Barrett Esophagus:

Improving Surveillance Strategies

Marjon Kerkhof

Financial support for this thesis was kindly given by:
Altana Pharma B.V.
AstraZeneca B.V.
Janssen-Cilag B.V.
Pentax Nederland B.V.
Schering-Plough B.V.
Department of Gastroenterology and Hepatology, Erasmus MC Rotterdam

Cover photo: Villin staining of the brush border in Barrett epithelium (Marjon Kerkhof)

Layout by: Sander Herfst
Printed by: PrintPartners Ipskamp B.V.
ISBN: 978-90-9021953-0

Barrett Esophagus: Improving Surveillance Strategies

Barrett oesofagus: verbetering van surveillance strategieën

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de rector magnificus
Prof.dr. S.W.J. Lamberts
en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op
vrijdag 29 juni 2007 om 11.00 uur

door

Marjon Kerkhof
geboren te Deventer

PROMOTIECOMMISSIE

Promotor:	Prof.dr. E.J. Kuipers
Overige leden:	Prof.dr. E.W. Steyerberg Prof.dr. J.J.B. van Lanschot Prof.dr. H.W. Tilanus
Copromotoren:	Dr. P.D. Siersema Dr. J.G. Kusters

CONTENTS

1

General introduction and outline of this thesis

BARRETT ESOPHAGUS

Barrett esophagus (BE) is the replacement of the pale-coloured squamous epithelium of the distal esophagus by red-coloured, velvet-like columnar epithelium of any length that can be recognized at endoscopy. According to the guidelines of the American College of Gastroenterology, this endoscopic diagnosis has to be confirmed by the presence of intestinal metaplasia (columnar epithelium with goblet cells) in biopsies of the distal esophagus.[1, 2] Cardiac-type mucosa is also frequently observed in the columnar-lined esophagus,[3] this is in contrast to intestinal metaplasia not regarded as a premalignant condition.[4, 5] BE is caused by chronic gastroesophageal reflux,[6] and predisposes to the development of esophageal adenocarcinoma (EAC). The incidence of both BE and EAC have increased rapidly over the past two decades in most Western countries and now comprises at least 60% of all esophageal cancer cases.[7-10] The development of EAC in BE is a gradual process in which important biological processes become disrupted. This process is classified into different stages, i.e., low-grade dysplasia (LGD), high-grade dysplasia (HGD), and finally EAC.[5, 11-13]

Currently, histopathologic assessment is the standard to judge to what stage the neoplastic process has progressed in an individual patient, and based on this, to determine the interval of endoscopic surveillance in patients with BE. The aim of surveillance is to detect progression of dysplasia in an early, curable stage.[1] Since cardiac-type mucosa is not regarded to be a premalignant condition, patients with only cardiac-type mucosa in their biopsies are currently excluded from a surveillance program. However, as intestinal metaplasia and cardiac-type mucosa are endoscopically indiscernible, sampling error can occur, and exclusion from endoscopic follow-up might be incorrect.[14]

Patients without dysplasia (no dysplasia; ND) in biopsies of BE are regarded to be at a lower risk of neoplastic progression than patients with LGD.[15] Therefore, according to current guidelines, the presence or absence of dysplasia in intestinal metaplasia determines the frequency of surveillance upper endoscopy (e.g. yearly for LGD, and every 3 years for ND).[1, 15, 16] However, a considerable interobserver variability in the interpretation of dysplasia has been demonstrated,[17] which may lead to superfluous follow-up endoscopies (in case of overdiagnosis), or insufficient control (in case of underdiagnosis).

It seems relevant to perform a straightforward risk stratification to define which patients with a columnar-lined esophagus with or without intestinal metaplasia should undergo endoscopic follow-up, and at which frequency. As the risk of developing EAC in BE is low,[15, 18, 19] the majority of these patients will not benefit from a burdensome endoscopic surveillance program. Further stratifying the risk of neoplastic progression in BE might permit more effective surveillance of high-risk patients and in addition improve the cost-effectiveness of surveillance.

Aim of this thesis

The aim of this thesis is to assess the currently used criteria for performing endoscopic surveillance in patients with BE, and to evaluate which clinical characteristics and biomarkers can contribute to risk stratification in patients with a columnar-lined esophagus, in order to refine surveillance strategies in these patients.

Outline of this thesis

In chapter 2, the presence of markers for intestinal metaplasia, i.e. CDX2, MUC2 and villin, in a columnar-lined segment of the esophagus in the absence of a histological diagnosis of intestinal metaplasia are evaluated. Furthermore, the predictive value of these markers for the presence of undetected intestinal metaplasia in the columnar-lined esophagus is investigated. In chapter 3, a model based on clinical characteristics is developed to estimate the probabilities of the presence of intestinal metaplasia, ND and LGD in biopsies of the columnar-lined esophagus. In chapter 4, the interobserver variability in establishing the grade of dysplasia in BE is assessed, and compared between non-expert general pathologists and expert gastrointestinal pathologists on the one hand, and between expert gastrointestinal pathologists on the other hand. In chapter 5, the existing literature is reviewed regarding the so far evaluated candidate biomarkers for improving risk stratification of patients with BE. In chapter 6, the most promising biomarkers, i.e. Ki67, p53, and DNA ploidy, are examined in more detail regarding their usefulness in identifying the subgroup of BE patients at highest risk for subsequent progression to EAC. In chapter 7, the perceived burden of upper gastrointestinal endoscopy in BE patients is explored. In the final chapter, chapter 8, the results described in this thesis are summarized and discussed.

REFERENCES

1. Sampliner RE. Updated guidelines for the diagnosis, surveillance, and therapy of Barrett's esophagus. *Am J Gastroenterol* 2002; 97: 1888-95.
2. Guillem PG. How to make a Barrett esophagus: pathophysiology of columnar metaplasia of the esophagus. *Dig Dis Sci* 2005; 50: 415-24.
3. Paull A, Trier JS, Dalton MD*, et al.* The histologic spectrum of Barrett's esophagus. *N Engl J Med* 1976; 295: 476-80.
4. Falk GW. Barrett's esophagus. *Gastroenterology* 2002; 122: 1569-91.
5. Hamilton SR, Smith RR. The relationship between columnar epithelial dysplasia and invasive adenocarcinoma arising in Barrett's esophagus. *Am J Clin Pathol* 1987; 87: 301-12.
6. Avidan B, Sonnenberg A, Schnell TG*, et al.* Hiatal hernia size, Barrett's length, and severity of acid reflux are all risk factors for esophageal adenocarcinoma. *Am J Gastroenterol* 2002; 97: 1930-6.
7. Bollschweiler E, Wolfgarten E, Gutschow C*, et al.* Demographic variations in the rising incidence of esophageal adenocarcinoma in white males. *Cancer* 2001; 92: 549-55.
8. Bytzer P, Christensen PB, Damkier P*, et al.* Adenocarcinoma of the esophagus and Barrett's esophagus: a population-based study. *Am J Gastroenterol* 1999; 94: 86-91.
9. Blot WJ, Devesa SS, Kneller RW*, et al.* Rising incidence of adenocarcinoma of the esophagus and gastric cardia. *Jama* 1991; 265: 1287-9.
10. van Soest EM, Dieleman JP, Siersema PD*, et al.* Increasing incidence of Barrett's oesophagus in the general population. *Gut* 2005; 54: 1062-6.
11. Haggitt RC. Barrett's esophagus, dysplasia, and adenocarcinoma. *Hum Pathol* 1994; 25: 982-93.
12. Hameeteman W, Tytgat GN, Houthoff HJ*, et al.* Barrett's esophagus: development of dysplasia and adenocarcinoma. *Gastroenterology* 1989; 96: 1249-56.
13. Buttar NS, Wang KK. Mechanisms of disease: Carcinogenesis in Barrett's esophagus. *Nat Clin Pract Gastroenterol Hepatol* 2004; 1: 106-12.
14. Oberg S, Johansson J, Wenner J*, et al.* Endoscopic surveillance of columnar-lined esophagus: frequency of intestinal metaplasia detection and impact of antireflux surgery. *Ann Surg* 2001; 234: 619-26.
15. Hage M, Siersema PD, van Dekken H*, et al.* Oesophageal cancer incidence and mortality in patients with long-segment Barrett's oesophagus after a mean follow-up of 12.7 years. *Scand J Gastroenterol* 2004; 39: 1175-9.
16. Mashimo H, Wagh MS, Goyal RK. Surveillance and screening for Barrett esophagus and adenocarcinoma. *J Clin Gastroenterol* 2005; 39: S33-41.
17. Montgomery E, Bronner MP, Goldblum JR*, et al.* Reproducibility of the diagnosis of dysplasia in Barrett esophagus: a reaffirmation. *Hum Pathol* 2001; 32: 368-78.
18. Falk GW. Barrett's esophagus-is it bad for your health? *Am J Gastroenterol* 2005; 100: 2622-3.
19. Conio M, Blanchi S, Lapertosa G*, et al.* Long-term endoscopic surveillance of patients with Barrett's esophagus. Incidence of dysplasia and adenocarcinoma: a prospective study. *Am J Gastroenterol* 2003; 98: 1931-9.

2

Does CDX2 expression predict Barrett's metaplasia in esophageal columnar epithelium without goblet cells?

Aliment Pharmacol Ther 2006 Dec; 24(11-12):1613-21

Marjon Kerkhof [1]
Dorine A Bax [1]
Leon M G Moons [1]
Adriana J van Vuuren [1]
Herman van Dekken [2]
Ewout W Steyerberg [3]
Ernst J Kuipers [1]
Johannes G Kusters [1]
Peter D Siersema [1]
for the CYBAR study group

Departments of [1]Gastroenterology and Hepatology, [2]Pathology, and [3]Public Health, Erasmus MC - University Medical Center Rotterdam, Rotterdam, the Netherlands

ABSTRACT

Background: Intestinal metaplasia (IM; Barrett's esophagus), but not cardiac-type mucosa (CM) in columnar-lined esophagus (CLE), is regarded as premalignant. Since IM and CM are endoscopically indiscernible, it is difficult to take targeted samples from CLE with consequently a risk of having undetected IM.

Aim: To investigate whether the intestinal markers CDX2, MUC2 and villin can predict the presence of undetected IM in CLE.

Methods: Presence of IM or CM was identified in 122 biopsy sets of CLE from 61 patients, collected at two subsequent follow-up upper endoscopies. CDX2, MUC2 and villin expression were determined by immunohistochemistry.

Results: All IM samples (55) were positive for CDX2 and MUC2 and 32/55 for villin. CDX2 expression was detected in 23/67 (34%) samples with only CM. Detection of CDX2 in CM increased the likelihood of finding IM in another biopsy set of CLE (OR 3.5, 95% CI=1.2-10, p=0.02). MUC2 was positive in13/23 (57%) of CDX2 positive CM samples, whereas villin was detected in 7/23 (30%).

Conclusions: CDX2 expression in CM likely predicts the presence of undetected IM in CLE, and thus may be a putative marker for the presence of IM in absence of goblet cells.

INTRODUCTION

Barrett's esophagus (BE) is a premalignant condition caused by chronic gastro-esophageal reflux,[1] which can progress from low-grade dysplasia to high-grade dysplasia, and subsequently to esophageal adenocarcinoma.[2-5]

BE is characterized by the replacement of the squamous epithelium of the esophagus by columnar epithelium with goblet cells (specialized intestinal metaplasia (IM)).[6] IM is associated with the expression of intestinal markers such as MUC2,[7] and villin.[8] Cardiac-type mucosa (CM) is also frequently observed in the columnar-lined esophagus (CLE),[9] with the absence of goblet cells as the only histological difference compared to IM.[10] CM, in contrast to IM, is not regarded as a premalignant condition.[2, 11] Therefore, only patients with IM are currently advised to undergo periodic endoscopic surveillance to detect progression to dysplasia in an early, potentially curable stage.[12] Others have reported that patients with biopsies from CLE without IM were at an increased risk of having undetected IM. This was explained by either sampling error or developing IM over time. According to current guidelines, these patients would have been falsely excluded from a surveillance program.[13]

The homeobox protein CDX2 is a transcription factor involved in the early intestinal differentiation of the epithelium of the intestines,[14-16] and its expression is also linked with BE,[17-19] suggesting that CDX2 is an early marker for the development of intestinal metaplasia in the esophagus as well. CDX2 regulates transcription of several intestinal genes, encoding proteins such as MUC2, alkaline phosphatase, and sucrase-isomaltase.[20, 21] It has been reported that intestinal phenotypic modifications may also be detected in the absence of goblet cells by CDX2 expression in CLE.[22, 23] This epithelium has been regarded as being early-stage BE, but these studies were cross-sectional and therefore provided not enough evidence for this hypothesis.

The aim of this longitudinal study was to investigate whether intestinal markers for IM, i.e.: CDX2 (early intestinal marker), MUC2 (goblet cell marker) and villin (late intestinal marker), were present in the columnar-lined segment of the esophagus in the absence of a histological diagnosis of IM (defined by the presence of goblet cells). Furthermore, we investigated whether these markers were predictive for the presence of IM in CLE, not detected due to sampling error or to IM developing over time.

MATERIALS AND METHODS

Patients and materials

In this multicenter study, 108 patients were evaluated for this retrospective study for the presence of an endoscopic CLE of at least 2 cm, and at least two follow-up endoscopies with biopsies being performed. Based on these inclusion criteria, 47 patients were excluded, and consequently 61 patients could be included in this study. Biopsies were taken at different levels from the CLE and embedded together in one paraffin block. In this study, sections of these paraffin embedded biopsy sets were used for evaluation. These slides were reviewed for the presence of IM by an expert gastrointestinal pathologist (HvD). Based on the presence of IM, patients were divided into three

Table 1. Classification of patients in groups, based on histology results from two subsequent endoscopies

	IM-group	Discordant-group		CM-group
1st endoscopy	IM	IM	CM	CM
2nd endoscopy	IM	CM	IM	CM
no. of patients	15	16	9	21

groups (Table 1): patients with IM in both biopsy sets (IM-group), patients with IM in one biopsy set and with only CM in the other biopsy set (discordant-group), and patients with only CM in both biopsy sets (CM-group). Patients with CM in the first endoscopy and IM in the second endoscopy, and visa versa, were taken together as the discordant-group.

Histology and immunohistochemistry

Six consecutive sections of 4 µm each from every biopsy set were mounted on adhesive slides, dried overnight at 37°C, and deparaffinized with xylene. The first of these serially sectioned slides was stained with haematoxylin and eosin (H&E) to determine the type of columnar epithelium (CM or IM). Alcian Blue and periodic acid-Schiff (PAS) stainings in consecutive slides were performed to facilitate the detection of mucin producing goblet cells. The next three slides were used for immunohistochemistry.

For immunohistochemistry, antigen retrieval was performed by boiling the deparaffinized samples in 10 mM monocitric acid buffer (pH 6.0) for 15 min, and slowly cooling down to room temperature (RT). Prior to immune staining, endogenous peroxidase activity was blocked by incubating the slides in a 0.5% solution of H_2O_2 in phosphate-buffered citric acid for 15 minutes at RT. Samples were washed for 5 minutes with TRIS-buffered saline (TBS) (pH 7.4). This was repeated 2 times. The samples were incubated in TBS buffer containing 10% rabbit non-immune serum (DAKO, Glostrup, Denmark) and 10% normal human plasma (DAKO) for 20 minutes. Sections were incubated for 16 hours at 4°C with respectively primary antibody anti-CDX2 (clone 392M, Biogenex, San Ramon

Table 2. Patients characteristics

	IM-group	Discordant-group	CM-group	p-value
No. of patients	15	25	21	
Mean age at 1st endoscopy in years (range)	59 (28-82)	58 (39-78)	52 (27-74)	0.30
Mean length of the CLE in cm (range)	4 (2-8)*	3 (2-7)	3 (2-5)	0.016
Mean number of biopsies (range)	5 (1-17)	4 (1-7)	4 (1-8)	0.27
Interval between subsequent endoscopies in months (range)	42 (12-158)	31 (4-112)	30 (4-117)	0.54
Proton-pump inhibitor use (%)	11/13 (85%)	20/22 (91%)	12/19 (63%)	0.18

IM: intestinal metaplasia, CM: cardiac-type mucosa, CLE: columnar-lined esophagus

* responsible for the significant difference

CA, USA) in a 1:100 dilution, anti-MUC2 (clone Ccp58, Novocastra, Newcastle upon Tyne, UK) in a 1:100 dilution or anti-villin (clone CWWB1, Lab Vision, Fremont CA, USA) in a 1:2000 dilution. Samples were again washed 3 times for 5 minutes with TBS (pH 7.4). Subsequently, biotin-labeled rabbit-anti-mouse antibody (DAKO) was used as second antibody, followed by the addition of a streptavidin-horseradish peroxidase complex (DAKO) using 3-amino-9-ethylcarbazole as substrate. Slides were analyzed for nuclear CDX2 staining, cytoplasmic MUC2 staining and brush border villin staining by two independent investigators (MK, DAB) who were blinded for the presence or absence of IM. CDX2 expression was considered positive if a clear red staining of at least five adjacent nuclei in the same gland was seen, to exclude incidental false positive nuclei. MUC2 expression was present if a red staining in the cytoplasm of (goblet) cells was observed. Villin expression was visualized as a red staining near the apical border of cells.

Statistical analysis

The Chi-square test, Mann-Whitney test, and Kruskal-Wallis test were used to compare the patient characteristics and the immunohistochemical stainings between the three patient groups. A p-value < 0.05 was considered significant. Odds ratios (ORs) with a 95% confidence interval were used as an estimate of the relative risk for the presence of IM. Calculations were initially done with upper endoscopies as the unit of analysis, ignoring the statistical dependency of endoscopies within the same patients. Subsequently, analyses were repeated with the consideration of only one endoscopy per patient. Statistic analyses were conducted using SPSS software (SPSS version 11.0, Chicago, Illinois, USA).

RESULTS

The presence of IM

IM was defined as the presence of goblet cell containing glands. In addition to goblet cells, non-goblet cells can also stain positive with alcian blue. Therefore, the presence of IM was evaluated by light-microscopic examination of H&E stained slides. Consecutive alcian blue, and PAS stained slides were only used to confirm a diagnosis of IM.

IM was observed in 55/122 (45%) biopsy sets. In 67/122 (55%) biopsy sets, only CM was present. The mean number of biopsies taken was five in the IM-group and four in the discordant- and CM-group (similar at the two endoscopies in each group), which was not significantly different (Table 2). When correcting for the length of the columnar segment, the mean number of biopsies taken per centimetre was not different in IM and CM biopsy sets (respectively 1.5/cm (range 0.1-4.3), and 1.5/cm (range 0.3-4.0), p=0.68). Based on the presence of IM, the IM-group consisted of 15 patients, the discordant-group of 25 patients, and the CM-group of 21 patients. Of all patient characteristics, only the length of the CLE differed significantly between the three groups (p=0.016), with CLE being longer in the IM-group, compared to the discordant- and the CM-group (Table 2).

Table 3. Results of immunohistochemical stainings of all biopsy sets

group (no.pts)	IM			CM		
	IM (30)	Discordant (25)	total (55)	Discordant (25)	CM (42)	total (67)
CDX2-positive	30 (100%)	25 (100%)	55 (100%)	13 (52%)**	10 (24%)	23 (34%)
MUC2-positive	30 (100%)	25 (100%)	55 (100%)	11 (44%)	5 (12%)	16 (24%)
villin-positive	22 (73%)	10 (40%)	32 (58%)	4 (17%)*	3 (7%)	7 (10%)

IM: intestinal metaplasia, CM: cardiac-type mucosa

* One sample could not be evaluated since not enough tissue was available

** p=0.019 (compared to CDX2 expression in biopsy sets with CM of the CM-group

CDX2 expression

To investigate the expression of CDX2 protein in IM and CM, CDX2 staining was evaluated. CDX2 expression was observed in all IM positive biopsy sets (Table 3; Figure 1a -b), i.e., in 30 of the IM-group and in 25 of the discordant-group. In addition, CDX2 expression was also observed in

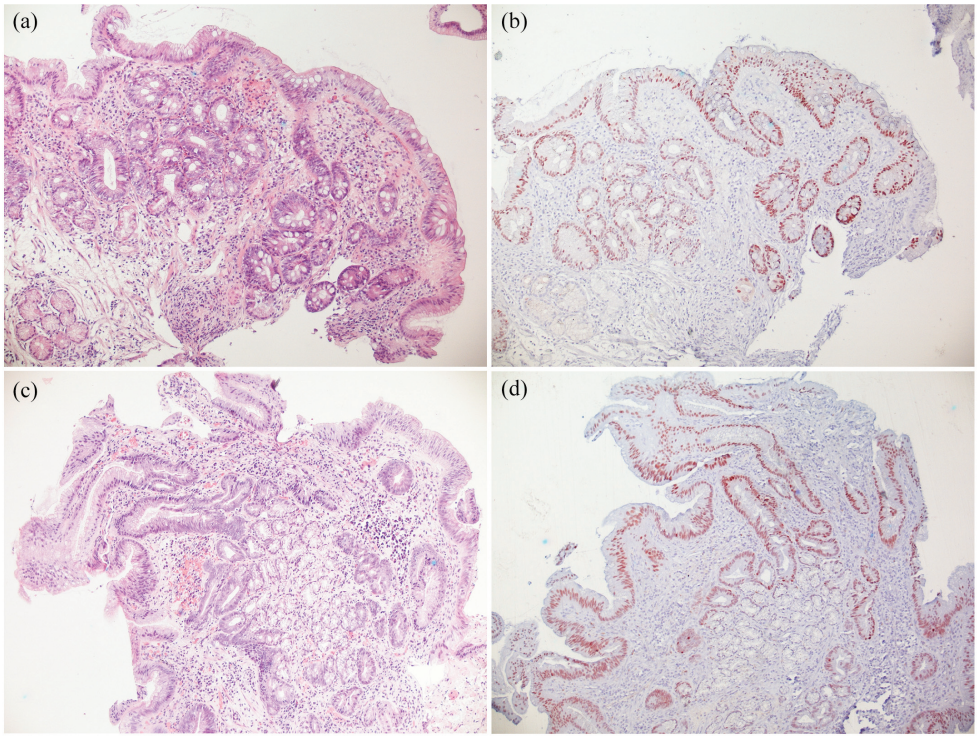


Figure 1. CDX2 expression in columnar epithelium of the esophagus. (A) Intestinal-type columnar epithelium with goblet cells (hematoxylin-eosin). (B) Nuclear staining (red) for CDX2 in intestinal-type columnar epithelium in a serial section of the same patient as in (A). (C) Cardiac-type columnar epithelium without goblet cells (hematoxylin-eosin). (D) CDX2 expression in cardiac-type columnar epithelium in a serial section of the same patient as in (C). Original magnifications x100.

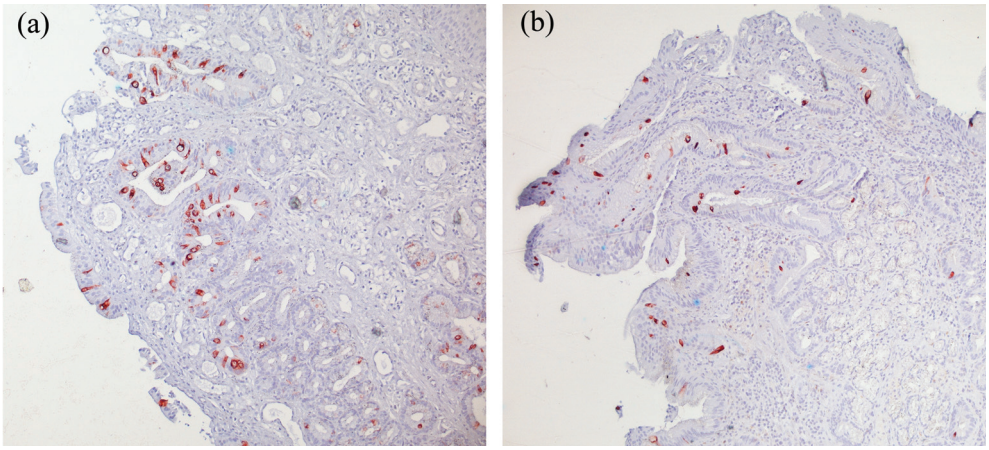


Figure 2. MUC2 expression in columnar epithelium of the esophagus. (A) MUC2 staining in goblet cells (red) in intestinal-type columnar epithelium. (B) MUC2 expression in cardiac-type columnar epithelium without goblet cells in a serial section of the same patient as in 1C. Note that the MUC2 expression is not associated with goblet cells. Original magnifications x100.

23/67 (34% (95% CI: 23-47) biopsy sets without IM (Table 3; Figure 1c-d).

CDX2 was more frequently observed in IM negative biopsy sets of patients of the discordant-group, in which the other biopsy set was positive for IM (13/25; 52%), than in patients of the CM-group, in which IM was absent in both biopsy sets (10/42; 24%) (p=0.019). The presence of CDX2 in CM therefore significantly increased the likelihood of observing IM in another biopsy set of the CLE (OR 3.5, 95% CI: 1.2-10, p=0.021), regardless if CM in the discordant-group was present in biopsies from the first or second upper endoscopy. When we calculated the predictive value of CDX2 expression in biopsies with CM taken during the first endoscopy, for the presence of IM in

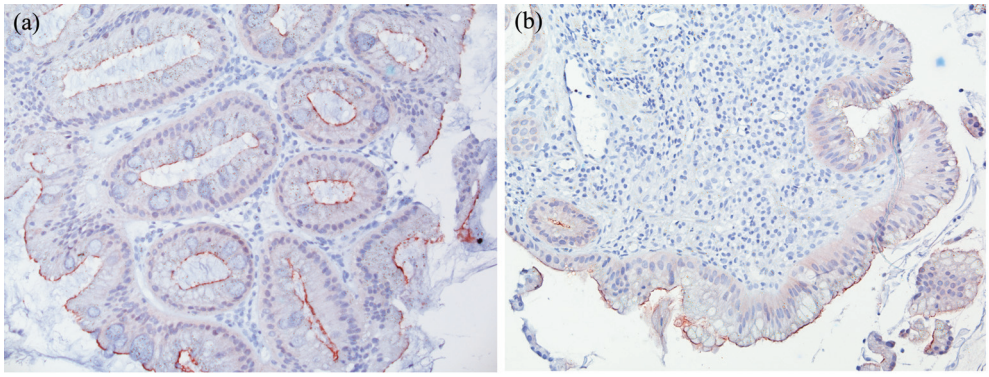


Figure 3. Villin expression in columnar epithelium of the esophagus. (A) Villin staining of the brush border (red) in intestinal-type columnar epithelium. (B) Villin expression in cardiac-type columnar epithelium without goblet cells. Note that the villin expression is not associated with goblet cells. Original magnifications x200.

biopsies of the next endoscopy, and visa versa, the ORs were similar (respectively 4.0, 95% CI:0.8-21, p=0.10 and 3.2, 95% CI:0.8-13, p=0.10). In one patient of group 3, both IM negative biopsies were positive for CDX2.

A longer segment of CLE was not associated with a higher change of CDX2 being present in CM (p=0.135). There was no correlation between the use of proton-pump inhibitors and the presence of CDX2 in CM (p=0.42).

MUC2 expression

Mucins are large glycoproteins forming the main components of the gel-like mucous layer on the surface of the intestine, protecting the mucosa against damaging luminal contents, such the gastro-esophageal refluxate.[24] MUC2 is a mucin specific for IM.[25-27] Since CDX2 regulates the transcription of MUC2,[20] we evaluated the expression of MUC2 in IM and CM. MUC2 staining in goblet cells was found in all biopsy sets with IM (Table 3, Figure 2a), and was mainly localized in the cytoplasm alongside the membrane. Moreover, MUC2 was also expressed in CM in 16/67 (24%) samples without IM. In CM, MUC2 was expressed in the entire cytoplasm of non-goblet columnar cells that did not stain positive with alcian blue (Figure 2b). Thirteen of 16 (81%) MUC2 positive CM samples were also positive for CDX2 in the same region.

Villin expression

Villin is an actin-binding cytoskeletal protein essential for brush border formation (microvilli) in normal end-differentiated epithelial cells of the intestine.[28] Therefore, the presence of a brush border of the esophageal columnar epithelium can be demonstrated by villin expression. We investigated whether villin protein was also expressed in CM in addition to the intestinal markers CDX2 and MUC2. One CM sample could not be evaluated, as there was not enough tissue available for staining. Villin expression was observed in 32/55 (58%) of IM positive biopsy sets (Figure 3a). In 7/66 (11%) CM samples, villin expression was found (Figure 3b), of which five were also CDX2 positive. Four CM samples (6%) were positive for CDX2 and MUC2, as well as for villin.

DISCUSSION

Patients with CM in CLE are currently excluded from surveillance endoscopy, as they are regarded as IM negative and thus as not having a premalignant condition.[13] This study shows a significant relationship between the intestinal marker CDX2 in CM and the presence of IM in biopsies taken at another time point, as CDX2 stained positive in 52% of CM biopsy sets of the discordant-group (with an OR of 3.5), in which the biopsy set of the other endoscopy was positive for IM (Table 2). In our opinion it is unlikely that, despite the two-dimensional analysis of the biopsies, goblet cells have been missed in these CM biopsy sets, as the CDX2 expression was often observed in large areas without goblet cells (Figure 1c-d), and, in addition, in the six consecutive slides also no goblet cells were observed. Therefore, CDX2 staining may represent a useful histological marker for the presence of IM in CLE despite the absence of goblet cells suggestive for IM.

CDX2 expression in CM as an indicator for the presence of IM has been reported in previous

studies.[17, 22] These studies were however cross-sectional, which means that biopsies were only evaluated at one time-point. In contrast, this study was a longitudinal study, in which biopsy sets of two subsequent endoscopies were compared.

Previously, it has been suggested that there are two possible reasons for not detecting IM.[13] Firstly, several authors have proposed that IM may develop over time in a two-step process. It has been suggested that multilayered epithelium, with morphological and immunohistochemical characteristics of both squamous and columnar epithelium, may represent a transitional stage in the development of Barrett's esophagus.[29] Others have suggested that IM develops from previously induced CM in the esophagus under influence of chronic inflammation.[13, 30-32] According to this theory, the finding of CDX2 expression in CM, and in a subset also expression of MUC2 and villin, could indicate early intestinal differentiation prior to morphologic changes such as goblet cells,[17, 33] and in this way being an intermediate stage in the differential shift of CM towards IM.[28, 32]

The second possibility for not detecting IM is sampling error. Although IM is predominantly present in the proximal end of the CLE,[34] IM and CM may have a patchy distribution. Since IM and CM are endoscopically indiscernible from each other, and the presence of IM can be very focal,[35] sampling error for the detection of IM may occur.[13] Sixteen of the twenty-five patients of the discordant-group had IM in their first, and CM in their second biopsy set (Table 1). It seems likely that in these cases the finding of no IM can be contributed to sampling error. The likelihood of detecting IM increased with the number of biopsies taken, and therefore taking not enough biopsies could be a reasonable explanation for missing IM in this group. Since in this study the mean number of biopsies taken per cm was similar in the IM samples and the CM samples, the possibility of sampling error seems to be ruled out. However, since IM has a patchy appearance in the CLE but is predominantly located at the proximal end of the CLE,[12, 34] it is possible that despite taking the same numbers of biopsies, IM could be missed due to taking proportionally less biopsies of the proximal part of the CLE. A similar explanation can be given for the other nine patients of the discordant-group who had CM detected at their first endoscopy, whereas IM was found in biopsies from the second endoscopy. Since the mean interval between two subsequent endoscopies in the discordant-group was with 30 months relatively long, and the development of IM is thought to be a slow process, it is also possible that IM in this subgroup has developed over time.

Although a final conclusion on the cause of not detecting IM in one set of biopsies cannot be given, the ORs for the predicting value of CDX2 in CM in the different subgroups were similar, and thus it is reasonable to assume that CDX2 expression in CM represents a reliable marker for the detection of the premalignant IM in CLE at another time point. In line with this assumption, it is likely that the 24% with CDX2 expression in CM biopsy sets in whom IM was not detected in both biopsy sets taken at different time points, will show IM in biopsies taken at a next endoscopy. Unfortunately, due to exclusion from the surveillance program, these patients have currently not undergone another follow-up upper endoscopy to evaluate this.

