2017

Frequency of Chronic Gastrointestinal Distress in Runners: Validity and Reliability of a Retrospective Questionnaire

Patrick B. Wilson
Old Dominion University, pbwilson@odu.edu

Follow this and additional works at: https://digitalcommons.odu.edu/hms_fac_pubs

Part of the Nutrition Commons, and the Sports Sciences Commons

Repository Citation
Wilson, Patrick B., "Frequency of Chronic Gastrointestinal Distress in Runners: Validity and Reliability of a Retrospective Questionnaire" (2017). Human Movement Sciences Faculty Publications. 44.
https://digitalcommons.odu.edu/hms_fac_pubs/44

Original Publication Citation

This Article is brought to you for free and open access by the Human Movement Sciences at ODU Digital Commons. It has been accepted for inclusion in Human Movement Sciences Faculty Publications by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.
Frequency of Chronic Gastrointestinal Distress in Runners: Validity and Reliability of a Retrospective Questionnaire

Patrick B. Wilson

Gastrointestinal (GI) symptoms may affect up to 90% of competitors during endurance races. Studies have typically assessed GI symptoms retrospectively or only over an acute timeframe, and information on the validity and reliability of the questionnaires employed is lacking. This investigation aimed to estimate the frequency of GI distress experienced by runners over 30 days and to establish the validity and reliability of a retrospective GI symptom questionnaire. Runners (70 men, 75 women) recorded GI symptoms with a prospective journal for 30 days. Retrospective GI symptom data were then collected after the 30-day period on two occasions within one week. GI symptoms were rated on a 0–10 scale. Descriptive statistics for GI symptoms are reported as medians (interquartile ranges) because of nonnormal distributions. Men and women experienced at least one GI symptom on 84.0% (59.8–95.1%) and 78.3% (50.0–95.2%) of runs, respectively. Moderate-to-severe GI symptoms (score of ≥5) were experienced on 13.8% (6.7–37.3%) and 21.7% (5.3–41.2%) of runs for men and women. Spearman’s rho correlations between journal ratings and retrospective questionnaire ratings ranged from 0.47 to 0.82 (all \( p < .001 \)), although they were highest when journal ratings were quantified as mean 30-day values (all rho ≥ 0.59). Reliability of the retrospective questionnaire ratings was high (rho = 0.78–0.92; \( p < .001 \)). In comparison with tracking GI symptoms with a daily journal, retrospective questionnaires seem to offer a convenient and reasonably valid and reliable method of quantifying GI symptoms over 30 days.

Keywords: endurance, exercise, nutrition, sport
reflect peak discomfort. However, this phenomenon has not been yet been observed in the context of GI distress experienced during exercise.

Given the wide variance in methods employed to assess GI symptoms in past literature and the lack of information on the validity and reliability of the questionnaires used, the purpose of this study was threefold: 1) establish the validity and reliability of a retrospective GI symptom questionnaire; 2) longitudinally estimate the frequency of GI distress experienced by endurance runners over a prolonged time period (30 days); and 3) determine whether retrospective GI symptom reports are more representative of peak or average GI symptoms.

Methods

General Design

This study employed an observational design with a combination of prospective and retrospective data collection. Running sessions and GI symptoms were recorded prospectively over approximately 30 days using a journal (described in detail later). At the end of this 30-day period, an initial retrospective questionnaire collected data about GI symptoms experienced over the 30-day period. Validity of this retrospective questionnaire was assessed by comparing it to the prospective journal, which was considered the reference method. A second retrospective questionnaire was completed within a week of the first retrospective GI questionnaire to establish the reliability of the retrospective GI symptom reports.

