Biomarkers In Patients With Barrett's Esophagus And Esophageal Adenocarcinoma

L. Suzuki

TO BE OR NOT TO BE:
Biomarkers In Patients With Barrett's Esophagus And Esophageal Adenocarcinoma

Lucia Suzuki

Colofon

Suzuki, L.
TO BE OR NOT TO BE: Biomarkers In Patients With Barrett's Esophagus And Esophageal Adenocarcinoma

Cover design: Inge de Voogt

ISBN: 978-94-6423-359-9

Layout and printing: ProefschriftMaken || www.proefschriftmaken.nl

Financial support for the printing of this thesis was kindly provided by Erasmus University Rotterdam, the department of Pathology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands and Chipsoft.

TO BE OR NOT TO BE:
Biomarkers In Patients With Barrett's Esophagus And Esophageal Adenocarcinoma

Zijn of niet zijn: biomarkers in patiënten met een Barrett slokdarm
en adenocarcinoom van de slokdarm

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de rector magnificus

Prof. dr. F.A. van der Duijn Schouten

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

donderdag 9 september 2021 om 10.30 uur

door

Lucia Suzuki
geboren te Rotterdam

Erasmus University Rotterdam

Promotiecommissie:

Promotoren:	Prof. dr. L.H.J. Looijenga
	Prof. dr. F.J. van Kemenade
Overige leden:	Prof. dr. M.P. Peppelenbosch
	Prof. dr. W.N.M. Dinjens
	Prof. dr. K. K. Krishnadath
Copromotor:	Dr. K. Biermann

Index

1

Introduction

Barrett's esophagus

Barrett's esophagus (BE) is a condition in which the normal squamous cells that line the esophagus are replaced by a columnar cell type mucosa that is both visible endoscopically and confirmed histologically by a pathologist (Figure 1). Most countries also require the presence of intestinal metaplasia to make a diagnosis of BE, more specifically the presence of goblet cells. Although the requirement of their presence is debated, most observers acknowledge goblet cell presence correlates better with an increased risk for progression to cancer i.e. esophageal adenocarcinoma (EAC).[1, 2]

Barrett's esophagus epidemiology and risk factors

BE is often present in patients with chronic gastroesophageal reflux disease (GERD) due to repeated exposure of their esophagus to excessive amounts of stomach acid and bile. Worldwide, GERD incidences are increasing due to changes in lifestyle accompanied by a rapid increase in prevalence of BE.[3, 4] As BE itself is asymptomatic and can only be diagnosed during endoscopy, leaving the vast majority unrecognized, the exact prevalence of BE is unknown. Estimates vary from 1 to over 20 percent, depending on the population studied and the criteria used.[5-8] Apart from chronic GERD, age above 50 years, male gender, Caucasian race and central adiposity are common predisposing risk factors for BE.[9]

Pathology: metaplasia-dysplasia-adenocarcinoma sequence

EAC develops after a sequential accumulation of molecular events, reflected by the histological presence of metaplasia and various stages of dysplasia, to adenocarcinoma (Figure 2).[10, 11] However, the time course between different stages is very variable and not fixed.

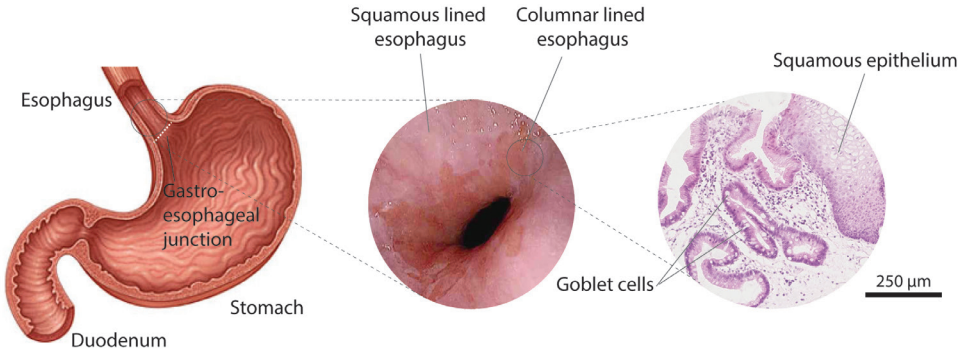


Figure 1. Endoscopic and histological appearance of Barrett's esophagus. Macroscopically visible with endoscopy, the normal white-looking squamous esophageal lining is replaced by a salmon-pink colored columnar mucosa. Microscopy shows simple glands lined with columnar epithelium containing goblet cells (intestinal metaplasia).

Metaplastic columnar cells lining the glands in non-dysplastic BE show minimal cytological atypia and surface maturation is preserved. Glands show frequent goblet cells and are round with abundant stroma (lamina propria) in between (Figure 2b). Non-dysplastic BE (NDBE) has a low estimated absolute risk of progression to cancer of less than 1% per patient per year.[12, 13]

Dysplasia is defined as neoplastic epithelium confined to the basement membrane. Grading allows for stratification and management of patients according to the associated estimated progression risk. The revised Vienna classification system attempted to minimize interobserver variation between pathologists and consists of five categories based on the degree of cytological and architectural abnormalities present.[14, 15] If a biopsy is neither obviously dysplastic nor negative for dysplasia the term indefinite for dysplasia can be used. Still, the diagnosis of low-grade dysplasia (LGD) in particular can be difficult.[16] Poor interobserver agreement is a known issue and led to the recommendation of having a

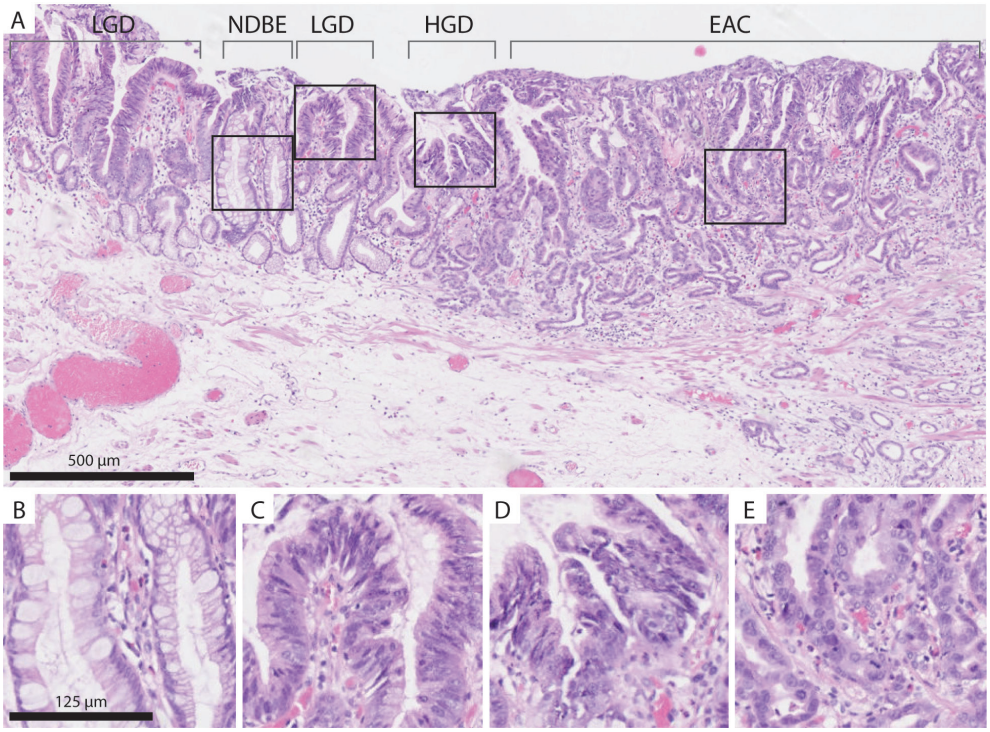


Figure 2. Histological example of the so-called metaplasia-dysplasia-carcinoma sequence. A) Various grades of (non-) dysplastic Barrett's esophagus and progression to adjacent esophageal adenocarcinoma (EAC) B) Non-dysplastic BE (NDBE) is present as well as foci of C) low- (LGD) and D) high-grade dysplasia (HGD). Histologically a continuous spectrum can be appreciated, in which hard limits with regard to the start and the end of the relevant dysplasia grade can be very difficult to determine. For example, HGD may be indistinguishable from E) EAC, in particular on a biopsy.

second, preferably an expert, gastro-intestinal pathologist confirm the diagnosis of dysplasia.[17, 18] In LGD the architecture remains preserved, but the epithelium shows cytological abnormalities, including enlarged nuclei, mucin depletion, frequent mitoses and a lack of surface maturation (Figure 2c).[19] These four easily reproducible LGD criteria may improve patient stratification .[20]

High-grade dysplasia (HGD) shows an even greater degree of cytological atypia, with marked pleomorphism, frequent (atypical) mitosis and loss of polarity, with haphazardly arranged nuclei in relation to the basement membrane (Figure 2d). Importantly, HGD involves architectural abnormalities, such as irregular size and shape of glands, crowded, back-to -back glands with only sparse intervening stroma. A biopsy with a diagnosis of HGD may be very suspicious and sometimes indistinguishable from a (superficial) adenocarcinoma (Figure 2e).[21] Moreover, a significant proportion of cases with HGD in the biopsy sample, 12.7% in a meta-analysis,[22] show evidence of carcinoma in the resected specimen. Frequently, a definite diagnosis on a superficial biopsy sample is not possible rendering the pathologist to use the diagnosis "HGD, cannot exclude/ suspicious for invasion".

Endoscopic screening & surveillance

Considering the high BE prevalence, a gastroscopy is only justified in white men, older than 50 years with long-term (more than 5 years) and severe (daily) reflux symptoms to prevent overdiagnosis and overtreatment.[23] Further endoscopic surveillance depends on BE segment length, presence of confirmed dysplasia and associated estimated progression risk. It involves targeted biopsies of visible abnormalities, followed by four-quadrant biopsies every 1-2 cm.[24] Yet, missing high-risk lesions (i.e. sampling error) may still occur.

Repeated endoscopy after six months is recommended in patients with confirmed LGD in the absence of macroscopically visible changes.[1, 25, 26] Apart from the increased risk of progression, persistent LGD is also associated with a risk of presence of adjacent HGD/EAC, believed to have been missed by biopsy. Therefore, endoscopic ablative therapy can be considered in patients with persistent LGD. A histologically confirmed diagnosis of HGD is indicative of endoscopic resection followed by ablative therapy.[26]

Despite screening and surveillance, a significant proportion of patients present with (advanced) EAC without previous knowledge of having (dysplastic) BE. A recent study among 8,564 EAC patients reported prior BE diagnosis in only 4.9% of patients.[27] Importantly, on the other hand, most patients with BE (>90%) don't die from EAC.[28]

Established EAC: early vs. advanced

Traditionally, complete esophagectomy (removal of the esophagus together with cardia and lesser curve of the stomach) including lymphadenectomy was the only treatment option with possibility to curation for EAC of all stages (Figure 3). As this procedure is related with high mortality and morbidity, increasingly, endoscopic mucosal resection is being used when EAC is detected at an early stage (EAC limited to the mucosa or submucosa) as well as for patients with HGD. Because no lymph nodes can be removed by this procedure, the risk of nodal metastases has to be negligible. Despite advances in clinical imaging modalities, this risk is still difficult to determine.[29-31]

Even with extensive radical surgery, survival rates in advanced EAC remained poor, caused by frequent local and systemic recurrences. A significant survival benefit is now widely recognized for patients with advanced EAC receiving neo-adjuvant treatment.[32] Accordingly, multimodal therapy, i.e. neo-adjuvant chemoradiationtherapy (nCRT) prior to surgery, has now become standard of care. Moreover, it is currently being addressed if apparent complete responders after nCRT can be kept under close surveillance and only undergo surgery, if evident localized tumor re-presents (SANO trial: surgery as needed in oesophageal cancer).[33]

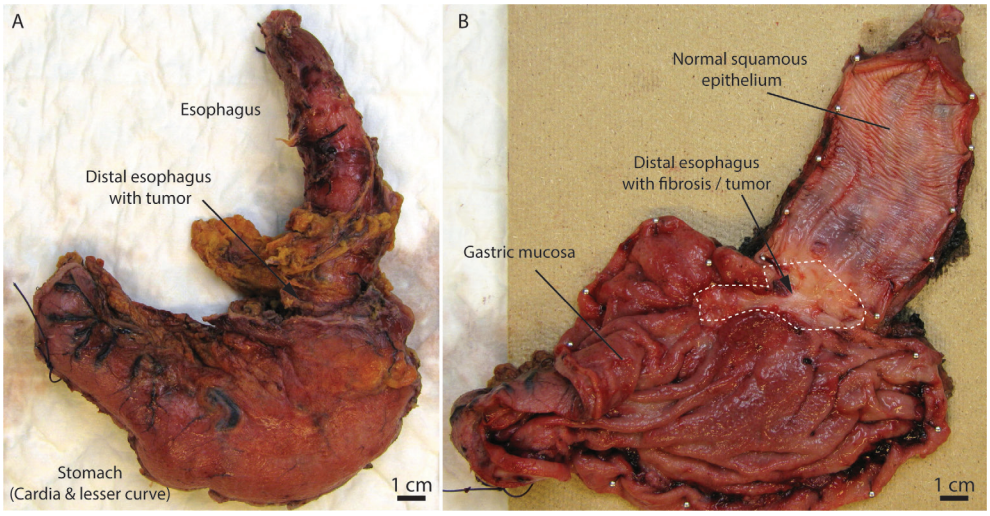


Figure 3. An esophageal adenocarcinoma specimen resected by esophagectomy after treatment with neo-adjuvant chemoradiationtherapy (nCRT). A) Esophagectomy specimen including lower esophagus and proximal part of the stomach with surrounding fatty tissue containing regional lymph nodes. B) After inking of the outer (circumferential) margins, the specimen is opened near the staple line and pinned onto cork to allow for optimal fixation.

Histopathological classification and current risk stratification

Histopathological examination, usually of a biopsy, is the principal means of a cancer diagnosis. Increasingly, this biopsy may provide the only "details" about the cancer as neo-adjuvant therapy (nCRT, ablation) may prohibit future assessment of the primary cancer. The TNM classification is the most important tool for prognostic assessment of patients in clinical practice and research as well as for epidemiological studies.[34] To evaluate response to nCRT, tumor regression grading (TRG) systems aim to categorize the amount of regressive changes. In general, the amount of therapy induced fibrosis in relation to the amount of residual tumor is evaluated using slightly different approaches.[35-37] Ulceration, mucin lakes and areas with fibrosis are used to identify areas where tumor was located (Figure 3). The TRG system first described by Mandard is the most widely used system (Table 1).[36] Numerous studies show the prognostic relevance of TRG, in which patients with complete tumor regression do best. Furthermore, the prognostic value of TRG may even exceed those of currently used staging systems (e.g., TNM staging) which are originated from data from untreated tumors. Post-nCRT staging (ypTNM) and grading response in additional treatment decision-making is debated, although addition of adjuvant chemotherapy (i.e. after surgery) has been reported to associate with a survival benefit, particularly in patients with residual nodal disease (ypN+).[38-40] Of great interest is the expanding research on identifying therapy response before surgery, making optimized personalized therapy according to the respective tumor behavior possible.[33]

Other variables of important prognostic significance include tumor grade and presence of tumor in vessels (lymphovascular invasion, LVI). Grade and LVI significantly affect survival and are particularly of clinical importance in patients with early stage EAC to define optimal treatment strategy.[41, 42]

Still, it remains impossible to accurately predict prognosis for an individual patient. However survival rates i.e. the percentage of patients still alive five years after surgery, may be useful to estimate the likelihood that treatment will be successful. In general, the outlook for EAC patients is poor, as EAC often recurs. Despite progress in treatment and surveillance programs, the prognosis has only increased slowly over the past 40 years, with a five-year survival rate that rarely exceeds 40%.[32, 43]

Biomarkers

Biomarker studies aim to improve both risk stratification and treatment protocols. Furthermore, they may provide inside on the biological behavior of disease. A biomarker is defined as a measurement indicative of a biological process, or a response to a therapeutic intervention.[44] Among these, the p53 protein encoded on the TP53 gene, is probably the most frequently studied biomarker. TP53 gene mutation is widely recognized as a driver mutation in EAC and frequently occurs early in malignant progression, followed by

Table 1. Tumor regression grade (TRG) as reported by Mandard et al.[36]

Tumor Regression Grade	Histological description
TRG 1	Complete response with absence of residual cancer
TRG 2	Rare residual (isolated) cancer cells scattered throughout fibrosis
TRG 3	Increase in residual cancer, but fibrosis still predominated
TRG 4	Residual cancer outgrowing fibrosis
TRG 5	Absence of regressive changes

further genomic instability.[11] Sequencing is the gold standard to detect mutations, but is labor intensive, time consuming and not available in all laboratories. As a surrogate for mutational analysis, p53 immunohistochemistry has been shown to improve the diagnostic reproducibility of a dysplasia diagnosis and increasingly its independent prognostic value regarding progression risk is being recognized.[16, 45, 46] Hence, p53 status analysis by immunohistochemistry is recommended by the Dutch and British guidelines as an adjunct to routine clinical diagnosis of dysplasia.[25, 26]

Interestingly, several studies showed pathogenic TP53 mutations and epigenetic alterations, like distinct methylation profiles, may precede histological dysplastic changes and potentially could be used to predict progression in BE patients.[47-49] However, despite an enormous amount of research, the ideal biomarker for BE to determine the most accurate progression risk has yet to be discovered.[50-53]

MicroRNAs are small non-protein-coding RNAs able to regulate the functionality of a third of all human genes, underscoring the potential influence on almost every genetic pathway. They can function as both tumor suppressors and oncogenes. MicroRNA dysregulation indeed frequently associates with human cancer.[54, 55] Accordingly, the amount of miRNA research related to BE and EAC has grown rapidly.[56-58] However, relatively few studies focused on miRNA related to progression.[59]

An important prognostic indicator in colorectal cancer is tumor budding. Tumor budding has been shown to be a risk factor for nodal metastasis and international guidelines now recommend reporting tumor budding due to its implications for management, such as additional resection in early cancers.[60] To date, its utility in esophageal cancer is uncertain as most recent studies in EAC did not include tumor budding.[41, 61, 62] The few available studies in EAC often used different methods to assess tumor budding, making comparison and interpretation of results difficult.[63-65] Besides, digital pathology is a rapidly involving field with the potential to improve staining interpretation and image analysis. Therefore, deep learning algorithms are increasingly being used to score potential biomarkers as objectively as possible.

Olfactomedin 4 (OFLM4) is considered as another promising biomarker. OLFM4 is an intestinal stem cell marker and stains intestinal metaplasia in non-dysplastic BE in a similar way as in normal colon crypts.[66] Low OLFM4 expression has been correlated with poor differentiation grade, nodal metastases and adverse survival in a variety of cancers, but data on OLFM4 in EAC are lacking.[67-72]

In addition to its value in BE, aberrant immunohistochemical expression of p53 in EAC has also been shown to be associated with response to neo-adjuvant treatment.[73, 74] Similar results were found for Sex Determining Region Y-Box 2 (SOX2), a transcription factor involved in the regulation of embryonic development and important for stem-cell maintenance.[74-76] These data imply an important underlying biological role for these proteins in the pathogenesis of EAC. Importantly, many biomarker studies are based on historical cohorts of patients treated with surgery alone to exclude potential effect from neo-adjuvant treatment, although some include mixed patient populations. The effect of neo-adjuvant treatment on immunohistochemical expression of biomarkers, and relevance hereof, is frequently unknown. With increased usage of neo-adjuvant treatment this warrants further research.

2

Aim and outline

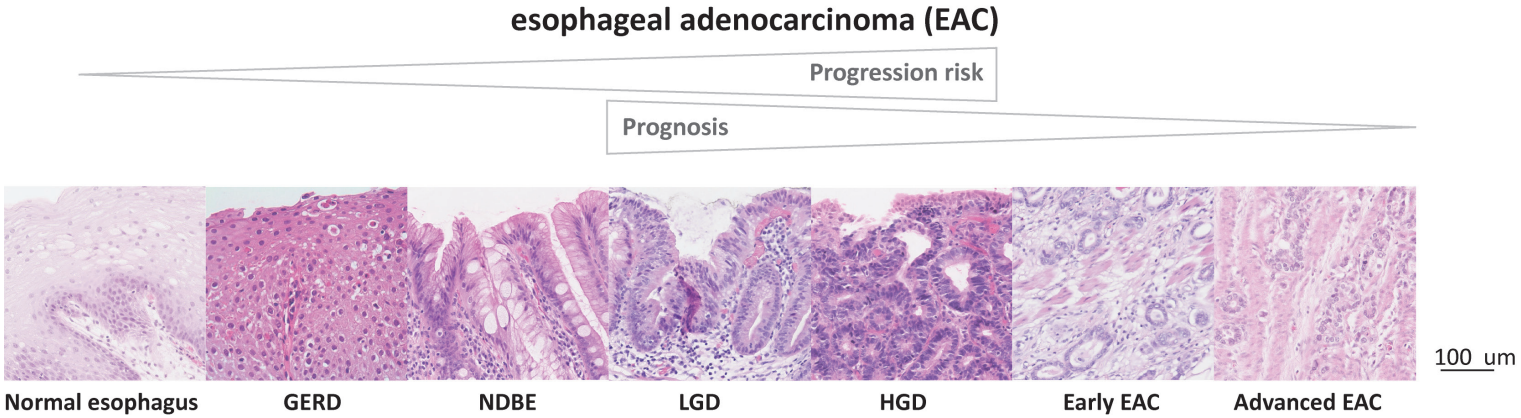
The general aim of this thesis is to improve risk stratification in patients with Barrett's esophagus (BE) and early esophageal adenocarcinoma (EAC) and improve survival prediction in advanced EAC. The key questions of this thesis were:

- Which patients with newly diagnosed BE are at increased risk for developing high-grade dysplasia (HGD) or EAC?
- Which patients with (early) EAC are at risk for developing (nodal) metastases?
- Can we further improve prediction of prognosis in patients with advanced EAC?

To approach these questions we used new and previously described biomarkers including microRNAs, tumor budding, Olfactomedin 4, p53 and SOX2 (Figure 1).

In **Chapter 3** we used a histology-based microRNA approach in non-dysplastic BE to discriminate those patients at risk for progression to HGD or EAC compared to non-progressors.

In **Chapter 4** the prognostic value of tumor budding in early (pT1b) EAC was examined, whereas **Chapters 5 and 6** describe the prognostic value of biomarkers in advanced EAC. As neo-adjuvant treatment may influence biomarker expression, only EAC patients treated with surgery alone were included. In **Chapter 5,** Olfactomedin 4 an intestinal stem cell marker, was examined. In **Chapter 6,** p53 immunohistochemistry was evaluated, including comparison with molecular analyses like DNA sequencing and high-throughput methylation analyses to reveal underlying molecular changes. In **Chapter 7** the immunohistochemical expression of both p53 and SOX2 expression before and after neo-adjuvant treatment were compared. In addition their association with survival and comparison with results found in patients treated with surgery alone is being addressed. At last, the results, concluding remarks and future prospects are discussed in **Chapters 8-9** and summarized in **Chapter 10**.



Chapter		Chapter 3			Chapters 4-5	Chapters 5-7
Question?		Who is at risk for progression?			Who is at risk for (nodal) metastases? Can we improve prognosis prediction?	
Biomarker studied?		miRNA			TB, OLFM4, p53, SOX2	
Method?		High-throughput miRNA profiling (histology, biopsy based)			H&E, IHC (conventional & digitilized) Automated digital TB count DNA sequencing, methylation analysis	

Figure 1. Aim and outline of this thesis.
GERD, gastro-esophageal reflux; NDBE, non-dysplastic Barrett's esophagus; LGD, low-grade dysplasia; HGD, high-grade dysplasia; miRNA, microRNA; TB, tumor budding; H&E, hematoxylin & eosin; IHC, immunohistochemistry

Exploring histology-based microRNA profiles in the progression of Barrett's esophagus to esophageal adenocarcinoma

Suzuki L, Gillis AJM, Dorssers LCJ, Henriquez J, Nieboer D, Bruno MJ, Spaander MCW, Looijenga LHJ, Biermann K.

Submitted.

Background: Better risk stratification is essential to identify patients with Barrett's esophagus (BE) at risk for progression to esophageal adenocarcinoma (EAC), to avoid overdiagnosis and subsequent overtreatment. MicroRNAs (miRs) could potentially serve as prognostic biomarkers for improved risk stratification. This study aims to identify miRs able to detect patients at risk for EAC prior to the development of dysplasia.

Methods: Non-dysplastic BE (NDBE) samples of patients who developed high-grade dysplasia (HGD) or EAC (progressors, n=16) were compared with NDBE samples from non-progressors (n=18). Both high-throughput miR-profiling, including 381 different miRs, and individual assays were used.

Results: The discovery set revealed 11 miRs with significantly higher levels in NDBE tissues from progressors (n=9) compared to non-progressors (n=9), while no significant differences were found between paired NDBE and HGD/EAC samples. A validation set (n=16), could not confirm these results. In addition, our data could not confirm previously described prognostic miRs (miR-192, miR-194 and miR-196b) to be associated with progression.

Conclusions: This study could not identify progression related miRs in NDBE samples and illustrates independent validation is crucial for (molecular) biomarker validation studies.

Abstract

INTRODUCTION

Patients with Barrett's esophagus (BE), the precursor lesion of esophageal adenocarcinoma (EAC), have a 30- to 60-fold increased risk of developing EAC compared to normal individuals.[77] Surveillance programs in patients with non-dysplastic BE (NDBE) aim to diagnose and treat EAC at an early stage. Although the (cumulative) annual risk of developing EAC is only 0.5%, BE prevalence ranges from 2 to 7% and these numbers are almost certainly an underestimation, making cost-effective surveillance a major challenge.[5, 78-80] Importantly, histology of BE is under discussion, partly due to low interobserver agreement for dysplasia.[81-83] We hypothesize that it might be possible to determine the progression risk before any dysplastic changes are detectable on histological examination using microRNAs (miRs). This would improve risk-stratification of BE patients under surveillance and circumvent the problem of interobserver variation in BE samples.

MiRs are small non-coding RNAs of ~22 nucleotides in length able to control translation or cause degradation of messenger RNA transcripts of protein-coding genes. They are involved in various biological processes, such as cell cycle, differentiation, development, and as such, deregulation has been reported in human diseases including cancer. High-throughput miR-profiling can differentiate between tissue types and distinctive miR signatures are recognized for specific cancer types.[55] Moreover, by inhibiting translation or facilitating degradation, miRs can either function as oncogenes or tumor suppressors. Consequently, miRs are increasingly studied as cancer biomarkers.[84, 85] Different miR-expression profiles have been observed in NDBE compared with normal esophageal tissue with squamous epithelial lining.[86] Selected miRs detectable in serum have been proposed as circulating biomarkers for NDBE.[56] In addition, differentially expressed miRs were found in NDBE compared to high-grade dysplasia (HGD) and/or EAC, but results are not found to be consistent.[58]. Nevertheless, various studies tried to identify specific miRs associated with progression, although the majority of these studies were cross-sectional and without longitudinal follow-up or only a few selected miRs were involved.[56-58, 87] Because of these shortcomings, the present study aims to clarify whether or not miRs exist that are differentially expressed in NDBE tissues of progressors and non-progressors.

This study uses a high-throughput miR profiling approach and endoscopic biopsy material with NDBE histology from patients with long-term follow-up.

METHODS

Patients & study design

Patients were retrospectively selected from the PROBAR study, a large Dutch multicenter cohort study which originally included more than 720 patients, with newly diagnosed or known BE of > 2cm, between November 2003 and December 2004. More elaborate details have been described previously.[46, 76] In accordance with previous publications, patients with a history of HGD/EAC were excluded in the present study, as well as those with a diagnosis of HGD/EAC within nine months after the first biopsy upon inclusion in the cohort study to prevent inclusion of prevalent cases (possible sampling error).[46, 76] In total, 51 patients progressed to HGD (n=37) or EAC (n=14) and were classified as progressors, whereas 584 patients were classified as non-progressors.

Only those cases with all stained slides from all follow-up moments available (for histopathological review) and sufficient paraffin material containing (as determined by two observers) were included, hence 16 progressors and 18 non-progressors were identified from the PROBAR cohort. Other reasons for exclusion included marked presence of abundant inflammatory cells or high volume of either normal stroma or squamous epithelium in comparison with Barrett's mucosa or adenocarcinoma. All histological samples of these patients were again critically revised and diagnosis confirmed by an expert gastrointestinal pathologist. Median follow-up time was 95 months (Table 1). Patients were divided in a discovery (series 1, n=18) and validation (series 2, n=16) set.

Initially high-throughput miR profiling was performed on the discovery set to identify up to 377 possible relevant miRs that showed statistically significant higher miR levels in NDBE from progressors compared to non-progressors, after which validation was performed using individual assays on selected miRs in the same set. Both methods were also applied in the validation set to validate the results found in the discovery set. Statistically significant differential miRs with lower levels in progressors were not further investigated as measuring downregulation was considered unpractical in a future diagnostic setting. Corresponding HGD/EAC samples (n=9) from the same progressors were similarly examined for the discovery set in order to compare miR levels from HGD/EAC versus NDBE samples from the same patients (i.e. paired samples) as well as in comparison with NDBE of non-progressors (i.e. unpaired samples). A flowchart for this study is depicted in Figure 1.

miR isolation and profiling

Glass slides with 4 µm sections from corresponding formalin fixed paraffin embedded blocks were baked at 60°C for 20 minutes and stained with hematoxylin after deparaffinization. The areas of interest were macrodissected and dissolved in 100 µl RNA*Later* solution (Cat# 7030, ThermoFisher). The AllPrep DNA/RNA FFPE Midi kit (Cat No./ID: 80234, QIAGEN, Venlo, The

Table 1. Patient characteristics.

		All patients (n =34)			Progressors (n=16)			Non-Progressors (n=18)		
		Series 1 (n=18)	Series 2 (n=16)	p-value	Series 1 (n=9)	Series 2 (n=7)	p-value	Series 1 (n=9)	Series 2 (n=9)	p-value
Age (y)	Mean (SD)	60 (10)	64 (15)	0.313	63 (10)	65 (10)	0.606	57 (10)	64 (9)	0.222
Gender (n)	Female	5	6	0.545	3	4	1.00	2	3	1.00
	Male	13	10		6	3		7	6	
Body mass index (kg/m2)	Median (IQR)	26.2 (6)	26.9 (7)	0.986	27.8 (10)	25.6 (8)	0.681	24.9 (4)	26.9 (7)	0.546
Smoking (n)	No, never	7	8	0.435	4	1	0.293	3	7	0.155
	Yes, former	7	3		3	2		4	1	
	Yes, current	4	5		2	4		2	1	
PPI use (n)	No	1	3	0.323	1	2	0.550	0	1	1.00
	Yes	17	13		8	5		9	8	
Length BE segment (cm)	Median (IQR)	4 (5)	5 (6)	0.721	4 (5)	5 (5)	0.681	3 (5)	4 (5)	0.796
Esophagitis (n)	No	12	12	0.458	4	7	0.034	8	5	0.294
	Yes	6	3		5	0		1	3	
Biopsies taken (n)	Median (IQR)	8 (7)	10 (6)	0.695	8 (9)	12 (4)	0.606	8 (8)	8 (4)	0.796
Follow-up time (m)	Median (IQR)	95 (86)	112 (97)	0.251	27 (48)	60 (59)	0.681	112 (38)	157 (46)	0.077
Progression type (n)	HGD	5	4	1.00	5	4	1.00	0	0	1.00
	EAC	4	3		4	3		0	0	

y, years; SD, standard deviation, n, number; IQR, interquartile range; m, months; HGD, high-grade dysplasia; EAC, esophageal adenocarcinoma

Netherlands) was used to isolate total RNA, including miRs, according to the manufacturer's instructions. The sample volumes were reduced by evaporation to 10 µl with an Eppendorf™ Concentrator Plus. All reagents for RT- qPCR were obtained from ThermoFisher (www.thermofisher.com). TaqMan miR reverse transcription kit in combination with Megaplex™ RT primers was used for cDNA synthesis. Subsequently, cDNA was pre-amplified and quantitated on Taqman Low Density miR array a cards (TLDA) containing 384 targets (including 7 controls) on a TaqMan 7900Fast thermocycler using supplier protocols. Individual reverse transcriptase (RT)-PCR miR assays were performed to confirm results found in TLDA analysis using a TaqMan 7500Fast thermocycler. A more detailed description of miR isolation and profiling has been previously described. [88]

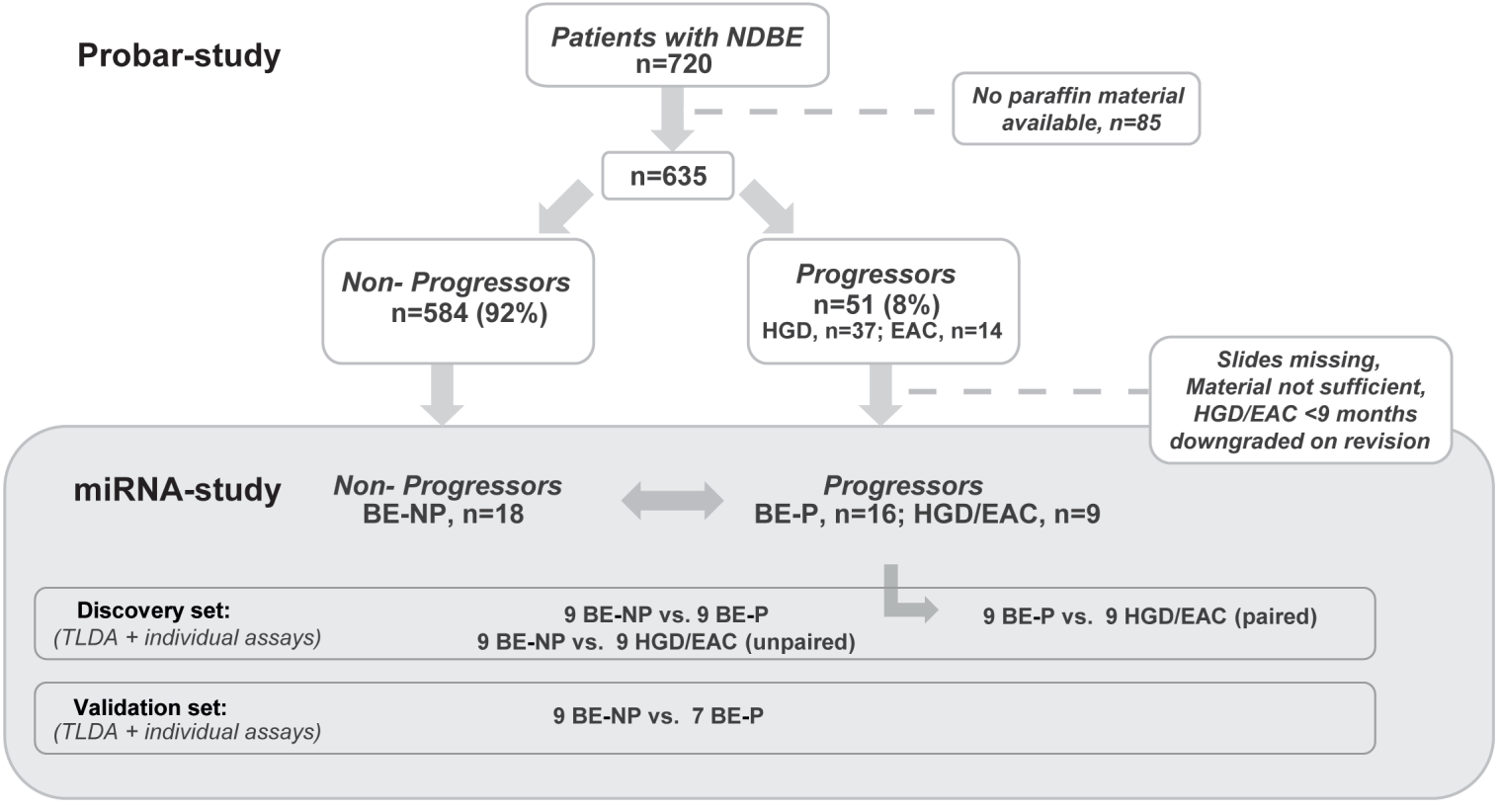


Figure 1. Flowchart of patients included in the study and study design.

Statistical & bioinformatical analysis

Descriptive statistics of general patient characteristics were analyzed with SPSS statistical software package (SPSS 25, IBM Corp). The Mann-Whitney U test or T-Test was used to compare continuous variables, the chi-squared or Fisher exact test for categorical variables.

Taqman miR array output data were uploaded in the ThermoFisher Cloud Software and analyzed using defined threshold settings for each individual miR. Cycle threshold (Ct) values, denoting the number of cycles required to achieve a defined level of signal, were exported (Supplementary Table 1), quality filtered using R-tools and normalized using QbasePlus software (Biogazelle N.V., Zwijnaarde, Belgium) as described previously.[89] Lower levels of miR are represented by higher Ct values and vice versa. The miRs which failed to give a Ct value were given the maximum number of cycles used (*"40"*). Normalization of single assay miR levels was performed using the levels of miR-26a, miR-26b, miR-28-5p and miR-374a which displayed comparable levels between the tissues analyzed with TLDA cards. Comparisons of the calibrated normalized relative quantities (CNRQ) levels between groups were performed using the Mann-Whitney U test for unpaired samples and the Wilcoxon signed rank test for paired samples in the QbasePlus software. A p-value of <0.05 was considered statistically significant. Statistics regarding miR analysis were corrected for multiple testing in all cases by the false discovery rate multiple comparison method.[90]

Normalization is crucial to correct for technical variation affecting miR levels. QbasePlus software allows for two types of normalization; "global mean normalization on common targets" and "global mean normalization". Both methods were used sequentially. While the first is based on the approach that a single given sample is normalized based on the geometric mean of the relative quantities of miRs measured in all samples, the second normalization is based only on all targets expressed in that sample.

RESULTS

Discovery set

In total, 18 patients were included in the discovery set (series 1); nine progressors and nine non-progressors. Minimum and maximum time to progression was 11 and 88 months (median 27 months), while non-progressors were followed for at least 92 up to 164 months (median 112 months). Patient characteristics are listed in Table 1. No significant differential levels of miRs were found between NDBE and HGD/EAC samples from the same patients (i.e. paired samples, online resource Supplementary Table 1). A total of 39 miRs showed statistically significant differential levels between NDBE from progressors and non-progressors (Table 2). From these, 11 miRs (miR-18a, miR-23, miR-93, miR-331-5p, miR-483-5p, miR-492, miR-494, miR-548b-5p, miR-548c-5p, miR-548d-5p, miR-576-3p)

Table 2. MiRs with statistically significant differential levels found in discovery set using TLDA and "global mean normalization on common targets".∞

	NDBE (NP) vs NDBE (P) (n=9 vs n=9)		*NDBE (NP) vs HGD/EAC (unpaired, n=9 vs n=9)*	
	Ratio	*p-value**	*Ratio*	*p-value**
hsa-miR-523	*0,09*	***0,027***	*0,30*	*0,647*
hsa-miR-548d-5p	*0,12*	***0,004***	*0,21*	*0,157*
hsa-miR-576-3p	*0,14*	***0,038***	*0,14*	*0,113*
hsa-miR-494	*0,19*	***0,004***	*0,61*	*0,563*
hsa-miR-548c-5p	*0,21*	***0,004***	*0,43*	*0,361*
hsa-miR-548b-5p	*0,22*	***0,038***	*0,46*	*0,325*
hsa-miR-492	*0,27*	***0,027***	*0,99*	*0,974*
hsa-miR-331-5p	*0,27*	***0,038***	*0,22*	***0,039***
hsa-miR-483-5p	*0,28*	***0,027***	*0,40*	*0,325*
hsa-miR-18a	*0,60*	***0,027***	*0,43*	***0,010***
hsa-miR-93	*0,63*	***0,020***	*0,51*	***0,042***
hsa-miR-331-3p	*1,39*	***0,046***	*1,07*	*0,798*
hsa-miR-374a	*1,46*	***0,020***	*1,03*	*0,647*
hsa-miR-125b	*1,49*	***0,027***	*1,64*	*0,319*
hsa-miR-29a	*1,52*	***0,020***	*0,83*	*0,828*
hsa-miR-199a-3p	*1,57*	***0,027***	*1,11*	*0,798*
hsa-miR-125a-5p	*1,62*	***0,027***	*0,80*	*0,857*
hsa-miR-30b	*1,64*	***0,006***	*1,45*	*0,157*
hsa-miR-376a	*1,66*	***0,027***	*1,60*	*0,647*
hsa-miR-127-3p	*1,70*	***0,032***	*1,33*	*0,828*
hsa-miR-99b	*1,73*	***0,027***	*1,26*	*0,528*
hsa-miR-365	*1,77*	***0,038***	*1,34*	*0,563*
hsa-miR-186	*1,85*	***0,027***	*1,09*	*0,828*
hsa-miR-193b	*1,92*	***0,046***	*1,42*	*0,563*
hsa-miR-411	*1,94*	***0,020***	*1,87*	*0,201*
hsa-miR-150	*1,94*	***0,046***	*0,88*	*0,798*
hsa-miR-30c	*1,95*	***0,004***	*1,30*	*0,249*
*hsa-miR-9**	*2,06*	***0,046***	*2,69*	***0,039***
hsa-miR-410	*2,13*	***0,020***	*1,21*	*0,681*
hsa-miR-342-3p	*2,15*	***0,020***	*1,17*	*0,897*
hsa-miR-139-5p	*2,17*	***0,020***	*1,93*	*0,062*
hsa-miR-539	*2,28*	***0,046***	*3,80*	*0,242*
hsa-miR-133a	*2,35*	***0,032***	*2,38*	*0,319*
hsa-miR-487b	*2,44*	***0,032***	*3,87*	*0,319*
hsa-miR-204	*2,50*	***0,038***	*2,76*	*0,054*

Continued

	NDBE (NP) vs NDBE (P) (n=9 vs n=9)		NDBE (NP) vs HGD/EAC (unpaired, n=9 vs n=9)	
	Ratio	*p-value**	*Ratio*	*p-value**
hsa-miR-218	*2,67*	***0,004***	*1,98*	*0,287*
hsa-miR-197	*2,70*	***0,032***	*1,42*	*0,619*
hsa-let-7c	*2,86*	***0,032***	*2,25*	*0,325*
hsa-miR-139-3p	*58,87*	***0,020***	*7,31*	*0,094*
hsa-miR-20b	*0,77*	*0,330*	*0,49*	***0,036***
hsa-miR-19a	*0,80*	*0,638*	*0,51*	***0,029***
hsa-miR-335	*0,84*	*0,919*	*0,54*	***0,022***
hsa-miR-135a	*0,82*	*0,501*	*0,56*	***0,039***
hsa-miR-21	*0,89*	*0,557*	*0,60*	***0,042***
hsa-miR-95	*0,91*	*0,743*	*0,37*	***0,039***
hsa-miR-185	*0,95*	*0,786*	*0,70*	***0,039***
hsa-miR-491-5p	*0,98*	*0,840*	*0,62*	***0,022***
hsa-miR-25	*1,18*	*0,557*	*0,65*	***0,042***
hsa-miR-885-5p	*1,84*	*0,291*	*3,08*	***0,039***
hsa-miR-135b	*1,08*	*0,657*	*0,40*	***0,042***

*No significant differential levels were found in miRs from paired NDBE and HGD/EAC samples (not shown). In total, 11 miRs showed higher levels (ratio <1) and 28 miRs showed lower levels (ratio >1) in NDBE from progressors. Three miRs were not only more abundant in NDBE from progressors but also in paired HGD/EAC samples compared to NDBE of non-progressors (miR-331-5p, miR-18a and miR-93). MiR-9 was lower in both NDBE from progressors and HGD/EAC. *FDR adjusted p-value. *NDBE, non-dysplastic Barrett's esophagus; NP, non-progressor; P, progressor; HGD, high-grade dysplasia; EAC, esophageal adenocarcinoma, n, number; hsa-miR, homo sapiens microRNA*

showed increased levels in progressors (ratio <1, Figure 2), whereas 28 miRs demonstrated a reduced level (ratio >1). Overall, 15 miRs were differential in levels between NDBE of non-progressors vs HGD/EAC (i.e. unpaired samples, Table 2), of which 13 showed higher and two lower levels in HGD/EAC. Interestingly, four miRs showed "shared" differential levels. Specifically, three miRs were not only more abundant in HGD/EAC, but also increased in NDBE of progressors (miR-331-5p, miR-18a, and miR-93). In contrast, miR-9 was lower in both NDBE of progressors and paired HGD/EAC samples in comparison with NDBE of non-progressors.