CDX2 is a transcription factor for MUC2, which is a mucin specific for IM.[25-27] In our study, as expected, all IM biopsies stained positive for MUC2. In 13/23 (57%) of the CDX2 positive CM biopsies, MUC2 staining was also positive. Villin expression was observed in 58% of the IM-positive

samples. This lower result of villin expression in IM compared to CDX2 expression and MUC2 expression has been suggested to be caused by the fact that the quantity of villin protein needs to have a sufficient level to result in a mature brush border.[36, 37] In addition to villin expression in IM, 5/23 (22%) of the CDX2 positive CM samples also showed villin expression, suggesting the presence of end-differentiated intestinal characteristics in CM. Although less frequent, the presences of MUC2 and villin expression in CM are supportive for the value of CDX2 as indicator of IM in CLE.

A possible limitation of this study is the use of one single technique to detect CDX2 in the biopsies. The major reason that we only used immunohistochemistry was that additional techniques such RT-PCR,[19] could not be performed on our paraffin-embedded tissue, but only on fresh snap frozen biopsies, which were not available in this retrospective study. However, as we performed the CDX2 immunohistochemical stainings with a commonly used dilution,[18, 19] which showed only very specific nuclear staining without background staining in the cytoplasm of cells, it is unlikely that the immunohistochemistry may have resulted in false positive results.

In conclusion, this study shows that the presence of CDX2 in CM might be able to predict the presence of IM in CLE, which was otherwise not detected due to sampling error or developing of IM over time. This suggests that CDX2 staining could be used as an additional marker for the presence of IM in CLE in the absence of goblet cells. A prospective follow-up study on patients with CM in their biopsies should be performed to confirm the predictive value of CDX2. Nonetheless, as the presence of IM is still the gold standard for the presence of premalignant BE, we suggest an additional endoscopy in patients in whom CDX2 expression in CM is demonstrated. This should include the taking of extensive biopsies for the detection of IM (especially near the squamo-columnar junction) to evaluate if endoscopic surveillance is indeed indicated in these patients.

REFERENCES

1. Avidan B, Sonnenberg A, Schnell TG*, et al.* Hiatal hernia size, Barrett's length, and severity of acid reflux are all risk factors for esophageal adenocarcinoma. *Am J Gastroenterol* 2002; 97: 1930-6.
2. Hamilton SR, Smith RR. The relationship between columnar epithelial dysplasia and invasive adenocarcinoma arising in Barrett's esophagus. *Am J Clin Pathol* 1987; 87: 301-12.
3. Haggitt RC. Barrett's esophagus, dysplasia, and adenocarcinoma. *Hum Pathol* 1994; 25: 982-93.
4. Hameeteman W, Tytgat GN, Houthoff HJ*, et al.* Barrett's esophagus: development of dysplasia and adenocarcinoma. *Gastroenterology* 1989; 96: 1249-56.
5. Buttar NS, Wang KK. Mechanisms of disease: Carcinogenesis in Barrett's esophagus. *Nat Clin Pract Gastroenterol Hepatol* 2004; 1: 106-12.
6. Guillem PG. How to make a Barrett esophagus: pathophysiology of columnar metaplasia of the esophagus. *Dig Dis Sci* 2005; 50: 415-24.
7. Chinyama CN, Marshall RE, Owen WJ*, et al.* Expression of MUC1 and MUC2 mucin gene products in Barrett's metaplasia, dysplasia and adenocarcinoma: an immunopathological study with clinical correlation. *Histopathology* 1999; 35: 517-24.
8. Regalado SP, Nambu Y, Iannettoni MD*, et al.* Abundant expression of the intestinal protein villin in Barrett's metaplasia and esophageal adenocarcinomas. *Mol Carcinog* 1998; 22: 182-9.
9. Paull A, Trier JS, Dalton MD*, et al.* The histologic spectrum of Barrett's esophagus. *N Engl J Med* 1976; 295: 476-80.
10. Chandrasoma PT, Lokuhetty DM, Demeester TR*, et al.* Definition of histopathologic changes in gastroesophageal reflux disease. *Am J Surg Pathol* 2000; 24: 344-51.
11. Falk GW. Barrett's esophagus. *Gastroenterology* 2002; 122: 1569-91.
12. Sampliner RE. Updated guidelines for the diagnosis, surveillance, and therapy of Barrett's esophagus. *Am J Gastroenterol* 2002; 97: 1888-95.
13. Oberg S, Johansson J, Wenner J*, et al.* Endoscopic surveillance of columnar-lined esophagus: frequency of intestinal metaplasia detection and impact of antireflux surgery. *Ann Surg* 2001; 234: 619-26.
14. Silberg DG, Swain GP, Suh ER*, et al.* Cdx1 and cdx2 expression during intestinal development. *Gastroenterology* 2000; 119: 961-71.
15. Beck F, Erler T, Russell A*, et al.* Expression of Cdx-2 in the mouse embryo and placenta: possible role in patterning of the extra-embryonic membranes. *Dev Dyn* 1995; 204: 219-27.
16. Silberg DG, Sullivan J, Kang E*, et al.* Cdx2 ectopic expression induces gastric intestinal metaplasia in transgenic mice. *Gastroenterology* 2002; 122: 689-96.
17. Phillips RW, Frierson HF, Jr., Moskaluk CA. Cdx2 as a marker of epithelial intestinal differentiation in the esophagus. *Am J Surg Pathol* 2003; 27: 1442-7.
18. Eda A, Osawa H, Satoh K*, et al.* Aberrant expression of CDX2 in Barrett's epithelium and inflammatory esophageal mucosa. *J Gastroenterol* 2003; 38: 14-22.
19. Moons LM, Bax DA, Kuipers EJ*, et al.* The homeodomain protein CDX2 is an early marker of Barrett's oesophagus. *J Clin Pathol* 2004; 57: 1063-8.
20. Yamamoto H, Bai YQ, Yuasa Y. Homeodomain protein CDX2 regulates goblet-specific MUC2 gene expression. *Biochem Biophys Res Commun* 2003; 300: 813-8.
21. Houde M, Laprise P, Jean D*, et al.* Intestinal epithelial cell differentiation involves activation of p38 mitogen-activated protein kinase that regulates the homeobox transcription factor CDX2. *J Biol Chem* 2001; 276: 21885-94.
22. Groisman GM, Amar M, Meir A. Expression of the intestinal marker Cdx2 in the columnar-lined esophagus with and without intestinal (Barrett's) metaplasia. *Mod Pathol* 2004; 17: 1282-8.
23. Steininger H, Pfofe DA, Muller H*, et al.* Expression of CDX2 and MUC2 in Barrett's mucosa. *Pathol Res*

Pract 2005; 201: 573-7.
24. Corfield AP, Myerscough N, Longman R*, et al.* Mucins and mucosal protection in the gastrointestinal tract: new prospects for mucins in the pathology of gastrointestinal disease. *Gut* 2000; 47: 589-94.
25. Warson C, Van De Bovenkamp JH, Korteland-Van Male AM*, et al.* Barrett's esophagus is characterized by expression of gastric-type mucins (MUC5AC, MUC6) and TFF peptides (TFF1 and TFF2), but the risk of carcinoma development may be indicated by the intestinal-type mucin, MUC2. *Hum Pathol* 2002; 33: 660-8.
26. Bax DA, Haringsma J, Einerhand AW*, et al.* MUC4 is increased in high grade intraepithelial neoplasia in Barrett's oesophagus and is associated with a proapoptotic Bax to Bcl-2 ratio. *J Clin Pathol* 2004; 57: 1267-72.
27. Guillem P, Billeret V, Buisine MP*, et al.* Mucin gene expression and cell differentiation in human normal, premalignant and malignant esophagus. *Int J Cancer* 2000; 88: 856-61.
28. Niwa T, Ikehara Y, Nakanishi H*, et al.* Mixed gastric- and intestinal-type metaplasia is formed by cells with dual intestinal and gastric differentiation. *J Histochem Cytochem* 2005; 53: 75-85.
29. Shields HM, Rosenberg SJ, Zwas FR*, et al.* Prospective evaluation of multilayered epithelium in Barrett's esophagus. *Am J Gastroenterol* 2001; 96: 3268-73.
30. Oberg S, Peters JH, DeMeester TR*, et al.* Inflammation and specialized intestinal metaplasia of cardiac mucosa is a manifestation of gastroesophageal reflux disease. *Ann Surg* 1997; 226: 522-30; discussion 30-2.
31. Hamilton SR, Yardley JH. Regnerative of cardiac type mucosa and acquisition of Barrett mucosa after esophagogastrostomy. *Gastroenterology* 1977; 72: 669-75.
32. Chandrasoma P. Controversies of the cardiac mucosa and Barrett's oesophagus. *Histopathology* 2005; 46: 361-73.
33. Glickman JN, Shahsafaei A, Odze RD. Mucin core peptide expression can help differentiate Barrett's esophagus from intestinal metaplasia of the stomach. *Am J Surg Pathol* 2003; 27: 1357-65.
34. Chandrasoma PT, Der R, Dalton P*, et al.* Distribution and significance of epithelial types in columnar-lined esophagus. *Am J Surg Pathol* 2001; 25: 1188-93.
35. Cassaro M, Di Mario F, Leandro G*, et al.* The dark side of the gastric biopsy. *Hum Pathol* 1999; 30: 741-4.
36. Fitzgerald RC, Omary MB, Triadafilopoulos G. Dynamic effects of acid on Barrett's esophagus. An ex vivo proliferation and differentiation model. *J Clin Invest* 1996; 98: 2120-8.
37. Kumble S, Omary MB, Fajardo LF*, et al.* Multifocal heterogeneity in villin and Ep-CAM expression in Barrett's esophagus. *Int J Cancer* 1996; 66: 48-54.

3

Predicting presence of intestinal metaplasia and dysplasia in columnar-lined esophagus: a multivariable analysis

Endoscopy (in press)

Marjon Kerkhof [1]
Ewout W Steyerberg [2]
Johannes G Kusters [1]
Ernst J Kuipers [1]
Peter D Siersema [1]

Departments of [1]Gastroenterology and Hepatology, and [2]Public Health,
Erasmus MC - University Medical Center Rotterdam, Rotterdam, the Netherlands

ABSTRACT

Background and study aims: We evaluated clinical risk factors in patients with presumed Barrett esophagus that could predict the presence of IM and dysplasia in biopsies of columnar-lined esophagus (CLE) independent of histological results.

Patients and methods: In 908 patients with a CLE of ≥ 2 cm, data on age, gender, reflux symptoms, tobacco and alcohol use, medication use and upper gastrointestinal endoscopy findings were prospectively collected. Multivariable logistic regression analysis was performed, and a model for predicting the histological results was developed.

Results: In 127/908 patients, biopsies of CLE did not contain IM (No IM). Of the 781 patients with IM, 663 (85%) patients had no dysplasia (ND), and 118 (15%) low-grade dysplasia (LGD). Most important predictors for the presence of IM were length of CLE, size of hiatal hernia and male gender, while among those with IM age and male gender were most important for the presence of LGD. Multivariable combinations of these predictors yielded reliable models, which were able to discriminate IM well from No IM (area under ROC curve: 0.82), but only reasonably discriminated LGD from ND (area: 0.65).

Conclusions: A simple model based on clinical findings is able to predict the presence of IM in biopsies from CLE. In contrast, predicting the presence of LGD versus ND in IM is more difficult. Predictions from these models may aid in the decision-making on whether surveillance should be performed in a patient with CLE in view of the known sampling error at endoscopy and interobserver variability at histology.

INTRODUCTION

Barrett esophagus (BE) is a premalignant condition which is predominantly caused by chronic gastro-esophageal reflux,[1] and is characterized by the replacement of the squamous epithelium of the esophagus by columnar epithelium with goblet cells (specialized intestinal metaplasia (IM)).[2]

Cardiac-type mucosa (CM) is also frequently observed in the columnar-lined esophagus (CLE),[3] with the absence of goblet cells as the only histological difference from IM.[4] CM, in contrast to IM, is not regarded as a premalignant condition.[5, 6] Therefore, according to the guidelines of the American College of Gastroenterology (ACG), only patients with IM in CLE are currently advised to undergo periodic endoscopic surveillance to detect progression to dysplasia in an early, potentially curable stage.[7] It has however been suggested that patients with CM in biopsies from CLE are at an increased risk of having undetected IM due to either sampling error or the development of IM over time, and thus for developing esophageal adenocarcinoma (EAC) on the long term.[8, 9] Current guidelines would incorrectly exclude these patients, being also at risk for neoplastic progression, from a surveillance program.[7] Therefore, it is relevant to perform risk stratification by using easily available clinical characteristics, which are able to define which patients with CLE with or without IM should undergo endoscopic follow-up. The most prominent risk factors for the presence of IM in CLE have been suggested to be male gender, frequent reflux episodes, presence and size of a hiatal hernia (HH), and length of CLE.[1, 10-13] However, in most of these studies, patients with IM in CLE were compared with controls without having CLE and not with controls with CLE without IM.

In a subset of patients, BE progresses via a stepwise process from low-grade dysplasia (LGD) to high-grade dysplasia, and subsequently to EAC.[6, 14-16] Patients with no dysplasia (ND) in biopsies of BE are regarded to be at a lower risk of neoplastic progression than patients with LGD.[17] Therefore, the presence or absence of dysplasia determines the frequency of surveillance upper endoscopy (e.g. yearly for LGD, and every 3 years for ND).[7, 17, 18] Previous studies have demonstrated considerable interobserver variability in the interpretation of dysplasia,[19, 20] which may lead to superfluous follow-up endoscopies (in case of overdiagnosis), or insufficient control (in case of underdiagnosis).[19] Age, length of the BE, and size of the HH have previously been suggested as risk factors for developing EAC in BE.[21, 22] Supplementary risk stratification for the presence of LGD in patients with CLE based on clinical characteristics could be of additional value for surveillance strategies.

The aim of this study was to investigate which clinical characteristics were predictive for the presence of IM and dysplasia in CLE, and to develop a simple model to estimate the probabilities of the presence of IM, ND and LGD in biopsies of CLE.

MATERIALS AND METHODS

Patients

This study was part of an ongoing, prospective clinical trial (CYBAR study) assessing the value of flow cytometry for individualizing the frequency of follow-up upper endoscopy in patients with BE. From October 2003 to December 2004, consecutive incident and prevalent patients with an endoscopic CLE of at least 2 cm were included in 15 Dutch hospitals. Patients were excluded if they

were unwilling to give informed consent, or if HGD/EAC was present at the index-endoscopy. The institutional review boards of all participating hospitals approved this study and informed consent was obtained from all patients prior to endoscopy.

Data collection and endoscopy

Several data were recorded at enrollment in the study (Table 1), as were endoscopical and histological findings. Endoscopy was performed by experienced endoscopists. The presence of reflux esophagitis was graded according to the Los Angeles classification.[23] The length of CLE was determined by measuring the distance between the squamocolumnar junction (the most proximal

Table 1. Patient characteristics per patient group

	No IM	**IM+ND**	**IM+LGD**	
	n=127	n=663	n=118	p-value
Age, mean ± SD (y)	55 ± 12	60 ± 11	65 ± 12	<0.001
Male gender (%)	79 (62%)	479 (72%)	97 (82%)	<0.001
CLE length, mean ± SD (cm)	2.7 ± 1.2	4.4 ± 2.5	5.0 ± 3.0	<0.001
HH present	88 (70%)	572 (87%)	102 (86%)	<0.001
If HH present: mean size ± SD (cm)	2.6 ± 1.3	3.5 ± 1.7	4.0 ± 1.8	<0.001
BMI, mean ± SD	27 ± 3.7	27 ± 3.9	27 ± 3.7	0.99
Smoking	26 (21%)	141 (21%)	18 (15%)	0.85
Alcohol	108 (78%)	507 (76%)	84 (71%)	0.03
Medication use				
PPI	110 (87%)	595 (90%)	107 (91%)	0.61
other antacidic med	15 (12%)	85 (13%)	11 (9%)	0.55
NSAID	13 (10%)	30 (5%)	5 (4%)	0.03
aspirin	6 (5%)	83 (13%)	18 (15%)	0.02
COX-2	2 (2%)	11 (2%)	0	0.37
GERD symptoms				
heartburn	48 (38%)	205 (31%)	31 (26%)	0.13
regurgitation	45 (36%)	165 (25%)	20 (17%)	<0.001
dysphagia	26 (21%)	89 (13%)	19 (16%)	0.12
Reflux esophagitis present	14 (11%)	63 (10%)	15 (13%)	0.53
grade A	6 (43%)	19 (30%)	8 (53%)	
grade B	4 (29%)	35 (56%)	6 (40%)	0.26
grade C	3 (21%)	7 (11%)	1 (7%)	
grade D	1 (7%)	2 (3%)	0	

IM, intestinal metaplasia; LGD, low-grade dysplasia; SD, standard deviation; CLE, columnar-lined esophagus; HH, hiatal hernia; BMI, body mass index; PPI, proton pump inhibitor; NSAID, nonsteroidal anti-inflammatory drug; COX-2, cox-2 inhibitor; GERD, gastroesophageal reflux disease

location where the light-pink mucosa of the squamous-lined esophagus joins the red mucosa of the CLE) and the endoscopic lower esophageal sphincter (diaphragma indentation) or, in case of a HH, the proximal margin of the longitudinal gastric folds.[24] The size of a HH was determined by measuring the distance between the proximal margin of the longitudinal gastric folds and the diaphragma indentation. According to current guidelines,[7] four-quadrant biopsies were obtained at 2-cm intervals along the length of the circumferential CLE, to identify the presence of IM and if so, the presence of dysplasia at histological investigation. Two biopsies at 2-cm intervals were taken of tongues of CLE comprising less than 50% of the circumference.

Histology

Biopsy specimens were fixed in 10% formalin, and embedded in paraffin. Sections of 4 μm each from every biopsy set were mounted on adhesive slides, dried overnight at 37°C, and deparaffinized with xylene. These slides were stained with haematoxylin and eosin (H&E) to determine the type of columnar epithelium (CM or IM). Pathologists of all participating hospitals (see Appendix) assessed the H&E stained slides of biopsies of patients included in their own center for the presence of IM and the grade of dysplasia. IM was defined as the presence of goblet cells in CLE.[7] Dysplasia was graded according to the Consensus Criteria of 1988, with adjustments made in 2001.[20, 25] Subsequently, one of a panel of five expert (panel expert) pathologists randomly reviewed the slides. These panel experts were blinded for age, sex, identity of the patient, and diagnosis of the initial pathologist. If there was disagreement on the presence of IM or grade of dysplasia between the initial pathologist and the panel expert, slides were blindly reviewed by a second member of the expert panel. A final diagnosis was established by a majority diagnosis, meaning that 2 of 3 pathologists had to agree upon the diagnosis, as previously reported.[19]

Statistical analysis and model performance

Chi-square tests, Mann-Whitney tests, and Kruskal-Wallis tests were used to compare the clinical predictors for presence of IM, ND or LGD. A p-value < 0.05 was considered statistically significant. Odds ratios (ORs) with 95% confidence intervals were used as measure of association. Associations between predictors and outcomes were first estimated univariately. Multivariable logistic regression analysis was applied to estimate the probability of the presence of IM in CLE. A second logistic regression model estimated the relative probability of LGD among patients with IM. The use of these two models was motivated by the clinical notion that the probability of the presence of IM is of predominant importance for the decision to perform endoscopic surveillance,[7] and that the presence of ND or LGD in IM will further determine the frequency of endoscopic follow-up.

Predictors were selected with a backward stepwise elimination procedure with $p \leq 0.20$ for inclusion.[26] The relevance of predictors was expressed by partial R^2 statistics, which account for the strength of effect (odds ratio) and the prevalence of a predictor. An infrequent predictor with high odds ratio has a lower R^2 than a more frequent predictor with the same odds ratio. The partial R^2 was calculated as the difference in Nagelkerke's R^2 statistics between a model including all predictors and a model that excluded one predictor at a time.[27]

Table 2. Relationship of predictors with the histological results of biopsies of CLE (ORs with 95% Confidence Interval and Partial R2 value)

	IM[#] vs. No IM			LGD vs. ND		
Variable	**univariate**	**multivariable [##]**	**Partial R^2**	**univariate**	**multivariable [##]**	**Partial R^2**
Age (decades)**	1.49 (1.27-1.75)*	1.30 (1.07-1.57)*	1.2%	1.39 (1.16-1.67)*	1.37 (1.14-1.66)*	2.6%
Male gender	1.71 (1.15-2.53)*	2.34 (1.46-3.76)*	2.1%	1.77 (1.08-2.93)*	2.55 (1.47-4.42)*	2.5%
Alcohol use	1.91 (1.12-1.85)*	2.00 (1.09-3.68)*	0.9%	1.43 (0.92-2.11)	1.61 (1.00-2.58)	0.7%
Heartburn	0.72 (0.48-1.06)	-		0.78 (0.50-1.21)	-	
Regurgitation	0.57 (0.38-0.85)*	0.55 (0.35-0.86)*	1.1.%	0.61 (0.37-1.02)	0.62 (0.36-1.05)	0.7%
Dysphagia	0.63 (0.39-1.01)	-		1.22 (0.71-2.09)	-	
NSAID	0.41 (0.21-0.80)*	0.38 (0.17-0.86)*	0.9%	0.93 (0.35-2.45)	-	
Aspirin	2.99 (1.29-6.97)*	2.16 (0.88-5.31)	0.6%	1.26 (0.72-2.18)	-	
Length of CLE (cm)***	1.85 (1.55-2.21)*	1.63 (1.37-1.95)*	8.2%	1.08 (1.01-1.16)*	-	
Size of HH (cm)***	1.65 (1.42-1.93)*	1.53 (1.28-1.83)*	4.4%	1.12 (1.01-1.23)*	1.13 (1.02-1.26)*	1.0%

* $p < 0.05$

** increase in risk for presence of IM or LGD for every 10 year increase in age

*** increase in risk for presence of IM or LGD for every cm increase

[#] ND and LGD together

[##] multivariable ORs only calculated if $p \leq 0.20$

The predictive accuracy of multivariable models is reflected in discrimination and calibration. Discrimination was assessed using receiver operating characteristic (ROC) analysis. The area under the ROC curve represents the proportion in which patients with a certain outcome (e.g. IM or LGD) had a higher probability than patients without that outcome.[28] Calibration (or reliability) refers to the degree of agreement between predicted and observed outcomes. Calibration was assessed by the Hosmer-Lemeshow goodness-of-fit test,[28] and graphically by plotting observed frequencies of the outcome (IM or LGD) against predicted probabilities.

Internal validity of model performance indicates that the results of the analysis hold for the data under study. Internal validity was assessed with bootstrapping techniques.[29] Moreover, bootstrap estimates were used to derive the final predictive models by correcting the logistic regression coefficients for overoptimism.[26]

Statistical analyses were conducted using SPSS software (SPSS version 11.0, Chicago, IL, USA), and R software (version 2.2.0, http://www.r-project.org/).

RESULTS

Patient characteristics

A total of 956 consecutive patients were enrolled in this study. Forty-eight patients were excluded (8 patients decided shortly after inclusion to withdraw, in 25 patients HGD/EAC was found at index endoscopy and 15 patients had a CLE segment <2 cm). Consequently, 908 patients (665 men and

253 women) were included in this study, with a mean age of 60 ± 12 years (range 19-88 years). Mean and median lengths of the columnar-lined segment were 4 cm, with a range of 2-16 cm, and mean number of biopsies taken per 2 cm CLE length was 4. Of these 908 patients, 127 (14%) had a CLE with columnar epithelium without goblet cells (No IM). BE was histologically confirmed by the presence of IM in CLE in 781 (86%) patients, of which 663 (85%) had IM without dysplasia (ND), and 118 (15%) IM with LGD. Patient characteristics of the three patient groups are summarized in Table 1.

There was a gradual increase in mean age, proportion of males, mean CLE length, and mean HH size in patients with respectively No IM, IM with ND and IM with LGD. A HH was significantly more often found in patients with IM compared to those without IM.

Univariate and multivariable analysis

Table 2 shows the results of the univariate and multivariable analyses of No IM vs. IM, and if IM was present, of ND vs. LGD. At IM vs. No IM analysis, the multivariable OR for length of CLE was 1.63, which indicates that with every centimeter increase in length of CLE, the risk of having IM in the segment increased by 63%. In contrast, length of CLE was not a significant predictor for the presence of LGD in CLE with IM at multivariable analysis (univariate OR 1.08). Length of CLE was the strongest predictor for the presence of IM (R^2 8.2%), followed by size of HH (R^2 4.4%), and male gender (R^2 2.1%). Age (R^2 2.6%) and male gender (R^2 2.5%) were most predictive for the presence of LGD in CLE with IM.

Model evaluation and application

The discriminative ability of the multivariable model for predicting the presence of IM in CLE was good (area=0.82, internally validated: 0.81) (Figure 1), whereas the model for LGD in CLE with IM had less discriminative ability (area=0.65, internally validated 0.64).

The two multivariable models were presented in a predictive score chart (Table 3). This score chart was intended to facilitate the estimation of the probabilities of IM in CLE, and the presence

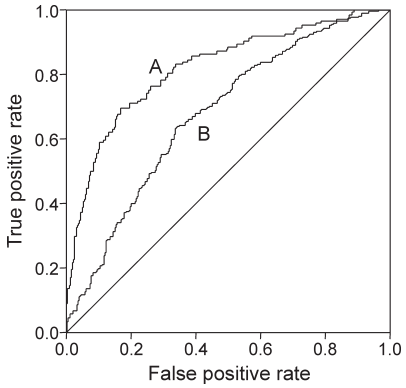


Figure 1. ROC curves of the models predicting IM (A) and distinguishing LGD from ND (B), indicating discriminative ability. The areas under the curves are 0.82 for IM vs. No IM and 0.65 for LGD vs. ND.

Table 3. Predictive score chart for the probability of IM and the relative probability of LGD in biopsies of CLE

Predictor	IM	LGD
Age (years):		
< 50	0	0
50-69	+1	+2
>= 70	+2	+4
Male gender	+2	+2
Alcohol use	+1	+1
Regurgitation	-1	-1
NSAID	-2	-
Aspirin	+2	-
Length CLE (cm):		
2	+2	-
3	+3	-
4	+4	-
5	+5	-
6-7	+6	-
8+	+7	-
Size of HH (cm)*:		
1	+1	0
2	+2	0
3	+3	0
4	+4	+1
5+	+5	+1
Sumscore: add relevant scores **	….	….

* If no HH was present, a size of 1 cm was assumed

** the exact formulas to calculate the predicted probabilities with the sumscore are:

lpIM= -2.19+0.5 * score IM; predicted probability of IM = 1 / [1+e- (lpIM)]
lpLGD= -3.21 + 0.35 * score LGD; relative probability of LGD = 1 / [1+e- (lpLGD)]
Predicted probability of LGD = relative probability of LGD * predicted probability of IM
Predicted probability of ND = Predicted probability of IM - predicted probability of LGD
Predicted probability of No IM = 1 – predicted probability of IM

of LGD in biopsies of CLE with IM in clinical practice. Scores for each predictor were derived from the logistic regression coefficients (multiplied by 2 in the IM model and by 3 in the LGD model). Values for continuous predictors were presented in such a way that the scores show small steps between intervals, but scores for intermediate values can be estimated by linear interpolation.

For individual patients, the scores corresponding to the values of the predictors can be filled in on the score chart. An individual sumscore consists of the sum of all scores. Figure 2 shows the

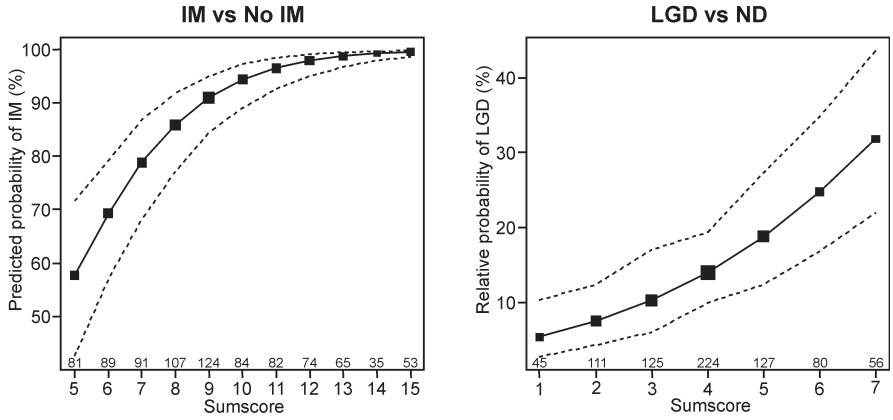


Figure 2. Predicted probabilities for IM vs No IM and relative probabilities for LGD vs ND among those with IM in CLE, corresponding to the sumscore as calculated with the score chart (Table 3). Numbers of patients per score are shown at the bottom of the graph. For example, a sumscore of +10 for IM corresponds to a probability of 94% [95% confidence interval 89% – 97%] in the left panel. To extract the predicted (absolute) probability for LGD in IM from the right panel, the score has to be multiplied with the score of IM (left panel). For example, a sumscore of +5 for LGD and +10 for IM corresponds with a predicted probability for LGD of 19%*94% = 18%.

probabilities corresponding to this sumscore.

We used the score chart for two hypothetical 70-year-old males (Table 4). Patient 1 had a CLE of 4 cm, and a HH of 2 cm. His medication use consisted of aspirin, but no NSAIDs. He did not have GERD symptoms, nor did he regularly consume alcohol. Patient 2 had different characteristics. The sumscores for patient 1 for IM and LGD were +13 and +7, respectively, and those for patient 2 were +5 and +6, respectively. According to Figure 2, patient 1 had a probability of finding IM of 98%, while patient 2 had a probability of 58%. Their relative probabilities of finding LGD were 31% and 24%, resulting in absolute probabilities of LGD of approximately 31% * 98% = 30%, and 24% * 58% = 14%, respectively. The exact probabilities can be calculated using the formulas shown under Table 3.

DISCUSSION

We evaluated the value of various clinical characteristics for the prediction of the presence of IM and LGD in a large prospective study of patients with CLE. This enabled us to construct a simple but powerful predictive model that may allow further refinement of the surveillance strategy for individual patients with CLE.

A longer length of CLE, a larger size of HH, and male gender were the most important predictors for the presence of IM in CLE. In previous studies from other groups, these factors were also reported to be associated with CLE.[1, 10-13] Although the length of CLE has been shown to be associated with the size of HH,[11, 30] we found in this study that the size of HH was also an independent risk factor for the presence of IM. In the multivariable analysis, age and male gender were also the most

Table 4. Illustration of the application of the predictive score chart in two hypothetical patients

	patient 1	patient 2	IM: pt 1	IM: pt 2	LGD: pt 1	LGD: pt 2
Age (years)	70	70	+2	+2	+4	+4
Male gender	yes	yes	+2	+2	+2	+2
Alcohol use	yes	no	+1	0	+1	0
Regurgitation	no	yes	0	-1	0	-1
NSAID	no	yes	0	-2	-	-
Aspirin	yes	no	+2	0	-	-
Length CLE (cm)	4	4	+4	+4	-	-
Size of HH (cm)	2	0	+2	0	0	0
Sumscore: add relevant scores			+13	+5	+7	+6
Predicted probabilities:						
Using Figure 2			~98%	~58%*	~30%	~14%*
Using the Formulas (Table 3)						
• No IM	1 %	42 %				
• IM	99 %	58 %				
- ND	67 %	44 %				
- LGD	32 %	14 %				

* Figure 2 provides relative probabilities for LGD. Predicted probability of LGD = relative probability of LGD * predicted probability of IM (see formula in Table 3). For patient 1: 31% * 98%=30%, for patient 2: 24% * 58%=14%.

prominent predictors for the presence of LGD in CLE with IM. These findings are in line with previous studies evaluating risk factors for EAC.[21, 31]

In a recently published study [32] a significant difference in BMI was found between patients with BE and a control group with a normal esophagus. The authors concluded that overweight was associated with an increased risk of developing BE. Furthermore, a relationship between a high BMI and the risk of EAC development has been reported in various case-control studies.[31, 33, 34] In our study, the mean BMI was high (category obese)[10] in all three groups, without significant differences between these groups. Therefore our results confirm the findings of a recent study of Bu et al. that not IM per se, but the presence of CLE (regardless the type of epithelium) may be associated with a high BMI,[35] and consequently BMI seems to be not discriminative for improving surveillance strategies.

