GASTROENTEROLOGY 1994;107:1741–1745

The Symptom-Association Probability: An Improved Method for Symptom Analysis of 24-Hour Esophageal pH Data

BAS L. A. M. WEUSTEN,* JAN M. M. ROELOFS,† LOUIS M. A. AKKERMANS,† GERARD P. VAN BERGE-HENEGOUWEN,* and ANDRÉ J. P. M. SMOUT*
Departments of *Gastroenterology and †Surgery, University Hospital Utrecht, Utrecht, The Netherlands

See editorial on page 1898.

Background/Aims: **All methods currently used to quantify the temporal relationships between symptoms and episodes of gastroesophageal reflux, as assessed by 24-hour pH monitoring, have major shortcomings. The aim of this study was to develop and validate a simple, all-comprising statistical method to calculate the probability that gastroesophageal reflux episodes and symptoms are associated.** ***Methods:*** **The 24-hour pH signal was divided into consecutive 2-minute periods. These periods and the 2-minute periods preceding the onset of symptoms were evaluated for the occurrence of reflux. Fisher's Exact Test was then applied to calculate the probability (*P* value) that reflux and symptom episodes were unrelated. Finally, the symptom-association probability (SAP) was calculated as (1.0 − *P*) × 100%. The SAP values found in 184 24-hour esophageal pH tests were compared with the symptom index and the symptom sensitivity index.** ***Results:*** **Discordance between the SAP and the symptom index was found in 21 patients (11%) and discordance between the SAP and the symptom-sensitivity index in 28 (15%). False-positive and false-negative symptom index values occurred preferentially in patients with small and large numbers of symptom episodes during the test, respectively ($P < 0.05$).** ***Conclusions:*** **The SAP is a single, simple, quantitative measure of the strength of the association between symptoms and reflux episodes that is devoid of the disadvantages inherent to previously used methods.**

In the diagnosis of gastroesophageal reflux disease, ambulatory 24-hour esophageal pH monitoring is a valuable tool not only for the quantitative measurement of esophageal acid exposure but also for the assessment of the association in time between symptoms and reflux. Several attempts have been made to quantify this association, all with their own restrictions.[1–5] The current most frequently used parameter is the symptom index (SI), defined as the percentage of reflux-related symptom episodes.[1] However, this parameter does not take into account the total number of reflux episodes. Hence, in patients with frequent gastroesophageal reflux, a high SI may well be caused by random associations between reflux and symptoms. The symptom-sensitivity index (SSI), defined as the percentage of symptom-associated reflux episodes, was introduced as an additional parameter to overcome the drawbacks of the SI.[2] However, this index fails to take into account the total number of symptom episodes, rendering its use of limited value. Other techniques aiming to correlate symptoms with reflux are either very complex or do not use adequate time windows.[3–5] Therefore, we wished to develop a practical and simple statistical method that would take all relevant observations into account, using an optimal time window, and would result in a single parameter, expressing the probability that observed associations are not caused by chance. This report describes this method and its evaluation in a group of patients with reflux symptoms.

Materials and Methods

Patients and Measurements

The 341 24-hour ambulatory esophageal pH monitoring studies performed in our laboratory from July 1991 to December 1993 were reviewed. From these, 184 studies (in 184 patients) were selected in which at least one symptom episode had occurred during the study. This group of 184 patients comprised 98 males and 86 females. Their mean age was 49 years (range, 16–79 years). All patients had been referred for investigation because of heartburn, acid regurgitation, or noncardiac chest pain occurring alone or in combination. Because the aim of our study was to analyze the relationships between symptoms and gastroesophageal reflux, the use of drugs, previous surgery, or the presence of esophagitis were not considered exclusion criteria.

In all patients, the lower esophageal sphincter was identified manometrically using an eight-channel water-perfused catheter with a Dent Sleeve (Dent Sleeve Pty. Ltd., Belair, Australia) and a pneumohydraulic low compliance perfusion system.

Abbreviations used in this paper: SAP, symptom-association probability; SI, symptom index; SSI, symptom-sensitivity index.

