PAPERS

Endoscopic assessment of oesophagitis: clinical and functional correlates and further validation of the Los Angeles classification

L R Lundell, J Dent, J R Bennett, A L Blum, D Armstrong, J P Galmiche, F Johnson, M Hongo, J E Richter, S J Spechler, G N J Tytgat, L Wallin

Department of Surgery, Sahlgrenska University Hospital, Sweden
L R Lundell

Gastrointestinal Medicine, Royal Adelaide Hospital, Adelaide, South Australia
J Dent

Remenham House, Hull, UK
J R Bennett

Division de Gastro-entérologie, Centre Hospitalier, Universitaire Vaudois, Lausanne, Switzerland
A L Blum

Division of Gastroenterology, McMaster University Medical Centre, Hamilton, Canada
D Armstrong

Department of Gastroenterology, Hotel-Dieu—CHU Nantes, France
J P Galmiche

Department of Surgery, Lund's University Hospital, Sweden
F Johnson

Tohoku University, School of Medicine, Sendai, Japan
M Hongo

Department of Gastroenterology, Division of Medicine, Cleveland Clinic Foundation, Cleveland, Ohio, USA
J E Richter

Department of Medicine, Beth Israel Hospital, Boston, USA
S J Spechler

Abstract

Background—Endoscopic oesophageal changes are diagnostically helpful and identify patients exposed to the risk of disease chronicity. However, there is a serious lack of agreement about how to describe and classify the appearance of reflux oesophagitis

Aims—To examine the reliability of criteria that describe the circumferential extent of mucosal breaks and to evaluate the functional and clinical correlates of patients with reflux disease whose oesophagitis was graded according to the Los Angeles system.

Methods—Forty six endoscopists from different countries used a detailed worksheet to evaluate endoscopic video recordings from 22 patients with the full range of severity of reflux oesophagitis. In separate studies, Los Angeles system gradings were correlated with 24 hour oesophageal pH monitoring (178 patients), and with clinical trials of omeprazole treatment (277 patients).

Results—Evaluation of circumferential extent of oesophagitis by the criterion of whether mucosal breaks extended between the tops of mucosal folds, gave acceptable agreement (mean κ value 0.4) among observers. This approach is used in the Los Angeles system. An alternative approach of grouping the circumferential extent of mucosal breaks as occupying 0–25%, 26–50%, 51–75%, 76–99%, or 100% of the oesophageal circumference, gave unacceptably high interobserver variation (mean κ values 0–0.15) for all but the lowest category of extent (mean κ value 0.4). Severity of oesophageal acid exposure was significantly ($p<0.001$) related to the severity grade of oesophagitis. Pretreatment oesophagitis grades A–C were related to heartburn severity ($p<0.01$), outcomes of omeprazole (10 mg daily) treatment ($p<0.01$), and the risk for symptom relapse off therapy over six months ($p<0.05$).

Conclusions—Results add further support to previous studies for the clinical utility of the Los Angeles system for endoscopic grading of oesophagitis.

(*Gut* 1999;45:172–180)

The endoscopic oesophageal changes caused by reflux disease are not only helpful diagnostically, but also identify patients exposed to a significant risk of disease chronicity.[1–4] Furthermore, the severity of oesophagitis gives useful guidance as to the likelihood of success of particular treatments.[5]

Unfortunately, there is a serious lack of agreement about how to describe and classify the appearances of reflux oesophagitis. There are shortcomings in the many different published systems. For example, many of these systems use ambiguous terminology that is variably interpreted by different endoscopists; some systems suffer from illogicalities; and many doctors grade appearances that are now known to be unreliable for the diagnosis of reflux oesophagitis, such as oedema, increased vascularity, etc.[6 7] It is probably these deficiencies that lead some endoscopists to use "personal" systems. The lack of agreement about how to describe the endoscopic appearances of reflux oesophagitis arises from a deficiency of critical evaluation of approaches to endoscopic grading of oesophagitis. This hampers both accurate communication about the clinical status of individual patients and the interpretation of data from clinical trials of the treatment of the disease.[8] Our international working group on the endoscopic classification of oesophagitis, which is supported by the World Organisation of Gastroenterology (OMGE), developed and proposed the Los Angeles Classification System in 1994[6] and reported its first evaluation of the criteria used. The system was so named because it was presented at the Los Angeles World Congress of Gastroenterology. The working group has continued to meet regularly in order to evaluate the proposed Los Angeles classification and as a result of these discussions has agreed that the definitive version of the classification should be as given in table 1. This version has been revised slightly from the previously published proposal,[6] in response to the concern that it is sometimes difficult to

Table 1   The Los Angeles classification of oesophagitis

Grade A	One (or more) mucosal break no longer than 5 mm, that does not extend between the tops of two mucosal folds
Grade B	One (or more) mucosal break more than 5 mm long that does not extend between the tops of two mucosal folds
Grade C	One (or more) mucosal break that is continuous between the tops of two or more mucosal folds but which involves less than 75% of the circumference
Grade D	One (or more) mucosal break which involves at least 75% of the oesophageal circumference

determine with confidence whether a mucosal break is completely circumferential or not. The change of grade D from fully circumferential to involvement of more than 75% of the circumference has required a matching modification of the definition of grade C.

