Assessment of reflux symptom severity: methodological options and their attributes

P Bytzer

Gut 2004;53(Suppl IV):iv28–iv34. doi: 10.1136/gut.2003.034298

Despite major advances in our understanding of reflux disease, the management of this disorder still presents many challenges. Reduction of heartburn is the most readily apparent objective for the patient with reflux disease. Thus the ability to measure heartburn accurately is of fundamental importance to clinical research in reflux disease. Here, the available data on the assessment of reflux symptoms—predominantly heartburn—in clinical trials of symptomatic reflux disease are examined.

still presents many challenges. Reduction of heartburn is the most readily apparent objective for the patient with reflux disease. Thus the ability to measure heartburn accurately is of fundamental importance to clinical research in reflux disease. Obstacles to interpreting the patient's subjective assessment include lack of agreed definitions of symptoms, arbitrary gradations of symptom severity and frequency, and lack of validated rating scales. An important requirement for future clinical research will be to define guidelines for the assessment of symptoms and, hence, of treatment success. This review will examine the available data on the assessment of reflux symptoms—predominantly heartburn—in clinical trials of symptomatic reflux disease.

Reflux disease may be associated with many symptoms but the major ones assessed in clinical trials are "heartburn" and "regurgitation". Here, I will concentrate on heartburn to illustrate the problems associated with symptom assessment but it should be recognised that heartburn is probably the best characterised reflux symptom and that difficulties in assessment are even greater for other symptoms that occur less frequently, are associated less clearly with reflux disease, and are often even less well defined than heartburn.

Most importantly, there is no universally accepted definition of heartburn. A definition of heartburn as "a burning feeling rising from the stomach or lower chest towards the neck" leads to improved recognition of reflux symptoms and is predictive of a good symptomatic response to acid suppression with a proton pump inhibitor (PPI).^{1–2} This description of heartburn is also important because it has been the enrolment criterion for many gastro-oesophageal reflux disease (GORD) treatment studies in patients with either endoscopy negative reflux disease or erosive oesophagitis.^{3–11} However, despite this, many patients do not consider heartburn and retrosternal burning to be synonymous.¹

DEFINITIONS OF HEARTBURN SEVERITY

The vast majority of clinical trials have graded heartburn severity using an ordinal scale (for example, a modified Likert scale) and most have used a four grade modified Likert scale, with word anchors defining heartburn severity by its impact on daily life (for example, "causing interference with normal activities") (table 1).

Abbreviations: GORD, gastro-oesophageal reflux disease; PPI, proton pump inhibitor; VAS, visual analogue scales

SUMMARY

Heartburn is usually assessed by measuring severity and frequency using modified Likert scales, usually with four, five, or seven grades. The various grades are not always defined and they frequently differ among trials. Severity is measured as either the most severe episode of heartburn over the past day, week, or month, or by the overall average of symptoms. Heartburn frequency is usually assessed at trial entry but not always at the end. Furthermore, frequency of heartburn is seldom a part of the definition of treatment success unless it is incorporated into a description that defines absence of symptoms. The number of days with heartburn over the past week, or the numbers of hours with heartburn over a 24 hour period, have been used to measure frequency. Patients who report frequent symptoms also seem to suffer from more severe grades of heartburn. Clinical trials suggest that the severity and frequency of heartburn improve in parallel during medical therapy. Diverse symptom response measures have been used, many studies reporting the proportion of patients who experienced absence of reflux symptoms or the number of symptom free days as primary outcome measures. Complete absence of heartburn is a very attractive outcome measure because it is unambiguous. Validation studies are lacking and it is not clear what the most appropriate outcome is in patients with heartburn. In short term studies, a strict end point, such as "absence of heartburn for the last seven days", appears attractive. In long term studies, the phrase, "sufficient control of heartburn", may be a suitable outcome measure although it too requires appropriate validation.

INTRODUCTION

Despite major advances in our understanding of reflux disease, the management of this disorder

Correspondence to:
Dr P Bytzer, Associate
Professor of Medicine,
Head, Department of
Medicine, Division of
Gastroenterology,
Glostrup University
Hospital, DK-2600
Glostrup, Denmark;
Peter.Bytzer@
DADLNET.DK

Table 1 Examples of definitions of heartburn severity used in clinical trials			
	Definition	Comment	Studies where definition was used
Four grade modified Likert scale	None: no heartburn	Most intense episode of heartburn during the previous 24 hours rated by patient diaries	Castell ¹³
	Mild: awareness of heartburn, but easily tolerated		
	Moderate: discomforting heartburn sufficient to cause interference with normal activities, including sleep	Most intense episode of heartburn during the previous 7 days rated by the investigator	Johnson ¹⁴
Four grade modified Likert scale	Severe: incapacitating heartburn, with inability to perform normal activities, including sleep		
	None	Overall heartburn intensity over each 24 hour period rated by patient diaries	Johnsson ⁴³
	Mild: awareness of symptom but easily tolerated		
Five grade modified Likert scale	Moderate: discomfort sufficient to cause interference with normal activities	Overall heartburn symptoms in the previous 7 days	Bate ⁴⁴
	Severe: incapacitating, with inability to perform normal activities		
	None		Schwizer ³²
Seven grade modified Likert scale	Mild: rare symptoms, no medication needed		
	Moderate: not interfering with daily activities, medications used occasionally, no dietary restriction		
	Severe: interfering with daily activities, chronic medications, dietary restrictions		
Visual analogue scale	Very severe: incapacitating, preventing daily activities, partial relief on medication, weight loss		
	None, minor, mild, moderate, quite severe, severe, very severe		Brun ⁴⁵
	0–100 mm scale ranging from no symptoms to unbearable symptoms		Faaij ⁴⁰

