



ELSEVIER



2

Endoscopic grading of reflux oesophagitis: The past, present and future

John Dent*

Director

*The University of Adelaide, Nerve-Gut Research Laboratory, Department of Gastroenterology & Hepatology,
Royal Adelaide Hospital, North Terrace, Adelaide, SA 5000, Australia*

The Los Angeles Classification for diagnosis and grading of reflux oesophagitis, first discussed at the 1994 World Congress of Gastroenterology, was published in its definitive form in 1999. The product of a demanding validation process, it consistently predicts the outcome of acid suppressant therapy, correlates with oesophageal acid exposure, and is the most reproducible and practical of oesophagitis grading systems. The attributes of the classification, which enhance the specificity of communication on reflux oesophagitis, are widely recognized, as it is now by far the most widely used method for description of reflux oesophagitis. Exclusion of minimal oesophageal mucosal change is, however, regarded as a significant limitation by some users, especially in Japan. Some data suggest that minimal changes may now be recognized with modern endoscopes; if this claim can be adequately validated, formal addition of criteria for minimal change could significantly improve the sensitivity of endoscopy for reflux disease.

Key words: reflux oesophagitis; gastro-oesophageal reflux disease; endoscopy; non-erosive reflux disease; diagnosis; therapy.

REASONS FOR USE OF STANDARDIZED, VALIDATED ENDOSCOPIC CRITERIA FOR REFLUX OESOPHAGITIS

When the oesophagus is assessed endoscopically for reflux oesophagitis, two key judgements are needed: is reflux oesophagitis present and if so, what is its severity? As explained further below, these judgements are key because they will influence clinical management, first by making a diagnosis and second by describing the oesophagitis in a way that helps to guide its treatment,^{1,2} defines the level of risk that the oesophagitis poses to the patient³ and enables assessment of the adequacy of treatment.

* Corresponding author. Tel.: +61 8 8222 5207; Fax: +61 8 8222 5867.
E-mail address: john.dent@health.sa.gov.au

Endoscopic findings also need to be described in terms that are well defined so that what the endoscopist saw can be communicated with acceptable precision to other doctors at that time or in the future. These significant practical challenges are frequently underestimated and can only be addressed adequately by use of validated criteria.

WHY THERE ARE SO MANY PUBLISHED APPROACHES TO ENDOSCOPIC ASSESSMENT OF REFLUX OESOPHAGITIS

A survey of the literature by the Los Angeles Classification Group in the early 1990s found over 30 different sets of criteria for the endoscopic assessment of reflux oesophagitis. The existence of such a large number indicated a lack of satisfaction with existing options and underscores the fact that mere publication should not be taken as an endorsement of the validity of endoscopic criteria. The Los Angeles Classification is the only system which has undergone a structured process of development and evaluation, with subsequent description of these processes and their outcomes as the primary aim of peer-reviewed, original research papers.^{4,5} Most of the other systems have been published as unvalidated by-products of peer-reviewed papers that have had another primary aim, mainly the reporting of evaluations of therapy for reflux oesophagitis.⁶ Other systems have been published after even less rigorous processes, most notably the Savary–Miller grading system which was first published in a book, essentially without any formal development, validation studies or peer review.⁷

LESSONS ABOUT OESOPHAGITIS GRADING SYSTEMS FROM REVIEW OF THE LITERATURE

The first step in the development of the LA Classification was the analysis of grading systems in current use (Figure 1). On the basis of frequency of citation in the then recent literature, the Savary–Miller Classification⁷ was used most, followed by the so-called Hetzel/Dent system,⁶ then by *ad hoc*, simple and not formally defined descriptors of grade, such as ‘mild’, ‘moderate’ or ‘severe’. The MUSE Classification was published just as the formal LA Classification development process was starting.⁸ These and other published systems had useful attributes, but all suffered from at least one of the following main problems: inclusion of minimal changes, inappropriate influence on oesophagitis grade by presence of Barrett’s oesophagus, use of ambiguous or difficult to apply descriptors to guide grading, criteria that were essentially histopathological and could not be judged reliably by endoscopy alone, no validation data on interobserver agreement and clinical relevance, a level of complexity that made the system difficult to memorize and use in clinical practice, and existence of several versions of the ‘same’ classification system.

AN APPROACH TO DEVELOPMENT OF OPTIMIZED VALID ENDOSCOPIC CRITERIA

The major steps used by the International Working Group for the Classification of Oesophagitis (IWGCO) for development of the Los Angeles Classification are shown schematically in Figure 1. This process is described in generic terms because it has general applicability. It was the process also used by the IWGCO for development

