Progress report

Technical aspects of intraluminal pH-metry in man: current status and recommendations

In April 1986 a discussion meeting entitled ‘Technical aspects of intraluminal pH-metry using pH-electrodes in man’ was held in Zurich, Switzerland. The meeting brought together specialists in theoretical aspects of pH-recording, design and manufacture of equipment, and clinical application of this technique. The aim was to define the actual level of knowledge of pH monitoring in the upper gastrointestinal tract and to provide recommendations for its practical use. In view of the increasing use of intraluminal pH-metry it was considered timely to hold a discussion with the focus primarily on equipment and experimental design because understanding these factors is essential for application of pH-metry to research and routine clinical investigation.

The meeting was organised in a format designed to maximise discussion and arrive at a consensus view. Before the meeting, participants submitted up to three concise statements on assigned topics, which were then precirculated to all 41 delegates. Fifty nine topics ranging from ‘Interference of external factors with pH-measurements – temperature’ to ‘Specific problems of duodenal pH-metry – normal values’ were addressed. There were no formal presentations and the entire meeting was devoted to discussion leading to modification, addition or replacement of these statements. Finally, delegates voted (agree, disagree, no opinion) on all statements and on the overall importance of a particular topic. These statements form the basis of this text with the voting being used to define the balance and emphasis of the article. Thus, the purpose of this article is to reflect the conclusions of the meeting and to incorporate a series of precise and unambiguous recommendations. As such certain of the statements may appear more dogmatic than in a conventional review article.

Equipment

pH-electrode

Intraluminal pH can be measured by either a glass, plastic or ISFET (ion sensitive field effect transistor) combination electrode (containing both sensing and reference elements) or a unipolar (sensing electrode alone) monocryalline antimony electrode connected to a pH-meter via cable. Combination glass telemetry capsules are also available which avoid potential problems associated with the cable but are more susceptible to errors caused by electromagnetic fields. Combination glass electrodes predominate although monocryalline antimony electrodes are being used increasingly for oesophageal pH-metry. Currently, combination glass
telemetry capsules and ISFET electrodes are rarely used, while plastic electrodes based on solvent polymeric membranes are not available for routine use.

The price of glass and ISFET electrodes is relatively high (about £200), whereas monocrystalline antimony electrodes are now available for about £30. Polymeric membrane electrodes are not yet manufactured commercially, but in theory they should prove even less expensive than antimony electrodes.

When selecting the appropriate type of electrode, both electrochemical and mechanical properties need to be considered.

(i) Electrochemical aspects
The response of glass and ISFET electrodes is essentially linear over the pH range 0 to 12. Monocrystalline antimony electrodes have a linear range from pH 1 to 8 in the absence of complex-forming ligands. At present, plastic electrodes only cover a selected pH range depending on the membrane composition and the other ions contained in the sample. In the case of biological fluids, a pH range up to 6.5 can be achieved. When using plastic electrodes for duodenal pH-metry the possible interaction of bile with plastic polymers needs to be considered.

Glass, ISFET and polymeric membrane electrodes have a response time (defined as the time required to attain 90% of the final value in unstirred solution) of less than two seconds in the pH range between 1 and 12. Monocrystalline antimony electrodes have a comparable response time in the pH range from 1 to 4, but they may have a longer response time for pH changes above 4 in the direction of neutrality.

In all types of electrodes except ISFETs minor drift of the electrode (defined as less than 0.2 pH units) is caused by liquid junction potentials, a greater drift is because of electrode failure. Ion sensitive field effect transistor electrodes have, independent from liquid junction potentials, a systematic drift of about 0.1 to 0.2 pH units per 24 hours which may be corrected for. In all other types of electrodes the time course of drift is unknown and thus complete correction is not possible. Correction at the end of an experiment, however, assuming a linear drift is probably superior to no correction. Furthermore, drift is a particular problem with unipolar electrodes using a skin reference, principally because an appropriate composition for the contact jelly has not yet been established.

Both the sensing and the reference electrodes are potential sources of error. This error can be compounded if the two elements are separated. At pH values up to 6, glass, ISFET and plastic electrodes have negligible error (less than 0.05 pH-units). The error of monocrystalline antimony electrodes with a skin reference electrode generally does not exceed 0.5 pH units. This may be unacceptable, however, when attempting to quantify the small differences in pH occurring with gastric and duodenal pH-metry.