Only a few studies have used five or seven grade modified Likert scales, despite their methodological advantages over scales with fewer grades.¹² More importantly, the time frame for symptom assessment has varied in different studies. Most questionnaires have rated heartburn severity as the overall intensity of the symptom over the previous day or week. However, other questionnaires have asked the patient to grade the severity by defining the most intense episode of heartburn during the previous day or week.^{13–14} There have been no studies determining whether these variations in the definition of severity result in differences in the classification of individual patients.

In one study, the investigators incorporated both frequency and severity into the same scale, assuming that mild symptoms, which did not interfere with normal activities, occurred only occasionally and that severe symptoms, interfering with normal activities, were likely to be present frequently.¹⁵

Heartburn severity has also been graded using visual analogue scales (VAS) (table 1). VAS are continuous, usually 10 cm long, often with the extremes labelled by specific terms like “worst possible symptom” and “no symptom”. Their reproducibility and responsiveness in upper gastrointestinal symptoms are well established.¹⁶ When used in serial measurements, patients should see their prior responses as this may increase sensitivity and thus the power of the trial.¹⁷ Outcome measures obtained from VAS may be difficult to interpret as small but statistically significant results do not necessarily indicate clinical relevance. Furthermore, the clinical relevance of equal measures or changes in outcome as assessed by continuous scales may differ between subjects.

DEFINITIONS OF HEARTBURN FREQUENCY

Most definitions assess the number of days with heartburn over the previous week or month (table 2). A few studies have rated frequency by the number of hours during the last 24 hour period with symptoms.¹⁸ The gradings used are arbitrary and have not been validated. Even so, the frequency of heartburn is obviously an important descriptor of reflux

disease severity. Patients with frequent reflux symptoms (occasionally versus one to three times daily versus almost constantly present) have a significantly greater oesophageal acid exposure on 24 hour pH monitoring compared with those with less frequent symptoms.¹⁹ Assessment of heartburn frequency is often used in clinical trials, for example as part of the eligibility criteria (see tables 3, 4) but is rarely used as part of the definition of treatment success,²⁰ unless

Table 2 Definitions of heartburn frequency used in clinical trials	
Definitions	Examples of studies where definition was used
Rare: less than once a week Moderate: twice a week Frequent: three to six times a week Daily	Schwizer ³²
Grade 0: absent Grade 1: less than 2 days a week Grade 2: 2–4 days a week Grade 3: more than 4 days a week	Vigneri ²⁰
No of hours during the last 24 hour period when experiencing heartburn: None <1 h 1–6 h >6 h	Johnsson ¹⁸
No of days with heartburn in the last 7 days (0–7)	Venables ⁷
No of days with heartburn episodes during the last 7 days: None 1 day 2–4 days 5–6 days 7 days	Venables ²⁶
Visual analogue scale 0–100 mm	Galmiche ³⁹

complete absence of symptoms (for example, during the previous week) is assumed to indicate complete symptom control.

WHAT IS MOST IMPORTANT TO REFLUX PATIENTS: SEVERITY, FREQUENCY, OR BOTH

Traditionally, severity, frequency, and duration have all been held to be important symptom qualities. Information about the relative importance of these symptom parameters for patients with heartburn has come from clinical trials. Not surprisingly, both the severity and frequency of heartburn seem to be important characteristics. Duration of the individual symptom episodes is probably also important to patients but this aspect has received very little attention in clinical research.

Data from clinical trials in erosive and non-erosive reflux disease^{6,7} have established an apparent relationship between heartburn severity, graded as mild, moderate, or severe, and quality of life impairment. It should be noted however that this may be a spurious observation as the gradation of heartburn severity was based on its impact on daily living and it may thus be little more than an indirect quality of life measure. In addition, less frequent symptoms of greater severity and duration may be perceived by some patients as representative of significant disease and thus worth treating. A key issue is thus the level of symptom severity or frequency at which a significant reduction in quality of life is seen. Self reported heartburn frequency in a population based survey is an important predictor of health care seeking.²¹

Patients who report frequent symptoms also seem to suffer from more severe grades of heartburn. Baseline assessments from three clinical trials in non-erosive reflux disease, comparing esomeprazole with omeprazole with a total of 2642 patients, showed that patients who reported severe heartburn were more likely to have daily heartburn than those with mild heartburn (see fig 2 in Dent and colleagues²² in this supplement (page iv1–iv24)) (AstraZeneca, data on file).

