Reproducibility of 24 hour oesophageal pH studies in infants

F J Hampton, U M MacFadyen, H Simpson

Abstract

Thirteen infants who had undergone 24 hour oesophageal pH monitoring to diagnose gastro-oesophageal reflux had a second study carried out to see if the results were reproducible. The studies were done without restricting the babies' activities. Appreciable differences were found, the percentage of the total time during which the pH was less than 4 varying by up to 3·7-fold between the two tests. The differences were largely the result of biological rather than technical variability. From these results estimates were made of the reliability of a single diagnostic study and the size of changes that would be necessary to show the effect of treatment.

These findings have a considerable impact on the diagnosis of abnormal gastro-oesophageal reflux and its response to treatment whether using 24 hour pH monitoring or any other method of measurement.

The number of different techniques for investigating gastro-oesophageal reflux indicates that none is perfect. Twenty four hour oesophageal pH monitoring, particularly when carried out at home, has the advantages of both a prolonged observation period and near normal conditions. It also has the potential disadvantages that it cannot measure the volume of refluxed material or detect any non-acidic reflux.

It has been accepted that individual subjects may have some investigations that show that gastro-oesophageal reflux is present and some that do not,[1–8] but this is usually ascribed to variations in the accuracy of the tests rather than the possibility of the same test showing reflux one day and not another. As symptoms of gastro-oesophageal reflux vary from day to day such a difference would not be surprising.

Twenty four hour pH studies cause relatively little discomfort or disruption to routine, and so are amenable to repetition to investigate their reproducibility. The results derived from pH monitoring are often subjected to detailed analysis and statistical comparisons, as opposed to simple grading as used for other methods of investigation. This makes it even more important that the reproducibility of the results is known, but there are little published data in adults and none at all in infants.[9–11]

The aim of this study was to obtain this information and use it to establish the reliability of a single normal or abnormal result in the diagnosis of gastro-oesophageal reflux and then to determine how helpful pH studies are when used to compare groups or to monitor the response of gastro-oesophageal reflux to treatment.

Department of Child Health, Clinical Sciences Building, Leicester Royal Infirmary, Leicester LE1 5WW
F J Hampton
U M MacFadyen
H Simpson

Correspondence to:
Dr MacFadyen.

Accepted 7 June 1990

Patients and methods

As part of several larger studies approved by the local ethics committee pH studies were being undertaken to assess the incidence of gastro-oesophageal reflux in infants presenting with a variety of problems. Infants studied included those presenting with chronic respiratory symptoms and acute events such as choking or apnoea as well as those with vomiting. Patients were selected for the duplicate study when it would cause no extra inconvenience to the baby or to the family or an unnecessary hospital stay. Eleven infants needed to stay in hospital for a series of investigations and had their studies on consecutive days so that leaving the pH electrode in position caused no additional discomfort. The other two infants had their studies repeated because of symptoms that persisted after four and six weeks. Four infants were on antireflux treatment, which was constant throughout. No restriction was placed on the positioning or handling of the babies or on their diet. The timing and nature of their intake of food and drink were recorded.

Acceptable duplicate studies lasting at least 15 hours and including both day and night were obtained in 13 infants with ages ranging from 7 days to 11 months (median 2 months) (table 1). No results had to be discarded because of too large a drift in the electrode reading in buffer, or because the infant removed the probe too soon. One infant who would have been studied twice pulled the probe out after the first study and it was not replaced.

pH MEASUREMENTS

The Synectics ambulatory system with semi-disposable monocrystant antimony pH elec-

Table 1 Details of infants

Age	*Presenting problem*	*History of vomiting*	*Receiving treatment*
11 Days	Apnoea	No	No
5 Months	Apnoea	Yes	Yes
6 Months	Apnoea	Yes	No
3 Weeks	Apnoea	Yes	Yes
3 Months	Apnoea (preterm)	Yes	Yes
8 Months	Apnoea	Yes	Yes
2 Weeks	Apnoea	Yes	No
1 Month	Respiratory (preterm)	Yes	No
2 Months	Apnoea	Yes	No
11 Months	Failure to thrive	Yes	No
7 Days	Apnoea	No	No
2 Months	Respiratory	Yes	No
2 Weeks	Apnoea	Yes	No

trodes with an external diameter of 2·1 mm was used. Each electrode was tested in Synectics buffers of pH 7 and 1 before and after the studies to ensure that the readings did not differ by more than 0·2 pH units. The electrodes were positioned in the middle third of the oesophagus as calculated from the infants' length and confirmed on chest radiograph. A silver/silver chloride reference electrode with Hellige electrode gel was placed on the chest. The pH was recorded every four seconds in a Digitrapper Mark 11 (Synectics) before being transferred to an Amstrad PC1512. Data analysis was done with dedicated software, 'EsopHogram' (Gastrosoft Inc).

