GASTROINTESTINAL CANCER

Observer variation in the diagnosis of superficial oesophageal adenocarcinoma

A H Ormsby, R E Petras, W H Henricks, T W Rice, L A Rybicki, J E Richter, J R Goldblum

Gut 2002;51:671-676

See end of article for authors' affiliations

Correspondence to:
Dr J E Goldblum,
Cleveland Clinic
Foundation, Department of
Anatomic Pathology, 9500
Euclid Avenue, Cleveland
Ohio 441095, USA;
goldblj@ccf.org

Accepted for publication
5 March 2002

Background and aims: When to perform oesophagectomy for neoplastic progression in Barrett's oesophagus is controversial. Some resect for high grade dysplasia whereas others defer treatment until intramucosal adenocarcinoma is diagnosed. Interobserver agreement for a diagnosis of high grade dysplasia or intramucosal adenocarcinoma remains unknown and may have therapeutic implications.

Methods: Histological slides from 75 oesophagectomy specimens with high grade dysplasia or T₁ adenocarcinoma were blindly reviewed by two gastrointestinal pathologists and one general surgical pathologist, and classified as high grade dysplasia, intramucosal adenocarcinoma, or submucosal adenocarcinoma. A subsequent re-review of all 75 cases by the same observers following establishment of uniform histological criteria was undertaken. Interobserver agreement was determined by kappa statistics. Coefficients <0.21, 0.21-0.40, 0.41-0.60, 0.61-0.80, and >0.80 were considered poor, fair, moderate, good, and very good agreement, respectively.

Results: Interobserver agreement among all pathologists and between gastrointestinal pathologists when comparing high grade dysplasia with intramucosal adenocarcinoma was only fair ($k=0.42$; 0.56, respectively) and did not substantially improve on subsequent re-evaluation following establishment of uniform histological criteria ($K=0.50$; 0.61, respectively).

Conclusions: When evaluating resection specimens and after implementation of uniform histological criteria, even experienced gastrointestinal pathologists frequently disagree on a diagnosis of high grade dysplasia versus intramucosal adenocarcinoma. Treatment strategies based on the histological distinction of high grade dysplasia from intramucosal adenocarcinoma using limited biopsy specimens should be re-evaluated.

Barrett's oesophagus can be associated with a variety of complications including ulcer, stricture, bleeding, and most importantly, the development of adenocarcinoma. Histologically detected dysplasia arising in a background of specialised columnar epithelium (Barrett's oesophagus) frequently precedes the development of oesophageal adenocarcinoma. This observation has formed the basis of cancer surveillance protocols using endoscopy and biopsy in an attempt to identify patients at increased risk of developing oesophageal adenocarcinoma.

The exact point at which oesophagectomy should be performed remains controversial. While some authors advocate oesophagectomy for a biopsy diagnosis of high grade dysplasia¹⁻⁴ others prefer to wait until intramucosal adenocarcinoma is detected.⁵⁻⁶ We have observed in our own institution that the distinction between high grade dysplasia (HGD) and intramucosal adenocarcinoma (IMC), particularly when evaluating limited endoscopic biopsy material, can be difficult. In fact, even when the evaluation is undertaken by experienced gastrointestinal pathologists, considerable divergence of opinion is frequently observed.

Reid *et al* have shown a high degree of interobserver agreement among pathologists when attempting to distinguish low grade dysplasia from HGD and/or intramucosal carcinoma.⁷ This study however involved only highly trained gastrointestinal pathologists, did not attempt to separate HGD from intramucosal carcinoma, and did not evaluate the effects of chance agreement using kappa statistics, a statistical method which calculates the extent to which the effects of chance contribute to interobserver agreement. Without taking into account the extent of chance agreement, the percentage agreement between observers can appear deceptively high when in reality only fair interobserver agreement may exist. To our knowledge, evaluation of interobserver agreement among

pathologists in the distinction of Barrett's related HGD, IMC, and submucosal adenocarcinoma (SMC) has not been adequately addressed using kappa statistics.

The present study was designed to: (1) evaluate inter- and intraobserver agreement among gastrointestinal pathologists and general surgical pathologists using kappa statistics in the distinction of Barrett's related HGD, IMC, and SMC when evaluating optimal histological material—namely, oesophageal surgical resection specimens, and (2) to determine if intervention using uniform histological criteria can improve interobserver agreement among pathologists in the assessment of Barrett's related HGD and superficial adenocarcinoma.