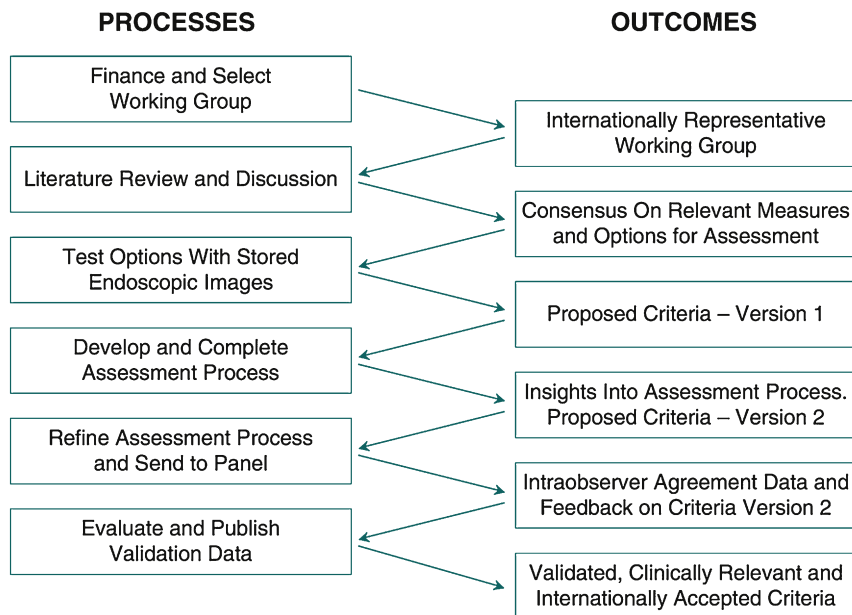


Figure 1. Schematic summary of the several processes and outcomes that led to development of the definitive Los Angeles Classification.

of the Prague C&M criteria (see chapter by P. Sharma). The likelihood of the development of criteria that are well matched to need was enhanced by recruiting members of the LA Classification Working Group from at least several regions of the world, in order to take into account a spectrum of views. The relatively high cost of bringing Working Group members together for several meetings in order to drive the processes and outcomes shown in Figure 1 is, however, a significant burden. The IWGCO has been fortunate in receiving support from AstraZeneca for the last 17 years, with additional funds being available for three years from the European Community, via the World Organisation of Gastroenterology. The multi-step iterative process summarized in Figure 1 is described in greater detail in the two original publications which present the development and validation of the Los Angeles Classification.^{4,5} Live endoscopy has been used for interobserver variability studies into the diagnosis and grading of reflux oesophagitis⁹ and has the major positive attribute of presenting endoscopic visualisation at the image quality normally experienced at endoscopy, but the substantial logistic and practical challenges of this approach led the IWGCO to use stored images for the LA Classification development work. Stored images also present significant practical challenges; only video recordings can come close to emulating a clinical diagnostic endoscopy, but it is only in the last few years that equipment has been available that preserves the image quality of conventional resolution endoscopes, especially when copies are made. Complete preservation of the quality of high resolution images as a video recording has been possible only very recently with the development of high definition (HD) recorders. Even with the best possible recording equipment, it is remarkably difficult to build a collection of high quality video recordings.

MAJOR DECISIONS ABOUT THE STRUCTURE OF THE LOS ANGELES CLASSIFICATION

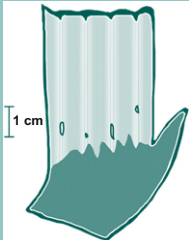
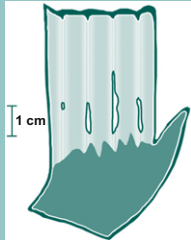
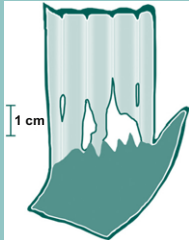
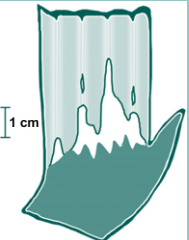
The Classification was named after Los Angeles since its major structures were first presented to a large audience at an interactive symposium at the 1994 World Congress of Gastroenterology in Los Angeles by members of the IWGCO. The audience gave valuable feedback for the final development process of the Classification.

The literature review and iterative process adopted resulted in the definitive version of the Los Angeles Classification.⁵ A graphic which illustrates the Classification can be downloaded from the IWGCO website (<http://www.iwgco.org>). To distinguish the LA Classification from most others, letters were chosen to code the severity grades from A to D. Definitions of these grades are given in Table I. The reasoning behind the major structures of the Classification is summarized briefly below.

Major structure I: do not use minimal changes as a criterion

It is important that endoscopic classifications correctly reflect the proven capacity of endoscopy for the recognition of reflux oesophagitis; that is, the criteria applied have good specificity, but naturally, it is also important to make use of the least severe changes that are diagnostic, to provide the best possible sensitivity. It is a vexed question as to whether so-called 'minimal changes' can be used to diagnose the presence of non-erosive reflux oesophagitis endoscopically with acceptable accuracy.^{4,5,9,10} There is a widespread belief that the so-called 'traditional' minimal changes of mucosal oedema, friability and erythema are indicative of non-erosive reflux oesophagitis. Our everyday experience of how the skin responds to injury without breakage of its surface has probably been influential in driving the widespread acceptance

Table I. Word description of the definitive version of the Los Angeles Classification.⁵

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

An explanatory graphic suitable for wall display is available at <http://www.iwgco.com>.

of minimal changes as being diagnostically valid. Savary and Miller implied that these changes are diagnostically useful without any supporting evidence.⁷ The most rigorous study of recognition of traditional minimal changes evaluated the ability of three endoscopists to recognize specific oesophageal non-erosive changes at live endoscopy with early 1990s endoscopes and found that there was unacceptably poor agreement on presence or absence of these (Figure 2).⁹ The IWGCO also evaluated the ability of endoscopists to recognize traditional minimal changes on video recordings as part of the LA Classification development work and found a similar lack of interobserver agreement.⁵ As discussed later in this article, it is possible that technical developments of endoscopes could now enable recognition of minimal changes, but this needs further research.