Participants

Runners were recruited through contact with running groups and directors of endurance running races. Inclusion criteria required runners be age 18–65 years, be currently running ≥20 miles per week, and completed at least one 10-mile run over the past month. Informed consent was provided by all runners using the web-based Qualtrics software (Provo, Utah, USA). Two-hundred fifty-four runners initially provided consent to participate, of which 145 runners were included in the final analysis (Table 1). Eighty-seven runners were lost to follow-up, did not complete the prospective journal, or sustained an injury during training, while 22 runners were excluded due to a reported GI condition (e.g., IBS, Celiac disease) or incomplete data on either the first or second retrospective GI questionnaire (Figure 1).

Prospective Journal

Runners were emailed a journal that was used to track running sessions over the course of approximately 30 days (mean = 29.9 ± 2.0 days). Information collected in each journal included the date and time of each run, run duration, and the overall level of exertion for each run using Borg’s 6–20 Rating of Perceived Exertion (RPE) scale (Borg, 1990). The average duration and RPE score for each participant’s runs over the 30-day period were calculated. Running volume was quantified by calculating the sum of durations for all runs over the recording period and dividing by the total number of days for the recording period (min-day⁻¹). In an effort to limit the burden of recording for runners, environmental data (temperature, humidity, etc.) and nutritional intake during each run were not documented in the journal.

Prospective Gastrointestinal Symptoms

Runners reported GI symptoms associated with each run with the same prospective journal that was used for tracking daily running sessions. The exact instructions included in the journal were as follows: “After each run, rate the overall level of discomfort you experienced during the run for the following symptoms: nausea, regurgitation/reflux, stomach fullness, abdominal cramps, gas/flatulence, and urge to defecate.” Nausea, regurgitation/reflux, and stomach fullness were considered upper GI symptoms, while abdominal cramps, gas/flatulence, and urge to defecate were considered lower GI symptoms. The following standardized definitions were provided for each symptom to ensure consistency of reporting across runners.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Characteristics of the Runners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics</td>
<td>Men (n = 70)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>46.1 ± 10.7</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>23.9 ± 2.5</td>
</tr>
<tr>
<td>Experience running (years)</td>
<td>10 (5–20)</td>
</tr>
<tr>
<td>Run duration (min-run⁻¹)</td>
<td>82.5 (61.0–113.1)</td>
</tr>
<tr>
<td>Running volume (min-day⁻¹)</td>
<td>57.9 (44.5–73.2)</td>
</tr>
<tr>
<td>Run RPE (6–20)</td>
<td>12.4 ± 1.5</td>
</tr>
</tbody>
</table>

Note. BMI = body mass index; RPE = rating of perceived exertion. Data presented as mean ± SD or median (IQR).
Nausea: A feeling of sickness in the stomach marked by an urge to vomit.

Regurgitation/reflux: Sensation of food or fluid returning from the stomach to the esophagus or mouth.

Stomach fullness: A sensation of fullness or abdominal pressure in the upper abdomen.

Abdominal cramps: Pain or cramping sensation, often experienced in the mid- or lower-portion of the abdomen.

Gas/flatulence: Gas or flatus expelled through the anus.

Urge to defecate: Sensation of needing to pass a bowel movement.

Runners rated symptoms based on a 0–10 Likert scale with descriptors of ‘no discomfort’, ‘moderate discomfort’ and ‘unbearable discomfort’ anchored at 0, 5 and 10. Importantly, 0–10 Likert scales have been extensively validated in other settings for the assessment of pain and discomfort (Bijur, Latimer, & Gallagher, 2003; Farrar et al., 2001), and this specific scale has been used previously to assess exercise-associated GI distress (Wilson, 2016).

The prospective journals were returned approximately 30 days after recording the first run. Mean, median, and maximum values for each symptom over the recording period were calculated for each runner. Because frequency and severity of GI symptoms are important aspects of GI distress, a variable incorporating both was created by calculating the proportion of runs over the 30-day period that runners reported at least one GI symptom greater than the following cut-off values: ≥1, ≥3, and ≥5. These proportions were calculated for all six symptoms combined, as well as for upper and lower symptoms separately.

