Correlation between manometric and pH tests for gastro-oesophageal reflux

C. STANCIU, R. C. HOARE, AND J. R. BENNETT

From the Gastrointestinal Unit, Hull Royal Infirmary, Kingston upon Hull

SUMMARY A series of tests, including gastro-oesophageal sphincter pressure measurement, short-term pH tests, and 15-hour overnight oesophageal pH recording were applied to 42 normal subjects and 214 patients with typical reflux symptoms. The results were compared by multivariate discriminant analysis. Sphincter pressure measurements misclassified 32%, stressed provocative manoeuvres 14.5%, and the best single discriminator from the overnight pH study was time below pH 5, which misclassified 13%. However, a combination of the number of reflux episodes in 15 hours with their mean duration reduced misclassification to 8.8%. Using this function, a boundary between normal and reflux can be drawn, and the degree of abnormality can be expressed visually as well as numerically.

This study was made to establish which of several specific diagnostic tests best matched the clinical diagnosis of gastro-oesophageal reflux (GOR) in a group of patients with characteristic 'reflux' symptoms, and to determine what correlation existed between different tests.

Methods

NORMAL SUBJECTS

There were 42 subjects (26 male, 16 female), aged 18 to 51 years (mean 32.8 years) with no symptoms of upper digestive tract disease. This group was composed of 26 volunteers and 16 subjects with the irritable bowel syndrome. The tests were explained and informed permission obtained.

PATIENTS

There were 214 patients (116 male, 98 female) aged 17 years to 78 years (mean 47.2 years) with typical symptoms of GOR (heartburn and acid regurgitation related to posture and meals, and discomfort on drinking hot fluids).

MANOMETRY

Intraluminal pressure was measured with three polyvinyl catheters (1.2 mm ID) with a single side-hole, 5 cm apart. The catheters were continuously perfused with distilled water at 0.8 ml/min. With the patient supine the tips were withdrawn twice through

the gastro-oesophageal sphincter (GOS) in steps of 0.5 cm. Six measurements of resting end-expiratory sphincter pressure were thus made in each patient, taking the mean end-expiratory fundal pressure as zero.

SHORT-TERM PH TESTS

Intra-oesophageal pH was measured by a Radiometer combined glass and reference pH electrode (GK 282 C) with Radiometer pH meter (Type PHM 26) recording on a direct writing recorder (Devices M.19 or George Washington oscillograph) at a constant speed.

The glass/calomel electrode system was filled with saturated potassium chloride solution, care being taken to avoid air bubbles and crystals which could cause artificial 'spikes'. The calomel electrode reservoir was placed at a higher level than the glass electrode to ensure a constant flow of potassium chloride solution through the porous plug. The electrode was standardised in buffer solutions of pH 4.0 to 6.5 before and after each test.

Short-term pH measurements were performed during the manometric studies. The pH electrode was either attached to the manometric catheters at the level of the distal opening, or passed separately, its position being measured against landmarks known from the manometric studies. Once the pH electrode (and manometric catheters) were placed into the stomach, tests for gastro-oesophageal reflux were carried out.

Received for publication 20 December 1976

Test 1

This test involved withdrawal of the pH electrode (and catheters) from the stomach past the gastro-oesophageal junction with the patient supine.

Test 2

This was a second withdrawal of the assembly after the addition of 250-300 ml 0.1 N HCl to the stomach.

For each of these withdrawal tests the distance in cm above the GOS at which the pH reached 4 was measured.

Test 3

The pH electrode was positioned 5 cm above the GOS and the patient was instructed to carry out a series of manoeuvres—deep breathing, Valsalva, Muller, coughing—while lying supine, on the right side and on the left side, and supine with 15° head-down tilt. The time required to perform these 16 manoeuvres was less than 10 minutes. A 'reflux episode' was defined as any occasion on which oesophageal pH dropped by two pH units or more from the base line 6.0 to 6.5. If reflux was frequent, distilled water was infused through the distal catheter to wash the tip of the electrode. The number of reflux episodes during the 16 manoeuvres was counted.