DATA ANALYSIS

The time studied was divided in different ways to give six time periods for analysis: the total time, the time excluding mealtimes, time less than or greater than two hours after a meal, and night and day. Night was arbitrarily defined as 2200 to 0600 hours. For each time period data were analysed using pH <4 to define four measures of reflux: the number of episodes, the number lasting longer than five minutes, the duration of the longest episode, and the percentage of time during which the pH was <4—also called the 'reflux index'. The number of episodes was expressed as an average rate/hour to compensate for time differences between the studies.

Initially the definition of the end of a reflux episode was when the pH went above 5 with its duration being the time during which the pH was less than 4. This definition is not universal,[11 12] some investigators defining the end of an episode as the point at which the pH goes above 4,[13 14] so the tests were all reanalysed using this criterion for counting episodes and measuring their duration.

Several further variables were calculated for the total time period alone, and then compared between the two studies. The area of the curve under pH 4 was calculated by counting the number of minutes, to the nearest integer, spent between each 0·1 of a pH unit. The number of 'minute units' was calculated for each hour of study. The computer software provided a coarser definition of area (under pH 5), which was expressed as the Kaye score. This is derived by multiplying the percentage of time with pH between 4 and 5 by 1, that between pH 3 and 4 by 2, that between pH 2 and 3 by 3, and that between pH 1 and 2 by 4, and then adding them. This was also adapted to record the time that the pH was under 4 by multiplying the percentage of time with pH between 3 and 4 by 1, and so on. The percentage of the total time during which the pH was <5 was also measured. Finally, a score based on the principles of Johnson and Demeester[13 14] was calculated from the normal values in infants published by Vandenplas and Sacre-Smits.[15]