Technical validation

Individual targeted assays were performed for the 11 aforementioned miRs, as well as one underrepresented miR (miR-139-3p) with a remarkable high ratio in TLDA analysis (Supplementary Tables 2 and 3). The miRs with lower levels in progressors were not

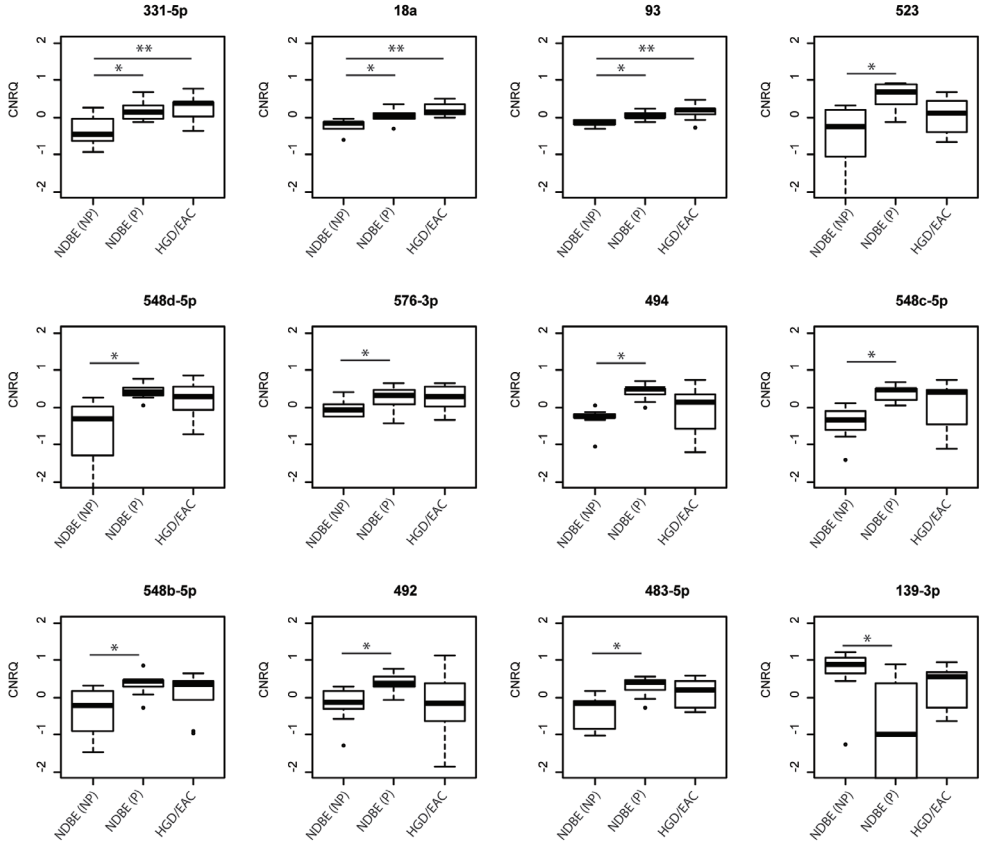


Figure 2. Boxplots with normalized levels of statistically significant differential microRNAs (miRs) for sample series 1 (TLDA). Sample groups are presented on the x-axis and calibrated normalized relative quantities (CNRQ) are shown on the y-axis. In total, 11 miRs showed higher levels in NDBE from progressors compared to non-progressors (FDR adjusted p-value <0.05 in *) NDBE from progressor vs. non-progressor and **) HGD/EAC vs. NDBE from non-progressor). Three miRs (miR-331-5p, miR-18a, and miR-93) showed similarly higher levels in both NDBE from progressors and HGD/EAC compared to NDBE from non-progressors. One miR (miR-139-3p) was downregulated in NDBE from progressors but showed a large spread in CNRQ.

further investigated as measuring downregulation was considered unpractical in a future diagnostic setting. All 11 miRs were confirmed as being upregulated in NDBE from progressors with comparable ratios as found in TLDA experiments (Figure 3, Supplementary Table 1). Single assays also confirmed the presence of similarly "shared" higher levels of the three previously mentioned miRs, i.e., miR-331-5p, miR-18a, and miR-93. In addition, miR-576-3p was similarly upregulated using single assay analysis but not with previous TLDA analysis. Only miR-139-3p could not be confirmed to be downregulated in the individual assay. Overall, based on these observations, data from TLDA experiments were considered reliable and accurate.

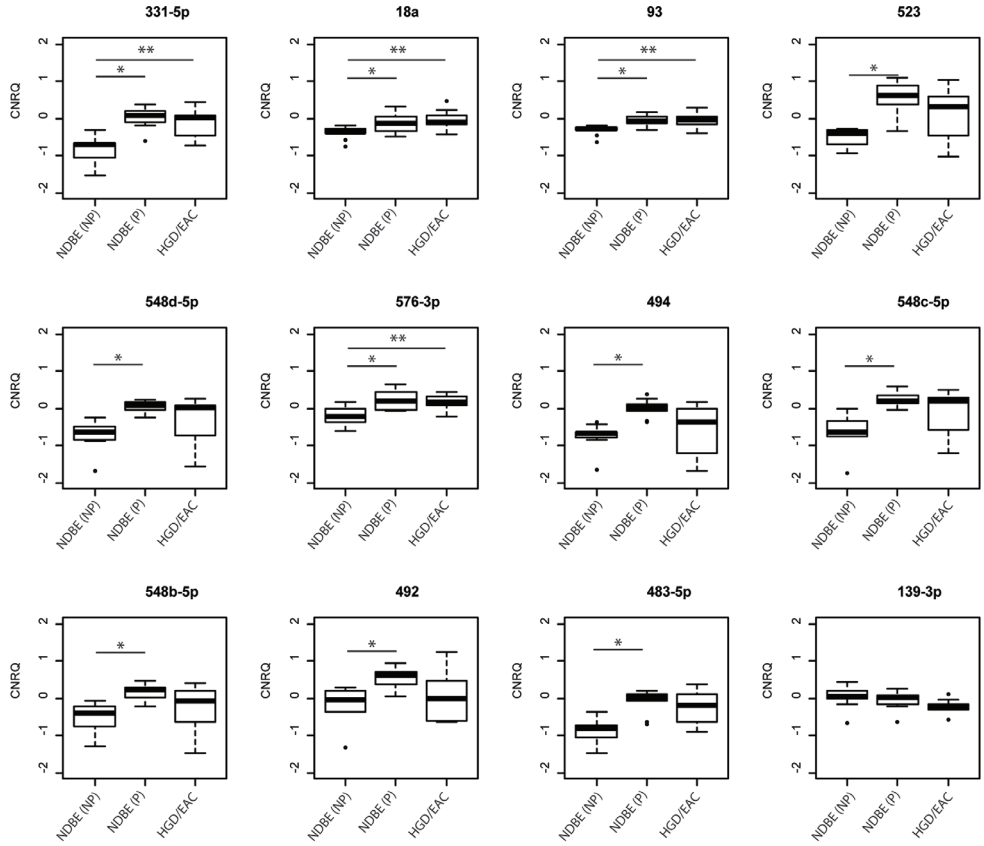


Figure 3. Boxplots with normalized miR levels of targeted individual assays for sample series 1. All 11 previously upregulated miRs from TLDA analysis were confirmed and showed similar ratio's. The previously downregulated miR 139-3p showed similar levels in all groups.

Independent validation set

Patient characteristics between series 1 and series 2 were comparable (Table 1). Apart from increased duration of longitudinal follow-up time, progressors did not differ in clinical characteristics from non-progressors (Supplementary Table 4). However, none of the miRs found in series 1 were significantly different using the same individual assays in series 2 (Supplementary Table 1). Subsequently, the RNA samples of series 2 were also subjected to high-throughput miR profiling using the same TLDA cards. Again, no miRs were found to be discriminative (Supplementary Table 1). Sample quality was similar between the two groups (Supplementary Figures 1 and 2).

Effect of different normalization strategy

A second normalization strategy was used to assess the effect thereof on the results. Using "global mean normalization", overall less miRs were found to be differentially expressed

(Supplementary Table 5). Again, no statistically significant differential levels were found between paired NDBE of progressors and HGD/EAC. In total, 19 miRs were statistically significant differentially expressed between NDBE from progressors and non-progressors for series 1. For series 2, again no significant differences were found between NDBE from progressors versus non-progressors.

DISCUSSION

Molecular and epigenetic alterations have been shown to precede histologically detectable dysplasia and potentially could be used to predict progression in non-dysplastic Barrett's esophagus.[47, 91-93] MiRs are able to repress the translation of mRNA expression of important cancer-related genes and therefore might be of use in the diagnosis and treatment of cancer. [54, 55, 94] The present study aimed to identify miRs that are associated with malignant progression in NDBE prior to the occurrence of histological changes, but failed to identify miRs with significant differential levels.

Both high-throughput miR-profiling and individual assays were used to investigate up to 381 miRs. In the discovery set, the results were promising, and 11 miRs showed a statistically significant higher level in NDBE of progressors versus non-progressors. These were technically validated and confirmed by individual assays, of which four miRs showed a similarly higher (miR-331-5p, miR-18a, miR-93) or lower (miR-9) level in both NDBE and HGD/EAC samples compared to NDBE from non-progressors. As such, these miRs in particular could potentially play an important role in BE-carcinogenesis, with altered miR regulation preceding histological progression. However, these promising findings were not confirmed in an independent validation set. In fact, both individual assays of the differential miRs found in the discovery set, as well as high-throughput profiling (TLDA) revealed no statistically significant differences in miR levels.

The outcome of miR research highly depends on the quality of the sample, the technical method applied for miR detection and data normalization. Normalization is crucial to correct for technical variation affecting miR levels between samples, often based on the measurement of one or more so-called reference genes ("housekeeping" genes), especially in the context of single assay experiments. For screening studies involving a large number of targets such as the TLDA methods applied in the described experiments, global mean normalization is considered as an appreciated alternative. When a relatively large number of targets are measured simultaneously, their average level reflects the input quantity and thus may be used for normalization. Nevertheless, to date there is no consensus on optimal normalization strategy, making comparison between studies difficult.[95, 96] Here, two slightly different global normalization strategies were used, with minimal effects on

overall outcome, demonstrating its limited impact in the results reported. Because clinical characteristics and sample quality between groups were comparable, our results indicate that, despite the use of so-called multiple testing corrections, the differential levels of miRs in the discovery set might be incidental findings. The presence of small (technical) variations in sample quality that could not be adequately corrected for using normalization could provide another explanation. For these considerations, it cannot be ruled out with certainty that particular miRs in NDBE from progressors actually are differentially regulated, but their practical value being limited. Given the heterogeneity in BE, it might be that the sampling protocol or timing (up to 88 months before histologically documented progression) was inappropriate for detection of molecular changes related to progression. To overcome the problem of sampling errors a recent study on miRs in samples obtained from the entire BE-segment using a non-endoscopic sampling device, called the *Cytosponge*, showed promising results, as they identified miRs that accurately discriminated NDBE from normal esophageal squamous tissue.[86] However, as longitudinal follow-up was lacking, progression related miRs could not be identified.

One other limitation of the present study is the relatively small sample size which is partly caused by the low progression rate in BE patients. Other clinical variables such as presence of esophagitis, history of smoking and PPI use (Table 1), were not considered in analysis although, in theory, these factors could be associated with changes in miR. Because of the small sample size, no multivariable model was created, due to possible overfitting. In addition, differences between progressors and non-progressors were only observed with regard to the duration of longitudinal follow-up time (Supplementary Table 4). Nevertheless, in an effort to increase patient numbers, combined analysis of TLDA data of series 1 and 2 was performed. This resulted in six miRs that were significantly different in NDBE (Supplementary Table 6). However, individual validation assays could only be confirmed for one miR (miR-548c-5p, Supplementary Table 7). Therefore, additional testing is essential to provide independent validation for these combined results (Supplementary Figure 3). Due to the small annual risk of developing EAC, this was not feasible in the available patient cohort.

Various studies previously described altered miR patterns in BE and EAC.[58] However, most of them were cross-sectional and did not make use of patient samples with known longitudinal follow-up. Only one previous study described progression related miRs in a small group of patients with NDBE (miR-192, miR-194, miR-196a and miR-196b, single assays, normalized with RNU-6B only).[59] Our data could not confirm these results by both TLDA and single assays (Supplementary Table 8), in both cohorts. Again, different normalization strategies might explain these differences, although no impact was identified in our study. The small non-coding RNA RNU-6B was used as normalizer, however, it has been argued that it is best

to normalize miRs with reference genes belonging to the same RNA class, as others may differ in their efficiency of extraction, reverse transcription and amplification.[97]

CONCLUSIONS

The present study does not provide evidence that histology based miRs from patients with NDBE could be informative to predict progression to HGD/ EAC. In addition, this study demonstrated the crucial role of an independent validation cohort to limit the number of unreproducible findings in literature.

AUTHOR CONTRIBUTIONS

Conceptualization, Leendert Looijenga and Katharina Biermann; Formal analysis, Lucia Suzuki, Ad JM Gillis, Lambert CJ Dorssers and Daan Nieboer; Investigation, Lucia Suzuki, Ad JM Gillis and Joitza Henriquez; Methodology, Lucia Suzuki, Lambert CJ Dorssers, Daan Nieboer, Leendert Looijenga and Katharina Biermann; Resources, Marco Bruno, Manon C. W. Spaander, Leendert Looijenga and Katharina Biermann; Software, Lambert CJ Dorssers; Supervision, Leendert Looijenga and Katharina Biermann; Validation, Lucia Suzuki, Ad JM Gillis and Lambert CJ Dorssers; Visualization, Lucia Suzuki and Lambert CJ Dorssers; Writing – original draft, Lucia Suzuki; Writing – review & editing, Lucia Suzuki, Ad JM Gillis, Lambert CJ Dorssers, Daan Nieboer, Marco Bruno, Manon C. W. Spaander, Leendert Looijenga and Katharina Biermann.

FUNDING/ CONFLICTS OF INTEREST

This study was funded by the Erasmus MC fellowship appointed to Dr. K. Biermann (KB) entitled "Barrett esophagus: improved prediction of progression by targeted risk stratification". The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results. The authors declare no conflict of interest.

ACKNOWLEDGMENTS

We thank Carlijn A. M. Roumans for helping us to collect the material and overall help regarding to the data obtained from the PROBAR cohort.

SUPPLEMENTARY MATERIALS

The following are available at request:

Supplementary Table 1. Raw data of TLDA results

Supplementary Table 2. Raw data of individual assay results

Supplementary Table 3. Individual assays confirm results obtained with TLDA in discovery set for 11 upregulated miRs in NDBE from progressors.

	NDBE (NP) vs NDBE (P) (n=9 vs n=9)		NDBE (NP) vs HGD/EAC (unpaired, n=9 vs n=9)	
	Ratio	p-value*	Ratio	p-value*
hsa-miR-523	0,09	***0,002***	0,235	0,149
hsa-miR-548d-5p	0,18	***0,000***	0,416	0,222
hsa-miR-576-3p	0,39	***0,016***	0,453	**0,029**
hsa-miR-494	0,17	***0,000***	0,716	0,600
hsa-miR-548c-5p	0,14	***0,000***	0,365	0,178
hsa-miR-548b-5p	0,22	***0,001***	0,668	0,233
hsa-miR-492	0,19	***0,001***	1,057	1,000
hsa-miR-331-5p	0,15	***0,000***	0,197	**0,015**
hsa-miR-483-5p	0,16	***0,001***	0,223	**0,015**
hsa-miR-18a	0,53	***0,040***	0,436	**0,015**
hsa-miR-93	0,57	***0,005***	0,533	**0,031**
hsa-miR-139-3p	1,25	0,515	1,911	0,075

All upregulated miRs identified with TLDA in series 1 were confirmed using targeted individual assays, including the three miRs (miR-331-5p, miR-18a and miR-93) with similarly higher levels in NDBE from progressors and HGD/EAC compared to NDBE from non-progressors. The results for the decreased miR 139-3p were not confirmed using the individual assay. *FDR adjusted p-value.[90]

Supplementary Table 4. Patient characteristics of progressors in comparison with non-progressors only differ in duration of longitudinal follow-up.

		All patiënts (n=34)		
		Non-progressors	Progressors	p-value*
		n=18	n=16	
Age (y)	Mean (SD)	61 (10)	64 (10)	0.211
Gender (n)	Female	5	6	0.545
	Male	13	10	
Body mass index (kg/m2)	Median (IQR)	26.9 (4)	26.7 (8)	0.313
Smoking (n)	No, never	10	5	0.278
	Yes, in the past	5	5	
	Yes, current	3	6	
PPI use (n)	No	1	3	0.233
	Yes	17	13	
BE length (cm)	Median (IQR)	3 (4)	5 (5)	0,070
Esophagitis (n)	No	13	11	0.708
	Yes	4	5	
Biopsies taken (n)	Median (IQR)	8 (5)	11 (4)	0.144
Follow Up Time (m)	Median (IQR)	121 (48)	31.5 (54)	**<0.001**

y, years; SD, standard deviation, n, number; IQR, interquartile range; m, months; HGD, high-grade dysplasia; EAC, esophageal adenocarcinoma

Supplementary Table 5. MiRs with statistically significant differential levels using the second normalization strategy ("global mean normalization").

	NDBE (NP) vs NDBE (P) (n=9 vs n=9)		NDBE (NP) vs HGD/EAC (unpaired, n=9 vs n=9)	
	Ratio	p-value*	Ratio	p-value*
hsa-miR-523	0,08	**0,049**	0,31	0,647
hsa-miR-487a*	0,08	**0,049**	0,57	0,630
hsa-miR-548d-5p	0,11	**0,008**	0,22	0,231
hsa-miR-494	0,17	**0,008**	0,64	0,647
hsa-miR-548c-5p	0,19	**0,019**	0,45	0,438
hsa-miR-548b-5p	0,20	**0,049**	0,48	0,438
hsa-miR-492	0,24	**0,028**	1,04	0,974
hsa-miR-597*	0,47	**0,042**	0,77	0,408
hsa-miR-18a	0,54	**0,022**	0,45	**0,033**
hsa-miR-93	0,56	**0,008**	0,54	0,106
hsa-miR-30b	1,47	**0,042**	1,51	**0,036**
hsa-miR-99b	1,55	**0,049**	1,31	0,438
hsa-miR-186	1,66	**0,042**	1,14	0,866
hsa-miR-411	1,74	**0,039**	1,95	0,135
hsa-miR-30c	1,75	**0,008**	1,36	0,161
hsa-miR-342-3p	1,93	**0,028**	1,22	0,894
hsa-miR-139-5p	1,95	**0,022**	2,02	**0,036**
hsa-miR-487b	2,24	**0,042**	4,04	0,135
hsa-miR-218	2,39	**0,008**	2,07	0,295
hsa-miR-19a	0,72	0,060	0,53	**0,019**
hsa-miR-335	0,75	0,478	0,56	**0,035**
hsa-miR-95	0,81	0,391	0,39	**0,035**
hsa-miR-339-5p	0,90	0,928	0,41	**0,035**
hsa-miR-885-5p	1,654	0,391	6,48	**0,035**

Ten miRs were upregulated (ratio <1) in NDBE from progressors compared to non-progressors in series 1, of which eight were similar as found with previous normalization strategy. Nine miRs were downregulated (ratio >1). Paired analysis of NDBE and HGD/EAC samples did not show any differences in miR levels (not shown). Overall, eight miRs were differential between HGD/EAC and NDBE from non-progressors. One miR (miR-18a) was similarly upregulated and two miRs (miR-30b, miR-139-5p) were similarly downregulated in NDBE from progressors and HGD/EAC compared to NDBE from non-progressors. No differential miRs were found in series 2 (not shown). *FDR adjusted p-value.[90]

Supplementary Table 6. Combined analysis of TLDA experiments from series 1 and 2 using both normalization strategies.

	Combined series; 1st normalization strategy		Combined series; 2nd normalization strategy	
	NDBE (NP) vs. NDBE (P) (n=18 vs n=16)		NDBE (NP) vs. NDBE (P) (n=18 vs n=16)	
	Ratio	p-value*	Ratio	p-value*
hsa-miR-548d-5p	0,20	**0,010**	0,20	**0,005**
hsa-miR-548c-5p	0,24	**<0,001**	0,24	**0,001**
hsa-miR-548b-5p	0,28	0,070	0,28	**0,035**
hsa-miR-523	0,08	0,070	0,08	**0,039**
hsa-miR-30b	1,42	**0,010**	1,43	**0,027**
hsa-miR-30c	1,57	**<0,001**	1,58	**0,008**
hsa-miR-99b	1,62	**0,040**	1,63	**0,034**
hsa-miR-139-5p	1,86	**<0,001**	1,87	**0,002**
hsa-miR-125a-5p	1,45	0,150	1,46	**0,048**
hsa-miR-186	1,54	0,060	1,54	**0,038**
hsa-miR-342-3p	1,57	0,100	1,58	**0,038**
hsa-miR-323-3p	1,73	0,060	1,71	**0,038**
hsa-miR-197	2,51	0,060	2,52	**0,034**
hsa-miR-19a	0,89	0,950	0,90	0,496
hsa-miR-135a	0,87	0,960	0,87	0,683
hsa-miR-135b	1,00	1,000	1,01	0,975
hsa-miR-335	1,01	0,970	1,02	0,885
hsa-miR-337-5p	1,40	0,820	1,41	0,666
hsa-miR-375	2,24	0,080	2,25	0,054
hsa-miR-885-5p	2,75	0,080	2,76	0,093
hsa-miR-215	1,33	0,890	1,34	0,776

Combined analysis of both series using both normalization strategies showed six significantly different miRs between BE from cases (BE-P) compared to controls (BE-NP), which were already found in analysis of series 1 alone, independent of normalization strategy used. *FDR adjusted p-value.[90]

Supplementary Table 7. Individual assays confirm significant difference for only one upregulated miR.

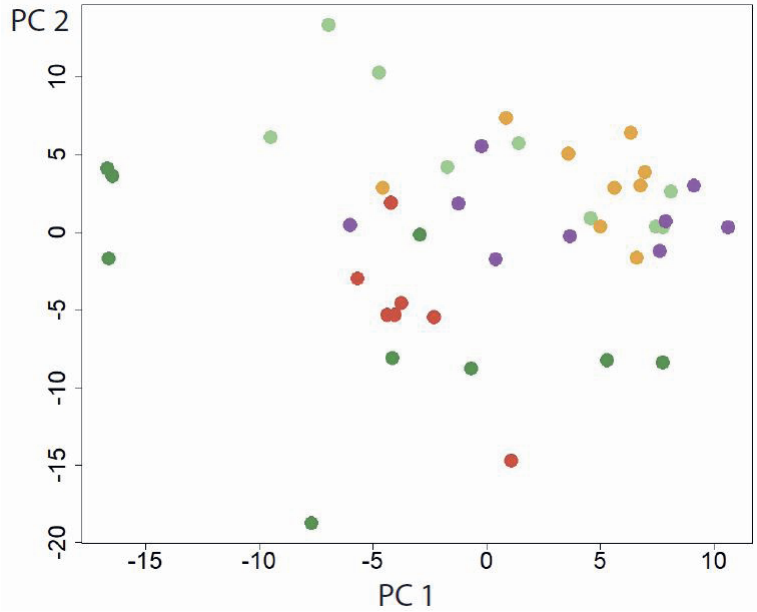
	Single Assay NDBE (NP) vs. NDBE (P) (n=18 vs n=16)	
	Ratio	p-value*
hsa-miR-523	0,26	0,166
hsa-miR-548d-5p	0,42	0,231
hsa-miR-548c-5p	0,33	**0,035**
hsa-miR-548b-5p	0,44	0,166

*FDR adjusted p-value.[90]

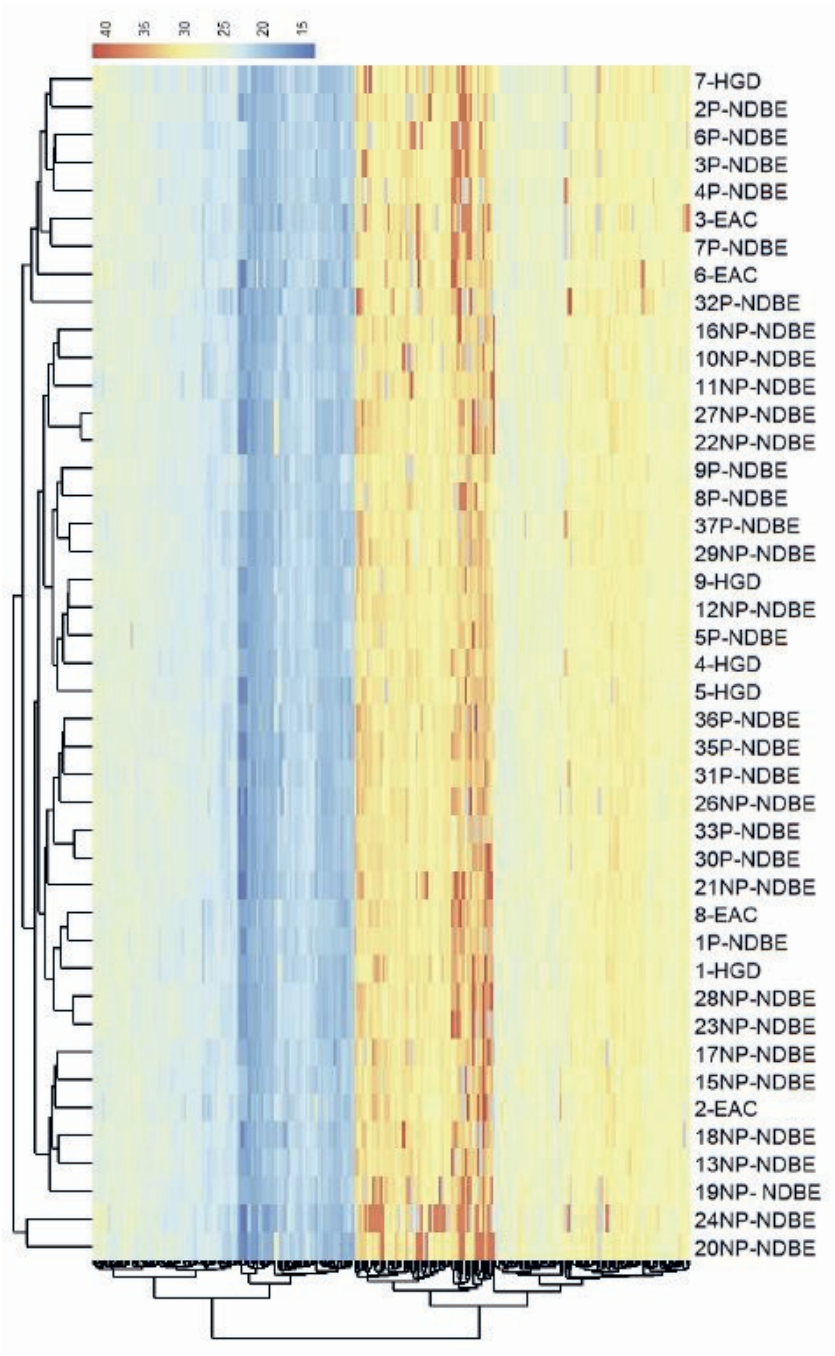
Supplementary Table 8. Discriminating miRs from literature were not confirmed** in the present study.[59]

	Series 1		Series 2		Combined series	
	NDBE (NP) vs. NDBE (P) (n=9 vs. n=9)		NDBE (NP) vs. NDBE (P) (n=9 vs. n=7)		NDBE (NP) vs. NDBE (P) (n=18 vs n=16)	
	Ratio	p-value*	Ratio	p-value*	Ratio	p-value*
TLDA						
hsa-miR-192	1,09	0,770	1,08	0,884	1,13	0,885
hsa-miR-194	0,94	0,928	1,15	1,000	1,08	0,975
hsa-miR-196b	1,01	0,928	0,82	0,751	0,96	0,885
Individual assay						
hsa-miR-192	1,16	0,515	1,26	0,468	1,20	0,397
hsa-miR-194	1,22	0,515	1,36	0,468	1,28	0,397
hsa-miR-196b	0,99	0,931	0,74	0,468	0,90	0,484

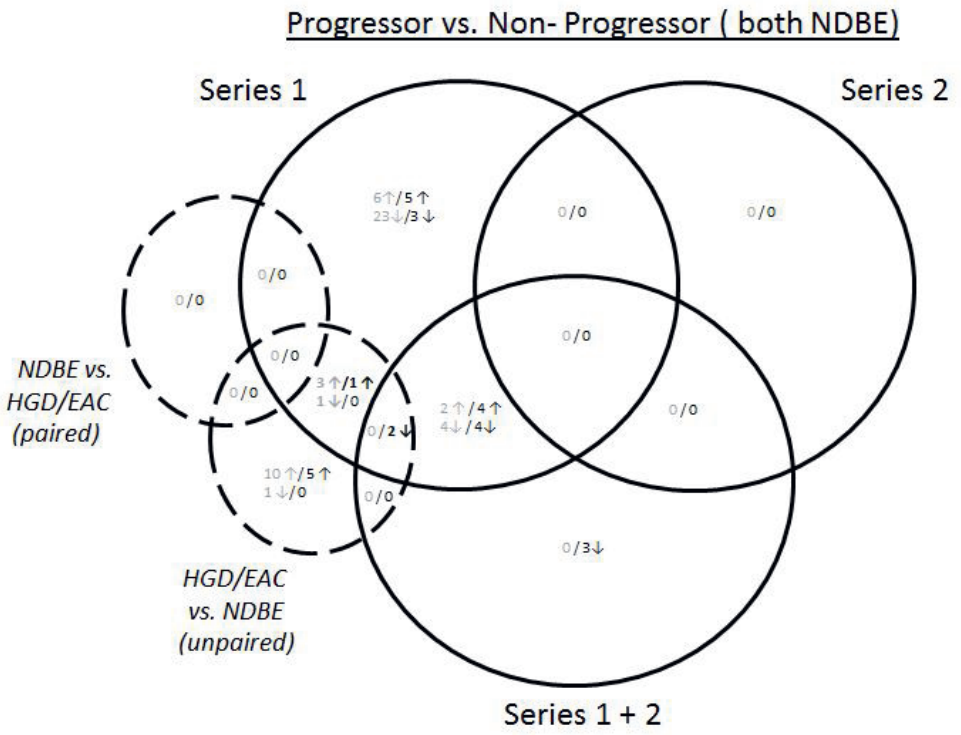
*FDR adjusted p-value.[90] ** One of four previously described miRs (miR-196a) was not present on our TLDA cards and was therefore not validated in the present study.



Supplementary Figure 1. Principal components analysis of normalized TLDA miR levels of both sample series with corresponding subgroups. The plot of the first two principle components (PC1 and PC2) reveal no separation of series 1 and 2 can be seen and HGD/EAC samples are located between BE samples, suggesting no differences in sample quality between both groups. Color codes for series 1: progressors, NDBE = orange, HGD/EAC = purple and non-progressors, NDBE = light green, series 2: progressors, NDBE = red, non-progressors, NDBE = forest green. Normalization was performed using the mean Ct of the top 197 measured miRs.



Supplementary Figure 2. Unsupervised clustering of 233 normalized miRTLDA levels shows no clear segregation between samples of series 1 and series 2. Series 1: progressors = 1-9P-NDBE; 1,4,5,7,9-HGD and 2,3,6,8-EAC, non-progressors = 10-13NP- and 15-20NP-NDBE. Series 2: non-progressors = 21-24NP- and 26-29NP-NDBE, progressors = 30-33P- and 35-37P-NDBE. Normalization was performed using the mean Ct of the top 197 measured miRs.



Supplementary Figure 3. Venn diagram of miRs with statistically significant differential levels identified using TLDA cards for both series 1, series 2 and the combined series. Large circles show miRs that are up- (↑) or down- (↓) regulated in NDBE of progressors versus non-progressors. Results from both normalization strategies are shown in grey (global mean normalization on common targets) and black (global mean normalization after exclusion of poor performing targets) respectively. Smaller dotted circles show miRs that are dysregulated in HGD/EAC versus NDBE of paired and unpaired samples for series 1.

4

Tumor budding is prognostic for metastases in patients with submucosal esophageal adenocarcinoma

Suzuki L*, ten Kate FJC*, Gotink AW**, van de Ven SEM**, Nieboer D, Weusten BLAM, Brosens LAA, van Hillegersberg R, Alvarez Herrero L, Seldenrijk CA, Alkhalaf A, Moll FCP, Schoon EJ, van Lijnschoten I, Tang T, van der Valk H, Nagengast WB, Kats-Ugurlu G, Plukker JTM, Houben MHMG, van der Laan J, Pouw RE, Bergman JJGHM, Meijer SL, van Berge Henegouwen MI, Wijnhoven BPL, de Jonge PJF, Doukas M, Bruno MJ, Looijenga LHJ, Koch AD***, Biermann K***

, **, and * Authors contributed equally*

Submitted

Background: Tumor budding (TB) is proposed to reflect epithelial-mesenchymal transition (EMT) and to be related to metastases formation. This study aimed to investigate whether TB assessment is of additional value in determining risk of metastases in pT1b esophageal adenocarcinoma (EAC).

Methods: Patients (n=252) from eight hospitals in the Netherlands, diagnosed with pT1b EAC and treated with endoscopic resection and/or surgery between 1989-2017 were included in this retrospective cohort study. Frequently reported methods of TB assessment (i.e., according to Ueno, Ohike and Thies) and automated digital methods were compared for interobserver reliability and prognostic relevance. Scoring in conventional H&E staining was compared to that in immunohistochemical staining. The risk of either lymph node metastases detected at surgery, or any metastases identified during follow-up, was used as clinical outcome measurement.

Results: Immunohistochemistry increased identification of tumor buds, although agreement between pathologists decreased compared with assessment on H&E for all methods. Prognostic power of pathologist-based TB methods was comparable regardless of the use of immunohistochemistry, and automated digitalized TB assessment was not prognostic of metastases. Only when assessed according to Ohike, high TB was associated with metastases, as well as with poor tumor grade, deep submucosal invasion and lymphovascular invasion both in the discovery (n=100), and validation (n=152) cohort. High TB remained an independent risk factor for metastases in multivariable analysis (n=252, HR 1.69, 95%CI 1.04-2.76).

Conclusions: In conclusion, TB assessment according to the Ohike method is a promising tool indicative for presence of metastases in pT1b EAC. However, external validation in additional studies is desirable.

INTRODUCTION

Patients with early stage esophageal adenocarcinoma (EAC) have a good prognosis, with a 5-year survival rate of 80%.[98] Survival is mainly determined by the absence or presence of regional lymph node metastases (LNM). The risk of LNM is low (0.2%) in tumors only invading the mucosa (pT1a) and therefore, these low-risk tumors are generally managed with endoscopic resection (ER) only.[99-101] No additional surgery is indicated for potential LNM that may be present.[42, 102]

Submucosal (pT1b) EAC harbor a higher risk for LNM, due to the presence of small vessels in the submucosa. This risk is estimated to be 21-78%,[103, 104] and consequently, their management is debated. Moreover, the risk increases when adverse characteristics, such as deeper tumor invasion, poor differentiation grade and lymphovascular invasion (LVI) are present.[41, 42, 61, 62, 102, 105-107]

Tumor budding (TB) is considered a promising biomarker for metastases in a variety of cancers.[63] A tumor bud is defined as a single tumor cell or a cluster of up to four tumor cells, without glandular differentiation.[60, 108] TB is proposed to reflect epithelial-mesenchymal transition (EMT). Presence of TB in a tumor has been associated with metastases and adverse survival, especially in early colorectal cancer (CRC).[60, 109-111] Hence, it is now common clinical practice to include TB in the pathological assessment of CRC.[60, 112, 113]

TB is described in EAC. However, the relevance of TB in pT1b EAC is unclear.[41, 61, 62, 105, 106] In addition, different methods to assess TB have been described, which might have an impact on the reported results.[60, 64, 65, 114-116] This study aims to determine whether TB is of additional value in risk stratification of patients with pT1b EAC.

METHODS

Patients

This retrospective multicenter cohort study has been approved by the medical ethical committee in the Erasmus MC and all participating hospitals (MEC-2016-050).

All patients diagnosed in the Netherlands between 1989-2017 with pT1b EAC and treated with ER and/or esophageal resection surgery were identified from Netherlands Cancer Registry. Patients were included from eight centers in the Netherlands. These were all expert centers with regard to the management of Barrett's esophagus and EAC. After meticulous review of the medical charts, patients with a pT1b EAC were included in the study when they met inclusion criteria (Figure 1). At least 12 lymph nodes in the resected specimen

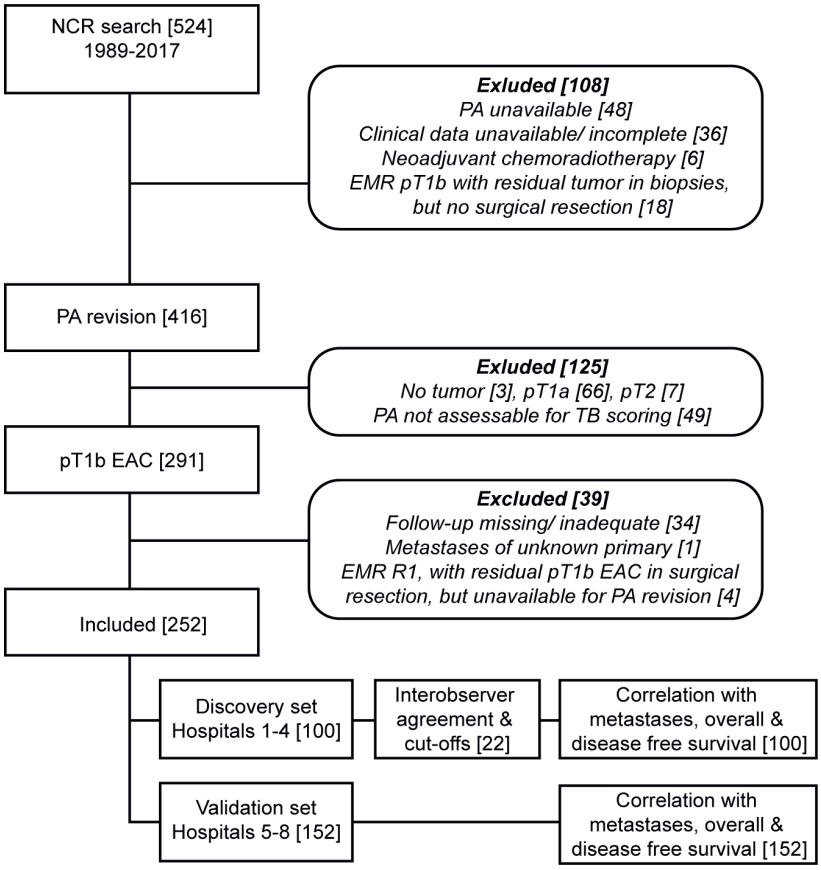


Figure 1. Flow diagram of the study. *NCR, National Cancer Registry; EAC, esophageal adenocarcinoma; HGD, high-grade dysplasia; pT1a, mucosal tumor only pT2, tumor invading muscularis propria; pT3, tumor invading adventitia; PA, pathology report/ slides/ tissue blocks; TB, tumor budding; EMR R1, endoscopic mucosal resection with positive vertical margins..*

were required to ensure a reliable pN-stage.[117] Patients treated with ER only, or those with fewer resected lymph nodes, were only included when they developed metastases during follow-up or, in the absence of metastases, complete follow-up data was available. Patients were excluded when treated with neoadjuvant chemoradiotherapy or when no PA material was available for revision. Furthermore, patients with positive vertical margins on EMR and biopsy proven residual tumor, but no subsequent surgery were excluded, as invasion depth was considered unreliable. Follow-up data was collected until November 2019. Patients without metastases, who were lost to follow-up within 1,5 years or those with an unknown cause of death within five years were also excluded.

Evaluation of histopathological characteristics

Three experienced GI-pathologists (FK, MD, KB) reviewed an initial 85 cases for tumor differentiation grade, presence of LVI, submucosal invasion depth and tumor stage based on WHO criteria,[113] as previously described in detail.[118] Interobserver agreement was good to excellent for differentiation grade and presence of LVI. Measurement of invasion depth showed moderate agreement.[118] Therefore, for the remaining patients included in this study, both LVI and grade were scored by one pathologist (FK). Invasion depth was determined during a consensus meeting (MD and KB), using both digitalized H&E, as well as double immunohistochemically pankeratin/ desmin stained slides (IHC). Invasion depth was further categorized in superficial (sm1, ≤500µm), intermediate (sm2, 501µm-1000µm), and deep (sm3, >1000µm) invasion according to the Paris classification.[119]

Study design

Patients were divided in a discovery and a validation cohort based on the hospital where they were treated. The discovery cohort was derived from the first four centers that had completed all data delivery.

Initially, two pathologists (KB, FK) evaluated TB using the most frequently reported methods to assess TB (Ueno, Ohike and Thies, see below for details) in a training set of 22 cases from the discovery cohort (n=100). TB was evaluated both on conventional H&E slides as well as IHC for all methods.[64, 65, 114-116] The most reliable methods, i.e. the two methods with highest interobserver agreement (as they showed comparable agreement), were evaluated in the entire discovery set by one pathologist (FK). In addition, all cases from the discovery set were subjected to different automated digital assessment methods (see below for details). Subsequently, only the method by which TB was independently associated with metastases, was scored by an independent investigator (LS) in the validation cohort (n=152), on digitalized (i.e. scanned) H&E slides. To ensure interobserver agreement was again reliable between this and previous observers, regardless the use of digitalized images, this observer assessed TB on the same training set using corresponding digitalized images.

Endpoint

Primary endpoint was the development of metastases defined as either regional LNM in the resection specimen after esophagectomy, or any metastases detected during follow-up. Secondary endpoints were disease free survival (DFS), defined as the time between surgery/endoscopic resection and the first recurrence of disease and overall survival (OS) defined as time between surgery/endoscopic resection and death.

Immunohistochemistry (IHC)

H&E stained slides and tissue blocks were retrieved from the archives of participating hospitals. From the tissue block corresponding with the H&E slide with deepest submucosal

invasion, 4 µm sections were cut and mounted on glass slides. The slides were deparaffinized, followed by heat-induced antigen retrieval using standard CC1 (Ventana Medical Systems) for 64 minutes in an automated slide staining system (BenchMark Ultra, Ventana Medical Systems). Subsequently, slides were incubated with pankeratin (32 minutes, pankeratin clone AE/AE3, dilution 1:800, Neomarkers, Fremont, CA, United States), Protease1 (eight minutes) and desmin (32 minutes, De-R-11 ready to use, Ventana Medical Systems, Roche, Tuscon, AZ, USA). For visualization, UltraView Universal Dap (Ventana Medical Systems) for keratin and UltraView Alkaline Phosphatase Red (Ventana Medical Systems) for desmin were used. Finally, slides were counterstained with hematoxylin.

Assessment of TB

The most frequently reported methods to assess TB in EAC, described by Ueno, Ohike and Thies, were compared in this study.[64, 65, 114-116] Based on the recommendations of an international consensus conference on TB in CRC (ITBCC) held in 2016[60], a tumor bud was defined as presence of a single tumor cell or a cluster of up to four tumor cells, completely surrounded by stroma and lacking any glandular formation. Ueno et al, first described the method recommended by the ITBCC for assessing and reporting TB in CRC.[60, 115] For this method, the invasive front is scanned with a 10x objective lens to determine the field (area of 0.785 mm^2, 20x objective) with the highest number of TB. The total number of TB is counted in this single hotspot area. In contrast, for the Ohike method, the entire invasive front is screened (at 10x magnification) and the number of fields (area of 0.785 mm^2, 20x objective) with at least five TB in the entire invasive front are counted.[114] For the Thies method, the number of TB is counted in one hotspot area of 0.189 mm^2 (40x objective) as well as in ten of these hotspots at the invasive front.[65] In Figure 2 these methods are all illustrated.

Digital TB assessment

All H&E and pankeratin stained slides were digitalized, by scanning them with a 40x lens in a single layer (Nanozoomer 2.0HT, Hamamatsu, the Netherlands). For all digital TB assessment methods (digital tumor bud count, DTBC), all scanned images of pankeratin stained slides were imported in Visiomorph (Visiopharm, Denmark). To ensure good quality images all images were checked and both the invasive border and the hotspot area of 0.785 mm^2 were manually delineated. In addition, a hotspot of 0.785 mm^2 was delineated by Visiomorph. Digital contrast was set to a minimum to identify epithelial areas only and to differentiate epithelium stained by pankeratin from the non-epithelial areas. In both selected hotspots (i.e. hotspot delineated by pathologist (FK), TB HS Path, and hotspot delineated by computer, TB HS Comp), and the entire delineated invasive front (TB/mm2), individually pankeratin stained areas of 60 μm^2 to 500 μm^2 were automatically marked and counted as TB. An example is shown in Figure 3. These aforementioned cut-offs were chosen after careful evaluation of cases in the training set where areas <60 μm^2 were considered

artifacts and areas >500 μm^2 did not qualify as TB. Large pankeratin positive areas as well as debris, macrophages were automatically excluded by dedicated image analysis software (at a 5x magnification).

Statistical analysis

High budding was defined according to previous literature: Ueno as five or more buds (H&E)[115], for Ohike as three or more budding fields (H&E)[114] and for Thies as 130 or more buds (ten fields, pankeratin).[65] Optimal cut-offs for the other, pankeratin based, methods were calculated by maximizing the Youden-index (Supplementary Table 1). Interobserver agreement between pathologists was analyzed using interclass correlation coefficients (ICC; two-way mixed model, absolute agreement). Strength of agreement was categorized as follows: poor, < 0.40; fair, 0.40–0.59; good, 0.60–0.74 and excellent, 0.75–1.00 (Supplementary Table 2).

Differences in categorical variables between groups were analyzed by chi-squared tests. Differences in continuous variables were evaluated with the Mann-Whitney U Test. To assess the prognostic value for risk of metastases and survival, Cox proportional hazard models were used. In multivariable analysis, adjustments were made for both clinical (i.e. age, gender) and histopathological characteristics including grade, LVI, invasion depth and for survival also pN-stage. Tumor size was not incorporated due to significant amount of missing or heterogeneous data. Moreover, when present, measurement was performed inconsistently during endoscopy, surgery, macro- or microscopy. Analyses were performed using SPSS-software (version 22, SPSS IBM Inc., Armonk, NY, USA) and R/R studio (version 3.6.2).

RESULTS

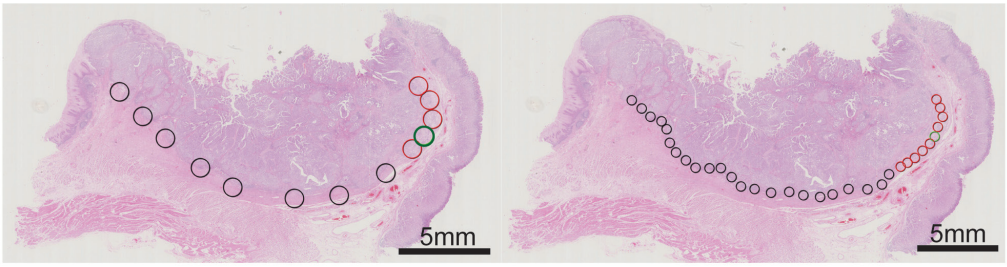
Patients

In total, 252 patients were eligible for inclusion in this study. Patient characteristics are listed in Table 1. Median age was 66 years (interquartile range, IQR: 58-72) and 87% was male. Patients were treated with either ER only, 36/252 (14%), ER followed by additional surgery, 50/252 (20%) or primary surgery, 166/252 (66%). When surgery was performed, LNM were found in 64 of 216 (30%) patients. The total percentage of patients with any metastases i.e. metastases detected at the time of surgery and metastases found during follow-up was 33% (83 out of 252 patients).