LONG-TERM PH TEST

The pH electrode, standardised before and after each test, was passed through the nose (or orally if the nasal route was uncomfortable) into the stomach to ascertain that the gastric pH was acid, and was then withdrawn to 5 cm above the GOS, the exact distance being known from the previous manometric studies. (In some patients the electrode position was also checked radiologically.) Once the pH electrode was positioned correctly, it was firmly anchored to the patient's cheek with adhesive tape, and continuous records were obtained for 15 hours (6 pm to 9 am). The patients, lying comfortably in bed, were allowed to eat, smoke, or talk, and most of them slept for six to eight hours, but were given no hypnotic drugs. Most of the patients lay flat with one or two pillows under their head, but were allowed to adopt the posture in which they were normally accustomed to sleep. Times of meals, and any symptoms occurring during the test, were noted on the record chart by trained staff. A 'reflux episode' was defined as any occasion on which oesophageal pH dropped by two pH units or more from the base-line of 6.0 to 6.5.

A reflux episode lasting less than one minute was defined as a 'spike'. The analysis of pH tracings was done in three ways:

1. *Total number of reflux episodes* This was calcu-

lated by counting the number of reflux episodes occurring over the total recording time.

2. *Mean duration of reflux episodes* Obtained by dividing the total duration of reflux episodes by their number. (Because the exact duration of a 'spike' could not be measured, each one was arbitrarily considered as lasting 30 seconds.)
3. *Percentage time with oesophageal pH less than 5, 4 and 3* The proportions of the period studied during which reflux episodes brought the pH below 5, 4, and 3 were each calculated as a percentage of total recording time.

STATISTICAL METHODS

A stepwise discriminant analysis (Kendall, 1957) programme (BMD070M) was used to compare GOS pressure, short-term pH tests, and long-term pH test. The analysis computes for each group a discriminant function which is a linear function of a subset of the recorded variables. For each group we have the discriminant functions (1) and (2) as follows:

for 'normal' group. $Y_N = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_n x_n$ (1)

for 'reflux' group. $Y_R = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n$ (2)

Where $a_0 a_1 \dots a_n$ and $b_0 b_1 \dots b_n$ are constants and $x_1 x_2 \dots x_n$ are discriminatory variables.

A subject is classified as 'normal' if $Y_N > Y_R$ and as 'reflux' if $Y_R > Y_N$.

Thus, the intersection of these planes gives a boundary region which can be used for classification of subjects. The boundary gives a 50% probability of classification into either group, and the probability of misclassification decreases on either side of this boundary.

In the analysis, logarithmic and square root transformations were used as required to normalise the data.

Not all variables were recorded for every subject, so that comparisons were made using subsets of the 256 subjects.

Results**GOS PRESSURE**

The mean sphincter pressure in 29 normal subjects was $17.8 \pm \text{SD } 6.39$ cm water, compared with $12.4 \pm \text{SD } 8.24$ in 210 patients with reflux symptoms ($P < 0.001$).

The discriminant analysis shows this test alone would 'misclassify' 31.8% of the subjects (Table 1).

SHORT-TERM PH TESTS

Results are based on 23 normal subjects and 108 reflux patients.

Table 1 *Results of discriminant analysis of gastro-oesophageal sphincter pressures*

Clinical group	Analysis group		Percentage misclassified
	Normal	Reflux	
Normal	18	11	37.9
Reflux	65	145	31.0
Overall			31.8

Calculated boundary value between normal subjects and reflux patients is 15.09 cm water.

Of the three short-term pH tests, the best discriminator between the two groups proved to be the provocative manoeuvres, which misclassified 14.5% (Table 2).

Table 2 *Discriminant analysis of provocative manoeuvre results*

Clinical group	Analysis group		Percentage misclassified
	Normal	Reflux	
Normal	16	7	30.4
Reflux	12	96	11.1
Overall			14.5

Calculated boundary value between normal subjects and reflux patients is 1.7 manoeuvres.

