Leading article

Twenty four hour intragastric acidity analysis for the future

Studies of 24 hour intragastric acidity have been useful in the investigation of effects of diet,[1][2] and drugs on gastric acidity in ulcer patients[3-13] and normal volunteers.[14-17] Indeed, such studies now seem mandatory in the development of new antisecretory drugs.[13][17-19] The obvious advantage of the technique is that acidity is measured during relatively long periods of time under conditions which approximate to daily living. These conditions are easily reproduced and thus comparative studies are possible. The technique is labour intensive, however, and has some disadvantages. During the day only acidity is measured, as it is not possible to assess the volume of gastric secretion. Duodenogastric reflux almost certainly occurs, but cannot be estimated. Complete collections of gastric juice are sometimes made at night, to allow assessment of nocturnal secretion. Nocturnal acidity assessed in this way, however, poorly reflects 'real life' acidity because of continual aspiration of gastric contents. When small numbers of replicate studies have been done the results have been reasonable[1][2][8][14] but the exact reproducibility and the magnitude of error implicit in the studies has not been accurately determined.

Unfortunately the relationship of acidity to the aetiology, clinical manifestations and management of peptic ulcer disease is unclear. The results from acidity measurements may have little practical bearing on care of patients. If one assumes, as most do, that acid is detrimental to the duodenal mucosa (at least in some individuals), then decreasing the duodenal acid load would seem logical. It is not known whether small volumes of high acidity or large volumes of lower acidity are more damaging. It is not even clear whether the magnitude or timing of inhibition of acidity, or secretion, is of importance in clinical terms. Indeed, inhibition of 24 hour intragastric acidity is not a prerequisite for clinical efficacy with antisecretory agents. Nevertheless, it might be anticipated that drugs with similar effects on acidity would have similar clinical activity. On the other hand, drugs which produce virtual anacidity – for example, omeprazole – seem to accelerate duodenal ulcer healing compared with H_2-receptor antagonists.[20][21]

Although one may question the relevance of the measurement of 24 hour intragastric acidity to clinical management, useful pharmacological comparisons between doses and drugs can be made. It has even been suggested that studies of nocturnal gastric acidity might identify those patients who respond poorly to H_2-receptor antagonists.[22] When drugs are used in the longer term, knowledge of the 24 hour acidity profile might alert one to potential problems related to hypoacidity such as bacterial colonisation and hypergastrinaemia (which could be predicted with omeprazole[23][11] but are unlikely with conventional H_2-receptor antago-

1

nists.[24-26] Although 24 hour gastric acidity measurement has generally been considered useful, different methods have been used to analyse these data and there has been no universally accepted method. With an increasing interest in the use of the technique, the time has probably come to unify the analyses and thus allow worthwhile between study comparisons to be made.

Assessment of gastric acidity

INDIVIDUALS
In most studies gastric contents are aspirated through a nasogastric tube and the pH of each sample measured *in vitro* using a suitably calibrated glass electrode. In some more recent investigations pH has been measured continuously *in vivo* by intragastric electrode,[17 27 30] but in all cases the datum recorded is pH. This measurement represents the negative logarithm of the hydrogen ion activity (which is actually what the electrode senses) and relates closely, but not exactly to hydrogen ion concentration.[28] (Hydrogen ion concentration only equals activity when low and is generally greater than hydrogen ion activity at the pH of gastric juice.) Coefficients to calculate the hydrogen ion concentration from gastric pH have been devised but are not appropriate when gastric contents include food.[28] Gastric acidity can therefore be described as pH or as hydrogen ion activity expressed in mmol/l calculated back from the formula:

$$pH = 1/\log H^+ \times 1000$$

The existence of these two acidity measures has resulted in the first of the unresolved problems of analysis. Although pH is convenient and well recognised, the logarithmic relationship is frequently ignored and linear plots of pH give distorted representations of acidity which can impede proper interpretation. Figure 1 shows a 24 hour acidity profile from a patient with a duodenal ulcer taking placebo (from 7). The same data are shown as pH and as H^+ activity on a linear scale: although the general pattern of acidity is similar no point coincides. The fluctuations of acidity are best shown by H^+ activity when acidity is high and are exaggerated by pH at low acidity. Hydrogen ion activity is effectively unmeasurable above pH 5·0 (0·01 mmol/l) and therefore poorly reflects changes of acidity above this level. The impact of these two different methods of describing acidity
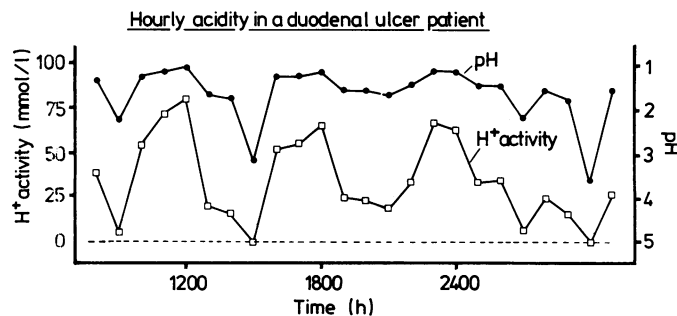

Fig. 1 *Hourly acidity over 24 hours in one duodenal ulcer patient taking placebo. Acidity is shown as H^+ activity (□) on left axis and as pH (●) on right axis.*

becomes clearer when an attempt is made to consider a number of data points together. Using the data in Figure 1, mean 24 h pH is 1·67 and mean 24 h H^+ activity 34·7 mmol/l. These two, commonly quoted, values should be equivalent as they are derived from the same data, but they differ by nearly 40% if the mean pH is converted to H^+ activity (pH 1·67=21·3 mmol/l). There is serious potential for confusion if comparisons are made between studies using these two methods of analysis. Further confusion arises where gastric juice is titrated to calculate hydrogen ion concentration.

GROUPS

Studies on individuals are rarely acceptable and experiments are generally done on groups, which introduces the problem of how to combine data from different subjects. There are three broad choices: all the individual data, or mean data (arithmetic or geometric), or median data can be shown. Where individual observations are not used, the range of observations can be shown as standard deviation, standard error or as a specified percentile range. Quantitative differences are seen when means or medians are used, but qualitatively the variations are usually small provided data are not notably skewed. Figures 2 and 3 show data from 10 duodenal ulcer patients receiving placebo and ranitidine 150 mg twice daily (from 7). The overall impression of drug effect is similar in either plot but substantial differences are seen (particularly during the night). These may have implications for comparisons with other agents, other studies, for marketing purposes and possibly for choice of 'appropriate' dosage.

Until recently means and standard errors have usually been published, but this is wrong. There is a strong statistical argument in favour of the use of medians and ranges, because 24 hour acidity data is rarely normally distributed either as pH or H^+ activity. If a plot of all the measurements used above is made, it is clear that the data are skewed (Fig. 4). The degree of 'skewness' can be calculated[29] but is easily visualised in frequency distribution plots (Fig. 4). If H^+ activity values are shown they too are skewed towards zero with ranitidine and towards high levels with placebo. The use of medians and ranges is more appropriate mathematically for these data. A single 'wild' pH measurement of 1·00 (100 mmol/l) will greatly weight the mean of nine other values of pH 3·00 (1·0 mmol/l) but
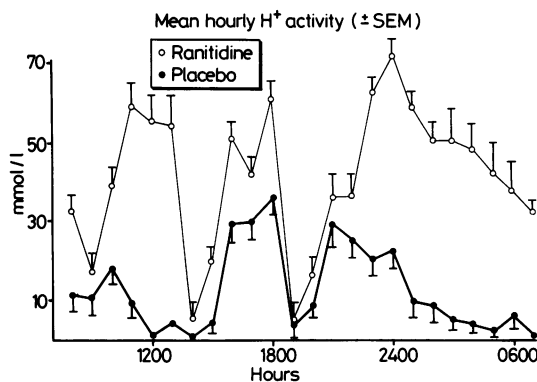

Fig. 2 *Mean hourly H^+ activity (±SEM) in 10 duodenal ulcer patients receiving placebo and ranitidine (replotted from 7 with permission).*

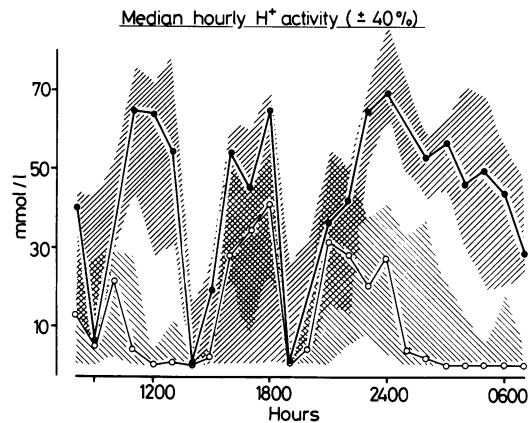Median hourly H⁺ activity (± 40%)