Based on the clinical predictors, we developed one model that could predict the finding of IM in biopsies taken from CLE, and a second model that could be able to distinguish ND from LGD in patients with IM in CLE. Since histological evaluation of biopsies is prone to sampling error and interobserver variability, these models could be of additional value in determining the

optimal surveillance frequency. The two models were found to be highly reliable. Internal validation procedures showed that IM could well be discriminated from No IM, however that LGD could only reasonably well be discriminated from ND. One should however realize that the absolute probability of LGD depends on the probability of the presence of IM, as was estimated with the first model.

The final models were presented in a predictive score chart. The probability of finding IM in biopsies from the CLE was high for the first presented patient (99%), but substantial lower for the second patient (58%) (Table 4), suggesting that even in the absence of IM in biopsies taken from CLE, at least the first patient should undergo endoscopic surveillance. It remains to be established what the frequency of this surveillance interval should be. The more general question is whether it is possible to define thresholds for the decision to perform surveillance and to determine optimal surveillance intervals. These thresholds should be determined in future studies and related to the expected benefits (detecting neoplastic progression at an early treatable stage) and cost-effectiveness of surveillance.[36, 37]

The strengths of our study are its prospective study design and size of the patient cohort. In addition, in order to minimize the known effect of substantial interobserver variation,[19, 20] biopsies in our study were reviewed by a panel of expert pathologists, resulting in a consensus diagnosis. In many reported studies, comparisons of clinical characteristics between patients with dysplasia versus no dysplasia were performed on the histological diagnosis of just one (expert) pathologist,[21, 22, 38] and these comparisons should therefore be interpreted with some caution.

There are also some limitations to our study. First, our predictive model was based on the comparison of three histologically different patient groups. In patients of the No IM group, IM could have been missed because of sampling error. It is however reasonable to assume that this occurred at a low frequency, since the endoscopists involved in this study had to follow a strict biopsy protocol for this prospective study. In addition, including patients with undetected IM in the No IM group would primarily have diluted the differences between both groups. Second, there may have been a selection bias. Patients with No IM in CLE at a previous endoscopy are currently not undergoing endoscopic surveillance in countries where the guidelines of the ACG are followed.[7] Therefore, the No IM group was proportionally smaller than it would have been if all these patients would have been advised to undergo follow-up endoscopy in the study period. The number of patients that could be analyzed was however still large enough to make this study sufficient statistically powered. Third, nonparticipation among patients may also have introduced bias. However, as nonparticipation was very low (data not shown) it is unlikely that these few patients could have influenced the results. Finally, intra- and interobserver variability in the interpretation of the length of CLE and size of HH is obviously a limitation, since different endoscopists performed the endoscopies. In a previous study, it was shown that the intraobserver agreement for length of CLE was only fair (κ=0.40), but agreement between true and measured length of CLE was considerable (κ=0.72).[39]

In conclusion, we found that the most important predictors for the presence of IM were length of CLE, size of HH and male gender, while age and male gender were the predominant determinants for the differentiation between LGD and ND among those with IM in CLE. The proposed predictive models were able to estimate the histology of biopsies (especially IM), based on easily available

clinical predictors. Although these models cannot replace histology-based surveillance yet, the predicted probabilities may aid in the decision on whether surveillance should be performed in a patient with CLE in view of the known sampling error at endoscopy and interobserver variability at histology. Future studies are needed to determine thresholds in the models for the decision to perform surveillance and at which interval, based on the expected benefits and cost-effectiveness of such a program.

REFERENCES

1. Avidan B, Sonnenberg A, Schnell TG*, et al.* Hiatal hernia and acid reflux frequency predict presence and length of Barrett's esophagus. *Dig Dis Sci* 2002; 47: 256-64.
2. Guillem PG. How to make a Barrett esophagus: pathophysiology of columnar metaplasia of the esophagus. *Dig Dis Sci* 2005; 50: 415-24.
3. Paull A, Trier JS, Dalton MD*, et al.* The histologic spectrum of Barrett's esophagus. *N Engl J Med* 1976; 295: 476-80.
4. Chandrasoma PT, Lokuhetty DM, Demeester TR*, et al.* Definition of histopathologic changes in gastroesophageal reflux disease. *Am J Surg Pathol* 2000; 24: 344-51.
5. Falk GW. Barrett's esophagus. *Gastroenterology* 2002; 122: 1569-91.
6. Hamilton SR, Smith RR. The relationship between columnar epithelial dysplasia and invasive adenocarcinoma arising in Barrett's esophagus. *Am J Clin Pathol* 1987; 87: 301-12.
7. Sampliner RE. Updated guidelines for the diagnosis, surveillance, and therapy of Barrett's esophagus. *Am J Gastroenterol* 2002; 97: 1888-95.
8. Oberg S, Johansson J, Wenner J*, et al.* Endoscopic surveillance of columnar-lined esophagus: frequency of intestinal metaplasia detection and impact of antireflux surgery. *Ann Surg* 2001; 234: 619-26.
9. Kerkhof M, Bax DA, Moons LMG, van Vuuren AJ, van Dekken H, Steyerberg EW, Kuipers EJ, Kusters JG, Siersema PD. CDX2 expression predicts Barrett's metaplasia in oesophageal columnar epithelium without goblet cells. *Aliment Pharmacol Ther* 2006; 24: 1613-21.
10. Campos GM, DeMeester SR, Peters JH*, et al.* Predictive factors of Barrett esophagus: multivariate analysis of 502 patients with gastroesophageal reflux disease. *Arch Surg* 2001; 136: 1267-73.
11. Cameron AJ. Barrett's esophagus: prevalence and size of hiatal hernia. *Am J Gastroenterol* 1999; 94: 2054-9.
12. Conio M, Filiberti R, Blanchi S*, et al.* Risk factors for Barrett's esophagus: a case-control study. *Int J Cancer* 2002; 97: 225-9.
13. Aste H, Bonelli L, Ferraris R*, et al.* Gastroesophageal reflux disease: relationship between clinical and histological features. GOSPE. Gruppo Operativo per lo Studio delle Precancerosi dell'Esofago. *Dig Dis Sci* 1999; 44: 2412-8.
14. Haggitt RC. Barrett's esophagus, dysplasia, and adenocarcinoma. *Hum Pathol* 1994; 25: 982-93.
15. Hameeteman W, Tytgat GN, Houthoff HJ*, et al.* Barrett's esophagus: development of dysplasia and adenocarcinoma. *Gastroenterology* 1989; 96: 1249-56.
16. Buttar NS, Wang KK. Mechanisms of disease: Carcinogenesis in Barrett's esophagus. *Nat Clin Pract Gastroenterol Hepatol* 2004; 1: 106-12.
17. Hage M, Siersema PD, van Dekken H*, et al.* Oesophageal cancer incidence and mortality in patients with long-segment Barrett's oesophagus after a mean follow-up of 12.7 years. *Scand J Gastroenterol* 2004; 39: 1175-9.
18. Mashimo H, Wagh MS, Goyal RK. Surveillance and screening for Barrett esophagus and adenocarcinoma. *J Clin Gastroenterol* 2005; 39: S33-41.
19. Kerkhof M, van Dekken H, Steyerberg EW, Meijer GA, Mulder AH, de Bruine A, Driessen A, ten Kate FJ, Kusters JG, Kuipers EJ, Siersema PD. Grading of dysplasia in Barrett's oesophagus:substantial interobserver variation between general and gastrointestinal pathologists. *Histopathology* 2007 (in press).
20. Montgomery E, Bronner MP, Goldblum JR*, et al.* Reproducibility of the diagnosis of dysplasia in Barrett esophagus: a reaffirmation. *Hum Pathol* 2001; 32: 368-78.
21. Gopal DV, Lieberman DA, Magaret N*, et al.* Risk factors for dysplasia in patients with Barrett's

esophagus (BE): results from a multicenter consortium. *Dig Dis Sci* 2003; 48: 1537-41.
22. Avidan B, Sonnenberg A, Schnell TG*, et al.* Hiatal hernia size, Barrett's length, and severity of acid reflux are all risk factors for esophageal adenocarcinoma. *Am J Gastroenterol* 2002; 97: 1930-6.
23. 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: 44-9.
24. Sharma P, McQuaid K, Dent J*, et al.* A critical review of the diagnosis and management of Barrett's esophagus: the AGA Chicago Workshop. *Gastroenterology* 2004; 127: 310-30.
25. Reid BJ, Haggitt RC, Rubin CE*, et al.* Observer variation in the diagnosis of dysplasia in Barrett's esophagus. *Hum Pathol* 1988; 19: 166-78.
26. Steyerberg EW, Eijkemans MJ, Harrell FE, Jr.*, et al.* Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. *Stat Med* 2000; 19: 1059-79.
27. Nagelkerke NJD. A note on a general definition of the coefficient of determination. *Biometrika* 1991; 78: 691-92.
28. Harrell FE. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer, 2001: xxii, 568.
29. Steyerberg EW, Harrell FE, Jr., Borsboom GJ*, et al.* Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* 2001; 54: 774-81.
30. Dickman R, Green C, Chey WD*, et al.* Clinical predictors of Barrett's esophagus length. *Gastrointest Endosc* 2005; 62: 675-81.
31. de Jonge PJ, Steyerberg EW, Kuipers EJ*, et al.* Risk factors for the development of esophageal adenocarcinoma in Barrett's esophagus. *Am J Gastroenterol* 2006; 101: 1421-9.
32. Stein DJ, El-Serag HB, Kuczynski J*, et al.* The association of body mass index with Barrett's oesophagus. *Aliment Pharmacol Ther* 2005; 22: 1005-10.
33. Lagergren J, Bergstrom R, Nyren O. Association between body mass and adenocarcinoma of the esophagus and gastric cardia. *Ann Intern Med* 1999; 130: 883-90.
34. Chow WH, Blot WJ, Vaughan TL*, et al.* Body mass index and risk of adenocarcinomas of the esophagus and gastric cardia. *J Natl Cancer Inst* 1998; 90: 150-5.
35. Bu X, Ma Y, Der R*, et al.* Body Mass Index is Associated with Barrett Esophagus and Cardiac Mucosal Metaplasia. *Dig Dis Sci* 2006; 51: 1589-94.
36. Provenzale D, Schmitt C, Wong JB. Barrett's esophagus: a new look at surveillance based on emerging estimates of cancer risk. *Am J Gastroenterol* 1999; 94: 2043-53.
37. Shaheen NJ, Provenzale D, Sandler RS. Upper endoscopy as a screening and surveillance tool in esophageal adenocarcinoma: a review of the evidence. *Am J Gastroenterol* 2002; 97: 1319-27.
38. Rudolph RE, Vaughan TL, Storer BE*, et al.* Effect of segment length on risk for neoplastic progression in patients with Barrett esophagus. *Ann Intern Med* 2000; 132: 612-20.
39. Guda NM, Partington S, Vakil N. Inter- and intra-observer variability in the measurement of length at endoscopy: Implications for the measurement of Barrett's esophagus. *Gastrointest Endosc* 2004; 59: 655-8.

4 Grading dysplasia in Barrett esophagus: substantial interobserver variation between general and gastrointestinal pathologists

Histopathology (in press)

Marjon Kerkhof [1]
Herman van Dekken [2]
Ewout W Steyerberg [3]
Gerrit A Meijer [4]
Andries H Mulder [5]
Adriaan de Bruïne [6]
Ann Driessen [6]
Fiebo J ten Kate [7]
Johannes G Kusters [1]
Ernst J Kuipers [1]
Peter D Siersema [1]
for the CYBAR study group

Departments of [1]Gastroenterology and Hepatology, [2]Pathology, and [3]Public Health, Erasmus MC - University Medical Center Rotterdam, Rotterdam, Departments of Pathology, [4]VU University Medical Center, Amsterdam, [5]Rijnstate Hospital, Arnhem, [6]Academic Medical Center, Maastricht, [7]Academic Medical Center, Amsterdam, the Netherlands.

ABSTRACT

Aims: To determine interobserver variation in grading of dysplasia in Barrett esophagus (BE) between non-expert general pathologists and expert gastrointestinal pathologists on the one hand and between expert pathologists on the other hand.

Methods and Results: In this prospective multicenter study, non-expert and expert pathologists graded biopsies of 920 patients with endoscopic BE, which were blindly reviewed by one member of a panel of expert pathologists (panel experts), and by a second panel expert in case of disagreement on dysplasia grade. Agreement between 2 of 3 pathologists was established as final diagnosis. Analysis was by kappa statistics (κ). Due to absence of intestinal metaplasia, 127/920(14%) patients were excluded. The interobserver agreement for dysplasia (no dysplasia (ND) versus indefinite for dysplasia/low-grade dysplasia (IND/LGD) versus high-grade dysplasia (HGD)/adenocarcinoma (EAC)) between non-experts and first panel experts and between initial experts and first panel experts was fair (κ=0.24 and κ=0.27, respectively), and substantial for differentiation of HGD/EAC from ND/IND/LGD (κ=0.62 and κ=0.58, respectively).

Conclusions: We observed considerable interobserver variability in the interpretation of ND or IND/LGD in BE between non-experts and experts, but also between expert pathologists. This implicates that less subjective markers are needed to determine the risk of developing EAC in BE.

INTRODUCTION

According to the guidelines the American College of Gastroenterology (ACG),[1] Barrett esophagus (BE) is characterized by the replacement of the normal stratified squamous epithelium of the distal esophagus by columnar epithelium with specialized intestinal metaplasia (IM),[2] which is characterized by the presence of goblet cells. BE is caused by chronic gastroesophageal reflux,[3] and predisposes to the development of esophageal adenocarcinoma (EAC). This development is a gradual process in which the accumulation of (epi)genetic changes causes disruption of important biological processes at the cellular level, which can ultimately cause these cells to behave as cancer cells, i.e., invade surrounding tissues and metastasize. The morphologic counterpart of these cell biologic changes is called dysplasia, which by convention is classified in different stages, i.e., low-grade dysplasia (LGD), high-grade dysplasia (HGD), and finally EAC.[4-7] Nowadays, histopathologic assessment is the standard read out to judge to what stage the neoplastic process has progressed in an individual patient, and based on this, to determine the interval of endoscopic surveillance in patients with BE. The aim of surveillance is to detect progression of dysplasia in an early, and therefore likely curable stage.[1]

In 1988, histologic criteria for grading dysplasia were established by a group of experts in gastrointestinal pathology,[8] and twelve years later the Vienna classification was made to resolve the discrepancies in nomenclature between Western and Japanese pathologists.[9] Various groups have demonstrated that the use of these criteria is still accompanied by considerable interobserver variability, although all these studies only included a relatively small number of patients, and most studies used percentage of concordance between observers in their evaluation.[8, 10-12] However, as a part of this observed agreement was only explained by chance, it is preferable to use Cohen κ values to account for agreement beyond chance. Montgomery et al.[13] published a study including two series of 125 slides and used kappa (κ) statistics for analyzing the histologic grade of dysplasia in BE. A moderate agreement (κ = 0.46) was found when clinically relevant diagnostic categories (ND;IND/LGD;HGD;EAC) were used. However, interobserver variation was only assessed between expert gastrointestinal (GI) pathologists. In daily practice general pathologists review the majority of BE biopsies. In this large prospective multicenter study we assessed interobserver variability in establishing the grade of dysplasia in BE and compared this between non-expert general pathologists and expert GI-pathologists, and between expert GI-pathologists. We performed this study in a large group of 920 patients with ≥ 2 cm columnar-lined epithelium of the distal esophagus.

MATERIALS AND METHODS

Case selection and histological material

This study was part of an ongoing, prospective clinical trial (CYBAR study) assessing the value of flow cytometry for individualizing the frequency of follow-up of upper GI endoscopy in patients with BE. From October 2003 to December 2004, 920 prevalent and incident consecutive patients with a columnar-lined segment of ≥ 2 cm in the distal esophagus were included in 15 Dutch hospitals. Four-quadrant biopsies were obtained at intervals of 2 cm, according to the current guidelines.[1] The

biopsies were fixed in formalin. The Medical Ethical Review Boards of all hospitals approved this study and written informed consent was obtained from all patients prior to endoscopy.

Histologic evaluation

Pathologists of all participating hospitals (see Appendix), including 9 non-expert general (non-expert) pathologists and 6 expert GI (expert) pathologists, assessed haematoxylin and eosin (H&E) stained slides of paraffin-embedded biopsies of patients included in their own center for the presence of IM and the grade of dysplasia. IM was defined as the presence of goblet cells in columnar epithelium of the esophagus. In line with the definition of BE according to the ACG guidelines,[1] patients with biopsies without IM were excluded from this study. This is unlike the guidelines of the British Society of Gastroenterology (BSG) in the United Kingdom, which state that IM is not required for the diagnosis BE.[14] Dysplasia was graded according to the Consensus Criteria of 1988, with adjustments made in 2001 (Table 1).[8, 13]

Subsequently, one member of a panel of five expert (panel expert) pathologists randomly reviewed the slides. These panel experts were blinded for age, sex, identity of the patient, and diagnosis of the initial pathologist. Furthermore, specific areas of interest were not marked on the slides. The diagnostic selections were IM or no IM, and, if IM was present, 'no dysplasia' (ND), 'indefinite for dysplasia' (IND)/LGD, HGD or EAC. It was recognized that in some slides it was difficult to

Table 1. Criteria for grading dysplasia in Barrett esophagus based on the Consensus Criteria of 1988 and 2001 [8, 12]

Negative for dysplasia	Architecture normal Surface maturation (nucleur-to-cytoplasmic ratio of surface cells is lower than of deeper glands) Cytology normal (nuclear polarity normal, which means size not different and located basally)
Indefenite for dysplasia	Architecture normal Surface maturation Cytology mild alterations (nuclear membrane irregularities, increased mitoses in deeper glands, inflammation)
Low-grade dysplasia	Architecture mildly alterated (glandular crowding, but identificable lamina propria) Surface maturation distorted (surface similar as deeper glands) Cytology mild, diffuse alterations (nuclear hyperchromasia, nuclear membrane irregularities, nuclear polarity normal)
High-grade dysplasia	Architecture marked alterated (crowding of cytological abnormal glands) Surface maturation lacking Cytology marked alterations (nuclear hyperchromasia, prominent irregular nucleoli with clumped chromatin, loss of nuclear polarity)
Intramucosal carcinoma	Architecture marked alterated (lamina propria effacing, penetration through basement membrane into lamina propria, syncytial growth pattern, extensive back-to-back microglands) Surface maturation lacking Cytology marked alterations

distinguish between the diagnoses 'indefinite for dysplasia' or LGD, and for practical purposes we combined these two diagnoses into a diagnosis of IND/LGD. All participating pathologists agreed to commit themselves to one of these diagnostic choices. If there was disagreement on the presence of IM or grade of dysplasia, slides were blindly reviewed by a second member of the expert panel. A final diagnosis was established by a majority diagnosis, meaning that 2 of 3 pathologists had to agree upon the diagnosis. In a few cases, in which disagreement on the grade of dysplasia was present after three opinions, a third member of the expert panel reviewed the slides, after which a final majority diagnosis could be established in all cases. A meeting was held a few months after the start of the study, in which the histological criteria were re-emphasized.

Statistical analysis

Interobserver agreement was determined by using Cohen kappa (κ) statistics, which are widely used mathematical coefficients adjusting for agreement by chance alone.[15] A value of zero indicates no agreement better than that which would be expected by chance alone. Values of < 0.21, 0.21-0.40, 0.41-0.60, 0.61-0.80 and >0.80 correspond with a poor, fair, moderate, substantial, and very good interobserver agreement, respectively.[15] The histological diagnoses were categorized as ND, IND/LGD, and HGD/EAC for the main analyses (3 categories). Further categorizations included ND, IND/LGD, HGD, and EAC (4 categories), ND/IND/LGD, and HGD/EAC (2 categories), and ND, and IND/LGD/HGD/EAC (2 categories). In some analyses, κ values could not be evaluated, since these values can only be computed if a symmetric 2-way table is present, which means that for example the second (expert) pathologist had to establish at least one time the same diagnosis as the initial (non-expert or expert) pathologist. Statistic analyses were conducted using SPSS software (SPSS version 11.0, Chicago, Illinois, USA).

RESULTS

In total, 920 patients, of whom 662 were men and 258 women, were included with a mean (± standard error of the mean (SEM)) age of 60 ± 0.4 years (range 19-88 years). Mean (± SEM) length of the columnar-lined segment was 4.3 ± 0.1 cm (range 2-16 cm), and the median was 4.0 cm.

In 127/920 (14%) patients, no IM was diagnosed. In 23 of these 127 (18%) patients, a false positive diagnosis of IM was made by the initial non-expert pathologist, which was corrected to no IM by 2 panel expert pathologists. Since in many countries the presence of IM is required for the diagnosis BE[1], these 127 patients were excluded from further analysis.

Of the remaining 793 patients with histologically confirmed BE, an initial diagnosis on the grade of dysplasia was made in the center where the patient was included, and was performed in 698 cases by a non-expert pathologist, and in 95 cases by an expert pathologist.

Grading dysplasia by the initial pathologists

The diagnoses of the initial pathologists (both non-experts and experts) were in 567/793 (72%) cases ND, in 210 (26%) IND/LGD, and in 16 (2%) HGD/EAC. These initial diagnoses and also the final diagnoses are shown in Table 2. If 3 clinically relevant diagnostic categories for statistical

Table 2. Initial diagnoses and final diagnoses of the grade of dysplasia in 793 patients with Barrett esophagus after reviewing of the slides by the expert panel

	Final diagnosis			
Initial diagnosis	*ND*	*IND/LGD*	*HGD/EAC*	*Totals*
ND	546	19	2	*567*
IND/LGD	111	94	5	*210*
HGD/EAC	0	1	15	*16*
Totals	*657*	*114*	*22*	*793*

κ value = 0.25
ND: no dysplasia, IND: indefinite for dysplasia, LGD: low-grade dysplasia, HGD: high-grade dysplasia, EAC: esophageal adenocarcinoma

analysis were used (ND, IND/LGD and HGD/EAC), the interobserver reproducibility between the initial pathologists (non-experts and experts) and second pathologists (first panel experts) for dysplasia was fair (κ=0.25, 95% confidence interval (CI):0.18-0.32, p=0.00) (Table 3). When HGD and EAC were analyzed as 4 separate categories (ND; IND/LGD; HGD; EAC), the interobserver reproducibility was similar. Since only the presence of HGD and EAC in BE has important therapeutic consequences for patients, we also analyzed the interobserver agreement when 2 broad diagnostic categories (ND/IND/LGD and HGD/EAC) were evaluated. The κ value for interobserver agreement between the initial and first panel expert pathologist when employing these 2 categories was substantial (κ=0.61). We also evaluated the interobserver reproducibility between no dysplasia and dysplasia (ND and IND/LGD/HGD/EAC), which was just as fair (0.27) as using 3 or 4 categories.

Initial diagnosis of non-expert pathologists

The diagnoses in patients with histological BE assessed by non-expert pathologists were in 500/698 (71%) cases ND, in 187 (27%) IND/LGD, and in 11 (2%) HGD/EAC.

In 502/698 (72%) patients, there was concordance on the dysplasia grade between the non-expert and the first panel expert pathologist. In 196/698 (28%) patients, a third opinion was needed because of disagreement between non-expert and first panel expert pathologists. In 98/196 (50%) cases, the initial (non-expert) diagnosis was downgraded, mostly from IND/LGD to ND, in 22 (11%) it was upgraded, and in 76 (39%) the final diagnosis was similar to the initial one. The interobserver reproducibility between non-expert and first panel expert pathologists for dysplasia was fair if analyzed in 3 categories as described above (κ=0.24) (Table 3). This interobserver reproducibility remained fair when analyzing in 4 categories and in 2 categories (no dysplasia vs. dysplasia), but was again substantial better when analyzing the categories HGD/EAC versus ND/IND/LGD (κ=0.62).

Initial diagnosis of expert pathologists

The diagnoses of patients with histological BE initially assessed by expert pathologists, were in 67/95 (71%) cases ND, in 23 (24%) IND/LGD, and in 5 (5%) HGD/EAC.

The expert pathologist establishing an initial diagnosis and the first panel expert pathologist agreed on the dysplasia grade in 66/95 (69%) patients. In 29 (31%) patients, a third opinion was needed because of disagreement between both expert pathologists. Comparable to the results of the non-expert pathologists, the initial diagnosis was downgraded in 12/29 (41%) cases, again mostly from IND/LGD to ND. In 4 (14%) cases it was upgraded, and in 13 (45%) the final diagnosis was similar to the initial one. The interobserver reproducibility for dysplasia between both expert pathologists (initial judgment and first revision), analyzed in 3 categories as well in 4 categories, was fair (κ=0.27) (Table 3). When the interobserver agreement in 2 categories (no dysplasia or presence of dysplasia) was analyzed, the κ value was somewhat better, although still fair (k=0.33). The κ value for interobserver reproducibility of the categories ND/IND/LGD versus HGD/EAC was moderate (κ=0.58).

Agreement between panel expert pathologists

In 225 of 793 cases the initial pathologist (both non-experts and experts) and first panel expert pathologist disagreed, and consequently a second panel expert had to give a third opinion. In 103/225 (46%) of these cases there was concordance on the grade of dysplasia between the two panel expert pathologists, which corresponds with a poor interobserver reproducibility between the first and second panel expert pathologist (κ=0.12), when using the 3 categories (Table 3). When analyzing the interobserver agreement between no dysplasia and the presence of any dysplasia, the κ value remained poor (0.16). Using 2 categories, in which HGD/EAC was separated from ND/IND/LGD, the κ value was moderate (0.41). The κ value of 4 dysplasia categories could not be calculated between the two panel expert pathologists, due to methodological limitations of the test.

Table 3. Calculated κ values (95% CI) to establish interobserver variation of dysplasia grading in Barrett esophagus for different categories of dysplasia

	3 categories:	**4 categories:**	**2 categories:**	**2 categories:**
	1. ND	*1. ND*	*1. ND*	*1. ND/IND/LGD*
	2. IND/LGD	*2. IND/LGD*	*2. IND/LGD/HGD/ EAC*	*2. HGD/EAC*
	3. HGD/EAC	*3. HGD*		
		4. EAC		
first vs. second path.				
total group	0.25 (0.18-0.32)	0.25 (0.18-0.32)	0.27 (0.20-0.34)	0.61 (0.44-0.78)
first opinion non-expert	0.24 (0.16-0.32)	0.25 (0.17-0.33)	0.26 (0.18-0.34)	0.62 (0.42-0.82)
first opinion expert	0.27 (0.08-0.46)	0.27 (0.08-0.46)	0.33 (-0.32-0.98)	0.58 (0.21-0.95)
Second vs. third path.	0.12 (0.00-0.24)	#	0.16 (0.03-0.29)	0.41 (0.17-0.65)

p<0.05 in all categories and groups

#: κ value could not be calculated

ND: no dysplasia, IND: indefinite for dysplasia, LGD: low-grade dysplasia, HGD: high-grade dysplasia, EAC: esophageal adenocarcinoma

Finally, in 27 cases a final diagnosis could not be established after the third opinion, as there was still no majority diagnosis on the grade of dysplasia. In these cases, a third panel expert pathologist blindly reviewed the slides. After four opinions, a final diagnosis was established in all cases.

DISCUSSION

Histopathologic evaluation of BE biopsies is currently the only diagnostic tool that is used to determine the risk of progression to HGD or EAC. In line with previous studies, our study shows a high interobserver variability in interpretation of the grade of dysplasia in BE, which was calculated by Cohen κ values which corrects for agreement by chance alone. Remarkably, the interobserver variability was not different between non-expert and expert pathologists and between expert pathologists. This implicates that the subjective component in the evaluation of dysplasia in BE can hardly be improved by the initial reviewing of BE biopsies by expert pathologists.

Our study also showed a poor agreement between two panel expert pathologists. It should be pointed out however that a second panel expert pathologist was only involved in cases where disagreement was present between the initial and the first panel expert. It may well be that these slides were more complex, which is demonstrated by the 27 cases in which a final diagnosis could not be established after the third opinion, as there was still no majority diagnosis on the grade of dysplasia.

The interobserver variability was mainly based on disagreement between the presence of either ND or IND/LGD. The histologic grade of dysplasia determines the frequency of surveillance with upper endoscopy, corresponding with the risk of progression (yearly for LGD and every 3 years for ND).[1, 16, 17] Misclassification may lead to unnecessary, burdensome follow-up endoscopies[18] (in case of overdiagnosis), or to possibly insufficient control (in case of underdiagnosis).

In contrast to the fair interobserver agreement in establishing a distinction between ND and IND/LGD, interobserver reproducibility was substantial, although still not perfect, in distinguishing HGD/EAC from ND/IND/LGD. This is an important finding which is in line with previous results,[8, 13] as mainly the presence of HGD or EAC in BE has at least short term clinical consequences for the treatment and prognosis of patients.[19] Depending on the condition of the patient and the extent of HGD or early EAC, an individualized strategy is employed, which implicates endoscopic surveillance every three months, endoscopic mucosectomy or a surgical resection.[1, 19]

For practical purposes, we did not make separate categories for biopsies with either IND or LGD. It seems unlikely that this has influenced the results of our study. Montgomery et al.[13] found a similar result for interobserver variability if IND was separated from LGD compared to the situation where both groups were combined (κ=0,43 vs. κ=0,48 respectively).

The most important explanation for histology disagreement in grading dysplasia in BE is that the system of grading morphological progression of dysplasia in discrete categories (ND, IND/LGD, HGD) actually means an artificial segmentation of a morphologic continuum.[5] It seems apparent that cases with morphology close to one of these artificial boundaries are more difficult to agree upon. In a previous study,[13] it was concluded that using formalin-fixative caused less disagreement

than Hollande-fixed specimens. We excluded this cause of interobserver disagreement by using only formalin fixation of the specimens in all participating centers.

Currently, there is still no reliable method to improve the problem of interobserver variability in diagnosing dysplasia in BE.[20] Montgomery et al.[13] systematically investigated the effect of a consensus meeting. It was shown that hardly any improvement of interobserver variability was established after the consensus meeting. A strategy to overcome difficulties with subjective grading is using morphometry, as has been performed by Polkowski et al.[21] However, this approach required dedicated equipment and expertise, which is not universally available. Furthermore, morphometry has difficulties to compensate for technical issues as tangential cutting or severe inflammation.[10] Another study [22] suggested that p53 protein quantification may improve interobserver variability in assessing the presence of LGD, however this was established in biopsies that were already histologically evaluated by an expert panel of pathologists, who agreed upon the level of dysplasia. In addition, enhanced magnification endoscopy has been evaluated to increase the likelihood that indeed biopsies with a high *a priori* risk of containing dysplasia in BE are present. It has been reported however that interobserver agreement remained poor using this method ($\kappa<0.4$).[23]

As there is currently no solution for decreasing the interobserver variability in histological grading of BE biopsies, the focus should be targeted on techniques that are additional to histology and less prone for interobserver variability, to identify those patients with ND or LGD who will most likely progress towards malignancy and really need frequent surveillance. This seems important, since approximately 75-90 % of patients with LGD will probably never develop EAC during their lifetime.[17, 24, 25] Several studies have been performed to investigate the value of such additional biomarkers. Firstly, in addition to the above described study, in which p53 quantification was evaluated for decreasing interobserver variability, it has also been proposed to determine the presence of p53 mutations as an adjunct to histology, to establish a more reliable cancer risk. Patients with LGD and p53 expression in their biopsies seem more likely to progress towards HGD/EAC (56-60%) than patients with LGD without p53 expression (0-25%)[26-29] Secondly, flow cytometric analysis of DNA content has been reported to be promising in determining the risk of progression towards malignancy.[30-32] Reid et al.[31] demonstrated in a population of 322 BE patients, that patients with ND, IND or LGD in their biopsies, and an aneuploid or tetraploid nuclear DNA content, had a 28% 5-yr cumulative risk of EAC development compared to 0% for those with normal flow cytometric results. Large prospective studies need to confirm these results. Finally, new endoscopy techniques have been developed to visualize early dysplasia, such as autofluorescence endoscopy, high resolution endoscopy or magnification chromendoscopy, to make targeted biopsies possible. Although these techniques seem promising for the visualization of areas with HGD in the columnar lined segment, LGD is still difficult to visualize.[33-35]

In conclusion, this study demonstrates a high interobserver variability in the interpretation of dysplasia in a large cohort of BE patients between non-expert and expert GI pathologists, but also among experts together. Since morphology is still the best available marker to determine cancer risk, we recommend reviewing BE slides by at least two pathologists, and, when indicated, to consult a third (expert) pathologist to establish a final diagnosis. In the meantime studies are

required which evaluate additional, less subjective markers to establish cancer risk in patients with BE, in order to determine the optimal surveillance frequency.