The boundary value between normal and reflux was calculated as 1.7. Since only integers can be measured in practice, two or more reflux episodes during manoeuvres are required to classify a patient as 'reflux'.

LONG-TERM (OVERNIGHT) PH STUDIES

Values for 29 normal subjects and 210 reflux patients are given in Table 5. There are generally good corre-

lations between all the measured variables in these recordings¹, but a discriminant analysis (Table 3) shows the best single discriminator to be percentage time below pH 5. By this test alone 13% of the group would be misclassified as normal subjects, though none of them gave pH values in the reflux range.

Table 3 *Discriminant analysis of percentage time less than pH 5 in 15-hour recordings*

Clinical group	Analysis group		Percentage misclassified
	Normal	Reflux	
Normal	29	0	0
Reflux	31	179	14.8
Overall			13.0

Calculated boundary value between normal subjects and reflux patients is 2.43.

Table 4 *Discriminant analysis of mean duration of reflux episodes combined with number of reflux episodes in 15-hour recording*

Clinical group	Analysis group		Percentage misclassified
	Normal	Reflux	
Normal	29	0	0
Reflux	21	189	10.0
Overall			8.8

The number of reflux episodes and their mean duration combined (Table 4) give the best two variable-discrimination and misclassification is reduced to 8.8% (Figure).

¹Statistical details may be obtained from J.R.B.

Table 5 *Published results of long pH tests*

Author	Duration of study (hr)	No. and type of subjects	% time pH 4	% time pH 5	Mean Duration	No. of reflux episodes per hour
Spencer (1969)	18	11 N 15 R	2.2 - 5.3 24.5 - 26.5	—	—	—
Woodward (1970)	15	33 R	0.5 - 30.6	2.6 - 66	—	—
Habibulla <i>et al.</i> (1971)	10	5 N 30 R	0 15.9 - 19.2	0 23 - 33	0 2.7 - 13.6	0 0.7 - 2.2
Rendall (1972)	6	11 N	0 - 6	—	—	—
Lichter (1974)	13½	41 R	0 - 17.3	—	—	—
DeMeester <i>et al.</i> (1974)	24	15 N 16 R	1.5 ± 1.4 11.1 ± 4.6	—	—	0.86 ± 0.61 2.6 ± 1.08
Boesby (1975)	12	26 N	0 - 2.91	0 - 6.63	—	—
Present paper	15	42 N 210 R	0.17 ± 0.23 7.0 ± 7.7	0.25 ± 0.41 13.3 ± 12.6	0.54 ± 0.48 14.7 ± 19.9	0.18 ± 0.16 0.84 ± 0.71

N: normal subject. R: reflux.

Figure Boundary between normal subjects and reflux patients using number of reflux episodes and their mean duration in 15-hour oesophageal pH recordings.

Discussion

This study has been confined to a comparison of tests which measure acid reflux directly, together with an assessment of the strength of the sphincter barrier to reflux. Tests which suggest indirectly that reflux occurs have not been included.

There is ample published evidence that all the tests we used give some indication of the likelihood of pathological gastro-oesophageal reflux. But in the practice of clinical medicine a test is valuable only if it reliably discriminates between two groups of subjects—in this case normal subjects and those with gastro-oesophageal reflux. When the discrimination of the tests is poor, false-negative and false-positive results are common and a point is reached at which the test is no better than the clinician's unaided acumen; at this point the test is valueless. Our attempt has been to define the discriminatory accuracy of each test.

It has been long known that there is a relationship between a low GOS pressure and reflux (Atkinson *et al.*, 1957), but, although some workers have found a good correlation (Pope, 1967; Winans and Harris, 1967; Cohen and Harris, 1971), others have found too much overlap between the groups to satisfy the needs of a diagnostic test (Benz *et al.*, 1972; Haddad, 1970). Technical details of the method used to measure sphincter pressure are obviously important. We used a standard technique for measuring unstressed sphincter squeeze but, like others, found it a poor test for discriminating between those who reflux and those who do not.

