to WL5, WL100 and LM. Subjects also underwent gastric emptying and EGG and completed the NDI, SF-36, PGWB, and SCL-90.

Statistical analysis. Data were expressed as means ± SD. Normality of all data sets was determined using the Kolmogorov-Smirnov test. For normally distributed data, correlations were calculated using the Pearson method, and differences between groups were determined using paired or unpaired t-tests. In several of the symptom comparisons, all values in the control group were zero. In these instances, comparisons were made using a one-sample t-test for the FD group compared with a hypothetical mean of zero. Correlations between non-normally distributed data sets were determined using the Spearman method, whereas differences between median group scores were determined using the Mann-Whitney test. Statistical significance was set at $P < 0.05$. Statistical calculations were made using GraphPad Prism version 3.00 for Windows (GraphPad Software, San Diego, CA).

RESULTS

Normal values for WL5. Seventy-three healthy subjects were evaluated. This group included 29 males and 44 females with a mean age of 29 ± 5 yr. The mean volume required to produce fullness was 648 ± 204 ml. Males (703 ± 217 ml) drank more than females (611 ± 188 ml), but the difference did not reach statistical significance ($t = 1.907; P = 0.0605$). These data are shown in Fig. 1. No healthy subject consumed <300 ml of water.

Significant correlations also existed between the volume to produce fullness and age ($r = 0.3536; P = 0.0070$) and height ($r = 0.2625; P = 0.0486$) but not BMI ($r = 0.1377; P = 0.3069$). Gastric emptying studies were performed in 26 subjects, and EGG was performed in 41 subjects. There was no correlation between volume to fullness and $T_{lag}$ ($r = -0.1532; P = 0.4549$) or $T_{1/2}$ ($r = -0.1489; P = 0.4679$). Abnormal EGGs were seen in 4 of 41 healthy subjects, all of whom manifested bradygastria. No differences existed between subjects with normal or abnormal EGGs and the volume required to produce fullness ($t = 0.004; P = 0.9965$).

In healthy subjects, there was no significant correlation between volume to produce fullness and NDI ($r = \ldots$).
Reproducibility of WL5. Forty-three healthy subjects underwent WL5 on two occasions separated by at least 2 wk, and 20 of these subjects were studied on a third session at least 2 mo after the initial test. For the 20 subjects assessed on 3 occasions, volumes to fullness were slightly but not significantly lower at subsequent sessions (570 ± 199 vs. 540 ± 154 vs. 515 ± 146; \( P = 0.3684 \)). Volume to fullness between the first and second session (WL51 and WL52) was highly correlated (\( r = 0.7729; P < 0.0001 \)). The mean difference between WL51 and WL52 was 33 ± 170 ml. Correlation between WL51 and the third measurement (WL53) was poor (\( r = 0.3263; P = 0.1693 \)). The mean difference between WL51 and WL53 was 55 ± 204 ml.

Comparison of WL5, WL100, and LM. Eighteen healthy subjects (9 female/9 male) underwent both WL5 and WL100, and 15 of these subjects also underwent LM. Testing by each method was separated by at least 2 wk. Results are shown in Fig. 3. The two water-based methods were highly correlated (\( r = 0.7929; P < 0.0001 \)), whereas the correlation between WL5 and LM was poor (\( r = 0.1995; P = 0.4760 \)). The mean volume to fullness for WL100 was 1128 ± 355 ml compared with 889 ± 270 ml for WL5 (\( P = 0.0002 \)). Mean volume for LM (688 ± 187 ml) was significantly lower than either WL100 (\( P = 0.0003 \)) or WL5 (\( P = 0.0360 \)).

FD. Eighty-seven FD patients were evaluated. The FD population was slightly older than the control population (36 ± 14 vs. 29 ± 5 yr; \( t = 3.624; P = 0.004 \)). The group contained 21 males and 66 females, which was a greater representation of females than the control group (76 vs. 60%; \( P = 0.0405 \)). Sixty-nine of the FD subjects were considered motility-like with dominant symptoms of nausea, bloating, or early satiety. Eighteen subjects were considered ulcer-like because they had a dominant symptom of upper abdominal pain. FD subjects required significantly less water to produce fullness than did controls (333 ± 235 ml vs. 648 ± 204 ml; \( t = 8.953; P < 0.0001 \)). No difference in volume to fullness was seen between ulcer-like dyspeptics (357 ± 231 ml) and dysmotility like dyspeptics (325 ± 237 ml; \( P = 0.6126 \)). Males and females with FD required significantly less water to produce fullness than did healthy subjects of the same sex (Fig. 4).
FD subjects were more symptomatic during the study than were healthy subjects (Fig. 5). Symptoms at baseline and at all three postingestion periods were greater in the FD subjects than in controls (nausea, \( P < 0.0001 \); fullness, \( P < 0.0009 \); bloating, \( P < 0.006 \)).