0016-5085/94/$3.00

A

Reflux	Symptoms +	Symptoms -	
+	S^+R^+	S^-R^+	R^+_{tot}
-	S^+R^-	S^-R^-	R^-_{tot}
	S^+_{tot}	S^-_{tot}	Total

B

Reflux	Symptoms +	Symptoms -	
+	3	22	25
-	1	694	695
	4	716	720

Figure 1. (*A*) Contingency table containing the number of symptomatic reflux-positive 2-minute periods (S^+R^+), the number of asymptomatic reflux-positive 2-minute periods (S^-R^+), the number of symptomatic 2-minute periods without reflux events (S^+R^-), the number of asymptomatic 2-minute periods without reflux events (S^-R^-), the total number of reflux-positive 2-minute periods (R^+_{tot}), the total number of 2-minute periods without reflux events (R^-_{tot}), the total number of symptomatic 2-minute periods (S^+_{tot}), the total number of asymptomatic 2-minute periods (S^-_{tot}), and the total number of recorded 2-minute periods (total). (*B*) Example of a contingency table in a patient who had four symptom episodes during the pH monitoring. In this example, 3 of 4 symptom episodes were associated with a reflux event; in total, 25 2-minute periods were reflux-positive. Application of Fisher's Exact Test results were $P = 0.0001$. The SAP is 99.99%.

Thereafter, a pH glass electrode (model LOT440; Ingold A. G., Urdorf, Switzerland) was inserted through the nose into the esophagus. The pH electrode was positioned 5 cm above the upper border of the lower esophageal sphincter. The pH signal was recorded in a digital data logger (MMS, Enschede, The Netherlands) using a sampling rate of 0.2 Hz.

The patients were instructed to fill in diary cards regarding the time of the meals and beverages, recumbent time, and the exact times of symptoms experienced during the 24-hour study. In addition, they were asked to press a button on the digital data logger at the beginning of each symptom episode. The diary data regarding the onset of symptoms were used primarily as a check of appropriate use of the event button. The time recorded in the diary was used for analysis only when a symptom was recorded in the diary but not with the event button. No restrictions were imposed on food or beverage intake or smoking behavior.

After 24 hours, stored data were retrieved from the data logger by a personal computer and analyzed automatically as described previously.[6] For periods of upright and supine body position and total recording time, the classical reflux variables (percentage of time with pH < 4, the number of episodes with pH < 4, the number of episodes with pH < 4 lasting 5 minutes or more, and the number of pH drops of more than one pH unit while pH > 4) were calculated. The SI (the number of reflux-related symptom episodes divided by the total number of symptom episodes $\times$ 100%) was calculated according to Wiener et al.[1] In addition, the SSI (defined as the number of symptom-associated reflux episodes divided by the total number of reflux episodes $\times$ 100%) was calculated according to Breumelhof et al.[2]

Calculation of the Symptom-Association Probability

The calculation of the symptom-association probability (SAP) consisted of the following computerized procedures. First, the 24-hour esophageal pH signal was divided into consecutive 2-minute periods. Thereafter, all 2-minute periods (on average 704 in number) were evaluated for the occurrence of gastroesophageal reflux. A 2-minute period was considered reflux-positive if either a decrease in pH below 4 lasting 5 seconds or more or a decrease in pH of more than 1 pH unit (within 5 seconds) had occurred during that period. Likewise, all 2-minute periods preceding the onset of symptom episodes were analyzed for the presence of reflux and classified as reflux-positive or reflux-negative. Subsequently, a contingency table was constructed containing four fields: one field containing the number of symptomatic reflux-positive 2-minute periods (S^+R^+), one with the number of asymptomatic reflux-positive 2-minute periods (S^-R^+), one with symptomatic 2-minute periods without reflux events (S^+R^-), and one with the number of asymptomatic 2-minute periods without reflux events (S^-R^-) (Figure 1).

Fisher's Exact Test was then applied to calculate the probability (*P* value) that the observed association between reflux

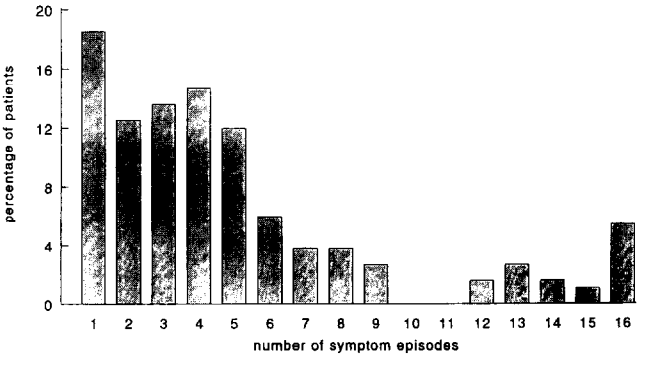

Figure 2. Distribution of the number of symptom episodes experienced during the 24-hour esophageal pH monitoring. The digital recording device allowed for the recording of 16 symptom episodes maximally.