Table 2 Median (range) values of all results

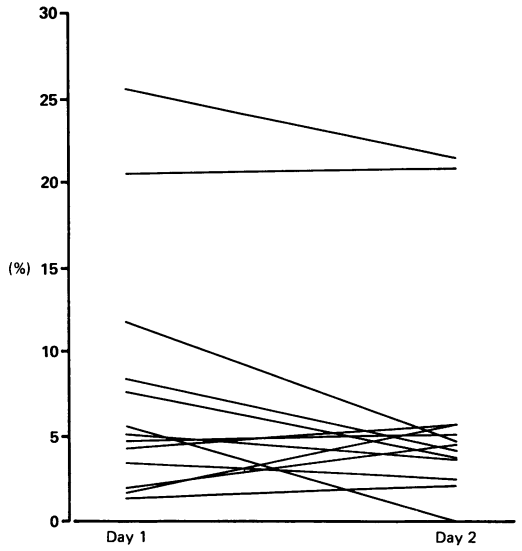
	Firsty study	*Second study*
Original analysis		
Total time:		
No of episodes	0·99 (0·17–7·21)	0·91 (0–6·24)
No lasting >5 mins	0·13 (0–0·63)	0·10 (0–0·71)
Longest episode (mins)	40 (3–88)	15 (0–47)
Time that pH <4(%)	5·1 (1·2–25·6)	4·5 (0–21·4)
Without meals:		
No of episodes	0·97 (0·32–7·16)	0·97 (0·6–6·05)
No lasting >5 mins	0·14 (0–0·58)	0·14 (0–0·76)
Longest episode (mins)	16 (3–73)	15 (5–44)
Time that pH <4(%)	4·7 (0·9–25·9)	4·7 (2–21·6)
≤120 Min after meal:		
No of episodes	0·88 (0–6·27)	1·21 (0·54–12·21)
No lasting >5 mins	0·09 (0–0·72)	0·08 (0–0·86)
Longest episode (mins)	8 (0–49)	7 (2–44)
Time that pH <4(%)	2·5 (0–38·5)	4·3 (0·6–39·6)
>120 Min after meal:		
No of episodes	0·95 (0·4–7·81)	1·15 (0·69–3·38)
No lasting >5 mins	0·10 (0–0·62)	0·20 (0–0·57)
Longest episode (mins)	11 (2–65)	13 (2–38)
Time that pH <4(%)	5·7 (0·8–21·1)	5·2 (1·0–14·4)
During the day:		
No of episodes	0·90 (0·06–8·76)	1·13 (0–9·68)
No lasting >5 mins	0·06 (0–0·75)	0·07 (0–0·94)
Longest episode (mins)	11 (0–73)	10 (0–44)
Time that pH <4(%)	2·9 (0–34·6)	4·3 (0–33·1)
During the night:		
No of episodes	0·75 (0·13–4·25)	0·63 (0–1·25)
No lasting >5 mins	0·13 (0–0·59)	0·13 (0–0·38)
Longest episode (mins)	10 (3–88)	11 (0–47)
Time that pH <4(%)	7·8 (0·7–21·5)	4·0 (0–10·2)
Using pH 4 to end reflux		
Total time:		
No of episodes	1·75 (0·65–11·93)	1·73 (0–11·04)
No lasting >5 mins	0·09 (0–0·54)	0·09 (0–0·5)
Longest episode (mins)	10 (3–49)	11 (0–36)
Time that pH <4(%)	5·1 (1·2–25·6)	4·5 (0–21·4)
≤120 Min after meal:		
No of episodes	1·37 (0–13·37)	2·42 (0·37–19·77)
No lasting >5 mins	0 (0–0·96)	0 (0–0·69)
Longest episode (mins)	5 (0–43)	4 (2–36)
Time that pH <4(%)	2·5 (0–38·5)	4·3 (0·6–39·6)
>120 Min after meal:		
No of episodes	2·23 (0·53–11·23)	1·84 (0·78–7·15)
No lasting >5 mins	0·07 (0–0·64)	0·16 (0–0·57)
Longest episode (mins)	7 (2–44)	9 (0–20)
Time that pH <4(%)	5·7 (0·8–21·1)	5·2 (1–14·4)
Total time other measures		
Reflux index (pH 4)	5·1 (1·2–25·6)	4·5 (0–21·4)
Reflux index (pH 4·2/3·8)	4·6 (1·8–24·9)	4·9 (0–23·5)
Reflux index (pH 5)	24·9 (7·3–53·6)	16·9 (0–40·6)
Kaye score <pH 5)	30·9 (11·3–75·3)	22 (0–64·9)
Kaye score (<pH 4)	6·6 (1·3–39)	6·7 (0–33·7)
Area under pH 4	1·7 (0·4–17·4)	2·4 (0–13·9)
Total score	21·7 (4·7–51·9)	15·3 (3·9–41·1)

STATISTICAL METHODS

Natural logarithms of each reflux variable were taken and the difference between each of these results for each infant on the two days was calculated. A mean (SD) of the differences were thus found for each variable. By taking the antilog of 1·96 multiplied by the SD of the differences, 95% confidence intervals were calculated. Because the original results were logged these confidence intervals are expressed as the limits within which the ratio of the two results for each reflux variable for one infant should lie.

Results

Table 2 shows the medians and ranges for all the reflux variables used in the following analysis and shows no trend between the first and second studies. The reflux indices for each infant on the two occasions are illustrated in the figure. Some infants had abnormal gastro-oesophageal reflux when it was defined as a value of greater than 2SD above the mean of Vandenplas and Sacre-Smits' normal values[15] and some using a reflux index of greater than 5% (table 3). Neither method produced the same results for all infants. The mean differ-



Changes in reflux index from day 1 to day 2.

ences between the paired results for all the variables were close to zero, confirming that there was no tendency for the first or second test to give larger results.

CALCULATED RESULTS

Table 4 gives the upper confidence limits derived from the antilog of 1·96 multiplied by the SD of the logged differences; the lower limit is the inverse of this. For example, the confidence intervals for the ratio of the two reflux indices for the total time are 1/3·7 (0·27) to 3·7: that is to say that a second study will have a reflux index somewhere between 27% and 370% of the first.