The criteria used by the Los Angeles classification focus on the description of the extent of visible mucosal breaks in the belief that this is of greatest diagnostic and prognostic value. In the previous study from this group,[6] which used video recording and endoscopic photographs, we found that endoscopists were able to identify mucosal breaks confined to the tops of mucosal folds, lesions that extended around the entire oesophageal circumference, and complications of oesophagitis such as stricture and columnar lined oesophagus. However, our original assessment of the utility of criteria for the assessment of the radial extent of mucosal breaks was inconclusive because of deficiencies in the design of the sheet used to score particular endoscopic features.[6] This was a major limitation, as definition of circumferential extent was judged to be a key measure of oesophagitis severity. As a result of this experience, the score sheet for assessing this variable was developed substantially and the revised version used in the further evaluations presented in this report.

A system that classifies the endoscopic changes of reflux oesophagitis should not only be unambiguous and simple to use, but should also be shown to distinguish between clinically relevant grades of severity of reflux disease. Thus the grading should distinguish among groups of patients with differing responses to treatment and with differing severity of reflux, as assessed by other measurements. These correlates have been assessed for the Los Angeles system for the first time in the present study.

The study had therefore two major aims: to examine the reliability of criteria that describe the circumferential extent of mucosal breaks; and to evaluate the functional and clinical correlates of patients with reflux disease whose oesophagitis was graded according to the Los Angeles system.

Academisch Medisch Centrum, AFDLG Maas, Darm-Enleverzeiken, Amsterdam, The Netherlands
G N J Tytgat

Department of Surgical Gastroenterology D2, KAS Glostrup, Glostrup, Denmark
L Wallin

Correspondence to:
Dr L Lundell, Department of Surgery, Sahlgrenska University Hospital, S-413 45 Göteborg, Sweden.

Methods

INTEROBSERVER AGREEMENT

Forty six endoscopists from different countries and continents (Europe, USA, Australia, and Japan) participated. Both trainee endoscopists, defined as those who had performed less than 500 upper gastrointestinal endoscopies, and experts who had performed more than 3000 upper gastrointestinal endoscopies were invited to take part. Of the participants, 25 were classified as experts and 15 as trainees. Information about the level of experience was unavailable for six contributors. As we were unable to reveal any significant differences in outcome between these groups, we present data retrieved when analysing the entire study group in the Results section. Each participant received a videotape which contained edited recordings of endoscopic images from 22 cases, each lasting approximately 30 seconds. The full range of severity of reflux oesophagitis was covered by the recordings. Endoscopists used a detailed worksheet (shown in full in the Appendix), which was a development of the one used in previous studies. The sheet scored the full range of findings relevant to the evaluation of an oesophagitis classification system, according to predefined criteria. The previously agreed definition of mucosal breaks, "an area of slough or erythema with a discrete line of demarcation from the adjacent more normal looking mucosa" was used.[6] The original videotapes were acquired with Olympus video endoscopes and either a Super-VHS or NT video recorder. The copies that were distributed were made directly from these original recordings. Five centres participated in producing original video material.

The most widely used coefficient of agreement in clinical studies, the κ statistic,[9] was used in its original version in order to evaluate the degree of agreement among observers. The range of possible values for κ is from -1, which indicates perfect disagreement, to $+1$, which is reached with perfect agreement. Agreement by chance alone gives a value of 0. Though far from perfect, values as low as 0.4 are considered to indicate acceptable agreement. In the κ statistics used, p0 is the observed proportion of agreement and pC is the expected agreement by chance in the relevant contingency table.

$$\kappa = \frac{p0 - pC}{1 - pC}$$

We analysed the data from the large number of observers as a multiple of comparisons between each pair of observers. That is, if 46 observers recorded the presence or absence of a particular feature, κ statistics were calculated for a total of $[n \times (n - 1)]$ comparisons. This means that a theoretical number of 2070 κ values have been computed for each item. Each κ value measured the agreement between two observers in the endoscopic videotape recordings in the 22 patients; κ values are given as medians and interquartile ranges.