Furthermore, results from these and other trials suggest that the severity and frequency of heartburn improve in parallel during medical therapy. For example, pooled data from three controlled trials comparing esomeprazole with omeprazole showed a relationship between the improvement in heartburn severity (scored on a four grade modified Likert scale) and reduction in the number of days per week with heartburn.²² Thus patients with a pronounced reduction in

Figure 1 Heartburn frequency and severity improves in parallel during medical therapy. Proportion of patients reporting a change in heartburn severity score of 3 (from severe to none) or 2 (from severe to mild or from moderate to none) according to the change in the number of days with heartburn (–7, from daily to none; –6, from daily to one day per week or from six days per week to none, etc). Results are pooled data from three controlled trials comparing esomeprazole with omeprazole (n = 2629) (AstraZeneca, data on file).

heartburn frequency also reported a more marked reduction in heartburn severity (fig 1). Comparable findings were reported from a clinical trial in which the symptomatic responses to omeprazole and ranitidine were evaluated in erosive reflux disease.²³

Even though these data might suggest that measures of symptom response could be restricted to either severity or frequency, we do not know if different treatment modalities, other than acid inhibitory drugs, might have a different impact on symptom patterns resulting in skewed or differential changes in these parameters. Furthermore, a randomised placebo controlled study comparing two doses of omeprazole suggested that both severity and frequency of heartburn are important independent determinants of patient satisfaction with therapy.⁴ New data support this, showing that most patients are willing to accept mild heartburn during treatment, but only for up to one day per week, whereas almost none is willing to accept severe or even moderate heartburn (see fig 1 in Dent and colleagues²² in this supplement (page iv1–iv24)).²⁴

WHAT IS MOST IMPORTANT IN CLINICAL TRIALS: SEVERITY, FREQUENCY, OR BOTH

A large number of different symptom response measures have been reported in the literature. Outcome measures in non-erosive reflux disease focus almost exclusively on symptom reduction and are usually more detailed and sophisticated than in trials for erosive reflux disease, which tend to concentrate on endoscopic signs of healing. Consequently, this review has focused mainly on methodology reported in non-erosive reflux disease trials.

Outcome measures should be validated in well designed studies designed for that purpose before they are used in clinical trials.²⁵ This ideal requirement has not been satisfied for the symptom outcome measures used in reflux disease and there is a remarkable lack of validation studies in the area. Thus it is not clear which outcome measure is most appropriate in GORD patients.

Only a minority of clinical trials in symptomatic reflux disease offer sufficient methodological details on the recording and definition of heartburn severity and frequency and outcome measures. Table 3 lists a number of different symptom outcome measures reported in major clinical trials in reflux disease. In table 4, eligibility criteria, relevant to heartburn symptoms, are listed together with a summary of outcome measures in trials, which have examined the symptomatic response to antisecretory medication in non-erosive reflux disease.

WHICH SYMPTOMS SHOULD BE MEASURED

Reflux patients often describe several different symptoms. Assessment of treatment effect for each individual symptom in clinical trials may thus lead to problems with false positive results as a result of multiple statistical testing. Furthermore, reflux patients may be disappointed if they expect reduction of all gastrointestinal symptoms when in fact the investigator focuses mainly on reduction of heartburn. In the study by Carlsson and colleagues⁶ which compared the effects of two doses of omeprazole in patients who had symptoms compatible with reflux disease, the primary outcome measure was “complete upper gastrointestinal symptom relief”. Belching and bloating were among the most common individual symptoms recorded at entry, and because these symptoms are probably not associated with gastric acid secretion or gastroesophageal reflux episodes, they would not be expected to improve on acid inhibitory drugs. Not surprisingly, this very broad definition of symptom reduction resulted in a response rate of only 35–41%, much lower than in other studies for which the primary outcome was reduction of heartburn.

Table 3 Summary of various symptom outcome measures in clinical trials of patients with heartburn