Most EAC were well (G1) or moderately (G2) differentiated (n=163, 65%) and in 57 cases (23%) LVI was identified. Patients in the discovery cohort (n=100) had more invasive tumors compared to the validation cohort (median 1700 vs. 1025 μm; p=0.007). Likewise, presence

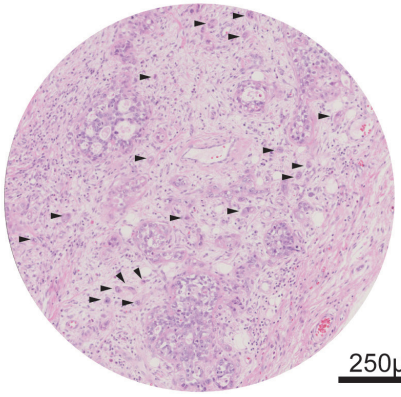
A **Methods used to assess tumor budding**

Scan invasive front at medium power (10x)



Ueno: Count tumor buds in "hotspot" area with most buds

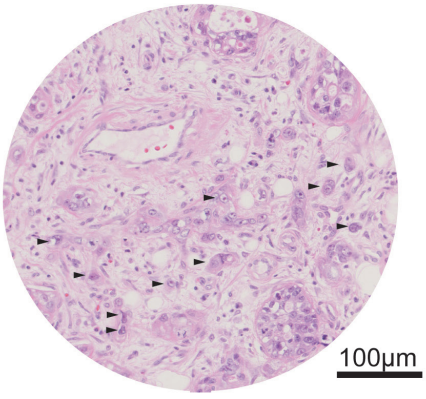
Ohike: Count budding fields with at least five buds



Area/ field: 0.785 mm² (20x objective)

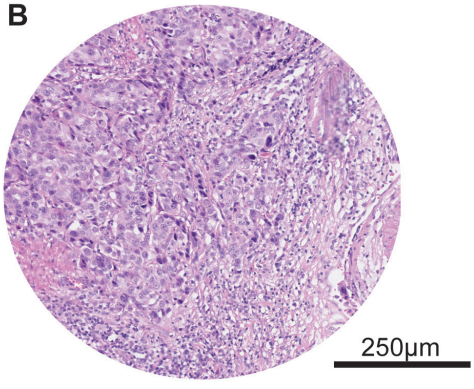
Thies (1): Count tumor buds in "hotspot" area with most buds

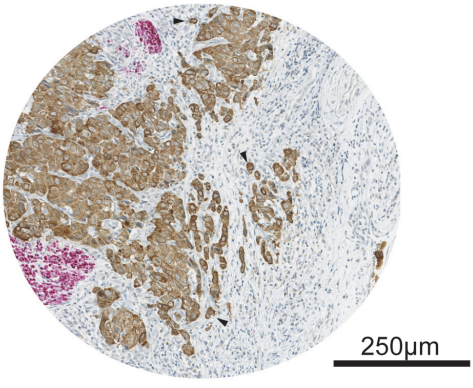
Thies (2): Count tumor buds in ten "hotspot" areas with most buds



Area: 0.189 mm² (40x objective)

B





< Figure 2. A) For all methods, the entire invasive front is scanned at medium power. For the Ueno method (recommended by the ITBCC[60] the total number of tumor buds (TB) in the hotspot area (green circle) with most buds are counted, whereas for the Ohike method, budding fields (of similar size as the hotspot area in the Ueno method) with at least five buds are counted (green and red circles).[60, 64, 114, 115] The Thies method is more or less comparable, although the area size in which buds are counted is smaller.[65] B) All previously described methods, were assessed both on H&E (left) as well as double pankeratin (epithelium= brown)/ desmin (smooth muscle= red) stained slides (right). Immunohistochemistry can be useful in cases where the tumor-stroma interface may be obscured (i.e. by inflammatory infiltrates) as TB are better identifiable on pankeratin stained slides compared to H&E slides. Given variable field diameters, application of a correction factor may be necessary to report the number of tumor buds in the equivalent of a 0.785-mm2 field. For this purpose, a conversion table may be used, as proposed by the ITBCC.[60] *ITBCC, international tumor budding consensus conference; H&E, hematoxylin and eosin*

of metastases was higher in the discovery cohort (38 % vs. 30%, p=0.165), and recurrences were more frequently observed in the discovery cohort (20% vs. 11%, p=0.049). Both cohorts were comparable regarding other clinicopathological characteristics.

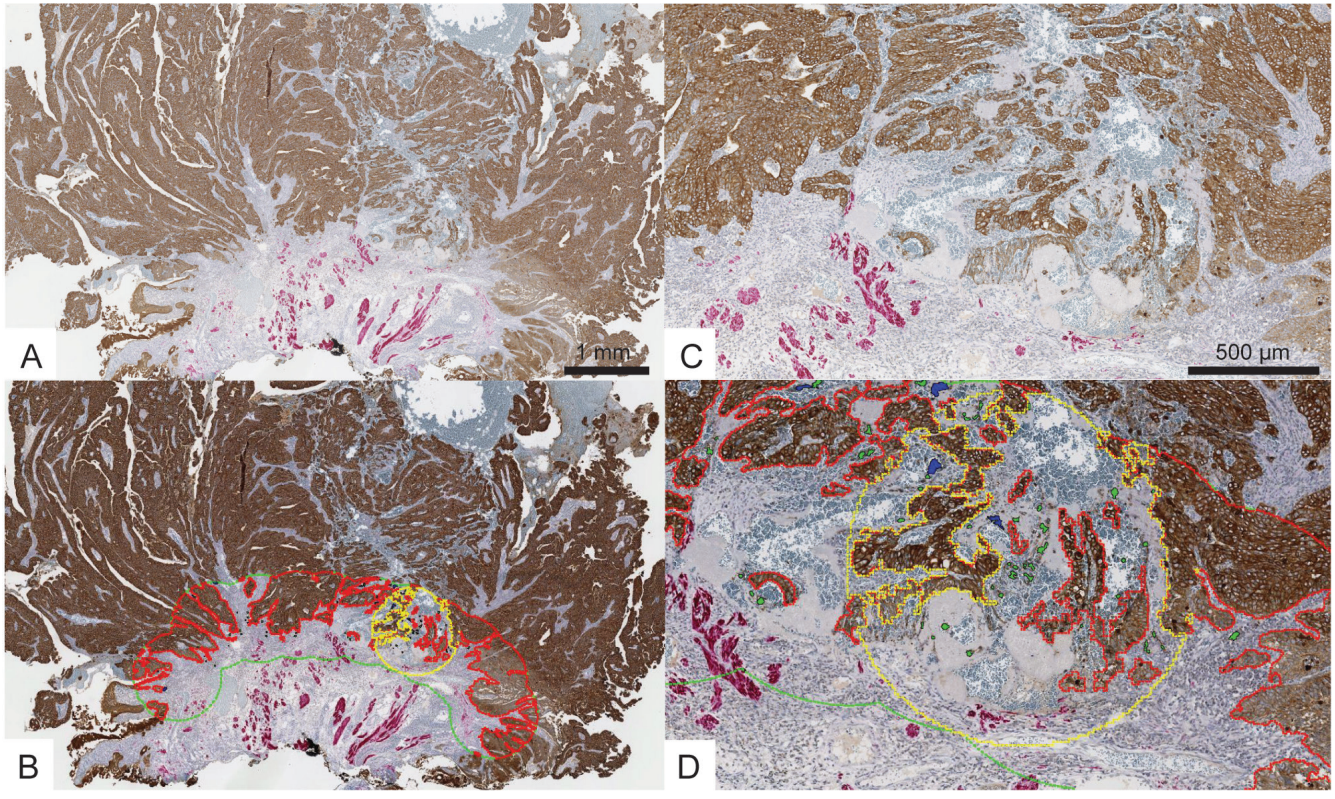


Figure 3. Example of automated digitalized tumor budding assessment. A) Double pankeratin/ desmin stained slides were digitalized and imported in Visiomorph (Visiopharm, Denmark). B) The entire invasive front was manually delineated (green area), whereas large fields of pankeratin positive epithelium were automatically recognized by the program (red areas). The hotspot area (field of 0.785 mm^2 at 20x magnification, with most tumor buds (TB)) was both manually selected by a pathologist (TB Hotspot Path, not shown), as well as by the computer (TB Hotspot Comp, yellow area). C, D) Detail of the hotspot area selected by the computer (yellow in D) in which the green dots are the tumor cells identified and counted as TB. Furthermore, automatically marked TB were counted in the entire invasive front (TB/mm^2, not shown).

Table 1. Clinicopathological characteristics.

Patient characteristics		Entire cohort (n=252)		Discovery cohort (n=100)		Validation cohort (n=152)		
		n	%	n	%	n	%	p-value*
Age, years	Median, IQR	66	58-72	65	58-72	66	59-72	0.699¥
Gender	Male	218	87	86	86	132	87	0.848
	Female	34	13	14	14	20	13	
Treatment	ER	36	14	12	12	24	16	0.343
	ER + Surgery	50	20	24	24	26	17	
	Surgery	166	66	64	64	102	67	
Grade	Well/ Moderate	163	65	66	66	97	64	0.723
	Poor	89	35	34	34	55	36	
LVI	No	195	77	7 6	76	119	78	0.671
	Yes	57	23	24	24	33	22	
Invasion depth, µm Median, IQR		1335	641-2500	1700	913-2775	1025	445-2407	**0.007¥**
Invasion depth, µm	Sm1, <500	52	21	12	12	40	26	**0.002**
	Sm2, 501-1000	49	19	15	15	34	22	
	Sm3, >1000	151	60	73	73	78	51	
cN-stage	cNx	22	9	5	5	17	11	0.076
	cN0	178	71	76	76	102	67	
	cN1	51	20	19	19	32	21	
	cN2	1	0	0	0	1	1	
	cN3	0	0	0	0	0	0	
pN-stage	pNx	38	15	13	13	25	16	0.442
	pN0	150	60	62	62	88	58	
	pN1	43	17	14	14	29	19	
	pN2	18	7	9	9	9	6	
	pN3	3	1	2	2	1	1	
Metastases∞	No	169	67	62	62	107	70	0.165
	Yes	83	33	38	38	45	30	
	No	169		62	62	107	70	0.058
	LNM	51		21	21	30	20	
	Distant	10		8	8	2	1	
	LNM + distant	22		9	9	13	9	
Recurrence	No	214	85	79	79	135	89	**0.049**
	Yes	37	15	20	20	17	11	
OS, months	Median, IQR	58	28-89	58	26-82	58	30-91	0.916¥
DFS, months	Median, IQR	56	25-89	56	20-82	56	27-90	0.959¥
Tumor budding	Low	153	61	48	48	105	69	**0.001**
(Ohike H&E)	High	99	39	52	52	47	31	

*Chi-squared test unless indicated otherwise. ¥Mann-Whitney U test; ∞Lymph node metastases at surgery and/ or metastases during follow-up; *N, number of cases; IQR, interquartile range; ER endoscopic resection; LVI, lymphovascular invasion; sm, submucosal; OS, overall survival; DFS, disease free survival; H&E, hematoxylin and eosin.*

Interobserver agreement

In all three methods assessed by the two pathologists, application of IHC increased the identification of TB, but decreased agreement between pathologists compared to assessment on H&E. The Ueno (ICC H&E, 0.962; IHC, 0.755) and Ohike (ICC H&E, 0.907; IHC, 0.859) methods both showed comparable, highest agreement in the training set (see Supplemental Table 2) and were further evaluated in the entire discovery cohort. Both Thies methods showed lower agreement and were therefore disregarded from further analysis in the discovery cohort. The independent observer used for the validation cohort, assessed the same training set, and again excellent interobserver agreement was reached (ICC H&E, 0.809).

TB and metastases risk

We first evaluated the association of primary clinicopathological characteristics including TB with prevalence of metastases in the discovery cohort (n=100). In univariable analysis poor differentiation grade, submucosal invasion depth, presence of LVI and TB according to all but the Ueno (H&E) and TB HS Path methods, were all associated with higher prevalence of overall metastases (Supplementary Table 3). In multivariable analysis, high TB according to the Ohike method was the only method independent of other risk factors and associated with a 2.5 fold increased risk (95% CI 1.03-5.98) of metastases when assessed on IHC and 1.8 fold increased risk (95% CI 0.80-3.92) when assessed on H&E (Supplementary Table 4).

In the validation cohort (n=152), LVI, increased invasion depth and high TB according to Ohike (H&E), but not poor grade, were associated with presence of metastases. In multivariable analysis, only LVI and invasion depth remained prognostic when corrected for other variables (Supplementary Table 5).

When both cohorts were analyzed together (n=252), TB was associated with all previously described risk factors (Table 2) and remained prognostic for metastases (HR 1.7, 95% CI 1.04-2.76) when corrected for LVI (HR 3.3, 95%CI 2.11-5.29), invasion depth (HR 1.1, 95% CI 1.03-1.17) and grade (HR 0.96, 95%CI 0.61-1.51) (Table 3).

TB and survival

Median DFS was 56 months (IQR 25-89 months) and median OS was 58 months (IQR 28-89). When patients were stratified based on TB, patients with high TB according to Ohike had shorter DFS ($p=0.023$) and OS ($p<0.001$) (Figure 4). Indeed, high TB as well as increased age, poor grade, LVI, invasion depth and presence of LNM at time of surgery (pN+) were all associated with adverse OS. However, in multivariable analysis only age, presence of LVI and pN+ were independent prognostic factors. Similarly, although LVI, pN+ and TB (but not grade and invasion depth) were significant adverse prognostic factors for DFS in univariable

analysis, only LVI and pN+ remained independently associated when correcting for other variables (Supplementary Table 6).

Table 2. Pathologic features and metastases according to tumor budding status assessed by the Ohike method (H&E).

	Tumor budding		
	Low, n=153 n (%)	High, n=99 n (%)	p-value*
Grade			
Well/ Moderate	118 (77)	45 (45)	**<0.001**
Poor	35 (23)	54 (55)	
LVI			
No	129 (84)	66 (67)	**0.001**
Yes	24 (16)	33 (33)	
Invasion depth, µm			
Sm1, <500	40 (26)	12 (12)	**<0.001**
Sm2, 501-1000	41 (27)	8 (8)	
Sm3, >1000	72 (47)	79 (80)	
Metastases**			
No	119 (78)	50 (51)	**<0.001**
Yes	34 (22)	49 (49)	

*Chi-squared test. ** Lymph node metastases at surgery and/ or metastases during follow-up; *n, number of cases; LVI, lymphovascular invasion; sm, submucosal; H&E, hematoxylin and eosin.*

Table 3. Tumor budding according to the Ohike method is associated with an increased risk of metastases independent of other unfavorable histological characteristics in the entire cohort.

	Metastases*		Univariable COX regression		Multivariable COX regression	
	No, n (%)	Yes, n (%)	HR (95% CI)	p-value	HR (95% CI)	p-value
Grade						
G1/ G2	117 (72)	46 (28)	Ref.	**0.036**	Ref.	0.594
G3	52 (58)	37 (42)	1.59 (1.03-2.45)		0.88 (0.55-1.41)	
LVI						
No	151 (77)	44 (23)	Ref.	**<0.001**	Ref.	**<0.001**
Yes	18 (32)	39 (68)	3.70 (2.39-5.73)		3.34 (2.11-5.29)	
Invasion depth**						
Median, IQR	1010 (470-2195)	2100 (1140-3450)	1.12** (1.06-1.19)	**<0.001**	1.10** (1.03-1.17)	**0.004**
Tumor budding						
Low	119 (78)	34 (22)	Ref.	**0.001**	Ref.	**0.034**
High	50 (51)	49 (49)	2.52 (1.62-3.91)		1.69 (1.04-2.76)	

*Age and gender were not significant in univariable analysis (p > 0.2). *Lymph node metastases at surgery and/ or metastases during follow-up. **Continuous variable, HR invasion depth per 500 µm increase. n, number of cases; HR, hazard ratio; CI, confidence interval; Ref., reference; LVI, lymphovascular invasion.*

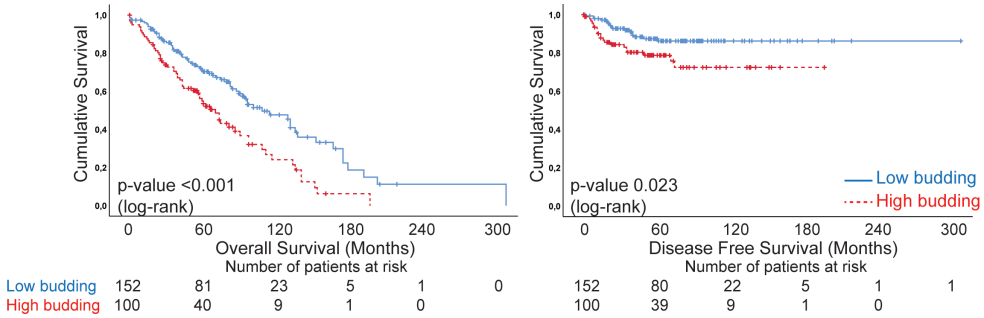


Figure 4. Kaplan-Meier curves for overall survival and disease free survival in the entire combined cohort stratified by tumor budding according to the Ohike method[114] indicate high budding associates with adverse survival.

DISCUSSION

Treatment strategy and outcome in patients with pT1b EAC are determined by the risk of metastases weighed against the risk of harm due to additional surgical treatment. Our results support the use of TB for improved risk stratification. In early pT1a CRC, high TB is an increasingly recognized predictor of LNM status (relative risk 5.1, 95% CI 3.6-7.3) and associated with poor survival in stage II CRC (HR: 3.7, 95% CI 2.2-6.3).[60, 63, 113, 115, 120]

In this study, frequently reported methods to assess TB were compared for their reliability and association with metastases and survival in a national cohort of patients with pT1b EAC. The Ueno and Ohike method were considered most reliable as they showed comparable, highest interobserver agreement. Both methods were associated with overall metastases in the discovery cohort, regardless the use of IHC, but only the Ohike method (IHC-based) was independently associated when correcting for other risk factors.

To date, the use of IHC is debated as relatively few have evaluated TB by IHC. It is currently recommended to be used as an adjunct only. For example, when a prominent inflammatory infiltrate is present, as this may obscure TB (Figure 2B).[60] In this study, prognostic strength was comparable regardless of the use of IHC. However, although IHC increased the identification of TB, it decreased interobserver agreement for all methods. This finding has been described previously,[121] and may be explained by staining of apoptotic tumor cells or small residual ductal structures destroyed by inflammation, and highlighted by IHC. These problems are usually well interpreted by a pathologist, but a computer will falsely identify these as TB, explaining the lack of association with metastases of the digital approaches tested. In line with this hypothesis, a recent study based on presentation of 3000 TB identified by pankeratin-based digital image analysis algorithms, evaluated by seven members of the ITBCC, showed only full agreement in 34%.[122]

Therefore, we aimed to validate the Ohike method based on conventional H&E instead of IHC stained slides. Generally, the same results were found as in the discovery cohort, except that TB was not significantly prognostic anymore. This could be due to the use of H&E, as well as the sample size and small differences between both cohorts. Patients in the discovery cohort had more invasive tumors, more metastases, more recurrences and more often high TB. In the entire cohort, high TB according to the Ohike method remained an independent prognostic predictor for metastases. Interestingly, the increased risk of high TB is most apparent in the absence of LVI.

Only one previous study described TB in early EAC and showed high TB (according to the Ohike method) to be prognostic for LNM and survival.[64] However, this study included patients with both mucosal (pT1a) and submucosal (pT1b) EAC, all treated with surgical

resection. Our study excluded patients with pT1a cancers, because of the known lower risk of metastases, for which ER is recognized as a safe and efficient treatment. In addition, our study included patients treated with ER only, as exclusion of these patients may result in selection bias towards patients with poor prognostic histological features. Because no lymph nodes were resected in patients who were only treated endoscopically, complete follow-up was required.

Surprisingly, the Ueno method, recommended by the ITBCC to be used for CRC reporting, was not associated with metastases when correcting for other risk factors in our study. The main difference between the Ueno and the Ohike method is the area size in which TB is assessed. Furthermore, the ITBCC recommends a 3-tiered system (low, 0-5; intermediate, 6-9; high, 10 or more TB), while we used the dichotomous system (high: five or more TB) that was originally described by Ueno et al.[108, 115] However, no difference in results were identified using either of these approaches. In the Ohike method, the number of budding fields, of similar size as a "hotspot" in the Ueno method, are counted with at least five buds. This method is therefore considered more representative of the entire invasive front. Conversely, the ITBCC recommends the hotspot method, as counting multiple fields may 'dilute' final TB score in cases with extensive focal TB. Consequently, this is the method used in the vast majority of CRC studies.[60] To our knowledge, the Ohike method has not been evaluated in CRC.

Limitations of the present study include its retrospective study design. Furthermore, there is a chance of overestimating the number of metastases by excluding patients without metastases but with incomplete follow-up. In addition, tumor size was not taken into account in multivariable analysis due to large amount of missing data. Tumor size may be particularly important for the Ohike method as in small tumors obviously, less field areas are present. Furthermore, even more methods have been described to assess TB, especially in CRC.[60, 63, 110, 123, 124] On the other hand, this is the first large study on TB in pT1b EAC, comparing different frequently described methods to assess TB, both assessed on H&E and IHC and including digitalized automated methods.

In summary, TB assessment according to the Ohike method (H&E) is reproducible and improves risk stratification in pT1b EAC, independently from other commonly used risk factors. The clinical relevance of TB in pT1b EAC needs to be confirmed in independent studies.

SUPPLEMENTARY MATERIALS

Supplementary Table 1. Cut-off values for tumor budding (TB) on conventional H&E and immunohistochemically (IHC, pankeratin) stained slides for the different methods used to assess TB, as well as for the automated digital TB count.

	High tumor budding	Positive if greater than or equal to	Sensitivity	1-Specificity	Specificity	Youden index
H&E						
Ueno	≥ 5 buds	Based on previous literature[115]				
Ohike	≥ 3 budding fields	Based on previous literature[114]				
Thies, 1 field	≥ 14 buds	13	0,71	0,067	0,933	0,647
Thies, 10 fields	≥ 52 buds	51,5	0.71	0,067	0,933	0,647
IHC						
Ueno	≥ 38 buds	37,75	0,71	0,133	1	0,581
Ohike	≥ 7 budding fields	6,5	0,71	0,133	1	0,581
Thies, 1 field	≥ 15 buds	14,25	0.86	0,2	0,8	0,657
Thies, 10 fields	≥ 130 buds	Based on previous literature[65]				
TB/mm^2	≥ 21 buds	20,5	0,71	0,267	1	0,447
TB Hotspot Comp	≥ 34 buds	33	0,71	0,2	1	0,514
TB Hotspot Path	≥ 50 buds	49,5	0,71	0,2	1	0,514

H&E, hematoxylin and eosin; IHC, immunohistochemistry (pankeratin)

Supplementary Table 2. Interobserver agreement for assessing tumor budding using different methods.[65, 114, 115]

Two observers (KBI, FK)	ICC	95% CI	Interpretation
H&E			
Ueno, buds	0.962	0.909-0.984	excellent
Ohike, budding fields	0.907	0.791-0.960	excellent
Thies, 1 field	0.914	0.803-0.964	excellent
Thies, 10 fields	0.737	0.469-0.881	good
IHC			
Ueno, buds	0.755	0.498-0.890	excellent
Ohike, budding fields	0.859	0.525-0.949	excellent
Thies, 1 field	0.739	0.470-0.882	good
Thies, 10 fields	0.673	0.357-0.850	good
Three observers (KBI, FK, LS)			
Ohike, budding fields H&E	0.809	0.654-0.908	excellent

ICC, intraclass correlation coefficient; CI, confidence interval; H&E, hematoxylin and eosin; IHC, immunohistochemistry (pankeratin)

Supplementary Table 3. Univariable COX regression analysis for metastases in the discovery set.

	Metastases		Discovery set (n=100)		
Variable (first = ref.)	No, N (%)	Yes, N (%)	HR	95% CI	p-value
Age, years*	65 (58-73)	66 (57-72)	0.98	0.84-1.16	0.842
Gender, male vs. female	52 (60) 10 (71)	34 (40) 4 (29)	0.60	0.21-1.69	0.334
H&E					
Grade, Well/ Moderate vs. Poor	48 (73) 14 (41)	18 (27) 20 (59)	2.86	1.50-5.45	**0.001**
LVI, absent vs. present	60 (79) 2 (8)	16 (21) 22 (92)	7.72	3.92-15.2	**<0.001**
Invasion depth, µm* median, IQR	1500 (830-2425)	2095 (1275-3645)	1.15	1.05-1.26	**0.002**
Ueno, ≤4 vs ≥ 5 buds	26 (74) 36 (55)	9 (26) 29 (45)	2.01	0.95-4.25	0.069
Ohike, ≤2 vs ≥ 3 budding fields	38 (79) 24 (46)	10 (21) 28 (54)	3.10	1.50-6.40	**0.002**
IHC					
Ueno ≤37 vs ≥ 38 buds	41 (72) 21 (49)	16 (28) 22 (51)	2.14	1.11-4.08	**0.022**
Ohike ≤6 vs ≥ 7 budding fields	43 (83) 19 (39)	9 (17) 29 (60)	4.73	2.23-10.1	**<0.001**
TB/mm^2 ≤20 vs ≥ 21 buds	35 (76) 27 (55)	11 (24) 22 (45)	2.23	1.08-4.62	**0.030**
TB HS Comp ≤33 vs ≥ 34 buds	35 (73) 27 (54)	13 (27) 23 (46)	2.08	1.05-4.13	**0.035**
TB HS Path ≤49 vs ≥ 50 buds	39 (65) 23 (61)	21 (35) 15 (39)	1.13	0.58-2.20	0.710

*Continuous variable, HR age per 5 years. HR invasion depth per 500 µm increase. *Ref., reference; N, number of cases; HR, hazard ratio; CI, confidence interval; LVI, lymphovascular invasion; H&E, hematoxylin and eosin; IHC, immunohistochemistry (pankeratin); TB, tumor budding.*

Supplementary Table 4. Multivariable COX regression analysis for metastases in the discovery set.*

Variable (first = ref.)	HR	95% CI	p-value
H&E			
Ueno, ≤4 vs ≥ 5 buds	1.25	0.57-2.73	0.573
Ohike, ≤2 vs ≥ 3 budding fields	1.77	0.80-3.92	*0.162*
IHC			
Ueno ≤37 vs ≥ 38 buds	1.06	0.52-2.16	0.869
Ohike ≤6 vs ≥ 7 budding fields	2.48	1.03-5.98	**0.038**
TB/mm^2 ≤20 vs ≥ 21 buds	1.10	0.48-2.52	0.817
TB HS Comp ≤33 vs ≥ 34 buds	1.53	0.72-3.25	0.274

*Adjusted for grade, lymphovascular invasion and invasion depth as continuous variable. *Ref., reference; HR, hazard ratio; CI, confidence interval; H&E, hematoxylin and eosin; IHC, immunohistochemistry (pankeratin)*

Supplementary Table 5. Univariable and multivariable COX regression analysis for metastases in the validation set.

	Metastases		Validation set (n=152)		
Variable (first = ref.)	No, n (%)	Yes, n (%)	HR	95% CI	p-value
Univariable COX regression analysis					
Age, years* median, IQR	66 (58-72)	67 (60-72)	1.04	0.89-1.22	0.606
Gender, male vs. female	91 (69) 16 (80)	41 (31) 4 (20)	0.64	0.23-1.78	0.392
Grade, Well/ Moderate vs. Poor	69 (71) 38 (69)	28 (29) 17 (31)	1.09	0.59-1.98	0.792
LVI, absent vs. present	91 (76) 16 (48)	28 (24) 17 (52)	2.29	1.25-4.19	**0.007**
Invasion depth, µm* median, IQR	865 (348-2140)	2100 (925-3230)	1.10	1.02-1.19	**0.013**
Ohike, ≤2 vs ≥ 3 budding fields	81 (77) 26 (55)	24 (23) 21 (45)	2.06	1.14-3.70	**0.016**
Multivariable COX regression analysis					
Grade, Well/ Moderate vs. Poor			0.71	0.37-1.36	0.303
LVI, absent vs. present			2.23	1.19-4.21	**0.013**
Invasion depth, µm*			1.10	1.01-1.19	**0.030**
Ohike, ≤2 vs ≥ 3 budding fields			1.61	0.84-3.09	0.151

*Continuous variable, HR age per 5 years. HR invasion depth per 500 µm increase. *Ref., reference; N, number of cases; HR, hazard ratio; CI, confidence interval; H&E, hematoxylin and eosin; IHC, immunohistochemistry (pankeratin); LVI, lymphovascular invasion.*

Supplemental Table 6. Tumor budding is not prognostic for survival.

Variable (first = ref.)	Overall Survival (n=252)			Disease Free Survival (n=252)		
	HR	95% CI	p-value	HR	95% CI	p-value
Univariable COX regression analysis						
Age, years*	1.21	1.10-1.33	**<0.001**	1.01	0.85-1.19	0.949
Gender, male vs. female	0.67	0.37-1.22	0.187	0.50	0.15-1.63	0.252
Treatment			0.211			0.247
ER vs. ER + Surgery	0.74	0.36-1.51	0.407	1.78	0.44-7.11	0.417
ER vs. Surgery	1.21	0.72-2.04	0.477	2.57	0.78-8.45	0.121
Grade, Well/ Moderate vs. Poor	1.48	1.04-2.11	**0.028**	1.65	0.86-3.17	0.130
LVI, absent vs. present	2.66	1.82-3.90	**<0.001**	4.58	2.40-8.75	**<0.001**
Invasion depth, µm*	1.05	1.00-1.11	**0.033**	1.08	0.99-1.17	0.088
pN-stage,			**<0.001**			**<0.001**
pN0 vs. pN+	2.93	1.99-4.29	**<0.001**	4.31	2.18-8.52	**<0.001**
pN0 vs. pNx	1.21	0.70-2.08	0.498	0.69	0.20-2.39	0.558
Ohike, ≤2 vs ≥ 3 budding fields	1.88	1.33-2.66	**<0.001**	2.08	1.09-3.98	**0.026**
Multivariable COX regression analysis						
Age, years*	1.23	1.12-1.35	**<0.001**			
Grade, G1/2 vs. G3	1.28	0.87-1.88	0.217			
LVI, absent vs. present	1.90	1.26-2.88	**0.002**	3.11	1.54-6.26	**0.002**
Invasion depth, µm*	1.02	0.96-1.07	0.594			
pN-stage			**0.001**			**0.007**
pN0 vs. pN+	2.27	1.46-3.52	**<0.001**	2.94	1.39-6.20	**0.005**
pN0 vs. pNx	1.04	0.59-1.81	0.898	0.69	0.20-2.37	0.551
Ohike, ≤2 vs. ≥ 3 budding fields	1.10	0.72-1.68	0.672	1.12	0.56-2.26	0.752

*Continuous variable, HR age per 5 years. HR invasion depth per 500 µm increase. *Ref., reference; N, number of cases; HR, hazard ratio; CI, confidence interval; ER, endoscopic resection; LVI, lymphovascular invasion.*

5

Olfactomedin 4 (OLFM4) expression is associated with nodal metastases in esophageal adenocarcinoma

Suzuki L, Ten Kate FJC, Gotink AW, Stoop H, Doukas M, Nieboer D, Spaander MCW, van Lanschot JJB, Wijnhoven BPL, Koch AD, Bruno MJ, Looijenga LHJ, Biermann K.

PLoS One. 2019;14(7):e0219494.

Background: To date no informative biomarkers exist to accurately predict presence of lymph node metastases (LNM) in esophageal adenocarcinoma (EAC).

Methods: We studied the discriminative value of Olfactomedin 4 (OLFM4), an intestinal stem cell marker, in EAC. Patients who had undergone esophagectomy as single treatment modality for both advanced (pT2-4) and early (pT1b) adenocarcinoma of the esophagus or gastro-esophageal junction were selected for this study from an institutional database (Erasmus MC University Medical Center, Rotterdam, The Netherlands). Surgical resection specimens of 196 advanced and 44 early EAC were examined. OLFM4 expression was studied by immunohistochemistry and categorized as low (<30%) or high (>= 30%) expression.

Results: Low OLFM4 was associated with poor differentiation grade in both advanced (60% vs. 34.8%, p= 0.001) and early EAC (39.1% vs. 9.5%, p= 0.023). LNM were present in 161 (82.1%) of advanced and 9 (20.5%) of early EAC respectively. Low OLFM4 was independently associated with the presence of LNM in advanced EAC in multivariable analysis (OR 2.7; 95% CI, 1.16-6.41; p= 0.022), but not in early EAC (OR 2.1; 95% CI, 0.46-9.84; p= 0.338). However, the difference in association with LNM between advanced (OR 2.7; 95% CI, 1.18-6.34; p= 0.019) and early (OR 2.3; 95% CI, 0.47-11.13; p= 0.302) EAC was non-significant (p= 0.844), suggesting that the lack of significance in early EAC is due to the small number of patients in this group. OLFM4 was not of significance for the disease free and overall survival.

Conclusions: Overall, low expression of intestinal stem cell marker OLFM4 was associated with the presence of LNM. Our study suggests that OLFM4 could be an informative marker with the potential to improve preoperative assessment in patients with EAC. Further studies are needed to confirm the value of OLFM4 as a biomarker for LNM.

INTRODUCTION

Esophageal cancer is a common cancer with high incidence and mortality rate, with an estimated 456 000 new cases and 400 000 deaths worldwide in 2012,[125] mostly due to diagnosis at advanced incurable stages with limited treatment options. Different histologic types exist, with esophageal squamous cell carcinoma (ESCC) and adenocarcinoma (EAC) being the most frequently encountered types. While ESCC incidences decline, EAC has been one of the fastest rising malignancies in Western countries.[126, 127] Highest incidence rates per 100.0000 person-years for EAC have been observed in the UK (7.2 in men, 2.5 in women) and the Netherlands (7.1 in men and 2.8 in women).[127]

Metastases to the regional lymph nodes is the most important prognostic factor in EAC patients undergoing treatment with curative intent.[128-130] Accurate pretreatment assessment of nodal status is thus important for both advanced and early lesions. In early EAC, patients eligible for endoscopic treatment only (i.e. not followed by surgical resection) should have a minor risk of LNM, because of the inability to perform a lymphadenectomy during endoscopic resection. However, despite all currently available clinical diagnostic modalities (especially EUS, CT and PET) clinical assessment of nodal status is still suboptimal. [29-31] Therefore, a more reliable tool is urgently needed in both advanced and early EAC.

Olfactomedin 4 (OLFM4, formerly known as hGC-1 or GW112) might be an interesting candidate biomarker in this context. It is a secreted glycoprotein, originally identified as a glycoprotein expressed in the olfactory neuroepithelium of bullfrogs. [131] OLFM4 was first cloned from human myeloblasts and is mainly expressed in the gastro-intestinal tract (stomach, small intestine and colon), prostate and bone marrow.[132] In human colon crypts, OLFM4 co-localizes with LGR5+ intestinal stem cells.[133] OLFM4 positive cells are also found in gastric intestinal metaplasia and Barrett's esophagus (BE), where it is confined to the base of metaplastic glands, in a similar way as in colon crypts, with gradually increased expression during dysplastic progression.[66] OLFM4 is regulated by G-CSF,[132] the transcription factor NF-kappaB,[134, 135] and the Wnt/β-catenin pathway[136] and can mediate cell adhesion through its interactions with extracellular matrix proteins such as cadherins and lectins.[137]

In gastric cancer, low OLFM4 expression is correlated with poor differentiation grade and the presence of LNM, as well as with adverse survival.[68, 72, 138, 139] Similarly, decreasing frequencies of expression along with cancer progression have been found in breast, endometrial, prostate and colon carcinoma amongst others.[67, 69, 71, 140] Because no data on OLFM4 in EAC are available yet, this study was undertaken to investigate the association between OLFM4 and presence of LNM and prognosis in both advanced and early EAC. We hypothesized low OLFM4 expression in EAC is associated with the presence of LNM and could be a potential biomarker for stratification of patient treatment.

METHODS

Patient selection & study design

Patients who underwent esophagectomy with curative intent for pathologically confirmed pT2-pT4 adenocarcinoma of the esophagus or gastro-esophageal junction between 1995 and 2016 in the Erasmus MC University Medical Center, Rotterdam were selected for this study. Patients were identified from a prospectively collected institutional database. To assure accurate pathological LNM status, patients treated with surgical resection and at least 12 lymph nodes in the resection specimen were included. Patients with concurrent cancer(s) in other organs and/or those dying from surgical complications (survival < 1 month) were excluded as well as patients that received (neo-) adjuvant chemoradiation therapy ((n)CRT). In addition, all patients with early (pT1b) EAC, treated between 1992 – 2014 at the Erasmus MC, were investigated. These patients were treated by either primary esophagectomy or endoscopic resection followed by esophagectomy because of poor prognostic criteria found in the endoscopic resection specimen. To increase patient numbers in the early EAC group, patients with early EAC and less than 12 pathologically examined lymph nodes, but available follow-up for more than 60 months were also included.

Clinical and pathological data had been prospectively collected, including age at surgery, sex, tumor location and size, surgical technique, resection margin status, differentiation grade, presence (pN-/ pN+) and number (pN0-3) of pathologically confirmed lymph node metastasis and disease free survival (DFS) and overall survival (OS). Resection margin positivity was defined as presence of tumor cells in the (inked) resection margin (definitions according to the College of American Pathologists (CAP)).[141] Recurrence was defined as either locoregional or distant during follow-up, which was either a clinical diagnosis and sometimes pathologically confirmed. DFS was defined as time between the date of surgery and first occurrence of disease progression. OS was defined as time between surgery and death. Patients lost to follow-up were censored at the time of the last visit to the outpatient clinics. The TNM system according to the UICC seventh edition was used for pathological grading and staging.[128] However, corresponding to the eighth edition, which shows no changes in the definitions of the T, N, and M categories, only carcinomas with their epicenter within the proximal 2 cm of the cardia (Siewert types I and II tumors) were included.[34]

Specimen characteristics

The hematoxylin-eosin stained slides and tissue blocks were retrieved from the archives of the Department of Pathology at the Erasmus MC University Medical Center and re-assessed for tumor staging, grading and additional immunohistochemical staining (IHC) for OLFM4. From the most representative slide with deepest tumor invasion, the FFPE block was selected and 4 µm thick sections were cut from this block. OLFM4 (clone DIE4M, Cell Signalling ref. 14369) staining was performed using an automated immunostainer (BenchMark Ultra,

Ventana Medical Systems, Roche, Tuscon, AZ, USA). In brief, deparaffinization according to BenchMark Ultra protocol and antigen retrieval by CC1 antigen retrieval solution (64 min, ref. 950-124, Ventana Medical Systems) were performed. Tissues were incubated with the primary antibody OLFM4 (32 min, dilution 1:400). Detection was performed with UltraView-DAB (ref. 760-500, Ventana Medical Systems) and amplification with Amplification Kit (ref. 760-080 Ventana Medical Systems). Next, the slides were counterstained with hematoxylin (ref: 790-2208, Ventana Medical Systems) and coverslipped. Each slide contained normal colon tissue as a positive control. Furthermore, normal tissue surrounding the tumor was evaluated for its physiological expression of OLFM4 and to assess background staining. OLFM4 expression was scored based on the percentage of tumor cells with cytoplasmic OLFM4 staining. In addition, the H-score based on predominant staining intensity (no / weak/ moderate/ strong staining) was initially scored in a discovery set (n=57). When present in the same slide (adjacent to tumor) OLFM4 expression was also evaluated in non-dysplastic Barrett's esophagus (NDBE). Barrett's esophagus was defined by metaplasia of the pre-existent squamous epithelium into columnar epithelium containing goblet cells.[1] All OLFM4 stained slides were reviewed independently by two investigators (LS and FK), blinded to the clinical and pathological outcome. In case of disagreement, a consensus was reached by review by both investigators. Specifically, 126 out of 240 cases showed a relatively small difference (1-10%) in scoring, of which the numbers were averaged. In 32 cases a difference of more than 10% was found, and a consensus was reached in a consensus meeting.

Statistical analysis

The optimal cut-off value of OLFM4 expression was based on receiver operating characteristic (ROC) curve analysis in advanced EAC, and corresponding Youden index (Supplementary Figure 1). Based on this evaluation, low OLFM4 was defined as < 30% expression, otherwise OLFM4 was considered to be high. The interobserver variation for the assessment of OLFM4 staining between the two observers was calculated using the intraclass correlation coefficient. Strength of agreement was categorized as follows: 0.00–0.20, poor; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, good; and 0.81–1.00, excellent.

Required sample size was not calculated a priori as no pilot data on OLFM4 in EAC was available to determine an expected effect size and it was also predetermined by study constraints. Differences between the advanced and early EAC cohorts were analyzed using Student's t test for normal distributions and the Mann–Whitney U test for non-normal distributions of continuous variables, and Pearson's chi-squared ($\chi 2$) test for categorical variables. Normality of distributions were assessed using the Shapiro-Wilk Test of Normality and by looking at the histogram plot. Correlations between clinicopathological variables and OLFM4 expression were analyzed using $\chi 2$- test or Fisher's exact test. Multivariable logistic regression was used to calculate independent associated factors for LNM in the

resection specimen (pN+). Only variables that were statistically significant in univariable analysis were included in multivariable analysis. To investigate whether the association of OLFM4 was different in advanced and early EAC we performed a logistic regression analysis containing all relevant confounders, OLFM4 status, early or advanced EAC and the interaction between OLFM4 status and early or advanced EAC.

Kaplan Meier curves were used to plot the 5-year DFS and OS by OLFM4 status and the distribution was analyzed using the log-rank test. Uni- and multivariable Cox proportional hazard models were applied to calculate the association between OLFM4 and survival. In multivariable analysis all clinical and pathological factors which proved to be prognostic for survival in univariable analysis were included ($p<0.05$). Statistical analysis was performed using SPSS-software (version 22, SPSS IBM inc, Armonk, NY, USA). A p-value of <0.05 (two-sided) was considered statistically significant. This study was reported according to the Reporting recommendations for tumor marker prognostic studies (REMARK, Supplementary Table 1).[142]

Ethical approval

This study was approved by the institutional review board (medical ethical committee) from the Erasmus Medical Center (Rotterdam, The Netherlands).

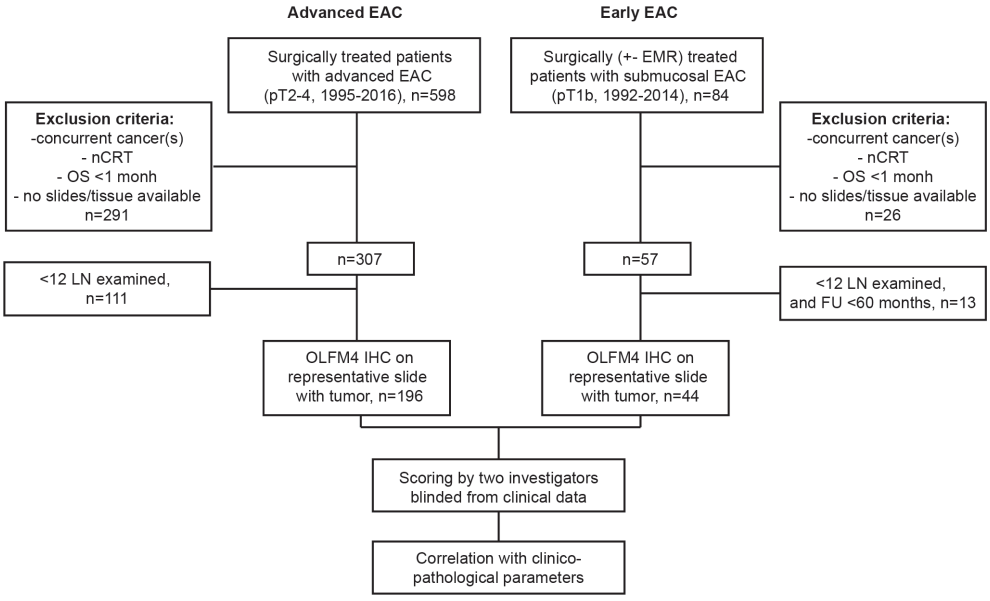


Figure 1. Flow diagram depicting the flow of patients throughout the study. *EAC, esophageal adenocarcinoma; nCRT, neo-adjuvant chemo-radiation therapy; OS, overall survival; FU, follow-up, LN, lymph nodes; OLFM4, Olfactomedin 4; IHC, immunohistochemistry*

RESULTS

Patient characteristics

A diagram depicting the flow of patients throughout the study is shown in Figure 1. Out of 240 EAC patients investigated in this study, 196 had advanced EAC (pT2-4) and 44 early EAC (pT1b). Clinicopathological characteristics are listed in Table 1.

Pattern of OLFM4 expression

In total, 240 EAC resection specimens were assessed for OLFM4 expression. The interobserver agreement for OLFM4 assessment was "good" to "excellent" between the two observers with an intraclass correlation co-efficient of 0.871 (95% CI, 0.782–0.918). However, the H-score resulted in a poor interobserver agreement (Cohen's kappa 0.2) and was disregarded from further analysis. In normal esophageal tissue (without presence of Barrett's esophagus), OLFM4 expression was absent (Supplementary Figure 2). In total, 87 (36.2%) EACs showed high OLFM4 expression and 153 (63.8%) EACs showed low OLFM4 expression (Figure 2). Mostly, expression of OLFM4 was homogeneous, but occasionally, heterogeneous OLFM4 expression was observed, with predominantly high OLFM4 expression towards the lumen and absence of OLFM4 expression towards the invasive front (Supplementary Figure 3). Non-dysplastic Barrett's esophagus (NDBE) showed a similar staining pattern as normal human colon, with cytoplasmic OLFM4 expression in the crypt basis (Figure 2). As in NDBE, OLFM4 expression was noted in the cytoplasm of the EAC cells (Figure 2C).