REFERENCES

1. Sampliner RE. Updated guidelines for the diagnosis, surveillance, and therapy of Barrett's esophagus. *Am J Gastroenterol* 2002; 97: 1888-95.
2. Guillem PG. How to make a Barrett esophagus: pathophysiology of columnar metaplasia of the esophagus. *Dig Dis Sci* 2005; 50: 415-24.
3. Avidan B, Sonnenberg A, Schnell TG*, et al.* Hiatal hernia size, Barrett's length, and severity of acid reflux are all risk factors for esophageal adenocarcinoma. *Am J Gastroenterol* 2002; 97: 1930-6.
4. Hamilton SR, Smith RR. The relationship between columnar epithelial dysplasia and invasive adenocarcinoma arising in Barrett's esophagus. *Am J Clin Pathol* 1987; 87: 301-12.
5. Haggitt RC. Barrett's esophagus, dysplasia, and adenocarcinoma. *Hum Pathol* 1994; 25: 982-93.
6. Hameeteman W, Tytgat GN, Houthoff HJ*, et al.* Barrett's esophagus: development of dysplasia and adenocarcinoma. *Gastroenterology* 1989; 96: 1249-56.
7. Buttar NS, Wang KK. Mechanisms of disease: Carcinogenesis in Barrett's esophagus. *Nat Clin Pract Gastroenterol Hepatol* 2004; 1: 106-12.
8. Reid BJ, Haggitt RC, Rubin CE*, et al.* Observer variation in the diagnosis of dysplasia in Barrett's esophagus. *Hum Pathol* 1988; 19: 166-78.
9. Schlemper RJ, Riddell RH, Kato Y*, et al.* The Vienna classification of gastrointestinal epithelial neoplasia. *Gut* 2000; 47: 251-5.
10. Baak JP, ten Kate FJ, Offerhaus GJ*, et al.* Routine morphometrical analysis can improve reproducibility of dysplasia grade in Barrett's oesophagus surveillance biopsies. *J Clin Pathol* 2002; 55: 910-6.
11. Alikhan M, Rex D, Khan A*, et al.* Variable pathologic interpretation of columnar lined esophagus by general pathologists in community practice. *Gastrointest Endosc* 1999; 50: 23-6.
12. Sagan C, Flejou JF, Diebold MD*, et al.* [Reproducibility of histological criteria of dysplasia in Barrett mucosa]. *Gastroenterol Clin Biol* 1994; 18: D31-4.
13. Montgomery E, Bronner MP, Goldblum JR*, et al.* Reproducibility of the diagnosis of dysplasia in Barrett esophagus: a reaffirmation. *Hum Pathol* 2001; 32: 368-78.
14. Playford RJ. New British Society of Gastroenterology (BSG) guidelines for the diagnosis and management of Barrett's oesophagus. *Gut* 2006; 55: 442.
15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159-74.
16. Mashimo H, Wagh MS, Goyal RK. Surveillance and screening for Barrett esophagus and adenocarcinoma. *J Clin Gastroenterol* 2005; 39: S33-41.
17. Hage M, Siersema PD, van Dekken H*, et al.* Oesophageal cancer incidence and mortality in patients with long-segment Barrett's oesophagus after a mean follow-up of 12.7 years. *Scand J Gastroenterol* 2004; 39: 1175-9.
18. Kruijshaar ME, Kerkhof M, Siersema PD*, et al.* The burden of upper gastrointestinal endoscopy in patients with Barrett's esophagus. Endoscopy (in press). In.
19. Stein HJ, von Rahden BH, Feith M. Surgery for early stage esophageal adenocarcinoma. *J Surg Oncol* 2005; 92: 210-7.
20. Montgomery E. Is there a way for pathologists to decrease interobserver variability in the diagnosis of dysplasia? *Arch Pathol Lab Med* 2005; 129: 174-6.
21. Polkowski W, Baak JP, van Lanschot JJ*, et al.* Clinical decision making in Barrett's oesophagus can be supported by computerized immunoquantitation and morphometry of features associated with proliferation and differentiation. *J Pathol* 1998; 184: 161-8.
22. Lorinc E, Jakobsson B, Landberg G*, et al.* Ki67 and p53 immunohistochemistry reduces interobserver

variation in assessment of Barrett's oesophagus. *Histopathology* 2005; 46: 642-8.
23. Mayinger B, Oezturk Y, Stolte M*, et al.* Evaluation of sensitivity and inter- and intra-observer variability in the detection of intestinal metaplasia and dysplasia in Barrett's esophagus with enhanced magnification endoscopy. *Scand J Gastroenterol* 2006; 41: 349-56.
24. Falk GW. Barrett's esophagus-is it bad for your health? *Am J Gastroenterol* 2005; 100: 2622-3.
25. Conio M, Blanchi S, Lapertosa G*, et al.* Long-term endoscopic surveillance of patients with Barrett's esophagus. Incidence of dysplasia and adenocarcinoma: a prospective study. *Am J Gastroenterol* 2003; 98: 1931-9.
26. Weston AP, Banerjee SK, Sharma P*, et al.* p53 protein overexpression in low grade dysplasia (LGD) in Barrett's esophagus: immunohistochemical marker predictive of progression. *Am J Gastroenterol* 2001; 96: 1355-62.
27. 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.
28. Skacel M, Petras RE, Rybicki LA*, et al*. p53 expression in low grade dysplasia in Barrett's esophagus: correlation with interobserver agreement and disease progression. *Am J Gastroenterol* 2002; 97: 2508-13.
29. Younes M, Lebovitz RM, Lechago LV*, et al.* p53 protein accumulation in Barrett's metaplasia, dysplasia, and carcinoma: a follow-up study. *Gastroenterology* 1993; 105: 1637-42.
30. Reid BJ, Blount PL, Rubin CE*, et al.* Flow-cytometric and histological progression to malignancy in Barrett's esophagus: prospective endoscopic surveillance of a cohort. *Gastroenterology* 1992; 102: 1212-9.
31. Reid BJ, Levine DS, Longton G*, et al.* Predictors of progression to cancer in Barrett's esophagus: baseline histology and flow cytometry identify low- and high-risk patient subsets. *Am J Gastroenterol* 2000; 95: 1669-76.
32. Menke-Pluymers MB, Mulder AH, Hop WC*, et al.* Dysplasia and aneuploidy as markers of malignant degeneration in Barrett's oesophagus. The Rotterdam Oesophageal Tumour Study Group. *Gut* 1994; 35: 1348-51.
33. Bruno MJ. Magnification endoscopy, high resolution endoscopy, and chromoscopy; towards a better optical diagnosis. *Gut* 2003; 52 Suppl 4: iv7-11.
34. Sharma P, Weston AP, Topalovski M*, et al.* Magnification chromoendoscopy for the detection of intestinal metaplasia and dysplasia in Barrett's oesophagus. *Gut* 2003; 52: 24-7.
35. Kara MA, Peters FP, Ten Kate FJ*, et al.* Endoscopic video autofluorescence imaging may improve the detection of early neoplasia in patients with Barrett's esophagus. *Gastrointest Endosc* 2005; 61: 679-85.

5

Biomarkers for risk stratification of neoplastic progression in Barrett esophagus

Submitted

Marjon Kerkhof [1]
Johannes G Kusters [1]
Herman van Dekken [2]
Ernst J Kuipers [1]
Peter D Siersema [1]

Departments of [1]Gastroenterology and Hepatology and [2]Pathology,
Erasmus MC - University Medical Center Rotterdam, Rotterdam, the Netherlands

ABSTRACT

Barrett esophagus (BE) is caused by chronic gastroesophageal reflux and predisposes to the development of esophageal adenocarcinoma through different histological grades of dysplasia. Only a subset of BE patients will finally develop esophageal adenocarcinoma. The majority will therefore not benefit from an endoscopic surveillance program, based on the histological identification of dysplasia. Several studies have been performed to find biomarkers that can be used to detect the subgroup of patients at an increased risk of developing malignancy in BE. In this review, we will summarize the most promising tissue biomarkers, i.e. proliferation/cell cycle proteins, tumor suppressor genes, adhesion molecules, DNA ploidy status and inflammation associated markers, that can be used for risk stratification in BE, and discuss their respective clinical application.

INTRODUCTION

Barrett esophagus (BE) is characterized by the replacement of the normal stratified squamous epithelium of the distal esophagus by columnar epithelium with specialized intestinal metaplasia (IM),[1] which is characterized by the presence of goblet cells. Chronic gastroesophageal reflux is the most important factor in the development of BE.[2] BE is a premalignant condition predisposing to the development of esophageal adenocarcinoma (EAC). This development is a gradual process in which the accumulation of (epi)genetic changes causes disruption of important biological processes at the cellular level, which can ultimately cause these cells to behave as cancer cells, i.e., invading surrounding tissues and metastasize. The morphologic counterpart of these molecular changes is called dysplasia. Dysplasia is commonly sub-classified into three distinct morphological stages, each thought to represent a subsequent step in tumor progression, i.e., low-grade dysplasia (LGD), high-grade dysplasia (HGD), and finally EAC.[3, 4]

Nowadays, histopathologic assessment is the standard read out to assess whether and to what stage neoplasia in BE has progressed in an individual patient, and based on this, to determine the interval of endoscopic surveillance in these patients. The aim of surveillance is to detect progression of dysplasia at an early, and therefore likely curable stage.[1]

Although EAC is frequently accompanied by Barrett's metaplasia, only approximately 5% of patients who present with EAC are known with a prior diagnosis of BE.[5, 6] Moreover, the risk of developing EAC in BE is low and has been suggested to be approximately 0.5% on a yearly basis.[7-9] This means that the majority of patients with BE will not benefit from an endoscopic surveillance program.[7-9] Further stratification of the risk of progression of BE to EAC might permit more effective targeting of repeated endoscopy to patients with an increased risk of progression.

At present, patients with BE are only risk stratified by the grade of dysplasia as assessed by histological evaluation of endoscopically taken biospies.[10] In 1988, histologic criteria for grading dysplasia were established by a group of experts in gastrointestinal pathology.[11] Histological grading according to these criteria is however accompanied by considerable interobserver variability, especially for the discrimination between no dysplasia (ND) and LGD.[12] Considerable effort went in the identification of a biomarker that could distinguish patients with high risk for EAC development from those with low risk for EAC. A biomarker can be defined as an indicator of pathological processes. The ideal biomarker for this would probably be a molecule that shows a variation in expression that is associated with neoplastic progression and is already detectable at an early stage in this process.[13] In this review, the most promising tissue biomarkers known so far will be discussed.

POTENTIAL BIOMARKERS FOR RISK STRATIFICATION

The transformation from a normal cell into a tumor cell requires several alterations, each of them leading to the induction of proteins involved in tumorigenesis or downregulation of proteins protecting the cell.[14] These alterations comprise usually genetic lesions or altered methylation patterns of genes, resulting in changes in mRNA and protein expression. The molecules involved

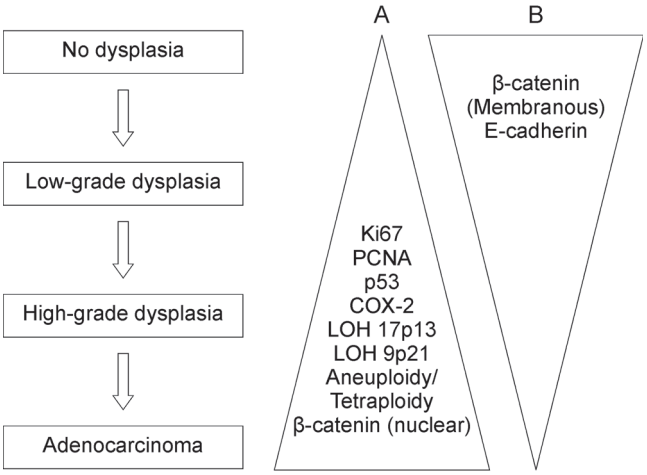


Figure 1. Schematic overview of the expression of the discussed biomarkers in the progression from Barrett's metaplasia towards esophageal adenocarcinoma. Biomarkers are grouped for those with an increased (A) or decreased (B) expression in the metaplasia-dysplasia-carcinoma sequence.

in these processes may therefore provide markers for the detection of early malignant progression. Based on the molecular alterations these markers can be divided in different groups, which will consecutively be described in this review: proliferation/cell cycle proteins, tumor suppressor genes, adhesion molecules, DNA content, and inflammation associated markers. In Figure 1, the pattern of expression of these biomarkers is shown in a schematic overview.

Proliferation/cell cycle proteins

Tissue damage by gastroesophageal reflux will lead to proliferation in order to replace the injured cells by new ones. In order to proliferate, a cell has to progress from the G1 to the S phase in the cell cycle (Figure 2). Progression to a next stage in the cell cycle requires the action of cyclin-Cdk (cyclin-dependent kinase)-complexes. When this proliferation runs out of control, neoplastic lesions will occur. Abnormalities of proteins that play a role in the progression from the G1 to the S phase can be observed during carcinogenesis. These proteins (i.e. PCNA, Ki67, and Cyclin D1) could therefore possibly serve as biomarker in predicting the risk of neoplastic progression.

PCNA

Proliferating cell nuclear antigen (PCNA) is a cofactor of DNA synthase and an indicator of cell cycle progression at the G1/S transition (Figure 2).[15] PCNA was the first proliferation marker that could be used for immunohistochemical staining of formalin-fixed paraffin tissue. As a consequence most of the initial proliferation marker work focused on PCNA, simply because there were no practical alternatives.[16] Several studies have shown that PCNA staining is increased in HGD/EAC, with an increase in intensity of PCNA expression with extension of the proliferative compartment upwards to the superficial layers of the glands as is seen in dysplasia.[17-19] This was not confirmed in another study, in which PCNA was found to be of limited value in differentiating between ND, LGD and

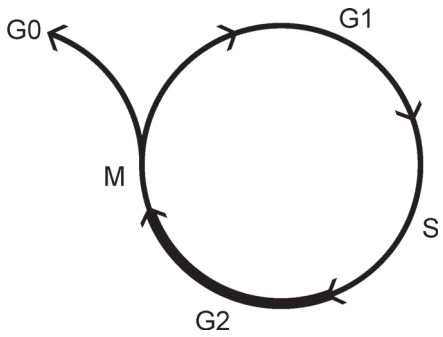


Figure 2. Cell cycle. G1 = gap 1, cells in resting phase (DNA = 2N); S = DNA synthesis; G2 = gap 2, cells are duplicated (DNA = 4N); M = mitosis, cells are divided in 2 daughter cells (DNA = 2N); G0 = resting phase, cells that cease division.

'indefinite for dysplasia' (IND) in BE.[20] In addition, PCNA staining is affected by the fixation method of the tissue, and antigen retrieval does result in staining of quiescent cells (G0 phase) (Figure 2).[16] Therefore, PCNA is probably not a good candidate marker that can be used for the prediction of patients at risk of neoplastic progression in BE.

Ki67

The human Ki67 protein is present during all active phases of the cell cycle (G1, S, G2, M), but is absent in resting cells (G0) (Figure 2). Although some of its features have been characterized, such as phosphorylation and nuclear transport, the function of the Ki67 protein is still largely unknown.[16] Expression of the Ki67 protein is strictly associated with cell proliferation. The fraction of Ki67 positive cells been demonstrated to correlate with the clinical course of the disorder.[16] No other protein has been shown to have an expression pattern that is so closely associated with the proliferative status of the cell. With the development of the Ki67 equivalent MIB-1, Ki67 immunostaining can, just like PCNA, be easily performed on formalin-fixed paraffin-embedded tissue. In contrast to PCNA (see above), the Ki67-antibody does not stain quiescent cells, thus Ki67 is a superior proliferation marker (Figure 3, page 65).[21]

The extent of immunohistochemical Ki67 expression is associated with each histological grade, showing a stepwise increase in Ki67 expression with neoplastic progression of BE.[22] In a study by Hong et al., statistical differences in expression levels between no dysplasia (ND), LGD and HGD were found. The category IND however had a great variety in expression pattern, sometimes even resembling HGD. These authors concluded therefore that Ki67 better can be used as an additional parameter to differentiate between BE patients with or without dysplasia.[23] In contrast, Olvera et al. concluded that Ki67 was able to differentiate LGD from HGD, but could not distinguish LGD from reactive changes (IND). The number of cases in this study was however probably too small (n= 25) for this conclusion to be made.[24] Currently, only cross-sectional studies on Ki67 expression in BE have been performed and longitudinal follow-up studies for evaluating the value of Ki67 as biomarker for risk prediction are therefore indicated. In a study by Polkowski et al.,[25] using

morphometry with assessment of the percentage of nuclei positive for Ki67 per 100 counted nuclei, it was shown that Ki67 was a valuable marker to overcome difficulties with subjective grading.[26]

Most studies on Ki67 expression in Barrett epithelium have been performed with immunohistochemistry. Detection of Ki67-positive cells in Barrett biopsies can also be performed with flow cytometry, making rapid quantification possible.[27] In this study, fresh frozen biopsies were required for flow cytometric evaluation, while immunohistochemistry can be performed on more easily available paraffin-embedded biopsies. In contrast to immunohistochemistry, the identity of Ki67-positive cells cannot be determined with flow cytometry. The usefulness of flow cytometry for Ki67 lies in the possibility to distinguish Ki67-positive G1 cells from quiescent G0 cells, which is important if combined with evaluation of the ploidy status (see further in 'DNA ploidy').

Cyclin D1

The Cyclin D1 gene is known to regulate the G1/S-checkpoint in the normal cell cycle (Figure 2), and may therefore play a role in carcinogenesis.[28] The role of cyclin D1 in cell cycle control is mediated through cyclin D1-cyclin-dependent kinase (cdk) complexes.[29] In a prospective study by Bani-Hani et al., immunohistochemically detected cyclin D1 was found to be significantly overexpressed in 92% of samples with EAC. In addition, in 67% of biopsies of these patients taken at earlier time points cyclin D1 overexpression was present, compared to in 29% of biopsies of controls without malignant progression in BE. Based on these results, it was suggested that cyclin D1-staining could be a useful biomarker in identifying BE patients at increased risk of neoplastic progression.[29] These results are however contradictory to more recent studies, in which cyclin D1 was not significantly associated with risk of malignant progression.[30, 31] Additional studies are clearly warranted.

Geddert et al. found that *cyclin D1* polymorphisms in patients with EAC were not significantly different from those of healthy controls, and therefore were unlikely to be associated with an increased risk of EAC.[28] In contrast, Casson et al found that the CCND1 A/A genotype was associated with an increased risk of developing BE and EAC, however no association was found between this genotype and cyclin D1 overexpression.[32]

Tumor suppressor genes

Tumor suppressor genes control cell proliferation by preventing cells from uncontrolled expanding. Proteins that activate the tumor suppressor gene behave as tumor suppressors. In a mutated tumor suppressor gene, the function may be lost due to inactivation, and consequently the protein has become an oncogene, leading to uncontrolled growth of mutated cells, and finally to malignancy. It has been suggested that mutated tumor suppressor genes may have the ability to predict neoplastic progression. In BE, particularly the role of p53 en p16 has been explored.

p53

p53 is a tumor suppressor gene, located on the 17p13 chromosome. The gene is involved in controlling cell proliferation.[33] Normally, cells contain low levels of wild-type p53. Wild-type p53 regulates two common responses to oncogenic stress, i.e., cell cycle arrest/DNA repair and apoptosis. In cells that are early in the G1-phase, p53 triggers a checkpoint blocking further progression through the

cell cycle, allowing the damaged DNA to be repaired before the cell enters the S-phase (Figure 2).[34] If the DNA damage cannot be repaired, p53 induces apoptosis.[35] This suggests that failure of p53 to respond to DNA damage will increase the susceptibility to mutational, oncogenic changes. Mutated p53 is dominant negative, as it overwhelms the wild-type protein and prevents it from functioning.[34] These p53 mutations are associated with an increased half-life of the p53 protein, leading to its accumulation in the cell nucleus to levels that can be detected by immunohistochemistry (Figure 3, page 65).[36] In contrast, wild-type p53 has a short half-life, and as a consequence these proteins do not accumulate and are therefore usually below the detection threshold of immunohistochemistry.[37] Approximately 90% of the p53 mutations are point mutations.[38]

As a consequence of DNA damage, the percentage of cells in G0/G1 or G2/M-phase that require DNA repair is increased.[34] This can be accompanied by p53 mutation and protein accumulation.[34] Several studies have shown a stepwise overexpression of p53 with increasing grades of dysplasia in BE.[22, 39-41] Younes et al. suggested that p53 accumulation might even occur before the phenotypic changes characteristic of dysplasia and malignancy become obvious, since normal-appearing nondysplastic glands adjacent to dysplastic glands or carcinoma were also positive for p53.[42] p53 as a biomarker of malignant progression in BE was confirmed in other studies, but the sensitivity of this marker alone in these studies was too low to predict cancer risk.[29, 30] Also when combined with other biomarkers, such as cyclin D1, β-catenin, and COX-2, p53 was found to be of limited value.[30]

Although immunohistochemistry for detecting p53 is cheap, quick, and easy to apply compared with other techniques, there are some limitations that are important to consider. The p53 antibodies that are commonly used do not only stain the mutant p53, but also detect wild-type p53. Thus, overexpression of the p53 protein does not correlate with p53 mutation *per se*.[37, 38] A second limitation of p53-based immunohistochemistry is that mutations for this tumor suppressor gene may exist without protein overexpression, and therefore will not be detected by immunohistochemistry.[37, 38, 43]

Another mechanism of inactivation of the wild-type p53 is loss of heterozygosity (LOH) for one or two alleles of the 17p13 gene.[37] LOH has been shown to occur in 0-6% of BE patients with ND, in 20-27% with LGD, in 57% with HGD,[44, 45] and in 54-92% with EAC,[45, 46] and sometimes coexists with a p53 mutation.[45, 46] It has been shown that clones of 17p13 LOH show variable expansion within the Barrett segment,[47] and a larger size of the LOH clone seems to be associated with a higher risk of progression to EAC.[48] A strong association has also been found between 17p13 LOH and an abnormal flow cytometric DNA content in BE.[47, 49] In 91% of flow cytometrically detected aneuploid/tetraploid cases, LOH at 17p13 was also detected, in contrast to only 17% of diploid cases.[47] In another study by the same group, LOH at 17p13 was found in 91% of diploid cases, in which aneuploidy developed during follow-up. Thus LOH preceded the development of aneuploidy during neoplastic progression in BE.[49] Recently, these investigators showed in a prospectively followed cohort that 37% of patients with LOH at 17p13 progressed over time from ND to EAC, compared to 3% of patients without LOH, suggesting that 17p13 LOH is an early event in the neoplastic cascade of BE.[44] Since the technique for 17p13 LOH is not routinely available, it is not commonly being

applied yet.[50]

p16

p16 is a tumor suppressor gene, which is located on chromosome 9p21. This gene is also known as cyclin-dependent kinase inhibitor 2 (CDKN2), INK4, or multiple tumor suppressor 1 (MST1).[47] Normally, the expression of p16 leads to G1 arrest by inhibiting the cyclin-dependent kinases that are responsible for phosphorylation of the retinoblastoma protein (Figure 2). Inactivation of p16 will lead to uncontrolled cell proliferation.[15] LOH is the predominant mechanism for inactivation of one of the p16 alleles, occurring in approximately 75% of samples taken from EAC.[51] Clones of cells with LOH at 9p21 have been shown to expand along the Barrett segment, creating a condition in which other mutations may arise that are able to induce EAC.[47, 52] CpG island methylation, mutation or homozygous deletions have also been suggested to be responsible for inactivation of the remaining p16 allele.[45, 52-54] Epigenetic modification of genes may already take place in normal mucosa of patients at risk of developing EAC, since hypermethylation was found to be present in 56% of biopsies of squamous epithelium of patients with EAC,[54] with no differences being found in the prevalence of p16 abnormalities (i.e. p16 CpG island methylation, p16 mutation, and 9p21 LOH) with advancing grades of dysplasia (88% in ND, 87% in LGD, and 86% in HGD).[52] It was shown that both LOH of 9p21 and p16 mutation occur as early lesions in diploid cell populations, prior to the development of aneuploidy and cancer.[47, 51] Although LOH of 9p21 is a common event in BE, large-scale studies have not been performed and the technique is not routinely available in most centers.

Adhesion molecules

Epithelial cells are tightly connected (cell-cell adhesion) with each other and one of the functions of this adhesion is to prevent development of malignancies by inhibition of proliferation. If cell-cell adhesion is loosened, penetration of toxic compounds, pathogenic organisms and inflammatory cells may occur which can cause DNA damage e.g. through the formation of oxygen radicals.[55] These oxygen radicals can cause DNA mutations leading to carcinogenesis. In addition, the loosened cell-cell connections make it easier for the neoplastic process to invade neighbouring tissues. Changes in adhesion proteins could therefore be valuable in predicting neoplastic progression of BE towards HGD/EAC. The most commonly reported adhesion proteins are E-cadherin and β-catenin.

E-cadherin & β-catenin

The transmembrane glycoprotein E-cadherin belongs to the family of calcium-dependent Wnt-related genes and plays a role in morphogenesis of tissues during embryogenesis. β-Catenin is directly linked to E-cadherin and together these proteins mediate cell-to-cell adhesion. The cell adhesion function of E-cadherin is frequently disturbed in carcinomas either by downregulation or mutation of the E-cadherin/catenin genes.[56] Adenomatous polyposis coli (APC) tumor suppressor gene (located at 5q21) regulates the intracellular concentration of β-catenin by causing its degradation. When the APC tumor suppressor gene is mutated, β-catenin accumulates in the nucleus and binds to transcription factors, resulting in the promotion of cellular proliferation and the prevention of

cellular death.[15] Normally, β-catenin is expressed in the membrane.[57] In BE, a decrease of both E-cadherin and membranous β-catenin on the one hand and an increase of nuclear β-catenin on the other hand has been observed during progression to EAC.[57-61] In a study by Murray et al., nuclear β-catenin expression was however not significantly associated with risk of malignant progression in BE.[30] As a result of these contradictory findings and the absence of large scale clinical cohort studies, the practical value of these proteins as biomarkers for predicting the risk of neoplastic progression in BE is still unclear.

DNA content

With the exception of germ-line cells, all other cells are normally diploid (2N). Human malignancies are associated with genomic instability, and many solid tumors show abnormalities of the cellular DNA content (aneuploidy or tetraploidy), which can be assessed by flow cytometry.[62] Aneuploidy is diagnosed if an increased number of cells are in the S phase of the cell cycle (Figure 2). This can be seen at flow cytometric analysis as a second discrete peak at > 2.7N in the histogram, comprising at least 2.5% of nuclei.[63, 64] Tetraploidy is present if > 6% of the nuclei are in the G2 phase, which is expressed by an increased 4N fraction (within range 3.85N-4.1N) at flow cytometry.[63-66] Another type of change of the DNA content is loss of heterozygosity (LOH) for one or two alleles of a gene, leading to inactivation of a protein,[37] as described above for p53 and p16.

DNA Ploidy

Neoplastic progression in BE is also associated with a process of genomic instability, leading to evolution of multiple aneuploid populations and finally to the development of a clone of cells capable of malignant invasion.[67]

A correlation between an increase in the percentage of biopsies with an abnormal DNA content (aneuploidy or tetraploidy) and an increase in the grade of dysplasia in BE has been reported.[27, 68, 69] The percentage of abnormal DNA content ranges from 0-13% in ND, 0-60% in LGD, 40-100%

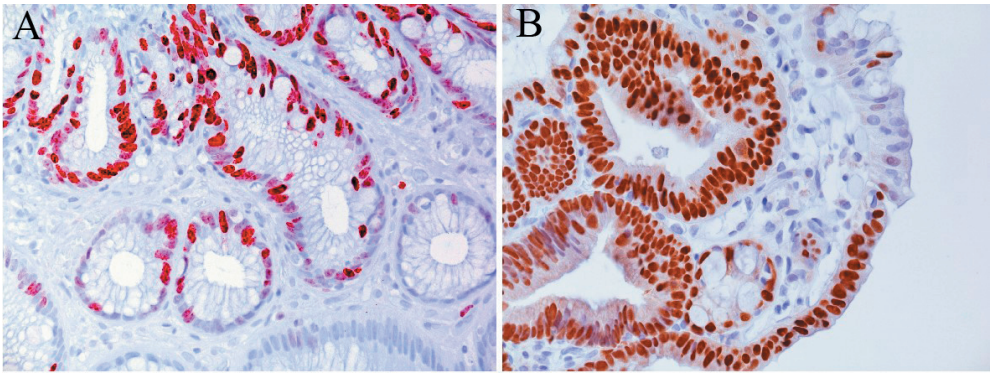


Figure 3. Typical examples of immunohistochemical staining for Ki67 and p53 expression in Barrett esophagus (specialized columnar epithelium). Original magnifications x400. (A) Ki67 overexpression. (B) p53 overexpression.

in HGD and 71-100% in EAC.[27, 66, 68, 70] Follow-up studies have suggested that the combination of histology and flow cytometry could be useful for the identification of BE patients at risk of developing EAC.[63, 71, 72] In contrast, Gimenez et al. found that DNA content as detected by flow cytometry was not able to predict progression in patients with ND or LGD. In this study, it was suggested that in the 'indefinite for dysplasia' group, abnormal DNA content could be used to differentiate between future neoplastic progression and reactive epithelial changes.[73] The majority of studies performed were based on the analysis of fresh material,[27, 63, 66, 71, 72] since when compared to analysis based on formalin fixed, paraffin-embedded biopsies material, the resulting histograms are of better quality due to a smaller amount debris, resulting in smaller peaks in the histogram. A disadvantage of fresh material is that immediate processing following biopsy is required to prevent the occurrence of false-positive DNA aneuploidy results.[64] This method is therefore not applicable in centers without an infrastructure to process fresh biopsy samples immediately. The technique of flow cytometry has been improved in such a way that the results on the more easily available formalin fixed, paraffin-embedded biopsies have become comparable with those on fresh tissue.[68, 70, 74] This suggests that DNA content as assessed by flow cytometry has the potential to become an easy to apply and useful biomarker for predicting neoplastic progression in BE. Additional large-scale prospective follow-up studies on formalin fixed, paraffin-embedded biopsies are however needed to confirm the value of the DNA-ploidy status as biomarker in BE.

Inflammation associated markers

Due to gastroesophageal reflux, injured epithelial cells will secrete inflammatory mediators such as cytokines and chemokines, leading to the attraction of inflammatory cells. These inflammatory cells produce reactive oxygen species, that can cause DNA damage and in this way induce tumor promoting mutations.[75] Cyclo-oxygenase-2 (COX-2) is the most well described inflammatory enzyme in relation to neoplastic progression in BE.

COX-2

COX-2 is an enzyme which is induced by inflammatory stimuli and cytokines, and catalyses the synthesis of prostaglandins from arachidonic acid. These prostaglandins stimulate cancer cell proliferation, inhibit apoptosis, and enhance cancer-induced angiogenesis and invasiveness.[76]

In most studies, a high expression level of COX-2 in HGD and EAC has been demonstrated.[77-80] There is however conflicting evidence as to whether COX-2 is involved in early development of EAC, since levels of COX-2 vary considerably in BE patients with ND or LGD.[81] Some studies showed no differences between ND and LGD,[77, 80] whereas others reported a progressive increase in COX-2 expression throughout the metaplasia-dysplasia-adenocarcinoma sequence.[78, 79] Cheong et al. found increased COX-2 expression in HGD compared to non-dysplastic BE, however COX-2 expression in EAC was decreased compared to HGD and not different from ND.[76] In a study by Murray et al., the combination of COX-2 expression and p53 expression was associated with a high risk of neoplastic progression (OR 27.3), although this combination was only present in 15% of patients who developed EAC.[30]

Table 1. Summary of the presently available biomarkers in Barrett esophagus and their pros and cons as biomarker for predicting an increased risk of cancer development in BE

Biomarker		Type of change	Pros	Cons	Potential Use
PCNA		increased expression with proliferation	easy to perform	also stains resting cells	+ / -
Ki67		increased expression with proliferation	easy, stains only proliferating cells	no large-scale longitudinal studies	+
p53	IHC	abnormal protein expression	easy to perform, cheap	stains also wild-type p53	+
	LOH	frequent LOH at 17p13	positive large prospective study	limited availability	+ / -
p16		LOH at 9p21, early lesion	common event	limited availability, no large-scale studies	-
Cyclin D1			easy to perform	contradictory findings, no large-scale studies	-
β-catenin		increased nuclear expression, decreased membranous expression		no large-scale studies	-
DNA ploidy		aneuploidy with progression	positive prospective studies performed on fresh tissue	more prospective studies needed on paraffin-embedded tissue	+
COX-2		increased expression		contradictory findings	-

IHC, immunohistochemistry; LOH, loss of heterozygosity; PCNA, proliferating cell nuclear antigen; COX-2, cyclo-oxegenase-2
+ probable, +/- possible, - unlikely

Different techniques have been used to evaluate COX-2 expression, such as immunohistochemistry,[76, 77, 80] Western-blotting [77] or reverse transcriptase/real time polymerase chain reaction,[79, 80, 82] but inconsistent results have been reported for all three techniques. Therefore, COX-2 is yet not reliable enough to be used as biomarker for stratifying neoplastic risk in BE.