FUNCTIONAL AND CLINICAL CORRELATES

Study I—Oesophageal pH monitoring

These data are a part product of a study into the use of omeprazole as a diagnostic test, which has been reported in detail elsewhere.[10] Patients were recruited into the study if they had heartburn as the main symptom for at least

six months, or if their major symptom was acid regurgitation, abdominal pain, or discomfort in association with heartburn. In total, 178 patients were entered into the study from 17 centres in Denmark, Norway, and Sweden. The patients scored their reflux induced symptoms in a structured manner during the last two days before enrolment into the study as either mild, moderate, or severe. At endoscopy, oesophagitis was graded according to the Los Angeles classification. Before entry into the therapeutic trial ambulatory 24 hour pH monitoring was carried out with a monocrystalline, antimony electrode positioned 5 cm above the oral margin of the lower oesophageal sphincter, which was located by stationary oesophageal manometry. During the pH study, the patients were asked to act as normally as possible in order to minimise any impact on reflux patterns that could result from the process of measurement.[8]

Study II—Clinical correlates
In the other clinical study, the major therapeutic outcomes of which have been reported elsewhere,[11] patients with upper gastrointestinal discomfort were recruited from the primary care setting. The patients were screened for reflux symptoms by use of a questionnaire,[12] which had been developed for the diagnosis of symptomatic reflux disease, and which focuses especially on objective recognition of heartburn, and on exacerbating or relieving factors. Patients with a questionnaire score of 4 or more, aged 18–80 years, were screened by endoscopy for enrolment into the study, provided that they had been experiencing upper gastrointestinal symptoms for at least three months with episodes occurring on at least four days during the previous week. At endoscopy, oesophagitis was graded according to the Los Angeles classification (table 1). For ethical reasons, patients found to have Los Angeles grade D oesophagitis, columnar lined oesophagus, or peptic strictures were excluded from further involvement in the study. Complete endoscopic and clinical data were available for 496 of the 538 patients randomised to treatment (see below). Of these, the 277 who had mucosal breaks were randomised to either omeprazole 20 or 10 mg daily. Patients without mucosal breaks were randomised to either omeprazole 20 or 10 mg daily or to placebo. Treatment was given for four weeks, under double blind conditions, after which it was stopped in those who had complete symptom relief. Those with persistent symptoms at four weeks were given open therapy with omeprazole 20 mg daily for another four weeks. All patients in whom symptoms were relieved at four or eight weeks entered a follow up period without treatment for up to six months, exiting the study if their symptoms relapsed according to predefined criteria before that time. There were 145 patients who had mucosal breaks at their pretreatment endoscopy who entered this phase of the study.
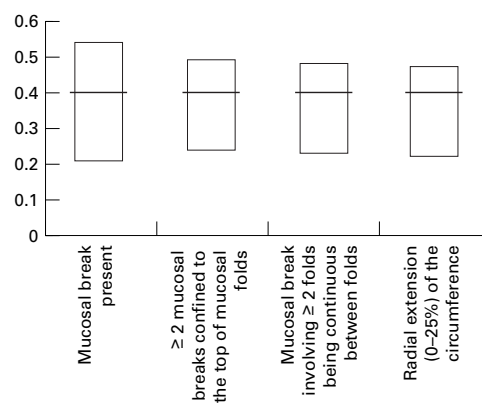

Figure 1   κ values for interobserver agreement on the presence and extent of mucosal breaks (median values and interquartile ranges).

STATISTICS
The relation between acid reflux variables and endoscopic grading of mucosal breaks was assessed with a logistic regression analysis in which the endoscopic grade was the dependent variable. The relation between symptom severity and endoscopic grading was evaluated by cross tabulation of the severity of heartburn and endoscopic grading and then a simple correlation analysis. A log rank test was applied to the proportion of patients in clinical remission during follow up with the pretreatment endoscopic grading as the dependent variable.[13]

Results
INTEROBSERVER AGREEMENT
There was no significant difference in the levels of agreement on judgements described below between experienced endoscopists and trainees, so the entire data set was pooled. The most important aspect of the endoscopic video sequence assessments was the reproducibility of evaluation of the circumferential extent of mucosal breaks as judged by the extension of mucosal breaks across the tops of two or more mucosal folds. Such mucosal breaks were assessed with the κ value of 0.4 (0.22–0.51) (fig 1).

When endoscopists were asked to evaluate the proportion of the circumference involved by mucosal breaks as 0–25%, 26–50%, 51–75%, 76–99%, or 100%, only the most limited extent (0–25%) was recorded with any level of accuracy (κ = 0.4, range 0.22–0.52). In the remaining groups, κ values equalled or were close to 0. Another aspect of assessing the circumferential extent was the assessment of whether one or more mucosal folds were involved by the mucosal breaks. The examiners again showed an acceptable agreement, scoring a κ value of 0.4 in this judgement.

The score sheets also allowed assessment of interobserver variation in the recognition of other aspects of the endoscopic oesophageal appearances already evaluated in our previous study.[6] Agreement on the presence or absence of a mucosal break in a particular endoscopic video sequence had a κ value of 0.4 (0.21–0.58). The presence of multiple mucosal breaks (two or more) was assessed with a similar level of accuracy (κ value 0.4 (0.25–0.54)).