OLFM4 expression and clinico-pathological characteristics in advanced and early EAC

In advanced EAC, 78 out of 130 (60%) cases with low OLFM4 expression were poorly differentiated, compared to 23 out of 66 (34.8%) EAC with high expression (p= <0.001, Table 2). A similar association between differentiation grade and OLFM4 expression was found in early EAC (9/23 (39%) vs. 2/21 (10%), p= 0.023). Low OLFM4 expression was also associated with presence of pathologically confirmed LNM at the time of resection in EAC (119/153 (78%) vs 51/87 (59%), p= 0.002). In advanced EAC OLFM4 was associated with LNM (113/130 (87%) vs 48/66 (73%), p= 0.014), but not in early EAC (6/23 (26%) vs 3/21 (14%), p= 0.332).

To identify the odds ratio (OR) of clinicopathological characteristics for presence of LNM in EAC, uni- and multivariable logistic regression analysis were performed (Table 3). In multivariable analysis, positive resection margin (OR 7.8, 95% CI, 1.70-35.68, p= 0.008), higher pT-stage (pT34, OR 4.0; 95% CI, 1.53-10.29, p= 0.005) and low OLFM4 expression (OR 2.7; 95% CI, 1.16-6.41; p= 0.022) were identified as independent prognostic variables for LNM in advanced EAC. In contrast, no independently prognostic variables were found in early EAC. However, in the combined cohort the interaction test showed no significant

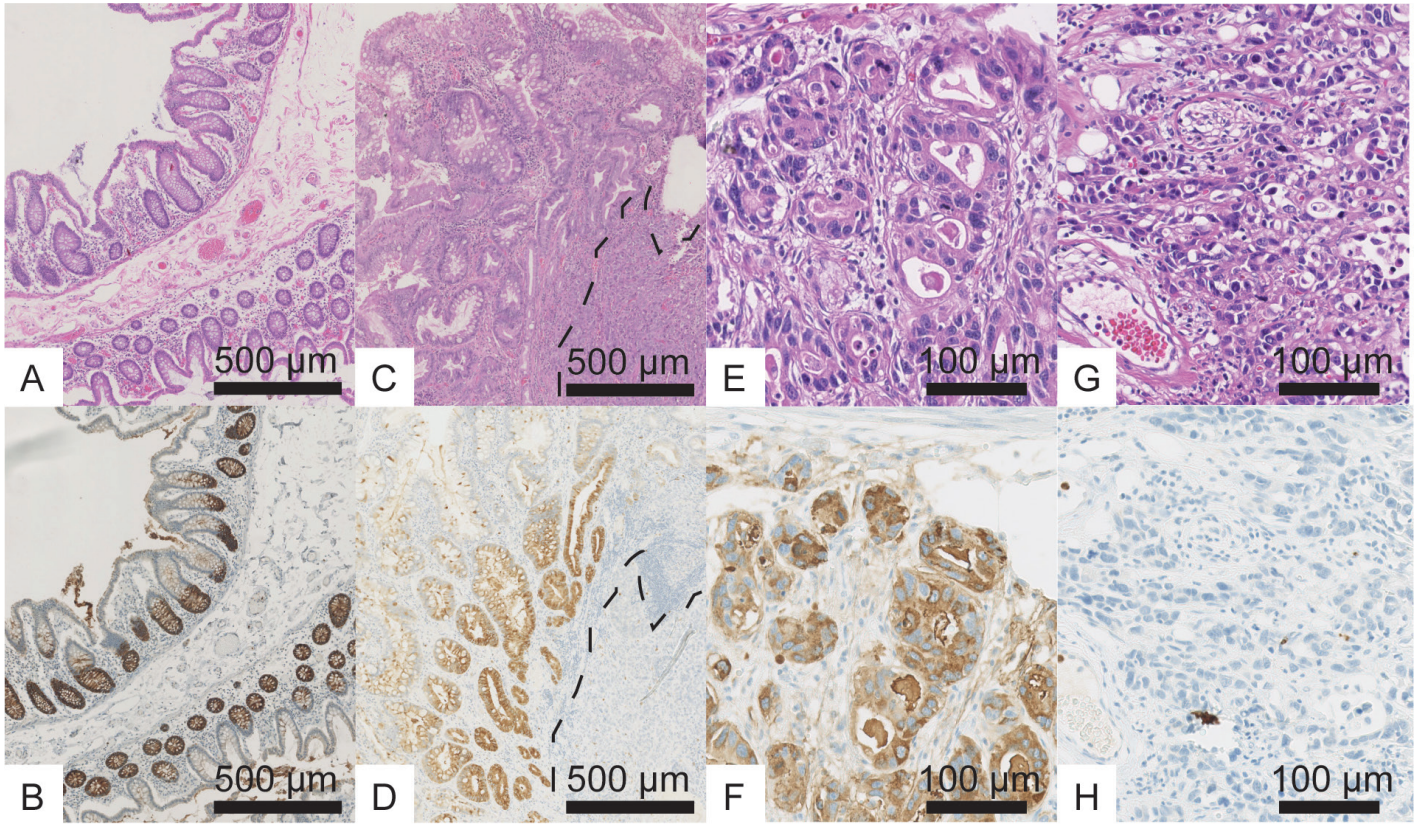


Figure 2. Examples of OLFM4 expression. OLFM4 expression in A, B) normal human colon tissue and C, D) non-dysplastic Barrett's esophagus overlying OLFM4 negative tumor cells (divided by dotted line). Representative cases of esophageal adenocarcinoma with E, F) high and G, H) low OLFM4 expression (A, C, E, G: hematoxylin- eosin; B, D, F, H: OLFM4).

Table 1. Patient characteristics.*

		All patients (Advanced + Early)	Advanced EAC (pT2-4)	Early EAC (pT1b)	Advanced vs Early
		n=240	n=196	n=44	p-value*
Age, years (mean [SD])		63 (10)	63 (10)	62 (9)	0.445§
Sex (n[%])	Male	199 (82.9)	165 (84.2)	34 (77.3)	0.271
	Female	41 (17.1)	31 (15.8)	10 (22.7)	
Surgery (n[%])	Transhiatal	150 (62.5)	120 (61.2)	30 (68.2)	**<0.001**
	Transthoracal	69 (28.8)	67 (34.2)	2 (4.5)	
	Total/ Partial Gastric	9 (3.7)	8 (4.1)	1 (2.3)	
	Unknown	12 (5.0)	1 (0.5)	11 (25)	
Siewert classification,¥ (n[%])	Type 1	114 (47.5)	80 (40.8)	34 (77.3)	**<0.001**
	Type 2	125 (52.1)	116 (59.2)	9 (20.5)	
Tumor size, mm▼ (mean [SD])		46.6 (24.2)	50.6 (23.8)	27.4 (15.8)	**<0.001§**
Radicality (n[%])	R0	179 (74.6)	135 (68.9)	44 (100)	**<0.001**
	R1	61 (25.4)	61 (31.1)	0 (0)	
Grade (n[%])	Well / Moderate	128 (53.3)	95 (48.5)	33 (75)	**0.001**
	Poor	112 (46.7)	101 (51.5)	11 (25)	
pT (n[%])	pT1b	44 (18.3)	0 (0)	44 (100)	**<0.001**
	pT2	25 (10.4)	25 (12.8)	0 (0)	
	pT3	168 (70.0)	168 (85.7)	0 (0)	
	pT4	3 (1.3)	3 (1.5)	0 (0)	
pN (n[%])	pN0	70 (29.2)	35 (17.9)	35 (79.5)	**<0.001**
	pN1	49 (20.4)	44 (22.4)	5 (11.4)	
	pN2	54 (22.5)	51 (26.0)	3 (6.8)	
	pN3	67 (27.9)	66 (33.7)	1 (2.3)	
pN- / pN+ (n[%])	pN-	70 (29.2)	35 (17.9)	35 (79.5)	**<0.001**
	pN+	170 (70.8)	161(82.1)	9 (20.5)	
Total LN (median [IQR])		18 (14-26)	19 (15-27)	14 (8-17)	**<0.001°**
LNM (median [IQR]		3 (0-7)	4 (1-8)	0 (0-0)	**<0.001°**
Recurrence,¶ n[%])	No	98 (40.8)	66 (33.7)	32 (72.7)	**<0.001**
	Yes	140 (58.3)	130 (66.3)	10 (22.7)	
Locoregional recurrence,¥ (n[%])					
	No	163 (67.9)	123 (62.8)	40 (90.9)	**<0.001**
	Yes	76 (31.7)	73 (37.2)	3 (6.8)	
Distant recurrence,¶ (n[%])	No	126 (52.5)	94 (48.0)	32 (72.7)	**<0.001**
	Yes	112 (46.7)	102 (52.0)	10 (22.7)	

Continue

Continued

		All patients (Advanced + Early)	Advanced EAC (pT2-4)	Early EAC (pT1b)	Advanced vs Early
		n=240	n=196	n=44	p-value*
pN+ and/ or recurrence,¥ (n[%])					
	No	49 (20.4)	22 (11.2)	27 (61.4)	**<0.001**
	Yes	190 (79.2)	174 (88.8)	16 (36.4)	
60 months survival (n[%])	Alive	79 (32.9)	46 (23.5)	33 (75.0)	**<0.001**
	Deceased	161 (67.1)	150 (76.5)	11 (25.0)	
Follow-up time, months (median [IQR]		25 (9-64)	19 (8-48)	38 (47-80)	**<0.001°**
DFS, months (median [IQR])		17 (7-60)	13 (6-35)	63 (32-77)	**<0.001°**
OS, months (median [IQR])		25 (9-64)	19 (8-48)	64 (47-80)	**<0.001°**
OLFM4 expression, ∞ (n[%])	Low	153 (63.8)	130 (66.3)	23 (59.0)	0.080
	High	87 (36.2)	66 (33.7)	21 (41.0)	

*P-values were based on Pearson's chi-squared test, unless indicated otherwise. All statistical tests were two-sided. *SD, standard deviation; R1, positive; R0, negative resection margins; IQR, interquartile range.* ***§*** *P-values were based on Student's t-test.* ***°*** *P-values were based on Mann-Whitney test. ¥ One sample (early EAC) had unknown data. ▼Eight samples (4 advanced, 4 early EAC) had unknown data. ¶ Two samples (early EAC) had unknown data. ∞ Low OLFM4 < 30% and high OLFM4 ≥30% immunohistochemical expression.*

difference in strength of the association of OLFM4 with LNM in advanced (OR 2.7.; 95% CI, 1.18-6.34; p= 0.019) and early (OR 2.3; 95% CI, 0.47-11.13; p= 0.302) EAC (p= 0.844, Table 4) . In other words, this test shows that there is no reason to assume that the association between OLFM4 and presence of LNM is different between both groups.

OLFM4 expression and prognosis

DFS was significantly better in patients with high OLFM4 expression (for advanced and early EAC cohorts combined, log-rank test, p= 0.024). This was confirmed by univariable COX regression analysis (HR 1.5; 95% CI, 1.05-2.15, p= 0.027). However, this observation did not hold in multivariable analysis (Supplementary Table 2). There was no significant difference in OS between EAC with low vs. high OLFM4 expression. Hence, OLFM4 expression was not prognostic for OS (Supplementary Table 3). Kaplan-Meier curves for both DFS and OS according to OLFM4 expression are depicted in Supplementary Figure 4.

Table 2. Distribution of OLFM4 expression according to clinicopathological characteristics in advanced and early EAC.

		All patients (Advanced + Early), n=240			Advanced EAC (pT2-4), n=196			Early EAC (pT1b), n=44		
		Low, n (%)	High, n (%)	p-value*	Low, n (%)	High, n (%)	p-value*	Low, n (%)	High, n (%)	p-value*
Age	<65	81 (52.9)	46 (52.9)	0.992	68 (52.3)	34 (51.5)	0.916	13 (56.5)	12 (57.1)	0.967
	>=65	72 (47.1)	41 (47.1)		62 (47.7)	32 (48.5)		10 (43.5)	9 (42.9)	
Sex	Male	129 (84)	70 (80.5)	0.446	111 (85.4)	54 (81.8)	0.518	18 (78.3)	16 (76.2)	0.870
	Female	24 (15.7)	17 (19.5)		19 (14.6)	12 (18.2)		5 (21.7)	5 (23.8)	
Surgery	Transhiatal	93 (60.8)	57 (65.5)	0.467	76 (58.5)	44 (66.7)	0.265	17 (73.9)	13 (61.9)	0.393
	Other	60 (39.2)	30 (34.5)		54 (41.5)	22 (33.3)		6 (26.1)	8 (38.1)	
Siewert Classification, ¥,	Type 1	75 (49.3)	39 (44.8)	0.501	57 (43.8)	23 (34.8)	0.226	18 (81.8)	16 (76.2)	0.650
	Type 2	77 (50.7)	48 (55.2)		73 (56.2)	43 (65.2)		4 (18.2)	5 (23.8)	
Tumor Size, ▼	<5 cm	85 (57.4)	44 (52.4)	0.457	66 (52.0)	28 (43.1)	0.243	19 (90.5)	16 (84.2)	0.550
	>= 5 cm	63 (42.6)	40 (47.6)		61 (48.0)	37 (56.9)		2 (9.5)	3 (15.8)	
Radicality	R0	115 (75)	64 (73.6)	0.784	92 (70.8)	43 (65.2)	0.422	23 (100)	21 (100)	NA
	R1	38 (24.8)	23 (26.4)		38 (29.2)	23 (34.8)		0 (0)	0 (0)	
Grade	Well/ Moderate	66 (43.1)	62 (71.3)	**<0.001**	52 (40.0)	43 (65.2)	**0.001**	14 (60.9)	19 (90.5)	**0.023**
	Poor	87 (56.9)	25 (28.7)		78 (60.0)	23 (34.8)		9 (39.1)	2 (9.5)	
pT	pT12	38 (24.8)	31 (35.6)	0.076	15 (11.5)	10 (15.2)	0.474	23 (100)	21 (100)	NA
	pT34	115 (75)	56 (64.4)		115 (88.5)	56 (84.8)		0 (0)	0 (0)	
pN	pN0	34 (22.2)	36 (41.4)	**0.008**	17 (13.1)	18 (27.3)	**0.040**	17 (73.9)	18 (85.7)	0.380
	pN1	33 (21.6)	16 (18.4)		29 (22.3)	15 (22.7)		4 (17.4)	1 (4.8)	
	pN2	42 (27.5)	12 (13.8)		40 (30.8)	11 (16.7)		2 (8.7)	1 (4.8)	
	pN3	44 (28.8)	23 (26.4)		44 (33.8)	22 (33.3)		0 (0)	1 (4.8)	
pN- / pN+	pN-	34 (22.2)	36 (41.4)	**0.002**	17 (13.1)	18 (27.3)	**0.014**	17 (73.9)	18 (85.7)	0.332
	pN+	119 (78)	51 (58.6)		113 (86.9)	48 (72.7)		6 (26.1)	3 (14.3)	
Recurrence (locoregional or distant), ¶	No	54 (35.8)	44 (50.6)	**0.025**	39 (30.0)	27 (40.9)	0.127	15 (71.4)	17 (81.0)	0.469
	Yes	97 (64.2)	43 (49.4)		91 (70.0)	39 (59.1)		6 (28.6)	4 (19.0)	
pN+ and/ or Recurrence, ¥	No	22 (14.5)	27 (31.0)	**0.002**	10 (7.7)	12 (18.2)	**0.028**	12 (54.5)	15 (71.4)	0.252
	Yes	130 (85.5)	60 (69.0)		120 (92.3)	54 (81.8)		10 (45.5)	6 (28.6)	

*Pearson's chi-squared test. *NA, not applicable, because all patients with early EAC had negative resection margins (R0) and were per definition staged pT1.* ¥ One sample (early EAC) had unknown data. ▼*Eight samples (4 advanced, 4 early EAC) had unknown data ¶ Two samples (early EAC) had unknown data.*

Table 3. Logistic regression analysis to evaluate the independent association of OLFM4 with LNM (pN+).*

	All patients (Advanced + Early EAC, n=240)				Advanced EAC (pT2-4, n=196)				Early EAC (pT1b, n=44)	
	Univariable		Multivariable		Univariable		Multivariable		Univariable	
	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value
Age (ref. = <65 y)										
>=65	1.1 (0.62-1.89)	0.785			1.0 (0.47-2.02)	0.936			1.1 (0.24-4.66)	0.932
Sex (ref. = Female)										
Male	1.3 (0.64-2.71)	0.442			1.1 (0.42-3.00)	0.813			1.0 (0.18-6.02)	0.968
Surgery (ref. = Transhiatal)										
Other	1.0 (0.58-1.82)	0.942			0.9 (0.45-1.98)	0.870			0.5 (0.10-3.06)	0.492
Siewert Classification, ¥ (ref. = Type 1)										
Type 2	1.9 (1.10-3.44)	**0.022**	0.9 (0.44-2.01)	0.867	1.0 (0.46-2.02)	0.914			1.1 (0.19-6.52)	0.915
Tumor Size, ▼ (ref. = <5 cm)										
>= 5 cm	2.5 (1.33-4.53)	**0.004**	1.9 (0.90-4.00)	0.095	1.2 (0.58-2.55)	0.609			2.7 (0.37-19.15)	0.330
Radicality (ref. = R0)										
R1	18 (4.28-76.37)	**<0.001**	8.3 (1.83-37.76)	**0.006**	9.5 (2.21-41.21)	**0.003**	7.8 (1.70-35.68)	**0.008**	NA	
Grade (ref. = Well/ Moderate)										
Poor	2.7 (1.49-4.87)	**0.001**	1.1 (0.50-2.37)	0.841	2.4 (1.10-5.09)	**0.027**	1.2 (0.48-2.78)	0.751	0.8 (0.14-4.73)	0.829
pT (ref. = pT12)										
pT34	14 (7.03-26.87)	**<0.001**	7.2 (3.27-15.8)	**<0.001**	5.9 (2.42-14.6)	**<0.001**	4.0 (1.53-10.29)	**0.005**	NA	
OLFM4 expression (ref. = High)										
Low	2.5 (1.39-4.38)	**0.002**	2.6 (1.22-5.62)	**0.013**	2.5 (1.19-5.25)	**0.016**	2.7 (1.16-6.41)	**0.022**	2.1 (0.46-9.84)	0.338

* Uni- and multivariable logistic regression analysis was performed to investigate the independent association between LNM and clinicopathological characteristics, only variables significant in univariable analysis were included in multivariable analysis. Hence, no multivariable analysis for early esophageal adenocarcinoma (EAC) was performed. *OR, odds ratio; CI, confidence interval; ref., reference; NA= not applicable, because all patients with early EAC had negative resection margins (R0) and were per definition staged pT1.*

Table 4. Multivariable logistic regression analysis to evaluate the association of OLFM4 with LNM (pN+) in all patients (left) with corresponding interaction model (right).*

	OR	p-value	OR	p-value
Siewert Classification, ¥ (ref. = Type 1)				
Type 2	1.0 (0.48-2.11)	0.992	1.0 (0.48-2.11)	0.992
Tumor Size, ▼ (ref. = <5 cm)				
>= 5 cm	1.6 (0.76-3.37)	0.217	1.6 (0.76-3.37)	0.217
Radicality (ref. = R0)				
R1	10.4 (2.32-46.73)	**0.002**	10.5 (2.33-47.20)	**0.002**
Grade (ref. = Well/ Moderate)				
Poor	1.1 (0.52-2.44)	0.762	1.1 (0.52-2.44)	0.767
pT (ref. = pT1)				
pT234	8.2 (3.17-21.12)	**<0.0001**	7.3 (1.75-30.77)	**0.007**
OLFM4 IHC (ref. = high expression)				
Early EAC (pT1), n=44			2.3 (0.47-11.13)	0.302
Advanced EAC (pT234), n=196			2.7 (1.18-6.34)	**0.019**
All patients (pT1234), n=240	2.6 (1.24-5.62)	**0.012**		
Interaction, § (Early vs. Advanced)			1.2 (0.21-6.91)	0.844

* Only variables significant in univariable analysis were included in multivariable analysis. *OR, odds ratio; ref., reference; R1, positive; R0, negative resection margins*
¥ One sample (early EAC) had unknown data.
▼ Eight samples (4 advanced, 4 early EAC) had unknown data.
§ The interaction variable indicates whether there is a difference in association of OLFM4 with LNM between early and advanced EAC. The model with separate effects of OLFM4 for early and advanced EAC did not give a better fit than the model with one effect only.

DISCUSSION

This is the first extensive study on OLFM4, an intestinal stem cell marker, in EAC and shows low OLFM4 expression is associated with positive LNM status.

Accurate pretreatment staging of patients with early and advanced EAC is important for optimal treatment selection and survival prediction.[128-130] Previous studies have shown that pretreatment staging is frequently inaccurate in EAC.[31, 143-146] In a recent publication on nCRT-naïve patients with standard pre-operative assessment only 35% of patients were preoperatively diagnosed with a correct T- and N-stage.[147] Particularly in patients with early (pT1) EAC, prevalence of LNM is highly variable and to date unpredictable, while positive LNM status is highly predictive for a poor 5-year survival.[107, 143]

The significance of OLFM4 in cancer is still controversial. OLFM4 is able to interact with cell surface proteins and known to facilitate cell-cell adhesion.[137, 148] OLFM4 has been attributed oncogenic properties as it was shown to promote tumor growth by acting as an anti-apoptotic protein and by increasing the mitotic activity of cancer cells.[149, 150] On the other hand, reduced OLFM4 expression was significantly associated with poor prognosis in patients with gastric,[139] colorectal[71, 151] and breast carcinoma[140] amongst others. In gastric carcinoma OLFM4 was also associated with metastasis.[72, 139]

The present study shows low OLFM4 expression was associated with poorly differentiated EAC, this is in concordance with the literature. In fact, in most cancers, a strong association between low OLFM4 and poor tumor differentiation grade was found, including gastric, colon, ovarian and prostate cancer[68, 69, 72, 137] These findings suggest tumor suppressive properties of OLFM4 and are in line with results found in various functional studies.[136, 152, 153] For example, in gastric cancer cell lines OLFM4 had an inhibitory effect on cell invasion via regulation of focal adhesion kinase (FAK) signaling.[153]

Furthermore, low OLFM4 expression, but not poor tumor differentiation, was independently associated with LNM in advanced EAC in the present study. Because LNM status is critical for the choice of treatment in early EAC, the investigation was extended to early EAC and 44 patients with pT1b tumors were separately analyzed. The overall incidence of LNM in the pT1b group (20.5%) was in line with previous reports.[143] Similar as in advanced EAC, loss of OLFM4 was associated with poor differentiation grade, but no association with LNM status was found. However, the interaction test in the combined cohort showed no significant difference in strength of the association of OLFM4 and LNM between the advanced and early EAC. Therefore, the result in the early EAC might be explained by the small sample size and overall low LNM incidence in this group of patients. Only one previous study studied

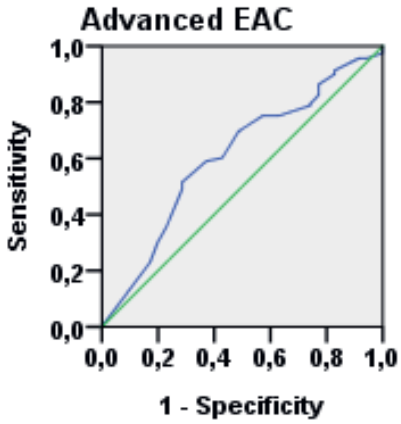
the role of OLFM4 in early cancer (pT1a and pT1b gastric cancer, n=105) and concluded that low OLFM4 expression was independently predictive for LNM.[72]

Despite the association with LNM, in contrast with results found in other types of cancer, OLFM4 seems to have no effect on clinical outcome. However, there are some important differences between the present study and the aforementioned previous studies on OLFM4 in other cancers. In the present study, only patients with at least 12 lymph nodes resected and identified were included, in order to reduce the percentage of patients with falsely negative pN0.[117] Although others may have included more cases, these studies were frequently based on patients with various tumor stages and mostly used tissue micro-arrays (TMAs) instead of whole tissue slides. In addition, different methods for scoring OLFM4 IHC were applied making comparison of results difficult.[68, 69, 71, 72, 138, 139] Importantly, TMAs may not accurately demonstrate tumor heterogeneity, which was observed in our study occasionally. In addition, whole tissue slides allow for simultaneous analyses of adjacent non-tumorous tissue and Barrett's esophagus. It would be very interesting to investigate OLFM4 expression in low-grade and high-grade dysplasia. However, in our samples, BE, with or without dysplasia, was present in only a limited number of cases. Therefore, investigation of OLFM4 expression patterns during neoplastic progression would require a separate study design using well defined sample criteria.

There are also some limitations to the present study. Specifically, all patients were from one academic center. Also, patients were treated with surgery alone, while current guidelines recommend nCRT prior to surgery for advanced EAC. However, additional treatment prior to surgery might influence OLFM4 expression and survival, hence it was decided to use a nCRT-naïve patient cohort.

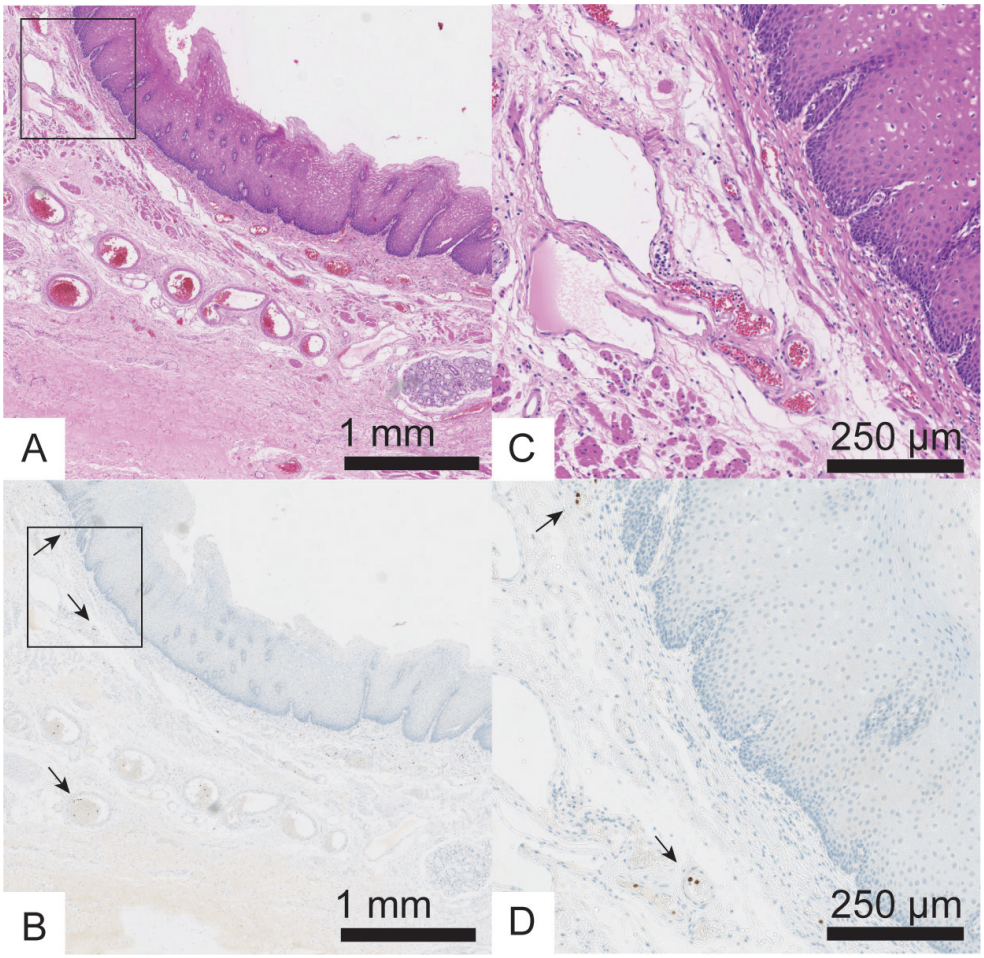
In conclusion, the present study shows that low OLFM4 expression was independently associated with LNM in EAC and hence might prove useful as a new biomarker. Improved prediction of LNM presence could benefit decision making in treatment of EAC patients. This is particularly important in early EAC where overtreatment can be avoided by endoscopic submucosal resection. More research is required to investigate whether OLFM4 is indeed biologically and clinically relevant in both advanced and early EAC.

SUPPLEMENTARY MATERIALS

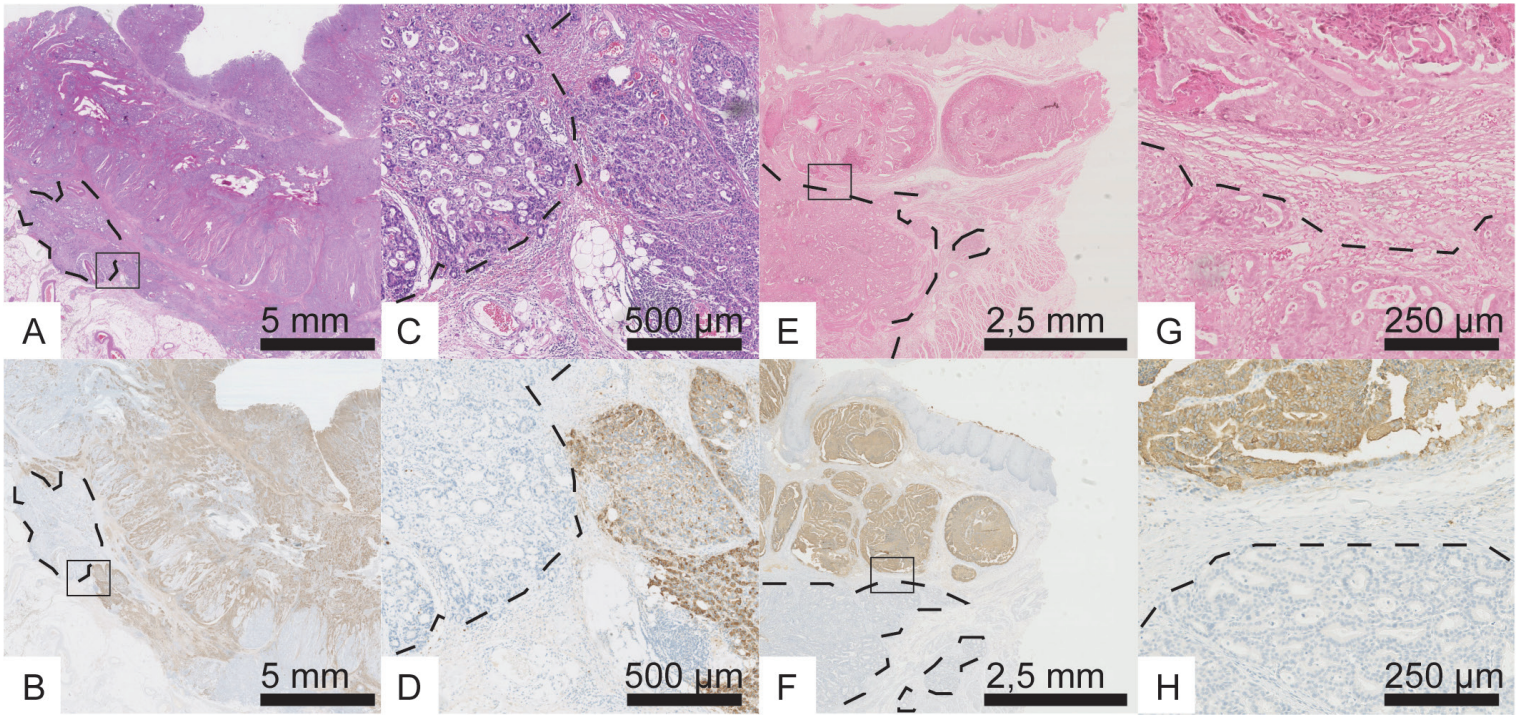


Cut-off (>=, %)	Sensitivity	Specificity	Youden index
10	0,516	0,714	0,23
20	0,602	0,571	0,17
30	**0,696**	**0,514**	**0,21**
40	0,752	0,371	0,12
50	0,795	0,257	0,05
60	0,826	0,229	0,05
70	0,863	0,229	0,09
80	0,913	0,171	0,08
90	0,957	0,057	0,01

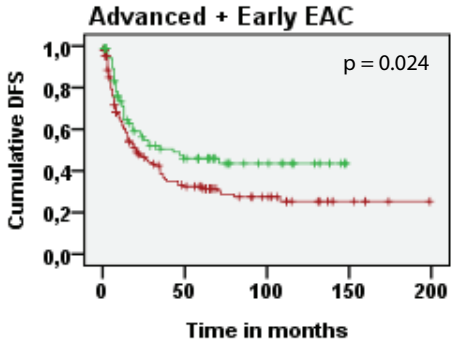
Supplementary Figure 1. Receiver operating characteristics – curve for the semi-quantative OLFM4 expression, according to the % of positive tumor cells (cytoplasm), and corresponding Youden index.

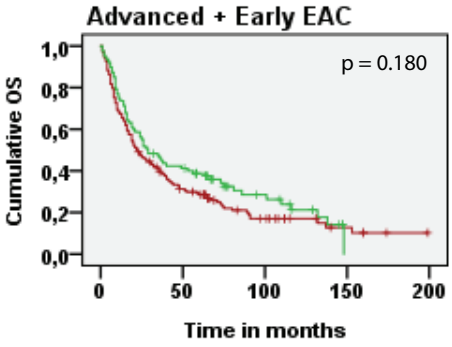


Supplementary Figure 2. OLFM4 expression in normal esophageal tissue. A, B) Normal esophageal tissue is negative for OLFM4. C, D) Magnification of A, B Only neutrophils are OLFM4 positive (brown dots indicated by arrows) and can be used as positive internal control (A, B: hematoxylin- eosin; C, D: OLFM4).



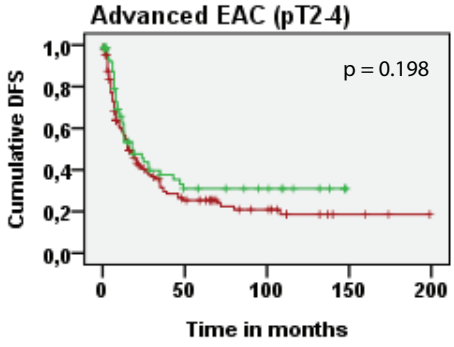
Supplementary Figure 3. Two examples of cases with heterogeneous OLFM4 expression. In A, B) tumor invading into the muscularis propria and adventitia can be seen. While the tumor in the mucosa, submucosa and muscularis propria is positive for OLFM4, two complete OLFM4 negative clones invading the surrounding fatty tissue can be seen (dotted line). C, D) Magnification of A, B. E, F) A well differentiated tumor with several OLFM4 positive tumor foci towards the lumen (squamous epithelium) and complete absence (below dotted line) of OLFM4 expression in tumor foci towards the invasive front. G, H) Magnification of E, F (A, C, E, G: hematoxylin- eosin; B, D, F, H: OLFM4).

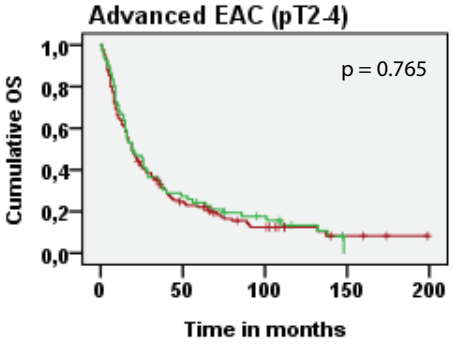




Number of patients at risk					
Low OLFM4	153	38	16	5	0
High OLFM4	87	29	13	0	0

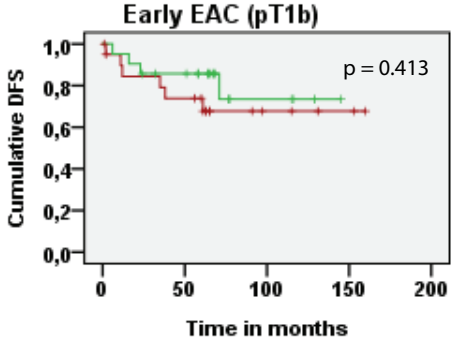
Number of patients at risk					
Low OLFM4	153	44	16	5	0
High OLFM4	87	35	13	0	0

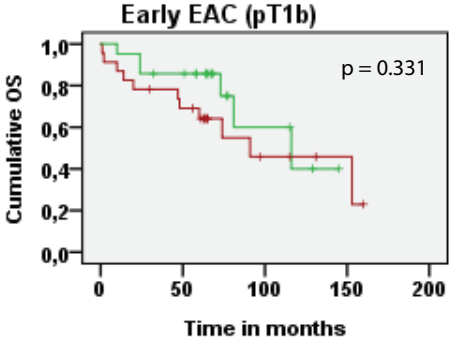




Number of patients at risk					
Low OLFM4	130	24	12	3	0
High OLFM4	66	13	9	0	0

Number of patients at risk					
Low OLFM4	130	29	12	2	0
High OLFM4	66	18	9	0	0





Number of patients at risk					
Low OLFM4	23	14	4	0	0
High OLFM4	21	16	4	0	0

Number of patients at risk					
Low OLFM4	23	15	4	0	0
High OLFM4	21	17	4	0	0

< Supplementary Figure 4. Kaplan-Meier curves for disease free survival (DFS, left) and overall survival (OS, right) according to OLFM4 expression. DFS and OS of both cohorts (upper two), patients with advanced (middle two) and patients with early (lower two) esophageal adenocarcinoma. Overall, DFS is better in patients with tumors with high OLFM4 expression, although this difference is only significant when both cohorts are combined (p=0.024, log-rank test). There is no significant difference in OS between EAC with low vs. high OLFM4 expression (log-rank test).

Supplementary Table 1. The REMARK checklist.[142]

Section/ Topic	#		Reported on page #
INTRODUCTION			
	1	State the marker examined, objectives, hypotheses.	3,4
MATERIAL AND METHODS			
Patients	2	Describe characteristics of study patients, including their source and inclusion and exclusion criteria.	5,6
	3	Describe treatments received and how chosen.	5,6
Specimen characteristics	4	Describe type of biological material used (incl. control samples) and methods of preservation and storage.	6
Assay methods	5	Specify assay methods and provide protocol, incl. specific reagents or kits used, quality control procedures, reproducibility assessments, quantitation methods, scoring and reporting. Specify whether and how assays were performed blinded to the study endpoint.	6
Study design	6	State the method of case selection (prospective/ retrospective, stratification/ matching). Specify the time period from which cases were taken, the end of follow-up period, and the median follow-up time.	5, 6, Fig 1, Table 1
	7	Precisely define all clinical endpoints examined.	5,6
	8	List all candidate variables initially examined or considered for inclusion in models.	5,6
	9	Give rationale for sample size; if the study was designed to detect a specified effect size give the target power and effect size	8
Statistical analysis methods	10	Specify all statistical methods, including details of any variable selection procedures and other model-building issues, how model assumptions were verified, and how missing data were handled.	7,8
	11	Clarify how marker values were handled in the analyses; if relevant, describe methods used for cutpoint determination.	7, S1 Fig
RESULTS			
Data	12	Describe the flow of patients through the study, including the number of patients included in each stage of the analysis (a diagram may be helpful) and reasons for dropout. Specifically, both overall and for each subgroup extensively examined report the number of patients and the number of events.	9, Fig 1, Table 1
	13	Report distributions of basic demographic characteristics (at least age and sex), standard (disease-specific) prognostic variables, and tumor marker, including numbers of missing values.	Table 1
Analysis and presentation	14	Show the relation of the marker to standard prognostic variables.	Table 2
	15	Present univariable analyses showing the relation between the marker and outcome, with the estimated effect (for example, hazard ratio and survival probability). Preferably provide similar analyses for all other variables being analyzed. For the effect of a tumor marker on a time-to event outcome, a Kaplan-Meier plot is recommended.	Table 3, 4 S2 and S3 Tables Fig 4
	16	For key multivariable analyses, report estimated effects (for example, hazard ratio) with confidence intervals for the marker and, at least for the final model, all other variables in the model.	Table 3 and 4 S2 and S3 Tables
	17	Among reported results, provide estimated effects with confidence intervals from an analysis in which the marker and standard prognostic variables are included, regardless of their statistical significance.	11-14, Table 3 and 4 S2 and S3 Tables
	18	If done, report results of further investigations, such as checking assumptions, sensitivity analyses, and internal validation.	-
DISCUSSION			
	19	Interpret the results in the context of the pre-specified hypotheses and other relevant studies; include a discussion of limitations of the study.	14-17
	20	Discuss implications for future research and clinical value.	14-17

Supplementary Table 2. Cox regression analysis to evaluate the risk for recurrence (DFS). Uni- and multivariable Cox regression analysis was performed to investigate the independent association between disease free survival (DFS) and clinicopathological characteristics, only variables significant in univariable analysis were included in multivariable analysis. Hence, no multivariable analysis for early esophageal adenocarcinoma (EAC) was performed.

	All patients (Advanced + Early EAC, n=240)				Advanced EAC (pT2-4, n=196)				Early EAC(pT1b, n=44)	
	Univariable		Multivariable		Univariable		Multivariable		Univariable	
	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
Age (ref. = <65 y)										
>=65 y	1.1 (0.80-1.56)	0.508			1.0 (0.74-1.48)	0.805			2.1 (0.59-7.39)	0.256
Sex (ref. = Female)										
Male	1.1 (0.70-1.69)	0.705			1.0 (0.60-1.51)	0.830			1.3 (0.28-6.30)	0.715
Surgery (ref. = Transhiatal)										
Other	0.9 (0.66-1.31)	0.668			1.0 (0.70-1.42)	0.984			0.0 (0.0-4.76)	0.165
Siewert Classification, ¥ (ref. = Type 1)										
Type 2	1.2 (0.87-1.71)	0.242			0.9 (0.63-1.27)	0.513			0.9 (0.20-4.35)	0.919
Tumor Size, ▼ (ref. = <5 cm)										
>= 5 cm	1.3 (0.90-1.76)	0.181			0.9 (0.65-1.30)	0.630			2.2 (0.47-10.53)	0.316
Radicality (ref. = R0)										
R1	2.6 (1.79-3.68)	**<0.001**	1.3 (0.86-1.88)	0.228	2.0 (1.37-2.85)	**0.001**	1.3 (0.88-1.93)	0.188		NA
Grade (ref. = Well/ Moderate)										
Poor	2.5 (1.77-3.48)	**<0.001**	1.8 (1.28-2.61)	**0.001**	1.9 (1.34-2.70)	**<0.001**	1.7 (1.15-2.39)	**0.006**	5.2 (1.48-17.91)	0.010
pT-stage (ref. = pT12)										
pT34	3.7 (2.34-5.87)	**<0.001**	1.9 (1.10-3.13)	**0.020**	2.3 (1.28-4.22)	**0.006**	1.5 (0.81-2.81)	0.197		NA
pN-stage (ref. = pN-)										
pN+	4.2 (2.61-6.81)	**<0.001**	2.6 (1.51-4.44)	**0.001**	3.2 (1.79-5.70)	**<0.001**	2.5 (1.38-4.64)	**0.003**	2.1 (0.53-8.02)	0.300
OLFM4 expression (ref. = High)										
Low	1.5 (1.05-2.15)	**0.027**	1.1 (0.76-1.57)	0.643	1.3 (0.87-1.85)	0.208			1.7 (0.48-5.98)	0.419

HR, hazard ratio; CI, confidence interval; ref., reference; NA= not applicable, because all patients with early EAC had negative resection margins (R0) and were per definition staged pT1. ¥ One sample (early EAC) had unknown data. ▼Eight samples (4 advanced, 4 early EAC) had unknown data.

Supplementary Table 3. COX regression analysis for OS. Uni- and multivariable Cox regression analysis was performed to investigate the independent association between overall survival (OS) and clinicopathological characteristics, only variables significant in univariable analysis were included in multivariable analysis. Hence, no multivariable analysis for early esophageal adenocarcinoma (EAC) was performed.

	All patients (Advanced + Early EAC, n=240)				Advanced EAC (pT2-4, n=196)				Early EAC (pT1b, n=44)	
	Univariable		Multivariable		Univariable		Multivariable		Univariable	
	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
Age (ref. <65 y)										
>=65 y	1.3 (1.01-1.79)	0.045			1.3 (0.96-1.75)	0.097			1.9 (0.72-5.00)	0.194
Sex (ref. = Female)										
Male	1.3 (0.86-1.87)	0.240			1.1 (0.76-1.71)	0.526			2.3 (0.52- 10.17)	0.270
Surgery (ref. = Transhiatal)										
Other	0.8 (0.58-1.06)	0.119			0.8 (0.58-1.08)	0.141			0.5 (0.15- 1.41)	0.172
Siewert Classification, ¥ (ref. = Type 1)										
Type 2	1.1 (0.81-1.45)	0.599			0.8 (0.59-1.08)	0.143			0.6 (0.13-2.63)	0.490
Tumor Size, ▼ (ref. = <5 cm)										
>= 5 cm	1.4 (1.02-1.82)	**0.038**	1.1 (0.81-1.47)	0.558	1.0 (0.75-1.38)	0.895			1.6 (0.43-5.88)	0.493
Radicality (ref. = R0)										
R1	2.4 (1.72-3.23)	**<0.001**	1.4 (0.98-1.96)	**0.065**	1.8 (1.34-2.55)	**<0.001**	1.4 (1.01-2.01)	**0.042**		NA
Grade (ref. = Well/ Moderate)										
Poor	1.9 (1.40-2.51)	**<0.001**	1.4 (1.05-1.96)	**0.023**	1.4 (1.04-1.92)	**0.025**	1.2 (0.90-1.69)	0.200	4.7 (1.67-13.31)	**0.003**
pT-stage (ref. = pT12)										
pT34	3.1 (2.14-4.45)	**<0.001**	1.7 (1.10-2.57)	**0.016**	2.0 (1.26-3.29)	**0.004**	1.4 (0.87-2.38)	0.158		NA
pN-stage (ref. = pN-)										
pN+	3.3 (2.28-4.78)	**<0.001**	2.1 (1.39-3.27)	**0.001**	2.5 (1.62-3.92)	**<0.001**	2.0 (1.28-3.25)	**0.003**	2.0 (0.68-5.89)	0.209
OLFM4 expression (ref. = High)										
Low	1.2 (0.91-1.66)	0.186			1.0 (0.77-1.44)	0.769			1.6 (0.60-4.53)	0.337

HR, hazard ratio; CI, confidence interval; ref., reference; y, years; NA= not applicable, because all patients with early EAC had negative resection margins (R0) and were per definition staged pT1. ¥ One sample (early EAC) had unknown data. ▼Eight samples (4 advanced, 4 early EAC) had unknown data.