For controls there was no correlation between ingested volume and symptoms of nausea, fullness, or bloating at any time point. For FD subjects, significant negative correlations existed between WL5 volume and baseline symptoms of nausea \( r_s = -0.3748; P = 0.0005 \) and fullness \( r_s = -0.2465; P = 0.0256 \) but not bloating. Post-WL5 symptoms of nausea demonstrated significant negative correlations at 10 \( r_s = -0.4272; P < 0.0001 \); 20 \( r_s = -0.3822; P = 0.0004 \), and 30 min \( r_s = -0.3834; P = 0.0007 \) after ingestion. Significant correlations were not seen for postprandial symptoms of fullness and bloating or for volume to fullness.

EGG was performed in 84 of 87 subjects and was abnormal in 36 of 84 (43%). The abnormalities seen consisted of bradygastria in 16, bradyarrhythmia in 12, and diffuse arrhythmia in 8. FD patients with an abnormal EGG reached fullness at significantly greater volumes than did patients with normal EGGs (416 ± 279 vs. 265 ± 171 ml; \( t = 3.056; P = 0.0030 \)). Gastric emptying studies were performed in 42 subjects. Volume to fullness was not correlated with \( T_{1/2} \) \( r = 0.2624; P = 0.0932 \) but was correlated with \( T_{1ag} \) \( r = 0.3336; P = 0.0330 \).

Compared with healthy subjects, FD subjects had significantly greater symptom severity as measured by NDI \( t = 12.03; P < 0.0001 \). As shown in Fig. 6, there is also significantly reduced quality of life as measured by both PGWB and SF-36 \( P < 0.0001 \) for both. These differences were seen for all subscale scores of the SF-36. FD subjects also displayed stronger responses to the psychological queries of the SCL-90. Again, the differences between the two groups were seen with all subscales of the SCL-90 but were most striking for the categories of somatization and depression.

Among the dyspeptics, there was a modest, negative correlation between volume to fullness and NDI score \( r = -0.2906; P = 0.0154 \). WL5 was not correlated with total PGWB score or any subscale score other than general health \( r = 0.3140; P = 0.0064 \). For SF-36, significant correlations existed on the physical functioning \( r = 0.2394; P = 0.04 \) and social functioning \( r = 0.2837; P = 0.0143 \) subscales. No correlation was seen with total SCL-90 score or any subscales other than interpersonal sensitivity \( r = 0.2810; P = 0.0185 \) and paranoid ideation \( r = 0.2936; P = 0.0136 \).

Because no healthy control consumed <300 ml of water, we compared functional dyspeptics with WL5 < 300 ml with WL5 ≥ 300 ml (Table 1). Forty-one functional dyspeptics had WL5 < 300 ml compared with 46 subjects with WL5 ≥ 300 ml. The groups did not differ with respect to sex, although there were more females in the WL5 < 300 ml group (83% vs. 71%; \( P = 0.2209 \)). The groups likewise did not differ in age \( t = 0.2655; \)
completed the unitless visual analog scale estimate of drinking capacity. The correlation between visual analog scale score and actual ingested volume was 0.2060 \((P = 0.2351)\). Controls, however, were able to reasonably predict drinking capacity \((n = 22; r = 0.58; P = 0.0051)\). To better understand this observation, we also asked a group of patients with documented diabetic \((n = 16)\) or idiopathic \((n = 3)\) gastroparesis and a group of 13 patients with GERD to complete the visual analog scale assessment and drink test. Gastroparesis was defined as scintigraphic evidence of delayed gastric emptying along with documented gastric stasis as well as a dominant symptom of vomiting. GERD was defined as a dominant symptom of heartburn with endoscopy demonstrating hiatal hernia and/or esophagitis and a clinical response to acid suppressive therapy. The volume to fullness for both gastroparesis and GERD patients was less than for controls \((288 \pm 197\) and \(406 \pm 142\) ml, respectively), but they were also able to accurately predict drinking capacity. For gastroparesis the correlation was \(r = 0.56 (P = 0.0124)\), and for GERD the correlation was \(r = 0.74 (P = 0.0041)\).