(ii) Mechanical properties
The smallest diameter of currently available combination glass electrodes is 3 mm. Ion sensitive field effect transistor, monocrystalline antimony and prototype plastic electrodes have diameters of about 2 mm. Theoretically, these types can be made smaller. When the diameter of the pH sensitive region of an electrode is small (below 0.5 mm) and located at its tip, the
probe may penetrate the mucus gel layer and record a high pH which exists at the mucosal surface. A pH-electrode immersed partly into the surface mucus coat, however, records the pH of the luminal content as long as there is a part of the pH sensitive membrane exposed to the luminal fluid.\(^3\)

When the electrode is to be passed by the preferred nasal route, length of the rigid part of the probe can be a problem in the case of glass electrodes. Plastic and antimony electrodes do not normally have any rigid parts.

**STERILISATION**

Without sterilisation, transmission of infectious diseases can never be totally excluded. As neither heat nor gas sterilisation is recommended for any type of electrode and no guidelines exist, disposable electrodes would be optimal to deal with disinfection problems. Whether an electrode can be disposed of after being used only once depends on its price.

It is recommended that urgent attention is given to defining conditions for sterilisation of all types of electrodes. In future, manufacturers should not market electrodes without providing adequate information on procedures for their sterilisation.

**RECORDING DEVICE**

When intraluminal pH-metry was introduced first, pH-tracings were written on stationary strip chart recorders in in-patients.\(^2\) Portable data loggers using magnetic tape\(^1\)\(^2\)\(^3\) followed by solid state memories\(^1\)\(^3\)\(^13\) have been subsequently developed. As ambulatory investigations are only possible using portable recorders and the majority are equipped with solid state memories, the ensuing discussion will be focused on these devices.

Solid state memory recorders require analogue-to-digital data conversion before storage. The way this is done determines the precision with which both pH-values and the velocity of pH changes are recorded. An accuracy of 1% is guaranteed for most of the recording devices available because of the commonly used 8-bit data format. The accuracy of recorded changes in pH depends largely on the sampling frequency and is therefore related to the storage capacity. For example, when an accuracy of 1% and a sampling interval of two seconds are desired, a capacity of 43200 kByte is required to obtain a one channel 24 hour recording.

To illustrate the influence of sampling frequency, ultrafrequent pH-recordings (256 samples/minute) were obtained in the stomach and oesophagus under basal conditions and after administration of an H\(_2\) antagonist. Recordings were done with a combination glass electrode (Type 440-M4, Ingold, Switzerland) connected to a programmable solid state recorder (LZ-105, Kaufhold, West-Germany). The original data sets and reduced versions (considering every second, fourth, eighth value, etc) were analysed. From the oesophageal recordings, percentage of time with pH below 4 and total number of reflux episodes were derived (Figs 1a, b). From the gastric recordings, medians were calculated (Fig. 1c). The data show that for means, medians, or threshold values, a sampling frequency of about 8/min is sufficient. For analysis of rapid fluctuations in pH occurring frequently in oesophagus and duodenum, however, a sampling frequency of 8/min would lead to excessive loss of information. For this purpose the sampling frequency must theoretically be in the range of the electrode response time – that is, about 60/min for a typical glass or plastic electrode.
This is a faster sampling rate than is normally available on commercial data loggers.

An event marker in a recording device is more convenient than a written diary for both clinical and experimental use. Multiple event markers may be confusing to patient and doctor, however, as well as being a potential source of error. The recording of events necessitates a certain memory space and this has to be considered when defining the required memory capacity. In clinical oesophageal pH-metry, the event marker is normally used to correlate pain attacks with reflux episodes. In experimental pH-metry, analysis of groups of recordings can be made more precisely in relation to specific events of a protocol by using a marker. Under all conditions, the interpretation of event markers must consider the subjective perception of the patient or volunteer.