CONCLUSION

It is generally accepted that the development of EAC in BE is a gradual process in which the disruption of biological processes at the cellular level is accumulating in the cascade from non-dysplastic BE, through LGD and HGD, and finally EAC.[83-86] At present, histological assessment of the degree of dysplasia is the gold standard for determining the risk of progression. This determines the frequency of endoscopic surveillance, according to the guidelines of the American College of Gastroenterology.[1] Several studies have evaluated various biomarkers that might assist in the risk stratification of progression from BE to EAC. In Table 1 the pros and cons of the biomarkers discussed in this review are briefly summarized. Although some biomarkers, such as DNA ploidy, p53 and Ki67, are promising candidates, either as an additional marker or even a substitute for histology, contradictory findings have been reported. Moreover, there is paucity of large prospective follow-up studies. For these reasons, biomarkers are not quite ready for use in clinical practice. One of the reasons that only a few large follow-up studies have been performed relates to the clinical observation that, although still increased, the incidence of EAC in the whole group of BE patients is relatively low.[7-9] Consequently, it is difficult to perform adequately powered prospective studies investigating the predictive value of biomarkers, unless performed in a multicenter setting.

Galipeau et al. recently showed that a panel of biomarkers (i.e. 17p13 LOH, 9p21 LOH, and DNA ploidy) improved the prediction of the subgroup of patients with BE at an increased risk of progression to EAC, compared to using only a single biomarker.[87] Therefore, future studies on risk stratification in BE should probably be performed in a multi-center setting in order to obtain large enough cohorts of BE patients that withstand rigorous statistical analysis, and should also be directed to the use of panels of biomarkers. We are convinced that in the future biomarkers will allow a more accurate prediction of the risk of neoplastic progression. And given the technological developments they will probably be determined in a (semi-)automated setup, eliminating observer-bias and thus replacing the ‘classic’, labour-intenstive histopathologic evaluation.

REFERENCES

1. Sampliner RE. Updated guidelines for the diagnosis, surveillance, and therapy of Barrett's esophagus. *Am J Gastroenterol* 2002; 97: 1888-95.
2. Avidan B, Sonnenberg A, Schnell TG*, et al.* Hiatal hernia size, Barrett's length, and severity of acid reflux are all risk factors for esophageal adenocarcinoma. *Am J Gastroenterol* 2002; 97: 1930-6.
3. Schlemper RJ, Riddell RH, Kato Y*, et al.* The Vienna classification of gastrointestinal epithelial neoplasia. *Gut* 2000; 47: 251-5.
4. Schlemper RJ, Kato Y, Stolte M. Review of histological classifications of gastrointestinal epithelial neoplasia: differences in diagnosis of early carcinomas between Japanese and Western pathologists. *J Gastroenterol* 2001; 36: 445-56.
5. Corley DA, Levin TR, Habel LA*, et al.* Surveillance and survival in Barrett's adenocarcinomas: a population-based study. *Gastroenterology* 2002; 122: 633-40.
6. Dulai GS, Guha S, Kahn KL*, et al.* Preoperative prevalence of Barrett's esophagus in esophageal adenocarcinoma: a systematic review. *Gastroenterology* 2002; 122: 26-33.
7. Falk GW. Barrett's esophagus-is it bad for your health? *Am J Gastroenterol* 2005; 100: 2622-3.
8. Hage M, Siersema PD, van Dekken H*, et al.* Oesophageal cancer incidence and mortality in patients with long-segment Barrett's oesophagus after a mean follow-up of 12.7 years. *Scand J Gastroenterol* 2004; 39: 1175-9.
9. Conio M, Blanchi S, Lapertosa G*, et al.* Long-term endoscopic surveillance of patients with Barrett's esophagus. Incidence of dysplasia and adenocarcinoma: a prospective study. *Am J Gastroenterol* 2003; 98: 1931-9.
10. Montgomery E, Goldblum JR, Greenson JK*, et al.* Dysplasia as a predictive marker for invasive carcinoma in Barrett esophagus: a follow-up study based on 138 cases from a diagnostic variability study. *Hum Pathol* 2001; 32: 379-88.
11. Reid BJ, Haggitt RC, Rubin CE*, et al.* Observer variation in the diagnosis of dysplasia in Barrett's esophagus. *Hum Pathol* 1988; 19: 166-78.
12. Montgomery E, Bronner MP, Goldblum JR*, et al.* Reproducibility of the diagnosis of dysplasia in Barrett esophagus: a reaffirmation. *Hum Pathol* 2001; 32: 368-78.
13. McManus DT, Olaru A, Meltzer SJ. Biomarkers of esophageal adenocarcinoma and Barrett's esophagus. *Cancer Res* 2004; 64: 1561-9.
14. Renan MJ. How many mutations are required for tumorigenesis? Implications from human cancer data. *Mol Carcinog* 1993; 7: 139-46.
15. Beilstein M, Silberg D. Cellular and molecular mechanisms responsible for progression of Barrett's metaplasia to esophageal carcinoma. *Gastroenterol Clin North Am* 2002; 31: 461-79, ix.
16. Scholzen T, Gerdes J. The Ki-67 protein: from the known and the unknown. *J Cell Physiol* 2000; 182: 311-22.
17. Kimos MC, Wang S, Borkowski A*, et al.* Esophagin and proliferating cell nuclear antigen (PCNA) are biomarkers of human esophageal neoplastic progression. *Int J Cancer* 2004; 111: 415-7.
18. Jankowski J, McMenemin R, Yu C*, et al.* Proliferating cell nuclear antigen in oesophageal diseases; correlation with transforming growth factor alpha expression. *Gut* 1992; 33: 587-91.
19. Gillen P, McDermott M, Grehan D*, et al.* Proliferating cell nuclear antigen in the assessment of Barrett's mucosa. *Br J Surg* 1994; 81: 1766-8.
20. Jaskiewicz K, Louw J, Anichkov N. Barrett's oesophagus: mucin composition, neuroendocrine cells, p53 protein, cellular proliferation and differentiation. *Anticancer Res* 1994; 14: 1907-12.
21. Cattoretti G, Becker MH, Key G*, et al.* Monoclonal antibodies against recombinant parts of the Ki-67 antigen (MIB 1 and MIB 3) detect proliferating cells in microwave-processed formalin-fixed paraffin

sections. *J Pathol* 1992; 168: 357-63.
22. Ohbu M, Kobayashi N, Okayasu I. Expression of cell cycle regulatory proteins in the multistep process of oesophageal carcinogenesis: stepwise over-expression of cyclin E and p53, reduction of p21(WAF1/CIP1) and dysregulation of cyclin D1 and p27(KIP1). *Histopathology* 2001; 39: 589-96.
23. Hong MK, Laskin WB, Herman BE*, et al.* Expansion of the Ki-67 proliferative compartment correlates with degree of dysplasia in Barrett's esophagus. *Cancer* 1995; 75: 423-9.
24. Olvera M, Wickramasinghe K, Brynes R*, et al.* Ki67 expression in different epithelial types in columnar lined oesophagus indicates varying levels of expanded and aberrant proliferative patterns. *Histopathology* 2005; 47: 132-40.
25. Polkowski W, Baak JP, van Lanschot JJ*, et al.* Clinical decision making in Barrett's oesophagus can be supported by computerized immunoquantitation and morphometry of features associated with proliferation and differentiation. *J Pathol* 1998; 184: 161-8.
26. Baak JP, ten Kate FJ, Offerhaus GJ*, et al.* Routine morphometrical analysis can improve reproducibility of dysplasia grade in Barrett's oesophagus surveillance biopsies. *J Clin Pathol* 2002; 55: 910-6.
27. Reid BJ, Sanchez CA, Blount PL*, et al.* Barrett's esophagus: cell cycle abnormalities in advancing stages of neoplastic progression. *Gastroenterology* 1993; 105: 119-29.
28. Geddert H, Kiel S, Zotz RB*, et al.* Polymorphism of p16 INK4A and cyclin D1 in adenocarcinomas of the upper gastrointestinal tract. *J Cancer Res Clin Oncol* 2005; 131: 803-8.
29. Bani-Hani K, Martin IG, Hardie LJ*, et al.* Prospective study of cyclin D1 overexpression in Barrett's esophagus: association with increased risk of adenocarcinoma. *J Natl Cancer Inst* 2000; 92: 1316-21.
30. Murray L, Sedo A, Scott M*, et al.* TP53 and progression from Barrett's metaplasia to oesophageal adenocarcinoma in a UK population cohort. *Gut* 2006.
31. Lao-Sirieix P, Brais R, Lovat L*, et al.* Cell cycle phase abnormalities do not account for disordered proliferation in Barrett's carcinogenesis. *Neoplasia* 2004; 6: 751-60.
32. Casson AG, Zheng Z, Evans SC*, et al.* Cyclin D1 polymorphism (G870A) and risk for esophageal adenocarcinoma. *Cancer* 2005; 104: 730-9.
33. Levine AJ, Momand J, Finlay CA. The p53 tumour suppressor gene. *Nature* 1991; 351: 453-6.
34. Merola E, Claudio PP, Giordano A. p53 and the malignant progression of Barrett's esophagus. *J Cell Physiol* 2006; 206: 574-7.
35. Haupt S, Berger M, Goldberg Z*, et al.* Apoptosis - the p53 network. *J Cell Sci* 2003; 116: 4077-85.
36. Hinds PW, Finlay CA, Quartin RS*, et al.* Mutant p53 DNA clones from human colon carcinomas cooperate with ras in transforming primary rat cells: a comparison of the "hot spot" mutant phenotypes. *Cell Growth Differ* 1990; 1: 571-80.
37. Ireland AP, Clark GW, DeMeester TR. Barrett's esophagus. The significance of p53 in clinical practice. *Ann Surg* 1997; 225: 17-30.
38. Keswani RN, Noffsinger A, Waxman I*, et al.* Clinical use of p53 in Barrett's esophagus. *Cancer Epidemiol Biomarkers Prev* 2006; 15: 1243-9.
39. Ramel S, Reid BJ, Sanchez CA*, et al.* Evaluation of p53 protein expression in Barrett's esophagus by two-parameter flow cytometry. *Gastroenterology* 1992; 102: 1220-8.
40. Krishnadath KK, Tilanus HW, van Blankenstein M*, et al.* Accumulation of p53 protein in normal, dysplastic, and neoplastic Barrett's oesophagus. *J Pathol* 1995; 175: 175-80.
41. Symmans PJ, Linehan JM, Brito MJ*, et al.* p53 expression in Barrett's oesophagus, dysplasia, and adenocarcinoma using antibody DO-7. *J Pathol* 1994; 173: 221-6.
42. Younes M, Lebovitz RM, Lechago LV*, et al.* p53 protein accumulation in Barrett's metaplasia, dysplasia, and carcinoma: a follow-up study. *Gastroenterology* 1993; 105: 1637-42.

43. Hamelin R, Flejou JF, Muzeau F*, et al.* TP53 gene mutations and p53 protein immunoreactivity in malignant and premalignant Barrett's esophagus. *Gastroenterology* 1994; 107: 1012-8.
44. Reid BJ, Prevo LJ, Galipeau PC*, et al.* Predictors of progression in Barrett's esophagus II: baseline 17p (p53) loss of heterozygosity identifies a patient subset at increased risk for neoplastic progression. *Am J Gastroenterol* 2001; 96: 2839-48.
45. Gonzalez MV, Artimez ML, Rodrigo L*, et al.* Mutation analysis of the p53, APC, and p16 genes in the Barrett's oesophagus, dysplasia, and adenocarcinoma. *J Clin Pathol* 1997; 50: 212-7.
46. Blount PL, Ramel S, Raskind WH*, et al.* 17p allelic deletions and p53 protein overexpression in Barrett's adenocarcinoma. *Cancer Res* 1991; 51: 5482-6.
47. Galipeau PC, Prevo LJ, Sanchez CA*, et al.* Clonal expansion and loss of heterozygosity at chromosomes 9p and 17p in premalignant esophageal (Barrett's) tissue. *J Natl Cancer Inst* 1999; 91: 2087-95.
48. Maley CC, Galipeau PC, Li X*, et al.* The combination of genetic instability and clonal expansion predicts progression to esophageal adenocarcinoma. *Cancer Res* 2004; 64: 7629-33.
49. Blount PL, Galipeau PC, Sanchez CA*, et al.* 17p allelic losses in diploid cells of patients with Barrett's esophagus who develop aneuploidy. *Cancer Res* 1994; 54: 2292-5.
50. Flejou JF, Svrcek M. Barrett's oesophagus--a pathologist's view. *Histopathology* 2007; 50: 3-14.
51. Barrett MT, Sanchez CA, Galipeau PC*, et al.* Allelic loss of 9p21 and mutation of the CDKN2/p16 gene develop as early lesions during neoplastic progression in Barrett's esophagus. *Oncogene* 1996; 13: 1867-73.
52. Wong DJ, Paulson TG, Prevo LJ*, et al.* p16(INK4a) lesions are common, early abnormalities that undergo clonal expansion in Barrett's metaplastic epithelium. *Cancer Res* 2001; 61: 8284-9.
53. Wong DJ, Barrett MT, Stoger R*, et al.* p16INK4a promoter is hypermethylated at a high frequency in esophageal adenocarcinomas. *Cancer Res* 1997; 57: 2619-22.
54. Hardie LJ, Darnton SJ, Wallis YL*, et al.* p16 expression in Barrett's esophagus and esophageal adenocarcinoma: association with genetic and epigenetic alterations. *Cancer Lett* 2005; 217: 221-30.
55. Naya MJ, Pereboom D, Ortego J*, et al.* Superoxide anions produced by inflammatory cells play an important part in the pathogenesis of acid and pepsin induced oesophagitis in rabbits. *Gut* 1997; 40: 175-81.
56. Behrens J. Cadherins and catenins: role in signal transduction and tumor progression. *Cancer Metastasis Rev* 1999; 18: 15-30.
57. Bian YS, Osterheld MC, Bosman FT*, et al.* Nuclear accumulation of beta-catenin is a common and early event during neoplastic progression of Barrett esophagus. *Am J Clin Pathol* 2000; 114: 583-90.
58. Bailey T, Biddlestone L, Shepherd N*, et al.* Altered cadherin and catenin complexes in the Barrett's esophagus-dysplasia-adenocarcinoma sequence: correlation with disease progression and dedifferentiation. *Am J Pathol* 1998; 152: 135-44.
59. Swami S, Kumble S, Triadafilopoulos G. E-cadherin expression in gastroesophageal reflux disease, Barrett's esophagus, and esophageal adenocarcinoma: an immunohistochemical and immunoblot study. *Am J Gastroenterol* 1995; 90: 1808-13.
60. Washington K, Chiappori A, Hamilton K*, et al.* Expression of beta-catenin, alpha-catenin, and E-cadherin in Barrett's esophagus and esophageal adenocarcinomas. *Mod Pathol* 1998; 11: 805-13.
61. Seery JP, Syrigos KN, Karayiannakis AJ*, et al.* Abnormal expression of the E-cadherin-catenin complex in dysplastic Barrett's oesophagus. *Acta Oncol* 1999; 38: 945-8.
62. Nowell PC. The clonal evolution of tumor cell populations. *Science* 1976; 194: 23-8.
63. Reid BJ, Blount PL, Rubin CE*, et al.* Flow-cytometric and histological progression to malignancy in

Barrett's esophagus: prospective endoscopic surveillance of a cohort. *Gastroenterology* 1992; 102: 1212-9.

64. 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.
65. Fennerty MB, Sampliner RE, Way D*, et al.* Discordance between flow cytometric abnormalities and dysplasia in Barrett's esophagus. *Gastroenterology* 1989; 97: 815-20.
66. Reid BJ, Haggitt RC, Rubin CE*, et al.* Barrett's esophagus. Correlation between flow cytometry and histology in detection of patients at risk for adenocarcinoma. *Gastroenterology* 1987; 93: 1-11.
67. Rabinovitch PS, Reid BJ, Haggitt RC*, et al.* Progression to cancer in Barrett's esophagus is associated with genomic instability. *Lab Invest* 1989; 60: 65-71.
68. Montgomery EA, Hartmann DP, Carr NJ*, et al.* Barrett esophagus with dysplasia. Flow cytometric DNA analysis of routine, paraffin-embedded mucosal biopsies. *Am J Clin Pathol* 1996; 106: 298-304.
69. Krishnadath KK, Tilanus HW, van Blankenstein M*, et al.* Accumulation of genetic abnormalities during neoplastic progression in Barrett's esophagus. *Cancer Res* 1995; 55: 1971-6.
70. Gimenez A, Minguela A, Parrilla P*, et al.* Flow cytometric DNA analysis and p53 protein expression show a good correlation with histologic findings in patients with Barrett's esophagus. *Cancer* 1998; 83: 641-51.
71. Reid BJ, Levine DS, Longton G*, et al.* Predictors of progression to cancer in Barrett's esophagus: baseline histology and flow cytometry identify low- and high-risk patient subsets. *Am J Gastroenterol* 2000; 95: 1669-76.
72. Teodori L, Gohde W, Persiani M*, et al.* DNA/protein flow cytometry as a predictive marker of malignancy in dysplasia-free Barrett's esophagus: thirteen-year follow-up study on a cohort of patients. *Cytometry* 1998; 34: 257-63.
73. Gimenez A, Minguela A, de Haro LM*, et al.* DNA ploidy status and proliferative activity as markers of malignant potential in Barrett's esophagus: flow cytometric study using routinely paraffin-embedded tissue. *World J Surg* 2000; 24: 72-7.
74. Hedley DW. Flow cytometry using paraffin-embedded tissue: five years on. *Cytometry* 1989; 10: 229-41.
75. Bax DA, Siersema PD, Van Vliet AH*, et al.* Molecular alterations during development of esophageal adenocarcinoma. *J Surg Oncol* 2005; 92: 89-98; discussion 99.
76. Cheong E, Igali L, Harvey I*, et al.* Cyclo-oxygenase-2 expression in Barrett's oesophageal carcinogenesis: an immunohistochemical study. *Aliment Pharmacol Ther* 2003; 17: 379-86.
77. Morris CD, Armstrong GR, Bigley G*, et al.* Cyclooxygenase-2 expression in the Barrett's metaplasia-dysplasia-adenocarcinoma sequence. *Am J Gastroenterol* 2001; 96: 990-6.
78. Shirvani VN, Ouatu-Lascar R, Kaur BS*, et al.* Cyclooxygenase 2 expression in Barrett's esophagus and adenocarcinoma: Ex vivo induction by bile salts and acid exposure. *Gastroenterology* 2000; 118: 487-96.
79. Kuramochi H, Vallbohmer D, Uchida K*, et al.* Quantitative, tissue-specific analysis of cyclooxygenase gene expression in the pathogenesis of Barrett's adenocarcinoma. *J Gastrointest Surg* 2004; 8: 1007-16; discussion 16-7.
80. Lagorce C, Paraf F, Vidaud D*, et al.* Cyclooxygenase-2 is expressed frequently and early in Barrett's oesophagus and associated adenocarcinoma. *Histopathology* 2003; 42: 457-65.
81. Mehta S, Boddy A, Johnson IT*, et al.* Systematic review: cyclo-oxygenase-2 in human oesophageal adenocarcinogenesis. *Aliment Pharmacol Ther* 2006; 24: 1321-31.
82. Shimizu D, Vallbohmer D, Kuramochi H*, et al.* Increasing cyclooxygenase-2 (cox-2) gene expression

in the progression of Barrett's esophagus to adenocarcinoma correlates with that of Bcl-2. *Int J Cancer* 2006; 119: 765-70.

83. Hamilton SR, Smith RR. The relationship between columnar epithelial dysplasia and invasive adenocarcinoma arising in Barrett's esophagus. *Am J Clin Pathol* 1987; 87: 301-12.
84. Haggitt RC. Barrett's esophagus, dysplasia, and adenocarcinoma. *Hum Pathol* 1994; 25: 982-93.
85. Hameeteman W, Tytgat GN, Houthoff HJ*, et al.* Barrett's esophagus: development of dysplasia and adenocarcinoma. *Gastroenterology* 1989; 96: 1249-56.
86. Buttar NS, Wang KK. Mechanisms of disease: Carcinogenesis in Barrett's esophagus. *Nat Clin Pract Gastroenterol Hepatol* 2004; 1: 106-12.
87. Galipeau PC, Li X, Blount PL*, et al.* NSAIDs modulate CDKN2A, TP53, and DNA content risk for progression to esophageal adenocarcinoma. *PLoS Med* 2007; 4: e67.

Aneuploidy and high expression of p53 and Ki67 predict neoplastic progression in Barrett esophagus

6

Submitted

Marjon Kerkhof [1]
Ewout W Steyerberg [2]
Johannes G Kusters [1]
Herman van Dekken [3]
Adriana J van Vuuren[1]
Ernst J Kuipers [1]
Peter D Siersema [1]

Departments of [1]Gastroenterology and Hepatology, [2]Public Health, and [3]Pathology,
Erasmus MC - University Medical Center Rotterdam, Rotterdam, the Netherlands

ABSTRACT

Background: Unless detected at an early stage, esophageal adenocarcinoma (EAC) has a poor prognosis. Changes in cellular DNA content and expression levels of p53 and Ki67 in Barrett esophagus (BE) are associated with the development EAC and might serve as markers to identify EAC at an early stage.

Aim and methods: To examine the presence of these three markers in various steps of neoplastic progression in BE towards EAC. Dysplasia was graded in 212 biopsy sets taken during follow-up upper endoscopy in 27 patients in whom ultimately high-grade dysplasia (HGD) or EAC was detected. Ploidy status was determined by flow cytometry, whereas Ki67 and p53 expression was determined by immunohistochemistry. Smoothing splines were used to analyze trends in time.

Results and conclusions: We found an increasing fraction of Ki67 overexpression and, to a lesser extent, abnormal DNA content in biopsies from BE patients, suggesting the potential value of these biomarkers in identifying patients at increased risk of progression towards HGD/EAC. Accumulation of p53 was seen several years before development of HGD/EAC, and may therefore be an early marker in BE at a stage when dysplasia is not yet detected with conventional histology. Prospective follow-up studies are needed to confirm these findings.

INTRODUCTION

Barrett esophagus (BE) is characterized by the replacement of the normal stratified squamous epithelium of the distal esophagus by columnar epithelium with specialized intestinal metaplasia (IM), which is characterized by the presence of goblet cells.[1] BE is caused by chronic gastroesophageal reflux,[2] and predisposes to the development of esophageal adenocarcinoma (EAC). The development of EAC is a gradual process in which the accumulation of (epi)genetic changes results in disruption of different biological processes at the cellular level. The morphologic counterpart of these cell biologic changes is called dysplasia, and is classified in different stages, i.e., low-grade dysplasia (LGD), high-grade dysplasia (HGD), and finally EAC.[3] Currently, histopathologic assessment is routinely used to stage the neoplastic progression in an individual patient and consequently to determine the interval of endoscopic surveillance in patients with BE. Surveillance aims to detect progression of dysplasia at an early, and therefore likely curable stage.[4]

The annual risk of EAC development in BE of was recently estimated to be approximately 0.5%. The majority of patients will therefore never progress towards EAC.[5-7] Substantial interobserver variation in the diagnosis of LGD makes histological evaluation of limited value in defining the subset of patients who need more frequent endoscopic surveillance.[8, 9]

It has been shown that aneuploidy was associated with progression towards EAC, and the combination of histology and flow cytometry was able to identify BE patients at risk for developing EAC.[10, 11] Another promising biomarker is the tumor suppressor gene p53, involved in controlling cell proliferation.[12] This marker showed a stepwise overexpression with increasing grade of dysplasia in cross-sectional settings.[13-16] On the other hand, a recent longitudinal study reported that the sensitivity of p53 as a biomarker was too low to predict HGD/EAC risk.[17] The human Ki67 protein, which is present during all active phases of the cell cycle (G1, S, G2, M), but is absent in resting cells (G0), is strictly associated with cell proliferation.[18] Ki67 could therefore serve as a potential biomarker in the identification of patients at risk of neoplastic progression. No longitudinal studies have so far evaluated Ki67 expression in BE patients during follow-up.

The aim of this study was to determine ploidy status and to examine expression of Ki67 and p53 in patients with BE who developed HGD or EAC after a prolonged period of endoscopic surveillance, and to evaluate if these markers could be used to identify a subgroup of patients at risk for subsequent progression to EAC.

METHODS

Patients and materials

In this retrospective, longitudinal study, 355 patients referred to the Erasmus MC - University Medical Center Rotterdam between January 1994 and December 2005 because of HGD or EAC in BE were evaluated. In order to be included, patients needed to have undergone at least one surveillance endoscopy with biopsies taken prior to the development of HGD/EAC, with a histological diagnosis of BE. As result, 328 patients were excluded and 27 patients were included into this study. All available paraffin blocks with biopsies taken at different levels of the columnar-lined esophagus

(CLE) from all previously performed follow-up upper gastrointestinal (GI) endoscopies pertaining to this group of patients were retrieved.

Histology and immunohistochemistry

Three consecutive sections of 4 µm each from every available biopsy set were mounted on adhesive slides, dried overnight at 37°C, and subsequently deparaffinized with xylene. The first of these serially sectioned slides was stained with haematoxylin and eosin (H&E) and evaluated by light microscopy for the presence of IM and grade of dysplasia by an expert gastrointestinal pathologist (HvD). IM was defined as the presence of goblet cell containing glands in columnar epithelium of the esophagus.[3] In line with the definition of BE according to the American College of Gastroenterology (ACG) guidelines,[4] blocks with biopsies with CLE but without IM were excluded from this study. Dysplasia in BE was graded according to the Consensus Criteria of 1988, with adjustments made in 2001.[9, 19]

The next two slides were used for immunohistochemistry to assess expression of p53 protein, and to estimate the proliferation rate by labeling the Ki67 antigen. Immunohistochemistry was performed as previously described,[20] except for p53 staining, where antigen retrieval was performed by boiling the deparaffinized samples for 15 minutes in 10mM Tris/EDTA buffer (pH 9.0). Sections were incubated for 16 hours at 4°C with the primary antibody anti-human Ki67 antigen (clone MIB-1, DAKO) in a 1:100 dilution, or for 1 hour at RT with the primary antibody anti-human p53 protein (Clone DO-7, DAKO) in a 1:1000 dilution. 3-Amino-9-ethylcarbazole was used as substrate for Ki67 and diaminobenzodine for p53. All slides were analyzed in a blinded fashion for nuclear Ki67 and p53 staining by two independent investigators. Only moderate to intense brown (p53) or red (Ki67) nuclear staining was considered positive. The percentage of positive cells was graded as previously described.[21] For p53, a percentage of positive cells ≤ 15% was regarded as normal expression (grade 0), 16-40% as moderate overexpression (grade 1) and >40% as strong (grade 2) overexpression. For Ki67, the percentage of positive cells was determined in longitudinally sectioned crypts and villi, and a percentage of positive cells ≤ 20% was regarded as normal expression (grade 0). A percentage of positive cells greater than 20% was regarded as increased proliferation (21-50% as moderate (grade 1), and >50% as strong (grade 2) overexpression).

Flow Cytometry

Sections of 50 µm from every biopsy set were deparaffinized with xylene and rehydrated in series of decreasing concentrations of ethanol, and finally phosphate-buffered-saline (PBS). Subsequently, the samples were incubated with 0.05% protease in a water bath at 37°C for 45 minutes, with intermittent vigorous mixing. After centrifugation at 800g for 5 minutes at 4°C, the solution was removed and the samples were resuspended in a 0.3% BSA in PBS solution with a 20G needle and filtered through a 50 µm nylon mesh. After centrifuging (5 min at 800g), the BSA-solution was removed, and a 400 µl 0.01% RNAse-solution was added to digest RNA. Finally, the DNA was stained with propidium iodide (0.1mg/ml).[22]

Subsequently the DNA content of the isolated nuclei was analyzed using a 4-color flow cytometer

(FACScalibur, Becton Dickinson, San Jose, CA). Data analysis was performed using CellQuest software (Becton Dickinson). At least 10.000 nuclei were counted for each sample. The DNA content of the nuclei was determined at a slow flow speed to avoid the possibility of interpreting doublets as aneuploid cell populations. Flow cytometric results were independently scored by two investigators (MK, AJvV) who were blinded for the histological and immunohistochemical data of the samples. In line with previous studies, aneuploidy was defined as the presence of a second discrete peak on the histogram at > 2.7N containing at least 2,5% of the nuclei.[10, 23] Similarly, tetraploidy was defined as the presence of a 4N fraction (range: 3.85N-4.1N) consisting > 6% of the nuclei, as described previously.[10, 23] Finally, diploidy (normal DNA content) was defined as the presence of a large peak at 2N containing the majority of nuclei, while the remaining nuclei did not fulfill the criteria of aneuploidy or tetraploidy.

Statistical analysis

Mann-Whitney tests and Kruskal-Wallis tests were used to compare the results of the different markers. A two-sided p-value < 0.05 was considered statistically significant. We considered p-values to be approximate, since correlations of marker results in individual patients were ignored. Smoothing splines were used to indicate trends over time. Statistic analyses were conducted using SPSS software (SPSS version 11.0, Chicago, Illinois, USA) and S-Plus software (S-Plus version 7, Insightful Inc., Seattle, WA, USA).

RESULTS

Patients and materials

Twenty-seven patients (24 males, 3 females) were included in this study, with a mean (± standard deviation (SD)) age of 59 ± 10 years (range 37-67 years) at the time of BE diagnosis. A total of 167 upper GI endoscopies (mean: 6 ± 3, range 2-18, per patient) was performed in the period from BE diagnosis until end of follow-up, which yielded 212 paraffin blocks with biopsies from the CLE with an histological diagnosis of BE. Mean age at development of HGD/EAC was 65 ± 9 years (range 47-78 years). A total of 108 endoscopies (mean: 4 ± 3, range 2-15, per patient) was performed before the diagnosis of HGD/EAC was made, which yielded 129 paraffin blocks with biopsies that could be evaluated. Mean time between first endoscopy and development of HGD/EAC was 75 ± 51 months (range 4-187 months). Individual results of patients are shown in the Appendix.

Histology

ND was found in 99/212 (47%) biopsy samples, whereas LGD was detected in 69 (32%) samples, HGD in 31 (15%), and EAC in 13 (6%). Of the biopsies taken prior to the diagnosis of HGD/EAC, 76/129 (59%) showed ND and 53 (41%) LGD. Although ND and LGD were alternately diagnosed in sequential biopsies of most patients until HGD/EAC developed (see Appendix), an increasing fraction of patients with LGD was observed when biopsies were taken at a closer time point prior to the moment of detecting HGD/EAC (Figure 1a).