Direct measurement of pH in the oesophagus has an obvious application in a disorder thought to be due largely to the effects of acid and pepsin on the oesophageal mucosa. Simply withdrawing the electrode from stomach to oesophagus and observing the rate of pH change in the lower oesophagus may be helpful (Tuttle and Grossman, 1958; Besancon *et al.*, 1962) but other workers, using techniques differing in detail but not in principle, have found less satisfactory results (Piccone *et al.*, 1965; Kantrowitz *et al.*, 1969; Skinner and Booth, 1970). We found that in practice it was hard to judge the 'end-point'—the exact distance above the sphincter at which the pH of 4 was reached, and that analysis of the data showed poor discrimination. This is not surprising, because reflux is rarely a continuous process but almost always intermittent (Edwards, 1973).

The pH electrode will undoubtedly detect a spurt of acid entering the oesophagus, and its use to detect reflux induced by standard manoeuvres designed to stress the GOS has shown promise (Piccone *et al.* 1965; Kantrowitz *et al.*, 1969; Haddad, 1970; Skinner and Booth, 1970; Jahadi and Chandler, 1972). We found this to be the most discriminating of our 'short-term' pH tests. A practical disadvantage is that many patients find the test uncomfortable or fail to perform the manoeuvres successfully.

Measurement of lower oesophageal pH over long periods has the attraction of appearing to be 'physiological'—the measurement should reflect what is daily happening in the oesophagus in the patient's normal circumstances. Such studies (Table 5) have shown significant differences between normal subjects and those with GOR (Spencer, 1969; Habibulla *et al.* 1971; DeMeester *et al.*, 1974). The clarity of distinction between the two groups has been less well clarified, and if the 'grey area' is large, the test is valueless as a clinical tool. Patrick (1970), for example, found considerable overlap between reflux patients and control subjects. Discriminant analysis enables the difference between the groups to be expressed accurately, and our measurements show that the proportion of the test time during which the lower oesophagus is at an acid pH—be it 3, 4, or 5—is a good indicator of the diagnosis of gastro-oesophageal reflux, and that percentage time below pH 5 is best. The mean percentage time below pH 5 in normal subjects was $0.26 \pm \text{SD } 0.40$. Discrimination is improved considerably if the number and mean duration of the reflux episodes is used instead, and, when both these variables are included, over 90% of subjects are accurately classified (Figure). The mean duration of episodes is partly dependent upon acid clearing from the oesophagus (Stanciu and Bennett, 1974b), not in itself a direct

measure of reflux, though probably an important element in determining the likelihood of symptoms and of oesophagitis.

This evaluation has been based on an entirely clinical differentiation between normal subjects and patients with gastro-oesophageal reflux. Error has been minimised by choosing as 'normal' those with no dyspeptic symptoms at all, and as 'reflux patients' only those with unequivocally typical symptoms. What is now required is the application of these tests to numbers of patients with more confusing clinical problems in the hope of establishing accurately their role in the clinical management. The apparatus required for such pH monitoring is relatively simple, and the results easy to calculate. It should prove a valuable aid in the diagnosis of reflux, as it already has in assessing provocative factors and the effects of therapy (Stanciu and Bennett, 1972, 1974a).

The work was done while C. S. was the holder of a Research Fellowship of the Leeds Regional Hospital Board. Skilled technical assistance from Mrs L. Brocklesby was invaluable, as was the secretarial help of Mrs M. Thomson. The Devices multi-channel recorder was on generous loan from Reckitt and Colman Limited.