If non-rechargeable batteries are used as the energy source, there should be a back up to ensure that recorded information remains in the memory even if the main batteries are removed or fail. New or fully recharged batteries should be used for every test. Ideally, recorders using rechargeable batteries should have a system which indicates the battery reserve power. At present, this is not technically feasible for Ni/Cd cells. The recorder should check the slope of the electrode after initial calibration. Because inbuilt checks of electrode-to-recorder connection are difficult to construct, the electrode plug should be lockable to prevent accidental disconnection. Several forms of electrode connectors are in use. There is a need for connectors to be standardised or adaptors made available so that different recorders and electrodes can be used interchangeably. In experimental situations, storage capacity is more important than external dimensions, but for routine clinical use, recorder size and weight must be kept to a minimum — for example, below 200×50×100 mm and 400 g. When comparing costs of different recording devices the need for additional software and service should be considered. Above a certain basic value, cost tends to reflect the number of facilities available in a particular system.
Conduct of studies

A two point calibration of the pH-electrode should be done using a neutral buffer and an acid buffer of pH 2 or less before and after each study. Normally, mucus does not collect on electrodes but even if present, it is not necessary to rinse the electrode before recalibration since H⁺-diffusion through mucus is extremely rapid in relation to even the fastest electrode response time.⁴⁴

Extreme caution should be exercised to ensure that automatic calibration procedures do not obscure information about electrode condition. At least one commercially available system equipped with automatic calibration is unreliable – for example, errors of up to 0.6 pH-units. This inadequacy was only apparent when the automatic calibration procedure was checked manually in a volunteer drug study (Fig. 2).

![Fig. 2](http://gut.bmj.com) Errors occurring in at least one of the commercially available longterm recording devices equipped with an automatic calibration procedure. In 18 consecutive recordings, the registered pH-values were checked manually at the start (0 h) and end (24 h) of the study.

The acid content and buffering capacity of food and drink can be considerable¹⁵ and thus have a major influence on intraluminal pH. Some food constituents also adhere to the electrode and may affect its response time. Comparison of pH-metries from different studies and from different centres is strictly possible only if food intake during pH-metry is standardised. Furthermore, valid interpretation of studies depends on
probe placement and the adequacy of measures to monitor its position. This has to be done using specific procedures for different locations.

Clinical pH-metry-studies should generally have a duration of 24 hours. This ensures prolonged periods of supine and upright body position are included as well as taking account of circadian variations in acid secretion. In the case of oesophageal pH-metry, studies of shorter duration – for example, three hours postprandial, have been proposed. A duration of less than 24 hours, however, cannot generally be recommended on the basis of present knowledge.

Data analysis

Acidity may be expressed as pH or H⁺ concentration. Interconversion from the logarithmic pH-scale to the linear H⁺-scale represents a monotone transformation which means that the ranking of values but not their relation to each other remains constant after transformation. Therefore, either can be used to express a single acidity value and it is a matter of convenience which one to use. When calculating summary variables from one recording or from groups of recordings, however, the choice of unit may have profound implications as means and standard deviations are dramatically influenced whereas percentile derived statistics (such as the median) are not affected.

For all statistical evaluation, a normal distribution which is homogenous among groups is the ideal. When evaluating selected time intervals from one individual, neither pH nor H⁺ values fulfil these criteria although pH-values are often closer and are therefore highly recommended. Nevertheless, there is some deviation from normal distribution when using pH. The pH median is thus the appropriate measure of the centre and should be used when calculating summary variables for selected time intervals from individual tracings. Availability of this type of data analysis should be taken into account when purchasing the equipment as it is not always a feature of commercial software supplied with data loggers.

When comparing different groups of recordings – for example, treated vs control, smoker vs non-smoker, non-parametric test are preferable when available. There is a gain in power when a non-parametric test is applied in the exact, as opposed to the approximated version. In all cases where parametric tests have to be used, it is necessary to establish that the distribution of the raw data is adequate for the procedure selected.

Intraluminal pH-metry generates a large amount of data and there is a tendency for the data to be subjected to multiple statistical testing – that is, evaluating more than one variable per conclusion. If this practice is followed, correct procedures must be undertaken which maintain the predetermined alpha error, otherwise differences will be overestimated. Appropriate methods for multiple comparisons are (i) repeated measures and/or multivariate analysis of variance, (ii) multiple one-way procedures which are tested against a corrected alpha-level, and (iii) reducing the bulk of data to a priori defined summary variables such as scores.

