VALUE IN HEALTH ▮ (2016) ▮▮▮–▮▮▮

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Comparison of Daily versus Weekly Recording of Gastroesophageal Reflux Disease Symptoms in Patients with a Partial Response to Proton Pump Inhibitor Therapy

Anna Rydén, PhD[1,*], *Olivia C. Leavy, MSc*[2], *Katarina Halling, MSc*[1], *Arthur A. Stone, PhD*[3]

[1]*AstraZeneca Gothenburg, Mölndal, Sweden;* [2]*Oxford Brookes University, Oxford, UK;* [3]*Department of Psychology and the USC Dornsife Center for Self-Report Science, University of Southern California, Los Angeles, CA, USA*

ABSTRACT

Background: The Reflux Symptom Questionnaire electronic Diary (RESQ-eD) and the Reflux Symptom Questionnaire 7-day recall (RESQ-7) are versions of a patient-reported outcome instrument that was developed and validated for measuring the frequency and intensity of symptoms in patients with gastroesophageal reflux disease (GERD) who have a partial response to proton pump inhibitor (PPI) therapy. **Objective:** The aim of these analyses was to assess the ability of the RESQ-7 to reproduce findings based on RESQ-eD reports of the same symptoms. **Methods:** These analyses are based on data from patients with GERD with a partial response to PPI (ClinicalTrials.gov identifier: NCT00703534). Participants completed the RESQ-eD twice daily for 7 days and the RESQ-7 on day 7. **Results:** Data from 446 patients were available for these analyses. Symptom-level analyses showed that, for intensity, mean domain scores were higher for the RESQ-7 (range 1.49–2.72) than for the RESQ-eD (range 1.45–2.57); for frequency, scores were lower for the RESQ-7 (range 2.58–4.82) than for the RESQ-eD (range 4.22–6.24). Correspondence analyses of RESQ-7 and RESQ-eD mean domain scores indicated excellent agreement for intensity (correlation-concordance coefficient 0.77–0.83) and fair agreement for frequency (correlation-concordance coefficient 0.40–0.58). Mean RESQ-eD subscale intensity scores for GERD symptoms were higher for symptoms experienced during the daytime than for those occurring at nighttime. Symptom recall was not associated with peak or recency effects. **Conclusions:** Patients with GERD slightly overestimated the intensity of their reflux symptoms and markedly underestimated the frequency on weekly recall compared with twice-daily reporting.

Keywords: ecological validity, frequency, gastroesophageal reflux disease, intensity, patient-reported outcome instruments.

Introduction

Having valid and reliable patient-reported outcome (PRO) instruments for measuring symptoms is paramount in medical research. Disease-specific PRO instruments are developed for particular target populations, and it is thus crucial that they capture the complete symptom pattern of the group of patients of interest [1]. The recall period is also an important aspect of a PRO instrument, and depends on the target population, symptoms, and the general setting in which the PRO instrument is to be implemented. Memory biases (cognitive heuristics) tend to affect retrospective evaluation of symptoms, with the majority of individuals being disproportionately influenced by the most intense and the most recent symptom events when asked to recall the average intensity of symptoms, an effect termed the "peak–end" rule [2–4].

In its guidance for industry, the US Food and Drug Administration recommends that questionnaires generally use short recall periods or ask patients to describe their current or recent state [5]. Asking patients to think back over a long period can reduce the accuracy of recall for all but the most memorable or stressful events [6], and may also make it more likely that the response is influenced by the patients' current state [7].

Patients with gastroesophageal reflux disease (GERD) typically have troublesome symptoms of heartburn and regurgitation [8], which usually resolve with proton pump inhibitor (PPI) therapy [9,10]. Results from a systematic review indicated that about 20% to 30% of patients with GERD who participated in primary care

Ethical approval: These post hoc analyses are based on data from the Patient-Reported Outcome Validation Study (ClinicalTrials.gov identifier: NCT00703534). All procedures performed in the study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Written informed consent was obtained from all individual participants included in the study.

Conflict of interest: A. Rydén and K. Halling are employees of AstraZeneca Gothenburg, Mölndal, Sweden. O. C. Leavy is an MSc graduate from Oxford Brookes University, Oxford, UK. Arthur A. Stone is a consultant for ERT Inc. and is a Senior Scientist with the Gallup Organization.

* *Address correspondence to:* Anna Rydén, AstraZeneca Gothenburg, S-431 83, Mölndal, Sweden.

E-mail: anna.ryden@astrazeneca.com.

1098-3015$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2016.05.007

studies had persistent heartburn or regurgitation during PPI therapy [11]. Symptoms confirmed by patients with a partial response to PPIs as being relevant to their GERD experience, in addition to heartburn and regurgitation, are hoarseness, cough, difficulty swallowing, and burping [12].

Most patients with GERD report having both daytime and nighttime symptoms [13]. Although the nature of daytime and nighttime GERD symptoms is similar in individuals with a partial response to PPI therapy, one study has reported a greater focus by patients on symptom intensity at night than during the day [14]. This increased attention paid to symptom intensity at night compared with during the day could potentially overestimate nighttime versus daytime symptom intensity on retrospective evaluation.