Major structure 2: define the minimum endoscopic lesion that is a reliable indicator of reflux oesophagitis

This lesion, called a 'mucosal break' was defined as 'an area of slough or erythema with a sharp line of demarcation from adjacent normal mucosa'.⁵ Essentially this word picture was used to distinguish an erosion, or near erosion, from less well-localized erythema. These lesions have also been described as 'red streaks'. No previous endoscopic classification of oesophagitis has attempted to give explicit guidance on a minimum lesion such as this. It was judged that it would be too restrictive to insist on the presence of visible slough before reflux oesophagitis could be diagnosed. This view is supported by demonstration of an elevated oesophageal acid exposure in symptomatic patients with red streaks.¹¹ Furthermore, a study of the histopathology of red streaks by IWGCO members showed major mucosal damage consistent with reflux-induced injury.¹²

Major structure 3: do not use 'ulcer' as a criterion for grading severity

Some grading systems use 'ulcer' as a measure of greater severity versus 'erosion'. The Working Group investigated their ability to agree on whether a mucosal break shown on endoscopic photographs was an ulcer or erosion (as defined pathologically). Whilst some mucosal breaks could be clearly identified as an ulcer on the basis of observer

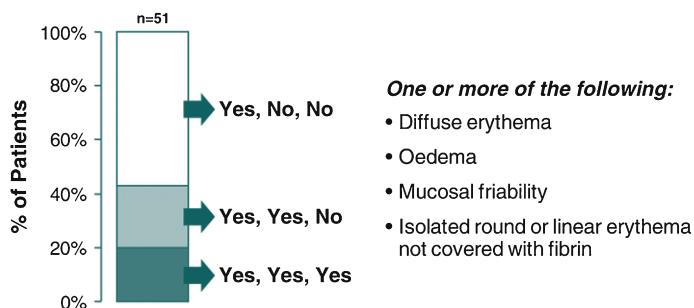


Figure 2. Summary of the major outcomes of the live endoscopy evaluation of traditional minimal changes from Bytzer et al.⁹ The bar chart plots patterns of agreement/disagreement of the three endoscopists on the presence of any one of the listed minimal changes. Agreement on the presence of any one type of the minimal changes assessed was virtually no better than chance.

agreement, in many no acceptable level of agreement was reached.⁵ This is therefore an insufficiently robust criterion for grading severity. The difficulty of judging consistently between 'ulcer' and 'erosion' and apparently loose use of these descriptors when applied to reflux oesophagitis led to the adoption of 'mucosal break'.

Major structure 4: use only the extent of oesophagitis as the measure of severity

This flows partly from major structure 2 above and recognition of the illogicality of a complication of oesophagitis determining oesophagitis grading,⁷ especially when the complication remains after healing of oesophagitis, as is the case with Barrett's oesophagus.

Major structure 5: four severity grades can be distinguished and there are circumstances when it is useful to do this

The Working Group was far from unanimous in its support for 4 grades, with a significant minority favouring three on the basis that the Classification should be as simple as possible. It is probably reasonable to conclude that the choice of four grades is supported by experience with the use of the Classification, as summarized below.

Major structure 6: use the radial extent of oesophagitis as the primary criterion for severity grading, judging this by the relationships of mucosal breaks to the tops of mucosal folds

This strategy performed best in exploratory interobserver agreement studies by the Group^{4,5} and enabled development of the grades A–D criteria. Measures of longitudinal extent were found to be less precise and so were minimized to keep criteria as simple and reproducible as possible.

PUBLICATION OF TWO VERSIONS OF THE LA CLASSIFICATION: AN ERROR OF JUDGEMENT

The substantial time taken to devise and carry out the several steps (Figure 1) that eventually resulted in the definitive version of the LA Classification led the Working Group to publish what was titled as 'A Progress Report on Observer Agreement' in 1996.⁴ The abstract of this paper concluded that more work was needed and in the discussion it was stated that '... this endoscopic classification system, which still represents a working model, is not yet ready for clinical application.' However, as feared by a prescient member of the Working Group this 'working model' was adopted quite widely and in advance of the final analysis published in 1999.⁵ One change was made to the severity grading criteria as a result of further work between 1996 and 1999 to the Grade D criterion from 'circumferential mucosal break' to 'one (or more) mucosal break which involves at least 75% of the oesophageal circumference.' Thus, the Group inadvertently breached its intention to publish only one version, in order to avoid the ambiguities that are now associated with there being several versions of the Savary–Miller Classification. Happily, the difference between the evolutionary and definitive version of the LA Classification is small and applies to the least common oesophagitis grade.

PRIMARY VALIDATION OF THE DEFINITIVE LA CLASSIFICATION

This involved 46 endoscopists from around the world, who evaluated 22 edited video recordings of the oesophageal examination in patients without and with oesophageal mucosal breaks of varying severity.⁵ There was fair interobserver agreement (mean kappa value 0.4) for endoscopic judgements that defined the severity grading of reflux oesophagitis according to the LA Classification.

SECONDARY VALIDATIONS OF THE DEFINITIVE LA CLASSIFICATION

A substantial amount of information has accumulated which provides indirect validation of the LA Classification. In a large number of clinical trials of the medical treatment of reflux disease there has been a remarkably consistent gradient of oesophagitis healing rates across the LA grades, with LA grade A having the highest rates^{13,14} (Figure 3). Pretreatment LA grade has also been a consistent prognostic factor for prevention of relapse of oesophagitis by proton pump inhibitor maintenance therapy.¹⁵ Oesophageal acid exposure values correlate with LA grade.^{5,16}

EVALUATIONS OF THE LA CLASSIFICATION BY THOSE OTHER THAN IWGCO MEMBERS

A comparison of the relative performances of a modified version of the Savary–Miller, and the MUSE and LA Classifications by Rath et al¹⁷ provides valuable insights. This study used assessment of video recordings by nine endoscopists with three defined levels of experience who graded the video recordings according to the three systems.