References

- Atkinson, M., Edwards, D. A. W., Honour, A. J., and Rowlands, E. N. (1957). Comparison of cardiac and pyloric sphincters. *Lancet*, **2**, 918-922.
- Besancon, F., Baujat, J. P., and Debray, C. (1962). Le reflux gastro-oesophagien: etude pH-graphique et électromanographique. *Semaine des Hôpitaux de Paris*, **38**, 1569-1576.
- Benz, L. J., Hootkin, L. A., Margulies, S., Donner, M. W., Cauthorne, R. T., and Hendrix, T. R. (1972). A comparison of clinical measurements of gastroesophageal influx. *Gastroenterology*, **62**, 1-5.
- Boesby, S. (1975). Gastro-oesophageal acid reflux and sphincter pressure in normal human subjects. *Scandinavian Journal of Gastroenterology*, **10**, 731-736.
- Cohen, S., and Harris, L. D. (1971). Does hiatus hernia affect competence of the gastroesophageal sphincter? *New England Journal of Medicine*, **284**, 1053-1056.
- DeMeester, T. R., Johnson, L. F., and Kent, A. H. (1974). Evaluation of current operations for the prevention of gastroesophageal reflux. *Annals of Surgery*, **180**, 511-525.
- Edwards, D. A. W. (1973). Symposium on gastro-oesophageal reflux and its complications. *Gut*, **14**, 233-253.
- Habibulla, K. S., Ammann, J. F., and Leigh Collis, J. (1971). Effect of posture on hiatus hernia as studied by oesophageal pH measurement. *Thorax*, **26**, 689-695.
- Haddad, J. K. (1970). Relation of gastroesophageal reflux to yield sphincter pressures. *Gastroenterology*, **58**, 175-184.
- Jahadi, M. R., and Chandler, J. P. (1972). Detecting gastro-oesophageal reflux by pH recording and acid reflux test. *American Surgeon*, **38**, 281-284.
- Kantrowitz, P. A., Corson, J. G., Fleischli, D. J., and Skinner, D. B. (1969). Measurement of gastroesophageal reflux. *Gastroenterology*, **56**, 666-674.
- Kendall, M. G. (1957). *A Course in Multivariate Analysis*, p. 144. Griffin: London.
- Lichter, I. (1974). Measurement of gastro-oesophageal acid reflux; its significance in hiatus hernia. *British Journal of Surgery*, **61**, 253-258.
- Patrick, F. (1970). Investigation of gastroesophageal reflux in various positions with a two-lumen pH electrode. *Gut*, **11**, 659-667.
- Piccone, V. A., Gutelius, J. R., and McCorriston, J. R. (1965). A multiphased esophageal pH test for gastro-oesophageal reflux. *Surgery*, **57**, 638-646.
- Pope, C. E. (1967). A dynamic test of sphincter strength: its application to the lower esophageal sphincter. *Gastroenterology*, **52**, 779-786.
- Rendall, M. (1972). Hiatus hernia and oesophageal reflux. *Proceedings of the Royal Society of Medicine*, **65**, 33-35.
- Skinner, D. B., and Booth, D. J. (1970). Assessment of distal esophageal function in patients with hiatus hernia and/or gastroesophageal reflux. *Annals of Surgery*, **172**, 627-637.
- Spencer, J. (1969). Prolonged pH recording in the study of gastro-oesophageal reflux. *British Journal of Surgery*, **56**, 912-914.
- Stanciu, C., and Bennett, J. R. (1972). Smoking and gastro-oesophageal reflux. *British Medical Journal*, **3**, 793-795.
- Stanciu, C., and Bennett, J. R. (1974a). Alginate/antacid in the reduction of gastro-oesophageal reflux. *Lancet*, **1**, 109-111.
- Stanciu, C., and Bennett, J. R. (1974b). Oesophageal acid clearing: one factor in the production of reflux oesophagitis. *Gut*, **15**, 852-857.
- Tuttle, S. G., and Grossman, M. T. (1958). Detection of gastro-oesophageal reflux by simultaneous measurement of intraluminal pressure and pH. *Proceedings of the Society for Experimental Biology and Medicine*, **98**, 225-227.
- Winans, C. S., and Harris, L. D. (1967). Quantitation of lower esophageal sphincter competence. *Gastroenterology*, **52**, 773-778.
- Woodward, D. A. K. (1970). Response of the gullet to gastric reflux in patients with hiatus hernia and oesophagitis. *Thorax*, **25**, 459-464.