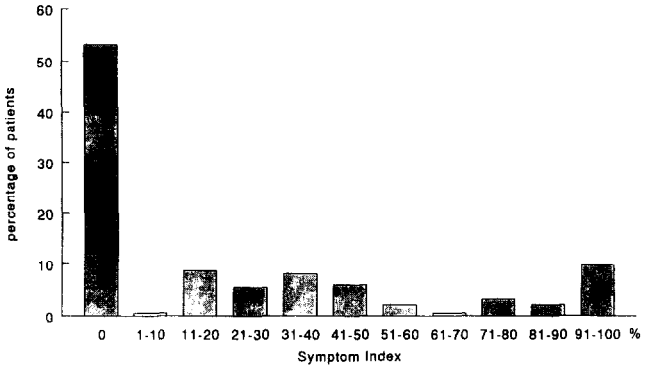

Figure 3. Distribution of the SI (percentage of reflux-related symptom episodes).

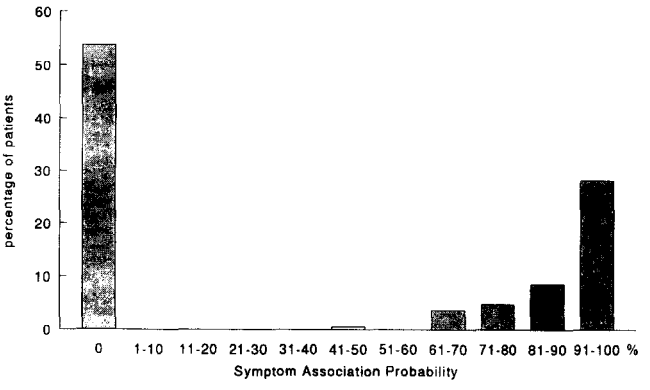

Figure 5. Distribution of the SAP. A value of 0 was assigned to the SAP in those patients in whom symptom episodes coincided significantly less than could be expected if the association between reflux and symptoms occurred by chance.

and symptoms occurred by chance.[7] The SAP was calculated as $(1.0 - P) \times 100\%$. In contrast to the χ^2 test, application of Fisher's Exact Test is also justified in case of low numbers in 1 of the 4 fields.[7] Significant SAP values were only considered when symptom episodes showed more associations with reflux than could be expected by chance, i.e., when the proportion of reflux positivity in the 2-minute periods preceding the symptom events was higher than that in the total recorded 2-minute periods. Significant SAP values (suggesting reflux-induced symptoms) were prevented in patients in whom symptom episodes coincided less often than could be expected to occur by chance. This was done by assigning the value 1 to the P value (resulting in a SAP of 0) when the proportion of reflux-positive presymptom periods was lower than the proportion of reflux-positive periods in general.

Statistical Analysis

In the statistical analysis, the nonparametric correlation coefficient (Spearman's rank, r_s) and the Mann–Whitney U test were used. All tests were performed using the statistical software package of NCSS 5.X series (Dr. Jerry L. Hintze, Kaysville, UT).

Results

During the measurement, the patients experienced 1–16 symptoms (median, 4). As shown in Figure 2, the distribution of the number of symptom episodes experienced during the 24-hour test was skewed in that most patients had only one symptom episode. The distribution of the SI and SSI values among the patients is shown in Figures 3 and 4. In 42 patients (23%), the SI was $\geq$ 50%. An SSI of $\geq$5% was observed in 18 patients (10%). The highest SSI encountered was 15.7%. The SAP ranged from 0% to 100% (Figure 5). A SAP of >95% was observed in 38 patients (21%). Significant correlations were found between SI and SAP (r_s = 0.9629, $P < 0.001$; Figure 6), and between the SSI and SAP (r_s = 0.9666; $P < 0.001$; Figure 7). However, discordance between the SI, SSI, and SAP values was observed in a substantial proportion of the patients. As shown in Table 1, in 13 of the 42 patients with an SI of $\geq$ 50%, the SAP was $\leq$ 95% (SAP, 83.1%–94.2%). In these 13 patients, the number of symptoms experienced during the measurement was significantly lower