**Retrospective Gastrointestinal Symptoms**

Runners also retrospectively rated GI symptoms at the end of the 30-day period. A web-based Qualtrics questionnaire was sent to runners, which had them report GI symptoms, gender, height, weight, years of running experience, and GI-related medical conditions. For retrospective GI symptoms, runners were asked to “rate the overall level of discomfort you have experienced during your training runs over the past month”. The same 0–10 scale and anchors (‘no discomfort’, ‘moderate discomfort’ and ‘unbearable discomfort’) were used. Approximately 24–36 hr after completion of the first retrospective GI questionnaire, the same questionnaire again was sent and completed within seven days.

**Statistics**

Data were analyzed using SPSS version 22 (IBM, Armonk, NY, USA). Normality of continuous data was assessed using the Kolmogorov-Smirnov test and by visually inspecting histograms. Data that were normally distributed were presented descriptively using means ± SD (SD), while medians with interquartile ranges (IQR) were used for nonnormal data.

Validity and reliability of the retrospective GI distress questionnaire were evaluated through the use of correlation coefficients. Because of nonnormal distributions that were resistant to normalization, Spearman’s rho correlations were used. A Mann-Whitney U test was used to examine if there were any differences.
in GI distress (proportions of runs with at least one GI symptom) between men and women, which was done to determine if correlations should be carried out in a sex-specific manner. A sample size calculation determined that approximately 47 runners were needed to detect at least modest correlations (rho = 0.40), assuming a beta of 0.20 and alpha of 0.05. A two-sided p-value £ 0.05 was used as the threshold for statistical significance.

**Results**

Overall, men and women experienced at least one GI symptom on 84.0% (59.8–95.1%) and 78.3% (50.0–95.2%) of their runs, respectively (Figure 2). Results from a Mann-Whitney U test indicated there was no difference between men and women for the proportion of runs with at least one GI symptom (Z = -1.1, p = .27). The proportions of runs for which runners experienced at least one GI symptom score ≥3 were 43.1% (16.5–71.0%) and 47.6% (16.7–69.2%) for men and women, respectively. At least one moderate-to-severe (≥5) GI symptom was experienced on 13.8% (6.7–37.3%) and 21.7% (5.3–41.2%) of men’s and women’s runs, respectively. Data specific to upper and lower GI symptoms are presented in Figure 2 as well.

Next, GI ratings from the prospective journal and retrospective questionnaire showed significant agreement (Table 2). All correlation coefficients were statistically significant (p < 0.001), ranging from 0.47 to 0.82. Without exception, the correlation coefficients for each GI symptom were highest when the prospective journal ratings were quantified as means, which was followed by maximum values and finally by median values. Thus, it appears that retrospective GI symptom reports show moderate-to-high validity when compared with daily journaling and most closely reflect overall mean values of GI distress over a 30-day period.

Reliability of the retrospective GI symptom reports was assessed by examining agreement between the first and second administrations of the retrospective questionnaire. The two questionnaires showed a high level of agreement, with correlations ranging from 0.78 to 0.92 (p < .001; Table 3). Nausea had the lowest test-retest repeatability (rho = 0.78), while urge to defecate showed the highest test-retest repeatability (rho = 0.92).

**Discussion**

Exercise has been known for decades to be associated with a transitory increase in GI dysfunction and subjective symptoms (Fogoros, 1980). Estimates to date have placed the incidence of GI distress during exercise from as low as 4% to over 90% depending on the study methodology and characteristics of the exercise bout (de Oliveira et al., 2014). Of all athletic endeavors, running is associated with the most pronounced increase in GI symptomology, which is likely due to a combination of factors that include reduced gut blood flow and increased mechanical jostling (de Oliveira, 2016; van Wijck et al., 2012). Much of the previous literature, however, has assessed GI symptoms over a single event or training session. Although assessing GI symptoms over an acute timeframe is more practical than assessing symptoms that reoccur chronically, the estimates from these studies may be more susceptible to being influenced by aberrant cases of GI distress. The present study, which was based on a 30-day assessment period, found that the majority of runners’ sessions were affected by at least one...
GI symptom (84.0% and 78.3% for men and women, respectively). With that said, experiencing a symptom like mild flatulence is clearly different than experiencing moderate-to-severe forms of symptoms like nausea or an urge to defecate. With that in mind, the results herein also show that moderate-to-severe GI symptoms (score of ≥5) were experienced fairly frequently, with 13.8% and 21.7% of men’s and women’s runs being affected.