DISCUSSION

In the present study we have demonstrated that, in healthy controls, WL5 is comparable with WL100 used by Boeckxstaens and colleagues \((2, 3)\). Our WL5 values for both controls and dyspeptics agree quite well with those reported by Koch et al. \((15)\). An advantage of WL5 over WL100 is that it can be administered in less time. With an average ingested volume of 1,128 ml administered at 100 ml/min, the WL100 method requires slightly more than 10 min to complete, which is twice as long as WL5. WL5 results in controls may not be repeatable over time. Test-retest values for healthy controls studied at least 2 wk apart were highly correlated, although volumes consumed at the second session were somewhat lower than initial volumes. Correlations, although volumes consumed at the second session were somewhat lower than initial volumes. Correlations, although volumes consumed at the second session were somewhat lower than initial volumes. Correlations, although volumes consumed at the second session were somewhat lower than initial volumes.

**Table 1. Characteristics of functional dyspeptics achieving satiety at <300 ml or >300 ml**

<table>
<thead>
<tr>
<th>M/F</th>
<th>WL5 &lt; 300 ml</th>
<th>WL5 ≥ 300 ml</th>
<th>(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>35 ± 15</td>
<td>36 ± 12</td>
<td>0.7912</td>
</tr>
<tr>
<td>EG5, % normal</td>
<td>70</td>
<td>44</td>
<td>0.0279</td>
</tr>
<tr>
<td>T lag, min</td>
<td>37 ± 7</td>
<td>45 ± 23</td>
<td>0.1241</td>
</tr>
<tr>
<td>T 1/2, min</td>
<td>90 ± 15</td>
<td>99 ± 20</td>
<td>0.1061</td>
</tr>
<tr>
<td>NDI, %</td>
<td>50 ± 15</td>
<td>44 ± 16</td>
<td>0.1092</td>
</tr>
<tr>
<td>PGWB</td>
<td>44 ± 22</td>
<td>59 ± 22</td>
<td>0.0039</td>
</tr>
<tr>
<td>SF-36 physical composite score</td>
<td>33 ± 12</td>
<td>39 ± 14</td>
<td>0.0468</td>
</tr>
<tr>
<td>SF-36 mental composite score</td>
<td>31 ± 11</td>
<td>35 ± 10</td>
<td>0.1075</td>
</tr>
<tr>
<td>SCL-90</td>
<td>64 ± 41</td>
<td>58 ± 47</td>
<td>0.5547</td>
</tr>
</tbody>
</table>

Values are means ± SE. WL5, 5-min water load test; EG5, electrogastrogram; T lag, lag phase for solid emptying; T 1/2, half-time for solid emptying; NDI, Nepean Dyspepsia Index; PGWB, Psychological General Well Being Index. Those reporting satiety at <300 ml of water had greater symptom severity and greater impairment in quality of life but less derangement of gastric function.
relation between initial WL5 volumes and volumes at a third session at least 2 mo later were not correlated, however. These data suggest that drink tests are prone to the intrasubject variation seen with other measures of gastric function, particularly with respect to handling of liquids (17, 23).

In contrast to Boeckxstaens et al. (2, 3) who found a significant correlation between the Nutridrink test and the WL100 method, we did not find a correlation between WL5 and LM. Certainly the use of caloric-based drinks introduces a variety of additional factors influencing satiety, including caloric composition, osmolarity, and palatability of the test meal. Although in certain situations the use of caloric meals might be useful, water would seem to offer a modality that restricts satiety determinants to gastric distension, visceral sensitivity, and psychological determinants of fullness.

No healthy control consumed <300 ml of water. We therefore defined an abnormal drink test as the ingestion of <300 ml using the WL5 method. With the use of this definition, 47% of FD subjects had impaired drinking capacity. This is consistent with observations made by other investigators (3, 15, 34). No correlation was found in healthy subjects or dyspeptics with respect to drink test volume and T1/2 for solid-phase gastric emptying. A modest but significant correlation was seen in dyspeptics for drink test volume and Tlag, but Tlag values did not differ significantly between WL5 < 300 ml and WL5 > 300 ml groups.