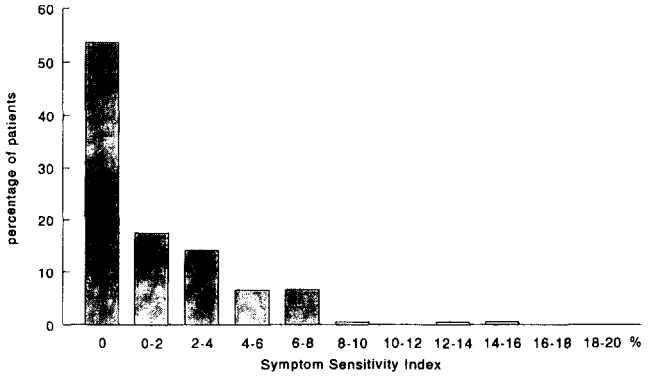

Figure 4. Distribution of the SSI (percentage of symptomatic reflux episodes).

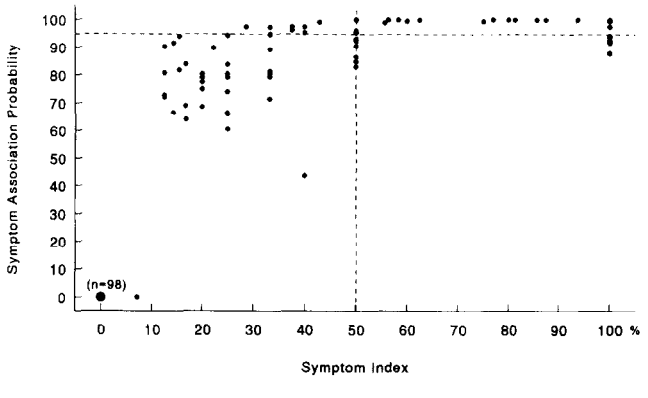

Figure 6. Correlation between the SAP and the SI. The *dotted lines* indicate the thresholds for the SAP (significant if >95%; *horizontal line*) and the SI (positive if $\geq$50%; *vertical line*). Note the discordance between the SI and the SAP in many patients, although the correlation between the SI and the SAP is significant (Spearman's r_s, 0.9629; $P < 0.001$).

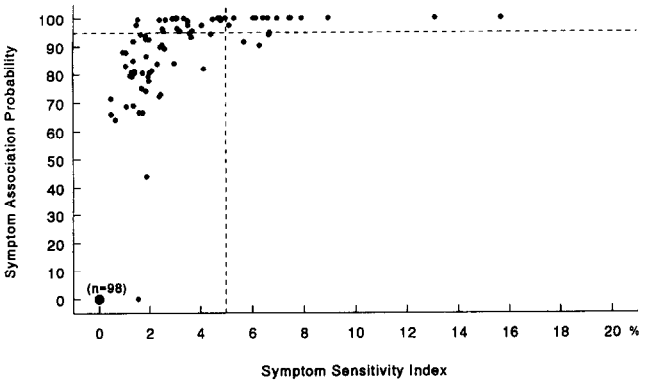

Figure 7. Correlation between the SAP and the SSI. The *dotted lines* indicate the thresholds for the SAP (significant if >95%; *horizontal line*) and the SSI (positive if ≥5%; *vertical line*). Note the discordance between the SSI and the SAP in many patients, although the correlation between the SSI and the SAP is significant (Spearman's r_s, 0.9666; $P < 0.001$).

than in the patients in whom the SI and SAP were in accordance ($P = 0.0001$). As shown in Table 2, 8 of the 141 patients with an SI of <50% (SI, 28.6%–42.9%) had an SAP > 95%. The number of symptoms experienced in these eight patients and the SSI were significantly higher than in the patients with accordance between SI and SAP values ($P < 0.05$ and $P < 0.001$, respectively). In the patients with high SI and low SAP values (Table 1), both the number of symptoms experienced during the test and the SSI were significantly lower as compared with the patients with low SI and high SAP values (Table 2) ($P < 0.0005$ and $P < 0.0005$). No significant differences in acid exposure times were detected between patients with concordance and discordance in SI and SAP values. In 4 of the 18 patients with a high SSI (≥5%), the SAP was ≤ 95% (SAP, 90.3%–94.8%), and in 24 of the 166 patients with an SSI of <5%, the SAP was > 95%. In only 13 patients

Table 1. Patients With High SI (≥50%) and Low SAP (≤95%)