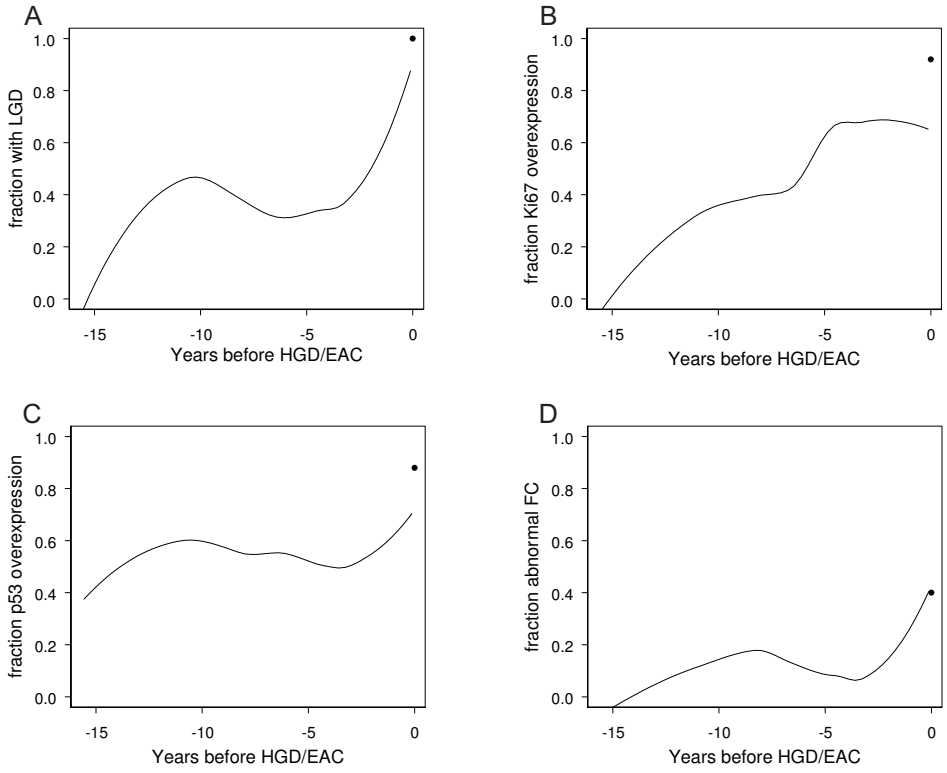


Figure 1. Fraction of patients with an abnormal result of biomarkers in their biopsies over time until HGD/EAC development. In figure (A) the fraction with LGD is shown, in figure (B) the fraction with Ki67 overexpression, in (C) the fraction with p53 overexpression, and in (D) the fraction with an abnormal DNA content (aneuploidy or tetraploidy). Smoothing splines were used to indicate the trend over time. The black dot at time point zero represents the fraction among the biopsy samples with HGD/EAC.

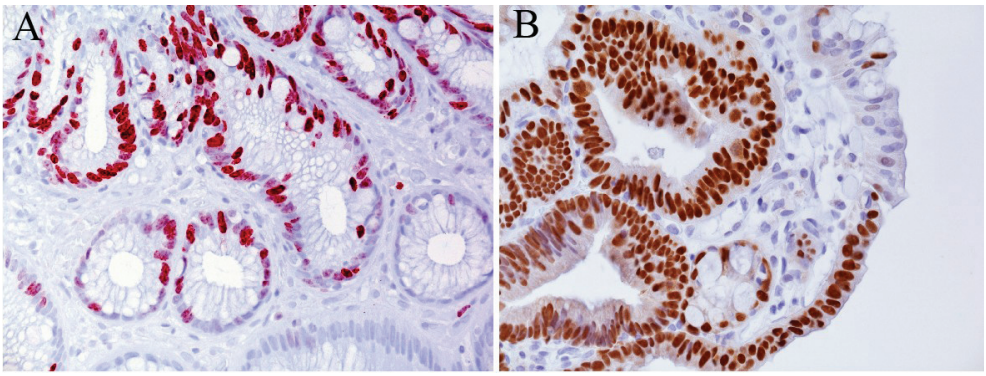


Figure 2. Examples of immunohistochemical staining for Ki67 and p53 expression (both grade 2) in biopsy samples of Barrett esophagus with high-grade dysplasia. Original magnifications x400. (A) Ki67 overexpression. (B) p53 overexpression.

Ki67 expression

Normal Ki67 expression (grade 0) was present in 66/212 (31%) biopsy samples, grade 1 overexpression in 81 (38%), and grade 2 overexpression in 65 (31%) (Figure 2a; Figure 3a). A stepwise increase in samples with Ki67 overexpression (grade 1 plus grade 2) was seen with increasing dysplasia grades (60/99 (61%) in ND, 46/69 (67%) in LGD, 28/31 (90%) in HGD, and 12/13 (92%) in EAC) (p=0.004) (Figure 3a). The proportion of biopsy samples with Ki67 overexpression was higher in biopsies with HGD or EAC compared to those with ND or LGD (p<0.001). To evaluate if Ki67 could be valuable as a predictive marker for neoplastic progression, we separately analyzed the biopsy samples before HGD/EAC development. The proportion of these samples (with ND or LGD) with grade 2 expression was 12% (15/129), while grade 0 and grade 1 were present in respectively 43% and 45%. In total, 73/129 (57%) biopsy samples of 23 patients showed Ki67 overexpression before HGD/EAC was diagnosed. The proportion of samples with Ki67 overexpression increased over time towards development of HGD/EAC (Figure 1b). The prevalence of LGD was similar in those with normal Ki67 expression (39%) or with overexpression (43%) (p=0.72). Separate analysis of Ki67 overexpression showed however that in biopsies with grade 2 expression, the percentage of biopsy samples with LGD was higher than in those with grade 1 expression (80% versus 33%, p=0.001).

p53 expression

Normal p53 expression (grade 0) was present in 97/211 (46%) biopsies samples, grade 1 overexpression in 29 (14%) samples, and grade 2 overexpression in 85 (40%) samples (Figure 2b; Figure 3b). One sample could not be evaluated, as there was not enough tissue available. The percentage of biopsies with p53 overexpression (grade 1 and 2) increased with increasing grades of dysplasia (38/98 (39%) in ND, 43/69 (62%) in LGD, 24/31 (77%) in HGD, and 9/13 (69%) in EAC, p<0.001), and was significantly higher in HGD/EAC compared to ND/LGD (p=0.002). To evaluate the usefulness of p53 as a predictive marker for neoplastic progression, we separately analyzed

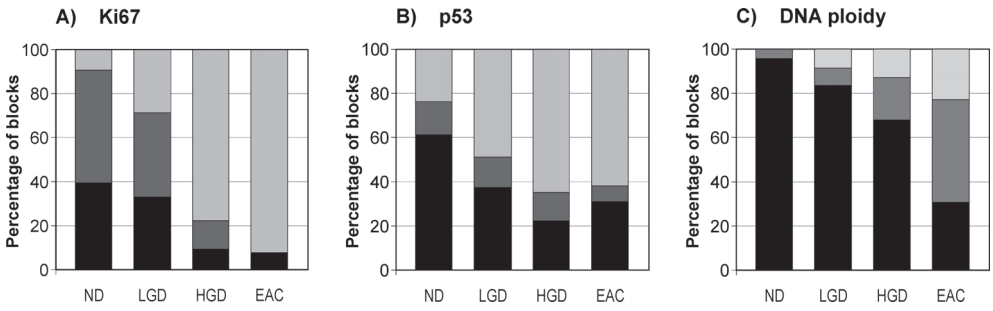


Figure 3. Results of biomarkers in biopsy samples with no dysplasia (ND), low-grade dysplasia (LGD), high-grade dysplasia (HGD) and esophageal adenocarcinoma (EAC). (A) Ki67 expression (■ 0-20% cells positive, ■ 21-50% cells positive, ■ ≥ 50% cells positive). (B) p53 expression (■ 0-15% cells positive, ■ 16-40% cells positive, ■ ≥ 40% cells positive). (C) DNA ploidy (■ diploid cell population, ■ aneuploid cell population, ■ tetraploid cell population).

biopsy samples before the development of HGD/EAC. The percentage of samples with ND or LGD with grade 2 expression was 35% (45/128), grade 1: 14%, and grade 0: 51%. In total, 63/128 biopsy samples of 22 patients showed p53 overexpression before the development of HGD/EAC. Overexpression was already present up to 15 years before HGD/EAC was diagnosed. The fraction of samples with p53 overexpression remained stable over time (Figure 1c). The percentage of biopsy samples with LGD was higher when p53 overexpression was present (51%), compared to samples with normal p53 expression (32%) (p=0.03).

Flow cytometric results

One-hundred seventy-six of 210 (84%) biopsy samples showed a normal, diploid DNA content, whereas in 34 (16%) samples aneuploidy or tetraploidy was detected. In 2/212 (1%) samples not enough nuclei could be isolated for flow cytometric evaluation. The percentage of samples with abnormal DNA content increased with increasing grades of dysplasia (4/98 (4%) in ND, 11/68 (16%) in LGD, 10/31 (32%) in HGD, and 9/13 (69%) in EAC, p<0.001) (Figure 3c). Abnormal DNA content was found in 43% (19/45 (63% aneuploidy and 37% tetraploidy)) biopsy samples with HGD or EAC, which was higher than in biopsies with ND or LGD (9% (15/166) abnormal DNA content) (p<0.001). When only biopsy samples obtained during endoscopies prior to a diagnosis of HGD/EAC were evaluated, aneuploidy/tetraploidy was observed in 13/128 (10%) samples of 10 patients. The percentage of samples with aneuploidy/tetraploidy increased over time towards HGD/EAC (Figure 1d). Seventy-seven percent (10/13) of the biopsy samples with aneuploidy/tetraploidy were graded as LGD, while LGD was only present in 37% (43/115) of samples with diploidy (p=0.006).

Predictive value of combined markers

Analyzing the results of the different biomarkers together, we found that with increasing grades of dysplasia, the number of abnormal markers also increased (p<0.001) (Table 1). If only 1 marker was abnormal, it was most often Ki67 overexpression (52/79, 66%). If 2 markers were abnormal, it was in the majority of cases (62/74, 84%) the combination of Ki67 and p53 overexpression. In all patients, at least one biomarker was abnormal in one or more biopsy samples before a histological diagnosis of HGD/EAC was made. In total, 101/129 (78%) biopsy samples expressed one or more

Table 1. Number of markers with an abnormal biomarker result (aneuploidy/tetraploidy, Ki67 overexpression or p53 overexpression) in the different grades of dysplasia

	Number of markers with abnormal result			
Grade of dysplasia	0	1	2	3
ND	25 (26%)	44 (45%)	28 (29%)	0
LGD	7 (10%)	31 (46%)	23 (34%)	7 (10%)
HGD	2 (7%)	3 (10%)	19 (61%)	7 (23%)
EAC	1 (8%)	1 (8%)	4 (31%)	7 (54%)

ND, no dysplasia; LGD, low-grade dysplasia; HGD, high-grade dysplasia, EAC, esophageal adenocarcinoma

abnormal biomarkers before the development of HGD/EAC, with a mean of 4 biopsy samples per patient with one or more abnormal biomarkers (range 1-17). To determine if these biomarkers could possibly replace current histology, we evaluated if abnormal results of biomarkers were present prior to or simultaneously with the presence of LGD. In total, in 19/27 (70%) patients p53 overexpression was detected prior to or simultaneous with the detection of LGD, compared to 12/27 (44%) cases in which LGD detection was present prior to or simultaneous with one or more other markers with an abnormal result. In 13/27 (48%) cases Ki67 overexpression was present prior to or simultaneous with the detection of LGD, and for FC abnormalities this was observed in 6/27 (22%) cases. In 15/27 (56%) patients, one or more biomarkers were abnormal prior to initial detection of LGD. Of these, p53 overexpression was most frequently found (9/15 (60%) cases).

DISCUSSION

An almost linear increase in the fraction of biopsy samples with LGD and/or Ki67 overexpression was observed at time points closer to the time of detecting HGD/EAC (Figure 1). Together with the fact that the interpretation of Ki67 expression has a low to moderate interobserver variation,[24] our findings suggest that Ki67 may be a reliable marker to predict risk of neoplastic progression in BE. In addition, in line with previous cross-sectional studies, a stepwise increase in the percentage of biopsy samples with Ki67 overexpression with increasing grade of dysplasia in BE was found.[13, 25, 26] The significant difference in Ki67 expression between biopsies with ND or LGD compared to those with HGD or EAC can be used to differentiate LGD from HGD, as was reported previously.[27] In contrast, others concluded that Ki67 overexpression could only be used as an additional parameter to differentiate between patients with or without dysplasia in BE.[28] In a study by Polkowski et al.,[29] it was demonstrated that cell morphometry with assessment of the Ki67-positive area resulted in a more objective assessment of BE grading.[30]

A similar, but less pronounced, trend was found for an abnormal ploidy status (Figure 1). This suggests that ploidy status alone as a biomarker may not be sensitive enough in predicting neoplastic progression in BE, which was in contrast with findings in previous studies.[10, 11, 31, 32] As expected, we found that the prevalence of abnormal DNA content as determined by FACS analysis paralleled an increase in grade of dysplasia.[31, 33, 34] Abnormal DNA content was significantly more common in biopsy samples with HGD or EAC, although still only 43% of these showed aneuploidy or tetraploidy. A separate analysis of HGD (32%) and EAC (69%) demonstrated however that our results were in the lower range of those reported in previous studies, in which an abnormal ploidy status ranged from 40-77% in HGD and from 71-100% in EAC.[31, 33-35] Our inferior results might have been caused by the fact that we used paraffin-embedded tissues instead of fresh material for flow cytometry. Although fresh material would have been preferable as histograms are often of better quality due to less debris and smaller peaks, the method for the more easily available paraffin-embedded biopsies has been improved over the years and results have been shown to be almost comparable.[36]

In 70% of patients, p53 expression was increased prior to or simultaneous with a first diagnosis of

LGD. As the percentage of samples with p53 overexpression remained stable over time (Figure 1), it confirms the observation in other studies that p53 could be particularly useful as an early marker in identifying patients with an increased risk of neoplastic progression some time before other markers (including LGD) become positive in BE.[37, 38] In a recent study, the sensitivity of p53 as a biomarker was however found to be too low to predict HGD/EAC risk.[17] As expected, the frequency of samples with p53 overexpression increased with the severity of dysplasia,[13-16, 38-40] although in these studies p53 overexpression was much lower in ND (0-5%) and LGD (9-36%) compared to our findings.[13-16, 39] An explanation for this could be that, in contrast to these studies, we compared biopsy samples of a patient cohort, in which all cases finally developed HGD or EAC.

Only 25% of patients with LGD have been reported to finally develop neoplastic progression.[6] Dysplasia may therefore have limited value as a marker for risk of progression to HGD/EAC. In 78% of our biopsies, one or more of the investigated biomarkers were abnormal prior to the development of HGD/EAC, suggesting that this set of biomarkers could be of additional value in identifying the relatively small group of patients at risk of developing malignancy in BE. In addition, in 56% of patients one or more biomarkers were abnormal prior to the development of LGD, suggesting a role in addition to dysplasia as marker for risk stratification. These biomarkers could also be helpful in improving the well-known interobserver variability in the histological differentiation of ND from LGD, since LGD was significantly more frequently observed when aneuploidy/tetraploidy, p53 overexpression, or strong (grade 2) Ki67 overexpression were present.

To our knowledge, this is the first study to evaluating Ki67 expression in a follow-up cohort study of BE patients developing HGD/EAC. Although this probably is the largest longitudinal study that evaluates various, relatively easy to apply biomarkers in the neoplastic progression of BE towards HGD/EAC, there are some limitations to our study. First, the number of patients included is still relatively small. This may have caused the plateau in the trend analysis in the period between 5 and 10 years before the development of HGD/EAC (Figure 1). We suppose that if a larger number of patients with consequently more biopsies had been available, the trend would have shown a more straight line over the years. This limitation is however hard to overcome, as in daily clinical practice up to 95% of patients in whom HGD/EAC is detected, BE was not previously diagnosed.[41, 42] Likewise, only 8% (27/355) of patients who developed HGD/EAC in our study participated in an endoscopic surveillance program for BE according to the recommendation of the ACG.[4] Secondly, although immunohistochemistry for detecting p53 is cheap, quick, and easy to apply compared to other techniques, there are also some drawbacks related to this method. The p53 antibody not only stains mutant p53, but also wild-type p53. Thus, overexpression of the p53 protein does not correlate with p53 mutation *per se*.[43] However, as the wild-type p53 has a short half-life, it is reasonable to assume that in most cases overexpression was caused by the mutated p53, which has a prolonged half-life.[44] In order to prevent overdiagnosis, we considered p53 expression up to 15% as normal. Another limitation is that a mutation in p53 may exist without protein overexpression, and will therefore remain undetected by immunohistochemistry.[43] Therefore, the results of our study comprise only IHC-detectable p53 mutations.

In conclusion, our results show that the grade of dysplasia in BE correlates with the proportion of

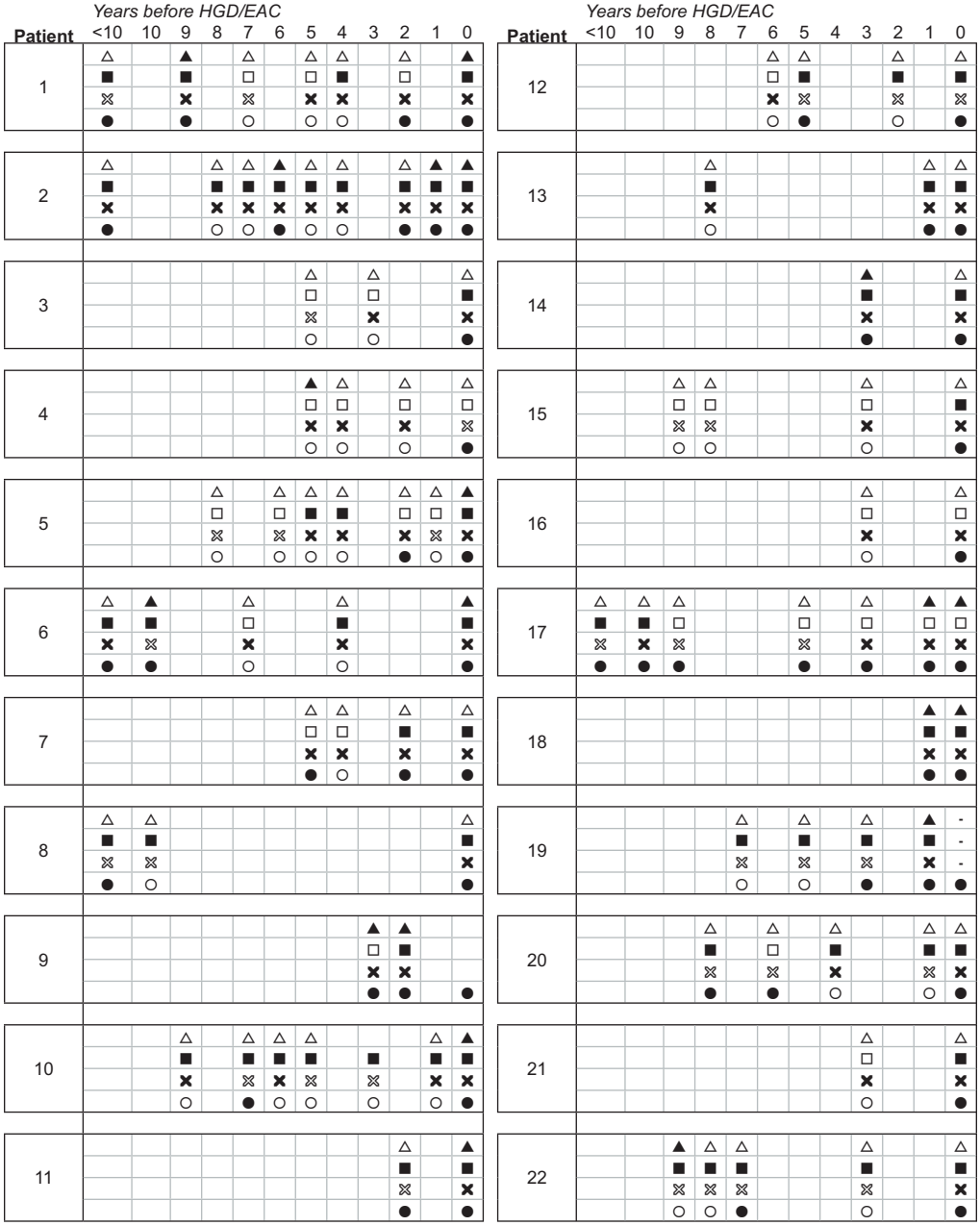
biopsies with abnormal ploidy status, and Ki67 and p53 overexpression. The increasing prevalence of Ki67 overexpression and, to a lesser extent, abnormal DNA-content over time, indicates that these markers may be useful in identifying patients at risk for neoplastic progression. Accumulation of p53, occurring several years before development of HGD/EAC and even often before the diagnosis of LGD, can be an early marker in predicting future progression in BE at a stage when other biomarkers are still negative. A case-control study or, preferably, a prospective follow-up study in a large cohort of patients is however warranted to confirm these findings.

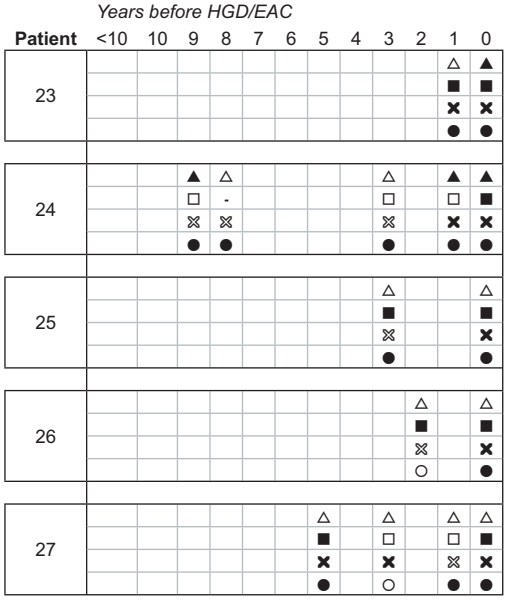
Acknowledgements

We thank Professor Fiebo J.W. ten Kate and Mr. Jos Mulder for their helpful suggestions to optimize the p53 immunostaining, and Enise Bagci for technical assistance.

APPENDIX

Results for histology grade (dysplasia), DNA-content detected with flow cytometry, p53 expression, and Ki67 expression, in individual patients with Barrett esophagus, prior to a diagnosis of HGD/EAC.





Time point zero indicates the moment of detection of HGD/EAC. If more than one endoscopy was performed >10 years before HGD/EAC was detected, biopsies with most abnormal biomarker-results are shown. Explanation of symbols: △, DNA-content normal (diploid); ▲, DNA-content abnormal (aneuploid/tetraploid); □, p53 expression normal; ■, p53 overexpression; ⊠, Ki67 expression normal; ✖, Ki67 overexpression; ○, no dysplasia; ●, dysplasia; -, not enough tissue available for evaluation.

REFERENCES

1. Guillem PG. How to make a Barrett esophagus: pathophysiology of columnar metaplasia of the esophagus. *Dig Dis Sci* 2005; 50: 415-24.
2. Avidan B, Sonnenberg A, Schnell TG*, et al.* Hiatal hernia size, Barrett's length, and severity of acid reflux are all risk factors for esophageal adenocarcinoma. *Am J Gastroenterol* 2002; 97: 1930-6.
3. Haggitt RC. Barrett's esophagus, dysplasia, and adenocarcinoma. *Hum Pathol* 1994; 25: 982-93.
4. Sampliner RE. Updated guidelines for the diagnosis, surveillance, and therapy of Barrett's esophagus. *Am J Gastroenterol* 2002; 97: 1888-95.
5. Conio M, Blanchi S, Lapertosa G*, et al.* Long-term endoscopic surveillance of patients with Barrett's esophagus. Incidence of dysplasia and adenocarcinoma: a prospective study. *Am J Gastroenterol* 2003; 98: 1931-9.
6. Hage M, Siersema PD, van Dekken H*, et al.* Oesophageal cancer incidence and mortality in patients with long-segment Barrett's oesophagus after a mean follow-up of 12.7 years. *Scand J Gastroenterol* 2004; 39: 1175-9.
7. Falk GW. Barrett's esophagus-is it bad for your health? *Am J Gastroenterol* 2005; 100: 2622-3.
8. Kerkhof M, van Dekken H, Steyerberg EW, Meijer GA, Mulder AH, de Bruine A, Driessen A, ten Kate FJ, Kusters JG, Kuipers EJ, Siersema PD. Grading of dysplasia in Barrett's oesophagus:substantial interobserver variation between general and gastrointestinal pathologists. *Histopathology* 2007 (in press).
9. Montgomery E, Bronner MP, Goldblum JR*, et al.* Reproducibility of the diagnosis of dysplasia in Barrett esophagus: a reaffirmation. *Hum Pathol* 2001; 32: 368-78.
10. Reid BJ, Blount PL, Rubin CE*, et al.* Flow-cytometric and histological progression to malignancy in Barrett's esophagus: prospective endoscopic surveillance of a cohort. *Gastroenterology* 1992; 102: 1212-9.
11. Reid BJ, Levine DS, Longton G*, et al.* Predictors of progression to cancer in Barrett's esophagus: baseline histology and flow cytometry identify low- and high-risk patient subsets. *Am J Gastroenterol* 2000; 95: 1669-76.
12. Merola E, Claudio PP, Giordano A. p53 and the malignant progression of Barrett's esophagus. *J Cell Physiol* 2006; 206: 574-7.
13. Ohbu M, Kobayashi N, Okayasu I. Expression of cell cycle regulatory proteins in the multistep process of oesophageal carcinogenesis: stepwise over-expression of cyclin E and p53, reduction of p21(WAF1/CIP1) and dysregulation of cyclin D1 and p27(KIP1). *Histopathology* 2001; 39: 589-96.
14. Ramel S, Reid BJ, Sanchez CA*, et al.* Evaluation of p53 protein expression in Barrett's esophagus by two-parameter flow cytometry. *Gastroenterology* 1992; 102: 1220-8.
15. Krishnadath KK, Tilanus HW, van Blankenstein M*, et al.* Accumulation of p53 protein in normal, dysplastic, and neoplastic Barrett's oesophagus. *J Pathol* 1995; 175: 175-80.
16. Younes M, Lechago J, Chakraborty S*, et al.* Relationship between dysplasia, p53 protein accumulation, DNA ploidy, and Glut1 overexpression in Barrett metaplasia. *Scand J Gastroenterol* 2000; 35: 131-7.
17. Murray L, Sedo A, Scott M*, et al.* TP53 and progression from Barrett's metaplasia to oesophageal adenocarcinoma in a UK population cohort. *Gut* 2006.
18. Scholzen T, Gerdes J. The Ki-67 protein: from the known and the unknown. *J Cell Physiol* 2000; 182: 311-22.
19. Reid BJ, Haggitt RC, Rubin CE*, et al.* Observer variation in the diagnosis of dysplasia in Barrett's esophagus. *Hum Pathol* 1988; 19: 166-78.
20. Kerkhof M, Bax DA, Moons LM*, et al.* Does CDX2 expression predict Barrett's metaplasia in oesophageal columnar epithelium without goblet cells? *Aliment Pharmacol Ther* 2006; 24: 1613-21.

21. Hage M, Siersema PD, Vissers KJ*, et al.* Molecular evaluation of ablative therapy of Barrett's oesophagus. *J Pathol* 2005; 205: 57-64.
22. Wang N, Pan Y, Heiden T*, et al.* Improved method for release of cell nuclei from paraffin-embedded cell material of squamous cell carcinomas. *Cytometry* 1993; 14: 931-5.
23. 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.
24. Lorinc E, Jakobsson B, Landberg G*, et al.* Ki67 and p53 immunohistochemistry reduces interobserver variation in assessment of Barrett's oesophagus. *Histopathology* 2005; 46: 642-8.
25. Feith M, Stein HJ, Mueller J*, et al.* Malignant degeneration of Barrett's esophagus: the role of the Ki-67 proliferation fraction, expression of E-cadherin and p53. *Dis Esophagus* 2004; 17: 322-7.
26. Halm U, Tannapfel A, Breitung B*, et al.* Apoptosis and cell proliferation in the metaplasia-dysplasia-carcinoma-sequence of Barrett's esophagus. *Hepatogastroenterology* 2000; 47: 962-6.
27. Olvera M, Wickramasinghe K, Brynes R*, et al.* Ki67 expression in different epithelial types in columnar lined oesophagus indicates varying levels of expanded and aberrant proliferative patterns. *Histopathology* 2005; 47: 132-40.
28. Hong MK, Laskin WB, Herman BE*, et al.* Expansion of the Ki-67 proliferative compartment correlates with degree of dysplasia in Barrett's esophagus. *Cancer* 1995; 75: 423-9.
29. Polkowski W, Baak JP, van Lanschot JJ*, et al.* Clinical decision making in Barrett's oesophagus can be supported by computerized immunoquantitation and morphometry of features associated with proliferation and differentiation. *J Pathol* 1998; 184: 161-8.
30. Baak JP, ten Kate FJ, Offerhaus GJ*, et al.* Routine morphometrical analysis can improve reproducibility of dysplasia grade in Barrett's oesophagus surveillance biopsies. *J Clin Pathol* 2002; 55: 910-6.
31. Fang M, Lew E, Klein M*, et al.* DNA abnormalities as marker of risk for progression of Barrett's esophagus to adenocarcinoma: image cytometric DNA analysis in formalin-fixed tissues. *Am J Gastroenterol* 2004; 99: 1887-94.
32. Teodori L, Gohde W, Persiani M*, et al.* DNA/protein flow cytometry as a predictive marker of malignancy in dysplasia-free Barrett's esophagus: thirteen-year follow-up study on a cohort of patients. *Cytometry* 1998; 34: 257-63.
33. Reid BJ, Sanchez CA, Blount PL*, et al.* Barrett's esophagus: cell cycle abnormalities in advancing stages of neoplastic progression. *Gastroenterology* 1993; 105: 119-29.
34. Montgomery EA, Hartmann DP, Carr NJ*, et al.* Barrett esophagus with dysplasia. Flow cytometric DNA analysis of routine, paraffin-embedded mucosal biopsies. *Am J Clin Pathol* 1996; 106: 298-304.
35. Gimenez A, Minguela A, Parrilla P*, et al.* Flow cytometric DNA analysis and p53 protein expression show a good correlation with histologic findings in patients with Barrett's esophagus. *Cancer* 1998; 83: 641-51.
36. Hedley DW. Flow cytometry using paraffin-embedded tissue: five years on. *Cytometry* 1989; 10: 229-41.
37. Younes M, Lebovitz RM, Lechago LV*, et al.* p53 protein accumulation in Barrett's metaplasia, dysplasia, and carcinoma: a follow-up study. *Gastroenterology* 1993; 105: 1637-42.
38. 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.
39. Symmans PJ, Linehan JM, Brito MJ*, et al.* p53 expression in Barrett's oesophagus, dysplasia, and adenocarcinoma using antibody DO-7. *J Pathol* 1994; 173: 221-6.
40. Gimenez A, de Haro LM, Parrilla P*, et al.* Immunohistochemical detection of p53 protein could improve the management of some patients with Barrett esophagus and mild histologic alterations. *Arch Pathol*

Lab Med 1999; 123: 1260-3.

41. Cameron AJ, Zinsmeister AR, Ballard DJ*, et al.* Prevalence of columnar-lined (Barrett's) esophagus. Comparison of population-based clinical and autopsy findings. *Gastroenterology* 1990; 99: 918-22.
42. Dulai GS, Guha S, Kahn KL*, et al.* Preoperative prevalence of Barrett's esophagus in esophageal adenocarcinoma: a systematic review. *Gastroenterology* 2002; 122: 26-33.
43. Keswani RN, Noffsinger A, Waxman I*, et al.* Clinical use of p53 in Barrett's esophagus. *Cancer Epidemiol Biomarkers Prev* 2006; 15: 1243-9.
44. Hinds PW, Finlay CA, Quartin RS*, et al.* Mutant p53 DNA clones from human colon carcinomas cooperate with ras in transforming primary rat cells: a comparison of the "hot spot" mutant phenotypes. *Cell Growth Differ* 1990; 1: 571-80.

7

The burden of endoscopy of upper gastrointestinal endoscopy in patients with Barrett esophagus

Endoscopy 2006 Sep; 38(9):873-8

Michelle E Kruijshaar [1]
Marjon Kerkhof [2]
Peter D Siersema [2]
Ewout W Steyerberg [1]
Marjolein Y V Homs [1]
Marie-Louise Essink-Bot [1]
for the CYBAR study group

Departments of [1]Public Health, and [2]Gastroenterology and Hepatology,
Erasmus MC - University Medical Center Rotterdam, Rotterdam, the Netherlands

ABSTRACT

Background and study aims: Patients with Barrett's esophagus (BE) are recommended to undergo regular surveillance with upper gastrointestinal (GI) endoscopy, an invasive procedure that may cause anxiety, pain and discomfort. We assessed patients' perceived burden of this procedure.