	Definition	Comment	Example of studies where outcome measure was used
Heartburn resolution	No symptoms in the last 7 days of therapy. ≤ 1 day with no more than mild heartburn episode in the past 7 days	Patient assessment by diary cards Heartburn frequency and severity recorded by diary card	Castell ¹³ Lind ⁴
Adequate control (or relief) of heartburn	Patient's assessment in response to the question: "does the study medication give sufficient control of your heartburn?"		Lind ⁴
Response	Symptoms improved by at least two grades (on 4 grade modified Likert scale) or from mild to none Symptoms improved by at least three grades (on 7 grade modified Likert scale) and no more than mild intensity		Bate ⁴⁴ Brun ⁴⁵
Willingness to continue	Willingness to continue maintenance treatment for heartburn with omeprazole or placebo (blinded)	End point analysed according to reason (inadequate relief of heartburn or other reasons)	Venables ²⁶
Length of time to acceptable symptom relief	Mean time (in min) up to acceptable symptom relief (no or mild symptoms) after single dose ranitidine treatment for heartburn episodes Time to onset of pain relief (VAS <75% of baseline for ≥30 min)	Diary card completed during the first hour after each study medication Electronic diary gave instructions to take medication and provided ratings scales for heartburn at frequent time intervals activated by an alarm clock	Elm ⁴⁶ Faaij ⁴⁰
Symptom scores	Median time in days for patients to achieve their first 24 hour interval without heartburn Scores calculated by adding daily severity values and multiplying by daily frequency values obtained during each week of symptom recording Heartburn scores calculated by multiplying the severity grade (0–3) by the frequency grade (0–3)	A reduction in heartburn score of at least 50% after treatment with omeprazole, compared with placebo treatment, was considered indicative of reflux disease	Miner ³⁰ Fass ³³ Vigneri ²⁰
Frequency of heartburn	Average heartburn severity per day or per night Percentage of heartburn free days and nights. Mean daily number of heartburn episodes. Number of days per week with heartburn. Reflux symptoms for less than four consecutive days	Heartburn severity scored as 0 = none, 1 = mild, 2 = moderate, 3 = severe Diary card data Diary card data Only frequency of symptoms part of the outcome definition—severity of symptoms not defined	Richter ⁴⁷ Richter ⁴⁸ Richter ²⁹ Hatlebakk ²⁷ Vigneri ²⁰

As the reflux symptoms of heartburn and regurgitation are part of the definition of symptomatic reflux disease, they have been the focus of treatment trials and are the primary concern in everyday clinical practice. Epigastric pain is not considered a specific symptom of gastro-oesophageal reflux but it often improves with active treatment in reflux patients,^{6 26 27} as does regurgitation.^{7 27 28} Other upper gastrointestinal symptoms, such as belching, bloating, nausea, and vomiting, have been evaluated in some placebo controlled trials.^{6 8 26 27 29 30} Some studies reported improvement, independent of treatment allocation, for many of these symptoms.^{27 29}

HOW SHOULD SYMPTOMS BE MEASURED

Ordinal scales are often used to evaluate the effect of treatment on reflux symptoms. This can be done by the use of single state scales or by transition scales. Single state scales, for example four or five grade modified Likert scales, are used to establish a patient's state at various time points (for example, at entry and completion). Scale scores should be composed of elements that are clearly defined, mutually exclusive, and ranked in a hierarchical manner. Furthermore, scale scores should be easy to translate into a clinical context.³¹ To optimise responsiveness to change, at least five to seven points should be included in the scale.¹² Furthermore, the scale must be able to detect improvement and deterioration equally in the patients under study. If patients are clustered at one end of the scale at entry, then

the scale may be unable to detect a change occurring in one direction—for example, deterioration in patients—who all score maximum severity at entry.

Transition scales measure the change in symptoms directly (for example, improved, unchanged, worse) and these scales should be symmetrical in their structure. Asymmetric designs—for example, with more grades for improvement than for deterioration—could potentially bias the results.³²

In reflux disease, composite or global symptom scores are usually developed by adding or multiplying daily heartburn severity by heartburn frequency.^{20 33} A predefined reduction in heartburn score may be used as an outcome criterion.³³ Unfortunately, categorisation in such scales is often ambiguous and the categories are not necessarily exhaustive or graded in equal intervals. Differences in scores may be difficult to evaluate unless the investigators provide a clinical context for interpretation.

Many recent studies have reported the proportion of patients who obtain total absence of symptoms or the number of symptom free days as primary outcome measures (table 4). These measures are easily understood, they make clinical sense, and they are not biased by the methodological difficulties associated with measuring subtle changes over time in symptom severity, frequency, and duration. On the other hand, these measures may underestimate treatment effect for subjects whose symptoms are reduced although not completely absent.

Table 4 Summary of assessments of heartburn in randomised treatment trials in endoscopy negative reflux disease