SI	SAP	pH < 4 (%)	No. of symptoms	SSI
50	92.4	11.6	4	2.0
50	86.5	1.8	2	1.9
50	93.2	3.1	2	3.6
50	90.5	6.0	2	2.5
50	83.1	7.4	2	1.1
50	84.9	6.3	2	1.4
100	93.8	7.4	1	1.8
100	87.8	15.0	1	1.1
100	92.6	11.7	1	1.9
100	91.8	7.1	1	1.4
100	91.8	1.7	1	1.4
100	88.0	12.4	1	0.9
100	94.2	4.1	1	1.6

Table 2. Patients With Low SI (<50%) and High SAP (>95%)

SI	SAP	pH < 4 (%)	No. of symptoms	SSI
37.5	97.5	8.7	16	3.5
37.5	96.2	4.4	8	3.1
28.6	97.4	1.2	7	4.1
42.9	99.2	4.7	7	4.8
33.3	97.3	3.1	6	4.0
40	97.4	15.0	5	5.1
40	95.3	8.2	5	3.2
40	95.4	5.1	5	3.7

(7.1%), the SI, SAP, and SSI exceeded the threshold values.

Discussion

In patients with symptoms suggestive of gastroesophageal reflux, especially in those without esophagitis at endoscopy, the analysis of the association in time between symptoms and reflux is of crucial importance.[8] It has been recognized that the absence of pathological reflux does not exclude gastroesophageal reflux as a possible cause of the patient's symptoms.[9,10] Twenty-four–hour esophageal pH monitoring provides the opportunity to analyze the esophageal pH profile for reflux preceding the onset of symptom episodes. In the literature, several attempts have been made to quantify the associations between symptoms and reflux.[1–5] However, comparison between these different parameters is difficult because a gold standard for symptom association in gastroesophageal reflux disease is lacking. Such a comparison can only be based on theoretical grounds.

Currently, the most commonly used parameter is the SI, defined as the number of reflux-related symptom episodes divided by the total number of symptom episodes multiplied by 100%.[1] There are at least three important shortcomings in this parameter, rendering its use of limited value. First, as already mentioned by several investigators,[2,11] the SI does not take into account the total number of reflux episodes. The higher the frequency of gastroesophageal reflux, the greater the chance that a symptom is found to be associated in time with reflux. Therefore, in patients with frequent gastroesophageal reflux who report only one symptom episode during the 24-hour pH monitoring, an SI of 100% may well occur by chance. Second, the cutoff between a positive and a negative SI is arbitrary, and proposed thresholds vary from 25% to 75%.[1,12–14] Third, the fact that the SI expresses the number of reflux-related symptom episodes as a percentage may obscure valuable information. An SI of 100% can express a situation of one reflux-related symptom episode as well as one of dozens of reflux-related symptom episodes.

The latter situation is much more convincing of reflux as a cause of the symptoms. In our study population, 13 patients (7%) had a high SI in combination with a low SAP (Table 1). The number of symptom episodes experienced during the measurement in this group of patients was relatively low, making them more susceptible for random symptom-reflux associations as indicated by the low SAP values. Hence, a low number of symptom episodes during the pH monitoring results in a high rate of false-positives for the SI measurement. In the patients with low SI and high SAP values (Table 2), SSI values were relatively high (indicating that a high percentage of reflux episodes is associated with symptoms), resulting in significant SAP values, although <50% of symptom episodes were associated with reflux events.

In an attempt to overcome the first-mentioned shortcoming, Breumelhof and Smout introduced an additional parameter, the SSI, defined as the percentage of symptom-associated reflux episodes.[2] However, in this parameter, the optimal threshold between normal and abnormal is also unclear, and the index has been criticized.[11] Other more complex methods have been proposed to quantify associations between reflux and symptoms. The binomial formula used by Ghillebert et al.[3] suffers from the shortcoming that the chance of random symptom-reflux associations is improperly defined. The method proposed by Armstrong et al. using the Kolmogorov–Smirnov test[4] has the disadvantage that an incorrect time window is used and that one-sided testing is impossible.

In the SAP, a contingency table is used as proposed by Orr.[11] In this table, the frequency of occurrence of all four possible combinations (asymptomatic and symptomatic 2-minute episodes with and without reflux) is recorded. We have recently shown that in the symptom analysis of 24-hour esophageal pH data, a time window that begins 2 minutes before the onset of the symptom event and ends at its onset provides optimal results.[15] Accordingly, in this study, the total recording period was divided into 2-minute periods that were analyzed for the occurrence of reflux. Using Fisher's Exact Test, significant differences in proportions of reflux-positive 2-minute periods between symptom-positive and symptom-negative 2-minute periods can be detected. Although a statistically significant result does not prove a causal relationship, the SAP provides objective information on the probability that observed associations in time between reflux and symptoms occur by chance. For instance, in the case of a SAP of >95%, the probability that the observed associations occurred by chance is <5%.