Fig. 3   *Median hourly acidity (range 10th–90th percentiles) in 10 duodenal ulcer patients receiving placebo and ranitidine. (Data from 7).*

will not alter the median value. It is also important to look closely at the ranges rather than to follow the line of the median or mean. For example in Figure 3, ranitidine use decreased median acidity to zero between 0200 and 0700, but the range shows that there were at least some patients with substantial acidity. The span of the ranges shown should be adjusted in line with the size of the sample. It should be large with data from a small number of subjects while the minimum interquartile range would suffice for larger groups. In Figure 3 the range covers 80% of the measurements showing data between the 10th and 90th centiles. By contrast, in the arithmetic mean graph (Fig. 2) ranitidine appears less effective between 0200 and 0700 but the standard errors show that there were patients with negligible acidity. While studies continue to be analysed in different ways great care must be taken when making comparisons to ensure like is compared with like. The onus is on authors to adequately explain their methods of calculation.

OVERALL 24 HOUR MEASUREMENTS

Investigations of 24 hour intragastric acidity generally compare the effects of one or more treatments with placebo or with no treatment. The effect of ranitidine on gastric acidity throughout the study period is easily visualised from either of the graphs (Figs. 2 and 3). When there are numerous
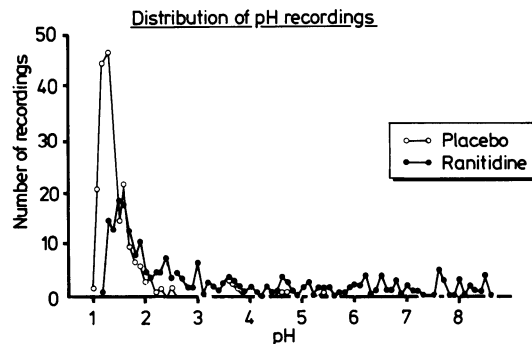
Distribution of pH recordings

Fig. 4   *Distribution of pH recordings from 10 duodenal ulcer patients receiving placebo and ranitidine. (Data from 7).*

comparisons, however, graphs of this type become dificcult to disentangle and numerical values are traditionally calculated to assess all, or part of the data. These values are often presented as a percentage of each other, allowing an assessment of percentage inhibition of acidity. The magnitude of such percentage change not only varies with the method of calculation (from means or medians) but also with the units used (pH or H^+ activity). The expression of all the data as one value is logical only if this has some clinical, pharmacological or physiological importance. An overall measure of 24 hour acidity should therefore be recognised for what it is, a mathematical simplification of data. It is probably irrelevant whether this value of pH, H^+ activity, or even area under the curve is calculated as a median or mean. Because the magnitude of observed changes varies with the different calculations, however, one must know how these values are arrived at. In some studies measurements have been made at half hourly intervals during the day and less often at night and overall measures would therefore be biased towards daytime acidity, unless account is taken of discrepancies in sampling intervals. Measures of mean or median 24 hour acidity cannot be compared between studies unless the sampling and analytical techniques as well as timing and frequency of meals are identical. This problem will increase as more use is made of the newer reliable intragastric electrodes which can measure acidity continuously.[17] [30] Computerised analysis of these records is labour saving, but programs need to be written to take account of the difficulties discussed above. At present, such analysis should be broad based and at least allow calculation of hourly or half-hourly median hydrogen ion activity with appropriate ranges. (Means may be available for those who dispute the value of the median.) The percentage of time with an intragastric pH of any level could be computed and the study period analysed in previously defined time intervals where relevant.