Table 2 κ values for the assessments of minimal changes in the distal oesophagus and the gastro-oesophageal junction

Endoscopic features	κ value	
Excessive reddening of the cardia	0.09	(0.00–0.29)
Erythema of the squamocolumnar junction	0	(−0.08–0.06)
Friability of the squamocolumnar junction	0.04	(0.00–0.25)
Blurring of the squamocolumnar junction	0.08	(0.00–0.22)
Diffuse erythema of the distal oesophagus	0.08	(0.00–0.28)
Patchy erythema of the distal oesophagus	0.00	(−0.07–0.09)
Increased vascularity of the distal oesophagus	0.00	(0.00–0.09)
Oedema or accentuation of the mucosal folds	0.00	(−0.02–0.15)

Median κ values are given, as well as the first and third quartile.

Table 3 Acid reflux parameters in relation to endoscopic grading of mucosal breaks

	Number of patients	% time pH <4	Number of reflux episodes
Normal	40	6.7 (1.0)	59.5 (12.8)
Grade			
A	50	9.3 (1.3)	106.9 (12.8)
B	50	13.7 (2.1)	138.3 (20.0)
C	9	11.7 (3.0)	135.8 (29.4)
D	10	19.1 (6.6)	163.0 (41.7)
C + D	19	15.6 (3.8)	150.1 (25.5)

Results expressed as mean (SEM).
Normal versus grade A, B, C, D, p=0.0007; normal + grade A versus grade B, C, D, p=0.003; normal + A, B versus grade C, D, p=0.01.

The presence of a stricture was assessed with a κ value of only 0.33 (0–0.46) but there were only two cases of stricture among the 22 cases. For the four cases of columnar metaplasia (all histologically confined), recognition of tongues that extended up the oesophagus had a κ value of 0.38 (0.13–0.58) and estimates of the circumferential extent of this metaplastic epithelium had a κ value of 0.42 (0.10–0.65). The investigators were specifically asked to assess the number and depth of ulcers, as opposed to mucosal breaks, with an ulcer being defined as a mucosal break greater than 3 mm deep. Reliable judgement on the presence of an ulcer was not achieved as this assessment had a κ value of −0.01 and the corresponding κ value for the assessment of the number of ulcers only reached −0.02.

The presence and absence of "minimal endoscopic changes" as listed in table 2 was evaluated close to the oesophagogastric mucosal junction; there was no agreement on the presence or absence of these findings.

Figure 3 The relation between pretreatment endoscopic grading of mucosal breaks and healing of these mucosal breaks after four weeks of omeprazole 10 or 20 mg daily.
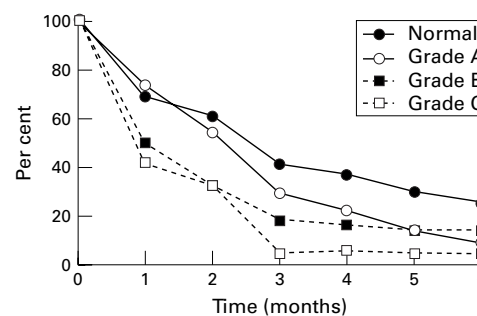
Figure 4 Percentage of patients in clinical remission during six months, follow up after initial treatment with omeprazole, which had healed the oesophagitis and controlled symptoms. The symptom relapse curves are given for each pretreatment endoscopic grade. Normal versus grade A, p=0.04; normal versus grade A + B + C, p=0.002; normal + grade A + B versus grade C, p=0.003.

FUNCTIONAL AND CLINICAL CORRELATES

Oesophageal pH monitoring and symptom assessments were completed in 178 patients as a prelude to inclusion in the short term therapeutic trial with omeprazole (table 3). There was a significant relation between the Los Angeles classification grade and the 24 hour oesophageal acid exposure values (p<0.01).

There was also a significant correlation between pretreatment symptom severity and endoscopic grading of mucosal breaks (fig 2).
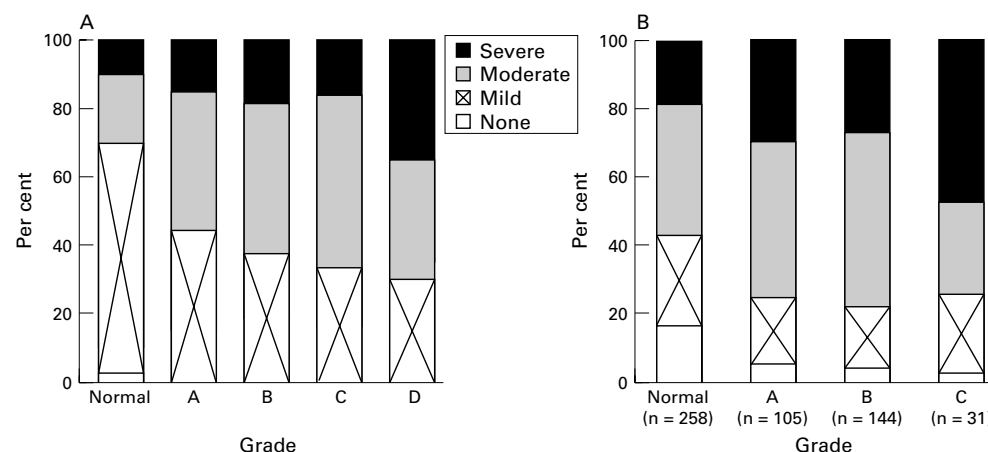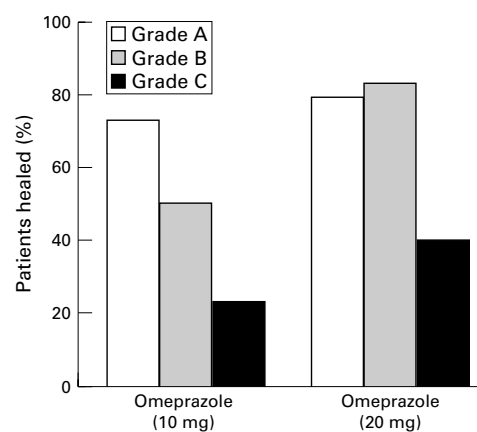