Study	Inclusion criteria Heartburn severity	Heartburn frequency	Assessment of heartburn severity	Assessment of heartburn frequency	Assessment of symptom outcome
Bate ³	Moderate to severe heartburn	Not specified	4 grade modified Likert scale averaged over last 7 days	Proportion of patients without heartburn Number of heartburn free days	Proportion of patients without heartburn
Lind ⁴	Mild, moderate, or severe heartburn	Heartburn at least 2 days weekly	4 grade modified Likert scale	Number of days with heartburn in the last week (0, 1, 2–4, 5–6, or 7 days)	Complete absence of heartburn in the last 7 days Resolution of heartburn (≤ 1 day with no more than mild heartburn in the 7 days before the 4 week visit)
Havelund ¹⁵	Mild, moderate, or severe heartburn	Mild, moderate, or severe heartburn	4 grade modified Likert scale	Number of days with heartburn	Heartburn score (average severity multiplied by number of days with heartburn)
Venables ⁷	Mild, moderate, or severe heartburn	2 out of 7 days with heartburn	4 grade modified Likert scale averaged over last 7 days	Number of days with heartburn in the last week	Relief of heartburn (≤ 1 day with no more than mild heartburn in the 7 days before the visit) Change in severity of heartburn Proportion of patients reporting heartburn
Venables ²⁶	Successful control of heartburn after 4–8 weeks of initial therapy	Not applicable	4 grade modified Likert scale averaged over last 7 days	Number of days with heartburn in the last week	Number of patients reporting heartburn
Galmiche ⁵	Mild, moderate, or severe heartburn	Heartburn during at least 2 days last week, despite alginic acid therapy	4 grade modified Likert scale	Number of days with heartburn	Resolution of heartburn (≤ 1 day with no more than mild heartburn in the 7 days before the visit) Complete absence of heartburn
Galmiche ³⁹	Not specified	At least 4 episodes of heartburn last week	100 mm visual analogue scale	100 mm visual analogue scale	Proportion of patients who had relief of at least 75% of heartburn episodes during the trial period Changes in heartburn symptoms Outcome of heartburn intensity and frequency assessed by visual analogue scale Total number of heartburn episodes Time to onset and duration of relief of heartburn
Carlsson ⁶	Score of 4 or more on Carlsson–Dent scale	Symptoms on at least 2 days during the last 7 days	4 grade modified Likert scale		Complete relief of upper gastrointestinal symptoms Sufficient control of upper gastrointestinal symptoms Proportion of patients without reflux symptoms per day over 4 weeks Willingness to continue (adequate relief of heartburn)
Lind ³⁵	Resolution of heartburn following short term treatment	None	Not applicable	Not applicable	
Richter ⁴⁸	Moderate or severe heartburn	Heartburn on $\geq 50\%$ of 7–10 pre-treatment days	4 grade modified scale	Number of days or nights with heartburn	Percentage of days and nights with heartburn Average heartburn severity/day
Hatlebakk ²⁷	Mild, moderate, or severe heartburn	>2 days/week in a 2 week run-in period	4 grade modified Likert scale	Number of days with heartburn in the 7 days before each visit	Adequate control of heartburn (≤ 1 day with no more than mild heartburn in the 7 days before the 4 week visit) Number of days per week with heartburn
Richter ²⁹	Moderate or severe heartburn	Heartburn on at least 4 of the last 7 days	4 grade modified Likert scale	Daily number of heartburn episodes	Complete resolution of heartburn every day during a full week Proportion of patients with no heartburn episodes on each day over 4 weeks Mean daily number of heartburn episodes
Richter ⁴⁷	Moderate or severe heartburn	Heartburn on $\geq 50\%$ of 7–10 pretreatment days	4 grade modified Likert scale	Number of days or nights with heartburn	Average heartburn severity per day Percentage of days and nights with heartburn
Talley ³⁶	No heartburn	Absence of heartburn in last 7 days of 4 week treatment period	4 grade modified Likert scale	Number of days with heartburn last week	Severity of heartburn last week Sufficient control of heartburn
Miner ³⁰	Moderate or severe heartburn	A minimum of five moderately severe GORD episodes in the 7 days before the end of a placebo run-in	5 grade modified Likert scale	Number of days or nights with heartburn	Median time in days to first 24 hour interval without heartburn

GORD, gastro-oesophageal reflux disease.

VARIOUS GRADES OF SYMPTOM CONTROL

Complete abolition of symptoms is a primary aim when treating patients with reflux disease from both methodological and clinical standpoints. From a methodological standpoint, the absence of heartburn is, intuitively, an attractive outcome measure and this is probably why it is one of the most widely used end points in reflux treatment trials. Complete absence of symptoms may not however be the primary long term aim of all patients. Indeed, complete absence of symptoms may seem to be a very ambitious and unrealistic goal, leading to a reduction in symptom severity to levels below those found in a healthy background population. In practice, many patients who are prescribed continuous treatment take their medication only when symptoms become troublesome. Thus reflux patients who were prescribed long term daily PPI therapy took their medication on only 50% of treatment days.³⁴ In controlled trials of on-demand treatment strategies in reflux disease, patients take a PPI on average every second to every third day.^{35–36} Similar findings have been reported from follow up studies outside the framework of a clinical trial.³⁷ Thus although abolition of reflux symptoms might seem to be an ideal outcome measure, many patients are prepared to accept a recurrence of reflux symptoms before they resume therapy.

In trials that have measured different levels of outcome (for example, “complete absence of heartburn”, “resolution of heartburn”, “adequate control of heartburn”), there is usually a hierarchy of treatment response rates, with the lowest response rate being reported for those in which complete absence of symptoms was the primary outcome. Interestingly, a significantly larger proportion of patients are willing to continue a treatment strategy even if it does not provide absolute symptom control.^{4–6–8}

Some clinical trials have used “sufficient control” or “resolution of heartburn” as an end point (tables 3, 4). This has usually been defined as no more than one day with no more than mild heartburn in the preceding week. In one study, “resolution of heartburn” corresponded well with the overall assessment reported by the patients in response to the question “Does the medication give sufficient control of your heartburn?”.⁴ A recent study reported on the relationship between complete absence of heartburn symptoms and quality of life.³⁸ Patients with complete absence of heartburn reported improved functioning and well being compared with patients with continuing heartburn problems. Unfortunately, the study did not report any comparisons between patients with complete absence of heartburn and patients with incomplete but acceptable control of their symptoms.