It is clear that different analytical methods can exaggerate (or underrate) an observed change in acidity. The logarithmic relationship together with the use of averages produces different curves and the effect of a drug is, seemingly, substantially smaller if hydrogen ion activity is considered (Fig. 5). If median data representing the same drug effect is shown, however,
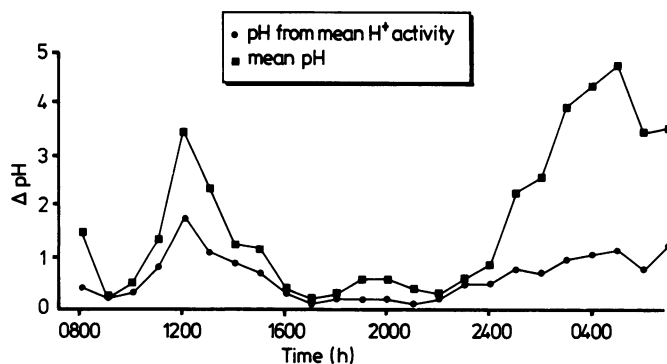

Fig. 5 *Mean change of pH ($\triangle$ pH from placebo) in 10 duodenal ulcer patients receiving ranitidine. Drug effect appears greater if mean pH (■) is considered than if pH derived from mean H^+ activity (●) is used. (Data from 7).*

there is minimal difference between analyses (Fig. 6). Where inhibition of acidity generally results in pH values above 5·0, conversion to hydrogen ion activity becomes meaningless because of the logarithmic relationship. Nevertheless, changes in pH above this level might be considered clinically or pharmacologically important. Frequency distribution curves of pH (Fig. 7) enable such differences to be shown and offer reasonable additional information in publications.

SIGNIFICANCE TESTING

Where data are not normally distributed, parametric statistics are inappropriate. The hourly measurements are also related to each other rather than being independent measures and statistical comparisons from hour to hour are therefore also incorrect.[31] Generally, significance testing has been applied to overall 24 hour values using individual differences or to differences in area under the curve by parametric tests. Paired Student's *t* tests are even more inappropriate because more than one comparison is often made. Residual variance assessed by analysis of variance for significance testing can help, but these parametric methods should probably be replaced by non-parametric tests. The Wilcoxson's signed rank test is simple to apply and will satisfactorily allow between group comparisons when there are only two, but non-parametric analysis of variance should be used when more comparisons are made (Friedman or Kruskal-Wallis).[32] As these tests all require the formation of individual 24 hour values any time related difference is obscured. To circumvent this problem it is reasonable to split the 24 hour period into logical predetermined sections which depend on the aims of the study. For example, if the aim of the investigation were to assess the duration of drug effect it would be reasonable to separate the study into pre- and postadministration periods which could be analysed separately. If this manoeuvre is done, however, the level at which 'significance' is accepted should be adjusted to take account of multiple significance testing. If the data do happen to be normally distributed (for example if hypoacidity due to one potent agent is compared only with hypoacidity due to another), simple direct parametric tests would seem reasonable, if unnecessary.

Occasionally, a set goal might be identified, such as maintaining intragastric pH at an arbitrary level. The success or failure of such a treatment would best be displayed using a frequency distribution of all
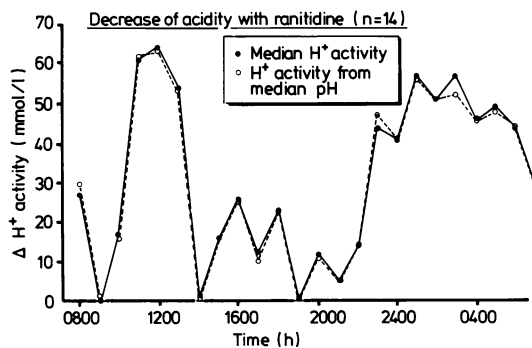

Fig. 6 *Median change in H^+ activity ($\triangle$ H^+ from placebo) in 10 duodenal ulcer patients receiving ranitidine. All values are negative and describe decreased acidity. The drug effect is similar irrespective of the mode of analysis (median pH $\bigcirc$; median H^+ activity $\bullet$). (Data from 7).*
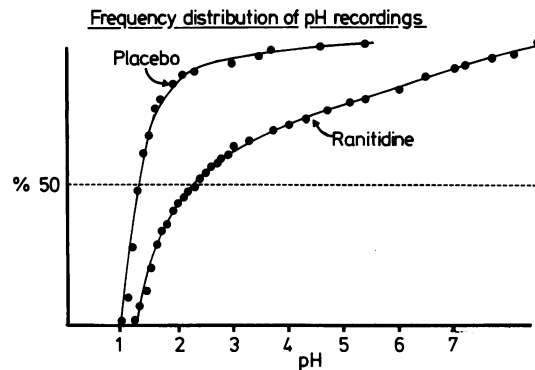
Fig. 7   *Percentage of recordings above or below any given pH in 10 duodenal ulcer patients receiving placebo and ranitidine. (Data from 7).*

measurements (Fig. 7), as strict temporal relationships are unimportant. Such data could also be statistically assessed by non-parametric analysis of variance of the frequency of relevant pH measurements.