Figure 2 Symptom severity as related to endoscopic grading of mucosal breaks in study I (A) and study II (B). Symptoms were assessed with focus on the severity of heartburn.

The severity of heartburn correlated significantly with the pretreatment endoscopic grade in both studies (study I: $r=0.31$, p<0.001; study II: $r=0.23$, p<0.01).

After four weeks of treatment with omeprazole 10 mg daily, there was a gradation of endoscopic healing rates from 77% of grade A patients to only 20% in grade C patients (fig 3). There was a gradation of healing rates with omeprazole 20 mg, but this time only between grades B and C.

Symptomatic relapse after initial short term treatment occurred in 83% of the patients during a six month follow up period. The proportion of patients still in clinical remission during these six months showed a significant relation to the pretreatment Los Angeles classification grade of mucosal breaks (fig 4).

Discussion

This report provides data which support the approaches used by the Los Angeles classification to categorise the extent of oesophagitis. Of special interest was the evaluation of the circumferential extent of the mucosal breaks. The data were generated by a newly structured worksheet and new endoscopic images better directed to the aim of the assessment than the one used previously.[6] These results indicate that endoscopists can score the joining of mucosal breaks between the tops of two mucosal folds with fair agreement, thus allowing scoring of circumferential extent with this method. Determination of radial extent without reference to mucosal folds had very poor levels of agreement. This report gives clinical and physiological data, which support the practical relevance of the Los Angeles classification and also provide the definitive description of the finally agreed form of the Los Angeles classification.

The version of the Los Angeles system described in the previous report of this group[6] was a proposed system still under evaluation, which has now been modified slightly from the original description (table 1). If, in the future, further data suggest that the Los Angeles classification could be improved by some evolution of the criteria, it is important that even a minor modification should be given an entirely different name, in order to avoid any ambiguity as to what criteria have been used. This is an important lesson, which has not been learnt from the Savary and Miller grading system.[14–17] The first version of this grading system differs substantially from the second, yet it is rarely stated which version is being used, either in clinical practice or even in published clinical trials. If the version in use is stated, the reader must be familiar with the detail of each version.

The major aim of the assessments of the endoscopic images presented in this report was to evaluate further how reliably endoscopists could assess the circumferential extent of mucosal breaks by localising the peaks of the mucosal folds. It was believed that the findings of the assessment from our previous study[6] could have been influenced by technical limitations of the assessment method. The present study gives a more positive indication that the peaks of the oesophageal mucosal folds are useful endoscopic landmarks for determination of extent. Another important message, related to the clinical usefulness of the system, was that no apparent differences emerged between experienced endoscopists and trainees. We would have liked to have obtained a stronger κ value than 0.4 to support our recommendation for use of the peaks of the mucosal folds as a primary method for determining radial extent. Determination of radial extent is, however, a necessary judgement for endoscopic classification of oesophagitis, and the peaks of the mucosal folds appear to be the best option on the basis of the data obtained.

The quality of the stored images continued to be a significant technical limitation. The stored images did not fully emulate the sensitivity of a live endoscopy done by an observer orientated to the use of the peaks of the mucosal folds as radial landmarks. Furthermore, despite our best efforts, the resolution of some of the endoscopic images was not optimal. Image quality was also impaired by the processes of image storage and copying. The stored images sent out for evaluation did not allow the observer to "explore" the oesophagus in order to clarify appearances. Such clarification should always include partial deflation of the oesophagus to recognise the position of the folds. Similar limitations with the current methodology might explain the differences in κ values between the present investigation and that performed by Bytzer *et al*.[7] However, significant advances in the optical resolution of video endoscopes, and in the capture, storage, and reproduction of video endoscopic images in digital form are currently being introduced. We believe that these technical developments will largely overcome the problems caused by suboptimal resolution of images as currently used. Our working group plans to use these developments for preparation of educational material, but judged that it was inappropriate to delay formal introduction and evaluation of the Los Angeles classification system pending the results of a further assessment done with technically better images. It needs, however, to be emphasised that the intraobserver variability in these situations also requires further validation.