PRO instruments for evaluating GERD symptoms are used to measure symptom intensity and frequency in clinical trial settings, and are valuable tools when assessing symptoms and deciding on disease management in primary care [6]. When asking patients to recall their GERD symptoms, the duration of the recall period used may be influenced by practical considerations. While daily electronic diaries provide an effective way of assessing conditions such as GERD, in which symptoms fluctuate from day to day [15], this short recall period may not be practical for use in routine clinical practice or in pragmatic clinical trials, for which questionnaires with a 7-day recall period are commonly used. The Reflux Symptom Questionnaire electronic Diary (RESQ-eD) and the Reflux Symptom Questionnaire 7-day recall (RESQ-7) are PRO instruments that were developed and validated for measuring the intensity and frequency of GERD symptoms in patients with a partial response to PPIs [12,16]. Both versions of the instrument are conceptually and structurally identical. However, the RESQ-eD uses a twice-daily electronic recording, which makes it preferable in the clinical trial setting, whereas the RESQ-7 uses a 1-week recall period, making it more practical in routine clinical care [12,16].

The aim of the post hoc analyses presented in this article was to assess the ability of the RESQ-7 to reproduce faithfully findings based on twice-daily RESQ-eD reports of the same symptoms (i.e., its ecological validity [17]), using data from a large clinical study [16]. First, basic questions about the comparability and correlation of mean levels of aggregated real-time and recall measures were considered separately for intensity and frequency. Next, cognitive heuristics (peak and recency effects) that may be at work in the recall measures were assessed. Last, the role of nighttime symptoms in recall was examined.

Methods

Participants

These post hoc analyses are based on data from part 1 (8–12-day screening phase) of the PRO Validation Study, which included patients with GERD who had a partial response to PPI therapy (N = 580; mean age 48 years [range 19–70 years]; 58% women; ClinicalTrials.gov identifier: NCT00703534) [16]. To be eligible for inclusion in the PRO Validation Study, patients had to have a history of GERD symptoms for at least 6 months and a minimum of 4 weeks of PPI therapy. Patients must have also reported (using the RESQ-7) a minimum of 3 days on which they experienced an at least mildly intense burning feeling behind the breastbone, and/or unpleasant movement of material upward from the stomach over the previous 7 days.

Measures

Participants completed the RESQ-eD twice daily during the screening phase (in the morning and before bedtime) and the RESQ-7 at the end of the screening phase. The RESQ-eD and the RESQ-7 were developed and validated for use in clinical trials in patients with a partial response to PPI therapy [12,16]. Both versions have the same 13 symptom items and use the same intensity scoring. However, the RESQ-eD is a twice-daily diary, whereas the RESQ-7 has a 7-day recall period.

For each version of the instrument, the 13 symptom items combine into four separate domains: Heartburn (five items: burning feeling behind breastbone; pain behind breastbone; heartburn; burning feeling in upper stomach; and pain in upper stomach); Regurgitation (four items: acid taste in mouth; bitter taste in mouth; unpleasant movement of material upward from the stomach; and stomach contents [liquid or food] moving upward to throat or mouth); Hoarseness, cough, difficulty swallowing (three items: hoarseness; cough; and difficulty swallowing); and Burping (one item: burping). The intensity of each symptom item is recorded using a six-point scale (0 = did not have; 1 = very mild; 2 = mild; 3 = moderate; 4 = moderately severe; and 5 = severe). The RESQ-7 also captures information on symptom frequency during the past week (have not had; 1 day; 2 days; 3–4 days; 5–6 days; daily [with a day being defined as a 24-hour period]).

Symptom Analyses

For RESQ-eD symptom analyses, data were assessed for the 7 days ending with the last day of the screening phase (i.e., the 7 days covered by the RESQ-7 recall period). Each day was defined as the 24-hour period covered by the symptom scores reported on the RESQ-eD in the evening (which covered symptoms experienced during the daytime) and those reported the next morning (which covered symptoms experienced during the preceding nighttime).

Daily (i.e., 24-hour) RESQ-eD symptom intensity was defined in the PRO Validation Study as the higher of the two daily item intensity scores (i.e., either the evening score or the following morning score, whichever was the higher). Mean weekly RESQ-eD symptom intensity was defined as the mean of the daily (24-hour) intensities during the week. Daily RESQ-eD domain intensity was defined as the mean (morning and evening) intensity values for all the items in that domain. Mean weekly RESQ-eD domain intensity was defined as the mean of the daily domain intensities during the week. RESQ-eD symptom frequency was computed on the basis of RESQ-eD intensity scores, with a daily item intensity score of at least very mild (score ≥ 1) indicating a symptomatic day for that symptom item. For domain frequency, an intensity score of at least very mild on a particular day for any item in that domain was taken as indicating a day with symptoms for that domain.

Statistical Analyses

To be included in the analyses, patients had to have completed both the RESQ-eD and the RESQ-7 during the screening phase. Imputation was performed on the RESQ-eD data set if a single observation was missing in the sequence of morning or evening registrations, by replacing the missing value with the larger of the two surrounding values. If two or more consecutive observations were missing, no imputation was performed and the patient was excluded from these analyses.

Mean ± SD domain scores were calculated for the RESQ-eD and the RESQ-7. For weekly frequency data from the RESQ-7, midpoints were created for the RESQ-7 response options 3 to 4 days (3.5 days) and 5 to 6 days (5.5 days). For the weekly frequency data from the RESQ-eD, the variable was categorized such that the 3-day response option and the 4-day response option were recoded into the 3- to 4-day interval, and the 5-day