WHEN SHOULD SYMPTOMS BE MEASURED

When single scale states are used as outcome measures, baseline measurements are needed. These are often expanded to summarise overall heartburn intensity over an appropriate time prior to treatment (for example, one week or one month). Even though the use of transition scales does not require a baseline assessment, it is generally recommended as it serves to document the patient’s symptom state at entry. At a minimum, outcome should be measured at completion of the trial and this should be the primary data point. Often, intermediary data points are also obtained but repeated measurements may lead to problems with false positive results due to multiple testing. Some outcome measures summarise symptom intensity over time (for example, number of days with or without heartburn) by using diary cards. This may be an important additional measure in clinical trials comparing different interventions which may be associated with differences in the onset of heartburn reduction but which seem to be equally effective when measured at the end of the trial.

Recent studies have focused on the time (in minutes or hours) to symptom reduction after single or repeated doses of antisecretory medication.^{39–41} This is of obvious importance for patients who take their medication on demand where prompt reduction of symptoms is important. Patients are usually instructed to record symptoms at regular intervals (minutes or hours), depending on the perceived speed of action. An electronic patient diary or an interactive voice response system can ensure more valid symptom recordings and prevent retrospective entries.^{40–42}

WHO SHOULD MEASURE SYMPTOMS

As a general rule, the patient should assess symptoms and symptom outcome directly. Patients and investigators may disagree when both evaluate symptom outcome. In many studies, it is not clear whether the final assessment was done by the patients directly or indirectly by way of a physician interview. Several studies have shown that investigators tend to be more optimistic than their patients in estimating the magnitude of treatment response at the final assessment. Thus in the study by Sandmark *et al*, the investigators rated approximately 75% of patients as completely symptom free after four weeks of omeprazole therapy. In comparison, only approximately 55% of patients felt that their symptoms were completely gone.²³

CONCLUSION

A large number of different symptom response measures have been reported in the literature. Many studies report the proportion of patients who obtain total absence of symptoms or the number of symptom free days as primary outcome measures. These measures are “hard” or more objective end points that are easily understood. Furthermore, they make clinical sense and they are not biased by methodological difficulties. On the other hand, such crude measures will underestimate treatment effects in those with incomplete but satisfactory symptom response. There is a general lack of validation studies in this area and it is not clear what the most appropriate outcome is in patients with heartburn. In short term studies (weeks to a few months), a strict end point such as “absence of heartburn for the last seven days” appears attractive since it is unambiguous and, therefore, methodologically sound. Furthermore, it will provide the patient with an “internal” standard of the best possible care with which to compare future therapies. In long term studies, a less strict end point, such as “sufficient control of heartburn”, may be more appropriate. However, a less strict end point based on predefined criterion such as “no more than mild symptoms on no more than one day per week”, or on the patient’s decision as to treatment adequacy, introduces other problems of subjectivity with respect to the definition of “mild” and “adequacy”.

The choice of symptom outcome measure depends also on the aim of the clinical trial. A study to compare two similar therapies may be best able to discriminate if it uses a “hard” end point, such as complete abolition of symptoms, whereas a study to assess the effect of treatment on patient quality of life or satisfaction may require a more detailed assessment of the magnitude of change in a patient’s symptoms. Given that studies may have different aims, it might be preferable if the primary aim were specified clearly, but if other outcome measures were also specified to allow comparability between different studies.

REFERENCES

- 1 Carlsson R, Dent J, Bolling-Sternevald E, *et al*. The usefulness of a structured questionnaire in the assessment of symptomatic gastroesophageal reflux disease. *Scand J Gastroenterol* 1998;**33**:1023–9.