Previously, no attempt has been made to define the endoscopic appearances that should be taken as indicative of erosion or ulceration. We have described these collectively as "endoscopic mucosal breaks" for reasons given previously.[6] The definition, "an area of slough or erythema with a discrete line of demarcation from the adjacent, more normal looking mucosa", still requires interpretation by the observer, but it is at least more specific than no definition at all, which has been the case until now. Currently, some members of the working group are conducting a study which correlates endoscopic mucosal breaks with the histology of biopsies directed precisely to sample these areas. The group will also consider other measures that may aid the distinction of

patches or tongues of mucosal columnar metaplasia from mucosal breaks. Methods that hold some promise for this include high resolution video endoscopy and dye spraying methods.[18]

Although the classification of the extent of the mucosal break was the primary aim of our group, we included a scoring of so called minimal changes as well. κ values close to or below 0 were obtained for recognition of "minimal changes", in keeping with results from another group.[7] Again we have to realise the pivotal role of technical quality of images in order to assess these subtle mucosal appearances. At present, we consider that current evidence indicates that reliable recognition of "minimal changes" is problematical and that the diagnostic significance of these changes for reflux disease is uncertain. Possibly, better image quality may make these changes more reliable in the future. Until "minimal changes" are shown to be of diagnostic value, they probably should not be recorded as present or absent, as to do so only causes confusion.

The Los Angeles classification was developed with the intention to provide a clinically relevant stratification of the severity of oesophagitis. Despite the lack of validation of other oesophagitis classification systems there is considerable data indicating that these recognise clinically relevant variations of the severity of oesophagitis. This is the case for both oesophageal pH monitoring data and response to medical treatment.[19–21] The oesophageal acid exposure data included in the present study support the clinical relevance of the Los Angeles classification with the possible exception of the acid reflux values in grade C patients. However, data were available from only nine patients with this oesophagitis grade, a number that is probably too small to draw a meaningful conclusion, given the obvious variability of pH monitoring acid exposure data.

Large prospective, randomised clinical trials gave us the opportunity to evaluate further the clinical relevance of the Los Angeles classification. Previous studies, which have used different endoscopic classification systems to correlate the severity of symptoms with endoscopic grade, have given negative results.[22 33] In the present study there was a significant correlation between the severity of heartburn and the Los Angeles grading. The clinical studies also found an impressive gradation of endoscopic healing relating to the endoscopic grading, when omeprazole was given in a dose of 10 mg daily for four weeks. The large numbers of patients involved and the statistically significant differences are arguments for the clinical relevance of the Los Angeles classification. On the other hand, the size of the r values (r=0.23–0.31) emphasises that factors other than endoscopic grading play a significant role in the clinical manifestation of the disease. It is worth noting that the data obtained show that the classification of patients into grade A and B seems to provide a clinically relevant subdivision. These patients represent the most numerous groups of patients suffering from reflux diseases, which have previously been grouped

somewhat ambiguously into one grade by the Savary and Miller classification.[14] Additional information relevant to the significance of the Los Angeles classification system was gained from following patients after a successful initial, short term, drug therapy. The symptomatic relapse curves separated patients, who before treatment had an absence of oesophageal mucosal breaks, from those with Los Angeles grades A–C.

The findings presented in this and our previous report,[6] represent the only rigorous assessment of an endoscopic grading system for reflux oesophagitis. The criteria used for the Los Angeles classification have been developed and agreed on by a truly international working group. It is hoped that this classification will be used widely and without ad hoc modification, with resultant improvement in communication about patient status for both clinical and research purposes.