6

Pattern of p53 protein expression is predictive for survival in chemotherapy-naïve esophageal adenocarcinoma

Ten Kate FJC, Suzuki L, Dorssers LCJ, Dinjens WNM, Jones DTW, Nieboer D, Doukas M, van Lanschot JJB, Wijnhoven BPL, Looijenga LHJ, Biermann K.

Oncotarget 2017;8;104123-104135.

Background: TP53 mutations are considered to be the driving factor in the initiation of esophageal adenocarcinoma (EAC). However, the impact of this gene and its encoded protein as a prognostic marker has not been definitely established yet.

Methods: In total, 204 chemoradiotherapy (CRT)-naive patients with EAC were included for p53 protein expression evaluation by immunohistochemistry (IHC) on the resection specimens, categorized as overexpression, heterogeneous or loss of expression, and correlated with disease free survival (DFS) and overall survival (OS) using multivariable Cox regression analysis. In a subset representing all three IHC subgroups mutational status of selected candidate genes (n=33) and high throughput methylation profiling (n=16) was assessed.

Results: Compared to heterogeneous p53 expression, loss and overexpression were both independently predictive for adverse DFS and OS. TP53 mutational status significantly correlated with the IHC categories (p=0.035). Most of the EAC with loss- or overexpression harbored TP53 mutations (18/20, representing nonsense and missense mutations respectively). In contrast, 6/13 EAC with heterogeneous expression were TP53 wild type, of which two demonstrated MDM4 or MDM2 amplification. Combined genomic hypomethylation and high frequency of intrachromosomal breaks was found in a selection of EAC without p53 overexpression.

Conclusion: P53 expression pattern is prognostic for DFS and OS in this historical cohort of CRT-naive EAC. P53 IHC is an informative readout for TP53 mutational status in EAC with either loss- or overexpression, but not in case of a heterogeneous p53 pattern. Different EAC pathogenesis might exist, related to p53 and other candidate gene status, DNA hypomethylation and intrachromosomal breaks.

Abstract

INTRODUCTION

Esophageal adenocarcinoma (EAC), being rare before the second half of the 20th century, is nowadays the predominant histological type of esophageal cancer in Western countries.[154-156] Presently the prognostication of patients with EAC is largely based on the TNM-classification supplemented with histological criteria [157]. Although this system has its value in the stratification of patients into prognostic groups,[158] the outcome for an individual patient is still difficult to predict. This is demonstrated by the fact that up to 27% of the patients with stage IB develop disease recurrence while up to 24% of the patients with stage IIIA EAC will have no disease recurrence after intentionally curative surgery.[158] Therefore, prognostic biomarkers complementing the TNM classification are urgently needed.

The *TP53* gene (OMIM# 191170), first discovered more than 30 years ago, has a cell- and context dependent biological function. It has been reported that p53 is deregulated in most cancer types. Given its central role in the control of proliferation and senescence, it can be assumed to be the driving force of cancers of various types, including EAC.[159-161] Several types of stress can lead to p53 dysregulation. In EAC, mutations in *TP53* are detected early in the pathogenesis, likely linked to severe DNA damage in Barrett esophagus (BE) due to the reflux of mixed gastric and duodenal juice into the esophagus.[162] Recent genome wide studies proposed that EAC precursor lesions containing *TP53* mutations rapidly develop extensive chromosomal instability with subsequent oncogene activation.[11, 163, 164]

Because of its dominant role in the development of EAC, p53 was also tested as a biomarker in EAC precursor lesions and in advanced EAC. There is growing evidence that p53 overexpression is related to dysplasia and independently predictive for progression in BE.[46, 81, 165-168] Overexpression is likely due to *TP53* mutations which stabilize the affected protein. "Absence" of p53 staining was described more recently in dysplastic BE.[169] This loss of expression is likely to be related to truncating mutations or to alternative, including epigenetic, mechanisms. Supporting the significance of the loss of expression, a recent IHC p53 study on a large prospective BE cohort revealed a significantly higher rate of progression to high grade dysplasia or EAC in low grade dysplasia harboring p53 overexpression and even higher in BE with absence of p53 expression.[46]

In parallel to the EAC precursors, the results of the earlier investigations also suggested significance of p53 in relation to prognosis in advanced EAC.[170-172] However, strong conclusions cannot be drawn because of several limitations, including heterogeneity related to p53 IHC interpretation and patient selection. This may have influenced the outcome of these studies and as such the true biological effect of p53 in the context of disease progression may remain unidentified.

Therefore, the aim of this study was to examine the prognostic value of p53 in a well-defined group of chemo- and radio-therapy-naive EAC, using a validated IHC approach. To further investigate the putative mechanism(s) involved, a combinatory investigation of expression pattern, mutational status of *TP53* and a selection of other (relevant) genes, as well as high-throughput profiling was performed in a subset of EAC.

METHODS

Patient selection

To evaluate the prognostic value of p53 in patients with EAC, a cohort of patients who underwent surgery with curative intent between 1995 and 2006, without prior (neo-) adjuvant treatment, was selected from the Department of Surgery at the Erasmus University Medical Center (Rotterdam, The Netherlands). All patients had pathologically proven pT2-pT4a adenocarcinoma of the esophagus or at the gastro-esophageal junction. Only patients who were alive one month after surgery were included in the analysis to correct for surgical mortality. Clinical and pathological data were prospectively collected, including anatomical tumor location according to Siewert,[173] tumor grade, pathological stage, age at surgery, comorbidities, OS and DFS. Tumor grading and staging was performed according to the TNM system as described by the UICC (Union Internationale Contre le Cancer, 2009, 7th edition).[157] Resection margin positivity was assessed on tumor cells in the resection margin. To ensure reliable classification, all slides were reviewed by an experienced GI pathologist (FK or KB) for depth of invasion.

The hematoxylin-eosin colored slides from the resection specimens were retrieved from the archive of the Department of Pathology at the Erasmus University Medical Center and a representative slide with EAC was selected. The corresponding FFPE block was retrieved and serial 4 µm sections for IHC and mutational analysis were mounted on glass slides.

Immunohistochemical analysis

The first slide of each selected FFPE block was stained for p53, ready to use kit (clone BP53-11, Ventana Medical Systems, Roche, Tuscon, AZ, USA). Staining was performed using an automated slide staining system (BenchMark Ultra, Ventana Medical Systems, Roche, Tuscon, AZ, USA), in which the slides were deparaffinized prior to the staining procedure and heat induced epitope retrieval at 97°C for 8 minutes. The primary antibody was incubated for 4 minutes, after which this was visualized using Ultraview (Ventana Medical Systems, Roche, Tuscon, AZ, USA) and counterstained with hematoxylin.

For optimal interpretation, representative tumor samples were evaluated by two experienced gastrointestinal (GI) pathologists (KB and FK) with specific knowledge on

p53, based on earlier published extensive IHC studies on EAC and its precursor lesions.[46, 74] A tumor sample with known overexpression of p53 was placed as positive control on each slide. Furthermore, normal tissue surrounding the tumor cells were evaluated for their physiological expression of p53, serving as internal control for the sample under investigation. If the positive control material or internal control was negative, the slide was disregarded for analysis. The pattern of p53 IHC was scored on all tumors cells present on the slide, based on the percentage of tumor cells with nuclear positivity on a semi-quantitative 7-point scale: 0%, 1-20%, 21-40%, 41-60%, 61-80%, 81-90% and 90-100% of the tumor cells. If the scores of the two pathologists were discordant, a third board certified pathologist evaluated the slides (MD), after which the final diagnosis was based on the consensus of two of the three pathologists. All pathologists were blinded for clinical and pathological data.

Mutational analysis and high-throughput methylation profiling

In total 34 EAC, among them 10 with no expression of p53, 14 with heterogeneous expression (1-60% of the tumor cells positive) and 10 with overexpression (61-100% positive tumor cells), were selected for targeted gene sequencing. Tumor area was manually macrodissected from the successive unstained slides, resulting in at least 30% tumor cells. DNA was extracted using proteinase K and 5% Chelex 100 resin.[174] An Ion AmpliSeq custom-made panel was created for selection of genes.[175] This consisted of primers for the entire TP53 gene supplemented with hotspots or the entire genes known to be frequently altered in EAC (ARID1A, PIK3CA, APC, DOCK2, ELMO1, CDKN2A and SMAD4).[11, 163, 176] Sequencing was performed on the Ion Torrent Personal Genome Machine or IonS5 system (Thermofisher Scientific, Hemel Hempstead, UK) according to the manufacturers protocol. In short, libraries were created using the ION AmpliSeq Library Preparation Kit. Template was prepared using the Ion Onetouch Template Kit and sequencing was performed with the Ion Sequencing Kit as described.[174] One sample was excluded from further analysis because of poor DNA quality and high frequencies of formalin artefacts. All other samples showed comparable and reliable sequence read coverage independent from sample age. The sequence variants with a read frequency of less than 5% (homozygous reference) or more than 95% (homozygous non-reference), with an amplicon coverage of less than 50, or a variant coverage of less than 10 reads were excluded from analysis, to eliminate formalin artefacts. All variants found in an intronic, intergenic, non-coding RNA or UTR3/5 region, and synonymous single nucleotide variations (SNV) were excluded.

Sixteen EAC, among them five tumors with loss of expression, five with overexpression and six with heterogeneous p53 expression, were selected for genomewide methylation analysis in addition to the targeted sequencing. Therefore, the Infinium MethylationEPIC BeadChip (Illumina, San Diego, CA, USA), targeting over 850,000 methylation sites, was applied according to the manufacturer's instruction at the Microarray unit of the Genomics

and Proteomics Core Facility of the German Cancer Research Center (DKFZ, Heidelberg, Germany). For a detailed description see earlier publication.[177]

For unsupervised clustering the most differential probes (with 0,22 SD difference from the mean) were selected. To assess copy number variation (CNV) methylation data were implemented in the R/Bioconducter packages Conumee. Intra-chromosomal breaks were calculated from the number of segments defined by the Conumee package (blue horizontal lines in Supplementary Figure 3). Segments are defined as chromosomal regions with distinct copy number changes to the adjacent region.

The number of segments relative to the median number of segments within this sample series was determined for each sample (presented in Figure 4). With this method amplification of genes were also assessed as described earlier.[178] To validate amplification of MDM2 immunohistochemistry staining (clone 1F2, Merck Milipore, Amsterdam, Holland) was performed on all samples in which no TP53 mutation was found.

Ethics

The investigational protocol was approved by the medical ethical committee in the Erasmus Medical Center (Rotterdam, The Netherlands) (MEC-12-469).

Statistical analysis

The primary endpoint of the study was 5-year DFS, defined as the time between surgery and the first clinical recurrence of disease, defined as clinical or radiological evidence of disease recurrence. Patients lost to follow-up were censored at the time of the last visit to the outpatient clinics. Secondary endpoint was OS, defined as time between surgery and death. The optimal cut-off for IHC was calculated using a ROC-curve and corresponding Youden-index (Supplementary Figure 1 and Supplementary Table 2).

The interobserver variation for the assessment of p53 staining between the two pathologists was calculated using Cohen's kappa. Strength of agreement was categorized as follows: 0.00–0.20, poor; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, good; and 0.81–1.00, excellent.

Kaplan Meier curves were used to plot the 5-year DFS by p53 status. Uni- and multivariable Cox proportional hazard models were applied to calculate the association between p53 IHC and survival. In the multivariable analysis adjustments were made for all clinical and pathological factors which proved to be prognostic for survival in the univariable analysis ($p<0,05$). The pN-stage was dichotomized in pN0 and a pN+ (pN1-3) group for the Cox regression analysis. The p53 status and mutational status were correlated using Fisher's

Exact test. The analysis was performed using SPSS-software (version 22, SPSS IBM inc, Armonk, NY, USA). A p-value of <0.05 was considered statistically significant.

RESULTS

Patient characteristics

Two hundred and sixteen (216) patients were initially identified to be eligible for this study. Of 12 patients, the formalin-fixed paraffin-embedded (FFPE) blocks could not be retrieved and were therefore excluded. From the remaining 204 patients with EAC the majority had a pT3-tumor (85.3%), tumor positive lymph nodes (79.4%) and negative resection margins (62.7%). Detailed patient and tumor characteristics are shown in Table 1 and Supplementary Table 1.

P53 expression correlates with overall–and disease free survival

The optimal cut-off for p53 expression was calculated, based on the receiver operating characteristics (ROC) curve and Youden-index (see Supplementary Table 2 and Supplementary Figure 1), into three groups, namely loss of expression (0% of tumor cells positive), heterogeneous expression (1-60% of tumor cells positive) and overexpression (61-100% of tumor cells positive). The interobserver variation for the assessment of p53 between the two observers was excellent (kappa 0.850, p<0.001). From the 204 patients, 55.9% (n=114) of the EAC showed overexpression, 26.5% (n=54) loss of expression, while 17.6% (n=36) had a heterogeneous expression. In all cases this was a homogeneous expression pattern throughout the cancer, of which representative examples are shown in Figure 1.

The pattern of p53 expression associated with disease free survival (DFS); overexpression–median DFS 14.6 months (95% CI 10.0-19.2), loss of expression–median DFS 14.2 months (95% CI 7.9-20.5) compared to the group with heterogeneous p53 expression–median DFS 37.1 months (95% CI 24.3-49.9). The corresponding Kaplan-Meier curves are shown in Figure 2.

Univariable analysis demonstrated a correlation between p53 expression and DFS (p=0.036). The risk of recurrence of EAC was increased for patient with p53 overexpression (hazard ratio (HR) 1.91; 95% CI 1.16- 3.14) as well as loss of p53 expression (HR 1.57; 95% CI 0.9-2.74) compared to heterogeneous p53 expression. This was also significant after multivariable analysis, adjusted for pT-stage, pN-stage, tumor differentiation and resection margin status (p=0.001). Patients with p53 overexpression/loss showed a significantly worse DFS compared to heterogeneous expression (HR 2.61; 95% CI 1.57-4.32; p= <0.001 and HR 2.75; 95% CI 1.55-4.9; p= <0.001, respectively) (Table 2 and Supplementary Table 3). A shorter

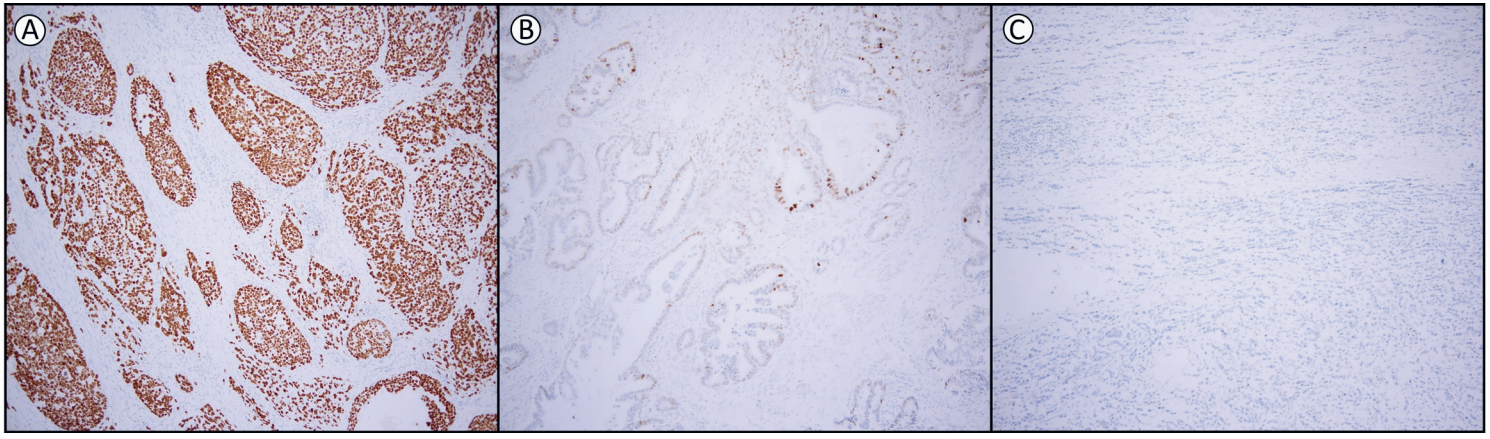


Figure 1. Examples of p53 expression in esophageal adenocarcinoma. A) overexpression (61-100% positive tumor cells) B) heterogeneous expression (1-60% positive tumor cells) and C) loss of expression (0% positive tumor cells). Magnification 1:100.

Table 1. Clinico-pathological characteristics for the 204 included patients with esophageal adenocarcinoma

	All patients n=204		p53 Loss (0%) n=54		p53 Heterogeneous (1- 60%) n=36		p53 Overexpression (61-100%) =114		p-value
	N	%	N	%	N	%	N	%	
Age at surgery									
Median	64.0		63.0		68.5		64.0		0.462
Range (IQR)	55.3-72.0		55.0-72.0		56.3-74.0		55.0-72.0		
Sex									
Male	174	85.3	51	25.0	29	14.2	95	46.6	0.337
Female	30	14.7	5	2.5	7	3.4	17	8.3	
Siewert classification									
Type 1	75	36.8	23	11.3	11	5.4	41	21.1	0.576
Type 2	129	63.2	33	16.2	25	12.3	71	34.8	
Pathologic T-stage									
pT2	27	13.2	9	4.4	3	1.5	16	7.8	0.556
pT3 or pT4	177	86.8	47	23.0	33	16.2	96	47.1	
Pathologic N-stage									
pN0	42	20.6	16	7.8	5	2.5	22	10.8	0.207
pN1 or more	162	79.4	40	19.6	31	15.2	90	44.1	
Histology grade									
Well	5	2.5	3	1.5	1	0.5	1	0.5	0.498
Moderate	80	39.2	22	10.8	15	7.4	43	21.1	
Poor	119	58.3	31	15.2	20	9.8	68	33.3	
Resection margin status									
pR0	128	62.7	33	16.2	22	10.8	73	35.8	0.714
pR1	76	37.3	23	11.3	14	6.9	39	19.1	
Alive after 60 months									
Yes	34	16.7	10	4.9	8	3.9	16	7.8	0.518
No	170	83.3	46	22.5	28	13.7	96	47.1	

overall survival (OS) was associated with p53 overexpression (median OS 19.4 months (95% CI 14.3-24.5)), and loss of expression (median OS 18.5 months (95% CI 15.3-21.7)) compared to the group with heterogeneous expression (median OS 32.4 months (95% CI 23.0-41.8)). Although no significance was identified in the univariable analysis (p=0.265), the multivariable analysis demonstrated that p53 expression was significantly associated with OS (p=0.003). Overexpression and loss of p53 expression were prognostic for a shorter survival period (HR respectively 1.99; 95% CI 1.29-3.07; p=0.002 and 2.17; 95% CI 1.33- 3.55; p=0.002) compared to heterogeneous expression (Table 2 and Supplementary Table 3).

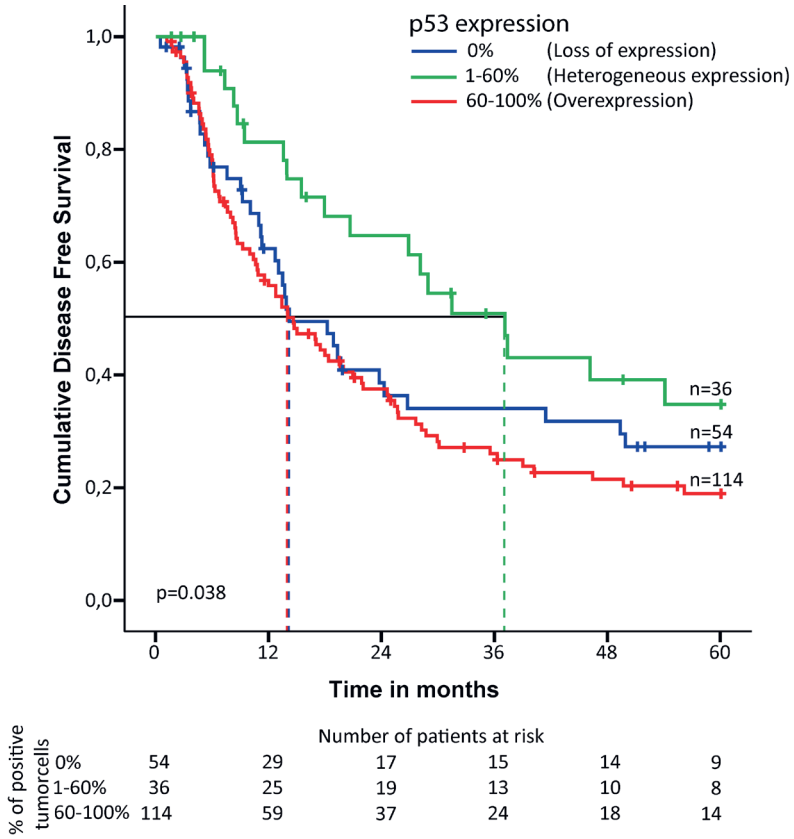


Figure 2. Kaplan-Meier curve for disease free survival in chemoradiotherapy-naive patients with esophageal adenocarcinoma. Expression pattern of p53 is subdivided into three groups: 0% of the tumor cells positive (loss of expression), 1-60% of the tumor cells positive (heterogeneous expression) and 61-100% of the tumor cells positive (overexpression). The dotted line indicates the median survival for each of the three groups. Number of patients at risk is indicated for each of the three groups at the bottom of the figure.

Targeted mutational analyses and high-throughput methylation profiling

To shed light on the possible mechanism(s) underlying the p53 staining patterns, sequencing of the whole *TP53* gene was performed using the Ion Torrent platform on 33 selected EAC (10 with overexpression, 10 with loss, and 13 with a heterogeneous expression) (Supplementary Table 4). Overall, 25 of 33 (76%) EAC showed a *TP53* mutation. *TP53* status correlated significantly with the IHC staining pattern (p=0.035) (Figure 3 and Supplementary Table 5). Of the 10 cases with loss of expression eight had non-sense mutations (splice site, frameshift mutation or stopgain) and two no mutation. All EAC with overexpression of p53 as detected by IHC had missense mutations. The EAC within the heterogeneous p53

Table 2. Multivariable Cox regression analysis for disease free survival and overall survival in patients with esophageal adenocarcinoma

	Multivariable Cox regression analysis					
	Disease free survival			Overall survival		
	HR	95% CI	p-value	HR	95% CI	p-value
Age	NA	NA	NA	1.026	1.010-1.042	0.001
pT-stage (ref. = pT2) pT3/4	2.152	1.156-4.005	0.016	2.010	1.168-3.459	0.012
pN-stage (ref. = pN0) pN+	3.445	1.981-5.990	<0.001	2.434	1.560-3.796	<0.001
Differentiation (ref. = Good/ Moderate)						
Poor	1.467	1.016-2.119	0.041	1.551	1.112-2.165	0.010
Resection margin (ref. = pR0)						
pR+	1.721	1.192-2.484	0.004	1.716	1.230-2.393	0.001
p53 (ref. = heterogeneous)						
loss of expression	2.754	1.547-4.903	0.001*	2.174	1.333-3.546	0.003*
overexpression	2.605	1.571-4.320		1.989	1.288-3.071	

*HR=Hazard Ratio, CI=Confidence Interval, NA = not available, excluding patients who died within one month after surgery. P53 expression, based on immunohistochemical expression, was classified as loss of expression (0% of the tumor cells positive), heterogeneous expression (1-60% of the tumor cells positive) and overexpression (61-100% of the tumor cells positive). *global p-value*

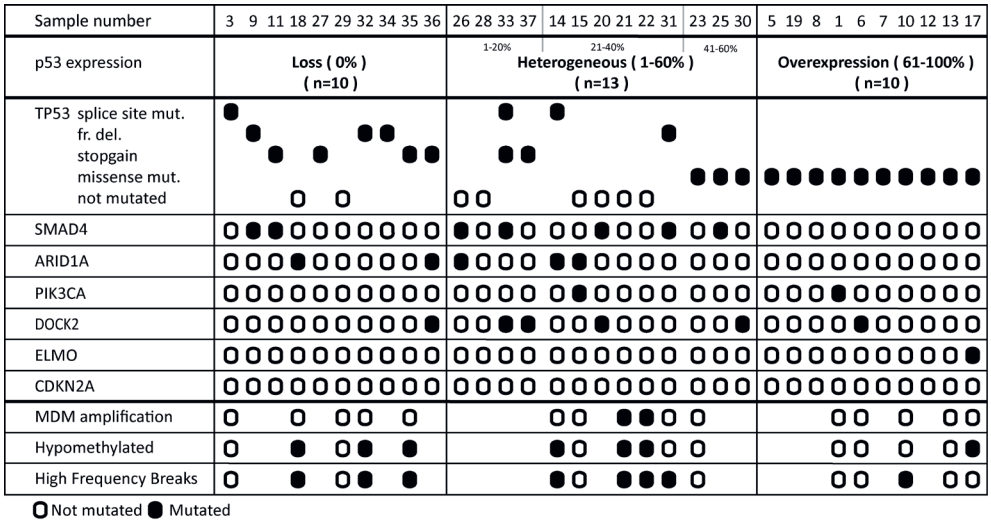
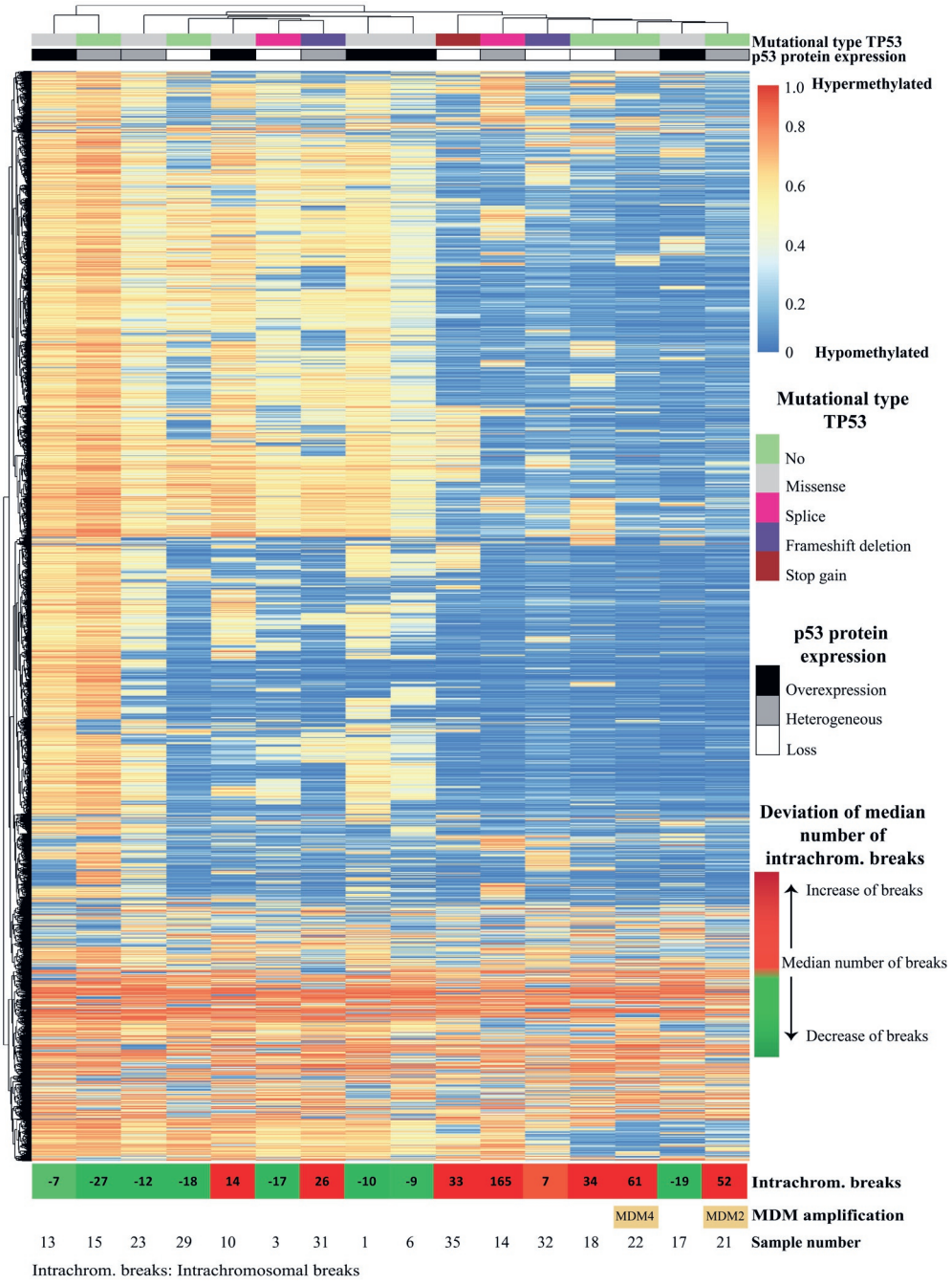


Figure 3. Molecular profile of esophageal adenocarcinoma. Mutational profile, as detected by targeted sequencing in 33 cases is categorized by p53 expression pattern. The order of samples is determined by the percentage of positive p53 tumor cells. The exact mutations found are displayed in Supplementary Table 4, ordered by case number. The CpG methylation-derived information (copy numbers, hypo-methylation status and relative high frequency breaks are summarized here. Further details are provided in the Supplementary Figures 2 and 3.



Intrachrom. breaks: Intrachromosomal breaks

< Figure 4. Unsupervised clustering of CpG methylation of 16 esophageal adenocarcinomas, using top 10,454 differential probes, in which every row is a methylation probe and every column is a sample. The corresponding violin plots for all methylation probes as well as top 10,454 differential probes are depicted in Supplementary Figure 3. At the top of the image, p53 expression and *TP53* mutational status are indicated for each sample. In the bottom of the image, the deviation of the median number of intra-chromosomal breaks (median number of intra-chromosomal breaks are calculated, samples with less breaks are depicted in green, and samples with more breaks depicted in red, numbers indicate deviation from median number of intra-chromosomal breaks) as well as the two samples with *MDM2* and *MDM4*-amplification are indicated. Five samples with an altered p53 pathway (either *TP53* mutations or *MDM2/4* amplifications) show extensive intra-chromosomal breaks (sample 31, 35, 14, 22 and 21), and only one sample without alteration of the p53 pathway (sample 18) showed extensive intra-chromosomal breaks. However none of the six samples (sample number 13, 23, 10, 1, 6 and 17) with a missense mutation showed an increase of intra-chromosomal breaks. Increased number of intra-chromosomal breaks correlates to the clustered hypomethylated CpG-sites.

expression group demonstrated a mixed picture, representing the three different patterns. Those with more than 40% p53 positive tumor cells all showed missense mutations (n=3), in analogy to EAC with overexpression, while in the lower percentage category two out of four showed a nonsense mutation (one containing both a splice site and stopgain mutation). EAC cases with heterogeneous p53 expression in the middle group (n=6, 21-40%) demonstrated no underlying *TP53* mutations in four and two nonsense mutations. Besides *TP53*, in total, 21 other proven pathogenic mutations in the following genes *SMAD4* (n=7), *ARID1A* (n=5), *PIK3CA* (n=2), *DOCK2* (n=6) and *ELMO* (n=1) were detected, significantly more in EAC with a heterogeneous p53expression (13/21; p=0.032) (Figure 3 and Supplementary Table 5). In these samples no mutations in CDKN2A were detected. Four cases of our series revealed no mutation in the investigated genes (cases 21, 22, 28 and 29). Multiple mutations were identified (including *TP53*) in 15 EAC, predominantly again in the heterogeneous p53 expression group (9/13 versus 3/10 and 3/10, respectively). In addition, a subset of these EAC (n=16) were investigated using high-throughput methylation profiling for the detection of chromosomal alterations between the three groups,[11] including five with overexpression, five with loss and six with a heterogeneous expression, all with known *TP53* mutational status (see Figure 4). No hypermethylation of the promotor region of *TP53* was detected in any of these EAC, including the two cases with loss of p53 expression and wild type (not mutated) *TP53* (cases 18 and 29).

Based on copy number variations (CNV) derived from these high-throughput methylation profiles (see Material and Methods section), regional chromosomal amplifications were identified, including those encompassing for example *MDM2* and *MDM4*, two genes of which amplification is known to be related to an alternative inactivation of p53 besides mutations. Two EAC showed such an amplification (cases 21 and 22, for *MDM2*, confirmed by immunohistochemistry, and *MDM4*, respectively, see Supplementary Figure 2). No other mutations were identified in these cases, and both showed a heterogeneous p53 expression

(21- 40% of positive tumor cells) (Figures 3 and 4). Besides these specific amplifications, an unsupervised clustering of the top 10,454 most differentiating CpG-sites was performed (see Figure 4 (heatmap) and Supplementary Figure 3 (Violin plots)). No difference was identified for the overall methylation distribution between the EAC investigated (Supplementary Figure 3, bottom panel), while a clear hypomethylation profile was identified for the most differentiation CpG-sites in seven EAC out of the 16 cases. These included three with absence, three with a heterogeneous and one with overexpression of p53. Only one showed no *TP53* anomaly (case 18, no p53 expression), while all others demonstrated either a mutation in *TP53* itself (three nonsense, one missense), or amplification of *MDM2* or *MDM4*. In addition, the number of intrachromosomal breaks per individual EAC was scored based on the CNV profile (see Supplementary Figure 2 and Figure 4). This analysis demonstrated that six out of the seven EAC with a hypomethylation profile showed a higher number of breaks compared to the group median, i.e., indicated in red boxes in Figure 4 (including those with the *MDM2* and *MDM4* amplification), while this was observed for only two of the EAC within the non-hypomethylated group. These data suggest that there is a correlation between p53 status (protein expression, mutational profile and *MDM2/4* amplification), accumulation of other mutations (preferentially in the p53 heterogeneous staining group), preferential presence of a hypomethylated profile in the loss and heterogeneous p53 group, and occurrence of intrachromosomal breaks.

DISCUSSION

This study primarily aimed to evaluate the relevance of p53 IHC for survival of patients with advanced EAC. A large, well defined cohort of CRT-naive surgically treated EAC was analyzed, and the pattern of p53 expression was shown to be significantly correlated with DFS and OS, independently from other clinico-pathological parameters including tumor stage. In addition, p53 expression patterns were correlated with the underlying *TP53* mutational status and genome wide methylation profile and derived information on chromosomal anomalies.

TP53 is one of the driving genes for the progression of BE into adenocarcinoma and whole genome sequencing studies have detected a high mutation frequency of *TP53* in EAC.[164, 176, 179] Conflicting results have, however, been reported so far on *TP53* and survival in patients with EAC.[171, 172, 180-185] Three previous systematic reviews analyzed the current literature and performed a meta-analysis of up to 16 different studies, employing IHC or sequencing of the *TP53* gene.[159-161] Although, overall, similar results were reported in all three meta-analyses suggesting a negative effect of mutated *TP53* on prognosis, the data should be interpreted with caution. First of all, many of the earlier studies did not consider the bias of patient selection and chemoradiotherapy (CRT) treatment.[170, 183, 186-188] Several

studies included patients who received surgery only as well as patients who underwent neoadjuvant treatment or definite CRT. This is of importance since p53 might modulate CRT response as suggested in earlier studies.[170, 186-193] Another important limitation of the published studies is the inconsistent methodology for detection and classification of p53 expression. From five studies using IHC on homogeneous EAC cohorts (total 384 patients), with surgery as single treatment modality and IHC approach, none qualified loss of expression as aberrant (see Supplementary Table 6).[171, 172, 181, 182, 184] This is significant since according to our interpretation, around 26% of EAC showed loss of p53 expression and had significantly worse outcome. In the present study based on evaluation of EAC resection specimens of 204 CRT-naive patients, with surgery as single modality, p53 was detected by IHC and categorized by experienced observers using optimized cut-off values. The pattern was classified as heterogeneous, overexpression or loss of expression.

Until now it is not clear whether p53 IHC or sequencing of *TP53* is the most optimal tool to improve risk stratification in EAC. Mutational status was suggested to be preferable by a recent meta-analysis.[160] Several previous EAC studies applied mutational status as single read out. [170, 172, 186] The assays used for gene sequencing in those older studies are likely to be suboptimal, since the *TP53* gene was only partly sequenced using PCR-based methods, which correlates with the low mutational rate (40-50%).[170, 172, 186] Although the efficacy of the gene sequencing techniques improved in recent years, they are still more time-consuming, labor intensive and expensive compared to IHC. Prediction of mutational status by IHC could be an alternative, but the prognostic accuracy might depend on the underlying cancer type.[194] To study the correlation between protein expression pattern and genetic status, a subset of 33 EAC was investigated using a targeted next generation sequencing approach. *TP53* mutational frequency rate was 76%, which is comparable to the recent investigations using whole genome or exome sequencing techniques.[176, 179] *TP53* status significantly correlated with the defined IHC categories (p=0.035). EAC with heterogeneous p53 expression was also heterogeneous in terms of the underlying *TP53* status, although it seems to be (again) subdivided into three groups, similar to loss of expression, similar to overexpression, and the (remaining) intermediate group.

Of interest is that most additional mutations in the other candidate genes investigated were identified in the group with heterogeneous p53 expression, including two cases with regional amplifications of *MDM2* or *MDM4* (Figure 3). These were identified in EAC without any other mutation. In contrast, all EAC with high percentage of p53 positive cells (more than 61%, n=10) showed missense mutations in *TP53*, which is in line with results of two earlier studies.[172, 195] EAC with loss of p53 expression demonstrated predominantly nonsense mutations, including splicing, stopgain and frameshift mutations (8/10). These nonsense mutations were also observed in a subset of EAC with a heterogeneous, but relatively low to modest p53 expression, in fact three out of five cases. In 4 out of five of the remaining

cases no *TP53* mutation was found. These observations warrant additional studies to be performed.

The putative difference in pathogenesis between these subgroups is supported by the results of the high-throughput methylation profiling performed. The hypomethylated profile of the most differentiating CpG sites combined with a high frequency of intrachromosomal breaks was predominantly observed in EAC with loss or a heterogeneous p53 pattern (either by a nonsense mutation (n=3) or *MDM2/4* amplification (n=2)). No apparent differences were observed using all CpG targets, demonstrating its specificity. EAC with a hypomethylated profile showed a higher frequency of intrachromosomal breaks, indicative for chromosomal instability. This is in line with the recently suggested role of DNA methylation as the newly identified guardian of the genome.[196] Based on this small subset of patients, these observations might be a potential explanation for the differences in DFS and OS as found in the present study, which warrants further investigations. Besides the prognostic effect of p53 expression, our results are clinically important. *TP53* status might be predictive for response to neoadjuvant chemotherapy.[74, 170] Clinical trials, such as the PANCHO trial, stratified for *TP53* status, are underway and have completed recruitment.[197] Other studies rely on new therapeutic agents created to restore the wild type activity of p53, one of the most promising compounds being APR-246.[198] Here we show that if IHC is used as a read-out for mutational status, results should be interpreted with caution especially in EAC with a heterogeneous p53 expression. In contrast, EAC with p53 overexpression or loss of expression are likely to have an underlying somatic mutation and extensive sequencing might not be necessary.

There are some limitations to this study. *TP53* sequencing was done in a single EAC area, and therefore potential intratumoral heterogeneity was not accounted for. However, this is considered unlikely to play an important role, since identical *TP53* mutations and homogeneous loss of heterozygosity of the *TP53* locus were detected across separated tumor regions in EAC previously,[175] and a homogenous IHC was identified in all cases. Furthermore, although p53 is stained using a proven informative automatic staining system and a standardized protocol, the scoring is subjective in nature. However, the interobserver variation for p53 IHC was excellent.

In summary, this study leads to various conclusions. First of all, we have demonstrated that p53 expression pattern is significantly correlated with DFS and OS. This finding stresses the biological role of p53 for the prognosis of patients with EAC. Secondly, we have shown that IHC is a good read out for the presence of *TP53* mutations mainly in EAC with p53 overexpression and probably in EAC with loss of expression but not in EAC with a heterogeneous p53 expression. This might be important for current and future studies in which patient treatment is stratified according to the *TP53*/p53 status. In addition,

our study could suggests existence of different pathogenesis of EAC, related to the p53 pathway (*TP53* mutational status and *MDM2/4* amplification), with downstream additional mutations of other candidate genes, as well as DNA methylation alterations and possibly related chromosomal instability. Yet, more work needs to be done for accurate genetic classification of EAC to fully reveal prognostic genetic signatures and involved mechanisms.

AUTHOR CONTRIBUTIONS

F.J.C. ten Kate: acquisition of data, administrative, technical, or material support, analysis and interpretation of data, statistical analysis, drafting of the manuscript; L. Suzuki, analysis and interpretation of data, critical revision of the manuscript; L.C.J. Dorssers: bioinformatics analysis and interpretation of data, critical revision of the manuscript; W.N.M. Dinjens, acquisition of data, critical revision of the manuscript; D.T.W. Jones: acquisition of data, critical revision of the manuscript; D. Nieboer, statistical analysis, critical revision of the manuscript; M. Doukas: acquisition of data, critical revision of the manuscript; J.J.B. van Lanschot, acquisition of data, critical revision of the manuscript; B.P.L. Wijnhoven, acquisition of data, critical revision of the manuscript; L.H.J. Looijenga: study concept and design analysis and interpretation of data critical revision of the manuscript study supervision; K. Biermann, study concept and design acquisition of data analysis and interpretation of data drafting of the manuscript study supervision.

CONFLICTS OF INTEREST

None of the participating authors have a conflicts of interest to declare.

FUNDING

This study was financed by Erasmus MC Fellowship appointed to Katharina Biermann.

SUPPLEMENTARY MATERIALS

Supplementary Table 1. Basic clinico-pathological characteristics for all patients subjected to mutational analysis.

	TP53 sequencing (n=33)	
	N	%
Age at surgery		
Median	63.00	
Range (IQR)	54.50-71.50	
Sex		
Male	31	93.9
Female	2	6.1
Siewert classification		
Type 1	12	36.4
Type 2	21	63.6
Pathologic T-stage		
pT2	3	9.1
pT3 or pT4	30	90.9
Pathologic N-stage		
pN0	6	18.2
pN1 or more	27	81.8
Histology grade		
Well	1	3.0
Moderate	10	30.3
Poor	22	66.7
Resection margin status		
pR0	22	66.7
pR1	11	33.3
Follow-up time, months		
Median	19.8	
Range (IQR)	9.63-41.33	
p53 expression		
0%	10	30.3
1-60%	13	39.4
61-100%	10	30.3

Supplementary Table 2. Calculation of optimal cut-off for % nuclear positive tumor cells for p53 immunohistochemistry.

% p53 positive tumor cells	Sensitivity	Specificity	Youden-index
1-20	1	0	0
21-40	0.944	0.127	0.071
41-60	0.894	0.206	0.1
61-80	0.866	0.27	0.136
81-90	0.775	0.317	0.092
91-100	0.69	0.397	0.087
0	0.254	0.698	-0.048

Supplementary Table 3. Univariate Cox regression analysis for disease free survival and overall survival in neoadjuvant treatment naïve patients with esophageal adenocarcinoma.

	Univariable Cox regression analysis					
	Disease free survival			Overall survival		
	HR	95% CI	p-value	HR	95% CI	p-value
Age	1.007	0.990-1.024	0.412	1.017	1.002-1.033	0.026
Sex (ref. Male)						
Female	1.172	0.750-1.833	0.486	0.942	0.612-1.451	0.786
Weight	0.990	0.977-1.003	0.142	0.992	0.981-1.004	0.171
Siewert (ref. Type 1)						
Type 2	0.859	0.612-1.205	0.379	0.871	0.639-1.186	0.380
pT-stage (ref. pT2)						
pT3/4	2.723	1.503-4.932	0.001	2.394	1.427-4.014	0.001
pN-stage (ref. pN0)						
pN+	3.504	2.044-6.007	<0.001	2.460	1.602-3.778	<0.001
Differentiation (ref. Well/ Moderate)						
Poor	1.716	1.216-2.421	0.002	1.544	1.134-2.104	0.006
Resection margin (ref. pR0)						
pR+	2.143	1.528-3.005	<0.001	2.101	1.540-2.867	<0.001
P53(ref. Heterogeneous)						
Loss of expression	1.569	0.897-2.743	0.036	1.333	0.831-2.138	0.265
Overexpression	1.909	1.161-3.139		1.420	0.931-2.165	

HR=Hazard Ratio, CI=Confidence Interval, patients who died within one month of surgery were excluded, p53 immunohistochemistry assessed as loss (0% of the tumor cells positive), heterogeneous expression (1-60% of the tumor cells positive) and overexpression (61-100% of the tumor cells positive).