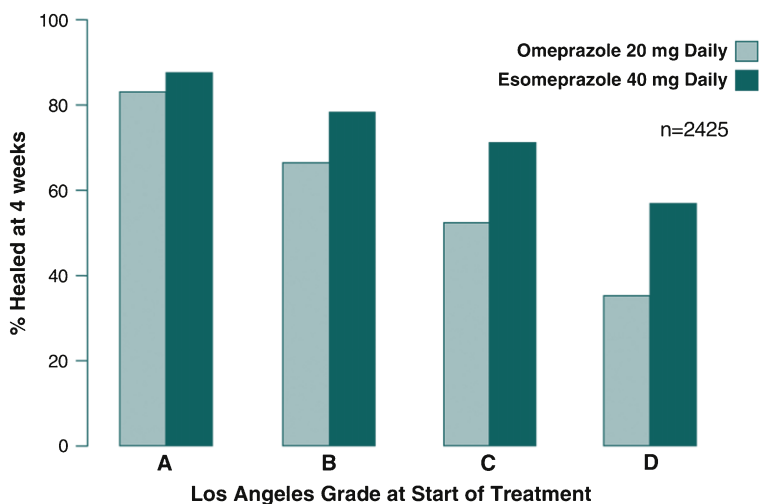


Figure 3. Healing rates of reflux oesophagitis according to LA grade after 4 weeks of treatment from a clinical trial which compared two proton pump inhibitors.¹³ The trial continued for a further 4 weeks. The 8-week cumulative healing rates showed less impact of LA grade, but still the same gradient of success rates by grade.

The LA Classification gave the best levels of agreement, with kappa values varying from 0.49 to 0.65 among the three levels of experience. Interestingly, the highest kappa values were achieved by the endoscopists in the middle range of experience. The Savary–Miller Classification performed poorly. In Japan, Kusano et al also studied interobserver agreement for LA grading by three groups of endoscopists with differing levels of experience, but only still endoscopic photographs were used.¹⁸ The kappa values for interobserver agreement were low ranging from 0.26 (least experienced) to 0.39 (most experienced). In a further study, Kusano et al found no improvement of interobserver agreement from modifications to the LA grading criteria, this study also being based on sets of still endoscopic photographs.¹⁹ These results are probably significantly determined by the use of still endoscopic images which cannot give a full emulation of the endoscopic examination required to recognize and grade oesophageal mucosal breaks.

THE LA CLASSIFICATION IN PATHOPHYSIOLOGICAL AND EPIDEMIOLOGICAL STUDIES AND IN POOLED DATA

A relatively small pH monitoring study from Japan¹⁶ has shown that patients with LA grades A and B oesophagitis have predominantly daytime reflux, whilst grades C and D have high levels of oesophageal acidification during the night as well as the day.

The LA Classification has been used recently in an important community epidemiological study in Sweden²⁰ and in the ProGERD study, a very large long-term observational study of reflux disease in Germany,²¹ yielding a more detailed and calibrated measure of the prevalence and severity pattern of reflux oesophagitis. A study of 11,945 patients with reflux oesophagitis was made possible by pooling of data from several large scale comparative clinical trials of proton pump inhibitor therapy in which reflux oesophagitis was graded with the LA Classification. This has provided authoritative data which define advancing age as a progressive risk factor for severe oesophagitis.²²

PATTERNS OF USE OF OESOPHAGITIS CLASSIFICATIONS IN RECENT ORIGINAL RESEARCH PUBLICATIONS

A literature search was carried out using Medline at PubMed. The search was limited to English language studies in humans, published in the years 2003–2006. The term ‘oesophagitis’ (including ‘esophagitis’) was used to search Medline, including MeSH terms. Reviews, editorials, letters and case reports were excluded from the search. Abstracts and titles of the papers identified were then reviewed for relevance, and articles were excluded if there was no mention of oesophagitis in their abstracts, or if only viral, fungal or eosinophilic oesophagitis, acute necrotizing oesophagitis, corrosive oesophagitis (caused by ingestion of caustic substances), endoscopy-negative or non-erosive reflux disease, and oesophagitis associated with stem cell transplantation or drug/radiation therapy were mentioned. Articles that described *in silico* modelling were excluded, as were any reviews, editorials, letters or case reports that escaped initial filtration. The oesophagitis grading system used in each study was noted. If the grading system was not identified in the abstract, the full article was reviewed. Articles were also categorized by geographical region and year of publication. The resulting data set was analysed to determine use of the LA Classification relative to

other grading systems throughout the study period, and to investigate possible temporal and/or geographical variations.

The searches yielded a final list of 467 articles. Of these, 305 (65%) reported use of a specific endoscopic grading system. The LA Classification was cited most frequently; it was used in more than half of all studies specifying a grading system (172/305; 56%; Figure 4A). It was also included in the three articles that used more than one system. The second most frequently used system was the Savary–Miller Classification, which was used in 36% (Figure 4A).