Graphic display of pH vs time (time courses), the time that pH was above or below a certain value (threshold curves), and the frequency of pH values within defined bands (frequency curves) are all used to summarise the data. Threshold and frequency curves contain the same information and it is a
matter of convenience which one to use. Time courses are the method of choice to show the time dependent pattern of acidity due to circadian rhythm, meal intake and for evaluating the effects of drugs. Threshold or frequency curves are the method of choice to summarise studies. Time courses and threshold or frequency curves combined provide the relevant graphic information on a pH-metry study. Graphs should be displayed in a uniform way with time on the x-axis and pH from low to high on the y-axis.

Data reduction becomes necessary for diagrammatic display of a complete 24 hour pH recording. For gastric pH-metry this can be achieved by successive determination of means or medians (medians are preferable but more time-consuming) over short time intervals – for example one minute. In the case of oesophageal pH-metry this procedure would lead to unacceptable alterations in the displayed form of gastroesophageal reflux episodes (Fig. 3). In order to avoid this, synoptic graphs of the time course from oesophageal pH-metry studies should display extreme values instead of means or medians. The best method is represented by initial determination of the median of each short time interval, selecting the minimum value if the median is below 7 or the maximum if the median is above 7. This procedure also ensures that acid as well as alkaline reflux episodes are displayed correctly.

Clinical aspects

Oesophageal pH-metry
In adults the pH-probe is normally positioned 5 cm proximal to the

Fig. 3 Examples of different methods of data reduction required for diagrammatic presentation of a full 24 hour oesophageal pH recording. The raw data plot is time consuming and may obscure information about single reflux episodes – for example, sampling frequency 30 min, above. When using means – for example, during intervals of 1 minute, short reflux episodes are misrepresented (middle). Appropriate representation of the raw data can be achieved by using extreme values instead of means (below).
Techniques for location of the pH-probe in order of preference are: manometry, pH-profile, endoscopy, fluoroscopy, and potential difference. Only in patients with normal anatomy of the cardia are manometry, pH-profile, or endoscopy all acceptable. When assessing the distance to the lower oesophageal sphincter endoscopically, it should be realised that nasal intubation of the electrode requires up to 5 cm more cable than oral intubation. Movement of the probe after proper placement is not a major problem in the oesophagus.

The sensory stimulation of the pharynx by the electrode cable stimulates salivary secretion which may result in faster oesophageal acid clearance. Because this effect subsides after four to six hours and swallowing frequency is not affected, 24 hour studies should not be substantially influenced. The electrode cable may also influence nocturnal reflux by interfering with normal sleep patterns. Both effects should be minimised by making the cable as smooth and thin as compatible with easy intubation and maintenance of electrode position. Simultaneous oesophageal and gastric pH-metry appears to be feasible as in volunteers a cable passing through the lower oesophageal sphincter with a diameter of 3 mm did not influence gastro-oesophageal reflux.

Liquids and solids with a pH less than 4 – for example, many soft drinks, citrus fruits, mixed pickles must not be ingested during oesophageal pH-metry. Temperature affects pH by various independent mechanisms – for example, probe performance, ionic mobility and dissociation constant. Ingestion of hot or cold food and drink can thus influence pH readings, particularly in the oesophagus, although this effect is generally small (less than 0.3 pH units).

For clinical purposes, the term ‘reflux time’ is defined as the percentage of time the probe registers a pH below 4. Because this variable is usually not normally distributed among groups of patients, means and standard deviations are inappropriate. The correct upper confidence limit of normal values should be determined using frequency analysis or by determining discriminatory power in a prospective way. Upper limits of normal reflux time from different centres are given in the Table.

Table 1  Upper limit of normal for percent time with pH below 4 during oesophageal long term pH-metry. Although this variable generally was not normally distributed, most authors expressed their results as mean plus a measure of variation (SEM, SD or 2 SD). For the sake of uniformity, these values have been transformed into mean plus two times standard deviation (+2 SD). In other studies, the upper limit of the range (RG) or the result of a receiver-operating-characteristic analysis (ROC) are given.