Supplementary Table 4. Summary of mutations found by Ion Torrent Sequencing of our custom made gene panel. See Supplementary File 1 (oncotarget-08-104123-s002.xlsx) on the following website: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5732792/

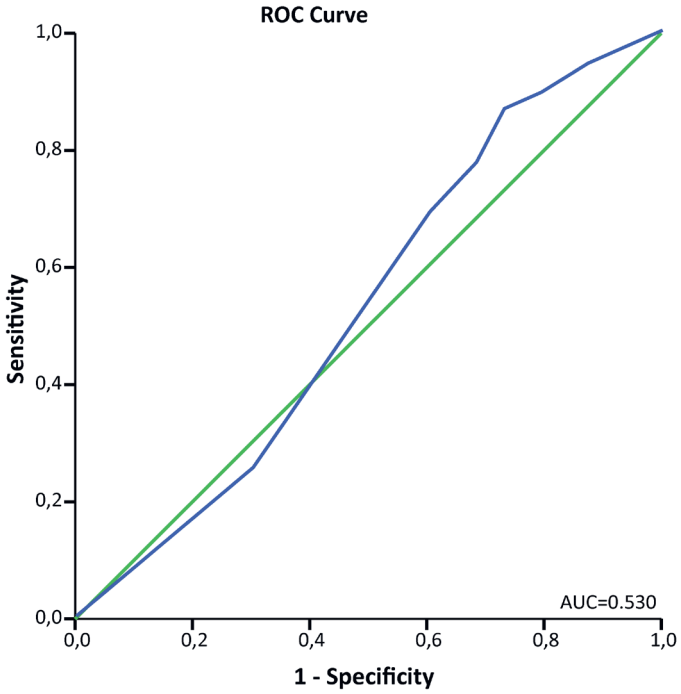
Supplementary Table 5. Number of *TP53* mutations and other mutations (*SMAD4*, *ARID1A*, *PIK3CA*, *DOCK2*, *ELMO* and *CDKN2A*) found by Next Generation Sequencing specified by tumors with aberrant (combined loss of expression and overexpression) and heterogeneous expression of p53 immunohistochemistry (IHC). Difference is calculated by Fisher exact test.

	p53 IHC expression		
	Loss and Overexpression	**Heterogeneous**	**p-value**
***TP53* mutated**	18	7	0.035
***TP53* not mutated**	2	6	
Other mutations	7	10	0.032
No other mutations	13	3	

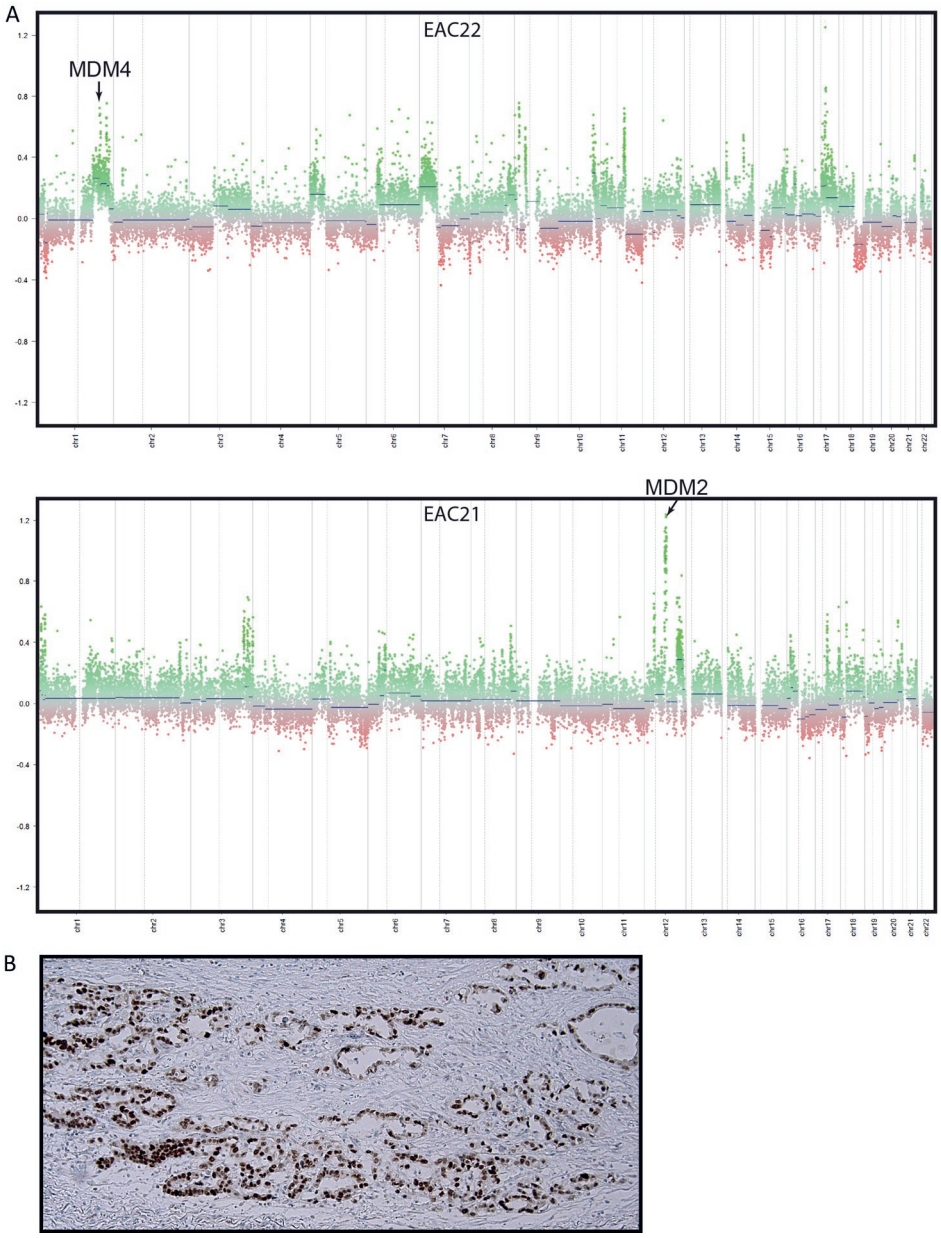
Supplementary Table 6. Characteristics of various earlier studies on the prognostic value of p53 in patients with esophageal adenocarcinoma (EAC).

First Author [Reference]	Year	n EAC (% of total included patients)	Specimen	CRT	p53 clone antibody used	Cut-off for over expression	Loss of expression considered?	Predictive
Flejou[184]	1994	62 (100%)	Resection	No	DO7	NA	No	No
Aloia[180]	2001	44 (72%)	Resection	No	PAb1801	NA	No	Yes
Falkenback[182]	2008	59 (100%)	Resection	No	DO7	5%	No	No
Madani[172]	2010	142 (100%)	Resection	No	DO7	1%	No	Yes
Cavazzola[181]	2009	46 (100%)	Resection	No	DO7	10%	No	No
Lehrbach[171]	2009	75 (100%)	Resection	No	DO7	<2 on scale of 5	No	No
Fareed[183]	2010	245 (94%)	TMA	Yes	NA	10%	No	Yes*
Duhaylongsod[199]	1995	42 (100%)	Resection	Yes	PAb1801	NA	No	No
Sauter[192]	1995	24 (100%)	Biopsy +resection	Yes	PAb1801	>5 adjacent cells in 1HPF	No	Yes
Moskaluk[185] §	1996	88 (100%)	Resection	Yes	DO7	50%	No	No
Wu[188] §	1998	92 (100%)	Resection	Yes	DO7	50%	No	No
Ribeiro[195]	1998	42 (74%	Resection	Yes	DO7	Weak positive	No	No

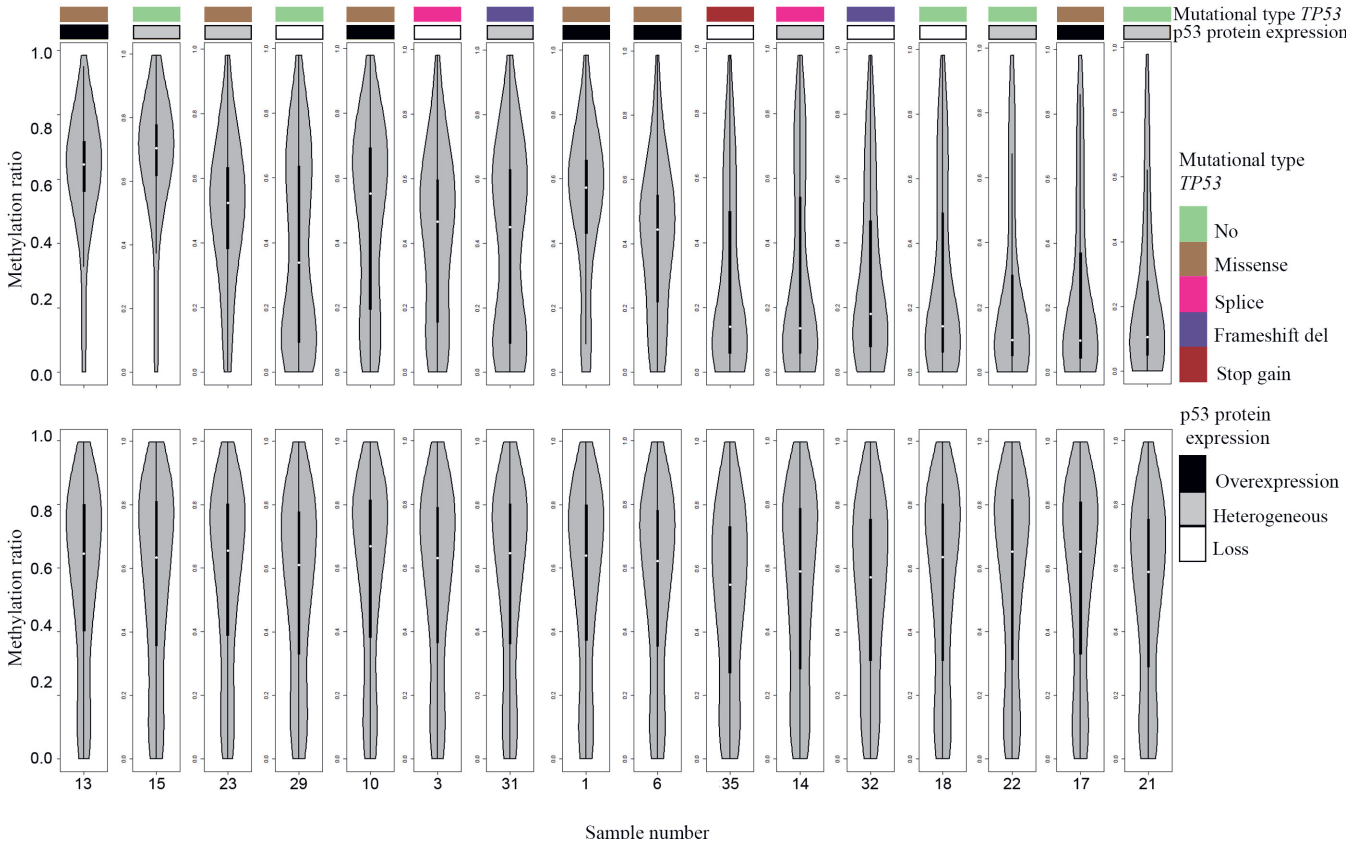
*: in patients treated with chemotherapy p53 was not predictive for survival. §: Overlap in patient group. NA: Not Available. CRT: chemoradiotherapy.



Supplementary Figure 1. Receiver operating characteristics-curve for the semi-quantitative p53 expression, according for the % of nuclear positive tumors cells, which is used to calculate the optimal cut-off value of p53 expression.



Supplementary Figure 2. A) Copy number profiles of selected esophageal adenocarcinoma (EAC) cases (sample number 21 and 22). B) MDM2 immunohistochemistry confirming amplification (sample number 21). A) Methylation intensity data were used to calculate relative copy numbers (output of the Conumee software package). Under- (red) and over-represented (green) regions are highlighted. The positions of the *MDM2* and *MDM4* amplicon peaks are indicated. The blue lines represent the regions within chromosomes (segments) with similar copy number. The total number of segments was determined per sample to estimate the relative frequency of intrachromosomal breaks in each case. B) MDM2 amplification of sample number 21 validated by immunohistochemical staining, magnification 100x.



Supplementary Figure 3. Violin plots of the selected and overall CpG methylation data. Upper panel: the methylation profile of the top 10,545 differential probes; lower panel: the methylation profile of all available methylation probes used, indicating no significant differences between cases based on all available methylation data. Abbreviation: Frameshift del: Frameshift deletion.

7

Effect of neoadjuvant chemoradiotherapy on p53 and SOX2 protein expression in esophageal adenocarcinoma

Suzuki L, Nieboer D, van Lanschot JJB, Spaander MCW, Looijenga LHJ, Biermann K.

Biomarkers in Medicine, 2020, 14(9), 785–793.

Background: To determine if neoadjuvant chemoradiotherapy (nCRT) affects p53 and SOX2 expression in esophageal adenocarcinoma (EAC).

Methods: Comparison of p53 and SOX2 expression in 100 paired pre- and post-nCRT EAC samples.

Results: Aberrant p53 was largely concordant (75/83, 90%), while 13/18 (72%) pre-nCRT samples with wild-type (WT) p53 staining, showed aberrant staining in paired post-nCRT samples. Similarly, 31/45 (69%) with previous WT SOX2 showed SOX2 loss in paired post-nCRT samples, whereas aberrant SOX2 loss was concordant in 50/55 (91%) cases. The prognostic values of both markers regarding survival differ before and after nCRT.

Conclusion: Aberrant expression of p53 and SOX2 staining in EAC tissue is unaffected by nCRT. Conversely, the WT-staining pattern frequently changed to aberrant expression.

Abstract

INTRODUCTION

Despite major advances in therapy, long-term survival rates of esophageal adenocarcinoma (EAC) remain moderate. Prognosis relies on both clinical and histopathological factors, although it remains a challenge to discriminate between patients who will show rapid progression or may have a more indolent course. Multimodal therapy, usually a combination of neoadjuvant chemoradiotherapy (nCRT) prior to surgery, has become standard of care since it was found to be superior compared to surgery alone in multiple trials.[32] Moreover, it is currently being addressed if clinically complete responders after nCRT can be kept under close surveillance and only undergo surgery, if evident localized tumor re-presents (SANO trial, Surgery As Needed in Oesophageal cancer, Netherlands Trial Register NTR6803).[33] Therefore, treatment of patients with EAC is increasingly based on small biopsy samples and in a subset of patients, i.e. patients with complete (clinical) response, a surgical sample may even not be available anymore.

Biomarkers may improve prognostication and additionally provide some insight in the biological behavior of EAC. For example, p53 expression was found to be independently prognostic for disease free survival (DFS) in chemotherapy-naïve EAC.[73] Similarly, Sex Determining Region Y-Box2 (SOX2) was found to be independently prognostic for survival in two independent EAC cohorts.[75] In addition, p53 and SOX2 expression in pre-treatment EAC biopsies were found to be predictive for response to nCRT.[74] These observations suggest major biological roles of p53 and SOX2 in the clinical behavior of EAC. However, most biomarker studies are based on historical cohorts of patients treated with surgery alone and to date, the effect of nCRT on p53 and SOX2 expression in EAC and the prognostic value of these markers pre- and post-treatment are unknown.

This study compares immunohistochemical p53 and SOX2 expression in paired pre-treatment (pre-nCRT) endoscopic biopsies and resection specimens after nCRT in one hundred EAC patients. As secondary outcome, both p53 and SOX2 expression before (endoscopic biopsies) and after nCRT (resection specimen) will be correlated with overall and disease-free survival.

METHODS

Patient selection

The total cohort of patients previously described by Van Olphen *et al*., was used for this study.[74] One of the patients has objected to the use of material for research and was therefore excluded from the present study. In short, 147 patients with histologically proven EAC who received at least 80% nCRT according to the CROSS regimen (Carboplatin and

Paclitaxel with concurrent radiation) followed by esophagectomy treated at the Erasmus University Medical Center with sufficient and representative material were included. Tumor regression grade (TRG) was evaluated according to the modified Mandard scoring system. Patients with TRG 1 and 2 were considered as major responders (n=74), and patients with TRG 3 and 4 as minor responders (n=73). Patients with TRG1 (no residual tumor cells, n=32) were excluded in the present study because no evaluation of post-treatment tumor cells could be performed. Both clinical and pathological data were thoroughly re-assessed and updated (recurrence data, last follow-up data) when needed. Study design is depicted in Figure 1.

Immunohistochemistry (IHC)

Pre-nCRT biopsies, already immunohistochemically stained for p53 and SOX2 as well as the hematoxylin-eosin stained slides from the paired resections were retrieved from the archives of the Department of Pathology at the Erasmus MC University Medical Center. From the most representative resection slide, 4 µm thick sections were cut from the corresponding block. From seven cases, material was missing. Both p53 and SOX2 immunohistochemical staining techniques have been previously described.[73, 74] In short, tissues were stained with the p53 ready to use kit (clone BP53-11, Ventana Medical Systems, Roche, Tuscon, AZ, USA) and SOX2 ready to use kit (clone Sp76, Cell Marque) using an automated slide staining system (Benchmark Ultra, Ventana Medical Systems, Roche, Tuscon, AZ, USA). Detection was performed with UltraView-DAB (ref. 760-500, Ventana Medical Systems) and amplification with Amplification Kit (ref. 760-080 Ventana Medical Systems). Next, the slides were counterstained with hematoxylin (ref. 790-2208, Ventana Medical Systems) and

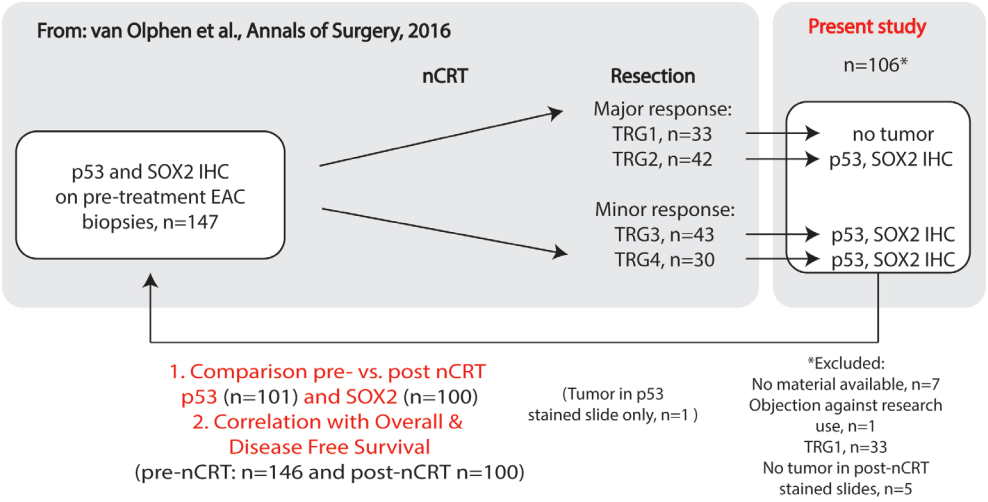


Figure 1. Study design. *nCRT, neoadjuvant chemoradiotherapy; TRG, tumor regression grade; EAC, esophageal adenocarcinoma; IHC, immunohistochemistry*

cover slipped. Each slide contained either embryonal cell carcinoma (for SOX2) or normal tonsil tissue (for p53) as a positive control.

Scoring

Scoring of pre-nCRT biopsies was already reported by Van Olphen *et al*., and was not repeated (raw data were used).[74] Scoring of post-nCRT resection specimen slides was performed in a similar way as previously has been described with either normal wild-type (WT) expression or aberrant p53 expression defined as either overexpression or complete loss of expression in >50% of tumor cells. Aberrant SOX2 expression was defined as absence of SOX2 expression in 50% of tumor cells (see Figure 2 for representative examples). Both investigators (LS, KB) were blinded for clinical data and pre-nCRT expression data. In case of disagreement the cases were reviewed simultaneously to reach a consensus diagnosis. One investigator (KB) also scored the pre-nCRT biopsies in the previous study.

Ethics

This study was approved by the institutional review board (medical ethical committee) from the Erasmus Medical Center (Rotterdam, The Netherlands).This study uses leftover patient material from regular patient diagnostics. The Ethics council of the Erasmus Medical Center approves research conducted on diagnostic tissue without special permission.

Statistical analysis

An interrater reliability analysis using the Cohen's Kappa statistic was performed to determine consistency among observers. The survival probability was estimated using Kaplan-Meier curves. A Cox proportional hazard model was used to estimate the association between both p53, SOX2 and survival. Only variables significant in univariable analysis were investigated further in multivariable analysis. Survival was defined as time from surgery until death (overall survival, OS) and time from surgery until first recurrence or death (disease free survival, DFS). P-values < 0.05 were considered statistically significant. Data were analyzed using SPSS statistical software (Version 25; SPSS, Chicago, Illinois, USA)

RESULTS

Baseline histopathological characteristics, other than immunohistochemical staining results are described in Supplementary Table 1. Paired samples of all cases with residual tumor in the post-nCRT esophageal resection specimen (i.e. TRG 2-4 cases, n=106) were reviewed. In five cases, no tumor cells could be identified in the post-nCRT stained slides. One case, with <10% remaining tumor cells after nCRT, i.e. TRG 2, only showed tumor cells in the slide stained with p53.

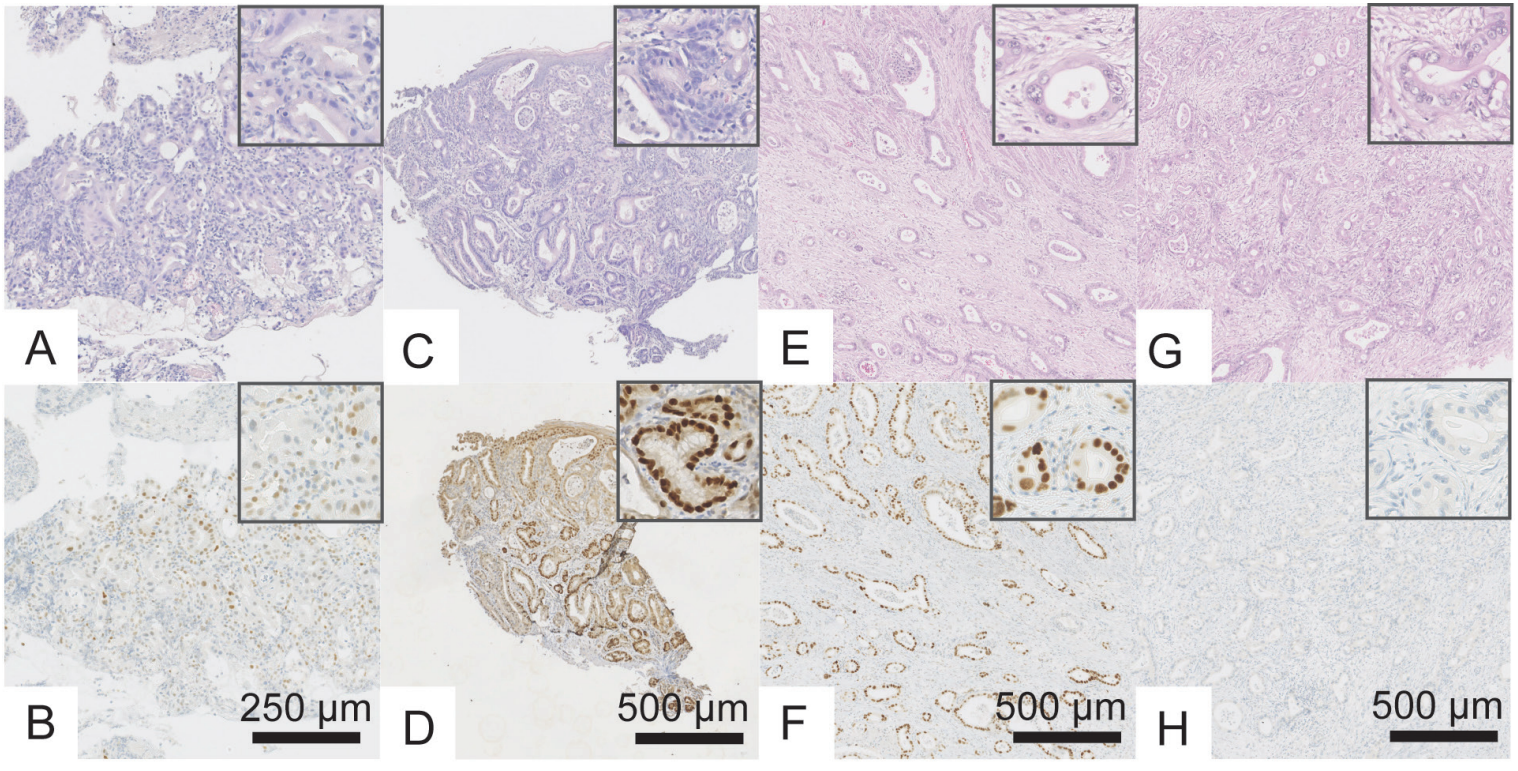


Figure 2. Representative example of discordant p53 and SOX2 expression patterns pre- and post- neoadjuvant chemoradiotherapy (nCRT) in a case with minor response (TRG 3-4). Wild-type p53 B) and wild-type SOX2 D) expression in the pre-nCRT endoscopic biopsy, whereas in the post-nCRT resection specimen, aberrant p53 F) and aberrant SOX2 H) in residual tumor is detected. A, C, E, G) corresponding hematoxylin & eosin-stained images.

Table 1. Comparison of p53 and SOX2 before and after neoadjuvant chemoradiotherapy (nCRT).

		post-nCRT (n, %)							
		p53			**SOX2**		**both**		**TRG1**
pre-nCRT (n, %)		WT	Loss	OE	WT	Loss	0/1 aberrant	2 aberrant	**(n=33)**
p53	WT	5 (28)	8 (44)	5 (28)	5 (28)	13 (72)	9 (50)	9 (50)	9
	Loss	2 (9)	18 (78)	3 (13)	4 (17)	19 (83)	5 (22)	18 (78)	2
	OE	6 (10)	1 (2)	53 (88)	10 (17)	49 (83)	15 (25)	44 (75)	22
SOX2	WT	6 (13)	13 (29)	26 (58)	14 (31)	31 (69)	18 (40)	27 (60)	9
	Loss	7 (13)	14 (25)	34 (62)	5 (9)	50 (91)	11 (20)	44 (80)	24
both	0 or 1	9 (16)	18 (33)	28 (51)	16 (29)	39 (71)	23 (42)	32 (58)	11
	2 aberrant	4 (9)	9 (20)	32 (71)	3 (7)	42 (93)	6 (13)	39 (87)	22

WT, wild-type; OE, overexpression; TRG, tumor regression grade.

p53 Immunohistochemistry (IHC)

In total, 75/83 (90%) cases with aberrant p53 IHC in pre-nCRT biopsy showed aberrant p53 IHC in its paired post-treatment samples. More specifically, 53/60 (88%) with pre-nCRT p53 overexpression and 18/23 (78%) with pre-nCRT loss of p53 remained concordant after nCRT (Table 1). In contrast, 13/18 (72%) with WT p53 staining in pre-nCRT samples showed aberrant staining pattern in paired post-nCRT samples. The interobserver agreement of post-treatment p53 was substantial with Kappa = 0.694 (95% CI 0.580-0.808).

SOX2 IHC

In 31/45 (69%) patients with WT SOX2 staining in pre-nCRT samples, paired post-nCRT samples showed SOX2 loss (Table 1). In contrast, 50/55 (91%) cases with SOX2 loss in pre-nCRT samples showed concordant SOX2 IHC in its paired post-nCRT samples. The interobserver agreement for post-nCRT SOX2 expression was moderate with Kappa = 0.493 (95% CI 0.416-0.644).

p53 and SOX2 combined analyses

Van Olphen *et al.* combined both markers and observed higher probability of a major response to nCRT in patients with aberrant p53 and loss of SOX2 expression in pre-nCRT samples.[74] In contrast, aberrant expression of both markers in post-nCRT samples was not associated with response to nCRT (Chi squared test, p=0.545; Supplementary Table 2). In

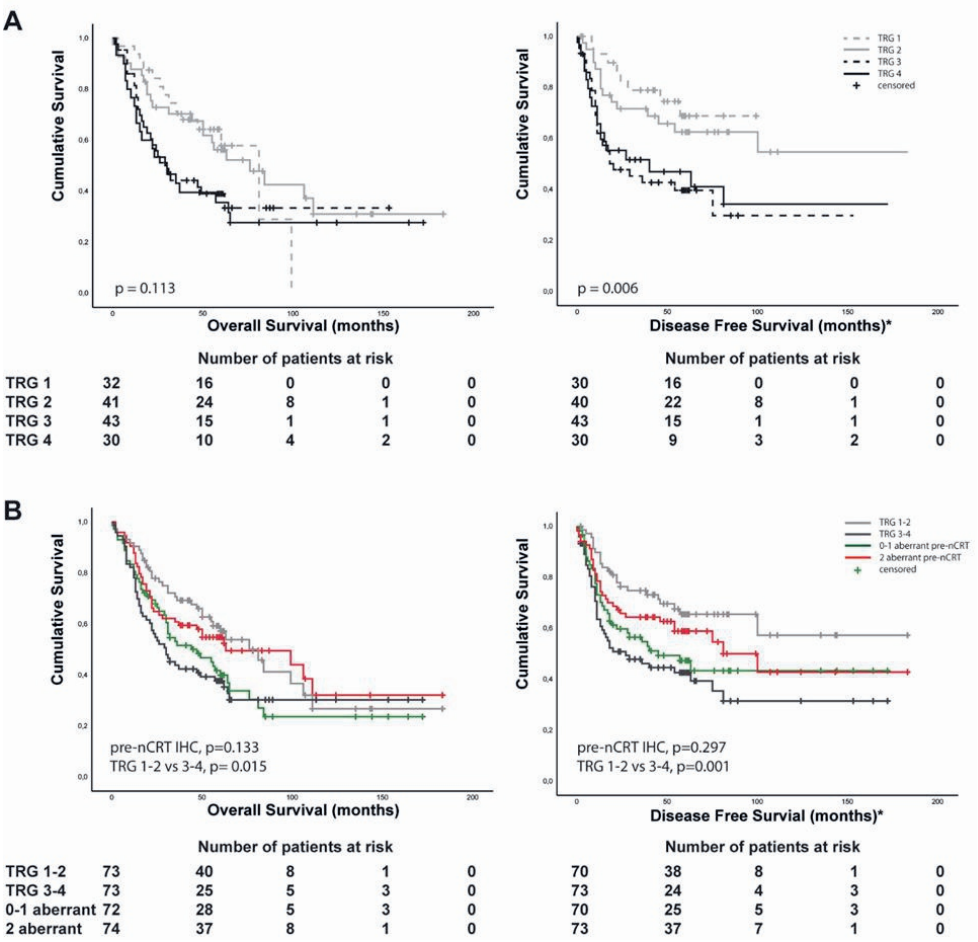
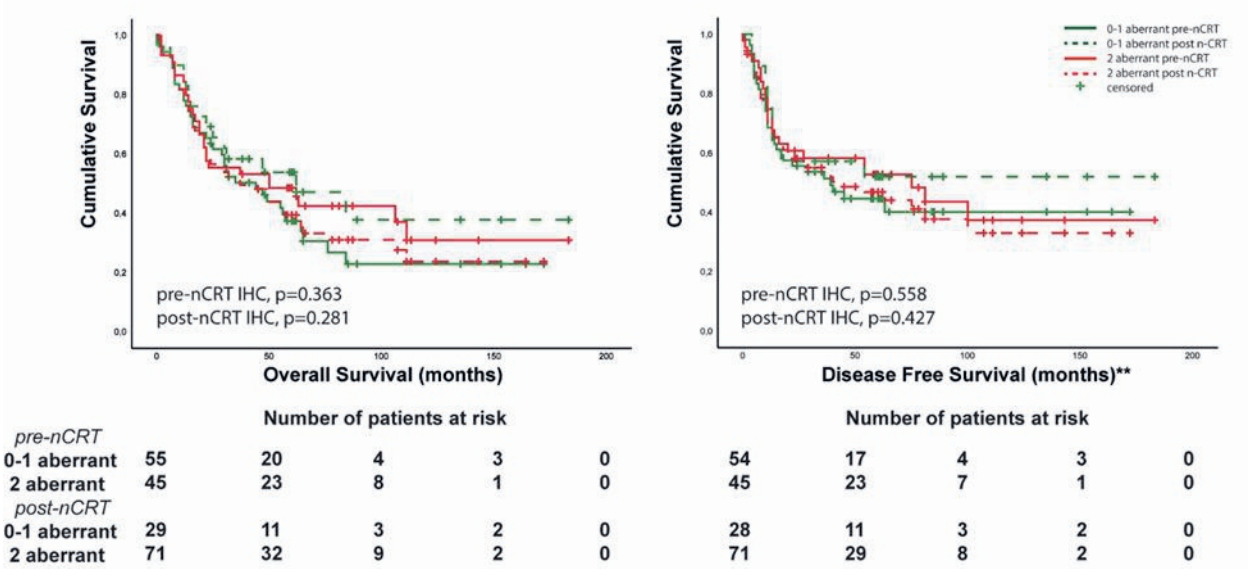


Figure 3. Response to neoadjuvant chemoradiotherapy (nCRT, i.e. tumor regression grade, TRG) clearly segregates patients (from entire cohort of 146 patients) into distinct prognostic groups for both A) overall (OS) and disease free survival (DFS). In B) response to nCRT is dichotomized into major (TRG1-2) and minor response (TRG3-4). Patients with aberrant expression of both p53 and SOX2 in their pre-treatment endoscopic biopsy (pre-nCRT IHC) have non-significantly better survival, compared with those with zero or one aberrant marker, in line with its association with major response. *Three cases had missing recurrence data.

39/45 (87%) cases with two aberrant markers in pre-nCRT samples, both markers were still aberrant post-nCRT. However, in cases with zero or one aberrant marker in pre-nCRT samples, both markers were aberrant in 32/55 (58%) cases post-nCRT.

Opposite prognostic value before and after nCRT

The degree of histopathological response to nCRT, i.e. TRG, shows a prognostic difference as depicted in Kaplan Meier curves (Figure 3). Patients with aberrant staining of both p53 and SOX2 in pre-nCRT samples showed a non-significantly improved overall and disease



Number of patients at risk										
pre-nCRT										
0-1 aberrant	55	20	4	3	0	54	17	4	3	0
2 aberrant	45	23	8	1	0	45	23	7	1	0
post-nCRT										
0-1 aberrant	29	11	3	2	0	28	11	3	2	0
2 aberrant	71	32	9	2	0	71	29	8	2	0

Figure 4. Kaplan-Meier curves stratified by immunohistochemistry (IHC) on endoscopic biopsy samples prior to neoadjuvant chemoradiotherapy (pre-nCRT) and on surgical resection specimens post-nCRT IHC (from 100 patients). Patients with aberrant expression of both p53 and SOX2 IHC in their post-nCRT resection samples (dashed lines, patients with TRG2-4 only) have adverse survival in comparison with patients with zero or one aberrant marker, in contrast to the results found in pre-nCRT biopsy samples (solid lines). However, both did not reach statistical significance. **One case had missing recurrence data.

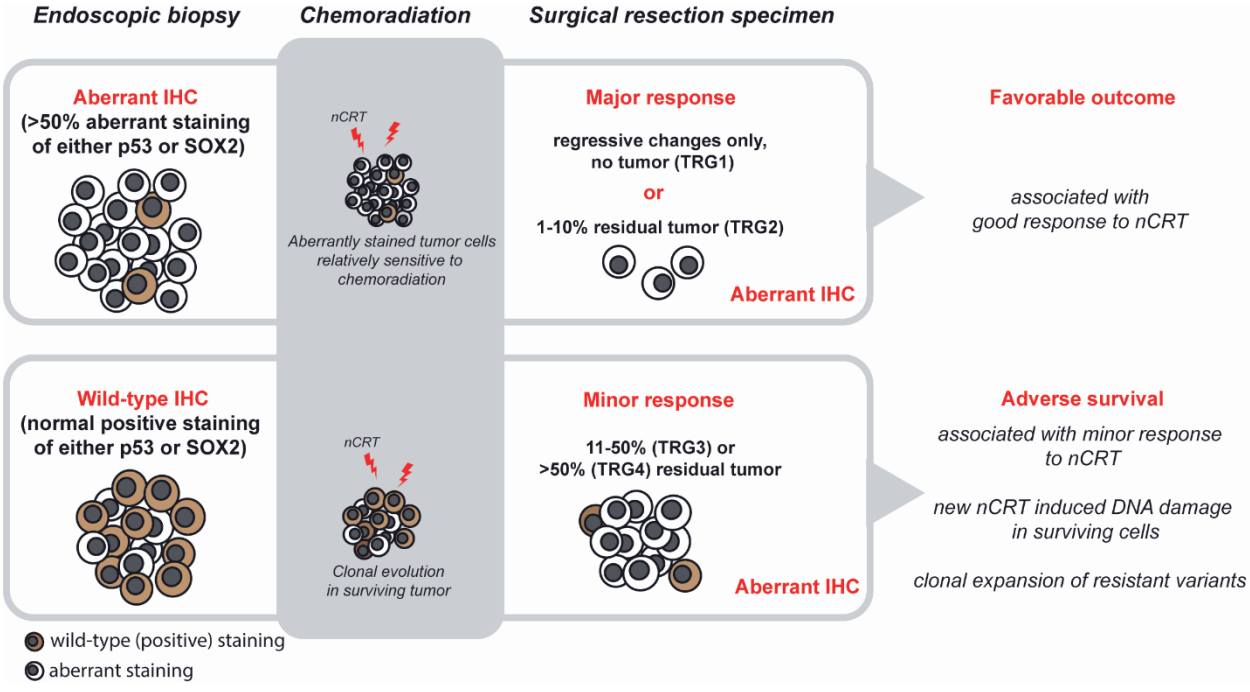


Figure 5. Representation and interpretation of the results found in the present study. Aberrant IHC patterns are largely unaffected by neoadjuvant chemoradiotherapy (nCRT), as the vast majority with aberrant SOX2 and p53 IHC in pre-nCRT biopsies showed concordant aberrant IHC in its paired post-nCRT samples (91 and 90% respectively). This is in contrast to the majority of patients with WT IHC in their biopsy samples showing aberrant staining pattern in paired post-nCRT samples (69 and 72%). Biopsies with (concurrent) aberrant p53 and SOX2 IHC were previously associated with a major response[74] and these patients showed associated favorable outcome in the present study. In contrast, those patients with WT IHC pattern on pre-treatment biopsies were associated with minor response[74] and often showed adverse survival. These results suggest that most tumor cells with an aberrant staining pattern on IHC are relatively sensitive to nCRT. In addition, surviving tumor cells may have mutated further resulting in cells with improved vitality and malignant potential. Furthermore, nCRT may also cause selective pressure with expansion of resistant variants.[207]

free survival compared to those with zero or one aberrant marker (Figure 3, C and D). In contrast, aberrant expression of both markers in the post-nCRT samples was associated with poor prognosis (Figure 4). Moreover, only differentiation grade, (clinical or pathological) N-stage and TRG were independently prognostic for DFS in multivariable COX regression analysis, whereas pathological N-stage was the only significant variable in multivariable analysis for OS (Supplementary Table 3).

DISCUSSION

In recent years, nCRT has become standard of care, significantly increasing survival in patients with advanced EAC. Yet, response to nCRT is highly variable with 23% of patients showing pathologically complete response, while survival rates remain moderate .[32, 43, 200] Biomarker studies aim to improve prognostication but are frequently based on historical cohorts with nCRT naïve patients (i.e. treated with surgery alone) as the effect of nCRT on biomarker expression has not been studied extensively yet.

The present study primarily compared two promising biomarkers in EAC, namely immunohistochemical p53 and SOX2 expression in paired pre- and post-nCRT EAC samples and studied its association with survival as a secondary outcome. Aberrant immunohistochemical expression of both p53 and SOX2 is largely concordant in paired samples before and after nCRT. Only one other study has compared immunohistochemical p53 staining patterns before and after therapy in patients with tubo-ovarian high-grade serous carcinoma and found similar results with high concordance of aberrant p53 before and after nCRT.[201] However, in the present study, when a WT pattern was found in pre-nCRT endoscopic biopsies, an aberrant pattern was frequently observed post-nCRT. These discordant IHC patterns may be the result of nCRT-induced DNA damage causing selective pressure of (new) mutated variants particularly in minor responders.[202]

Both pre- and post-nCRT p53 and SOX2 expression were not significantly associated with survival. Nevertheless, a trend could be observed in patients with aberrant staining of both p53 and SOX2 in the pre-nCRT samples performed better with regard to overall and disease free survival, compared to those with none or one aberrant marker. This is in line with the previously described association of concurrent aberrant expression with major response to nCRT.[74] Interestingly, an opposite trend could be observed when aberrant expression of both markers was observed in the post-nCRT samples, as these patients often showed adverse prognosis. These apparent discrepancies in survival may be the result of enrichment with minor responders (i.e. high volume of residual tumor) in our cohort, as cases without residual tumor could not be evaluated and were therefore excluded from the present study (selection bias). In addition, after nCRT, the vast majority of cases showed aberrant staining

patterns. A schematic representation and interpretation of the results found in this study and the previous study by van Olphen *et al.*[74] is depicted in Figure 5.

The management of patients with residual disease after nCRT is nowadays uncertain and consensus-based recommendations are lacking. If tumors with aberrant staining pattern indeed are more susceptible to chemoradiotherapy compared to those with WT staining, additional adjuvant therapy could be considered for patients showing WT before and an aberrant pattern after nCRT. Adjuvant therapy following nCRT has indeed been shown to improve survival in patients with locally advanced rectal cancer and is therefore increasingly being administered.[203-205] In addition, a potential survival benefit for adjuvant chemotherapy in EAC patients treated with nCRT followed by resection has been suggested in two retrospective analyses.[38, 206]

There are a number of limitations to the present study. Biopsies only represent a small fraction of the tumor, while EAC is known to be very heterogeneous at a molecular level.[175, 208] Therefore, sampling error must be taken into account. Nevertheless, this probably had little impact on the current study as high concordance was found before and after therapy in aberrantly stained cases. Yet, we did not perform sequencing in these particular cases, and although IHC is a good readout for at least TP53 mutational status, it is not known whether both similar staining patterns actually reflect the same mutation. Furthermore, despite automated staining and standardized protocols, IHC scoring remains subjective and a few residual negative tumor cells can easily be overlooked in a background of many regressive changes (i.e. fibrosis, inflammation). When no staining can be appreciated, albeit in the presence of your positive control, one should be sure there are actually negative tumor cells (i.e. indicating loss of p53 or SOX2) present and these cases required careful attention.

This is the first study comparing immunohistochemical p53 and SOX2 expression before and after nCRT in EAC. Despite the heterogeneous nature of EAC, both markers are largely concordant before and after nCRT, particularly when aberrant, which emphasizes the reliability of these diagnostic markers in small pre-treatment samples. This is important as treatment of patients with EAC is increasingly based on small biopsy samples as a result of which tumor from a surgical resection specimen may no longer be available. This is already the case in pathologically complete responders following nCRT. Our results also highlight the fact that results of biomarker studies regarding survival in nCRT naïve EAC patients cannot simply be extended to EAC patients treated with nCRT.

FUTURE PERSPECTIVE

It is important to realize that many studies on biomarkers in patients who were pre-treated with CRT suffer from selection bias. In such studies, it should be considered to evaluate the pre-treatment biopsies in patients with complete pathological response (i.e. no residual tumor), since these patients would otherwise be excluded. In vitro / in vivo (personalized) cancer models offer further options to study chemoresistance to therapy and selective pressure of cancer cells.

SUMMARY POINTS

- Aberrant expression of p53 and SOX2 staining in EAC tissue is largely unaffected by nCRT.
- These results emphasize the reliability of aberrant p53 and SOX2 expression in small pre-treatment samples, despite the heterogeneous nature of EAC.
- Conversely, the wild-type staining pattern of both markers frequently changed to aberrant expression after nCRT.
- Our findings support the model of enrichment of pre-existing mutant subclones after nCRT by selective pressure, particularly in minor responders.
- In the era of nCRT, treatment of patients with EAC is increasingly based on small biopsies; in pathologically complete responders, tumor from a surgical resection specimen is already no longer available.
- Results of biomarker studies regarding survival in nCRT-naïve EAC patients cannot simply be extended to EAC patients treated with nCRT.

SUPPLEMENTARY MATERIALS

Supplementary Table 1. Patient and tumor characteristics of present cohort and remaining cases with pre-treatment immunohistochemistry only.

Variable	Total cohort, n=146 (%)*	Cohort with pre- and post IHC, n=100 (%)*	Cases with pre-IHC only, n=46	*p -value*
Age				0.827
Median, years (IQR)	63 (56-70)	63 (56-69)	63 (56-70)	
Gender				0.859
Woman	18 (12)	12 (12)	40 (87)	
Male	128 (88)	88 (88)	6 (13)	
Grade				0.769
Well/Moderate	93 (64)	63 (63)	30 (65)	
Poor	53 (36)	37 (37)	16 (35)	
cT-stage				**0.034**
cT1	5 (3)	1 (1)	4 (9)	
cT2	29 (20)	18 (18)	11 (24)	
cT3	112 (77)	81 (81)	31 (67)	
cN-stage				0.550
cN0	46 (32)	31 (31)	15 (33)	
cN1	55 (38)	41 (41)	14 (30)	
cN2	41 (28)	25 (25)	16 (35)	
cN3	4 (3)	3 (3)	1 (2)	
ypT-stage				**<0.001**
ypT0	36 (25)	2 (2)	34 (74)	
ypT1	22 (15)	16 (16)	6 (13)	
ypT2	28 (19)	26 (26)	2 (4)	
ypT3	59 (40)	55 (55)	4 (9)	
ypT4	1 (1)	1 (1)	0 (0)	
ypN-stage				**<0.001**
ypN0	95 (65)	52 (52)	43 (93)	
ypN+	51 (35)	48 (48)	3 (7)	
Radicality				**0.014**
R0	134 (92)	88 (88)	46 (100)	
R1	12 (8)	12 (12)	0 (0)	
TRG				**<0.001**
TRG1	33 (23)	0 (0)	33 (72)	
TRG2	40 (28)	32 (32)	8 (17)	
TRG3	43 (30)	38 (38)	5 (11)	
TRG4	30 (21)	30 (30)	0 (0)	

Supplementary Table 2. Combined SOX2 and p53 immunohistochemical staining results before and after nCRT stratified by response to nCRT.