MATERIALS AND METHODS

Selection of patients and histological material

Oesophageal resection specimens dating from 1987 to 1997 with Barrett's related HGD, IMC, or SMC were retrieved from the files of the Department of Anatomic Pathology. Patients with clinically advanced oesophageal adenocarcinoma treated with preoperative chemoradiotherapy followed by oesophagectomy and found to have superficial oesophageal adenocarcinoma at postoperative pathological evaluation were excluded from the study. A total of 75 cases were retrieved from the files in which haematoxylin and eosin stained slides were available for review. All histological slides (mean 22 slides per case) were reviewed by one author (AHO) and sections demonstrating the most extensive degree of HGD or the deepest extent of tumour involvement were selected.

Abbreviations: HGD, high grade dysplasia; IMC, intramucosal adenocarcinoma; SMC, submucosal adenocarcinoma; K, kappa.

Table 1 Interobserver agreement (kappa statistics) for a diagnosis of high grade dysplasia, intramucosal adenocarcinoma, and submucosal adenocarcinoma—first histological assessment (n=75)

Observers	Overall kappa	Individual kappa	95% CI	Interpretation
All observers	0.59		0.49–0.68	Moderate
HGD		0.48	0.35–0.61	Moderate
IMC		0.50	0.37–0.63	Moderate
SMC		0.76	0.63–0.89	Good
GI-1/GI-2	0.68		0.51–0.85	Good
GI-1/GSP	0.59		0.42–0.76	Moderate
GI-2/GSP	0.49		0.32–0.65	Moderate

GI-1/GI-2, gastrointestinal pathologists 1 and 2; GSP, general surgical pathologist; HGD, high grade dysplasia; IMC, intramucosal adenocarcinoma; SMC, submucosal adenocarcinoma; Kappa, kappa coefficient; 95% CI, confidence interval.

Table 2 Interobserver agreement (kappa statistics) for a diagnosis of high grade dysplasia versus intramucosal adenocarcinoma*—first histological assessment (n=49)

Reading	Kappa	p Value	95% CI	Interpretation
All observers	0.42	<0.001	0.25–0.58	Moderate
GI-1/GI-2	0.56	<0.001	0.28–0.84	Moderate
GI-1/GSP	0.45	<0.002	0.17–0.73	Moderate
GI-2/GSP	0.26	0.071	-0.02–0.54	Fair

*Excludes any specimen diagnosed with submucosal adenocarcinoma.

GI-1/GI-2, gastrointestinal pathologists 1 and 2; GSP, general surgical pathologist; Kappa, kappa coefficient; 95% CI, confidence interval.

Figure 1 A case in which agreement for a diagnosis of intramucosal adenocarcinoma was seen among all observers. The figure demonstrates neoplastic cells entirely filling the lamina propria thus replacing the pre-existing Barrett's mucosa (haematoxylin and eosin, $\times 40$).

Histological specimens

The selected histological slides were reviewed independently by three pathologists, two gastrointestinal pathologists (22 and nine years' experience), and one general surgical pathologist (five years' experience), who were blinded as to the age, sex, race, and identity of the patient. Each case was categorised based on the deepest extent of oesophageal wall involvement as either HGD, IMC, or SMC by all three pathologists. To ensure that acceptable intraobserver agreement was present, one gastrointestinal pathologist (JRG) re-reviewed all histological slides 10 days following the first assessment.

Following independent review by all three pathologists, a consensus meeting was undertaken 18 months following the first histological review whereby uniform histological criteria for a diagnosis of HGD, IMC, and SMC were established. Using guidelines modified from the Inflammatory Bowel Disease-

Figure 2 A case in which agreement for a diagnosis of high grade dysplasia was seen among all observers. The figure demonstrates neoplastic cells characterised by pronounced nuclear pleomorphism and hyperchromasia involving both crypts and the luminal surface. The architectural complexity can be readily explained by the presence of pre-existing Barrett's mucosa. Some pathologists may not classify such a case as high grade dysplasia and would consider a diagnosis of low grade dysplasia (haematoxylin and eosin, $\times 100$).