Patients and methods: A total of 192 patients with BE were asked to fill out questionnaires one week before, on the endoscopy day, one week after and one month after the endoscopy. Four variables were assessed: 1) pain and discomfort experienced during endoscopy, 2) symptoms, 3) psychological: anxiety, depression and distress levels (Hospital Anxiety and Depression scale, Impact of Event Scale), and 4) perceived risk of developing adenocarcinoma.

Results: At least one questionnaire was sent back by 180 patients (94%), 151 filled out all four (79%). Of all patients, only 14% experienced the endoscopy as painful. However, 59% reported it to be burdensome. Apart from an increase in throat ache (47 % after endoscopy versus 12% before) the procedure did not cause physical symptoms. Patients' anxiety, depression and distress levels were significantly increased in the week before the endoscopy compared to the week after. Patients perceiving their risk of developing adenocarcinoma as high reported higher levels of psychological distress and more burden from the procedure.

Conclusions: Upper GI endoscopy is burdensome for many BE patients and causes moderate distress. Perceiving a high risk of adenocarcinoma may increase distress and the burden experienced from the procedure. The benefits of endoscopic surveillance for BE patients should be weighted against its drawbacks, including the short-term burden for patients.

INTRODUCTION

The incidence of adenocarcinoma of the esophagus (EAC) has increased rapidly over the past two decades in most Western countries and it now comprises two-thirds of esophageal cancers.[1-4] A major risk factor for adenocarcinoma is Barrett's esophagus (BE), a condition in which the normal squamous epithelium of the distal esophagus is replaced by columnar epithelium of the intestinal type. The risk of developing EAC for patients with BE has been estimated to be approximately 0.5% per year.[5] In order to detect adenocarcinoma at an earlier stage with more potential for curative treatment, patients diagnosed with BE are recommended to adhere to endoscopic surveillance.[6]

Upper gastrointestinal (GI) endoscopy is an invasive procedure that may be associated with anxiety, pain and discomfort. The number of subjects experiencing these side-effects is much larger than the number experiencing potential health benefits of the procedure, as only few will develop adenocarcinoma.[5, 7-10] The extent to which patients with BE are burdened by regular upper gastrointestinal (GI) endoscopy is unclear, as data on anxiety, pain, discomfort and symptoms related to upper GI endoscopy have not been reported previously.

The value of this potentially burdensome surveillance is furthermore uncertain, as evidence that it prolongs survival is still lacking.[11-14] A risk of developing adenocarcinoma of about 0.5% per year [5] may be too low to allow surveillance to be cost-effective.[15, 16] Given the uncertain value of endoscopic surveillance, it becomes even more important to take the burden of endoscopic surveillance on patients with BE into account.

The aim of this study was to explore patients' perceived burden of upper GI endoscopy. Using questionnaires we assessed pain and discomfort, symptoms, psychological distress, and perceived risk of developing EAC.

PATIENTS AND METHODS

Patients

This questionnaire study was part of an ongoing clinical trial (CYBAR) assessing the value of flow cytometry for individualizing frequency of follow-up of upper GI endoscopy. Criteria to include prevalent and incident BE patients were: BE segment of 2 cm or more confirmed by intestinal metaplasia, absence of high-grade dysplasia and of carcinoma, ability to speak and read Dutch, informed consent and willingness to undergo follow-up. Three hospitals participated in the present questionnaire study: one academic teaching hospital and two regional hospitals. We included 197 patients between November 2003 and December 2004.

Questionnaires and measurements

Patients were asked to fill out questionnaires at four time points: one week before the endoscopy (baseline), at the endoscopy day (just before undergoing it), one week after and one month after the endoscopy. The different components are discussed below.

Pain and discomfort. Pain and discomfort experienced during the procedure were measured one week after the endoscopy. Items were adapted from earlier studies[17, 18] assessing five stages of

the procedure (receiving sedation, introduction of the endoscope, undergoing the endoscopy, removing the endoscope, period directly after endoscopy) in three response options (not, quite and very painful or unpleasant). Discomfort was additionally assessed for fasting and, if applicable, for waking up after sedation. Finally, we asked patients to rate the overall burden of the procedure (very, somewhat, not burdensome), and state the most and least burdensome parts of the procedure.

Symptoms. To detect whether the endoscopy caused physical symptoms, we compared the occurrence of ten symptoms at baseline and one week: throat ache, heartburn, regurgitation, flatulence or feeling bloated, vomiting, hematemesis, dysphagia of solid foods, of fluid foods, diarrhea, and constipation. Questions were composed in analogy to a previous study[18] using four answer categories (not at all, one day, 2-3 days, 4 or more days).

Psychological distress. At each time point the Hospital Anxiety and Depression scale (HAD), a validated self-report instrument with good reliability and a validity sufficient for screening, assessed anxiety (7 items) and depression (7 items).[19, 20] Scores per subscale range from 0-21, scores of 11 or higher indicating clinical, and 8-10 borderline anxiety or depression.[19, 20] Scores from a Dutch general population sample (n=1901) of similar age (average age of 61 year) and sex distribution (51% female) are available.[19]

At baseline and one week the Impact of Event Scale (IES) measured psychological distress reflected in intrusion (7 items) and avoidance (8 items) of thoughts and feelings.[21, 22] The total scale ranges between 0 and 75, with scores of 26 or over indicating a high risk of developing a stress disorder.[23] At baseline we assessed distress associated with the upper GI endoscopy itself and at one week distress concerning the biopsy result.

Perceived risk of developing EAC. We assessed patients' subjective evaluation of their risk of developing adenocarcinoma (do they perceive their risk as high or low?) as a potential determinant of the perceived burden, in seven response options (very small, small, quite small, not small or big, quite big, big, very big).[24]

Demographics and other data. Demographic data were collected at baseline, including patients' classification of their own health, using the EuroQol-5D.[25-27] The EuroQol-5D contains five items: mobility, self-care, usual activities, pain, and anxiety and depression with 3 response options (no, some, severe/ complete limitations).

Patients were asked whether this was their first, second or a later endoscopy. Information regarding the endoscopy (e.g. grade of dysplasia, sedation) was recorded.

Analyses

Categorical data were analyzed in SPSS version 11.0.1. Symptoms were compared before and after endoscopy using a method analogous to the Wilcoxon test: all responses were ranked and analysis of variance (ANOVA) was applied to the differences in these ranks. This model also allowed us to study effects of determinants (age, sex, hospital, previous endoscopies, sedation and dysplasia; stepwise inclusion $p<0.1$) and perceived risk (as continuous variable). Effects of determinants and perceived risk on reported pain and discomfort were studied using ANOVA after combining the items into pain and discomfort summary scores, by adding the item responses (no = 0, quite = 1,

very = 2). After comparing Crohnbach's alphas we included four items per score: introduction of the endoscope, undergoing the endoscopy, removal of the endoscope and the period immediately afterwards.

The continuous HAD scores were compared over time with repeated-measures ANOVA, using 'proc mixed' with REML and a compound symmetry covariance structure (this performed better than unstructured) in SAS version 8.2. The models comprised the main effects of time, possible determinants (stepwise inclusion), and interactions between a determinant and time. The effect of

Table 1. Patient details and baseline data

Characteristic (measured at baseline)	% (except where indicated)	n	Responses, n
Personal details			
Sex: male	66.1	119	180
Age, years	mean 61.9 (SD 12.1)		179
Marital status: married/ living with partner	76.6	134	175
Employment status			175
- Pensioner/ early retirement	49.7	87	
- In paid work	33.9	59	
- Unemployed	16.4	29	
Education			170
- Secondary	54.7	93	
- Elementary	18.8	32	
- Tertiary and postgraduate	23.5	40	
Hospital			180
- Ikazia Hospital	45.6	82	
- Deventer Hospital	33.9	61	
- Erasmus MC	20.6	37	
Clinical details			
Previous endoscopies, n			171
- 0	0.6	1	
- 1	15.2	26	
- 2 or more	84.2	144	
Sedation: yes	25.0	43	172
Dysplasia			153
- none	77.8	119	
- low-grade	22.2	34	
Reflux esophagitis	10.7	19	178
PPI use	91.6	164	179
General health, EuroQol			169-172
- No mobility problems	78.9	135	
- No self-care problems	97.0	164	
- No daily activities problems	79.5	136	
- No pain or discomfort	56.5	96	
- No anxiety or depression	83.7	144	
Perceived risk			167
- Very small - small	62.8	105	
- Quite high - high	1.2	2	

SD, standard deviation; PPI, proton pump inhibitor

perceived risk (continuous) was studied in a separate model (main effects of risk perception, time, and their interaction).

RESULTS

Response and respondent characteristics

Of all 197 patients approached, 192 were eligible for the study: two were excluded because of not having BE and three for not attending their appointment. Ninety-four percent of these patients (180) filled out at least one questionnaire, 79% (151) all four.

Table 1 shows the respondents' characteristics. The majority was male (66%), older than 60 years (59%) and had undergone more than one endoscopy prior to this one (84%). Most respondents had no dysplasia (78%), while 22% had low-grade dysplasia. The majority did not report general health problems other than pain/ discomfort (41.8% reported some pain or discomfort on the EuroQol-5D). Most respondents perceived their own risk of developing esophageal cancer as very small or small (62.8%), almost none as quite high or high (1.2%).

Pain and discomfort

In Table 2 we show the percentage of respondents experiencing discomfort and pain from different parts of the procedure. Only 14% of patients experienced pain during the procedure (8% to 19% depending on the item). However, the majority reported discomfort (63% from introducing the

Table 2. Pain and discomfort reported by patients, for different stages of the endoscopic surveillance procedure

Measure	Stage of the procedure (no. of responses)	Not		Quite		Very		NA
		n	%	n	%	n	%	
Pain	introduction scope (172)	146	84.8	24	14.0	2	1.2	
	undergoing (170)	137	80.6	28	16.5	5	2.9	
	remove scope (171)	157	91.8	11	6.4	3	1.8	
	period after (171)	145	84.8	22	12.9	4	2.3	
	sedation (172)	43	100.0	0	0.0	0		129
Discomfort	introduction scope (172)	64	37.2	81	47.1	27	15.7	
	undergoing (172)	75	43.6	72	41.9	25	14.5	
	remove scope (171)	144	84.2	24	14.0	3	1.8	
	period after (169)	136	80.5	29	17.2	4	2.4	
	sedation (172)	162	93.9	10	6.1	0		124
	discomfort waking up (170)	44	97.8	1	2.2	0		125
	fasting (171)	138	80.7	28	16.4	5	2.9	
Burdensome	total procedure (167)	68	40.7	87	52.1	12	7.2	

NA, not applicable

endoscope and 56% from the endoscopy itself) and assessed the procedure on the whole as 'burdensome' (59%). Approximately half of the respondents chose 'the introduction of the endoscope' as the most burdensome part of the procedure and overnight fasting as the least burdensome (46% and 53%, respectively).

Older patients reported less discomfort and burden (p=0.05 and 0.01, respectively). Reported discomfort was higher among patients who did not receive sedation (p=0.00), and differed between hospitals (p=0.00). No differences were found for education, between patients with no or low grade dysplasia, or patients with no, one or two or more previous endoscopies. A higher perceived risk of developing adenocarcinoma was associated with reporting more discomfort and burden from the procedure (p=0.02 and 0.01, respectively).

Symptoms

Only a few patients reported symptoms before or after the endoscopy. Symptoms that were reported by more than 20% of respondents at baseline were: heartburn (34%), regurgitation (45%) and flatulence (56%). After endoscopy also throat ache (47%) and dysphagia for solid foods (23%) were frequently reported. Except for an increase in throat ache (12% before and 47% after endoscopy, p = 0.00), the procedure did not cause symptoms.

Psychological burden

Mean scores and 95% confidence intervals of the anxiety and depression (HAD), and distress scales (IES), are shown for each time point in Figure 1. Before endoscopy, HAD and IES scores were significantly higher than one-week after, indicating more anxiety, depression, and distress

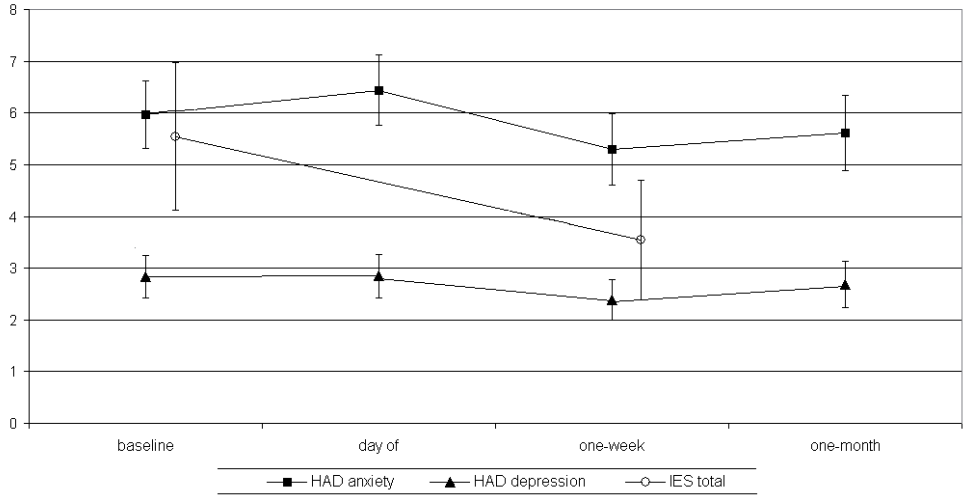


Figure1. Measurements at four time points using the Hospital Anxiety and Depression (HAD) scale, and at two time points using the Impact of Event Scale (IES). Mean scores and 95% confidence intervals are shown. The significant p values were as follows (corrected for significant determinants): HAD anxiety: baseline vs. 1 week, p=0.02; day of endoscopy vs. 1 week, p=0.00; day of endoscopy vs. 1 month, p=0.00. (Baseline vs. 1 month was borderline, p=0.075.) HAD depression: baseline vs. 1 week, p=0.00; day of endoscopy vs. 1 week, p=0.02. IES: baseline vs. 1 week, p=0.01.

before the endoscopy (for p-values see Figure 1). Anxiety scores were significantly higher than scores reported by the general population (mean anxiety scores of 6.0 and 6.4 before and 5.3 and 5.6 after endoscopy versus 3.90; p<0.0001). Depression scores were always significantly lower (2.8; 2.9; 2.4 and 2.7 versus 3.70, p=0.00). At the day of endoscopy, 23% of patients (39) had scores indicative of clinical anxiety and 17% (28) of borderline anxiety. Clinical depression scores were seen in 2% (3) and borderline in 5% (8). High distress scores were found in 6% (11).

At all time points, men reported less distress (IES, p=0.04). Patients with a high education showed lower depression scores (HAD, p=0.01). Between hospitals depression and distress differed (p=0.02 and 0.00, respectively). Scores did not differ by age, between first, second or later endoscopies, sedation, or grade of dysplasia. Respondents' perceived risk of developing EAC had a significant effect on the scores of all scales (anxiety p=0.01; depression <0.0001; distress 0.00), a lower perceived risk predicting lower anxiety, depression and distress. No significant interaction effects of any of the determinants with time were found, indicating that the pattern of the scores over time (e.g. higher depression scores before endoscopy) did not differ with perceived risk or any of the other determinants. This indicates that although levels of anxiety and depression differ by risk perception and some other determinants, their pattern over time does not.

DISCUSSION

This study showed that upper GI endoscopy associated with anxiety and distress before, and discomfort during the procedure. The procedure is not experienced as painful and causes few symptoms afterwards. Overall, the procedure was assessed as burdensome by over half the patients. In agreement with prior belief of physicians and nurses, the introduction of the endoscope was most often experienced as the most burdensome part of the procedure, while undergoing the endoscopy itself came second.

The procedure is clearly burdensome, as 60 percent of patients reported burden and discomfort. The higher levels of anxiety, depression and distress before the endoscopy may indicate that the procedure furthermore causes a psychological burden. However, these higher scores can also be explained by a positive reassurance effect of a favorable the test-result after the procedure. Nevertheless, we measured distress (IES) regarding the endoscopy itself before the endoscopy was performed and regarding the test-result afterwards. The higher distress scores before endoscopy thus indicate that patients were more distressed about undergoing the test than about the result of the biopsies. Also, most patients perceived their risk of developing esophageal cancer as very low. We therefore conclude that, although the procedure may also have a positive reassurance effect, the test definitely also has a negative psychological impact, causing anxiety and distress. The most important part of this psychological burden seems to be anxiety. Anxiety was significantly higher than in the general population, and nearly a quarter of patients had scores indicative of clinical anxiety at the day of endoscopy. Depression levels were lower than in the general population and hardly any patients had depression or distress levels in the clinical range.

Pain and discomfort from the procedure varied with age, hospital, whether sedation was given,

and perceived risk. Older patients reporting less discomfort and burden could possibly be explained by older patients having a less strong esophageal closure reflex, making it easier for the endoscope to be introduced.[28] Differences between hospitals may be caused by a variety of differences between patients, doctors and sedation in these hospitals. Sedation, as expected, decreased reported discomfort. Psychological scores also varied significantly with age, hospital and perceived risk, and by sex and education. These determinants did influence anxiety and distress levels in general, but not the pattern of anxiety and distress across measurements, i.e. the psychological burden of the endoscopy.

Patients who perceived their risk as higher were significantly more anxious and distressed at all measurements and reported more discomfort and burden from the endoscopy. It seems logical that the thought of having a high risk of developing cancer will give rise to worries and this distress is not necessarily related to the endoscopy itself, as argued before. An explanation for these patients also reporting more discomfort and burden may be found in their higher psychological distress. It is often suggested that patients who are able to relax during the endoscopy experience less pain and discomfort from it. To investigate this we studied the effect of the psychological scores (at baseline) on reported pain and discomfort from the procedure in an additional analysis. In fact, we found the opposite of what we expected: higher psychological scores were related to reporting *less* discomfort and burden from the procedure. Thus, it remains unexplained why patients with a higher risk perception report more discomfort and burden.

A drawback of this study is that only patients willing to undergo frequent surveillance were included in the study. This may have led to an underestimation of the burden. Patients not willing to adhere to (frequent) surveillance may be those patients who are very anxious and distressed and/ or patients who have experienced a very burdensome endoscopy in the past. Such patients are more likely to experience a higher burden. The fact that we found no effect of prior endoscopies on any of our outcome variables suggests that patients with a first or second endoscopy report the same burden. However, as the vast majority of patients had already experienced more than one endoscopy, we may have to study this in another population. Nevertheless, although all patients in this study were under surveillance, more than half of them experienced the endoscopy as burdensome, strengthening our conclusion that upper GI endoscopy is burdensome.

This study shows that a majority of patients with BE experience discomfort from upper GI endoscopy and are distressed beforehand. Although this may not seem like a high burden, a majority of patients experience the procedure as burdensome. Perceiving one's risk of developing adenocarcinoma of the esophagus as 'high' may increase a patient's distress and the discomfort he/she experiences from the procedure. Recommendations for endoscopic surveillance should take into account the short-term burden and distress of upper GI endoscopy for patients and studies aimed at tailoring the frequency of surveillance to individual patient characteristics are warranted.

ACKNOWLEDGEMENTS

We thank all staff at the departments of Gastroenterology and Hepatology of the three hospitals for their help, in particular we thank Corry Leunis, who coordinated the work in the Ikazia hospital. We are indebted to the gastroenterologists who performed the endoscopies and allowed this study to take place. Finally, we are grateful to Caspar Looman for his valuable statistical advice.

REFERENCES

1. Bollschweiler E, Wolfgarten E, Gutschow C*, et al.* Demographic variations in the rising incidence of esophageal adenocarcinoma in white males. *Cancer* 2001; 92: 549-55.
2. Botterweck AA, Schouten LJ, Volovics A*, et al.* Trends in incidence of adenocarcinoma of the oesophagus and gastric cardia in ten European countries. *Int J Epidemiol* 2000; 29: 645-54.
3. Bytzer P, Christensen PB, Damkier P*, et al.* Adenocarcinoma of the esophagus and Barrett's esophagus: a population-based study. *Am J Gastroenterol* 1999; 94: 86-91.
4. Devesa SS, Blot WJ, Fraumeni JF, Jr. Changing patterns in the incidence of esophageal and gastric carcinoma in the United States. *Cancer* 1998; 83: 2049-53.
5. Shaheen NJ, Crosby MA, Bozymski EM*, et al.* Is there publication bias in the reporting of cancer risk in Barrett's esophagus? *Gastroenterology* 2000; 119: 333-8.
6. Sampliner RE. Updated guidelines for the diagnosis, surveillance, and therapy of Barrett's esophagus. *Am J Gastroenterol* 2002; 97: 1888-95.
7. Cullen J, Schwartz MD, Lawrence WF*, et al.* Short-term impact of cancer prevention and screening activities on quality of life. *J Clin Oncol* 2004; 22: 943-52.
8. van der Burgh A, Dees J, Hop WC*, et al.* Oesophageal cancer is an uncommon cause of death in patients with Barrett's oesophagus. *Gut* 1996; 39: 5-8.
9. Hage M, Siersema PD, van Dekken H*, et al.* Oesophageal cancer incidence and mortality in patients with long-segment Barrett's oesophagus after a mean follow-up of 12.7 years. *Scand J Gastroenterol* 2004; 39: 1175-9.
10. Conio M, Blanchi S, Lapertosa G*, et al.* Long-term endoscopic surveillance of patients with Barrett's esophagus. Incidence of dysplasia and adenocarcinoma: a prospective study. *Am J Gastroenterol* 2003; 98: 1931-9.
11. Spechler SJ. Managing Barrett's oesophagus. *Bmj* 2003; 326: 892-4.
12. Chang JT, Katzka DA. Gastroesophageal reflux disease, Barrett esophagus, and esophageal adenocarcinoma. *Arch Intern Med* 2004; 164: 1482-8.
13. Tytgat GN, Van Sandick JW, Lanschot JJ*, et al.* Role of surveillance in intestinal metaplasia of the esophagus and gastroesophageal junction. *World J Surg* 2003; 27: 1021-5.
14. Craanen ME, Blok P, Meijer GA*, et al.* Surveillance in Barrett's oesophagus: a critical reappraisal. *Scand J Gastroenterol Suppl* 2002: 4-8.
15. Provenzale D, Kemp JA, Arora S*, et al.* A guide for surveillance of patients with Barrett's esophagus. *Am J Gastroenterol* 1994; 89: 670-80.
16. Inadomi JM, Sampliner R, Lagergren J*, et al.* Screening and surveillance for Barrett esophagus in high-risk groups: a cost-utility analysis. *Ann Intern Med* 2003; 138: 176-86.
17. Essink-Bot ML, Rijnsburger AJ, van Dooren S*, et al.* Women's acceptance of MRI in breast cancer surveillance because of a familial or genetic predisposition. *Breast* 2006; 15: 673-6.
18. Essink-Bot ML, de Koning HJ, Nijs HG*, et al.* Short-term effects of population-based screening for prostate cancer on health-related quality of life. *J Natl Cancer Inst* 1998; 90: 925-31.
19. Spinhoven P, Ormel J, Sloekers PP*, et al.* A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. *Psychol Med* 1997; 27: 363-70.
20. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983; 67: 361-70.
21. Brom D, Kleber RJ. De Schok Verwerkings Lijst. *Ned Tijdschr Psychol* 1985; 40: 164-68.
22. Horowitz M, Wilner N, Alvarez W. Impact of Event Scale: a measure of subjective stress. *Psychosom Med* 1979; 41: 209-18.
23. Chemtob CM, Tomas S, Law W*, et al.* Postdisaster psychosocial intervention: a field study of the

impact of debriefing on psychological distress. *Am J Psychiatry* 1997; 154: 415-7.
24. van Dooren S, Rijnsburger AJ, Seynaeve C*, et al.* Psychological distress in women at increased risk for breast cancer: the role of risk perception. *Eur J Cancer* 2004; 40: 2056-63.
25. TheEuroQolGroup. EuroQol-a new facility for the measurement of health-related quality of life. *Health Policy* 1990; 16: 199-208.
26. Brooks R. EuroQol: the current state of play. *Health Policy* 1996; 37: 53-72.
27. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997; 35: 1095-108.
28. Ren J, Xie P, Lang IM*, et al.* Deterioration of the pharyngo-UES contractile reflex in the elderly. Laryngoscope 2000; 110: 1563-6.

8 Summary and conclusions

SUMMARY

The aim of surveillance in patients with a Barrett esophagus (BE) is to detect progression of dysplasia at an early and therefore likely curable stage. The interval of endoscopic surveillance in patients with BE is currently based on the histopathological stage (i.e. grade of dysplasia). This approach is however known to have several pitfalls. First, only patients with intestinal metaplasia (IM) in the columnar-lined segment of the esophagus (CLE) have so far been regarded to have a premalignant condition and are enrolled in an endoscopic surveillance program, in contrast to patients with only cardiac-type mucosa (CM) in their biopsies. However, as IM and CM can be both present in the CLE and are endoscopically indiscernible, sampling error can occur, and exclusion of patients with only CM from endoscopic follow-up might therefore be incorrect. In addition, it has been suggested that IM in CLE may develop over time, and a follow-up endoscopy in the course of time may than detect IM. Secondly, in line with the risk of neoplastic progression, the presence or absence of dysplasia in IM determines the frequency of endoscopic surveillance, but the interpretation of dysplasia is subject to considerable interobserver variability, leading to both superfluous follow-up endoscopies in some patients and insufficient control of others. Therefore, it is relevant to perform risk stratification to define which subgroup of patients with CLE with or without IM should undergo endoscopic follow-up, and at which frequency.

The aim of the work described in this thesis was to assess the currently used criteria for performing endoscopic surveillance in patients with CLE, and to evaluate which clinical characteristics and biomarkers could contribute to risk stratification in patients with CLE, in order to refine surveillance strategies in these patients.

In **chapter 2**, the presence of various intestinal markers, i.e. CDX2, MUC2 and villin, in a CLE in the absence of a histological diagnosis of IM was evaluated. The aim of this study was to investigate whether these intestinal markers could predict the presence of undetected IM in CLE. In total, 122 biopsy sets of CLE from 61 patients were evaluated. CDX2 expression was detected in 23/67 (34%) samples with only CM. Detection of CDX2 in CM increased the likelihood of finding IM in another CLE biopsy set of the same patient (OR 3.5, 95% CI=1.2-10, p=0.02). On the basis of these data, we conclude that the presence of CDX2 in CM predicts the presence of IM in CLE, either simultaneously or over time. This suggests that CDX2 staining can be used as an additional marker for the presence of IM in CLE in the absence of goblet cells. A prospective follow-up study on patients with CM in their biopsies is however needed to confirm the predictive value of CDX2 in this regard.

In addition to predicting the presence of IM in CLE with histological markers, it is also of interest to determine if easily available clinical characteristics can be used to define which patients with CLE with or without IM should undergo endoscopic follow-up. In **chapter 3** a model based on clinical characteristics was developed to estimate the probabilities of the presence of IM and dysplasia in biopsies of CLE, regardless the histological results. In 908 patients with a CLE of ≥ 2 cm, data on age, gender, reflux symptoms, tobacco and alcohol use, medication use and upper gastrointestinal endoscopy findings were collected. Multivariable logistic regression analysis was performed to

develop a predictive model. Most important predictors for the presence of IM were length of CLE, size of hiatal hernia and male gender, while among those with IM, age, and male gender were most important for the presence of LGD. These results further showed that the proposed predictive models were able to estimate the histology of biopsies (especially IM), based on easily available clinical predictors. The predicted probabilities from these models may aid in the decision-making on whether surveillance should be performed in a patient with CLE in view of the known sampling error at endoscopy and interobserver variability at histology.

It has been reported that despite the use of established criteria for grading dysplasia in BE, a considerable interobserver variability between pathologists remains. This interobserver variation has however only been assessed between expert gastrointestinal (GI) pathologists. As in daily practice general pathologists review the majority of BE biopsies, we compared in **chapter 4** the interobserver variability in establishing the grade of dysplasia in BE between non-expert general pathologists and expert gastrointestinal pathologists on the one hand, and between expert gastrointestinal pathologists on the other hand. In this prospective multicenter study, non-expert and expert pathologists assessed esophageal biopsies of 920 patients with endoscopic BE. These biopsies were blindly reviewed by an expert gastrointestinal pathologist, and the slides were reviewed by a second expert in case of disagreement on the presence of IM and/or dysplasia grade. The results of this study showed a fair ($\kappa = 0.24$) interobserver agreement in the distinction between ND and IND/LGD, and a substantial ($\kappa = 0.62$) interobserver agreement, although still not perfect, in the distinction between HGD/EAC and lower stages. Remarkably, the interobserver variability was not different between non-expert and expert pathologists and between expert pathologists, implicating that the subjective component in the evaluation of dysplasia in BE can hardly be improved if only expert pathologists would review biopsies from CLE. Although morphology is quite subjective, it is presently the most commonly used marker to determine cancer risk. It is our opinion that additional, less subjective markers are required to establish cancer risk in patients with BE, in order to determine the optimal surveillance frequency.

Several studies have been performed to find an ideal biomarker for improving the risk stratification of BE patients. In **chapter 5** the existing literature was reviewed regarding the so far evaluated candidate biomarkers for improving risk stratification of patients with BE. Of these biomarkers, DNA ploidy, p53, and Ki67, seems most promising in being an additional marker or even a substitute for histology. As large prospective follow-up studies investigating the clinical value of these biomarkers have not been performed, these biomarkers are not ready for this purpose yet.

In **chapter 6** the most promising biomarkers, i.e. Ki67, p53, and DNA ploidy, were examined in more detail regarding their usefulness in identifying BE patients at highest risk for subsequent progression to EAC. The grade of dysplasia was determined in 212 biopsy sets taken during follow-up upper endoscopies in 27 patients in whom ultimately HGD or EAC was detected. Mean follow-up time before HGD or EAC was detected was 75 ± 51 months (range 4-187 months). The grade of dysplasia in BE highly correlated with the proportion of biopsies with abnormal ploidy status, Ki67 and p53 overexpression. Samples taken at a closer time point prior to the detection of HGD/EAC were more prone to contain LGD and Ki67 overexpression, and, to a lesser extent, to be aneuploid

or tetraploid, suggesting the value of these biomarkers in identifying patients at increased risk for neoplastic progression. Accumulation of p53 occurred up to 15 years prior to the development of HGD/EAC and was thus found to be a possible early marker in predicting future progression in BE at a stage when other biomarkers are still negative. Large, prospective follow-up studies are needed to confirm these findings.

The above-described chapters were directed to the improvement of identifying patients at risk for neoplastic progression in order to find an optimal endoscopic surveillance strategy. Upper gastrointestinal endoscopy is however an invasive procedure that may be associated with anxiety, pain and discomfort. The extent to which patients with BE are burdened by regular upper endoscopy is unclear. In **chapter 7**, the perceived burden of upper gastrointestinal endoscopy in BE patients was explored with questionnaires. Risk perception was taken into account as a potential determinant of a patients' perception of the burden of endoscopy. In total, 197 patients with BE were included in this questionnaire study. The results of this study showed that upper endoscopy caused anxiety and distress beforehand, probably because of the prospect of the endoscopic procedure. Although discomfort was experienced during the procedure, the endoscopy was usually not reported as painful and caused few symptoms afterwards. Patients with higher risk perceptions reported significantly more discomfort and burden from the endoscopy. Overall, the procedure was assessed as burdensome by over half the patients, and patients were more concerned about undergoing the endoscopy itself than about the possible negative result of histological evaluation of the biopsies.

CONCLUSIONS

Risk stratification in patients with BE is important for determination of the optimal surveillance strategy, which should result in the detection of neoplastic progression at an early, curable stage at the one hand, at limited costs and patient burden on the other hand. In this thesis the pitfalls of the currently used method of determining endoscopic surveillance intervals in patients with BE are discussed, and suggestions for improving these strategies are proposed.