In conclusion, we believe that the SAP currently is the best parameter to quantify the temporal association between reflux and symptoms. It uses optimal time windows and provides a measure of strength of the temporal association that takes all relevant variables into account. We advocate the use of the SAP in future studies in which associations between reflux events and symptom episodes are important and the incorporation of the SAP in the analysis of clinical 24-hour esophageal pH studies.

References

1. Wiener GJ, Richter JE, Copper JB, Wu WC, Castell DO. The symptom index: a clinically important parameter of ambulatory 24-hour esophageal pH monitoring. Am J Gastroenterol 1988; 83:358–361.
2. Breumelhof R, Smout AJPM. The symptom sensitivity index: a valuable additional parameter in 24-hour esophageal pH recording. Am J Gastroenterol 1991;86:160–164.
3. Ghillebert G, Janssens J, Vantrappen G, Nevens F, Piessens J. Ambulatory 24-hour intraoesophageal pH and pressure recordings v provocation tests in the diagnosis of chest pain of oesophageal origin. Gut 1990;31:738–744.
4. Armstrong D, Emde C, Inauen W, Blum AL. Diagnostic assessment of gastroesophageal reflux disease: what is possible vs. what is practical? Hepatogastroenterology 1992;39(suppl 1):3–13.
5. Johnston BT, McFarland RJ, Collins JSA, Love AHG. The binomial symptom index: correlating esophageal symptoms with acid reflux (abstr). Gastroenterology 1992;102:A92.
6. Smout AJPM, Breedijk M, van der Zouw C, Akkermans LMA. Physiological gastroesophageal reflux and esophageal motor activity studied with a new system for 24-hour recording and automated analysis. Dig Dis Sci 1989;34:372–378.
7. Lentner C, Diem K, Seldrup J, eds. Geigy scientific tables. Volume 2. 8th ed. Basel, Switzerland: Ciba-Geigy, 1982:227.
8. Weusten BLAM, Akkermans LMA, vanBerge-Henegouwen GP, Smout AJPM. Ambulatory combined oesophageal pressure and pH monitoring: relationships between pathological reflux, oesophageal dysmotility and symptoms of oesophageal dysfunction. Eur J Gastroenterol Hepatol 1993;5:1055–1060.
9. Vantrappen G, Janssens J, Ghillebert G. The irritable oesophagus—a frequent cause of angina-like pain. Lancet 1987;2: 1232–1234.
10. Howard PJ, Maher L, Pryde A, Heading RC. Symptomatic gastro-oesophageal reflux, abnormal oesophageal acid exposure, and mucosal acid sensitivity are three separate, though related, aspects of gastro-oesophageal reflux disease. Gut 1991;32:128–132.
11. Orr WC. Noncardiac chest pain: conundrum of cause and effect. Am J Gastroenterol 1991;86:1548–1550.
12. Hewson EG, Sinclair JW, Dalton CB, Richter JE. Twenty-four–hour esophageal pH monitoring: the most useful test for evaluating noncardiac chest pain. Am J Med 1991;90:576–583.
13. Johnsson F, Joelsson B, Gudmundsson K, Greiff L. Symptoms and endoscopic findings in the diagnosis of gastroesophageal reflux disease. Scand J Gastroenterol 1987;22:714–718.
14. Singh S, Richter JE, Bradley LA, Haile JM. The symptom index. Differential usefulness in suspected acid-related complaints of heartburn and chest pain. Dig Dis Sci 1993;38:1402–1408.
15. Lam HGT, Breumelhof R, Roelofs JMM, vanBerge-Henegouwen GP, Smout AJPM. What is the optimal time window in symptom analysis of 24-hour esophageal pressure and pH data? Dig Dis Sci 1994;39:402–409.

Received April 20, 1994. Accepted July 25, 1994.

Address requests for reprints to: Bas L. A. M. Weusten, M.D., Department of Gastroenterology, University Hospital Utrecht, P. O. Box 85500, 3508 GA Utrecht, The Netherlands. Fax: (31) 30-507371.

Supported in part (to B.L.A.M.W.) by Astra Pharmaceuticals, Rijswijk, The Netherlands.