Dysplasia Morphology Study Group,⁸ HGD was defined as intraepithelial neoplasia characterised by pronounced nuclear pleomorphism, hyperchromasia, and pseudostratification involving crypts and the luminal surface, also accompanied by architectural complexity. IMC was defined as penetration of neoplastic cells through the basement membrane to lie within the lamina propria or muscularis mucosa but not beyond. In addition, the presence of architecturally complex collections of neoplastic cells in the lamina propria that could not be explained by the presence of pre-existing Barrett's mucosa were categorised as IMC (see below). SMC was defined as infiltration of neoplastic cells into the submucosa, including

Figure 3 A case in which disagreement for a diagnosis of high grade dysplasia versus intramucosal carcinoma was seen among the two gastrointestinal pathologists. The figure demonstrates neoplastic cells with a back to back architectural glandular arrangement that cannot be readily explained by the presence of pre-existing Barrett's mucosa. However, penetration of neoplastic cells through the basement membrane to lie within the lamina propria cannot be identified with certainty (haematoxylin and eosin, $\times 100$).

penetration beyond a thickened double muscularis mucosa, a feature which is frequently encountered in Barrett's oesophagus.

To verify accurate learning and application of the uniform histological criteria, all three pathologists simultaneously evaluated microscopic sections demonstrating the above histological features at a multi-headed microscope during the consensus conference. Using these criteria, all histological sections were independently re-reviewed in a blinded fashion by the same three pathologists within one week of the consensus meeting. This re-review was undertaken 18 months following the first evaluation to ensure that any memory of individual histological specimens would not bias subsequent observations.

Statistical analysis

Inter- and intraobserver agreement was determined using kappa statistics (K). Kappa statistics are widely used and accepted mathematical coefficients which provide a measure of interobserver agreement.⁹ The method of calculating kappa statistics adjusts the observed agreement for that expected by chance alone, not just the "extent of chance agreement", as it is never known what proportion of the agreement is actually due to chance—that is,

$$\text{kappa} = P_o - P_e / (1 - P_e)$$

where P =proportion agreeing, P_o = proportion observed, and P_e =proportion expected. The proportion expected is calculated under the null hypothesis of no association between the two measures. The kappa statistic was first proposed by Cohen¹⁰ and can have values of between -1 and $+1$, where negative values represent less agreement than expected by chance and positive values more agreement. A value of zero indicates no agreement better than that which would be expected by chance alone. Using previously described kappa coefficient intervals, values of <0.21 , $0.21-0.40$, $0.41-0.60$, $0.61-0.80$, and >0.80 were designated as poor, fair, moderate, good, and very good interobserver agreement, respectively.⁹ While it is helpful to have some means of classifying kappa statistics in terms of whether they are "fair", "moderate", or "good", etc, no absolute definitions are possible.⁹ Thus in practice the difference between a kappa statistic of 0.59 and 0.62 , for instance, is minimal while such kappa statistics would be differentiated as "moderate" and "good", respectively, using the above schema.

Such cutoff points are more or less arbitrary. In practice, a kappa value much below 0.5 will indicate poor agreement although the degree of acceptable agreement must depend on the circumstances. Following both the first and second histological assessments by all three pathologists, interobserver agreement using kappa statistics, with accompanying p values and 95% confidence intervals, were calculated for all possible combinations of observers and diagnostic categories—namely, HGD versus IMC, IMC versus SMC, and HGD versus SMC. That is, in the assessment of HGD versus IMC, all cases with a diagnosis of SMC were simply excluded. Likewise, in the assessment of SMC versus IMC and HGD versus SMC, all cases with a diagnosis of HGD and IMC were excluded, respectively. The calculated Kappa coefficients at both histological assessments for all combinations of observers and diagnostic categories were compared to determine the extent of improvement in interobserver agreement attributable to the establishment of uniform histological criteria. p values demonstrate the statistical significance of the comparison against a null hypothesis value of $K=0$ (that is, no agreement beyond that expected by chance).

RESULTS

First histological assessment

The cohort included 71 men and four women, aged 46–83 years (mean 71). When taking into account the effects of chance agreement, overall interobserver agreement was only moderate between all observers ($K=0.59$; table 1) and good between the two gastrointestinal pathologists ($K=0.68$; table 1). The diagnostic category which demonstrated the lowest level of interobserver agreement was HGD ($K=0.48$ for all observers; table 1; $K=0.60$ for gastrointestinal pathologists). Notably, interobserver agreement was only fair between the second gastrointestinal pathologist (GI-2) and the general surgical pathologist for a diagnosis of HGD ($K=0.36$) or IMC ($K=0.36$). The table of frequencies corresponding to the calculated kappa values in table 1 are included in the appendix.