FUTURE RECOMMENDATIONS

Uniformity of analysis and expression of data would help investigators to compare different studies. Within any individual experiment providing like is compared with like, however, the qualitative conclusions are independent of the method of analysis. The logarithmic relationship of pH to hydrogen ion activity and the usual non-normal distribution of data has resulted in a confusion of methods of data presentations. No published investigation of 24 hour intragastric acidity (including this author's) has used all of the 'correct' analytical procedures outlined above. We have a choice; either we continue to use many different analyses, of which some are statistically incorrect, or a standard analytical method is used by all. If the latter were possible, the following recommendations seem sensible: (1) Median and percentile (minimum-interquartile) ranges of H^+ activity (mmol/l) should be used for graphs, calculations of overall acidity and percentage inhibition of acidity. A graph of median hourly acidity would allow the temporal relationship of drug administration and meals with inhibition of acidity to be shown; an indication of interindividual variation would be shown by the ranges. (2) Frequency distribution curves of pH measurements should be published to demonstrate all the data; these have the advantage of showing changes of pH above 5·0 which is not possible with median hourly hydrogen ion activity. (3) Non-parametric statistics should be used and multiple repeated significance testing avoided. (4) The exact methematical methods used to calculate overall 24 hour, nocturnal or other arbitrary measures of acidity should be described in detail, and (5) Analytical programs for assessment of continuous pH recording should incorporate the above.

R WALT

Department of Therapeutics,
University of Nottingham,
Queens Medical Centre,
Nottingham NG7 2UH

References

1 James AH, Pickering GW. The role of gastric acidity in the pathogenesis of peptic ulcer. *Clin Sci* 1949; **8:** 181–210.

2 Lennard-Jones JE, Babouris N. Effect of different foods on the acidity of the gastric contents in patients with duodenal ulcer. *Gut* 1965; **6:** 113–7.

3 Pounder RE, Williams JG, Milton-Thompson GJ, Misiewicz JJ. 24 hour control of intragastric acidity by cimetidine in duodenal ulcer patients. *Lancet* 1975; **2:** 1069–72.

4 Pounder RE, Hunt RH, Vincent SH, Milton-Thompson GJ, Misiewicza JJ. 24 hour intragastric acidity and nocturnal acid secretion in patients with duodenal ulcer during oral administration of cimetidine and atropine. *Gut* 1977; **18:** 85–90.

5 Peterson WL, Barnett C, Feldman M, Richardson CT. Reduction of twenty four hour gastric acidity with combination drug therapy in patients with duodenal ulcer. *Gastroenterology* 1979; **77:** 1015–20.

6 Milton-Thompson GJ, Wright B, Vincent D, Hunt RH. Comparisons of 24 hour intragastric acidity in duodenal ulcer patients on high dose antacid or cimetidine. *Gastroenterology* 1979; **76:** 1204.

7 Walt RP, Male PJ, Rawlings J, Hunt RH, Milton-Thompson GJ, Misiewicz JJ. Comparison of the effects of ranitidine, cimetidine and placebo on the 24 hour intragastric acidity and nocturnal acid secretion in patients with duodenal ulcer. *Gut* 1981; **22:** 49–54.

8 Walt RP, Gomes M, Wood EC, Logan LH, Pounder RE. Effect of daily oral omeprazole on 24 hour intragastric acidity. *Br Med J* 1983; **287:** 12–4.

9 Gledhill T, Buck M, Paul A, Hunt RH. Cimetidine or vagotomy? Comparison of the effects of proximal gastric vagotomy, cimetidine and placebo on nocturnal intragastric acidity and acid secretion in patients with cimetidine resistant duodenal ulcer. *Br J Surg* 1983; **70:** 704–6.

10 Gledhill T, Howard OM, Buck M, Paul A, Hunt RH. Single nocturnal dose of an H$_2$-receptor antagonist for treatment of duodenal ulcer. *Gut* 1983; **24:** 904–8.

11 Sharma BK, Walt RP, Pounder RE, Gomes M, Wood EC, Logan LH. Optimal dose of oral omeprazole for maximal 24 hour decrease of intragastric acidity. *Gut* 1984; **25:** 957–64.