When use of the LA Classification was analysed for each year from 2003 to 2006, a steady increase was revealed in the percentage of studies using the system, from 40% overall in 2003 to 73% in 2006 (Figure 4B). In the analysis by region, 41% of studies in Europe, 46% of studies in North America and 79% of studies in the rest of the world used the LA Classification between 2003 and 2006 (Figure 4B).

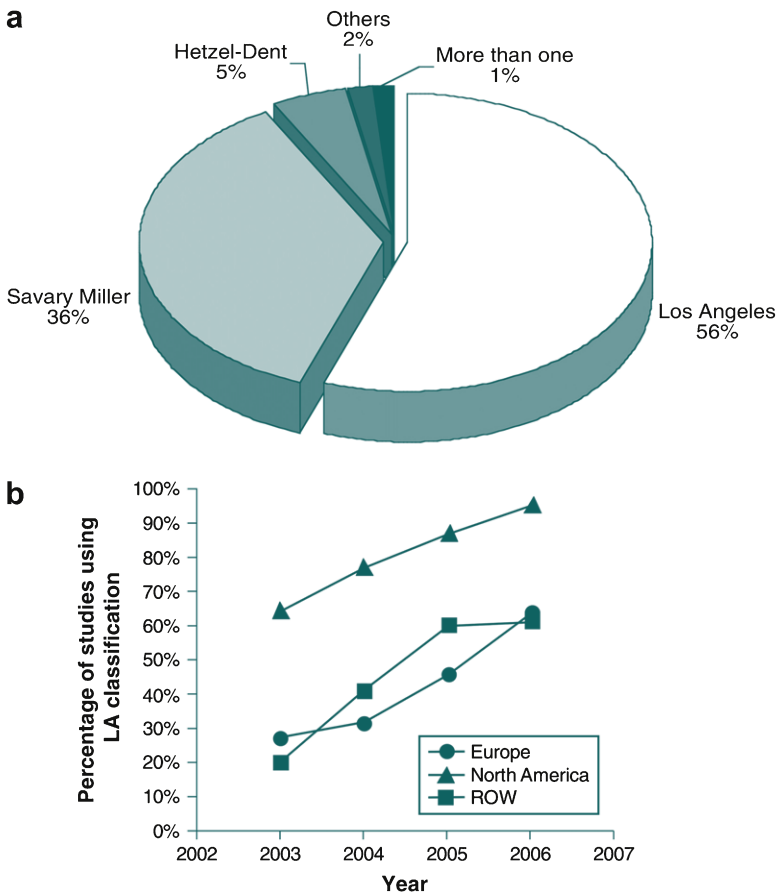


Figure 4. (A) Overview of the use of formal systems for grading of reflux oesophagitis in original peer-reviewed papers published from 2003 to 2006 inclusive ($n = 306$). (B) Patterns of use of the Los Angeles Classification year by year from 2003 to 2006 inclusive in original peer-reviewed papers ($n = 306$), according to three major areas of the world.

INTERNET-BASED SURVEY OF THE USE OF OESOPHAGITIS GRADING SYSTEMS IN ROUTINE CLINICAL PRACTICE

A survey of gastroenterologists who performed endoscopy was carried out over 3 weeks in mid-2007 via GastroSource, an educational website sponsored by Astra-Zeneca. Three-quarters of the 414 respondents were in predominantly routine clinical practice, the remainder had a major commitment to teaching and research. Of respondents, 38% were from Europe, 18% from North America and 45% from the rest of the world (5% no information). The LA Classification was used overall by 71%, but by only 52% of North American endoscopists. The Savary–Miller Classification was used by only 13% of respondents. Of endoscopists who performed more than 40 endoscopies a week, 85% used the LA Classification, compared to 57% who carried out five or less endoscopies per week.

Clinicians who had been involved in a clinical trial of reflux oesophagitis were slightly more likely to use the LA Classification (79%) than those who had had no such involvement (65%). Only 45% of respondents used a software-based endoscopy reporting system; for these respondents, 72% reported that the software supported reporting findings according to the LA Classification.

These data are remarkably similar to those obtained from the literature survey (Figure 4A and B) and indicate that there is also a high rate of uptake of the LA Classification in routine clinical endoscopic practice.

DEVELOPMENTS THAT MIGHT ENABLE USE OF MINIMAL CHANGES AS A DIAGNOSTIC CRITERION

The plausibility that non-erosive reflux-induced injury might be recognizable endoscopically is supported by the existence of histological changes in the distal oesophageal mucosa of probably a majority of non-erosive reflux disease patients.²³ The fact that non-erosive changes have not yet been proven to be of diagnostic value could be for several reasons other than that they are not an effective marker for non-erosive oesophagitis: the image resolution of endoscopes used for previous formal evaluations may have been inadequate, copies of still photographs or video recordings used for some assessments could have degraded image quality so that relevant changes were obscured, the best criteria for recognition of non-erosive injury might not have been identified and evaluated, or the endoscopists making the assessments might have had insufficient expertise to distinguish subtle mucosal abnormalities. A combination of these factors is also a real possibility.

Japanese endoscopists are generally convinced that they can recognize traditional minimal changes and because of this a modified version of the LA Classification is used by many Japanese endoscopists.¹⁰ This version keeps the criteria for the definition and grading of mucosal breaks unchanged, but has added grade 'M' for traditional minimal change, and grade 'N' to denote absence of both erosions and minimal changes. The 'M' grade, derived from the modified grading system developed by the Japanese Study Group for the Esophageal Disorders²⁴ requires presence of erythema and/or 'whitish turbidity' of the mucosa.