When interobserver agreement for a diagnosis of HGD versus IMC was evaluated, interobserver agreement was only moderate at best between all observers ($K=0.42$) and between the two gastrointestinal pathologists ($K=0.56$; table 2). Agreement for a diagnosis of IMC was seen in all cases where neoplastic cells filled the lamina propria and muscularis mucosa thus replacing the pre-existing Barrett's mucosa (fig 1). In addition, agreement for a diagnosis of HGD was seen in cases where the neoplastic cells did not transgress an intact basement membrane (fig 2). Interobserver disagreement however was demonstrated when there was architectural complexity such that the glandular architecture could not be readily attributable to the presence of pre-existing Barrett's mucosa (fig 3). That is, there was disagreement in cases with architectural complexity but lacking definite transgression of the basement membrane.

As expected, interobserver agreement with respect to a diagnosis of HGD versus SMC was very good for all combinations of observers ($K=1.0$). However, separation of IMC from SMC was less than perfect with an overall kappa value of 0.71 and a kappa of 0.76 between the gastrointestinal pathologists.

Intraobserver agreement was calculated for one gastrointestinal pathologist (GI-1) 10 days following the first histological assessment and was very good ($K=0.98$).

Second histological assessment

Overall interobserver agreement between all observers was good with a kappa coefficient of 0.65 (table 3), an increase from 0.59 in the first assessment. The improvement was attributable to increased interobserver agreement in the diagnostic category of HGD ($K=0.62$ v 0.48 ; tables 1, 3), which was largely due to improved interobserver agreement between the

Table 3 Interobserver agreement (kappa statistics) for a diagnosis of high grade dysplasia, intramucosal adenocarcinoma, and submucosal adenocarcinoma—second histological assessment (n=75)

Reading	Overall kappa	Individual kappa	95% CI	Interpretation
All observers	0.65		0.56–0.74	Good
HGD		0.62	0.49–0.75	Good
IMC		0.52	0.39–0.65	Moderate
SMC		0.82	0.69–0.95	Very good
GI-1/GI-2	0.68		0.52–0.84	Good
GI-1/GSP	0.62		0.46–0.78	Good
GI-2/GSP	0.65		0.49–0.81	Good

GI-1/GI-2, gastrointestinal pathologists 1 and 2; GSP, general surgical pathologist; HGD, high grade dysplasia; IMC, intramucosal adenocarcinoma; SMC, submucosal adenocarcinoma; Kappa, kappa coefficient; 95% CI, confidence interval.

Table 4 Interobserver agreement (kappa statistics) for a diagnosis of high grade dysplasia versus intramucosal adenocarcinoma*—second histological assessment (n=46)

Reading	Kappa	p Value	95% CI	Interpretation
All observers	0.50	<0.001	0.33–0.67	Moderate
GI-1/GI-2	0.61	<0.001	0.32–0.90	Good
GI-1/GSP	0.43	0.004	0.14–0.72	Moderate
GI-2/GSP	0.46	0.002	0.17–0.75	Moderate

*Excludes any specimen diagnosed with submucosal adenocarcinoma.

GI-1/GI-2, gastrointestinal pathologists 1 and 2; GSP, general surgical pathologist; Kappa, kappa coefficient; 95% CI, confidence interval.

second gastrointestinal pathologist (GI-2) and the general surgical pathologist ($K=0.36$ in round 1 v 0.58 in round 2). Interobserver agreement between the two gastrointestinal pathologists remained unchanged ($K=0.68$; tables 1, 3). The table of frequencies corresponding to the calculated kappa values in table 3 are included in the appendix.

When interobserver agreement for a diagnosis of HGD versus IMC was evaluated, kappa values were only modestly improved overall and between the gastrointestinal pathologists (tables 2, 4). As seen in the first histological assessment, agreement for a diagnosis of IMC was seen in cases where neoplastic cells filled the lamina propria and muscularis mucosa thus replacing the pre-existing Barrett's mucosa. Interobserver disagreement however was again demonstrated when there was less extensive infiltration by neoplastic cells into the lamina propria or muscularis mucosa and where the architectural complexity was such that the histological

appearance was not readily attributable to the presence of pre-existing Barrett's mucosa, as previously described (fig 3).