An abnormal EGG was seen in 45% of FD subjects, which is consistent with previously published observations (15, 18, 22). Significantly fewer FD subjects with abnormal drink tests also had abnormal EGGs. The EGG and gastric emptying data would suggest that an abnormal drink test in FD patients is not a marker for disturbed gastric emptying or myoelectrical activity.

Additionally, recent studies using a barostat have shown a poor correlation between drink test volume and threshold to discomfort or fundic accommodation (3). This is not surprising. Barostat testing most commonly examines only the proximal stomach, whereas drink tests distend both the proximal and distal stomach (9). Since gastric emptying begins within minutes of ingestion, the proximal small bowel is also likely stimulated by the drink test (10). In addition to proximal stomach hypersensitivity, duodenal and antral hypersensitivity have also been demonstrated in FD subjects (4, 27). Because gut perception is influenced by spatial summation, the volume of distribution may be a critical determinant of subject response (28).

FD subjects were significantly more symptomatic than healthy subjects both at baseline and after the drink test. No correlations existed between volume to satiety and symptoms for controls or for FD subjects with respect to bloating and fullness. FD subjects displayed a significant negative correlation between volume to satiety and scores for nausea, suggesting that patients with dominant symptoms of nausea were less likely to drink.

FD subjects also had greater symptom severity as measured by NDI and lower quality of life on both SF-36 and PGWB measures. Psychopathology was significantly more prevalent among FD subjects, and significant differences existed between dyspeptics and controls for all SCL-90 subscales.

Among FD subjects, however, the drink test was not particularly helpful in delineating meaningful clinical subgroups. Across the FD population, there was little correlation between volume to fullness and measures of quality of life or psychopathology. There was a significant negative correlation between NDI scores and drink test volume, but this was modest. For SF-36, significant, modest positive correlations existed on the
physical and social functioning subscales. Modest positive correlations were also seen on the interpersonal sensitivity and paranoid ideation subscales of the SCL-90. The significance of this observation is not clear, and this study was not designed to accurately identify personality characteristics of test subjects.

In contrast to correlations of WL5 with measured variables across the FD population, comparisons of FD subjects with normal and abnormal drink tests disclosed several significant differences between these groups. FD subjects with normal or abnormal drink tests did not differ significantly with respect to NDI scores, but the abnormal drink test group did have significantly lower PGWB scores as well as significantly lower SF-36 subscale scores for social functioning and mental health. SCL-90 scores were not different between the two groups. These data suggest that an abnormal drink test identifies subjects with greater impairment in quality of life.

The poor correlation between estimated and observed drinking capacity among dyspepsics was a surprising finding. If the FD subjects had a primary gastric abnormality that limited drinking capacity, one would logically expect them to not only drink less but to also anticipate the ability to drink less, as was seen in the patients with gastroparesis. These data support the concept of impaired visceral perception in dyspepsics. The FD subjects in this study seem uncertain as to exactly what their stomachs will allow them to accomplish.

In summary, WL5 is a simple measure that correlates well with other water-based drink tests but not with calorically-based methods. Intrasubject variability, at least in healthy controls, appears to be significant for WL5. An abnormal water load test, defined as the ingestion of <300 ml of water, is not a marker for abnormal gastric emptying or gastric myoelectrical activity. FD subjects with an abnormal drink test would appear to represent a clinical subset more likely to report nausea and with greater psychological distress and impairment of quality of life. Dyspepsics also appear less able to accurately predict drinking behavior compared with controls, gastroparetics, and patients with GERD. Whether this disturbance is caused by primary visceral pathology or more central mechanisms cannot be elucidated from these data.

Whether the water load test will emerge as a useful clinical tool remains to be determined. At present, insufficient evidence exists to support the routine use of the water load test in clinical practice. Available literature suggests that the test is not a surrogate for gastric emptying, accommodation, or visceral sensitivity. The water load test, originally intended to be a noninvasive measure of gastric physiology, may in fact be a better measure of personality characteristics and quality of life.

REFERENCES


22. Parkman HP, Miller MA, Trate D, Knight LC, Urbain JL, Maurer AH, and Fisher RS. Electrogastrography and gastric