TIME PERIODS

Review of the confidence intervals for all the variables showed that the total time was at least as reproducible as any other time period. Night time was the least reproducible.

REFLUX VARIABLES

The number of episodes that lasted longer than five minutes was the most reproducible measure, but several of the infants had at least some time periods during one of the studies with no such episodes. The ratio for these studies is not calculable and is excluded, possibly causing an underestimate of the variability (see table 4). The length of the longest reflux episode was the least reproducible measure in all time periods, varying up to 22-fold. The number of episodes and the reflux index have intermediate values.

Using pH 4 to define the end of the reflux episode (table 5) there was no consistent change in the confidence intervals: for some measures they increased and some decreased. This method of analysis produced more studies with no episodes lasting longer than five minutes, and so reduced the number of patients contributing to this calculation.

There is thus no apparent advantage in subdividing the total time nor in using variables other than the reflux index, which is independent of the definition of the end of a reflux episode. We therefore used this index for investigating the possible influence of electrode inaccuracies and used only the total time for further analyses.

ACCURACY OF ELECTRODES

During the testing of the pH catheters a maximum drift of pH readings in buffers of 0·2 pH units was allowed. To see whether this could be

Table 3 Results expressed as 'normal' or 'abnormal' either by Vandenplas and Sacre-Smits' results or as reflux index over 5%

	No of infants in whom both studies normal	*No of infants in whom both studies abnormal*	*No of infants in whom one study was normal and one abnormal*
Vandenplas and Sacre-Smits[15]:			
No of episodes	0	8	5
No lasting >5 mins	8	3	2
Longest episode	4	3	6
Time that pH <4(%)	4	4	5
All normal compared with any abnormal results	0	8	5
Reflux index >5%	3	2	8

Table 4 Upper limits of confidence intervals for 24 ratios. Figures in parentheses represent number of infants whose results were calculable

	No of episodes/ hour	*No of episodes/ hour lasting >5 min*	*Longest episode (min)*	*Reflux index (percentage of time that pH <4)*
Total time	3·8	2·4 (9)	11·9	3·7
Time excluding mealtimes	3·9	2·2 (8)	11·9	3·3
≤120 Min after meals	4·5	3·5 (5)	5·8	3·3
>120 Min after meals	4·4	4·3 (5)	17·1	8·5
During day (0600–2200)	5·4	4·8 (6)	8·2	5·6
During night	8·0	2·9 (6)	21·9	15·1

Table 5 Upper limits of confidence intervals after reanalysis using pH 4 to define end of reflux. Figures in parentheses represent number of infants whose results were calculable

	No of episodes/ hour	*No of episodes/ hour lasting >5 min*	*Longest episode (min)*	*Reflux index (percentage of time that pH <4)*
Total time	4·0	2·4 (6)	10·9	3·7
≤120 Min after meals	4·0	3·6 (3)	6·0	3·3
>120 Min after meals	6·3	1·4 (4)	17·1	8·5

Table 6 Upper limits of confidence intervals for all variables analysed for total time

Reflux index using pH 4 as cut off for both studies	3·7
Reflux index using pH 4·2 and 3·8 for cut off	2·9
Reflux index using pH 5 for both studies	1·9
Kaye score using pH 5 as baseline	2·0
Kaye score using pH 4 as baseline	4·4
Area under curve using pH 4 as cut off	4·9
Score for total time	3·9

responsible for some of the variability the studies were reanalysed. The study with the lower reflux index for each infant was reanalysed using a cut off value of pH 4·2 and that with the higher reflux index using a cut off of 3·8 to reduce the differences between the two studies. The confidence interval for the ratio for the reflux index for the total time was only reduced from 3·7 to 2·9-fold.

OTHER VARIABLES (table 6)

The computer software automatically calculated the Kaye score, and this result had a lower variability than the measures previously analysed. It was felt that this might be because the pattern of reflux was more accurately reflected by a measure of the 'area under the curve' so this was calculated for pH under 4. This however showed an increased variability, as did the Kaye score adapted for pH under 4. The percentage of total time during which the pH was <5 was less variable.