1 Spechler SJ. Epidemiology and natural history of gastro-oesophageal reflux disease. *Digestion* 1992;**51**(suppl 1):24–9.
2 Bennett JR, Chapman RWG, Harvey RF, *et al*. Management guidelines for gastro-oesophageal reflux disease. *The European Digestive Disease Week*, Amsterdam, 1991:46–54.
3 Spechler SJ. Complications of gastro-oesophageal reflux disease. In Castell DO, editor. *The oesophagus*. Boston: Little Brown and Co., 1992:543–56.
4 Dent J, Bremner CG, Collens MJ, *et al*. Barrett's oesophagus. Working party report to the World Congress of Gastroenterology, Sydney 1990. *J Gastroenterol Hepatol* 1991;**6**:1–22.
5 Lundell L. Long-term treatment of gastro-oesophageal reflux disease with omeprazole. *Scand J Gastroenterol* 1994;**29**(suppl 201):74–8.
6 Armstrong D, Bennett JR, Blum AL, *et al*. The endoscopic assesment of oesophagits: a progress report on observer agreement. *Gastroenterology* 1996;**111**:85–92.
7 Bytzer P, Havelund T, Möller Hansen J. Interobserver variation in the endoscopic diagnosis of reflux oesophagitis. *Scand J Gastroenterol* 1993;**28**:119–25.
8 Armstrong D, Emde C, Inauen W, Blum AL. Diagnostic assessment of gastrooesophageal reflux disease: what is possible vs what is practical? *Hepatogastroenterology* 1992;**39**:3–13.
9 Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Measurement* 1960;**20**:37–46.
10 Johnsson F, Weywadt L, Solhaug JH, *et al*. One week omeprazole treatment in the diagnosis of gastrooesophageal reflux disease. *Scand J Gastroenterol* 1998;**33**:15–20.
11 Carlsson R. Gastroesophageal reflux disease (GERD) in primary care: an international study of different treatment strategies with omeprazole. *Eur J Gastroenterol Hepatol* 1998;**10**:119–24.
12 Carlsson R, Dent J, Bolling-Sternevald E, *et al*. The usefulness of a structured questionnaire in the assessment of symptomatic gastroesophageal reflux disease. *Scand J Gastroenterol* 1998;**33**:1023–9.
13 Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. *BMJ* 1992;**304**: 1491–4.
14 Savary M, Miller G. L'oesophage. *Manuel at atlas d'endoscopie*. Soleure: Garsman Ab, 1977.
15 Monnier P, Savary M. Contribution of endoscopy to gastro-oesophageal reflux disease. *Scand J Gastroenterol* 1984;**106**: 26.
16 Ollyo JB, Gontollier C, Brossard E, *et al*. La nouvelle classification de Savary des ooesophagites de reflux. *Acta Endoscopica* 1992;**22**:307.
17 Savary M, Miller G. The oesophagus. *Handbook and atlas of endoscopy*. Solothorn, Switzerland: Gassman AG, 1978.
18 Endo M, Hideck-Ide H. *Endoscopic staining in early diagnosis of oesophageal cancer*. Tokyo: Japan Scientific Societis Press, 1991.
19 Lundell L. Acid suppression in the long term treatment of peptic strictures and Barrett's oesophagus. *Digestion* 1992; **51**(suppl 1):49–58.
20 Ghillebert G, Demeyere AM, Jansson J, *et al*. How well can quantitative 24-hour intraesophageal pH monitoring distinguish various degrees of reflux disease? *Dig Dis Sci* 1995; **40**:1317–24.
21 Hatlebakk JG, Berstad A. Endoscopic grading of reflux oesophagitis: what observations correlate with gastro-oesophageal reflux? *Scand J Gastroenterol* 1997;**32**:760–5.

22 Galmiche JP, Bruley des Varannes S. Symptoms and diseases severity in gastro-oesophageal reflux disease. *Scand J Gastroenterol* 1994;**29**:62–8.
23 Pace F, Santalucia F, Bianchi Porro G. Natural history of gastro-oesophageal reflux disease without oesophagitis. *Gut* 1992;**32**:848–58.
24 Schindlbeck NE, Klauser AG, Berghammer G, *et al*. Three year follow up of patients with gastro-oesophageal reflux disease. *Gut* 1992;**33**:1016–19.
25 Tytgat GN, Nio CY, Schotborgh RH. Reflux oesophagitis. *Scand J Gastroenterol* 1990;**25**(suppl 175):1–12.
26 Baldi F, Ferrarini F, Langanesi A, *et al*. Acid gastro-oesophageal reflux and symptom occurrence. Analyses of some factors influencing their association. *Dig Dis Sci* 1989;**34**:1890–3.
27 Hetzel DJ, Dent J, Reed WD, *et al*. Healing and relapse of severe peptic esophagitis after treatment with omeprazole. *Gastroenterology* 1988;**95**:903–12.
28 Sandmark S, Carlsson R, Fausa O, *et al*. Omeprazole or ranitidine in the treatment of reflux oesophagitis. Results of

double-blind, randomized, Scandinavian multicentre trial. *Scand J Gastroenterol* 1988;**23**:625–32.
29 Koelz HR, Birchler R, Bretholz A, *et al*. Healing and relapse of reflux oesophagitis during treatment with ranitidine. *Gastroenterology* 1986;**91**:1198–205.
30 Tytgat GN, Molenaar CBH. Grading endoscopic severity in gastroesophageal reflux disease. *Rumanian Journal of Gastroenterology* 1994;**3**:91–3.
31 Tytgat GNJ, Hansen OJA, Carling L, *et al*. Effect of cisapride on relapse of reflux oesophagitis, healed with an antisecretory drug. *Scand J Gastroenterol* 1992;**27**:175–83.
32 Dent J, Yeomans ND, Mackinnon M, *et al*. Omeprazole versus ranitidine for prevention of relapse in reflux oesophagitis. A controlled double-blind trial of the efficacy and safety. *Gut* 1994;**35**:590–8.
33 Blum AL, Adami B, Bouzo MH, *et al*. Effect of cisapride on relapse of oesophagitis. A multinational placebo-controlled trial in patients healed with an antisecretory drug. *Dig Dis Sci* 1993;**38**:551–60.