Japanese endoscopists have an impressive record of recognising very minor but clinically relevant mucosal abnormalities in the stomach, so it is important to revisit the conclusion made by the IWGCO in the early 1990s that minimal changes of the oesophageal mucosa are not recognized reliably, especially given the greatly

improved image resolution of modern endoscopes. A recent multicentre Japanese study²⁵ recruited 115 patients with troublesome heartburn, but no mucosal breaks. Centres and endoscopists involved had a special interest in the diagnosis of reflux disease. Grade 'M' oesophagitis was diagnosed at live endoscopy in 43% ($n = 49$). There were no significant differences in the demographics between patients graded 'N' and 'M'. Unfortunately, only 35% ($n = 17$) of grade 'N' and 21% ($n = 14$) of grade 'M' patients consented to oesophageal pH monitoring. The pH data in Table 2 suggest, but do not prove conclusively that grade 'M', as recognized by expert Japanese endoscopists, identifies patients with a greater median burden of acid reflux, but with major overlap of individual subject values. If grade 'M' truly reflects non-erosive reflux oesophagitis, patients so classified could reasonably be expected to have higher rates of response to acid suppressant therapy than those with a normal mucosa. Disappointingly, a study in Japan by Miwa et al showed no such differentiation between 115 non erosive reflux disease patients treated with rabeprazole, of whom 49 were classified as grade M.²⁶

A study by Amano et al²⁷ assessed the interobserver agreement of Japanese endoscopists for recognition of minimal changes in endoscopic still photographs. The kappa values were unacceptably low (maximum 0.26), even when endoscopists were grouped according to their training and experience. This negative finding should be regarded as inconclusive because of its use of still photographs, rather than video recordings.

Most recently, Miwa et al²⁸ evaluated interobserver agreement among 96 Japanese endoscopists for recognition of oesophageal mucosal breaks, 'M' grade minimal changes and normal oesophageal mucosa in a collection of 30 video recordings lasting from 30 to 40 s. There was remarkably poor interobserver agreement for each of these three judgements, which suggests that the study processes were not successful in emulating live oesophageal endoscopy. Several factors may have contributed, including use of video-projection and display of each video recording only once.

The reliability of the finding of grade 'M' for diagnosis of reflux disease in Japan remains unproven, despite some encouraging data. New studies are needed that overcome limitations and pitfalls discussed below.

Table 2. Selected oesophageal pH monitoring data from the study of Joh et al²⁵ in patients with troublesome heartburn and no mucosal breaks, according to whether 'traditional' minimal changes of the oesophageal mucosa were present.

	Minimal change absent ($n = 17$)	Minimal change present ($n = 14$)	<i>P</i> -value
% time oesophageal pH less than 4 (range)	1.5 (0.0–11.1)	6.4 (0.3–14.9)	0.080
Total number of acid reflux episodes	29 (0–497)	108 (9–399)	0.049

There is marked intersubject variability, but a suggestion that minimal changes detected at live endoscopy in this study are of some value in predicting higher levels of acid reflux. Values given are median and range.

BEST APPROACHES FOR STUDIES OF THE CLINICAL SIGNIFICANCE AND ENDOSCOPIC RECOGNITION OF MINIMAL CHANGES

The demands of a technically adequate study of the diagnostic value of traditional or novel minimal change criteria are considerable, since first, the levels of interobserver agreement on presence of any form of minimal change need to be determined in two study groups of adequate size: those with non-erosive reflux disease and in individuals proven not to have reflux disease. Because symptom evaluation has imperfect sensitivity for both diagnosis and exclusion of reflux disease, it is very desirable that all subjects be further defined diagnostically by 24-hour oesophageal pH monitoring. Without this, the diagnostic value of criteria will probably be underestimated because the healthy control group will contain some individuals with asymptomatic non-erosive oesophagitis and in the reflux disease group, some patients who have been misdiagnosed. Furthermore, endoscopists must be blinded to the presence and nature of symptoms experienced by study subjects, because Bytzer has shown that the interpretation of a video recording of an endoscopy was strongly influenced by knowledge of symptom patterns.²⁹ If video recordings are used rather than live endoscopy, these recordings need to be made with equipment that fully preserves the image viewed at live endoscopy. The method of display of images is important for this. Video projectors usually substantially degrade the original image quality, whilst computer screens have good resolution.

EVALUATIONS OF NOVEL MINIMAL CHANGE CRITERIA

A small study of novel minimal change criteria planned and carried out by IWGCO members in Europe and Australia has met several, but unfortunately not all of the demanding requirements given above. The study investigated the diagnostic value of several novel proposed criteria (Table 3) detected by high resolution magnification endoscopy.³⁰ Video recordings made in around 2000 were evaluated for these changes by a panel of trained endoscopists. As can be seen from Table 3, there was very poor interobserver agreement on the presence of the changes studied, so that even if these changes are specific for reflux disease, this study suggests that they would have poor diagnostic utility. The judgement that these changes were considered to be present relatively frequently in subjects without reflux disease argues further against the diagnostic utility of the changes evaluated. The outcomes of this study should however, be interpreted cautiously, as the video recordings that were assessed and their method of display, both using the best available equipment at the time, would still have resulted in some degradation of image quality compared to a live examination with the same endoscope.