- 2 Carlsson R, Bolling E, Jerndal P, *et al.* Factors predicting response to omeprazole treatment in patients with functional dyspepsia. *Gastroenterology* 1996;**110**:A76.3.
- 3 Carlsson R, Frison L, Lundell L, *et al.* Relationship between symptoms, endoscopic findings and treatment outcome in reflux esophagitis. *Gastroenterology* 1996;**110**:A77.
- 4 Lind T, Havelund T, Carlsson R, *et al.* Heartburn without oesophagitis: efficacy of omeprazole therapy and features determining therapeutic response. *Scand J Gastroenterol* 1997;**32**:974-9.
- 5 Galmiche J-P, Barthelemy P, Hamelin B. Treating the symptoms of gastro-oesophageal reflux disease: a double-blind comparison of omeprazole and cisapride. *Aliment Pharmacol Ther* 1997;**11**:765-73.
- 6 Carlsson R, Dent J, Watts DA, *et al.* Gastro-oesophageal reflux disease in primary care: an international study of different treatment strategies with omeprazole. *Eur J Gastroenterol Hepatol* 1998;**10**:119-24.
- 7 Venables TL, Newland RD, Patel AC, *et al.* Omeprazole 10 milligrams once daily, omeprazole 20 milligrams once daily, or ranitidine 150 milligrams twice daily, evaluated as initial therapy for the relief of symptoms of gastro-oesophageal reflux disease in general practice. *Scand J Gastroenterol* 1997;**32**:965-73.
- 8 Bate CM, Griffin SM, Keeling PWN, *et al.* Reflux symptom relief with omeprazole in patients without unequivocal oesophagitis. *Aliment Pharmacol Ther* 1996;**10**:547-55.
- 9 Talley NJ, Venables TL, Green JBR, *et al.* Esomeprazole 40 mg and 20 mg is efficacious in the long-term management of patients with endoscopy-negative gastro-oesophageal reflux disease: a placebo-controlled trial of "on-demand" therapy for 6 months. *Gastroenterology* 2000;**118**:A658.
- 10 Kahrilas PJ, Falk GW, Johnson DA, *et al.* Esomeprazole improves healing and symptom resolution as compared with omeprazole in reflux oesophagitis patients: a randomized controlled trial. *Aliment Pharmacol Ther* 2000;**14**:1249-58.
- 11 Richter JE, Kahrilas PJ, Johanson J, *et al.* Efficacy and safety of esomeprazole compared with omeprazole in GERD patients with erosive esophagitis: a randomized controlled trial. *Am J Gastroenterol* 2001;**96**:656-65.
- 12 Veldhuyzen van Zanten SJO, Tytgat KMAJ, Pollak PT, *et al.* Can severity of symptoms be used as an outcome measure in trials of non-ulcer dyspepsia and *Helicobacter pylori* associated gastritis? *J Clin Epidemiol* 1993;**46**:273-9.
- 13 Castell DO, Kahrilas PJ, Richter JE, *et al.* Esomeprazole (40 mg) compared with lansoprazole (30 mg) in the treatment of erosive esophagitis. *Am J Gastroenterol* 2002;**97**:575-83.
- 14 Johnson DA, Benjamin S, Vakil N, *et al.* Esomeprazole once daily for 6 months is effective therapy for maintaining healed erosive esophagitis and for controlling gastroesophageal reflux disease symptoms: a randomized, double-blind, placebo-controlled study of efficacy and safety. *Am J Gastroenterol* 2001;**96**:27-34.
- 15 Havelund T, Aalykke C. The efficacy of a peptic-based raft-forming anti-reflux agent in endoscopy-negative reflux disease. *Scand J Gastroenterol* 1997;**32**:773-7.
- 16 Nyrén O, Adami HO, Bates S, *et al.* Self-rating of pain in nonulcer dyspepsia. A methodological study comparing a new fixed-point scale and the visual analogue scale. *J Clin Gastroenterol* 1987;**9**:408-14.
- 17 Guyatt G, Berman LB, Townsend M, *et al.* Should study subjects see their previous responses? *J Chronic Dis* 1985;**38**:1003-7.
- 18 Johnsson F, Hatlebakk JG, Klintenberg A-C, *et al.* The symptom relieving effect of esomeprazole 40 mg daily in patients with heartburn. *Gastroenterology* 2001;**120**:A437.
- 19 Joelsson B, Johnsson F. Heartburn—the acid test. *Gut* 1989;**30**:1523-5.
- 20 Vignieri S, Termini R, Leandro G, *et al.* A comparison of five maintenance therapies for reflux esophagitis. *N Engl J Med* 1995;**333**:1106-10.
- 21 Oliveria SA, Christos PJ, Talley NJ, *et al.* Heartburn risk factors, knowledge, and prevention strategies: a population-based survey of individuals with heartburn. *Arch Intern Med* 1999;**159**:1592-8.
- 22 Dent J, Armstrong D, Delaney B, *et al.* Symptom evaluation in reflux disease: workshop background, processes, terminology, recommendations, and discussion outputs. *Gut* 2004;**53**(suppl IV):iv1-24.
- 23 Sandmark S, Carlsson R, Fausa O, *et al.* Omeprazole or ranitidine in the treatment of reflux esophagitis. Results of a double-blind, randomized, Scandinavian multicenter study. *Scand J Gastroenterol* 1988;**23**:625-32.
- 24 Junghard O, Carlsson R, Lind T. Sufficient control of heartburn in endoscopy negative GORD trials. *Scand J Gastroenterol* 2003;**38**:1197-9.