<table>
<thead>
<tr>
<th>Upper limit of normal (% time with pH below 4)</th>
<th>Duration of recording (h)</th>
<th>Adults/children</th>
<th>Remarks</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>Upright</td>
<td>Supine</td>
<td>15</td>
<td>A</td>
<td>+2SD</td>
</tr>
<tr>
<td>3-8</td>
<td>—</td>
<td>3-1</td>
<td>—</td>
<td>12</td>
<td>RG</td>
</tr>
<tr>
<td>1-9</td>
<td>2-8</td>
<td>—</td>
<td>24</td>
<td>A</td>
<td>+2SD</td>
</tr>
<tr>
<td>8-6</td>
<td>8-8</td>
<td>10-1</td>
<td>24</td>
<td>A</td>
<td>+2SD</td>
</tr>
<tr>
<td>4-2</td>
<td>6-3</td>
<td>1-2</td>
<td>24</td>
<td>A</td>
<td>+2SD</td>
</tr>
<tr>
<td>—</td>
<td>5-9</td>
<td>4-6</td>
<td>24</td>
<td>A</td>
<td>+2SD</td>
</tr>
<tr>
<td>5-9</td>
<td>10-5</td>
<td>6-0</td>
<td>24</td>
<td>A</td>
<td>ROC</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>—</td>
<td>24</td>
<td>C</td>
<td>+2SD</td>
</tr>
</tbody>
</table>
The duration of a period when pH is below 4 is termed a 'reflux episode'. Using this variable, the number of reflux episodes as a function of time, mean and maximal duration of reflux episodes, and the number of reflux episodes with a duration exceeding a certain threshold – for example, five minutes, can be calculated. It must be realised, however, that the number of reflux episodes depends to some degree on the sampling frequency of the recorder (Fig. 1a) and on the analysis procedures included in the evaluation program. For example, some programs make use of a hysteresis, with different threshold values for the onset and end of a reflux episode, in order to avoid an erroneously high count of reflux episodes when there are minor oscillations of pH about the threshold value. Consequently, 'normal' values for the number of reflux episodes and derived parameters can only be quoted for individual centres. Reflux time is probably a more sensitive parameter than reflux episode derived variables for diagnostic purposes.

**Gastric pH-metry**

As pH can markedly vary between different parts of the stomach, it is vital that the electrode is maintained in a known and reproducible position. The electrode position in the stomach should be referenced to the gastric topography rather than the lower oesophageal sphincter. Location of the probe should be determined frequently, particularly during experimental studies. Positioning the electrode under fluoroscopy is highly recommended. Initial radiology cannot always ensure accurate position, however, especially in the antrum or after food intake.

Movement of the probe into the fundic air bubble is a particular problem associated with gastric pH-metry. Although the electrode continues to register pH due to presence of a thin fluid over the electrode, it is important to realise that this may no longer reflect the pH of the luminal content. This applies equally to unipolar and combination electrodes.

The degree of standardisation that is desirable for gastric pH-monitoring depends on the aim of the study. A number of general recommendations can be given, however. Meals must be standardised with regard to composition, volume and time of administration. Alcoholic drinks, coffee, smoking, and drugs must be avoided if possible or else documented. Finally, subjects must record periods of work, physical exercise and sleep.

Comparable normal values of gastric intraluminal pH are not available as to date neither the protocol nor the evaluation procedures have been standardised.

**Duodenal pH-metry**

The techniques for duodenal pH-metry are least well defined and relatively few studies exist. As data for its validation are not available, it cannot be performed without repeated checks on electrode position. Ideally location of the electrode should be checked continuously. Changes in body position, food intake, and gastrointestinal motility can all cause electrode displacement so producing large pH-changes because of the extremely steep pH-gradient along the length of the proximal duodenum and across the pylorus. Large balloons which are used to anchor duodenal electrodes may alter motility and hence markedly influence pH in the bulb. Even when intragastric acidity is high (pH 2 or below), the gastroduodenal junction cannot always be localised continuously from the pH-gradient across the
pylorus. The gastroduodenal junction can be localised continuously by measuring electrical potential difference between duodenum and skin or buccal mucosa. For this purpose, intraluminal pH can only be measured by a combination glass electrode where the distance between glass and reference elements is reduced to the minimum. Fluoroscopy, ultrasonography or measurement of pressure profiles only enable intermittent localisation of the gastroduodenal junction.