Interobserver agreement in distinguishing HGD from SMC was again very good for all observer combinations ($K=1.0$). Overall interobserver agreement between IMC and SMC was essentially unchanged ($K=0.74$).

DISCUSSION

Barrett's oesophagus is a premalignant condition and as such most patients are followed in a surveillance programme by periodic endoscopy with biopsy. The goal of endoscopic surveillance is to identify a histological lesion which is a marker of increased risk of having, or subsequently developing, invasive adenocarcinoma, prompting therapy with curative intent. However, there is some disagreement about what the end point of endoscopic surveillance should be.

Some authors advocate oesophagectomy for a biopsy diagnosis of HGD.^{1–4} There are several arguments to support this approach. Superficially invasive adenocarcinoma has been found in up to 73% of patients undergoing oesophagectomy for a biopsy diagnosis of HGD.^{1–4 11–15} At our institution, even in those patients with Barrett's oesophagus followed regularly with four quadrant biopsy specimens taken at 2 cm intervals using jumbo biopsy forceps, 33% of patients who underwent oesophagectomy for a preoperative diagnosis of HGD without gross or microscopic evidence of carcinoma harboured an unsuspected adenocarcinoma.¹⁶ In contrast, Reid and colleagues⁵ and Levine and colleagues⁶ have reported the use of a vigorous biopsy protocol (four quadrant jumbo biopsies for every 1 cm of Barrett's mucosa as well as biopsies of any suspicious lesion) that reliably differentiates between HGD and invasive adenocarcinoma in patients with Barrett's oesophagus. Using this protocol, all 11 patients with a preoperative diagnosis of HGD that came to oesophagectomy were found to have HGD in their resected specimen.^{5 6} However, the practicality and cost of this intense endoscopic surveillance technique has been questioned.¹⁶ Furthermore, in the aforementioned studies,^{5 6} five of 16 (31%) patients with a

Figure 4 Graphical representation of kappa coefficients for a diagnosis of high grade dysplasia versus intramucosal adenocarcinoma for all combinations of observers before and after establishment of uniform histological criteria. The broken line represents a kappa coefficient of 0.80. GI-1/GI-2, gastrointestinal pathologists 1 and 2; GSP, general surgical pathologist; K, kappa coefficient; 95% CI, 95% confidence interval.

preoperative diagnosis of IMC that underwent oesophagectomy were found to have invasion of the submucosa in the resected specimen.^{5,6} Underdiagnosis of SMC is of clinical importance given the five year survival rates, which are 73–79% for IMC versus 16–44% for SMC.^{1,17–19} In addition, given the presence of lymphatic channels in the oesophageal mucosa, there is a 5–8% risk of lymph node metastasis even for those tumours limited to the intramucosal compartment.^{1,19–22} Certainly, the above arguments must be balanced by the morbidity and mortality associated with oesophagectomy. In studies with large numbers of patients with oesophageal cancer, operative mortality rates range from 0% to 6% in patients with early stage oesophageal carcinoma.^{1,2,4,11,22}

Given the above discussion, separation of HGD from IMC on preoperative biopsy specimens is of more than academic interest and thus we have sought to determine how well pathologists agree on distinguishing between these lesions. In a sense, we optimised the ability of pathologists to agree on these diagnoses by using oesophageal resection specimens and by establishing uniform histological criteria by a consensus conference. Even under these circumstances there was only moderate interobserver agreement in the separation of HGD from IMC and only modest improvement when evaluated by experienced gastrointestinal pathologists.

Although IMC is defined anatomically, histological recognition of penetration by neoplastic cells through the basement membrane is sometimes difficult. In cases in which the lamina propria and muscularis mucosa were essentially filled by neoplastic cells, there was excellent interobserver agreement in the recognition of IMC. All cases in which there was observer disagreement however were characterised by a complex glandular architecture which was not readily explained by the presence of pre-existing Barrett's mucosa. The absence of a stromal desmoplastic reaction in IMC, unlike that seen when neoplastic cells infiltrate the submucosa, makes this distinction difficult. In addition, accurate recognition of the basement membrane and the muscularis mucosa can be difficult to determine in the background of a complex glandular architecture which can distort the normal appearance of these landmark structures.