In case the use of a combined measure could provide a better reproducibility a score was calculated for each test using each of the four reflux variables for the total time. A result equal to the mean for the age group given by Vandenplas and Sacre-Smits scored 2 and the number of SDs away from this were added or subtracted as described by Johnson and Demeester.[13 14] The reproducibility of the total score was no better than that of any other variable.

Discussion

To investigate the reproducibility of any test, as when comparing two different methods of measurement, it is not satisfactory simply to correlate the pairs of results, as is so often done.[16] It would be surprising if there was not a correlation, because this simply means that the larger values on one occasion are likely to be larger on another. It is necessary to measure the differences between the two results, and the average of the two, and analyse these. If the differences are unrelated to the average result then they can be treated directly but if they are proportional to the average then logarithmic transformation is necessary. This transformation results in the expression of a ratio rather than an absolute difference between the results.

We used the logarithms because we found that the difference was proportional to the average for most of the reflux variables that we generated. Although not all the variables showed this pattern in our small study, the larger study of Weiner *et al* showed that the reflux index does follow this pattern and it seemed likely that all the variables would do so in a larger study.[10]

The results of reanalysing the studies at pH 3·8 and 4·2 only reduced the upper confidence limit by a fifth (3·7 to 2·9), suggesting that any variability associated with equipment was less important than biological variability. The absence of a trend for the first or second study to give higher results also militates against a deterioration in the reaction time of the electrodes. Excluding the two infants who did not have their studies on consecutive days increased the upper confidence interval for the ratio from 3·7 to 4·2.

Some investigators have used the area under the curve to reflect the intensity as well as the duration of acid exposure. Izquierdo *et al* found that this provided good discrimination between normal and abnormal reflux,[17] and Vandenplas *et al* found that it correlated better with oesophagitis.[18] In this study using the area under the curve increased the variability both when analysed accurately using 0·1 pH units and coarsely using 1 pH unit. This suggests that an increase in the time spent below pH 4 is accompanied by a proportional increase in time spent at even lower pH values. This supports our contention that minor variations in the recorded pH do not make an important contribution to the variability. It may also explain why the area under the curve provides a good discrimination, because it tends to exaggerate the differences between studies.

The use of pH 4 as a cut off value to define reflux events was initially derived on the theoretical grounds that pepsin is active at pH <4 and so oesophagitis would be likely to occur.[19 20] It was also shown that pain occurred at pH <4, although Tuttle *et al* reported that pain was caused by oesophageal distension as well as by acidity.[21] Later the use of 24 hour pH studies in subjects with and without symptoms showed that values derived using pH 4 provided the best discrimination between groups.[22 23] Some investigators disagreed, however, and found pH 5 provided a better discrimination.[24 25]

In this study the use of pH 5 as a cut off level either for the percentage of total time during which the pH was <5 or for the Kaye score representing the area of the curve under pH 5 reduced the variability. Because the area under

pH 4 had a higher variability this suggested that it was the use of pH 5 rather than the use of a measure reflecting area that reduced the variability. If the use of pH 5 could also be shown to provide adequate discrimination between normal and abnormal reflux in children then it might be a more appropriate value to use.

In this study few restrictions were placed on the infants in order to study them under 'normal' conditions. It may be that for studying responses to treatment a stricter regimen would give less variability. Adult studies have, however, shown that ambulatory monitoring tends to give results that more closely reflect the clinical picture and correlate better with the presence of oesophagitis than studies carried out in hospital, so thought is needed before imposing such restrictions.[26]

This study only included infants, and the results cannot be directly extrapolated to older children. In addition the study was small and only two infants had a reflux index over 20%. Weiner *et al* studied adults with a wide range of reflux indices, however, and found that the variability was similar over the whole range with the reflux index varying up to 3·2-fold compared with 3·7 in this study.[10] Johnsson and Joelsson analysed the results differently, but if the reflux indexes are extracted from their graph and treated in the same way the result is also 3·2.[9] This suggests that until shown otherwise it may be assumed that the same conclusions apply to older children.