Duodenal pH is characterised by predominantly alkaline values (between pH 7 and 8) with brief periods of very acidic values after gastric emptying.32-38 There is no general agreement on the normal duodenal pH or the specific changes occurring in different diseases.

Conclusions and recommendations

Oesophageal pH-metry has been extensively investigated and is regarded as a reliable diagnostic aid. We recommend carrying out 24 hour studies with an electrode positioned 5 cm above the manometrically determined lower oesophageal sphincter. Use of the percentage of time with pH below 4 (for which normal values exist) is probably the most reliable indicator for the detection of a pathological gastroesophageal reflux. An overall accuracy of at least 0.5 pH units and a sampling frequency of at least 10/min is regarded as sufficient. These features are available in nearly all devices on the market irrespective of the type of electrode used. Inexpensive and therefore disposable electrodes small enough for use in the newborn are also recommended.

In contrast, there is insufficient information on which to make firm recommendations for gastric and duodenal pH-metry. As a result, the recording device and electrode should have the greatest accuracy possible with an error of less than 0.1 pH-units. At present, this can be obtained only using glass, ISFET or plastic electrodes and precise calibration procedures. Some commercially available pH data loggers fail to meet the latter requirement. The utility of intraluminal pH-metry in diagnosis and management of disorders of the stomach and duodenum remains to be established.

The authors wish to acknowledge Dr R Bumm, Dr M Cucala and Ms J Poppen for their invaluable help in organising the meeting. This work was supported by ‘Deutsche Forschungsgemeinschaft’ grant # Em 36/1-3 and ‘Swiss National Foundation’ grant # 3.827.0.86. The meeting was supported by Glaxo and the following manufacturers of pH equipment: Ingold, Kaufhold, M & M, Oxford, Radiometer, Sandhill, Sitas and Synectics.

Original statements and data header for transfer of pH-data

The original statements with the results of the voting as well as a data header for transfer of data from pH-metry studies between different research groups developed during the meeting39 can be obtained from C Emde upon request.

C Emde, A Garner, and Al Blum

Division of Gastroenterology,
Klinikum Steglitz der FU Berlin,
Hindenburgdamm 30, D-1000 Berlin 45, FRG,
Bioscience Dept, ICI Pharmaceuticals Division,
Intraluminal pHmetry

Alderley Park, Macclesfield, Cheshire, and Division of Gastroenterology, Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland.

Edited by (alphabetical order): F Baldi, Bologna, Italy; T R DeMeester, Omaha, Nebraska, USA; J Dent, Adelaide, Australia; T Gasser, Mannheim, West Germany; J Jansen, Leuven, Belgium; R F McCoy, Manchester, UK; S A Muller-Lissner, Munich, West Germany, and S J Rune, Glostrup, Denmark.

Participants (alphabetical order): S Aggestrup, Copenhagen, Denmark; P Bauerfeind, Lausanne, Switzerland; C S Clark, Littleton, Colorado, USA; G VanDeventer, Los Angeles, California, USA; D F Evans, Nottingham, UK; C J Fimmel, Los Angeles, California, USA; K Fuchs, Kiel, West Germany; M Hannibal, Herlev/Copenhagen, Denmark; B C Hurst, Macclesfield, UK; A Holscher, Munich, West Germany; B Joelson, Lund, Sweden; B Kapur, Rotherham, UK; HJ Kauffhold, Berlin, West Germany; H R Koetz, Zurich, Switzerland; G McMachlan, Glasgow, UK; T Lendau, Chicago, Illinois, USA; J M G Melkert, Putegni Vitvoorde, Belgium; S Menghino, Padova, Italy; U Oesch, Zurich, Switzerland; F Pacic, Milano, Italy; B Rosenkilde, Glostrup, Denmark; J Rohmel, Berlin, West Germany; K Saksager, Copenhagen, Denmark; T Sanderson, Abingdon, Berks, UK; W Schweizer, Omaha, Nebraska, USA; W Simon, Zurich, Switzerland; H Smole, Dissenhefen, Switzerland; H Tschudin, Urdorf, Switzerland; H F Weiser, Munich, West Germany, and W Wu, Winston-Salem, North-Carolina, USA.

References