Given the inability of pathologists, including experienced gastrointestinal pathologists, to agree in these equivocal cases, even after evaluating abundant histological material (that is, oesophageal resection specimens) and following establishment of uniform histological criteria, several issues arise. Firstly, histological evaluation of lesions falling within the spectrum between HGD and IMC can be so complex that even "expert" opinion may be called into question. Secondly, given

the serious consequences to the patient of either undergoing oesophagectomy or failing to treat an undetected potentially curable invasive carcinoma, clinical treatment decisions which are reliant on histological distinction of HGD from IMC would require at the very least a very good level of interobserver agreement to be of any use in clinical practice. The results of the present study clearly show that the level of improvement both overall and for all combinations of observers, including the two gastrointestinal pathologists, in the distinction of HGD from IMC fell short of the degree of interobserver agreement needed in order to confidently make clinical treatment decisions. These results are summarised graphically in fig 4 where only the upper limit of the 95% confidence interval of the two gastrointestinal pathologists exceeded the very good interobserver agreement ($K > 0.80$). This problem is likely to be compounded in routine clinical practice where this distinction is made preoperatively using only limited biopsy specimens. Thus the reliability of histological assessment in these equivocal cases and the validity of histomorphology as the "standard" upon which treatment decisions are made requires re-evaluation. In view of the fact that the distinction between HGD and intramucosal carcinoma cannot reliably be made on histological evaluation, one could argue that both diagnostic entities should be regarded as one category that for practical purposes warrants consideration of the same treatment—that is, definitive surgical resection in surgical candidates.

These results may prompt consideration of biopsy reviews by specialist gastrointestinal pathologists at centres with demonstrated expertise and a multidisciplinary team approach to the evaluation of Barrett's related HGD and superficial adenocarcinoma. In fact, a stronger plea might be made for a qualified second opinion to cover all oesophageal dysplastic/carcinoma problems, as microscopy still remains the most widely available and well tested means of diagnosis. Adoption of a consensus terminology such as the Vienna classification of gastrointestinal epithelial neoplasia may also be helpful.²³ In a recent study, using this classification scheme, pathologists were able to increase interobserver agreement in the evaluation of neoplastic oesophageal lesions from very poor ($K=0.01$) to fair ($K=0.31$).

The results of the present study also highlight the need for further development and evaluation of sensitive and specific adjunctive biomarkers in an attempt to help resolve the shortcomings of histomorphology in the distinction of HGD from IMC. The most well studied biomarkers to date include ploidy assessed by flow cytometry,²⁴ mutated p53 tumour suppressor gene protein expression,^{25–28} and proliferative activity labelling

Table A1 Frequency table corresponding to the calculated kappa values in the first and second histological assessments by observer and diagnostic category combinations ($n=75$)

	First assessment			Second assessment		
	HGD	IMC	SMC	HGD	IMC	SMC
GI-1/GI-2						
HGD	8	0	0	17	6	0
IMC	8	34	3	3	20	1
SMC	0	3	19	0	6	22
GI-1/GSP						
HGD	7	1	0	14	9	0
IMC	9	32	4	4	19	1
SMC	0	4	18	0	5	23
GI-2/GSP						
HGD	8	8	0	13	7	0
IMC	8	25	4	5	24	3
SMC	0	4	18	0	2	21

GI-1/GI-2, gastrointestinal pathologists 1 and 2; GSP, general surgical pathologist; HGD, high grade dysplasia; IMC, intramucosal adenocarcinoma; SMC, submucosal adenocarcinoma.

index (Ki-67/MIB1)^{29,30} via immunohistochemistry. These biomarkers have been identified in up to 89% of patients with both low and high grade epithelial dysplasia and in up to 100% of patients with invasive carcinoma while being only rarely expressed in normal control oesophageal biopsies. In addition, these biomarkers have been correlated with progression from benign Barrett's mucosa to epithelial dysplasia and invasive carcinoma. More recently, other biomarkers including cyclin D1,³¹⁻³³ a cell cycle regulation protein, and telomerase,^{34,35} a protein that confers cell immortality, have been evaluated using both in situ hybridisation and immunohistochemical techniques. Both of these markers have also been correlated with progression along the Barrett's metaplasia-dysplasia-carcinoma sequence. However, there have been no significant differences in expression of any of these markers in HGD and IMC, and thus at this time there appears to be no clinical utility of these adjunctive techniques.