	IHC		TRG					TRG		
pre-nCRT (biopsy)	***p53***	***SOX2***	**1**	**2**	**3**	**4**	***Total (n)***	**1/2**	**3/4**	**p-value***
0 aberrant	WT	WT	0	0	4	5	9			
1 aberrant	Loss	WT	2	3	6	5	16	23 (31%)	49 (67%)	**<0.001**
	OE	WT	7	7	12	8	35			
	WT	Loss	2	2	8	1	13			
2 aberrant	Loss	Loss	7	10	2	2	21	50 (69%)	24 (33%)	
	OE	Loss	14	19	11	9	53			
	Total (n)		32	41	43	30	146	73	73	146
post-nCRT (resection)										
0 aberrant	WT	WT	-	1	1	1	3			
1 aberrant	Loss	WT	-	2	1	1	4	8** (25%)	21 (31%)	0.545
	OE	WT	-	2	7	3	12			
	WT	Loss	-	3	5	2	10			
2 aberrant	Loss	Loss	-	7	8	8	23	24** (75%)	47 (69%)	
	OE	Loss	-	17	16	15	48			
	Total (n)		-	32	38	30	100	32**	68	100

IHC, immunohistochemistry; TRG, tumor regression grade; nCRT, neoadjuvant chemoradiotherapy; WT, wild type; OE, overexpression. *Chi squared test ** TRG2 only

Supplementary Table 3. Both SOX2 and p53 expression, before or after treatment, do not predict survival in this cohort of EAC patients treated with neoadjuvant chemoradiotherapy (nCRT).

	Overall Survival				*Disease Free Survival*			
	Univariable		**Multivariable**		**Univariable**		**Multivariable**	
Variable	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
Age	1.2 (0.80-1.89)	0.346			1.0 (0.61-1.59)	0.938		
>= 63 y (<63 y = ref.)								
Gender	1.3 (0.61-2.60)	0.542			1.0 (0.46-2.03)	0.937		
Male (Female = ref.)								
Grade	1.7 (1.07-2.57)	**0.023**	1.5 (0.96-2.45)	0.077	1.9 (1.14-3.02)	**0.013**	1.8 (1.07-3.01)	**0.026**
Poor (Well/Moderate = ref.)								
cT-stage	2.1 (1.18-3.86)	**0.013**	1.7 (0.93-3.17)	0.086	2.4 (1.18-4.84)	**0.016**	1.8 (0.85-3.64)	0.131
cT3 (cT1/2 = ref.)								
cN-stage	1.6 (1.00-2.65)	**0.048**	1.4 (0.87-2.37)	0.157	2.2 (1.19-3.91)	**0.011**	2.0 (1.08-3.67)	**0.027**
cN+ (cN0 = ref.)								
ypT-stage	2.0 (1.25-3.16)	**0.004**	1.2 (0.69-2.05)	0.540	3.0 (1.68-5.33)	**<0.001**	1.5 (0.79-2.93)	0.212
ypT2-4 (ypT0-1 = ref)								
ypN-stage	2.2 (1.43-3.40)	**<0.001**	1.7 (1.04-2.64)	**0.033**	2.9 (1.78-4.71)	**<0.001**	1.9 (1.12-3.17)	**0.017**
ypN+ (ypN0 ref.)								
Radicality	1.5 (0.72-3.13)	0.275			1.9 (0.89-3.93)	0.099		
R1 (R0 ref.)								
TRG	1.7 (1.10-2.61)	**0.017**	1.5 (0.92-2.48)	0.103	2.4 (1.43-3.96)	**0.001**	2.1 (1.16-3.64)	**0.014**
TRG34 (TRG12 = ref.)								
pre-nCRT SOX2	0.7 (0.47-1.10)	0.131			0.8 (0.46-1.22)	0.243		
Loss (Wild-type = ref.)								

Continue

Continued

	Overall Survival				Disease Free Survival			
	Univariable		Multivariable		Univariable		Multivariable	
Variable	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
pre-nCRT p53	1.0 (0.55-1.88)	0.951			1.1 (0.54-2.19)	0.828		
Aberrant (Wild-type = ref.)								
Nr. of aberrant markers pre-nCRT	0.7 (0.47-1.11)	0.138			0.8 (0.48-1.26)	0.302		
2 markers (0/1 = ref.)								
post-nCRT SOX2 (n=100)	1.5 (0.74-2.86)	0.276			1.2 (0.60-2.52)	0.570		
Loss (Wild-type ref.)								
post-nCRT p53 (n=100)	1.3 (0.55-2.95)	0.582			1.4 (0.55-3.48)	0.497		
Aberrant (Wild-type ref.)								
Nr. of aberrant markers post-nCRT (n=100)	1.4 (0.77-2.44)	0.287			1.3 (0.69-2.40)	0.434		
2 markers (0/1 ref.)								

8

Discussion

In this thesis several biomarkers were studied as an attempt to improve risk stratification in patients with Barrett's esophagus (BE) and esophageal adenocarcinoma (EAC). These will be briefly discussed here in the context of the various key questions that formed the basis of this thesis.

Which patients with newly diagnosed BE are at increased risk for developing high-grade dysplasia (HGD) or EAC?

BE screening and surveillance programs aim to prevent the development of cancer or at least detect and treat EAC at an early stage with minimal burden and risks. Despite interobserver variation issues, these programs are primarily based on histological diagnosis of low-grade dysplasia (LGD) in endoscopic biopsy samples. Molecular and epigenetic changes can precede histological dysplastic changes, can be objectively measured, and may be used to predict progression in BE patients.[47-49]

MicroRNAs cannot be used to identify patients with non-dysplastic Barrett's esophagus at risk for progression to high-grade dysplasia (HGD) or EAC

MicroRNAs (miRs), small non-protein-coding RNAs, are gene regulatory molecules and influence the output of many protein-coding genes.[54] In **Chapter 3** we hypothesized miR profiles may be used to discriminate those patients at risk for progression by comparing these profiles in non-dysplastic BE (NDBE) tissue of endoscopic biopsy samples from progressors and non-progressors. However, patients who developed HGD or EAC after a long-term follow-up showed no significantly differential miR expression levels as patients who did not progress after a long-term follow-up. Interestingly, paired (i.e. from the same patient) NDBE and HGD/EAC samples, also did not show any significant differences. Only one previous study compared progressors with non-progressors, but our data could not confirm the miRs from this study as being significantly different in BE from progressors versus non-progressors.[59]

A frequent major limitation to studies regarding BE progression risk based on endoscopic biopsy material is the possibility of sampling error.[209] Moreover, sample size is limited by the low progression rate in NDBE patients.[13] Our study emphasizes the need for proper validation and, although promising results were found on miRs as biomarkers in testicular germ cell cancer and conjunctival melanoma, it hopefully will prevent further investigations on histology based miRs as a biomarker in NDBE.[210, 211]

Which patients with (early) EAC are at risk for developing (nodal) metastases?

Endoscopic mucosal resection (EMR) is recommended for patients with HGD or superficial EAC limited to the mucosa (pT1a). Endoscopic submucosal resection (ESR) is also increasingly used, but still debated, in tumors limited to the submucosal layer (pT1b).[100] Because no lymph nodes can be removed by this procedure, the risk of metastases has to be negligible.

To date, this risk is still difficult to determine. Depending on histopathological characteristics such as grade, presence of lymphovascular invasion, invasion depth and margin status of the EMR/ ESR, surgical resection usually follows endoscopic resection.[42]

Tumor budding (TB) according to the Ohike method and Olfactomedin 4 (OLFM4) are independent risk factors for metastases

TB is considered as a histological reflection of epithelial mesenchymal transition and of prognostic relevance, and increasingly also of clinical consequence, in early colorectal cancer (CRC).[100, 110, 123, 212, 213] Different methods to assess TB have been described. Therefore, a standardized method to assess and report TB was recommended following an International Tumor Budding Consensus Conference (ITBCC).[60] However, its relevance in pT1b EAC, is unclear. Moreover, studies comparing different TB methods are lacking.[63, 116] The study described in **Chapter 4** showed that TB assessed according to the Ohike method[114] is reproducible and informative. TB is prognostic for presence of metastases in pT1b EAC, independently from other previously described risk factors. Usage of pankeratin immunohistochemistry (IHC) showed similar prognostic value, increased identification of buds, but decreased interobserver agreement for all methods. Automated digitalized assessment had no added value. Our results were in line with the only previous study on TB in early (pT1a + pT1b) EAC using the same method.[64] The main difference between this method and the method recommended method by the ITBCC, which was not independently associated with metastases in our study, is the area size in which TB is assessed. In the recommended method, the total number of tumor buds are counted in a hotspot with most buds, whereas in the Ohike method, budding fields with at least five buds are counted along the entire invasive front. Furthermore, the recommendations from the ITBCC are solely based on studies in CRC. To our knowledge, the Ohike method has not been evaluated in CRC.

Another interesting biomarker, intestinal stem cell marker OLFM4 is described in **Chapter 5**. Patients with both advanced and early EAC were included and OLFM4 expression also associated with nodal metastases. Following the association with nodal status, patients whose tumors had high TB and low OLFM4 expression had adverse overall and disease free survival, although these results did not reach statistical significance. This may be explained by the relatively small sample size of the cohorts, especially in the early EAC group in the OLFM4 study, or due to the large proportion of deeply invading (pT34) tumors in the advanced group. In these patients, presence of occult distant metastases, that have spread hematogenously, is more likely and associated with poor survival when compared with early less invasive tumors.[214, 215] Other factors include the duration of follow-up, which may have been too short. Despite these limitations, this is the first extensive study on OLFM4 in EAC and similar results were found in a variety of other cancers, predominantly gastric adenocarcinoma.[68, 69, 72, 216]

Although more (prospective) research is required to evaluate the clinical impact of these new and promising biomarkers, we recommend to report TB according to the Ohike (H&E) method in the pathology report of patients with early pT1b EAC to guide clinical management, as it is a relatively simple method, does not require additional IHC and can be easily incorporated in daily diagnostic practice.

Can we further improve prediction of prognosis in patients with advanced EAC?

EAC survival rates vary depending on the cohort studied and whether or not they include a mixture of esophageal cancers subtypes, treatment regime and/or stages. Generally, once disease recurrence is diagnosed, the prognosis is poor. Furthermore, presence of metastases is the most important prognostic factor for adverse survival.

The most frequently mutated gene in human cancer, TP53, functions as a tumor suppressor. When mutated, the altered p53 protein is unable to trigger apoptosis in cells with damaged DNA, thereby losing its tumor-suppressing properties and facilitating uncontrolled cell proliferation. In EAC, TP53 gene mutation is considered a driver mutation, as it frequently occurs early in malignant progression.[11]

Aberrant p53 IHC is an independent risk factor for poor prognosis in EAC patients treated with surgery only

Being the most studied biomarker, many previous studies suggested the TP53 gene and p53 protein are linked to prognosis in EAC. However, the interpretation of p53 IHC has changed over time from dichotomous positive/negative to a three-tiered pattern based scoring: overexpression, loss of and heterogeneous, wild-type expression. Also, mixed patient cohorts have often been used, including patients with different types of esophageal cancer or treatment strategies amongst others. **Chapter 6** describes p53 IHC is prognostic for prognosis in a well-defined cohort of EAC patients. Furthermore, mutational analysis confirmed p53 IHC is an informative readout for TP53 mutational status in EAC with aberrant expression. Importantly, our study only included patients treated with surgery, as neo-adjuvant treatment may influence immunohistochemical biomarker expression. Yet, neo-adjuvant chemoradiotherapy (nCRT) is currently the gold standard in the treatment of patients with advanced EAC.

Although TNM classification remains the most important prognosticator, these studies provide further support for the presence of different molecular subtypes in EAC with distinct clinical outcomes.

Conclusions from biomarker studies in nCRT naïve EAC patients cannot simply be extended to patients treated with nCRT

In a first attempt to further address the influence of nCRT on immunohistochemical biomarker expression, immunohistochemical p53 and SOX2 expression before and after nCRT were compared in **Chapter 7**. Our results show expression of these markers is largely unaffected by nCRT, particularly when aberrant. In contrast, when pre-nCRT biopsies showed heterogeneous, wild-type staining, aberrant expression was frequently observed in the post-nCRT surgical resection specimen. In contrast to previous results found in naïve patients, (predominantly) treated with surgery only,[73, 75] these markers were not prognostic for survival in our cohort of patients who were all treated with nCRT. Moreover, aberrant expression of both markers was previously associated with major response (tumor regression grades, TRG 1-2) to nCRT.[74] Major response is indicative of improved survival opposed to those with minor response (TRG3-4). One explanation for these conflicting results is the unavoidable selection bias in our study, enriching minor responders with a poor prognosis as patients with major, complete response have no residual tumor cells. Hence, immunohistochemical staining for p53 or SOX2 in the post-nCRT resection specimen was not possible in cases with complete response (TRG1). Furthermore, nCRT may also cause selective pressure with expansion of resistant variants.[207] Thus, survival rates in heterogeneous patient cohorts should be interpreted with caution and biomarker results from naïve EAC patients cannot simply be extended to patients treated with nCRT.

9

Conclusions

Better identification of patients with Barrett's esophagus (BE) at risk of progression is essential to limit increases in healthcare costs and risk of both under- and overtreatment. We describe that the microRNA analyses performed in histologic non-dysplastic BE biopsy samples are of no value to predict BE progression and emphasize the need for proper validation of results. The impact of biopsy sampling errors on determining progression risk can be mostly circumvented by using new non-endoscopic methods. In addition, the majority of patients who may have a high progression risk, are not yet under surveillance and subsequently, are not biopsied at all, as prior BE diagnosis in reported in only 4.9% of patients with esophageal adenocarcinoma (EAC).[27] The greatest improvement in early detection of EAC can therefore be achieved by identifying these patients. In this regard, a new device called Cytosponge TM, which consists of a 'sponge on a string' pill, is one of the most promising methods.[217] When the pill is swallowed, it expands into a small rough textured sponge in the stomach, when pulled back up, it collects millions of cells lining the whole esophagus. Patients that test positive for the Trefoil Factor 3 (TFF3) marker protein are likely to have BE.[218] In addition, more and more research is focused on diagnosis and monitoring BE by means of so-called liquid biopsies (blood samples).[219, 220] These methods are less-invasive and cost-effective to diagnose and monitor people with BE and can be performed at a general practice level, instead of needing referral to a hospital for an endoscopy.[217, 221]

To date, histological diagnosis of dysplasia remains the most important prognosticator, for which problems with interobserver variation should be minimized by confirmation of the diagnosis by another, preferably expert, pathologist.[222] Yet, it is essential to obtain a more detailed understanding of the signaling networks involved in dysplastic progression to EAC. For that reason, it would be very interesting to further evaluate the expression of Olfactomedin 4 (OLFM4) in BE. At present, only limited research on OLFM4 in BE exists, but one previous study reported gradually increased expression during dysplastic progression in a small number of esophageal and gastric samples.[66, 68]

When EAC is diagnosed at an early stage, endoscopic resection is increasingly used. Apart from tumor grade, invasion depth and presence of lymphovascular invasion, tumor budding (TB) according to the Ohike method should be added in the pathology report from patients with early, pT1b EAC to guide treatment strategy and estimate whether the risk of metastases outweighs the risk of mortality and morbidity from additional surgery. Importantly, although associated, TB and tumor grade are not the same. Moreover, TB associates better with metastases than tumor grade. In EAC, grading is still based on the proportion of gland formation in the entire tumor. [113] In contrast, in CRC, grading should now be based on the least differentiated component according to most recent WHO criteria.[113] More importantly, the invasive front should not be taken into account when grading CRC, but this should be separately assessed for TB.[113] Grading EAC according to

these CRC criteria may further increase the prognostic relevance of tumor grade. OLFM4 expression also associated with nodal metastases but these results should be validated in larger cohorts.

Aberrant p53 expression is prognostic for adverse survival in patients treated with surgery alone (i.e. those who did not receive neo-adjuvant chemoradiotherapy, nCRT).[73] However, most patients with advanced EAC are now treated with nCRT prior to surgery. Therefore, we compared immunohistochemical expression of p53 and SOX2 before and after nCRT. Although a change from normal wild-type p53 and SOX2 expression to aberrant expression was frequently observed, aberrant expression of both markers was largely concordant before and after nCRT. Nevertheless, conclusions from biomarker studies in naïve patients cannot simply be extended to histological samples obtained from patients treated with nCRT. Still, these studies provide valuable, additional insight in the biological behavior of EAC. For example, our study on p53 in naïve EAC supports the presence of various molecular subtypes in EAC.[223]
This is important as current oncological practice is increasingly dependent on molecular diagnostics of tumor tissue, with brain tumors being the greatest example.[224] Furthermore, it can also be used to determine eligibility for personalized selection of cancer drugs based on the presence of actionable mutations. Some of the drugs like EGFR, BRAF, ALK, ROS1 and PARP inhibitors and their associated markers have only emerged in recent years, whereas immunohistochemically based hormone receptor status has been a mandatory part of a breast cancer diagnosis for decades, guiding the use of endocrine therapy.[225]

Recent work from The Cancer Genome Atlas indicate esophageal cancers indeed include multiple subtypes driven by distinct genetic alterations.[223] A striking similarity was found between genomic profiles of EAC and a subtype of stomach cancer, characterized by chromosomal instability, which suggests these cancers in fact are exactly the same cancer, but arose in different organs. Still, compared with other cancers (i.e. brain, breast amongst others), molecular diagnostics and targeted molecular based therapy in EAC is still in its infancy. Nevertheless, encouraging results have been reported. Evaluation of ERBB2 (human epidermal growth factor receptor 2 (HER2)) is now recommended and Trastuzemab (Herceptin) is now approved for the treatment of advanced or metastatic EAC.[113] Furthermore, alterations of the ERBB2 gene, present in one-third of EAC,[223] and the resulting overexpression of the ERBB2 (HER2) protein, has been shown to be predictive of response to this therapy.[113] Similarly, immune profile and immune checkpoint inhibitors, established therapeutic options for several cancer types, are increasingly evaluated in EAC, some of which with promising results.[226-228]

In conclusion, optimal treatment strategy in EAC remains a challenge and it is clear EAC pathogenesis is driven by multiple complex pathways. Analysis of one single marker

is unlikely to precisely predict progression, presence of metastases or recurrence in an individual patient; it is more likely that a combination of risk factors provides a better estimate.

10

Summary

The incidences of esophageal adenocarcinoma (EAC) and its precursor lesion, Barrett's esophagus (BE) are rapidly increasing. At an early stage, endoscopic treatment of EAC is possible with a relatively good chance of cure. In contrast, an advanced stage requires very invasive treatment, often includes neo-adjuvant chemoradiotherapy (nCRT) and holds a poor prognosis. To further improve risk stratification in these patients new and previously described biomarkers were evaluated in this thesis (Figure 1).

MicroRNA (miR) are small non-coding RNAs able to control translation of messenger RNA transcripts of protein-coding genes. This way, they can either function as oncogenes or tumor suppressors. High-throughput miR-profiling can differentiate between tissue types. Distinctive miR signatures are recognized for specific cancer types, both on tissue as well as body fluids. Consequently, miR are increasingly studied as potential cancer biomarkers, both regarding diagnosis as well as prognosis and prediction. Based on our results described in **Chapter 3**, miR profiles from endoscopic biopsy samples of patients with non-dysplastic BE cannot be used to predict progression to high-grade dysplasia and/or EAC.

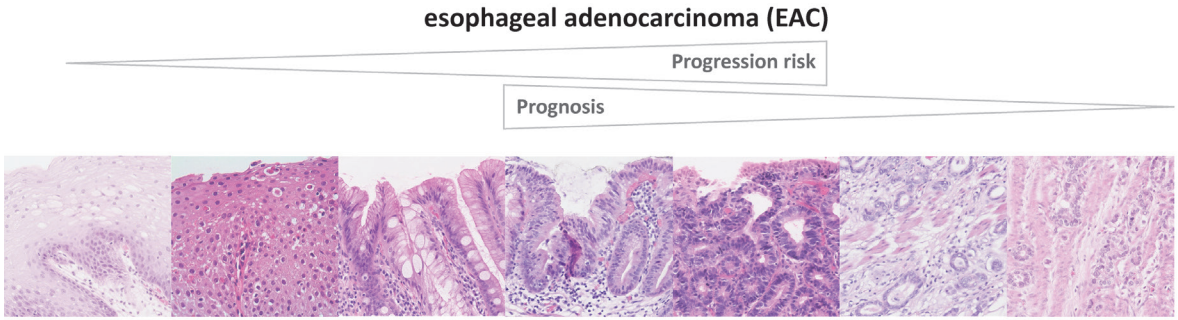
In patients with early, submucosal (pT1b) EAC, determining if the risk of metastases outweighs the risk of mortality and morbidity from surgical removal of the entire esophagus and part of the stomach is difficult. Frequently used risk indicators include lymphovascular invasion, tumor differentiation grade and depth of invasion. However, the estimated risk of metastases based on these indicators varies widely in the literature published. **Chapters 4 and 5** describe high tumor budding according to the Ohike method and low OLFM4 immunohistochemical expression are independent risk factors for (nodal) metastases. All these factors combined would likely lead to a better risk stratification for this group of patients. However, because of the limited studies related to these topics in EAC so far, additional research is needed to confirm these results.

Advanced EAC carries a poor prognosis, although reported survival rates vary greatly depending on the cohort studied. The *TP53* gene and its encoded protein p53, reported to be, at least one of, the guardian(s) of the genome, can be assumed to be the driving force of cancers of various types, including EAC. In spite of this knowledge, still a number of questions need to be answered. Aberrant p53 immunohistochemical staining pattern is an independent risk factor for poor prognosis in a well-defined cohort of EAC patients who were treated with surgery only (**Chapter 6**). Moreover, it correlates well with TP53 gene mutational status. Yet, neo-adjuvant chemoradiotherapy (nCRT) is currently the gold standard in the treatment of patients with advanced EAC. Moreover, it is currently being addressed if clinically complete responders after nCRT can be kept under close surveillance and only require surgery if evident localized tumor recurs. Therefore, treatment of patients with EAC is increasingly based on small biopsy samples and in a subset of patients, i.e.

patients with complete (clinical) response, a surgical tumor sample may even not be available anymore.

To date, the effect of nCRT on biomarker expression such as p53 and SOX2 in EAC and their prognostic value pre- and post-treatment are unknown. This is of relevance in the context of evaluation of these biomarkers in relation with clinical behavior. **Chapter 7** describes aberrant p53 and SOX2 IHC are largely concordant before and after nCRT. Nevertheless, results from biomarker studies in nCRT naïve EAC patients cannot simply be extended to patients treated with nCRT. This is most likely due to progression related events during the course of the disease and exposure to nCRT.

In conclusion, optimal treatment strategy in EAC remains a challenge and it is clear EAC pathogenesis is driven by multiple complex, and often interrelated, pathways. Although extern, prospective validation is desirable for most results, this thesis further emphasizes that analysis of one single biomarker is unlikely to precisely predict progression, presence of metastases, or recurrence in an individual patient. Hence, it is more likely that a combination of several risk factors, and as such also biomarkers, provides a better estimate. This thesis describes some additional promising markers that could be used for this purpose.



Chapter		Chapter 3			Chapters 4-5	Chapters 5-7
Question?		Who is at risk for progression?			Who is at risk for (nodal) metastases? Can we improve prognosis prediction?	
Biomarker studied?		miRNA			TB, OLFM4, p53, SOX2	
Method?		High-throughput miRNA profiling (histology, biopsy based)			H&E, IHC (conventional & digitilized) Automated digital TB count DNA sequencing, methylation analysis	
Conclusions?	miRNA in NDBE (biopsy) not useful to identify progressors				- invasion depth, LVI+, high TB, low OLFM4 are risk factors for metastases in pT1b EAC - aberrant p53 associates with adverse survival in naïve EAC patients - these results cannot simply be extended to patients treated with nCRT	
Future prospects?		↓ sampling error, ↑ sample size CytospongeTM, liquid biopsy			(prospective) external validation of risk factors molecular diagnostics	

Figure 1. Outline and conclusions of this thesis. *GERD, gastro-esophageal reflux disease; NDBE, non-dysplastic Barrett's esophagus; LGD, low-grade dysplasia; HGD, high-grade dysplasia; miRNA, microRNA; TB, tumor budding; H&E, hematoxylin & eosin; IHC, immunohistochemistry; LVI, lymphovascular invasion; pT1b, limited to the submucosa; nCRT, neo-adjuvant chemoradiotherapy*

11

Samenvatting (NL)

Het aantal patiënten dat jaarlijks gediagnosticeerd wordt met een tumor van het type adenocarcinoom in de slokdarm (EAC) en de voorloperlaesie hiervan, Barrett slokdarm (BE) neemt snel toe. In een vroeg stadium is endoscopische behandeling van EAC nog mogelijk. Hierbij wordt via de mond een buigzame slang met camera ingebracht en doorgeschoven tot in de slokdarm. Het kwaadaardige weefsel kan met behulp van deze slang verwijderd worden. Wanneer de tumor vroeg is ontdekt, heeft men een relatief goede kans op genezing. Daarentegen vereist een verder gevorderd stadium een zeer invasieve behandeling. Chirurgische, complete verwijdering van de slokdarm met hierbij vaak ook een stuk van de maag en omgevend weefsel met hierin lymfeklieren is dan nodig, omdat deze mogelijk uitzaaiingen van de tumor kunnen bevatten. Hiernaast is vaak voorbehandeling nodig, middels een combinatie van chemotherapie en lokale bestraling (nCRT). Een laat ontdekt EAC heeft vaak een slechte prognose. Om de verschillende risico's en de prognose voor deze patiënten beter te kunnen inschatten, werden in dit proefschrift nieuwe en eerder beschreven biomarkers geëvalueerd, zodat in de toekomst voor elke patiënt de meest optimale behandelingsstrategie kan worden gekozen.

MicroRNA's (miR) zijn kleine niet-coderende RNAs die de translatie van messenger-RNA-transcripten van eiwit coderende genen controleren of degradatie kunnen veroorzaken. Op deze manier kunnen ze ofwel als oncogenen (d.w.z. genen die kanker kunnen veroorzaken) ofwel als tumoronderdrukkers fungeren. Met behulp van miR-profilering kan men onderscheid maken tussen weefseltypen en kunnen zelfs onderscheidende miR-profielen worden herkend voor specifieke kankertypes. Dit kan zowel in weefsel als in lichaamsvloeistoffen. Om deze reden worden miR in toenemende mate bestudeerd als biomarkers. Op basis van onze resultaten beschreven in **Hoofdstuk 3**, kunnen miR profielen van endoscopisch verkregen biopten van patiënten met niet-dysplastische BE echter jammer genoeg niet worden gebruikt om de progressie naar hooggradige dysplasie of EAC te voorspellen.

Bij patiënten met vroege EAC, maximaal reikend tot in de submucosa (pT1b), is het moeilijk in te schatten of het risico op uitzaaiingen opweegt tegen het risico op sterfte en morbiditeit veroorzaakt door chirurgische verwijdering van de gehele slokdarm, een deel van de maag en de regionale lymfeklieren. Veelgebruikte risico-indicatoren zijn onder meer aanwezigheid van tumorcellen in (lymfe)vaten, differentiatiegraad en de diepste tumordoorgroei. Het geschatte risico op metastasen op basis van deze indicatoren varieert echter sterk in de tot nu toe gepubliceerde literatuur. In de **Hoofdstukken 4 en 5** wordt beschreven dat aanwezigheid van veel tumor buds, gescoord volgens de Ohike-methode en immunohistochemische (IHC) lage OLFM4- expressie, in deze groep patiënten met pT1b tumoren, ook onafhankelijke risicofactoren voor uitzaaiingen zijn. Al deze factoren gecombineerd zullen hoogstwaarschijnlijk leiden tot een betere risicostratificatie voor

deze groep patiënten. Er is hierover voor EAC echter nog te weinig bekend in de literatuur waardoor aanvullend onderzoek nodig is om deze resultaten te bevestigen.

In een later, vergevorderd stadium heeft EAC een erg slechte prognose, hoewel de gerapporteerde overlevingskansen sterk variëren afhankelijk van het bestudeerde patiënt cohort. Het TP53-gen en zijn gecodeerde eiwit p53, door sommigen ook wel, tenminste één van, de bewaker(s) van het genoom genoemd, kan, wanneer deze gemuteerd is, worden beschouwd als de drijvende kracht achter verschillende soorten kanker, waaronder EAC. Desondanks moeten nog een aantal vragen worden beantwoord. Een afwijkend immunohistochemisch aankleuringspatroon van het tumor suppressor eiwit p53 blijkt een onafhankelijke risicofactor voor slechte prognose in een goed gedefinieerd cohort van EAC-patiënten die alleen met chirurgie werden behandeld (**Hoofdstuk 6**). Bovendien correleert het aankleuringspatroon goed met de mutatiestatus van het TP53-gen. Momenteel worden de meeste patiënten met geavanceerde EAC echter behandeld met nCRT en chirurgie. Bovendien wordt momenteel onderzocht of patiënten met een klinisch complete respons, waarbij er dus radiologisch en bij endoscopie na afname van biopten geen tumor meer aanwezig is na nCRT, onder nauwlettend toezicht kunnen worden gehouden en alleen kunnen worden geopereerd als er opnieuw aanwijzingen zijn dat er tumor aanwezig is. Dit betekent dat de behandeling van patiënten met EAC in toenemende mate gebaseerd is op kleine hoeveelheden weefsel (biopten). Bij een subgroep van patiënten, wederom de patiënten met een volledige (klinische) respons, is het daarom goed mogelijk dat een chirurgisch verkregen tumormonster nooit meer beschikbaar komt.

Tot op heden is het effect van nCRT op de expressie van biomarkers zoals p53 en SOX2 in EAC en hun prognostische waarde voor en na de behandeling onbekend. Dit is relevant in de context van het evalueren van deze biomarkers in relatie tot klinisch gedrag. **Hoofdstuk 7** beschrijft dat afwijkende p53 en SOX2 IHC grotendeels hetzelfde is voor en na behandeling met nCRT. Toch kunnen resultaten van biomarker onderzoek in EAC patiënten zonder nCRT niet zomaar worden vergeleken met de resultaten in patiënten met EAC die wel zijn behandeld met nCRT. Dit is hoogstwaarschijnlijk te wijten aan progressie gerelateerde gebeurtenissen in de tumor gedurende de loop van de ziekte en de blootstelling van de tumor aan nCRT.

Al met al blijft het kiezen van de optimale behandelingsstrategie bij patiënten met EAC dus een uitdaging. Het is duidelijk dat het ontstaan en de progressie van EAC erg complex is en tussen patiënten sterk kan verschillen. Hoewel externe, prospectieve validatie voor de meeste resultaten wenselijk is, benadrukt dit proefschrift verder dat de analyse van slechts één enkele biomarker onwaarschijnlijk is om de progressie, aanwezigheid van metastasen of kans op een tumor recidief bij een individuele patiënt te voorspellen. Het is veel waarschijnlijker dat een combinatie van meerdere risicofactoren, ofwel biomarkers, een

betere schatting zal geven. Dit proefschrift beschrijft enkele (aanvullende) veelbelovende biomarkers die voor dit doel kunnen worden gebruikt.

References

1. Shaheen NJ, Falk GW, Iyer PG, Gerson LB, American College of G. Acg clinical guideline: Diagnosis and management of barrett's esophagus. *Am J Gastroenterol* 2016;**111**;30-50; quiz 51.
2. Smith J, Garcia A, Zhang R, DeMeester S, Vallone J, Chandrasoma P. Intestinal metaplasia is present in most if not all patients who have undergone endoscopic mucosal resection for esophageal adenocarcinoma. *Am J Surg Pathol* 2016;**40**;537-543.
3. Runge TM, Abrams JA, Shaheen NJ. Epidemiology of barrett's esophagus and esophageal adenocarcinoma. *Gastroenterol Clin North Am* 2015;**44**;203-231.
4. El-Serag HB, Sweet S, Winchester CC, Dent J. Update on the epidemiology of gastro-oesophageal reflux disease: A systematic review. *Gut* 2014;**63**;871-880.
5. Ronkainen J, Aro P, Storskrubb T *et al.* Prevalence of barrett's esophagus in the general population: An endoscopic study. *Gastroenterology* 2005;**129**;1825-1831.
6. van Soest EM, Dieleman JP, Siersema PD, Sturkenboom MC, Kuipers EJ. Increasing incidence of barrett's oesophagus in the general population. *Gut* 2005;**54**;1062-1066.
7. Ward EM, Wolfsen HC, Achem SR *et al.* Barrett's esophagus is common in older men and women undergoing screening colonoscopy regardless of reflux symptoms. *Am J Gastroenterol* 2006;**101**;12-17.
8. Gerson LB, Shetler K, Triadafilopoulos G. Prevalence of barrett's esophagus in asymptomatic individuals. *Gastroenterology* 2002;**123**;461-467.
9. Arora Z, Garber A, Thota PN. Risk factors for barrett's esophagus. *J Dig Dis* 2016;**17**;215-221.
10. Contino G, Vaughan TL, Whiteman D, Fitzgerald RC. The evolving genomic landscape of barrett's esophagus and esophageal adenocarcinoma. *Gastroenterology* 2017;**153**;657-673 e651.
11. Stachler MD, Taylor-Weiner A, Peng S *et al.* Paired exome analysis of barrett's esophagus and adenocarcinoma. *Nat Genet* 2015;**47**;1047-1055.
12. Hvid-Jensen F, Pedersen L, Drewes AM, Sorensen HT, Funch-Jensen P. Incidence of adenocarcinoma among patients with barrett's esophagus. *N Engl J Med* 2011;**365**;1375-1383.
13. Desai TK, Krishnan K, Samala N *et al.* The incidence of oesophageal adenocarcinoma in non-dysplastic barrett's oesophagus: A meta-analysis. *Gut* 2012;**61**;970-976.
14. Dixon MF. Gastrointestinal epithelial neoplasia: Vienna revisited. *Gut* 2002;**51**;130-131.
15. Schlemper RJ, Riddell RH, Kato Y *et al.* The vienna classification of gastrointestinal epithelial neoplasia. *Gut* 2000;**47**;251-255.
16. Kaye PV, Haider SA, Ilyas M *et al.* Barrett's dysplasia and the vienna classification: Reproducibility, prediction of progression and impact of consensus reporting and p53 immunohistochemistry. *Histopathology* 2009;**54**;699-712.
17. Vennalaganti P, Kanakadandi V, Goldblum JR *et al.* Discordance among pathologists in the united states and europe in diagnosis of low-grade dysplasia for patients with barrett's esophagus. *Gastroenterology* 2017;**152**;564-570 e564.
18. Krishnamoorthi R, Lewis JT, Krishna M *et al.* Predictors of progression in barrett's esophagus with low-grade dysplasia: Results from a multicenter prospective be registry. *Am J Gastroenterol* 2017;**112**;867-873.

19. Waters KM, Salimian KJ, Voltaggio L, Montgomery EA. Refined criteria for separating low-grade dysplasia and nondysplastic barrett esophagus reduce equivocal diagnoses and improve prediction of patient outcome: A 10-year review. *Am J Surg Pathol* 2018;**42**;1723-1729.
20. Ten Kate FJC, Nieboer D, Ten Kate FJW *et al.* Improved progression prediction in barrett's esophagus with low-grade dysplasia using specific histologic criteria. *Am J Surg Pathol* 2018;**42**;918-926.
21. Zhu W, Appelman HD, Greenson JK *et al.* A histologically defined subset of high-grade dysplasia in barrett mucosa is predictive of associated carcinoma. *Am J Clin Pathol* 2009;**132**;94-100.
22. Konda VJ, Ross AS, Ferguson MK *et al.* Is the risk of concomitant invasive esophageal cancer in high-grade dysplasia in barrett's esophagus overestimated? *Clin Gastroenterol Hepatol* 2008;**6**;159-164.
23. Richtlijnen barrett oesofagus. Barrett Expert Centra, 2018.
24. Weusten B, Bisschops R, Coron E *et al.* Endoscopic management of barrett's esophagus: European society of gastrointestinal endoscopy (esge) position statement. *Endoscopy* 2017;**49**;191-198.
25. Fitzgerald RC, di Pietro M, Ragunath K *et al.* British society of gastroenterology guidelines on the diagnosis and management of barrett's oesophagus. *Gut* 2014;**63**;7-42.
26. di Pietro M, Fitzgerald RC, group BSGBsgw. Revised british society of gastroenterology recommendation on the diagnosis and management of barrett's oesophagus with low-grade dysplasia. *Gut* 2018;**67**;392-393.
27. Wenker TN, Tan MC, Liu Y, El-Serag HB, Thrift AP. Prior diagnosis of barrett's esophagus is infrequent, but associated with improved esophageal adenocarcinoma survival. *Dig Dis Sci* 2018;**63**;3112-3119.
28. Sikkema M, de Jonge PJ, Steyerberg EW, Kuipers EJ. Risk of esophageal adenocarcinoma and mortality in patients with barrett's esophagus: A systematic review and meta-analysis. *Clin Gastroenterol Hepatol* 2010;**8**;235-244; quiz e232.
29. Bergeron EJ, Lin J, Chang AC, Orringer MB, Reddy RM. Endoscopic ultrasound is inadequate to determine which t1/t2 esophageal tumors are candidates for endoluminal therapies. *J Thorac Cardiovasc Surg* 2014;**147**;765-771: Discussion 771-763.
30. Betancourt Cuellar SL, Sabloff B, Carter BW *et al.* Early clinical esophageal adenocarcinoma (ct1): Utility of ct in regional nodal metastasis detection and can the clinical accuracy be improved? *Eur J Radiol* 2017;**88**;56-60.
31. Grotenhuis BA, Wijnhoven BP, Poley JW *et al.* Preoperative assessment of tumor location and station-specific lymph node status in patients with adenocarcinoma of the gastroesophageal junction. *World J Surg* 2013;**37**;147-155.
32. Shapiro J, van Lanschot JJ, Hulshof MC *et al.* Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (cross): Long-term results of a randomised controlled trial. *Lancet Oncol* 2015;**16**;1090-1098.
33. Noordman BJ, Spaander MCW, Valkema R *et al.* Detection of residual disease after neoadjuvant chemoradiotherapy for oesophageal cancer (presano): A prospective multicentre, diagnostic cohort study. *Lancet Oncol* 2018;**19**;965-974.

34. Bertero L, Massa F, Metovic J *et al.* Eighth edition of the uicc classification of malignant tumours: An overview of the changes in the pathological tnm classification criteria-what has changed and why? *Virchows Arch* 2018;**472**;519-531.
35. Becker K, Mueller JD, Schulmacher C *et al.* Histomorphology and grading of regression in gastric carcinoma treated with neoadjuvant chemotherapy. *Cancer* 2003;**98**;1521-1530.
36. Mandard AM, Dalibard F, Mandard JC *et al.* Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. Clinicopathologic correlations. *Cancer* 1994;**73**;2680-2686.
37. Thies S, Langer R. Tumor regression grading of gastrointestinal carcinomas after neoadjuvant treatment. *Front Oncol* 2013;**3**;262.
38. Burt BM, Groth SS, Sada YH *et al.* Utility of adjuvant chemotherapy after neoadjuvant chemoradiation and esophagectomy for esophageal cancer. *Ann Surg* 2017;**266**;297-304.
39. Drake J, Tauer K, Portnoy D, Weksler B. Adjuvant chemotherapy is associated with improved survival in patients with nodal metastases after neoadjuvant therapy and esophagectomy. *J Thorac Dis* 2019;**11**;2546-2554.
40. Greally M, Ku GY. Adjuvant chemotherapy for poor pathologic response after pre-operative chemoradiation in esophageal cancer: Infeasible and illogical. *J Thorac Dis* 2019;**11**;S1855-S1860.
41. Leggett CL, Lewis JT, Wu TT *et al.* Clinical and histologic determinants of mortality for patients with barrett's esophagus-related t1 esophageal adenocarcinoma. *Clin Gastroenterol Hepatol* 2015;**13**;658-664 e651-653.
42. Newton AD, Predina JD, Xia L *et al.* Surgical management of early-stage esophageal adenocarcinoma based on lymph node metastasis risk. *Ann Surg Oncol* 2018;**25**;318-325.
43. Jemal A, Ward EM, Johnson CJ *et al.* Annual report to the nation on the status of cancer, 1975-2014, featuring survival. *J Natl Cancer Inst* 2017;**109**.
44. Biomarkers. Springer Nature 2019.
45. Snyder P, Dunbar K, Cipher DJ, Souza RF, Spechler SJ, Konda VJA. Aberrant p53 immunostaining in barrett's esophagus predicts neoplastic progression: Systematic review and meta-analyses. *Dig Dis Sci* 2019;**64**;1089-1097.
46. Kastelein F, Biermann K, Steyerberg EW *et al.* Aberrant p53 protein expression is associated with an increased risk of neoplastic progression in patients with barrett's oesophagus. *Gut* 2013;**62**;1676-1683.
47. Stachler MD, Camarda ND, Deitrick C *et al.* Detection of mutations in barrett's esophagus before progression to high-grade dysplasia or adenocarcinoma. *Gastroenterology* 2018;**155**;156-167.
48. Chettouh H, Mowforth O, Galeano-Dalmau N *et al.* Methylation panel is a diagnostic biomarker for barrett's oesophagus in endoscopic biopsies and non-endoscopic cytology specimens. *Gut* 2018;**67**;1942-1949.
49. Dilworth MP, Nieto T, Stockton JD *et al.* Whole genome methylation analysis of nondysplastic barrett esophagus that progresses to invasive cancer. *Ann Surg* 2019;**269**;479-485.
50. Gregson EM, Fitzgerald RC. Biomarkers for dysplastic barrett's: Ready for prime time? *World J Surg* 2015;**39**;568-577.

51. Janmaat VT, van Olphen SH, Biermann KE, Looijenga LHJ, Bruno MB, Spaander MCW. Use of immunohistochemical biomarkers as independent predictor of neoplastic progression in barrett's oesophagus surveillance: A systematic review and meta-analysis. *PLoS One* 2017;**12**;e0186305.
52. Konda VJA, Souza RF. Biomarkers of barrett's esophagus: From the laboratory to clinical practice. *Dig Dis Sci* 2018;**63**;2070-2080.
53. Qureshi AP, Stachler MD, Haque O, Odze RD. Biomarkers for barrett's esophagus–a contemporary review. *Expert Rev Mol Diagn* 2018;**18**;939-946.
54. Esquela-Kerscher A, Slack FJ. Oncomirs–microrna with a role in cancer. *Nat Rev Cancer* 2006;**6**;259-269.
55. Lu J, Getz G, Miska EA *et al.* Microrna expression profiles classify human cancers. *Nature* 2005;**435**;834-838.
56. Clark RJ, Craig MP, Agrawal S, Kadakia M. Microrna involvement in the onset and progression of barrett's esophagus: A systematic review. *Oncotarget* 2018;**9**;8179-8196.
57. Kong X, Gong S, Su L, Li C, Kong Y. Expression signatures and roles of micrornas in human oesophageal adenocarcinomas. *J Cell Mol Med* 2018;**22**;123-130.
58. Mallick R, Patnaik SK, Wani S, Bansal A. A systematic review of esophageal microrna markers for diagnosis and monitoring of barrett's esophagus. *Dig Dis Sci* 2016;**61**;1039-1050.
59. Revilla-Nuin B, Parrilla P, Lozano JJ *et al.* Predictive value of micrornas in the progression of barrett esophagus to adenocarcinoma in a long-term follow-up study. *Ann Surg* 2013;**257**;886-893.
60. Lugli A, Kirsch R, Ajioka Y *et al.* Recommendations for reporting tumor budding in colorectal cancer based on the international tumor budding consensus conference (itbcc) 2016. *Mod Pathol* 2017;**30**;1299-1311.
61. Davison JM, Landau MS, Luketich JD *et al.* A model based on pathologic features of superficial esophageal adenocarcinoma complements clinical node staging in determining risk of metastasis to lymph nodes. *Clin Gastroenterol Hepatol* 2016;**14**;369-377 e363.
62. Manner H, Wetzka J, May A *et al.* Early-stage adenocarcinoma of the esophagus with mid to deep submucosal invasion (pt1b sm2-3): The frequency of lymph-node metastasis depends on macroscopic and histological risk patterns. *Dis Esophagus* 2017;**30**;1-11.
63. Berg KB, Schaeffer DF. Tumor budding as a standardized parameter in gastrointestinal carcinomas: More than just the colon. *Mod Pathol* 2018;**31**;862-872.
64. Landau MS, Hastings SM, Foxwell TJ, Luketich JD, Nason KS, Davison JM. Tumor budding is associated with an increased risk of lymph node metastasis and poor prognosis in superficial esophageal adenocarcinoma. *Mod Pathol* 2014;**27**;1578-1589.
65. Thies S, Guldener L, Slotta-Huspenina J *et al.* Impact of peritumoral and intratumoral budding in esophageal adenocarcinomas. *Hum Pathol* 2016;**52**;1-8.
66. Jang BG, Lee BL, Kim WH. Intestinal stem cell markers in the intestinal metaplasia of stomach and barrett's esophagus. *PLoS One* 2015;**10**;e0127300.
67. Duan C, Liu X, Liang S *et al.* Oestrogen receptor-mediated expression of olfactomedin 4 regulates the progression of endometrial adenocarcinoma. *J Cell Mol Med* 2014;**18**;863-874.