To derive practical information from this study the 'reliability' of a pH study and the usefulness of pH studies for comparisons were calculated using the reflux index for the total time. To estimate the 'reliability' of a normal or abnormal result the chance of getting the same answer (normal or abnormal) on a second day was calculated. Different investigators use different definitions of 'normal' but one of the values widely used to define the upper limit of normal is a reflux index of 5%. Taking this as an example it can be predicted that if a reflux index is close to 5% then there is a high chance of getting a normal result on one day and an abnormal one on another. If the patient's reflux index is less than 2·5% then there is at least a 68% chance of finding a reflux index that remains less than 5% on the second day. Similarly if the patient's reflux index is over 10% there is at least a 68% chance of a second reflux index being over 5%. If the patient's reflux index is less than 1·25% then there is virtually no chance of a second study showing a reflux index of over 5%. In the same way an original reflux index of over 20% means that a repeat study is unlikely to give a reflux index of less than 5%. This principle applies whatever value is chosen for the upper limit of the reference range. This is not the same as calculating the patient's 'true' reflux index, but because our contention is that this varies from day to day this is the best measure of 'reliability' that is possible.

When considering comparisons between groups or in individual subjects before and after treatment the pH study seems to be less useful. The extent of the variability is such that when comparing two groups it would require 80 patients/group to show an average 25% reduction in reflux index with a power of 90%. If an individual subject is treated and retested, then a 75% reduction in reflux index is necessary to be sure of a response to treatment. A lesser reduction, though seeming to be a response, might only represent the day to day variability.

CONCLUSIONS

These studies show that there was a considerable variation in the results obtained in two pH studies carried out under similar conditions in the same infants. Talking to parents indicates that their child's symptoms—at least vomiting, which is the easiest to define—vary considerably and there may be some days when it does not occur and others when it is severe. It is not surprising that the results of the pH study also varied, and it seems likely that other methods of investigating gastro-oesophageal reflux if similarly studied would also vary. We suggest that biological variability is responsible for our findings rather than any problems with the technique.

Although 24 hour pH studies provide useful information not available from other methods of investigating gastro-oesophageal reflux, the underlying variability of the amount of reflux must be taken into account when interpreting the results. If a result does not fit with the clinical picture then a repeat pH study should be considered instead of, or as well as, a different investigation. The variability that we have shown may obscure clinically important improvements in gastro-oesophageal reflux in small studies of treatment and means that new calculations for the power of such studies are needed. Forty eight hour studies would allow results to be obtained from smaller groups of patients, and even longer periods of study may be needed to provide a better picture in an individual subject.

This work was supported by a grant to FJH from Janssen Pharmaceutical Limited. We thank the paediatricians who referred infants under their care for investigation, the parents for agreeing to the investigations, and the parents and nursing staff for looking after the infants during the studies. Statistical discussion with John Thompson was much appreciated but he is not responsible for any possible errors.

1 Arasu TS, Wyllie R, Fitzgerald JF, *et al.* Gastro-oesophageal reflux in infants and children, comparative accuracy of diagnostic methods. *J Pediatr* 1980;**96**:798–803.
2 Blumhagen JD, Rudd TG, Christie DL. Gastro-oesophageal reflux in children: radionuclide gastroesophagography. *AJR* 1980;**135**:1001–4.
3 Christie DL. The acid reflux test for gastro-oesophageal reflux. *J Pediatr* 1979;**94**:78–81.
4 Euler AR, Ament ME. Detection of gastro-oesophageal reflux in the paediatric-age patient by oesophageal intraluminal pH probe measurement. (Tuttle test.) *Pediatrics* 1977;**60**:65–8.
5 Heyman S, Kirkpatrick JA, Winter HS, Treves S. An improved radionuclide method for the diagnosis of gastro-oesophageal reflux and aspiration in children. *Radiology* 1979;**131**:479–82.
6 Jona JZ, Sty JR, Glicklich M. Isotope scans for gastro-oesophageal reflux. *J Pediatr Surg* 1981;**16**:114–7.
7 Rudd TG, Christie DL. Demonstration of gastro-oesophageal reflux in children by radionuclide gastro-oesophageal. *Radiology* 1979;**131**:483–6.
8 Seibert JJ, Byrne WJ, Euler AR, Latture T, Leach M, Campbell M. Gastro-oesophageal reflux the acid test: scintigraphy or the pH probe. *AJR* 1983;**140**:1087–90.