12 Santana IA, Sharma BK, Pounder RE, Wood EC, Masters S, Talbot M. 24 hour intragastric acidity during maintenance treatment with ranitidine. *Br Med J* 1984; **289:** 1420.

13 Walt RP, Reynolds JR, Langman MJS et al. Intravenous omeprazole rapidly raises intragastric pH. *Gut* 1985; **26:** 902–6.

14 Pounder RE, Williams JG, Milton-Thompson GJ, Misiewicz JJ. Effect of cimetidine on 24 hour intragastric acidity in normal subjects. *Gut* 1976; **17:** 133–8.

15 Prichard PJ, Yeomans ND, Mihaly GW, Jones DB, Smallwood RA, Louis WJ. The effect of omeprazole, a (H$^+$K$^+$)ATPase inhibitor, on 24 hour gastric pH profile in man. *Aust NZ J Med* 1983; **13:** 432–3.

16 Prichard PJ, Yeomans ND, Mihaly GW et al. Omeprazole: a study of its inhibition of gastric pH and oral pharmacokinetics after morning or evening doses. *Gastroenterology* 1985; **88:** 64–9.

17 Etienne A, Fimmel CJ, Bron BA, Loizeau E, Blum A. Evaluation of pirenzepine on gastric acidity in healthy volunteers using ambulatory 24h intragastric pH monitoring. *Gut* 1985; **26:** 241–5.

18 Santana IA, Sharma BK, Orchard K, Pounder RE. 24 hour intragastric acidity before and during treatment with enprostil. [Abstract] *Gut* 1985; **26:** A545.

19 Deakin M, Ramage JK, Paul A, Gray S, Billings J, Williams JG. Effect of enprostil on 24 hour intragastric acidity and nocturnal acid and pepsin output. [Abstract] *Gut* 1985; **26:** A545.

20 Classen M, Damman HG, Domschke W et al. Omeprazole heals duodenal but not gastric ulcers more rapidly than ranitidine. [Abstract] *Gut* 1985; **26:** A558.

21 Lauritsen K, Rune SJ, Bytzer P et al. Effect of omeprazole and cimetidine on duodenal ulcer. *N Engl J Med* 1985; **312:** 958–61.

22 Gledhill T, Buck M, Hunt RH. Effect of no treatment, cimetidine 1G per day, cimetidine 2G per day and cimetidine combined with atropine on nocturnal gastric secretion in cimetidine non-responders. *Gut* 1984; **25:** 1211–6.

23 Sharma BK, Santana IA, Wood EC et al. Bacterial colonisation with omeprazole. Intragastric bacterial activity and nitrosation before, during and after treatment with omeprazole. *Br Med J* 1984; **289:** 717–9.

24 Hansky J, Stern AI, Korman MG, Waugh J. Effects of long-term cimetidine on serum gastrin in duodenal ulcer. *Dig Dis Sci* 1979; **24:** 468–70.

25 Muscroft TJ, Youngs DJ, Burdon DW, Keighley MRB. Cimetidine is unlikely to increase formation of intragastric N-nitroso-compounds in patients taking a normal diet. *Lancet* 1981; **1:** 408–10.

26 Milton-Thompson GJ, Ahmet Z, Lightfoot WF *et al*. Intragastric acidity, bacteria, nitrite, and N-nitroso compounds before during and after cimetidine treatment. *Lancet* 1982; **1:** 1091–5.

27 Kapur B, Mills JG, Glenny H, Burland WL, Lunt M, Bardhan KD. Evaluation of large single night time doses of cimetidine using continuous 24 hour ambulatory gastric pH monitoring. [Abstract] *Gut* 1985; **26:** A559.

28 Moore EW, Scarlatta RW. The determination of gastric acidity by the glass electrode. *Gastroenterology* 1965; **49:** 178–88.

29 Sokal RR, Rohlf FJ. *Biometry*. San Francisco: W H Freeman, 1981.

30 Fimmel CJ, Etienne A, Cilluffo T *et al*. Long-term ambulatory gastric pH monitoring: Validation of a new method and effect of H_2-receptor antagonists. *Gastroenterology* 1985; **88:** 1842–51.

31 Elashoff JD. Down with multiple *t* tests. *Gastroenterology* 1981; **80:** 615–20.

32 Siegel S. *Nonparametric statistics for the behavioural sciences*. New York: McGraw-Hill, 1956: 166–93.