Appendix
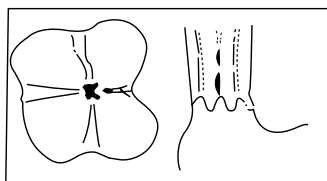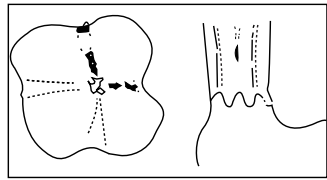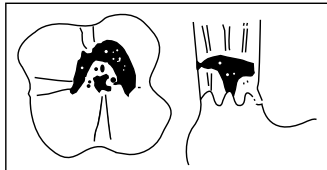
Endoscopic classification of reflux oesophagitis

Institution .... Investigator, Inititals .... Date .... / .... / ....

Patient no .... Country .... Expert: Yes/No
Trainee: Yes/No

Mucosal breaks

1. Is there a mucosal break?   1a .... Yes/No

If yes –

     How many breaks < 5 mm diameter?
         Only one   1b .... Yes/No
         Many (2 or more)   1c .... Yes/No
         Do all have overlying white/yellow membrane   1d .... Yes/No

     How many breaks > 5 mm diameter?
         Only one   1e .... Yes/No
         Many (2 or more)   1f .... Yes/No
         Do all have overlying white/yellow membrane   1g .... Yes/No

----

2. How many folds are involved by mucosal breaks?

Only one fold   2a .... Yes/No

More than one fold   2b .... Yes/No

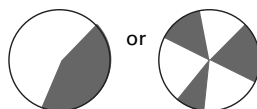Many folds (2 or more with mucosal breaks continuous between the tops of 2 or more folds, but not circumferential)   2c .... Yes/No

3. What proportion of the circumference is affected by mucosal breaks?

0–25%   or   3a .... Yes/No
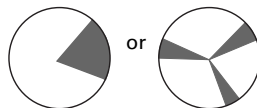
26–50%   or   3b .... Yes/No

51–75%   or   3c .... Yes/No

76–99%   or   3d .... Yes/No

100%   3e .... Yes/No
(circumferential)
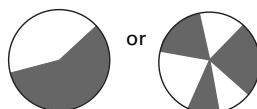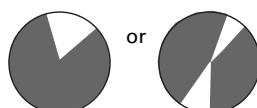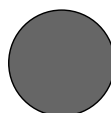
Ulcers
4. How many mucosal breaks are > 3 mm deep?

| | Only one | 4a .... Yes/No |
| | Many (2+) | 4b .... Yes/No |

----

5. Is there an ulcer?   5a .... Yes/No

How many ulcers are present?   Only one   5b .... Yes/No
                                 Many (2+)   5c .... Yes/No

Are the ulcers contiguous with more superficial breaks?   5d .... Yes/No

Complications
Stricture

6. Is there an oesophageal stricture?   6a .... Yes/No

If yes: 6b .... Yes/No

    Is the stricture narrower than the endoscope?
    (i.e. does the endoscope fail to pass the stricture)

Barrett's oesophagus/metaplasia
7. Is there excessive reddening at the cardia?   7 .... Yes/No
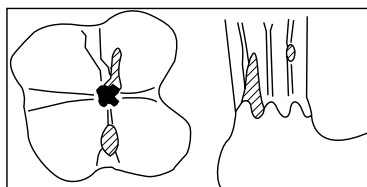
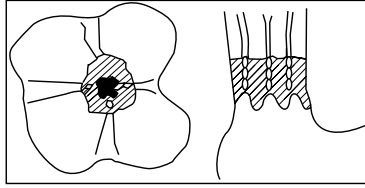8. Is columnar metaplasia (Barrett's epithelium) present?   8a .... Yes/No
   If yes:
       does it form fingers   8b .... Yes/No
         and/or islands?

Is it circumferential
(> 3 cm extent)?

8c .... Yes/No

(8a and 8b are NOT mutually exclusive)

Minor features suggestive of oesophagitis

9.   Is the distal oesophageal mucosa normal?   9a .... Yes/No

If it is not normal, is there;
— erythema in one segment at the squamocolumnar junction?   9b .... Yes/No

— blurring or indistinctness of the squamocolumnar junction?   9c .... Yes/No

— friability of the squamocolumnar junction?   9d .... Yes/No

— diffuse erythema in the distal oesophagus?   9e .... Yes/No

— patchy erythema in the distal oesophagus?   9f .... Yes/No

— increased vascularity in the distal oesophagus?   9g .... Yes/No

— oedema or accentuation of the mucosal folds?   9h .... Yes/No

Overall summary of endoscopic features of oesophagitis

10.   Mark relevant features of oesophagitis on diagrams:

cm	A	R	P	L
28				
24				
20				
16				
12				
8				
4				
0				