- 25 Talley NJ, Nyrén O, Drossman DA, *et al.* The irritable bowel syndrome: toward optimal design of controlled treatment trials. *Gastroenterol Int* 1993;**6**:189-211.
- 26 Venables TL, Newland RD, Patel AC, *et al.* Maintenance treatment for gastro-oesophageal reflux disease. A placebo-controlled evaluation of 10 milligrams omeprazole once daily in general practice. *Scand J Gastroenterol* 1997;**32**:627-32.
- 27 Hatlebakk JG, Hyggen A, Madsen PH, *et al.* Heartburn treatment in primary care: randomised, double blind study for 8 weeks. *BMJ* 1999;**319**:550-3.
- 28 Johnsson F, Moum B, Vilien M, *et al.* On-demand treatment in patients with oesophagitis and reflux symptoms: comparison of lansoprazole and omeprazole. *Scand J Gastroenterol* 2002;**37**:642-7.
- 29 Richter JE, Peura D, Benjamin SB, *et al.* Efficacy of omeprazole for the treatment of symptomatic acid reflux disease without esophagitis. *Arch Intern Med* 2000;**160**:1810-16.
- 30 Miner P Jr, Orr W, Filippone J, *et al.* Rabeprazole in nonerosive gastroesophageal reflux disease: a randomized placebo-controlled trial. *Am J Gastroenterol* 2002;**97**:1332-9.
- 31 MacKenzie CR, Charlson ME. Standards for the use of ordinal scales in clinical trials. *BMJ* 1986;**292**:40-3.
- 32 Schwizer W, Thumshirn M, Dent J, *et al.* *Helicobacter pylori* and symptomatic relapse of gastro-oesophageal reflux disease: a randomised controlled trial. *Lancet* 2001;**357**:1738-42.
- 33 Fass R, Ofman JJ, Gralnek IM, *et al.* Clinical and economic assessment of the omeprazole test in patients with symptoms suggestive of gastroesophageal reflux disease. *Arch Intern Med* 1999;**159**:2161-8.
- 34 Hungin APS, Rubin G, O'Flanagan H. Factors influencing compliance in long-term proton pump inhibitor therapy in general practice. *Br J Gen Pract* 1999;**49**:463-4.
- 35 Lind T, Havelund T, Lundell L, *et al.* On demand therapy with omeprazole for the long-term management of patients with heartburn without oesophagitis—a placebo-controlled randomized trial. *Aliment Pharmacol Ther* 1999;**13**:907-14.
- 36 Talley NJ, Lauritsen K, Tunturi-Hihnela H, *et al.* Esomeprazole 20 mg maintains symptom control in endoscopy-negative gastro-oesophageal reflux disease: a controlled trial of 'on-demand' therapy for 6 months. *Aliment Pharmacol Ther* 2001;**15**:347-54.
- 37 Schindlbeck NE, Klausner AG, Berghammer G, *et al.* Three year follow up of patients with gastroesophageal reflux disease. *Gut* 1992;**33**:1016-19.
- 38 Revicki DA, Crawley JA, Zedet MW, *et al.* Complete resolution of heartburn symptoms and health-related quality of life in patients with gastro-oesophageal reflux disease. *Aliment Pharmacol Ther* 1999;**13**:1621-30.
- 39 Galmiche J-P, Shi G, Simon B, *et al.* On-demand treatment of gastro-oesophageal reflux symptoms: a comparison of ranitidine 75 mg with cimetidine 200 mg or placebo. *Aliment Pharmacol Ther* 1998;**12**:909-17.
- 40 Faaij RA, van Gerven JMA, Jolivet-Landreau I, *et al.* Onset of action during on-demand treatment with Maalox suspension or low-dose ranitidine for heartburn. *Aliment Pharmacol Ther* 1999;**13**:1605-10.
- 41 Holtmann G, Bytzer P, Metz M, *et al.* A randomized, double-blind, comparative study of standard-dose rabeprazole and high-dose omeprazole in gastro-oesophageal reflux disease. *Aliment Pharmacol Ther* 2002;**16**:479-85.
- 42 Robinson M, Fitzgerald S, Hegedus R, *et al.* Onset of symptom relief with rabeprazole: a community-based, open-label assessment of patients with erosive oesophagitis. *Aliment Pharmacol Ther* 2002;**16**:445-54.
- 43 Johnsson F, Weywadt L, Solhaug J-H, *et al.* One-week omeprazole treatment in the diagnosis of gastro-oesophageal reflux disease. *Scand J Gastroenterol* 1998;**33**:15-20.
- 44 Bate CM, Riley SA, Chapman RWG, *et al.* Evaluation of omeprazole as a cost-effective diagnostic test for gastro-oesophageal reflux disease. *Aliment Pharmacol Ther* 1999;**13**:59-66.
- 45 Brun J, Sorngard H. High dose proton pump inhibitor response as an initial strategy for a clinical diagnosis of gastro-esophageal reflux disease (GERD). *Fam Pract* 2000;**17**:401-4.
- 46 Elm M, Hellke P, Andrén K, *et al.* Time to relief of episodic symptoms of gastro-oesophageal reflux disease. A crossover comparison of single doses of the effervescent and standard formulations of ranitidine. *Scand J Gastroenterol* 1998;**33**:900-4.
- 47 Richter JE, Campbell DR, Kahrilas PJ, *et al.* Lansoprazole compared with ranitidine for the treatment of nonerosive gastroesophageal reflux disease. *Arch Intern Med* 2000;**160**:1803-9.
- 48 Richter JE, Kovacs TO, Greski-Rose PA, *et al.* Lansoprazole in the treatment of heartburn in patients without erosive oesophagitis. *Aliment Pharmacol Ther* 1999;**13**:795-804.