The possibility that video recording image quality and the expertise of the endoscopist assessors influenced the results of the study of Edebo et al.³⁰ is suggested by another larger, but in some ways less rigorous, German study by Kiesslich et al.³¹ with high resolution magnification endoscopy which evaluated essentially the same novel criteria as Edebo et al.³⁰ At live endoscopy, one endoscopist who was blinded to the subject's symptom status found novel minimal changes in 27/39 non-erosive reflux disease patients and in 8/39 control patients, even though the categorisation of study subjects to reflux disease/not reflux disease was made solely on the basis of symptoms. Unfortunately, this study did not examine how well other endoscopists could agree on the presence of the novel minimal changes evaluated.

Table 3. Prevalence and interobserver agreement on novel proposed minimal change criteria assessed by magnification endoscopy for reflux disease patients without mucosal breaks and healthy control subjects, from the study of Edebo et al.³⁰

Criterion	Subjects' criterion present (%)		Kappa value ^b
	Controls (n = 8)	Reflux disease (n = 7)	
Triangular indentations at SCJ ^a	63	86	0.27
Apical mucosal break at indentation	38	57	0.20
Palisade vessels not seen	60	62	0.59
Pin-point/comma-shaped vessels	38	55	0.25
Branching blood vessels below SCJ	27	10	0.23
Serrated SCJ	8	26	0.11
Villiform mucosa below SCJ	98	93	0.24

Oesophageal pH monitoring was used to finally define the two populations. Interobserver agreement on the presence of the criteria was unacceptably low and no criterion appeared to identify the reflux disease patients.

^a SCJ, squamocolumnar junction.

^b Kappa value 0.2–0.4, poor–fair agreement.

It is clear that as the technology of endoscopic imaging advances, the potential for reliable detection of oesophageal minimal changes increases. This potential is expanded by novel imaging technologies discussed in the chapter by L. Gossner. These are developments for the future. At present, in routine clinical endoscopic practice with conventional endoscopic methods, it is still not appropriate to use minimal changes for the diagnosis of reflux disease because they remain inadequately validated.

SUMMARY

Since the introduction of the Los Angeles Classification of reflux oesophagitis, substantial data have accumulated which support its superior validity and clinical relevance compared to other endoscopic grading systems. The Los Angeles Classification is now by far the most commonly used system, with consequent benefits to the clarity of communication about reflux oesophagitis, both for purposes of patient management and research.

Oesophageal mucosal breaks are present in less than half of untreated patients with reflux disease, so it is an important question whether minimal changes can be used as a valid indication of non-erosive reflux oesophagitis. Such changes are not included in the Los Angeles Classification. Despite limited evidence that modern endoscopes enable some endoscopists to recognize minor mucosal changes that are correlated with reflux disease, studies that have tested the validity of these changes have been predominantly negative. Analysis of the studies published in the last 14 years suggests that the methods for storing and display of images used in interobserver agreement studies have failed to reproduce the quality of images obtained at live endoscopy. This is a significant deficiency for judgements on the presence of subtle, indefinitely demarcated mucosal changes. Technologies now exist that are capable of reliably emulating live endoscopy; it is essential that future evaluations of minimal change use these.

Practice points

- The Los Angeles Classification provides the best standardized description of reflux oesophagitis and is simple to use
- Reporting of the Los Angeles grade is useful for determining clinical management
- Currently, the finding of minimal oesophageal mucosal changes is of uncertain diagnostic value.

Research agenda

- There is a need for well-designed studies of the validity of oesophageal minimal change criteria
- Minimal change studies with stored endoscopic images must use methods that emulate the image quality and processes of live endoscopy
- Validation of minimal changes requires comparison of their prevalence in subjects with and without reflux disease, as established by symptom screening and oesophageal pH monitoring

ACKNOWLEDGEMENTS

Long-term financial support of the IWGCO by AstraZeneca and support from the World Organization of Gastroenterology has been essential for the work of this Group and is greatly appreciated. Fujinon provided magnification endoscopes for the study of Edebo et al.³⁰ Claire Mulligan and Chris Winchester of Oxford Pharmagenesis, UK are thanked for their substantial assistance with the review of patterns of use of endoscopic classifications and editorial assistance. Elizabeth Hernhage of AstraZeneca, Sweden organized and analyzed the web survey. The author is indebted to Esther Breed for her extensive support with drafting of the manuscript and in sourcing references, and to Liesl Ross who prepared the figures and tables.

REFERENCES

- *1. Dent J, Brun J, Fendrick AM et al. An evidence-based appraisal of reflux disease management – the Genval Workshop Report. *Gut* 1999; **44**(Suppl 2): I5–I35.
2. Tytgat GN, McColl K, Tack J et al. New algorithm for the treatment of gastroesophageal reflux disease. *Aliment Pharmacol Ther*. 2008; **27**: 249–256.
- *4. Armstrong D, Bennett JR, Blum AL et al. The endoscopic assessment of esophagitis: a progress report on observer agreement. *Gastroenterology* 1996; **111**: 85–92.
- *5. Lundell LR, Dent J, Bennett JR et al. Endoscopic assessment of oesophagitis: clinical and functional correlates and further validation of the Los Angeles classification. *Gut* 1999; **45**(2): 172–180.
6. Hetzel DJ, Dent J, Read WD et al. Healing and relapse of severe peptic esophagitis after treatment with omeprazole. *Gastroenterology* 1988; **95**: 903–912.
7. Savary M & Miller G. The oesophagus, *Handbook and atlas of endoscopy*. Solothorn, Switzerland: Gassman AG, 1978.
8. Armstrong D, Monnier P, Nicolet M et al. Endoscopic assessment of oesophagitis. *Gullet* 1991; **1**: 63–67.