In conclusion, even with the benefit of abundant histological material (oesophagectomy specimens) assessed by experienced gastrointestinal pathologists using uniform histological criteria, the present study demonstrated that only moderate overall interobserver agreement can be achieved in distinguishing HGD from IMC. It is likely that this discrepancy is compounded when evaluating endoscopic biopsy specimens, which in clinical practice are the types of specimens that pathologists must evaluate prior to definitive therapy. Re-evaluation of treatment strategies based on histological distinction of HGD from IMC in endoscopic biopsy specimens is warranted.

APPENDIX

The table of frequencies corresponding to the calculated kappa values in the first and second histological assessments by observer and diagnostic category combinations are included in table A1.

.....

Authors' affiliations

A H Ormsby, Department of Anatomic Pathology, Henry Ford Health System, Detroit, Michigan, USA

R E Petras, W H Henricks, J R Goldblum, Department of Anatomic Pathology, Cleveland Clinic Foundation, Cleveland, Ohio, USA

T W Rice, Department of Thoracic and Cardiovascular Surgery, Cleveland Clinic Foundation, Cleveland, Ohio, USA

L A Rybicki, Department of Biostatistics and Epidemiology, Cleveland Clinic Foundation, Cleveland, Ohio, USA

J E Richter, Center for Swallowing and Esophageal Diseases, and Department of Gastroenterology, Cleveland Clinic Foundation, Cleveland, Ohio, USA

REFERENCES

- Rice TW, Falk GW, Achkar E, et al. Surgical management of high-grade dysplasia in Barrett's esophagus. *Am J Gastroenterol* 1993;**88**:1832-6.
- Pera M, Trastek VF, Carpenter HA, et al. Barrett's esophagus with high-grade dysplasia: an indication for esophagectomy? *Ann Thorac Surg* 1992;**54**:199-203.
- Altorki NK, Sunagawa M, Little AG, et al. High-grade dysplasia in the columnar-lined esophagus. *Am J Surg* 1991;**161**:97-100.
- Edwards MJ, Gable Dr, Lentsch AB, et al. The rationale for esophagectomy as the optimal therapy for Barrett's esophagus with high-grade dysplasia. *Ann Surg* 1996;**223**:585-91.
- Levine DS, Haggitt RC, Blount PL, et al. An endoscopic biopsy protocol can differentiate high-grade dysplasia from early adenocarcinoma in Barrett's esophagus. *Gastroenterology* 1993;**105**:40-50.
- Reid BJ, Weinstein WM, Lewin KJ, et al. Endoscopic biopsy can detect high-grade dysplasia or early adenocarcinoma in Barrett's esophagus without grossly recognizable neoplastic lesions. *Gastroenterology* 1988;**94**:81-90.