9 Johnsson F, Joelsson BO. Reproducibility of ambulatory oesophageal pH monitoring. *Gut* 1988;**29**:886–9.
10 Weiner GJ, Morgan TM, Copper JB, *et al.* Ambulatory 24-hour oesophageal pH monitoring: reproducibility and variability of pH parameters. *Dig Dis Sci* 1988;**33**:1127–33.
11 Ward BW, Wu WC, Richter JE, Lui KW, Castell DO. Ambulatory 24 hour oesophageal pH monitoring. *J Clin Gastroenterol* 1986;**8**(supp 1):59–67.
12 Emde C, Garner A, Blum AL. Technical aspects of intraluminal pHmetry in man: current status and recommendations. *Gut* 1987;**28**:1177–88.
13 Johnson LF, Demeester TR. Twenty-four hour pH monitoring of the distal oesophagus. *Am J Gastroenterol* 1974;**62**: 325–32.
14 Johnson LF, Demeester TR. Development of the twenty-four hour intraoesophageal pH monitoring composite scoring system. *J Clin Gastroenterol* 1986;**8**(supp 1):52–8.
15 Vandenplas Y, Sacre-Smits L. Continuous 24 hour pH monitoring in 285 asymptomatic infants 0 to 15 months old. *J Pediatr Gastroenterol Nutr* 1987;**6**:220–4.
16 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;**i**:307–10.
17 Izquierdo M, Tovar J, Elizequirre I. L'exposition acide oesophagienne en un seul chiffre: la surface sous la courbe de pH. *Chir Pediatr* 1989;**30**:1–5.
18 Vandenplas Y, Goossens A, Marichal M. Area under pH 4: advantages of a new parameter in the interpretation of oesophageal pH monitoring data. *Pediatr Res* 1989;**26**:269.
19 Kinlaasho E, Fromm O, Silen W. Effect of bile salts and related compounds on isolated oesophageal mucosa. *Surgery* 1980;**87**:280–5.
20 Lillemoe KD, Johnson LF, Harmon JW. The role of the components of gastroduodenal contents in experimental acid oesophagitis. *Surgery* 1982;**92**:276–84.
21 Tuttle SG, Rufin F, Bettarello A. The physiology of heartburn. *Ann Intern Med* 1961;**55**:292–300.
22 Vitale GC, Cheadle WG, Sadek S, Michel ME, Cuschieri A. Computerised 24-hour ambulatory oesophageal pH monitoring and oesophagogastro-duodenoscopy in the reflux patient. *Ann Surg* 1984;**200**:724–8.
23 Wallin L, Madsen T. 12-hour simultaneous registration of acid reflux and peristaltic activity in the oesophagus. *Scan J Gastroenterol* 1979;**14**:561–6.
24 Evans DF, Haynes J, Jones JA, Stower MJ, Kapila L. Ambulatory oesophageal pH monitoring in children as an indicator for surgery. *J Pediatr Surg* 1986;**21**:221–3.
25 Stanciu C, Hoare RC, Bennett JR. Correlation between manometric and pH tests for gastro-oesophageal reflux. *Gut* 1977;**18**:536–40.
26 Branicki FJ, Evans DF, Ogilvie AL, Atkinson M, Hardcastle JD. Ambulatory monitoring of oesophageal pH in reflux oesophagitis using a portable radiotelemetry system. *Gut* 1982;**23**:992–8.

Indomethacin and the premature brain

Those of us who do not have a first class honours degree in physics will probably have difficulty in grasping the basic principles of near infrared spectroscopy (NIRS) and the Beer-Lambert law describing optical absorption in a highly scattering medium. Workers at the University College and Middlesex School of Medicine in London have, however, used NIRS to assess cerebral haemodynamics in premature babies given intravenous indomethacin (Edwards *et al, Lancet* 1990; **335**:1491–5) They found that cerebral blood flow, oxygen delivery, cerebral blood volume, and reactivity to carbon dioxide all fell sharply after giving the drug and it made no difference whether it was given rapidly or slowly. They do not, however, feel that the use of indomethacin to close patent ductus arteriosus should be abandoned but suggest that attempts should be made to ensure that oxygenation is maximal before giving the drug. That, of course, is a far from satisfactory solution, though it may be the best available at present. We badly need an alternative.

ARCHIVIST