Given the lack of research examining chronic GI symptoms in runners, direct comparisons to previous literature are somewhat challenging. The most comprehensive study to date that has examined the incidence of

### Table 2  Validity of the Retrospective Questionnaire Based on Spearman’s rho Correlations Between the Prospective Journal and Retrospective Questionnaire

<table>
<thead>
<tr>
<th>Journal reports</th>
<th>Nausea</th>
<th>Regurgitation/ reflux</th>
<th>Stomach fullness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean nausea</td>
<td>.59*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median nausea</td>
<td>.48*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max Nausea</td>
<td>.53*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean regurgitation / reflux</td>
<td>.76*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median regurgitation / reflux</td>
<td>.50*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max regurgitation / reflux</td>
<td>.72*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean stomach fullness</td>
<td></td>
<td>.74*</td>
<td></td>
</tr>
<tr>
<td>Median stomach fullness</td>
<td></td>
<td>.66*</td>
<td></td>
</tr>
<tr>
<td>Max stomach fullness</td>
<td></td>
<td>.69*</td>
<td></td>
</tr>
</tbody>
</table>

### Table 3  Reliability of the Retrospective Questionnaire Based on Spearman’s rho Correlations

<table>
<thead>
<tr>
<th>Administration 2</th>
<th>Nausea</th>
<th>Regurgitation/ reflux</th>
<th>Stomach fullness</th>
<th>Abdominal cramps</th>
<th>Gas/ flatulence</th>
<th>Urge to defecate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nausea</td>
<td>.78*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regurgitation/ reflux</td>
<td>.87*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stomach fullness</td>
<td></td>
<td>.85*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominal cramps</td>
<td></td>
<td></td>
<td>.83*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gas/ flatulence</td>
<td></td>
<td></td>
<td></td>
<td>.83*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urge to defecate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.92*</td>
<td></td>
</tr>
</tbody>
</table>

*Denotes p-values < .001. The two questionnaires were administered and completed no more than 7 days apart.
GI distress in endurance athletes comes from Pfeiffer et al. (2012), who found that roughly 4–32% of endurance competitors experienced at least one GI symptom ≥5 on a 0–9 scale. Specific to running, Pfeiffer et al. (2012) found that 4% of marathoners experienced at least one GI symptom ≥5. The slightly different scale (0–9) used by Pfeiffer et al. (2012), as compared with the present investigation (0–10), further complicates comparisons. The present study used a 0–10 scale because they are extensively validated in other areas of pain/discomfort research (Bijur et al., 2003; Farrar et al., 2001) and the numerical anchors at each end of the scale (0 and 10) have clear implicit meaning to most people. Regardless of the methodological differences, the results of the current study and Pfeiffer et al. (2012) suggest that upwards of 20–30% of endurance exercise bouts (including running) are affected by moderate-to-severe GI distress.

Although Pfeiffer et al. (2012) offer some of the most recently comparable data, the most directly comparable data to the current study come from Keeffe et al. (1984) and Riddoch and Trinick (1988). Using a single retrospective questionnaire, Keeffe et al. (1984) had 707 runners report whether they experienced a particular GI symptom during easy and hard runs. The presence of upper GI symptoms (either occasionally or frequently) ranged from 0.3% for vomiting during easy runs to 11.6% for nausea during hard runs. For lower GI symptoms, roughly 36–38% of runners reported either occasionally or frequently experiencing an urge to defecate (Keeffe et al., 1984). Likewise, Riddoch and Trinick (1988) reported that among 471 marathoners, 83% reported occasionally or frequently suffering from one or more GI symptoms during or immediately after running, which closely reflects the finding from the current study that runners experienced at least one GI symptom during 78–84% of runs. In totality, the current study’s findings, along with results from previous literature, provide robust evidence that most running sessions are affected in some form or another by GI symptoms.