- Reid BR, Haggitt RC, Rubin CE, et al. Observer variation in the diagnosis of dysplasia in Barrett's esophagus. *Hum Pathol* 1988;**19**:166-78.
- Riddell RH, Goldman H, Ransohoff DF, et al. Dysplasia in inflammatory bowel disease: standardized classification with provisional clinical applications. *Hum Pathol* 1983;**14**:931-66.
- Altman DG. *Practical statistics for medical research*. New York: Chapman and Hall, 1991.
- Cohen JA. A coefficient of agreement for nominal scales. *Educ Psychol Assess* 1960;**20**:37-46.
- Heitmiller RF, Redmond M, Hamilton SR. Barrett's esophagus with high-grade dysplasia: an indication for prophylactic esophagectomy. *Ann Surg* 1996;**224**:66-71.
- Cameron AM, Carpenter HA. Barrett's esophagus, high-grade dysplasia and early adenocarcinoma: a pathological study. *Am J Gastroenterol* 1997;**92**:586-601.
- Skinner EB, Walther BC, Riddell RH, et al. Barrett's esophagus. Comparison of benign and malignant cases. *Ann Surg* 1983;**198**:554-66.
- Lee RG. Dysplasia in Barrett's esophagus. A clinicopathologic study of six patients. *Am J Surg Pathol* 1985;**12**:845-52.
- Hamilton SR, Smith RR. The relationship between columnar epithelial dysplasia and invasive adenocarcinoma arising in Barrett's esophagus. *Am J Clin Pathol* 1997;**87**:301-12.
- DeMeester TR, Attwood SE, Smyrk TC, et al. Surgical therapy in Barrett's esophagus. *Ann Surg* 1990;**212**:528-42.
- Falk GW, Rice TW, Goldblum JR, et al. Jumbo biopsy forceps protocol still misses unsuspected cancer in Barrett's esophagus with high-grade dysplasia. *Gastrointest Endosc* 1999;**49**:120-6.
- Menke-Pluymers MBE, Schoute NW, Mulder AH, et al. Outcome of surgical treatment of adenocarcinoma in Barrett's esophagus. *Gut* 1992;**33**:1454-8.
- Nabeya K, Hanaoka T, Li S, et al. What is the ideal treatment for early esophageal cancer? *Endoscopy* 1993;**25**:670-1.
- Sugimachi K, Ikebe M, Kitamura K, et al. Long-term results of esophagectomy for early esophageal carcinoma. *Hepatogastroenterology* 1993;**40**:203-6.
- Goseki N, Koike M, Yoshida M. Histopathologic characteristics of early stage esophageal carcinoma. A comparative study with gastric carcinoma. *Cancer* 1992;**69**:1088-93.
- Bogomoletz WV, Molas G, Gatey B, et al. Superficial squamous cell carcinoma of the esophagus: a report of 76 cases and review of the literature. *Am J Surg Pathol* 1989;**13**:535-46.
- Schlemper RJ, Riddell RH, Kato Y, et al. The Vienna classification of gastrointestinal epithelial neoplasia. *Gut* 2000;**47**:251-5.
- Rabinovitch PS, Longton G, Blount PL, et al. Predictors of progression in Barrett's esophagus III: Baseline flow cytometric variables. *Am J Gastroenterol* 2001;**96**:3071-83.
- Klump B, Hsieh CJ, Holzmacher K, et al. Diagnostic significance of nuclear p53 expression in the surveillance of Barrett's esophagus—a longitudinal study. *Z Gastroenterol* 1999;**37**:1005-11.
- Moskaluk CA, Heitmiller R, Zahurak M, et al. p53 and p21 (WAF1/CIP1/SDI1) gene products in Barrett's esophagus and adenocarcinoma of the esophagus and esophagogastric junction. *Hum Pathol* 1996;**27**:1211-20.
- Younes M, Ertan A, Lechago LV, et al. p53 Protein accumulation is a specific marker of malignant potential in Barrett's metaplasia. *Dig Dis Sci* 1997;**42**:697-701.
- Krishnadath KK, Tilanus HW, van Blankenstein M, et al. Accumulation of p53 protein in normal dysplastic, neoplastic Barrett's oesophagus. *J Pathol* 1995;**175**:175-80.
- Peters FT, Ganesh S, Kuipers EJ, et al. Epithelial cell proliferation activity of Barrett's esophagus: methodology and correlation with traditional cancer risk markers. *Dig Dis Sci* 1998;**43**:1501-6.
- Polkowski W, van Lanschot JJ, Ten Kate FJ, et al. The value of p53 and Ki67 as markers for tumour progression in the Barrett's dysplasia-carcinoma sequence. *Surg Oncol* 1995;**4**:163-71.
- Kitahara K, Yasui W, Yokozaki H, et al. Expression of cyclin D1, CDK4 and p27KIP1 is associated with the p16MTS1 gene status in human esophageal carcinoma cell lines. *J Exp Ther Oncol* 1996;**1**:7-12.
- Arber N, Lightdate C, Rotterdam H, et al. Increased expression of the cyclin D1 gene in Barrett's esophagus. *Can Epidemiol Biomarkers Prev* 1996;**5**:457-9.
- Coppola D, Falcione R, Livingston S, et al. Cyclin D1 expression correlates with degrees of dysplasia in Barrett's esophagus. *Lab Invest* 1997;**76**:298-302.
- Morales CP, Lee EL, Shay JW. In situ hybridization for detection of telomerase RNA in the progression from Barrett's esophagus to esophageal adenocarcinoma. *Cancer* 1998;**83**:652-9.
- Mullick T, Tasch JE, Ormsby AH, et al. Telomerase upregulation in Barrett's esophagus precedes the development of esophageal adenocarcinoma. *Am J Gastroenterol* 2001;**75**:525.