The guidelines of the American College of Gastroenterology,[1] in which the presence of IM in CLE is required for being considered to have a premalignant condition and subsequently being enrolled in a surveillance program, are probably too strict. The British Society of Gastroenterology instead recently recommended endoscopic surveillance in all patients with CLE in the esophagus.[2] This will however cause an increase in work-load and a decrease of cost-effectiveness of surveillance. In this thesis, an alternative strategy is proposed, in which IM is still regarded as the type of epithelium that is most likely to undergo neoplastic progression, but CM is not immediately excluded from endoscopic follow-up due to the possibility of having undetected IM (due to sampling error or development over time) in CLE. The detection of the intestinal marker CDX2 in CM and a high probability of having IM in CLE based on the predictive model reported in chapter 3, can be used to decide which patients with CM should undergo a follow-up endoscopy.

Patients who require endoscopic follow-up, can be stratified into groups with different risks for malignant progression by using the clinical predictors from the predictive model for the presence of

LGD (chapter 3), and the presence of biomarkers such as p53, Ki67 and the flow cytometric finding of aneuploidy/tetraploidy in biopsies from CLE. The clinical usefulness of these biomarkers and the consequences for the frequency of surveillance intervals need however further investigation in large prospective studies before they can be used in clinical practice.

The future of improving risk stratification in BE will probably come from the use of panels of biomarkers. Further research in combinations of biomarkers should preferably be performed in a prospective, multicenter setting, as larger cohorts of BE patients are needed for an analysis with sufficient statistical power. In addition, new endoscopy techniques that may have the capacity of visualizing dysplasia, such as autofluorescence endoscopy, high resolution endoscopy or magnification chromendoscopy, could make targeted biopsies possible, which will consequently result in a reduction of sampling error. The first results with these techniques seem promising for the visualization of areas with HGD in CLE, the more difficult diagnosis LGD seems however still difficult to visualize.[3-5]

REFERENCES

1. Sampliner RE. Updated guidelines for the diagnosis, surveillance, and therapy of Barrett's esophagus. *Am J Gastroenterol* 2002; 97: 1888-95.
2. Playford RJ. New British Society of Gastroenterology (BSG) guidelines for the diagnosis and management of Barrett's oesophagus. *Gut* 2006; 55: 442.
3. Kuipers EJ, Haringsma J. Diagnostic and therapeutic endoscopy. *J Surg Oncol* 2005; 92: 203-9.
4. Sharma P, Weston AP, Topalovski M*, et al.* Magnification chromoendoscopy for the detection of intestinal metaplasia and dysplasia in Barrett's oesophagus. *Gut* 2003; 52: 24-7.
5. Kara MA, Peters FP, Ten Kate FJ*, et al.* Endoscopic video autofluorescence imaging may improve the detection of early neoplasia in patients with Barrett's esophagus. *Gastrointest Endosc* 2005; 61: 679-85.

Samenvatting en conclusies

SAMENVATTING

Het doel van surveillance bij patiënten met een Barrett-oesofagus (BO) is het in een vroeg stadium detecteren van progressie van dysplastische afwijkingen waardoor er meer curatieve mogelijkheden zijn. Het interval van endoscopische surveillance bij patiënten met BO is momenteel gebaseerd op het histopathologische stadium (de graad van dysplasie) van het Barrett epitheel. Deze benadering heeft echter enkele valkuilen. Ten eerste wordt alleen de aanwezigheid van intestinale metaplasie (IM) in het cilindrisch epitheel van de oesofagus beschouwd als een premaligne aandoening. Daarom doorlopen alleen patiënten met IM een endoscopisch surveillance programma, dit in tegenstelling tot patiënten met cardia-type mucosa (CM) in hun biopten. Zowel IM als CM kunnen in het cilindrisch epitheel aanwezig zijn. Deze typen epitheel zijn endoscopisch echter niet te onderscheiden, waardoor er sampling error kan optreden. Hierdoor worden patiënten met alleen CM mogelijk ten onrechte uitgesloten voor follow-up. Daarnaast zou IM zich in de loop van de tijd kunnen ontwikkelen in het cilindrisch epitheel waarin eerst alleen CM aanwezig was. Ten tweede wordt aan de hand van de graad van dysplasie het risico op neoplastische progressie in BO ingeschat, wat vervolgens de frequentie van endoscopische surveillance bepaald. Het is echter gebleken dat de interpretatie van dysplasie onderhevig is aan een aanzienlijke interobserver variabiliteit, wat kan leiden tot zowel te frequente follow-up endoscopieën als ook tot te weinig surveillance. Om deze reden is het relevant om met behulp van risico stratificatie te definiëren welke subgroep van patiënten met cilindrisch epitheel in de oesofagus (met of zonder IM) endoscopische follow-up zouden moeten ondergaan en met welke frequentie.

Het doel van het onderzoek dat wordt beschreven in dit proefschrift was om de huidige criteria voor endoscopische surveillance van patiënten met cilindrisch epitheel in de oesofagus te beoordelen. Tevens werd geëvalueerd welke klinische karakteristieken en biomarkers zouden kunnen bijdragen aan risico stratificatie, met als doel de surveillance strategie voor deze patiënten te verfijnen.

In **hoofdstuk 2** werd gekeken naar de aanwezigheid van intestinale markers als CDX2, MUC2 en villin in cilindrisch epitheel, zonder een histologische diagnose van IM. Het doel van deze studie was het onderzoeken van de bruikbaarheid van deze intestinale markers om de aanwezigheid van niet-gedetecteerde IM in cilindrisch epitheel te voorspellen. In totaal werden 122 samples met biopten afkomstig van 61 patiënten geëvalueerd. CDX2 expressie werd vastgesteld in 23/67 (34%) samples waarin alleen CM was aangetroffen. De detectie van CDX2 in CM bleek de kans op het aantreffen van IM in andere biopten van dezelfde patiënt te verhogen (OR 3,5; 95% CI=1,2-10; p=0,02). Op basis van deze gegevens werd geconcludeerd dat de aanwezigheid van CDX2 in CM de aanwezigheid van IM in cilindrisch epitheel (gelijktijdig of in de loop van de tijd) kan voorspellen. Dit suggereert dat de CDX2 kleuring kan worden gebruikt als aanvullende marker, om de aanwezigheid van IM in cilindrisch epitheel zonder slijmbekercellen aan te tonen. Een prospectieve follow-up studie naar patiënten met CM in hun biopten is echter geïndiceerd om de voorspellende waarde van CDX2 op de aanwezigheid van IM te bevestigen.

Naast het voorspellen van de aanwezigheid van IM in cilindrisch epitheel met behulp van histologische markers, is het eveneens interessant om te bepalen of eenvoudig beschikbare

klinische karakteristieken gebruikt kunnen worden om te bepalen welke patiënten met cilindrisch epitheel met of zonder IM endoscopische follow-up zouden moeten ondergaan. In **hoofdstuk 3** is een model ontwikkeld, dat op basis van klinische karakteristieken (zonder gebruik van histologie) kan inschatten wat de waarschijnlijkheid is dat IM of dysplasie aanwezig is in cilindrisch epitheel. Van 908 patiënten met cilindrisch epitheel in de oesofagus met een lengte van ≥ 2 cm werden gegevens verzameld zoals leeftijd, geslacht, reflux-symptomen, roken, alcoholgebruik, medicatie en gastroscopie-bevindingen. Multivariabele logistische regressie werd gebruikt om het predictie-model te ontwikkelen. De belangrijkste predictoren voor de aanwezigheid van IM zijn de lengte van het cilindrisch segment, de grootte van de hiatus hernia en mannelijk geslacht. Voor de aanwezigheid van laaggradige dysplasie (LGD) in IM zijn dit leeftijd en mannelijk geslacht. De resultaten lieten verder zien dat de voorgestelde predictie-modellen in staat zijn om met behulp van deze eenvoudig verkrijgbare klinische predictoren de histologische bevindingen (met name IM) in biopten uit de oesofagus te voorspellen. Met het oog op het risico van sampling error bij gastroscopie en interobserver variabiliteit bij histologie, zouden deze predictie-modellen nuttig kunnen zijn bij de besluitvorming of surveillance zou moeten plaatsvinden bij een patiënt met cilindrisch epitheel in de oesofagus.

Het is bekend dat, ondanks gevestigde criteria voor het graderen van dysplasie in BO, er een aanzienlijke interobserver variabiliteit tussen pathologen bestaat. Deze interobserver variabiliteit is echter alleen onderzocht tussen expert gastro-intestinale pathologen. Aangezien in de dagelijkse praktijk vooral algemeen pathologen de BO biopten beoordelen, is in **hoofdstuk 4** de interobserver variabiliteit vergeleken tussen zowel niet-expert en expert pathologen, als ook tussen expert pathologen onderling. In deze prospectieve multicentrische studie werden door niet-expert en expert pathologen oesofagusbiopten van 920 patiënten met een endoscopische BO beoordeeld. Deze biopten werden zonder voorkennis herbeoordeeld door een expert patholoog. Indien er geen overeenstemming was betreffende de aanwezigheid van IM en/of de graad van dysplasie, werden de coupes vervolgens beoordeeld door een tweede expert patholoog. De resultaten van deze studie laten zien dat er sprake is van een slechte interobserver overeenstemming (κ=0,24) in het onderscheid tussen geen dysplasie (ND) en IND (‘indefinite for dysplasia’)/LGD. Daarentegen is er een redelijke, maar zeker nog geen perfecte, interobserver overeenstemming (κ=0,62) in het onderscheid tussen HGD (hooggradige dysplasie)/EAC (oesofagus adenocarcinoom) enerzijds en lagere stadia van dysplasie anderzijds. Opvallend is dat er geen verschil is in de interobserver variabiliteit tussen niet-expert en expert pathologen en tussen expert pathologen onderling, wat impliceert dat de subjectieve component in de evaluatie van dysplasie in BO nauwelijks zou kunnen worden verbeterd wanneer alleen expert pathologen de biopten zouden beoordelen. Hoewel het histologisch onderzoek subjectief is, is het op dit moment de meest gebruikte marker om het risico op neoplasie te bepalen. Op grond van deze bevindingen kan geconcludeerd worden dat aanvullende, minder subjectieve markers nodig zijn om het risico op neoplasie en daarmee het surveillance interval te bepalen bij patiënten met BO.

Verscheidene studies zijn verricht om een biomarker te vinden waarmee risico stratificatie van patiënten met BO kan worden verbeterd. In **hoofdstuk 5** wordt beschreven wat er in de bestaande

literatuur bekend is over mogelijke kandidaat biomarkers. Van de beschreven biomarkers lijken DNA ploïdie, p53 en Ki67 het meest veelbelovend te zijn als aanvullende marker, of mogelijk zelfs als substituut voor histologie. Aangezien er nog geen grote prospectieve follow-up studies verricht zijn waarin de klinische waarde van deze biomarkers wordt onderzocht, zijn deze biomarkers nog niet klaar om gebruikt te worden voor bovenstaand doel.

In **hoofdstuk 6** zijn de meest veelbelovende biomarkers Ki67, p53 en DNA ploïdie onderzocht op hun bruikbaarheid bij het identificeren van BO patiënten met een hoog risico op progressie richting EAC. De dysplasie-graad werd vastgesteld in 212 follow-up biopten van 27 patiënten die in de loop der tijd HGD of EAC hadden ontwikkeld. De gemiddelde follow-up duur voordat HGD/EAC werd gedetecteerd was 75 ± 52 maanden (range 4-187 maanden). De dysplasie-graad in BO was zeer sterk gecorreleerd met het percentage biopten met een abnormale ploïdie-status, en met Ki67- en p53-overexpressie. Biopten afgenomen op een tijdstip dichter bij het moment waarop HGD/EAC werd gedetecteerd, lieten vaker LGD, Ki67-overexpressie en in minder mate ook aneuploïdie of tetraploïdie zien. Dit suggereert de waarde van deze biomarkers bij het identificeren van patiënten met een verhoogd risico op neoplastische progressie. Aangezien p53 overexpressie tot 15 jaar voor de ontwikkeling van HGD/EAC kan worden aangetoond, wordt p53 beschouwd als een potentiële vroege marker voor het voorspellen van progressie in BO in een stadium waarbij andere biomarkers nog negatief zijn. Grote prospectieve studies zijn echter noodzakelijk om deze bevindingen te bevestigen.

De hierboven beschreven hoofdstukken zijn gericht op de verbetering van het identificeren van patiënten *at risk* voor neoplastische progressie, met als doel de optimale endoscopische surveillance strategie vast te stellen. Een gastroscopie is echter een invasieve procedure die mogelijk is geassocieerd met angst, pijn en ongemak. De mate waarin patiënten met BO een gastroscopie als een belasting ervaren is onbekend. In **hoofdstuk 7** werd de last die ervaren werd tijdens de gastroscopie door patiënten met BO nagegaan met behulp van vragenlijsten. Risico perceptie werd beschouwd als een potentiële determinant van de ervaren belasting van een gastroscopie. In totaal werden 197 patiënten geïncludeerd in deze vragenlijst-studie. De resultaten van deze studie laten zien dat een gastroscopie reeds voorafgaand aan het onderzoek angst en ongemak veroorzaakt, waarschijnlijk door het vooruitzicht van het moeten ondergaan van de endoscopische procedure. Hoewel de procedure zelf als onprettig werd ervaren, werd de gastroscopie meestal niet gerapporteerd als pijnlijk en veroorzaakte het weinig klachten achteraf. Patiënten die hun eigen risico op neoplastische progressie hoog inschatten, rapporteerden meer ongemak en last van de gastroscopie. In het geheel werd de procedure als belasting ervaren door meer dan de helft van de patiënten. Daarnaast waren patiënten meer bezorgd over het ondergaan van de gastroscopie zelf, dan over de mogelijke negatieve uitkomsten van de histologische beoordeling van de biopten.

CONCLUSIES

Risico stratificatie bij patiënten met BO is belangrijk voor het vaststellen van de optimale surveillance strategie, wat zou moeten resulteren in de detectie van neoplastische progressie in een vroeg, en

daardoor mogelijk curatief stadium aan de ene kant, en in beperkte kosten en zo min mogelijk ongemak voor de patiënt aan de andere kant. In dit proefschrift worden de valkuilen in de huidig gebruikte methoden voor het bepalen van de endoscopische surveillance intervallen bediscussieerd en worden suggesties voor het verbeteren van deze strategieën voorgesteld.

Er kan geconcludeerd worden dat de richtlijnen van het American College of Gastroenterology (ACG)[1] mogelijk te strikt zijn, omdat daarin de aanwezigheid van IM in cilindrisch epitheel van de oesofagus als vereiste wordt gegeven voor het hebben van een premaligne aandoening. Alleen deze patiënten worden geïncludeerd in een surveillance programma. In tegenstelling tot de ACG heeft de Britsh Society of Gastroenterology[2] recent aanbevolen om bij alle patiënten met cilindrisch epitheel in de oesofagus endoscopische surveillance te verrichten. Dit zal echter een verhoging van de werkdruk geven met een verminderde kosteneffectiviteit van het surveillance programma. In dit proefschrift wordt een alternatieve aanpak voorgesteld. Hierbij wordt IM nog steeds beschouwd als het type epitheel dat het meest waarschijnlijk neoplastische progressie ondergaat, maar wordt CM niet direct uitgesloten van endoscopische follow-up, aangezien er sprake kan zijn van niet-gedetecteerde IM (door sampling error of door ontwikkeling van IM in de loop van de tijd) in cilindrisch epitheel. Zowel de aanwezigheid van de intestinale marker CDX2 in CM als ook een hoge waarschijnlijkheid op de aanwezigheid van IM in cilindrisch epitheel volgens het predictie-model zoals beschreven in hoofdstuk 3, kunnen als indicatie worden gebruikt om te bepalen welke patiënten met CM endoscopische follow-up zouden moeten ondergaan.

Patiënten bij wie endoscopische follow-up is geïndiceerd, kunnen worden gestratificeerd in groepen met verschillende risicograden van kans op maligne progressie in BO. Dit kan met behulp van de klinische predictoren betreffende de aanwezigheid van LGD (hoofdstuk 3) en de aanwezigheid van de biomarkers p53, Ki67 en DNA-ploïdie in biopten uit cilindrisch epitheel. De klinische bruikbaarheid van deze biomarkers en de consequenties voor de frequentie van surveillance intervals behoeven echter verder onderzoek in grote prospectieve studies, voordat ze kunnen worden toegepast in de klinische praktijk.

In de toekomst zullen waarschijnlijk panels met meerdere biomarkes moeten worden gebruikt ter verbetering van de risico stratificatie van patiënten met BO. Toekomstig onderzoek naar mogelijke combinaties van deze biomarkers zou bij voorkeur in een prospectief multicentrisch studieverband uitgevoerd moeten worden, omdat statistische analyses met voldoende power relatief veel patiënten met BO eisen. Daarnaast lijken nieuwe endoscopische technieken, waarmee dysplasie gevisualiseerd kan worden, veelbelovend. Hierdoor is gerichter biopteren mogelijk, waardoor de sampling error zou kunnen worden gereduceerd. Ondanks het feit dat de eerste resultaten met deze technieken bij het visualiseren van gebieden met HGD in cilindrisch epitheel veelbelovend lijken te zijn, blijkt het voor de histologisch moeilijker diagnose LGD nog niet behulpzaam te zijn.[3-5]

REFERENTIES

1. Sampliner RE. Updated guidelines for the diagnosis, surveillance, and therapy of Barrett's esophagus. *Am J Gastroenterol* 2002; 97: 1888-95.
2. Playford RJ. New British Society of Gastroenterology (BSG) guidelines for the diagnosis and management of Barrett's oesophagus. *Gut* 2006; 55: 442.
3. Kuipers EJ, Haringsma J. Diagnostic and therapeutic endoscopy. *J Surg Oncol* 2005; 92: 203-9.
4. Sharma P, Weston AP, Topalovski M*, et al.* Magnification chromoendoscopy for the detection of intestinal metaplasia and dysplasia in Barrett's oesophagus. *Gut* 2003; 52: 24-7.
5. Kara A, Peters FP, Ten Kate FJ*, et al.* Endoscopic video autofluorescence imaging may improve the detection of early neoplasia in patients with Barrett's esophagus. *Gastrointest Endosc* 2005; 61: 679-85.

Dankwoord

DANKWOORD

Hier ligt mijn proefschrift dan. Tot niet zo lang geleden kon ik me niet voorstellen dat het ooit af zou komen. Vele, vele uren zitten er in en ik ben er dan ook erg trots op. Dit proefschrift was echter nooit tot stand gekomen zonder de steun en interesse van velen, waarvoor iedereen héél hartelijk dank! Een aantal mensen wil ik graag specifiek noemen.

Allereerst wil ik mijn beide copromotoren, Peter Siersema en Hans Kusters, bedanken. Het was voor mij behoorlijk wennen om vanuit de kliniek ineens in het onderzoekswereldje terecht te komen, op een laboratorium nog wel, waarin ik me in eerste instantie als met twee linkerhanden voelde. Gelukkig groeide ik er dankzij jullie begeleiding razendsnel in. Beste Peter, ik heb ontzettend bewondering voor je tomeloze inzet, kennis en snelheid waarmee jij artikelen nakijkt, terwijl je nog 10 andere promovendi te begeleiden hebt! Altijd als ik dacht, dit wordt niks, wist jij mijn pessimistische gedachten weer om te switchen en zag ik het weer rooskleuriger in. En zie hier het resultaat! Beste Hans, ook jij was er altijd bij bij de vele werkbesprekingen, met vaak goede ideeën. Artikelen kwamen vaak behoorlijk rood weer bij me terug, echter wel met goede suggesties, wat het eindresultaat zeer zeker ten goede kwam.

Natuurlijk wil ik ook mijn promotor, Ernst Kuipers, heel hartelijk bedanken voor het creëren van de gelegenheid om promotie-onderzoek te mogen verrichten binnen de afdeling maag-darm-leverziekten. Beste Ernst, eens in de zoveel tijd ging ik de ‘stand-van-zaken’ met je doorspreken, waarna ik door je interesse en positief ingestelde woorden altijd weer met een goed gevoel wegging. Tevens wil ik je bedanken voor het vertrouwen in mij als toekomstig MDL-arts en ik kijk er dan ook naar uit om de kneepjes van het vak binnen jouw afdeling te mogen leren.

Geen enkel artikel kan zonder goede statistiek. En geen goede statistiek zonder goede statisticus, dus daarom wil ik graag Ewout Steyerberg bedanken. Ik heb mezelf veel aangeleerd in SPSS, maar het was altijd erg prettig om even te checken of ik de juiste statistische tests had gebruikt. Beste Ewout, hartelijk dank hiervoor, maar niet in de minste plaats ook voor de ingewikkelde smoothing splines, predictie-modellen en andere ‘hogere’ statistiek, wat ik echt niet zelf had gekund.

Mijn promotie-onderzoek is begonnen met het opstarten van de CYBAR-studie, waarbij in vele ziekenhuizen door het hele land veel Barrett-patiënten zijn geïncludeerd door MDL-artsen en research-verpleegkundigen en vele histologie-coupes zijn bekeken door de pathologen van deze ziekenhuizen. Hoewel deze studie uiteindelijk niet in mijn proefschrift terecht is gekomen, aangezien de follow-up nog niet is afgerond, wil ik toch graag al deze mensen heel hartelijk danken voor de fijne samenwerking (zie Appendix). Zonder jullie zouden we nooit zo’n groot Barrett-cohort hebben kunnen opzetten, wat erg belangrijk is voor een goed onderzoek. Tevens wil ik specifiek alle leden van het CYBAR-panel noemen, de pathologen die keer op keer al die ‘vreselijke’ doosjes vol met coupes van mij toegestuurd kregen om ze te herbeoordelen. Zonder jullie zou dit onderzoek niet zo goed zijn geweest als dat het nu is. Herman van Dekken, Dries Mulder, Gerrit Meijer, Adriaan de Bruïne, Ann Driessen, Arend Karrenbeld en Fiebo ten Kate, hartelijk bedankt!

Ook wil ik iedereen van het MDL-lab bedanken, mijn werkbasis van de afgelopen 3 jaar. Linda, jij bent een fantastische secretaresse en een goed luisterend oor. Jij hebt je hart echt

op de goede plek, dank je wel! Hanneke, bedankt voor het mij wegwijs maken in het lab. Die twee linkerhanden waren snel weer een linker en een rechter! En natuurlijk ook bedankt voor het meedenken tijdens en buiten alle werkbesprekingen. Enise, jou wil ik bedanken voor het uitvoeren van de vele immunohistochemische kleuringen in het kader van je afstudeerproject, maar ook van een stukje van mijn onderzoek. Özlem, het was erg fijn dat je me een tijd kwam helpen met het invoeren van die honderden inclusie-formulieren. Michelle, onze samenwerking heeft tot een mooi hoofdstuk in dit proefschrift geleid. Anthonie, ik vind het erg leuk dat we samen zijn gaan salsa-dansen. Inmiddels zijn we vergevorderd (ongelofelijk, maar waar) en hebben we andere collega's aangestoken. Bedankt voor de gezellige salsa-uurtjes! Aan alle MDL-lab-collega's: jullie bleken een bruisend vat vol gezelligheid te zijn. Alle vrijdagmiddagborrels in Dizzy, karaoke-avonden in het parlement, MDL-labdagen en overige leuke dingen waren een welkome afwisseling van de serieuze aangelegenheden. Iedereen erg bedankt hiervoor!

Helaas zaten we niet met alle promovendi op één afdeling, maar zat er nog een delegatie op de dakpoli en op de flex-plekken, waardoor we elkaar helaas niet vaak zagen. Gelukkig werd er af en toe geborreld en hadden we congressen, waardoor we toch weer gezellig konden bijkletsen. Lieve Jolanda, jij hebt je promotie al achter de rug. Ik ben zwaar onder de indruk hoe ontiegelijk hard jij hebt gewerkt om dat superboekje van je in zo'n korte tijd af te krijgen. En daarnaast nog altijd in de stemming voor een borrel! Lieve Sanna, samen co-schappen gedaan in het Martini, samen als poortarts gewerkt in Drachten, vervolgens beiden in Rotterdam aan een promotie-onderzoek begonnen, en nu tegelijk gestart met de opleiding MDL. Reuze-gezellig! Als alle drukte weer voorbij is, moeten we weer eens gaan skaten. Jij ook veel succes met de laatste zware loodjes bij het afronden van jouw promotie-onderzoek, het gaat je absoluut lukken! Lieve Sarwa, ook jou sprak ik helaas veel te weinig. Ik vind het superknap van je hoe je naast de drukke bezigheden omtrent je promotie-onderzoek ook nog eens al die ontzettend zware examens voor je toekomst in het buitenland wist voor te bereiden. Heel veel succes met alles! Ook alle andere mede-promovendi heel erg bedankt voor de leuke tijd.

Ook wil ik mijn nieuwe collega's van het Sint Franciscus Gasthuis bedanken voor alle steun en interesse in de vorderingen omtrent mijn promotie-onderzoek. En natuurlijk ook voor de gezellige borrels en etentjes! Zowel de maatschap als de collega-arts-assistenten creëren een erg prettige werkomgeving, ik kan me geen betere vooropleidingsplek toewensen!

Lieve Mirian en Brenda, ik vind het heel fijn dat jullie mij op deze grote dag als paranimfen ter zijde willen staan. Ik zal jullie psychische steun in het zweetkamertje hard nodig hebben! Lieve Marieke, bedankt dat je altijd oprecht interesse wist te tonen, ondanks alles wat je voor je kiezen hebt gekregen. Lieve Jeltje, jij bent een ster in dingen in een helder perspectief plaatsen, ik kon altijd heerlijk mijn hart luchten. Ladies, helaas was ons contact de laatste tijd alleen nog af en toe over de telefoon, jullie wonen ook zo ver weg! Gelukkig breekt nu de tijd aan waarop we de schade kunnen gaan inhalen! Ook al mijn andere lieve vriendinnen, Gea, Sietske, Jasperien, Sanneke, Andrea, Esther en Klaske, bedankt voor al jullie interesse, gezellige telefoontjes en ook wij gaan de schade inhalen!

Zo'n heel promotie-onderzoek vergt nogal wat van je en dat zou me niet gelukt zijn zonder de

goede basis die ik heb meegekregen van mijn ouders. Lieve papa en mama, voor jullie was het best moeilijk dat ik ineens zo ver weg ging wonen, in mijn eentje in die grote stad Rotterdam. Gelukkig is alles op z'n pootjes terecht gekomen, mede dankzij jullie eeuwige steun en geloof in mij, waarvoor ik jullie enorm wil bedanken. Ik kan me geen betere, lievere ouders wensen, dikke kus! Lieve Wilfred en Geppy, Hugo en Lisette, jullie ook bedankt voor jullie interesse. Hopelijk kunnen we vanaf nu elkaar wat vaker opzoeken. Lieve Ferry en Sjanie, Erwin en Liselotte, ik heb er een lieve schoonfamilie bijgekregen, jullie ook bedankt voor al jullie interesse en begrip.

Dit promotie-onderzoek heeft mij niet alleen een mooi proefschrift opgeleverd. Lieve Sander, ik ben ontzettend blij dat ik jou heb leren kennen. Jij bent alles wat ik me kan wensen. Je vermogen om mijn 'ik-zie-het-niet-meer-zitten' buien om te vormen naar een optimistischer kijk op mijn onderzoek is erg groot. En je onvoorwaardelijke steun in de laatste fase van mijn promotie-onderzoek was fantastisch! Jij bent mijn atje, ik hou van jou!

APPENDIX

CYBAR (Flow CYtometry for the detetion of BARrett's patients at risk for developing adenocarcinoma) study group

Erasmus MC, Rotterdam
Afdeling Maag-Darm-Leverziekten: M Kerkhof, E J Kuipers, J G Kusters, P D Siersema; *Afdeling Pathologie:* H van Dekken; *Afdeling Maatschappelijke Gezondheidszorg:* E W Steyerberg, M Kruijshaar, M-L Essink-Bot.

IJsselland Ziekenhuis, Capelle aan den IJssel
Afdeling Maag-Darm-Leverziekten: W A Bode, H Geldof; *Afdeling Pathologie:* H van der Valk.

Ikazia Ziekenhuis, Rotterdam
Afdeling Maag-Darm-Leverziekten: D J Bac, R J Th Ouwendijk, C Leunis; *Afdeling Pathologie:* R W M Giard.

VU Medisch Centrum, Amsterdam
Afdeling Maag-Darm-Leverziekten: E C Klinkenberg, H Akol; *Afdeling Pathologie:* G A Meijer.

Albert Schweitzer Ziekenhuis, Dordrecht
Afdeling Maag-Darm-Leverziekten: W Lesterhuis, R Beukers, P Honkoop, W van de Vrie; *Afdeling Pathologie:* R J Heinhuis.

Deventer Ziekenhuis, Deventer
Afdeling Maag-Darm-Leverziekten: F ter Borg; *Afdeling Pathologie:* J W Arends.

Streekziekenhuis Midden Twente, Hengelo
Afdeling Maag-Darm-Leverziekten: G Tan; *Afdeling Pathologie:* J van Baarlen.

Rijnstate Ziekenhuis, Arnhem
Afdeling Maag-Darm-Leverziekten: N Aparicio, R de Vries, P Wahab, P van Embden; *Afdeling Pathologie:* A H Mulder.

Sint Franciscus Gasthuis, Rotterdam
Afdeling Maag-Darm-Leverziekten: L Berk, A J P van Tilburg, H S L M Tjen; *Afdeling Pathologie:* H van der Valk.

Medisch Spectrum Twente, Enschede
Afdeling Maag-Darm-Leverziekten: J J Kolkman, P Mensink, R Veenstra; *Afdeling Pathologie:* J van Baarlen.

Maasland Ziekenhuis, Sittard
Afdeling Maag-Darm-Leverziekten: L Engels; *Afdeling Pathologie:* W Vos.

Universitair Medisch Centrum Groningen, Groningen
Afdeling Maag-Darm-Leverziekten: F T M Peters; *Afdeling Pathologie:* A Karrenbeld.

Isala Klinieken, Zwolle
Afdeling Maag-Darm-Leverziekten: M Oudkerk Pool, B E Schenk, J Kamp, W H Bos, F Veen; *Afdeling Pathologie:* F Moll.

De Heel Medisch Centrum, Zaandam
Afdeling Maag-Darm-Leverziekten: R Loffeld; *Afdeling Pathologie:* M J Flens.

Franciscus Ziekenhuis, Roosendaal
Afdeling Maag-Darm-Leverziekten: H van Roermund; *Afdeling Pathologie:* F Lockefeer.

Academisch Ziekenhuis Maastricht, Maastricht
Afdeling Pathologie: A de Bruïne, A Driessen.

Curriculum vitae

CURRICULUM VITAE

Marjon Kerkhof werd op 5 juli 1976 geboren te Deventer. Na het behalen van haar V.W.O. eindexamen aan de Rijksscholengemeenschap te Heerenveen in 1994, besloot ze geneeskunde te gaan studeren. Vanwege het ontbreken van natuurkunde en scheikunde in het vakkenpakket heeft zij in één jaar alsnog deze V.W.O. certificaten gehaald aan de avondscholengemeenschap 'De Friese Wouden' te Heerenveen en is zij in 1995 gestart met de studie geneeskunde aan de Rijksuniversiteit van Groningen. De coschappen heeft zij verricht in het Martini Ziekenhuis te Groningen, het keuze-coschap op de afdeling interne geneeskunde/maag-darm-leverziekten in het 'Wilhelmina Ziekenhuis' te Assen en in december 2001 behaalde zij vervolgens het arts-examen cum laude. Van februari 2002 tot september 2003 werkte zij als poortarts op de spoedeisende hulp van 'Nij Smellinghe' te Drachten. Vanaf september 2003 werkte zij onder begeleiding van haar promotor Prof.dr. E.J. Kuipers en haar copromotoren Dr. P.D. Siersema en Dr. J.G. Kusters aan haar promotie-onderzoek naar Barrett oesofagus. Dit onderzoek werd uitgevoerd op het laboratorium van de afdeling maag-darm-leverziekten van het Erasmus MC te Rotterdam en vormde de basis voor dit proefschrift. In december 2006 is zij gestart met haar opleiding tot maag-darm-leverarts via het Erasmus MC te Rotterdam (opleiders: Prof. Dr. E.J. Kuipers en Dr. R.A. de Man), waarbij de vooropleiding wordt verricht in het Sint Franciscus Gasthuis te Rotterdam (opleiders: Drs. A.P. Rietveld en Dr. H.C.T. van Zaanen).