Beyond establishing the prevalence of GI distress over a chronic timeframe, this study also addressed the important issue of whether a retrospective questionnaire is a valid and reliable tool for assessing GI symptoms. To the knowledge of the author, this is the first study to evaluate these properties of a GI symptom questionnaire in the context of exercise training. In terms of validity, the retrospective questionnaire performed reasonably well when considering prospective journaling as the reference method, as all correlation coefficients were ≥ 0.47. Moreover, correlations with mean values from the prospective journals were all above 0.7, with the exception of nausea (rho = 0.59). In regards to reliability properties, the Spearman’s rho correlations from Table 3 suggest the retrospective questionnaire was moderately-to-highly reliable (all rho ≥ 0.78) for measurements taken no more than seven days apart. The favorable reliability observed in this sample of runners reaffirms other research that has found retrospective GI symptom reports to have acceptable test-retest reliability in nonathletes (Adelstein et al., 2008). Taking the validity and reliability evaluations together, this investigation provides reassurance for investigators wishing to use retrospective questionnaires as a convenient way to measure chronic GI symptomology in endurance runners.

A practical issue closely related to validity and reliability is whether retrospective GI questionnaires more closely reflect peak or cumulative GI distress. As discussed previously, retrospective memories of discomfort may more accurately reflect real-time peak discomfort as opposed to average discomfort or the duration of discomfort (Redelmeier & Kahneman, 1996). Data from this study reveal that retrospective reports correlated most highly with mean ratings from prospective journals, which was followed by maximum ratings and finally by median ratings. Given that the prospective journal ratings exhibited a positive skew for most runners (which would result in a lower median than mean), it’s probable that peak or near peak values were primarily responsible for the higher correlations with mean ratings. These findings suggest that both peak and cumulative discomfort influence retrospective reports but that peak or near peak discomfort may play a more prominent role.

Some methodological weaknesses should be considered when interpreting the findings of this investigation. Runners volunteering for this study could have been more likely to regularly experience GI distress due to the fact that recruitment materials and the informed consent document made it clear that examining GI distress was a goal of the study. Thus, the occurrence of GI distress observed may not truly reflect that prevalence of these symptoms in all runners. As discussed, however, the estimates obtained from this study appear to be in congruence with several others (Keeffe et al., 1984; Riddoch & Trinick, 1988). While runners were asked to prospectively track GI symptoms over 30 days with a journal, they still reported these symptoms after each run. Thus, even the journals should be considered, to some degree, retrospective in nature. To mitigate this concern, runners were encouraged to record GI symptoms immediately after each run. Finally, only six symptoms were evaluated for this study, while others have evaluated additional symptoms such as vomiting, actual bowel movements, bloody bowel movements, belching, etc. (Keeffe et al., 1984; Pfeiffer et al., 2012). Most of these other GI symptoms, with the exception of belching, are relatively uncommon and thus do not pose a major issue for the generalizability of the current study’s findings.

To summarize, runners typically experience at least one GI symptom on the majority of their runs (78–84%), and of even more concern, up to 13.8% and 21.7% of men’s and women’s runs are affected by moderate-to-severe GI distress. In comparison with tracking GI symptoms on a daily basis with a journal, retrospective questionnaires seem to offer a convenient, valid, and reliable method of quantifying GI symptoms over 30 days. Additional research, however, may be warranted to examine how these measurement properties change under other conditions and for different athletes.
Acknowledgments

The study was designed by PBW. Data were collected by PBW and analyzed by PBW. Data interpretation and manuscript preparation was undertaken by PBW. No funding was used to support this project. The author has no conflicts of interest to report.

References