- *9. Bytzer P, Havelund T & Hansen JM. Interobserver variation in the endoscopic diagnosis of reflux esophagitis. *Scand J Gastroenterol* 1993; **28**(2): 119–125.
10. Hongo M. Minimal changes in reflux esophagitis: red ones and white ones. *J Gastroenterol* 2006; **41**: 95–99.
11. Hatlebakk JG & Berstad A. Endoscopic grading of reflux oesophagitis: what observations correlate with gastro-oesophageal reflux? *Scand J Gastroenterol* 1997; **32**: 760–765.
12. Vieth M, Haringsma J, Delarive J et al. Red streaks in the oesophagus in patients with reflux disease: is there a histomorphological correlate? *Scand J Gastroenterol* 2001; **36**(11): 1123–1127.
- *13. Richter JE, Kahrilas PJ, Johanson J et al. Efficacy and safety of esomeprazole compared with omeprazole in GERD patients with erosive esophagitis: a randomized controlled trial. *Am J Gastroenterol* 2001; **96**(3): 656–665.
14. Labenz J, Armstrong D, Lauritsen et al. A randomized comparative study of esomeprazole 40 mg versus pantoprazole 40 mg for healing erosive oesophagitis: the EXPO study. *Aliment Pharmacol Ther.* 2005; **21**(6): 739–746.
- *15. Lauritsen K, Deviere J, Bigard MA et al. Esomeprazole 2 mg and lansoprazole 15 mg in maintaining healed reflux oesophagitis: Metropole study results. *Aliment Pharmacol Ther.* 2003; **17**(3): 333–341.
16. Adachi K, Fujishiro H, Katsube T et al. Predominant nocturnal acid reflux in patients with Los Angeles grade C and D reflux esophagitis. *J Gastroenterol Hepatol* 2001; **16**(11): 1191–1196.
- *17. Rath HC, Timmer A, Kunkel C et al. Comparison of interobserver agreement for different scoring systems for reflux esophagitis: impact of level of experience. *Gastrointest Endosc* 2004; **60**(1): 44–49.
18. Kusano M, Ino K, Yamada T et al. Interobserver and intraobserver variation in endoscopic assessment of GERD using the “Los Angeles” classification. *Gastrointest Endosc* 1999; **49**(6): 700–704.
19. Kusano M, Sugimoto S, Kawamura O et al. Numerical modification of the Los Angeles classification of gastroesophageal reflux disease fails to decrease observer variation. *Dig Endosc* 2004; **16**(1): 9–11.
20. Ronkainen J, Aro P, Storskrubb T et al. High prevalence of gastroesophageal reflux symptoms and esophagitis with or without symptoms in the general adult Swedish population: a Kalixanda study report. *Scand J Gastroenterol* 2005; **40**(3): 275–285.
21. Labenz J, Nocon M, Lind T et al. Prospective follow-up data from the ProGERD study suggest that GERD is not a categorical disease. *Am J Gastroenterol* 2006; **101**(11): 2457–2462.
- *22. Johnson DA & Fennerty MB. Heartburn severity underestimates erosive esophagitis severity in elderly patients with gastroesophageal reflux disease. *Gastroenterology* 2004; **126**(3): 660–664.
23. Dent J. Microscopic esophageal mucosal injury in nonerosive reflux disease. *Clin Gastroenterol Hepatol* 2007; **5**(1): 4–16.
24. *The Japanese Society for Esophageal Disease. Criteria for endoscopic diagnosis of esophagitis.* Tokyo: Kanehara, 1973. in Japanese.
25. Joh T, Miwa H, Higuchi K et al. Validity of endoscopic classification of nonerosive reflux disease. *J Gastroenterol* 2007; **42**: 444–449.
26. Miwa H, Sasaki M, Furuta T et al. Efficacy of rabeprazole on heartburn symptom resolution in patients with non-erosive and erosive gastro-oesophageal reflux disease: a multicenter study from Japan. *Aliment Pharmacol Ther.* 2007; **26**: 67–77.
27. Amano Y, Ishimura N, Furuta K et al. Interobserver agreement on classifying endoscopic diagnoses of nonerosive esophagitis. *Endoscopy* 2006; **38**: 1032–1035.
28. Miwa H, Nori K, Sakagami T, et al. Interobserver agreement in endoscopic evaluation of reflux esophagitis using a modified Los Angeles classification incorporating grades N and M: a validation study in a cohort of Japanese endoscopists. *Dis Esophagus* (in press).
29. Bytzer P. Information bias in endoscopic assessment. *Am J Gastroenterol* 2007; **102**: 1585–1587.
- *30. Edebo A, Tam W, Bruno M et al. Magnification endoscopy for diagnosis of nonerosive reflux disease: a proposal of diagnostic criteria and critical analysis of observer variability. *Endoscopy* 2007; **39**(3): 195–201.
- *31. Kiesslich R, Kanzler S, Vieth M et al. Minimal change esophagitis: prospective comparison of endoscopic and histological markers between patients with non-erosive reflux disease and normal controls using magnifying endoscopy. *Dig Dis.* 2004; **22**: 221–227.