68. Jang BG, Lee BL, Kim WH. Olfactomedin-related proteins 4 (olfm4) expression is involved in early gastric carcinogenesis and of prognostic significance in advanced gastric cancer. *Virchows Arch* 2015;**467**;285-294.
69. Li H, Rodriguez-Canales J, Liu W *et al.* Deletion of the olfactomedin 4 gene is associated with progression of human prostate cancer. *Am J Pathol* 2013;**183**;1329-1338.
70. Liu C, Guo Y, Wu W *et al.* Plasma olfactomedin 4 level in peripheral blood and its association with clinical features of breast cancer. *Oncol Lett* 2017;**14**;8106-8113.
71. Liu W, Liu Y, Zhu J, Wright E, Ding I, Rodgers GP. Reduced hgc-1 protein expression is associated with malignant progression of colon carcinoma. *Clin Cancer Res* 2008;**14**;1041-1049.
72. Zhao J, Shu P, Duan F *et al.* Loss of olfm4 promotes tumor migration through inducing interleukin-8 expression and predicts lymph node metastasis in early gastric cancer. *Oncogenesis* 2016;**5**;e234.
73. Ten Kate FJC, Suzuki L, Dorssers LCJ *et al.* Pattern of p53 protein expression is predictive for survival in chemoradiotherapy-naive esophageal adenocarcinoma. *Oncotarget* 2017;**8**;104123-104135.
74. van Olphen SH, Biermann K, Shapiro J *et al.* P53 and sox2 protein expression predicts esophageal adenocarcinoma in response to neoadjuvant chemoradiotherapy. *Ann Surg* 2016.
75. Ten Kate FJC, van Olphen SH, Bruno MJ *et al.* Loss of sry-box2 (sox2) expression and its impact on survival of patients with oesophageal adenocarcinoma. *Br J Surg* 2017;**104**;1327-1337.
76. van Olphen S, Biermann K, Spaander MC *et al.* Sox2 as a novel marker to predict neoplastic progression in barrett's esophagus. *Am J Gastroenterol* 2015;**110**;1420-1428.
77. Lagergren J. Adenocarcinoma of oesophagus: What exactly is the size of the problem and who is at risk? *Gut* 2005;**54 Suppl 1**;i1-5.
78. Hayeck TJ, Kong CY, Spechler SJ, Gazelle GS, Hur C. The prevalence of barrett's esophagus in the us: Estimates from a simulation model confirmed by seer data. *Dis Esophagus* 2010;**23**;451-457.
79. Naini BV, Souza RF, Odze RD. Barrett's esophagus: A comprehensive and contemporary review for pathologists. *Am J Surg Pathol* 2016;**40**;e45-66.
80. Yousef F, Cardwell C, Cantwell MM, Galway K, Johnston BT, Murray L. The incidence of esophageal cancer and high-grade dysplasia in barrett's esophagus: A systematic review and meta-analysis. *Am J Epidemiol* 2008;**168**;237-249.
81. Kaye PV, Ilyas M, Soomro I *et al.* Dysplasia in barrett's oesophagus: P53 immunostaining is more reproducible than haematoxylin and eosin diagnosis and improves overall reliability, while grading is poorly reproducible. *Histopathology* 2016;**69**;431-440.
82. Kerkhof M, van Dekken H, Steyerberg EW *et al.* Grading of dysplasia in barrett's oesophagus: Substantial interobserver variation between general and gastrointestinal pathologists. *Histopathology* 2007;**50**;920-927.
83. Wani S, Falk GW, Post J *et al.* Risk factors for progression of low-grade dysplasia in patients with barrett's esophagus. *Gastroenterology* 2011;**141**;1179-1186, 1186 e1171.
84. Farazi TA, Hoell JI, Morozov P, Tuschl T. Micrornas in human cancer. *Adv Exp Med Biol* 2013;**774**;1-20.

85. Hayes J, Peruzzi PP, Lawler S. Micrornas in cancer: Biomarkers, functions and therapy. *Trends Mol Med* 2014;**20**;460-469.
86. Li X, Kleeman S, Coburn SB *et al.* Selection and application of tissue micrornas for nonendoscopic diagnosis of barrett's esophagus. *Gastroenterology* 2018;**155**;771-783 e773.
87. Slaby O, Srovnal J, Radova L *et al.* Dynamic changes in microrna expression profiles reflect progression of barrett's esophagus to esophageal adenocarcinoma. *Carcinogenesis* 2015;**36**;521-527.
88. Rijlaarsdam MA, van Agthoven T, Gillis AJ *et al.* Identification of known and novel germ cell cancer-specific (embryonic) mirs in serum by high-throughput profiling. *Andrology* 2015;**3**;85-91.
89. Mestdagh P, Van Vlierberghe P, De Weer A *et al.* A novel and universal method for microrna rt-qpcr data normalization. *Genome Biol* 2009;**10**;R64.
90. Benjamini Y HJ. Controlling the false discovery rate: A practical and powerful approach to multiple testing. . *Journal of the Royal Statistical Society* 1995;**57(No. 1)**;289–300. .
91. Dilworth MP, Nieto T, Stockton JD *et al.* Whole genome methylation analysis of nondysplastic barrett esophagus that progresses to invasive cancer. *Ann Surg* 2018.
92. Wang JS, Guo M, Montgomery EA *et al.* DNA promoter hypermethylation of p16 and apc predicts neoplastic progression in barrett's esophagus. *Am J Gastroenterol* 2009;**104**;2153-2160.
93. Reid BJ, Levine DS, Longton G, Blount PL, Rabinovitch PS. 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**;1669-1676.
94. Bartel DP. Micrornas: Genomics, biogenesis, mechanism, and function. *Cell* 2004;**116**;281-297.
95. D'Haene B, Mestdagh P, Hellemans J, Vandesompele J. Mirna expression profiling: From reference genes to global mean normalization. *Methods Mol Biol* 2012;**822**;261-272.
96. Schwarzenbach H, da Silva AM, Calin G, Pantel K. Data normalization strategies for microrna quantification. *Clin Chem* 2015;**61**;1333-1342.
97. Chugh P, Dittmer DP. Potential pitfalls in microrna profiling. *Wiley Interdiscip Rev RNA* 2012;**3**;601-616.
98. Coleman HG, Xie SH, Lagergren J. The epidemiology of esophageal adenocarcinoma. *Gastroenterology* 2017.
99. Marino KA, Sullivan JL, Weksler B. Esophagectomy versus endoscopic resection for patients with early-stage esophageal adenocarcinoma: A national cancer database propensity-matched study. *J Thorac Cardiovasc Surg* 2018;**155**;2211-2218 e2211.
100. Pimentel-Nunes P, Dinis-Ribeiro M, Ponchon T *et al.* Endoscopic submucosal dissection: European society of gastrointestinal endoscopy (esge) guideline. *Endoscopy* 2015;**47**;829-854.
101. Naveed M, Kubiliun N. Endoscopic treatment of early-stage esophageal cancer. *Curr Oncol Rep* 2018;**20**;71.
102. Gertler R, Stein HJ, Schuster T, Rondak IC, Hofler H, Feith M. Prevalence and topography of lymph node metastases in early esophageal and gastric cancer. *Ann Surg* 2014;**259**;96-101.

103. Sepesi B, Watson TJ, Zhou D *et al.* Are endoscopic therapies appropriate for superficial submucosal esophageal adenocarcinoma? An analysis of esophagectomy specimens. *J Am Coll Surg* 2010;**210**;418-427.
104. Bollschweiler E, Baldus SE, Schroder W *et al.* High rate of lymph-node metastasis in submucosal esophageal squamous-cell carcinomas and adenocarcinomas. *Endoscopy* 2006;**38**;149-156.
105. Gockel I, Sgourakis G, Lyros O *et al.* Risk of lymph node metastasis in submucosal esophageal cancer: A review of surgically resected patients. *Expert Rev Gastroenterol Hepatol* 2011;**5**;371-384.
106. Gamboa AM, Kim S, Force SD *et al.* Treatment allocation in patients with early-stage esophageal adenocarcinoma: Prevalence and predictors of lymph node involvement. *Cancer* 2016;**122**;2150-2157.
107. Dubecz A, Kern M, Solymosi N, Schweigert M, Stein HJ. Predictors of lymph node metastasis in surgically resected t1 esophageal cancer. *Ann Thorac Surg* 2015;**99**;1879-1885; discussion 1886.
108. Ueno H, Murphy J, Jass JR, Mochizuki H, Talbot IC. Tumour 'budding' as an index to estimate the potential of aggressiveness in rectal cancer. *Histopathology* 2002;**40**;127-132.
109. Grigore AD, Jolly MK, Jia D, Farach-Carson MC, Levine H. Tumor budding: The name is emt. Partial emt. *J Clin Med* 2016;**5**.
110. Koelzer VH, Zlobec I, Lugli A. Tumor budding in colorectal cancer—ready for diagnostic practice? *Hum Pathol* 2016;**47**;4-19.
111. Mitrovic B, Schaeffer DF, Riddell RH, Kirsch R. Tumor budding in colorectal carcinoma: Time to take notice. *Mod Pathol* 2012;**25**;1315-1325.
112. Sanjay Kakar CS, Mariana E. Berho, David K. Driman, Patrick Fitzgibbons, Wendy L. Frankel, Kalisha A. Hill, John Jessup, Alyssa M. Krasinskas, Mary K Washington Protocol for the examination of specimens from patients with primary carcinoma of the colon and rectum (4.0.1.0). College of American Pathologists (CAP), 2017.
113. Board WCoTE. *Digestive system tumours*. Lyon (France): International Agency for Research on Cancer, 2019.
114. Ohike N, Coban I, Kim GE *et al.* Tumor budding as a strong prognostic indicator in invasive ampullary adenocarcinomas. *Am J Surg Pathol* 2010;**34**;1417-1424.
115. Ueno H, Mochizuki H, Hashiguchi Y *et al.* Risk factors for an adverse outcome in early invasive colorectal carcinoma. *Gastroenterology* 2004;**127**;385-394.
116. Koelzer VH, Langer R, Zlobec I, Lugli A. Tumor budding in upper gastrointestinal carcinomas. *Front Oncol* 2014;**4**;216.
117. Dutkowski P, Hommel G, Bottger T, Schlick T, Junginger T. How many lymph nodes are needed for an accurate pn classification in esophageal cancer? Evidence for a new threshold value. *Hepatogastroenterology* 2002;**49**;176-180.
118. Gotink AW, Ten Kate FJ, Doukas M *et al.* Do pathologists agree with each other on the histological assessment of pt1b oesophageal adenocarcinoma? *United European Gastroenterol J* 2019;**7**;261-269.
119. The paris endoscopic classification of superficial neoplastic lesions: Esophagus, stomach, and colon: November 30 to december 1, 2002. *Gastrointest Endosc* 2003;**58**;S3-43.

120. Lang-Schwarz C, Melcher B, Haumaier F *et al.* Budding and tumor-infiltrating lymphocytes–combination of both parameters predicts survival in colorectal cancer and leads to new prognostic subgroups. *Hum Pathol* 2018;**79**;160-167.
121. Puppa G, Senore C, Sheahan K *et al.* Diagnostic reproducibility of tumour budding in colorectal cancer: A multicentre, multinational study using virtual microscopy. *Histopathology* 2012;**61**;562-575.
122. Bokhorst JM, Blank A, Lugli A *et al.* Assessment of individual tumor buds using keratin immunohistochemistry: Moderate interobserver agreement suggests a role for machine learning. *Mod Pathol* 2019.
123. Cho SJ, Kakar S. Tumor budding in colorectal carcinoma: Translating a morphologic score into clinically meaningful results. *Arch Pathol Lab Med* 2018;**142**;952-957.
124. Trotsyuk I, Sparschuh H, Muller AJ *et al.* Tumor budding outperforms ypt and ypn classification in predicting outcome of rectal cancer after neoadjuvant chemoradiotherapy. *BMC Cancer* 2019;**19**;1033.
125. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J *et al.* Cancer incidence and mortality patterns in europe: Estimates for 40 countries in 2012. *Eur J Cancer* 2013;**49**;1374-1403.
126. Abnet CC, Arnold M, Wei WQ. Epidemiology of esophageal squamous cell carcinoma. *Gastroenterology* 2018;**154**;360-373.
127. Coleman HG, Xie SH, Lagergren J. The epidemiology of esophageal adenocarcinoma. *Gastroenterology* 2018;**154**;390-405.
128. Sobin L, Gospodarowicz M, Wittekind C eds. *Tnm classification of malignant tumours*, 2009.
129. Akutsu Y, Matsubara H. The significance of lymph node status as a prognostic factor for esophageal cancer. *Surg Today* 2011;**41**;1190-1195.
130. Wang WP, He SL, Yang YS, Chen LQ. Strategies of nodal staging of the tnm system for esophageal cancer. *Ann Transl Med* 2018;**6**;77.
131. Snyder DA, Rivers AM, Yokoe H, Menco BP, Anholt RR. Olfactomedin: Purification, characterization, and localization of a novel olfactory glycoprotein. *Biochemistry* 1991;**30**;9143-9153.
132. Zhang J, Liu WL, Tang DC *et al.* Identification and characterization of a novel member of olfactomedin-related protein family, hgc-1, expressed during myeloid lineage development. *Gene* 2002;**283**;83-93.
133. van der Flier LG, Haegebarth A, Stange DE, van de Wetering M, Clevers H. Olfm4 is a robust marker for stem cells in human intestine and marks a subset of colorectal cancer cells. *Gastroenterology* 2009;**137**;15-17.
134. Chin KL, Aerbajinai W, Zhu J *et al.* The regulation of olfm4 expression in myeloid precursor cells relies on nf-kappab transcription factor. *Br J Haematol* 2008;**143**;421-432.
135. Kim KK, Park KS, Song SB, Kim KE. Up regulation of gw112 gene by nf kappab promotes an antiapoptotic property in gastric cancer cells. *Mol Carcinog* 2010;**49**;259-270.
136. Liu W, Li H, Hong SH, Piszczek GP, Chen W, Rodgers GP. Olfactomedin 4 deletion induces colon adenocarcinoma in apcmin/+ mice. *Oncogene* 2016;**35**;5237-5247.

137. Liu W, Chen L, Zhu J, Rodgers GP. The glycoprotein hgc-1 binds to cadherin and lectins. *Exp Cell Res* 2006;**312**;1785-1797.
138. Liu W, Zhu J, Cao L, Rodgers GP. Expression of hgc-1 is correlated with differentiation of gastric carcinoma. *Histopathology* 2007;**51**;157-165.
139. Luo Z, Zhang Q, Zhao Z, Li B, Chen J, Wang Y. Olfm4 is associated with lymph node metastasis and poor prognosis in patients with gastric cancer. *J Cancer Res Clin Oncol* 2011;**137**;1713-1720.
140. Xiong B, Lei X, Zhang L, Fu J. The clinical significance and biological function of olfactomedin 4 in triple negative breast cancer. *Biomedicine & Pharmacotherapy* 2017;**86**;67-73.
141. Edge SB BD, Compton CC, Fritz AG, Greene FL, Trotti A, editors. *Ajcc cancer staging manual (7th ed)*. . New York, NY: Springer, 2010.
142. Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting recommendations for tumor marker prognostic studies (remark): Explanation and elaboration. *BMC Med* 2012;**10**;51.
143. Cuellar SL, Carter BW, Macapinlac HA *et al.* Clinical staging of patients with early esophageal adenocarcinoma: Does fdg-pet/ct have a role? *J Thorac Oncol* 2014;**9**;1202-1206.
144. Khanna LG, Gress FG. Preoperative evaluation of oesophageal adenocarcinoma. *Best Pract Res Clin Gastroenterol* 2015;**29**;179-191.
145. Pech O, Gunter E, Dusemund F, Origer J, Lorenz D, Ell C. Accuracy of endoscopic ultrasound in preoperative staging of esophageal cancer: Results from a referral center for early esophageal cancer. *Endoscopy* 2010;**42**;456-461.
146. Tekola BD, Sauer BG, Wang AY, White GE, Shami VM. Accuracy of endoscopic ultrasound in the diagnosis of t2n0 esophageal cancer. *J Gastrointest Cancer* 2014;**45**;342-346.
147. Winiker M, Mantziari S, Figueiredo SG, Demartines N, Allemann P, Schafer M. Accuracy of preoperative staging for a priori resectable esophageal cancer. *Dis Esophagus* 2018;**31**;1-6.
148. Grover PK, Hardingham JE, Cummins AG. Stem cell marker olfactomedin 4: Critical appraisal of its characteristics and role in tumorigenesis. *Cancer Metastasis Rev* 2010;**29**;761-775.
149. Kobayashi D, Koshida S, Moriai R, Tsuji N, Watanabe N. Olfactomedin 4 promotes s-phase transition in proliferation of pancreatic cancer cells. *Cancer Sci* 2007;**98**;334-340.
150. Zhang X, Huang Q, Yang Z, Li Y, Li CY. Gw112, a novel antiapoptotic protein that promotes tumor growth. *Cancer Res* 2004;**64**;2474-2481.
151. Seko N, Oue N, Noguchi T *et al.* Olfactomedin 4 (gw112, hgc-1) is an independent prognostic marker for survival in patients with colorectal cancer. *Exp Ther Med* 2010;**1**;73-78.
152. Chen L, Li H, Liu W *et al.* Olfactomedin 4 suppresses prostate cancer cell growth and metastasis via negative interaction with cathepsin d and sdf-1. *Carcinogenesis* 2011;**32**;986-994.
153. Guo LL, He ZC, Yang CQ, Qiao PT, Yin GL. Epigenetic silencing of olfactomedin-4 enhances gastric cancer cell invasion via activation of focal adhesion kinase signaling. *BMB Rep* 2015;**48**;630-635.
154. Abrams JA, Sharaiha RZ, Gonsalves L, Lightdale CJ, Neugut AI. Dating the rise of esophageal adenocarcinoma: Analysis of connecticut tumor registry data, 1940-2007. *Cancer Epidemiol Biomarkers Prev* 2011;**20**;183-186.
155. Bosetti C, Levi F, Ferlay J *et al.* Trends in oesophageal cancer incidence and mortality in europe. *Int J Cancer* 2008;**122**;1118-1129.

156. Rubenstein JH, Shaheen NJ. Epidemiology, diagnosis, and management of esophageal adenocarcinoma. *Gastroenterology* 2015;**149**;302-317.
157. Sobin L.H. GMK, Wittekind C. *Tnm classification of malignant tumours, 7th edition*. 7th Edition ed: Wiley-Blackwell, 2010;336.
158. Talsma K, van Hagen P, Grotenhuis BA *et al.* Comparison of the 6th and 7th editions of the uicc-ajcc tnm classification for esophageal cancer. *Ann Surg Oncol* 2012;**19**;2142-2148.
159. Chen M, Huang J, Zhu Z, Zhang J, Li K. Systematic review and meta-analysis of tumor biomarkers in predicting prognosis in esophageal cancer. *BMC Cancer* 2013;**13**;539.
160. Fisher OM, Lord SJ, Falkenback D, Clemons NJ, Eslick GD, Lord RV. The prognostic value of tp53 mutations in oesophageal adenocarcinoma: A systematic review and meta-analysis. *Gut* 2017;**66**;399-410.
161. McCormick Matthews LH, Noble F, Tod J *et al.* Systematic review and meta-analysis of immunohistochemical prognostic biomarkers in resected oesophageal adenocarcinoma. *Br J Cancer* 2015;**113**;107-118.
162. Theisen J, Peters JH, Fein M *et al.* The mutagenic potential of duodenoesophageal reflux. *Ann Surg* 2005;**241**;63-68.
163. Ross-Innes CS, Becq J, Warren A *et al.* Whole-genome sequencing provides new insights into the clonal architecture of barrett's esophagus and esophageal adenocarcinoma. *Nat Genet* 2015;**47**;1038-1046.
164. Weaver JM, Ross-Innes CS, Shannon N *et al.* Ordering of mutations in preinvasive disease stages of esophageal carcinogenesis. *Nat Genet* 2014;**46**;837-843.
165. 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;**55**;1390-1397.
166. Sikkema M, Kerkhof M, Steyerberg EW *et al.* Aneuploidy and overexpression of ki67 and p53 as markers for neoplastic progression in barrett's esophagus: A case-control study. *Am J Gastroenterol* 2009;**104**;2673-2680.
167. 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**;2508-2513.
168. Younes M, Brown K, Lauwers GY *et al.* P53 protein accumulation predicts malignant progression in barrett's metaplasia: A prospective study of 275 patients. *Histopathology* 2017.
169. Kaye PV, Haider SA, James PD *et al.* Novel staining pattern of p53 in barrett's dysplasia—the absent pattern. *Histopathology* 2010;**57**;933-935.
170. Kandioler D, Schoppmann SF, Zwrtek R *et al.* The biomarker tp53 divides patients with neoadjuvantly treated esophageal cancer into 2 subgroups with markedly different outcomes. A p53 research group study. *J Thorac Cardiovasc Surg* 2014;**148**;2280-2286.
171. Lehrbach DM, Cecconello I, Ribeiro Jr U, Capelozzi VL, Ab'saber AM, Alves VA. Adenocarcinoma of the esophagogastric junction: Relationship between clinicopathological data and p53, cyclin d1 and bcl-2 immunoexpressions. *Arq Gastroenterol* 2009;**46**;315-320.

172. Madani K, Zhao R, Lim HJ, Casson AG. Prognostic value of p53 mutations in oesophageal adenocarcinoma: Final results of a 15-year prospective study. *Eur J Cardiothorac Surg* 2010;**37**;1427-1432.
173. Siewert RJ, Feith M, Werner M, Stein HJ. Adenocarcinoma of the esophagogastric junction: Results of surgical therapy based on anatomical/topographic classification in 1,002 consecutive patients. *Ann Surg* 2000;**232**;353-361.
174. Dubbink HJ, Atmodimedjo PN, Kros JM *et al.* Molecular classification of anaplastic oligodendroglioma using next-generation sequencing: A report of the prospective randomized eortc brain tumor group 26951 phase iii trial. *Neuro Oncol* 2016;**18**;388-400.
175. van Nistelrooij AM, van Marion R, Koppert LB *et al.* Molecular clonality analysis of esophageal adenocarcinoma by multiregion sequencing of tumor samples. *BMC Res Notes* 2017;**10**;144.
176. Dulak AM, Stojanov P, Peng S *et al.* Exome and whole-genome sequencing of esophageal adenocarcinoma identifies recurrent driver events and mutational complexity. *Nat Genet* 2013;**45**;478-486.
177. Sturm D, Witt H, Hovestadt V *et al.* Hotspot mutations in h3f3a and idh1 define distinct epigenetic and biological subgroups of glioblastoma. *Cancer Cell* 2012;**22**;425-437.
178. Fontebasso AM, Papillon-Cavanagh S, Schwartzentruber J *et al.* Recurrent somatic mutations in acvr1 in pediatric midline high-grade astrocytoma. *Nat Genet* 2014;**46**;462-466.
179. Nones K, Waddell N, Wayte N *et al.* Genomic catastrophes frequently arise in esophageal adenocarcinoma and drive tumorigenesis. *Nat Commun* 2014;**5**;5224.
180. Aloia TA, Harpole DH, Jr., Reed CE *et al.* Tumor marker expression is predictive of survival in patients with esophageal cancer. *Ann Thorac Surg* 2001;**72**;859-866.
181. Cavazzola LT, Rosa AR, Schirmer CC *et al.* Immunohistochemical evaluation for p53 and vegf (vascular endothelial growth factor) is not prognostic for long term survival in end stage esophageal adenocarcinoma. *Rev Col Bras Cir* 2009;**36**;24-34.
182. Falkenback D, Nilbert M, Oberg S, Johansson J. Prognostic value of cell adhesion in esophageal adenocarcinomas. *Dis Esophagus* 2008;**21**;97-102.
183. Fareed KR, Al-Attar A, Soomro IN *et al.* Tumour regression and ercc1 nuclear protein expression predict clinical outcome in patients with gastro-oesophageal cancer treated with neoadjuvant chemotherapy. *Br J Cancer* 2010;**102**;1600-1607.
184. Flejou JF, Paraf F, Potet F, Muzeau F, Fekete F, Henin D. P53 protein expression in barrett's adenocarcinoma: A frequent event with no prognostic significance. *Histopathology* 1994;**24**;487-489.
185. Moskaluk CA, Heitmiller R, Zahurak M, Schwab D, Sidransky D, Hamilton SR. P53 and p21(waf1/cip1/sdi1) gene products in barrett esophagus and adenocarcinoma of the esophagus and esophagogastric junction. *Hum Pathol* 1996;**27**;1211-1220.
186. Gibson MK, Abraham SC, Wu TT *et al.* Epidermal growth factor receptor, p53 mutation, and pathological response predict survival in patients with locally advanced esophageal cancer treated with preoperative chemoradiotherapy. *Clin Cancer Res* 2003;**9**;6461-6468.

187. Schneider PM, Stoeltzing O, Roth JA *et al.* P53 mutational status improves estimation of prognosis in patients with curatively resected adenocarcinoma in barrett's esophagus. *Clin Cancer Res* 2000;**6**;3153-3158.
188. Wu TT, Watanabe T, Heitmiller R, Zahurak M, Forastiere AA, Hamilton SR. Genetic alterations in barrett esophagus and adenocarcinomas of the esophagus and esophagogastric junction region. *Am J Pathol* 1998;**153**;287-294.
189. He C, Li L, Guan X, Xiong L, Miao X. Mutant p53 gain of function and chemoresistance: The role of mutant p53 in response to clinical chemotherapy. *Chemotherapy* 2017;**62**;43-53.
190. Hientz K, Mohr A, Bhakta-Guha D, Efferth T. The role of p53 in cancer drug resistance and targeted chemotherapy. *Oncotarget* 2017;**8**;8921-8946.
191. Kappel S, Janschek E, Wolf B *et al.* Tp53 germline mutation may affect response to anticancer treatments: Analysis of an intensively treated li-fraumeni family. *Breast Cancer Res Treat* 2015;**151**;671-678.
192. Sauter ER, Keller SM, Erner SM. P53 correlates with improved survival in patients with esophageal adenocarcinoma. *J Surg Oncol* 1995;**58**;269-273.
193. Yamasaki M, Miyata H, Fujiwara Y *et al.* P53 genotype predicts response to chemotherapy in patients with squamous cell carcinoma of the esophagus. *Ann Surg Oncol* 2010;**17**;634-642.
194. Logullo AF, de Moura RP, Nonogaki S, Kowalski LP, Nagai MA, Simpson AJ. A proposal for the integration of immunohistochemical staining and DNA-based techniques for the determination of tp53 mutations in human carcinomas. *Diagn Mol Pathol* 2000;**9**;35-40.
195. Ribeiro U, Jr., Finkelstein SD, Safatle-Ribeiro AV *et al.* P53 sequence analysis predicts treatment response and outcome of patients with esophageal carcinoma. *Cancer* 1998;**83**;7-18.
196. Hoffman RM. Is DNA methylation the new guardian of the genome? *Mol Cytogenet* 2017;**10**;11.
197. Kappel SB, C.; Wolf, B; Gacic, S. ; Schoppmann, S. F. ; Devyatko, Y.; Prager, G.; Ba-SSalamah, A.; Wrba, F.; Pluschnig, U.; Kührer, I.; Mittlböck, M.; Hejna, M.; Zacherl, J.; Kandioler, D.; for the Pancho Study Group*. Turning the tables on surgical oncology: The pancho trial unplugged. *European Surgery* 2008;**40**;277-283.
198. Liu DS, Read M, Cullinane C *et al.* Apr-246 potently inhibits tumour growth and overcomes chemoresistance in preclinical models of oesophageal adenocarcinoma. *Gut* 2015;**64**;1506-1516.
199. Duhaylongsod FG, Gottfried MR, Iglehart JD, Vaughn AL, Wolfe WG. The significance of c-erb b-2 and p53 immunoreactivity in patients with adenocarcinoma of the esophagus. *Ann Surg* 1995;**221**;677-683; discussion 683-674.
200. van Hagen P, Hulshof MC, van Lanschot JJ *et al.* Preoperative chemoradiotherapy for esophageal or junctional cancer. *N Engl J Med* 2012;**366**;2074-2084.
201. Casey L, Kobel M, Ganesan R *et al.* A comparison of p53 and wt1 immunohistochemical expression patterns in tubo-ovarian high-grade serous carcinoma before and after neoadjuvant chemotherapy. *Histopathology* 2017;**71**;736-742.
202. Sakai K, Kazama S, Nagai Y *et al.* Chemoradiation provides a physiological selective pressure that increases the expansion of aberrant tp53 tumor variants in residual rectal cancerous regions. *Oncotarget* 2014;**5**;9641-9649.

203. Khrizman P, Niland JC, ter Veer A *et al.* Postoperative adjuvant chemotherapy use in patients with stage ii/iii rectal cancer treated with neoadjuvant therapy: A national comprehensive cancer network analysis. *J Clin Oncol* 2013;**31**;30-38.
204. Lichthardt S, Zenorini L, Wagner J *et al.* Impact of adjuvant chemotherapy after neoadjuvant radio- or radiochemotherapy for patients with locally advanced rectal cancer. *J Cancer Res Clin Oncol* 2017;**143**;2363-2373.
205. Sun Z, Gilmore B, Adam MA *et al.* Adjuvant chemotherapy after preoperative chemoradiation improves survival in patients with locally advanced rectal cancer. *Dis Colon Rectum* 2017;**60**;1050-1056.
206. Mokdad AA, Yopp AC, Polanco PM *et al.* Adjuvant chemotherapy vs postoperative observation following preoperative chemoradiotherapy and resection in gastroesophageal cancer: A propensity score-matched analysis. *JAMA Oncol* 2018;**4**;31-38.
207. Greaves M, Maley CC. Clonal evolution in cancer. *Nature* 2012;**481**;306-313.
208. Secrier M, Li X, de Silva N *et al.* Mutational signatures in esophageal adenocarcinoma define etiologically distinct subgroups with therapeutic relevance. *Nat Genet* 2016;**48**;1131-1141.
209. Smith MS, Ikonomi E, Bhuta R *et al.* Wide-area transepithelial sampling with computer-assisted 3-dimensional analysis (wats) markedly improves detection of esophageal dysplasia and barrett's esophagus: Analysis from a prospective multicenter community-based study. *Dis Esophagus* 2019;**32**.
210. Boellaard WPA, Gillis AJM, van Leenders G *et al.* Cellular origin of microrna-371a-3p in healthy males based on systematic urogenital tract tissue evaluation. *Andrology* 2019;**7**;463-468.
211. van Ipenburg JA, Gillis Ing AJM, Dorssers LCJ *et al.* Microrna profiling in benign and malignant conjunctival melanocytic lesions. *Ophthalmology* 2020;**127**;432-434.
212. Hashiguchi Y, Muro K, Saito Y *et al.* Japanese society for cancer of the colon and rectum (jsccr) guidelines 2019 for the treatment of colorectal cancer. *Int J Clin Oncol* 2020;**25**;1-42.
213. Vieth M, Quirke P, Lambert R, von Karsa L, Risio M, International Agency for Research on C. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First edition—annotations of colorectal lesions. *Endoscopy* 2012;**44 Suppl 3**;SE131-139.
214. Oetzmann von Sochaczewski C, Haist T, Pauthner M *et al.* The overall metastatic rate in early esophageal adenocarcinoma: Long-time follow-up of surgically treated patients. *Dis Esophagus* 2019;**32**.
215. Wu SG, Zhang WW, He ZY, Sun JY, Chen YX, Guo L. Sites of metastasis and overall survival in esophageal cancer: A population-based study. *Cancer Manag Res* 2017;**9**;781-788.
216. Wang XY, Chen SH, Zhang YN, Xu CF. Olfactomedin-4 in digestive diseases: A mini-review. *World J Gastroenterol* 2018;**24**;1881-1887.
217. Januszewicz W, Tan WK, Lehovsky K *et al.* Safety and acceptability of esophageal cytosponge cell collection device in a pooled analysis of data from individual patients. *Clin Gastroenterol Hepatol* 2019;**17**;647-656 e641.

218. Paterson AL, Gehrung M, Fitzgerald RC, O'Donovan M. Role of tff3 as an adjunct in the diagnosis of barrett's esophagus using a minimally invasive esophageal sampling device-the cytosponge(tm). *Diagn Cytopathol* 2019.
219. Boldrin E, Rumiato E, Fassan M *et al.* Liquid biopsy as a novel tool to monitor the carcinogenesis of barrett's esophagus. *Transl Res* 2016;**176**;127-131.
220. Rumiato E, Boldrin E, Malacrida S *et al.* Detection of genetic alterations in cfdna as a possible strategy to monitor the neoplastic progression of barrett's esophagus. *Transl Res* 2017;**190**;16-24 e11.
221. Committee AT, Trindade AJ, Navaneethan U *et al.* Advances in the diagnosis and surveillance of barrett's esophagus (with videos). *Gastrointest Endosc* 2019;**90**;325-334.
222. Duits LC, Phoa KN, Curvers WL *et al.* Barrett's oesophagus patients with low-grade dysplasia can be accurately risk-stratified after histological review by an expert pathology panel. *Gut* 2015;**64**;700-706.
223. Cancer Genome Atlas Research N, Analysis Working Group: Asan U, Agency BCC *et al.* Integrated genomic characterization of oesophageal carcinoma. *Nature* 2017;**541**;169-175.
224. Merve A, Millner TO, Marino S. Integrated phenotype-genotype approach in diagnosis and classification of common central nervous system tumours. *Histopathology* 2019;**75**;299-311.
225. Coombes RC, Powles TJ, Berger U *et al.* Prediction of endocrine response in breast cancer by immunocytochemical detection of oestrogen receptor in fine-needle aspirates. *Lancet* 1987;**2**;701-703.
226. Wang VE, Grandis JR, Ko AH. New strategies in esophageal carcinoma: Translational insights from signaling pathways and immune checkpoints. *Clin Cancer Res* 2016;**22**;4283-4290.
227. Kelly RJ, Zaidi AH, Smith MA *et al.* The dynamic and transient immune microenvironment in locally advanced esophageal adenocarcinoma post chemoradiation. *Ann Surg* 2018;**268**;992-999.
228. Ku GY. The current status of immunotherapies in esophagogastric cancer. *Hematol Oncol Clin North Am* 2019;**33**;323-338.

Appendices

CURRICULUM VITAE

Lucia Suzuki werd geboren op 28 augustus 1988 te Rotterdam. In 2006 behaalde zij haar VWO diploma aan het Comenius College te Capelle a/d IJssel. In 2007 behaalde zij haar propedeuse van de studie Geneeskunde aan de Erasmus Universiteit Rotterdam. In 2010 behaalde zij haar doctoraal examen, terwijl zij in 2009 ook gestart was met een tweejarige Researchmaster Neuroscience aan dezelfde universiteit. In 2011 behaalde ze vervolgens haar Master of Science in Neuroscience (*Excellent*), waarna ze aan haar coschappen begon. In 2014 doorliep ze haar keuzecoschap bij de afdeling Pathologie en haar oudste coschap bij de Neurologie (beide in het Erasmus MC), waarna ze in augustus 2014 haar artsexamen behaalde (*Cum Laude*). Na een korte periode als ANIOS op de afdeling Neurologie van het Albert Schweitzer Ziekenhuis te Dordrecht begon zij eind december 2014 aan haar vervolgopleiding tot patholoog in het Erasmus Medisch Centrum (opleider: dr. G.J.L.H. van Leenders). In oktober 2016 onderbrak zij de opleiding voor een promotietraject gericht op Barrett en adenocarcinoom van de slokdarm onder supervisie van prof. dr. L.J.H. Looijenga, prof. dr. F.J. van Kemenade (promotoren) en dr. K. Biermann (copromotor). Per april 2020 ging zij weer verder met haar vervolgopleiding tot patholoog (opleider: dr. R.M. Verdijk). Deze verwacht zij, na perifere stages in o.a. het Maasstad Ziekenhuis (opleider: dr. M.A. den Bakker) te Rotterdam en het Reinier Haga Medisch Diagnostisch Centrum in Delft (opleider: dr. F.M.M. Smedts), in 2023 af te ronden.

LIST OF PUBLICATIONS

Suzuki L, Gillis AJM, Dorssers LCJ, Henriquez J, Nieboer D, Bruno MJ, Spaander MCW, Looijenga LHJ, Biermann K. ***Exploring histology-based microRNA profiles in the progression of Barrett's esophagus to esophageal adenocarcinoma.*** Submitted.

Suzuki L, ten Kate FJC, Gotink AW, van de Ven SEM, Nieboer D, Weusten BLAM, Brosens LAA, van Hillegersberg R, Alvarez Herrero L, Seldenrijk CA, Alkhalaf A, Moll FCP, Schoon EJ, van Lijnschoten I, Tang T, van der Valk H, Nagengast WB, Kats-Ugurlu G, Plukker JTM, Houben MHMG, van der Laan J, Pouw RE, Bergman JJGHM, Meijer SL, van Berge Henegouwen MI, Wijnhoven BPL, de Jonge PJF, Doukas M, Bruno MJ, Looijenga LHJ, Koch AD, Biermann K. ***Tumor budding is prognostic for metastases in patients with submucosal esophageal adenocarcinoma.*** Submitted.

Gotink AW, van de Ven SEM, Ten Kate FJC, Nieboer D, **Suzuki L**, Weusten BLAM, Brosens LAA, van Hillegersberg R, Alvarez Herrero L, Seldenrijk CA, Alkhalaf A, Moll FCP, Schoon EJ, van Lijnschoten I, Tang TJ, van der Valk H, Nagengast WB, Kats-Ugurlu G, Plukker JTM, Houben MHMG, van der Laan JS, Pouw RE, Bergman JJGHM, Meijer SL, van Berge Henegouwen MI, Wijnhoven BPL, de Jonge PJF, Doukas M, Bruno MJ, Biermann K, Koch AD. ***Individual risk calculator to predict lymph node metastases in patients with submucosal (T1b) esophageal adenocarcinoma: a multicenter cohort study.*** Endoscopy 2021 Feb 24. doi: 10.1055/a-1399-4989. Epub ahead of print.

Suzuki L, Ten Kate FJC, Gotink AW, Stoop H, Doukas M, Nieboer D, Spaander MCW, van Lanschot JJB, Wijnhoven BPL, Koch AD, Bruno MJ, Looijenga LHJ, Biermann K. ***Olfactomedin 4 (OLFM4) expression is associated with nodal metastases in esophageal adenocarcinoma.*** PLoS One 2019;14;e0219494.

Ten Kate FJC, **Suzuki L**, Dorssers LCJ, Dinjens WNM, Jones DTW, Nieboer D, Doukas M, van Lanschot JJB, Wijnhoven BPL, Looijenga LHJ, Biermann K. ***Pattern of p53 protein expression is predictive for survival in chemoradiotherapy-naïve esophageal adenocarcinoma.*** Oncotarget 2017;8;104123-104135.

Suzuki L, Nieboer D, van Lanschot JJB, Spaander MCW, Looijenga LHJ, Biermann K. ***Effect of neoadjuvant chemoradiotherapy on p53 and SOX2 protein expression in esophageal adenocarcinoma.*** Biomarkers in Medicine, 2020, 14(9), 785–793.

Other

Suzuki L, Coulon P, Sabel-Goedknegt EH, Ruigrok TJ**. *Organization of cerebral projections to identified cerebellar zones in the posterior cerebellum of the rat.*** J Neurosci 2012;32;10854-10869.

PHD PORTFOLIO

Name PhD student:	Lucia Suzuki
Research School:	Molmed
Department:	Pathology
PhD period:	2016-2020
Promotoren:	Prof. dr. L.H.J. Looijenga, Prof. dr. F.J. van Kemenade
Copromotor:	Dr. K. Biermann

PhD training	Year	ECTs
General courses		
Biomedical English Writing and Communication	2019-2020	2
Research Integrity	2019	0,3
Basic course on SPSS	2016	1
Basic course on R	2017	1,8
Survival analysis Course	2017	0,6
Photoshop & Illustrator for PhD-students and other researchers	2017	0,3
Specific courses		
Basis Opleiding Pathologie (BOP) Cursus Pathofysiologie	2020	1
NGS in DNA diagnostics course	2016	1
Basis Opleiding Pathologie (BOP) Cursus Oncologie	2016	1
LPAV courses	2016-2021	1
Seminars and workshops		
BDIAP Molecular Pathology Study Day	2020	0,3
ESP Webinar: precursors in epithelial malignancy in the digestive system: what's new?	2020	0,1
Endnote workshop	2019	0,2
Biomedical Research Techniques XV	2016	1,5
Basic and Translational Oncology	2016	1,8
Discipline Overstijgend Onderwijs (DOO) – Samenwerking	2016	0,3
9th BDIAP Seminar for Trainees in Histopathology on Approach to Cut-up	2016	0,3
Discipline Overstijgend Onderwijs (DOO) – Kwaliteitsvisitaties, beroepsvisitaties, communicatiestromen en praktijkorganisatie (LPAV)	2016	0,3

	Year	ECTs
(Inter)national conferences/ meetings		
110th United States and Canadian Academy of Pathology (USCAP) Annual meeting, virtual	2021	1
32st European Congress of Pathology (ECP) Annual meeting, virtual	2020	1
Molecular Medicine day, Rotterdam, The Netherlands	2019	0,3
10th Belgian Week of Pathology, Brussels, Belgium (poster, 2x)	2019	1
31st European Congress of Pathology, Nice, France (poster)	2019	1
Erasmus MC Cancer Institute Research Day	2018, 2019	0,6
28th ESDE Annual Meeting, Utrecht, The Netherlands	2017	0,3
29th European Congress of Pathology, Amsterdam, The Netherlands (oral)	2017	1
CRUK International Symposium on Oesophageal cancer, Cambridge, United Kingdom (poster)	2017	1
5th Daniel den Hoed day	2017	0,3
Bijeenkomst Moleculaire Diagnostiek in de Pathologie	2016	0,3
Week van de Pathologie	2016, 2017, 2019	3
LEPO lab meetings (oral)	2016-2019	1
JNI meetings (oral)	2016-2020	1
PALM meetings (oral)	2016-2018	1
Barrett research meetings (oral)	2016-2020	1
Teaching		
Teach the Teacher II	2019	0,3
Teaching medical students (Vaardigheidsonderwijs)	2016-2020	3
Tutor first year medical students	2016-2017	2
Basis Kwalificatie Onderwijs (BKO): Training "Omgaan met groepen"	2016	0,1
Other activities		
Board member LPAV	2016-2018	2

DANKWOORD

Bij deze wil ik iedereen bedanken voor hun steun en medewerking bij de totstandkoming van dit proefschrift. De leden van de kleine en grote commissie wil ik bedanken voor het beoordelen van mijn proefschrift en hun deelname aan de oppositie. (Ex-) Collega's van het Erasmus MC, zowel van de afdeling Pathologie, Chirurgie, Maag-, Darm- Leverziekten en Maatschappelijke gezondheidszorg en natuurlijk alle coauteurs, ontzettend bedankt voor jullie betrokkenheid en waardevolle input.

Een aantal personen wil ik in het bijzonder bedanken. Allereerst mijn copromotor Katharina Biermann. Dank voor het vertrouwen om mij als eerstejaars AIOS te vragen voor dit promotieonderzoek. Ik heb wetenschappelijk onderzoek altijd belangrijk gevonden en wist al snel na de start van mijn opleiding tot patholoog dat ik graag wilde promoveren. Jij bood mij die kans en hebt mij gedurende dit traject op een erg fijne manier begeleid. Ik hoop dat jullie inmiddels jullie draai hebben gevonden in Duitsland en weet zeker dat je niet alleen door mij wordt gemist in het Erasmus MC.

Mijn promotor prof. dr. L.H.J. Looijenga, beste Leendert, dank voor uw goede begeleiding en adviezen. Het was een prettige samenwerking ondanks dat deze de laatste periode vooral per mail verliep. Ik twijfel er niet aan dat u ook in het Prinses Máxima Centrum vele mooie onderzoeksresultaten zult behalen. Mijn andere promotor, prof. dr. F.J. van Kemenade, beste Folkert, dank voor uw betrokkenheid, steun en wijze raad tijdens de opleiding en het afronden van dit proefschrift.

Laboratorium voor Experimentele Patho-oncologie, ofwel LEPO-labgenootjes Fiebo, Berdine, Martine en Remko, dank voor de altijd gezellige lunches en hulp bij allerhande problemen waar ik meer dan eens tegenaan liep. Beste Fiebo, dank voor je gezelligheid op de kamer. Ik vond het erg fijn om met je te kunnen sparren over mijn onderzoeksresultaten. Daarnaast heb ik genoten van je verhalen over je bijen en andere huisdieren. Berdine, gefeliciteerd met jouw eerste kindje, ik hoop dat alles goed gaat en wens jullie veel geluk en gezondheid toe. Beste Martine, ik kijk ook erg uit naar jouw proefschrift, wat heb jij hier ontzettend veel werk en energie ingestopt! Beste Remko, zelfs toen je al lang en breed klaar was bij LEPO stond je altijd klaar om mij te helpen. Ad en Lambert, dank voor jullie hulp en geduld bij het labwerk en de interpretatie hiervan. Joitza, de vrolijkste en gezelligste stagiaire die er is, dank voor je inzet en harde werk bij het microRNA onderzoek, heel veel succes met je verdere studie!

Mijn (schoon)familie, vrienden en vriendinnen, dank voor jullie interesse in mijn werk en onderzoek, al ben ik blij dat ik het met jullie juist over andere dingen kan hebben. Door COVID-19 hebben we elkaar de afgelopen tijd helaas veel minder vaak gezien. Ik hoop dat we snel weer vaker en op een normale manier bij elkaar te kunnen komen!

Ik heb jullie gezelligheid de afgelopen periode tijdens deze pandemie enorm gemist. Inge de Voogt, dank dat je tijd wilde vrijmaken voor het ontwerp van de omslag van mijn proefschrift!

Liefste Levy, ontzettend bedankt dat je altijd achter mij staat. Mede dankzij jou sta ik hier en ben ik meer dan alleen een moeder. Het ouderschap is pittig, maar brengt focus en vooral ook heel veel moois. Wat zijn wij samen toch een goed team! Lieve Yuno, jouw vrolijkheid en nieuwsgierigheid geven mij altijd de kracht om gewoon door te blijven gaan. Je wordt vast een hele lieve grote zus voor jouw broertje. Nog maar een paar weekjes wachten!

Tot slot, mijn paranimfen Charlotte en Mareska. Charlotte, wat fijn dat we naast vriendinnen ook collega's zijn geworden! Vanaf het begin van onze opleiding Geneeskunde hebben we, in binnen- en buitenland, heel wat avonturen met elkaar beleefd. Ik weet zeker dat er de komende jaren nog vele herinneringen zullen volgen. Mareska, de laatste jaren is onze band, ondanks het feit dat je al jaren in Londen woont, juist hechter geworden. Hoewel we op vele fronten enorm verschillen, lijken we stiekem toch meer op elkaar dan ik vroeger dacht. De afgelopen jaren brachten naast veel mooie dingen helaas ook erg moeilijke periodes. Hierdoor besef ik des te meer hoe belangrijk familie is en hoe trots en blij ik ben dat jij mijn zus bent. Dankbaar en blij dat jullie op deze bijzondere dag mijn paranimfen willen zijn.