This is the published version of a paper published in *Scandinavian Journal of Gastroenterology*.

Citation for the original published paper (version of record):

https://doi.org/10.1080/00365521.2019.1656775

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-164536
Z-line alterations and gastroesophageal reflux: an endoscopic population-based prospective cohort study

Bengt Wallner, Ove Björk, Anna Andreasson, Michael Vieth, Peter T. Schmidt, Per M. Hellström, Anna Forsberg, Nicholas J. Talley and Lars Agreus

INTRODUCTION

Barrett’s esophagus (BE) is a premalignant condition in the distal esophagus [1]. In BE, the squamous epithelium of the distal esophagus is replaced by columnar epithelium, thus displacing the squamocolumnar junction, also called the Z-line, orally, and giving rise to a columnar lined esophagus (CLE). In BE, there is an increased risk of developing esophageal adenocarcinoma (EAC) [2]. Gastroesophageal reflux (GER) is known to be of primary etiological importance in both BE and EAC [3].

The definition of BE is controversial, and no uniformly accepted criteria exist [4]. Previously, the endoscopic definition was a change in the esophageal epithelium, of any length that could be recognized at endoscopy [5]. BE was also previously subdivided into short-segment BE (CLE <3 cm in length) and long-segment BE (CLE >3 cm in length) [5]. Lately, a minimum finding of 1 cm CLE has been suggested for the diagnosis of BE to avoid overdiagnosis or misdiagnosis of columnar epithelium in a hiatal hernia. Some definitions, but not all, include the presence of histologically verified intestinal metaplasia (IM) in the endoscopic segment of CLE as the required diagnostic criterion [6–8].

Even though BE can lead to EAC, the issue of surveillance is controversial. At present, there are no randomized controlled trials available. However, the on-going Boss Trial aims to study whether regular endoscopic surveillance is better than endoscopy at need for detecting early signs of esophageal cancer in BE [8]. Some studies have suggested that with adequate endoscopic surveillance, EAC can be detected at an earlier stage, and with improved survival [9].

Even under ideal conditions, it has been shown that the inter-observer agreement regarding the endoscopic length measurement in the esophagus is poor, with a κ value of 0.40, and a mean difference between true length and measured length of 1.10 cm [10]. This margin of error is hence larger than the proposed limit for the diagnosis of BE.
There are at present two classification systems used to report the extent of endoscopic CLE in a standardized manner, one qualitative and one quantitative. The qualitative ZAP (Z-line appearance) classification evaluates the endoscopic appearance of the Z-line, and it has been associated with the prevalence of IM, as well as with GER [11–13]. The classification does not involve any endoscopic length measurement and has shown a high reliability with interrater agreements ($\kappa$ values) in the range of 0.72–0.90 between endoscopists [14]. The quantitative Prague C & M classification, which has reached most international acceptance, is used to describe the circumferential (C) and maximum (M) extent of columnar metaplasia in the distal esophagus measured in centimeters [15]. It has also been found reliable, albeit not in a clinical, diagnostic setting, since the length measurement in the validation papers were done from an edited video image with the depth of endoscope insertion given as a numeric value on the footage, and not from the 5 cm hash marks on the endoscope [15,16].

The aim of the present study was to investigate which endoscopical alterations of the Z-line can be attributed to a previous history of GER. Originally, the aim was to compare the finding of an irregular Z-line (ZAPI), with the finding of obvious tongues of CLE (ZAPII/III). Because of the recent proposal of a 1 cm limit of the CLE in order to be diagnosed with BE, we chose to also compare with the finding of a total Prague C/M-value of at least 1 cm.

Materials and methods

A comprehensive description of the setting, population, esophagogastroduodenoscopy (EGD) procedure and symptom evaluation has been given elsewhere [17].

Study population

A population cohort in Östhammar, Uppsala County, Sweden, has been surveyed for over two decades regarding gastrointestinal symptoms, using the validated Abdominal Symptom Questionnaire (ASQ) [18] in 1988, 1989, 1995 and 2011. The study population is representative of the general Swedish population in terms of gender, age, income and other potential selection factors. In 2012, the population was invited to undergo an EGD. Exclusion criteria for EGD were angina pectoris, myocardial infarction (last 6 months), congestive heart failure, severe lung disease, severe liver disease, esophageal varices, treatment with anticoagulants, need for anesthesia for the endoscopy, earlier surgery of the stomach and age 80 years or above.

Endoscopy

The endoscopies were performed by five experienced endoscopists (among them the authors BW, PH, AF and LA), who at the time of endoscopy were blinded to the presence of any symptoms. The EGDs were performed according to a strict protocol. The gastroesophageal junction (GEJ) was defined as the proximal margin of the gastric mucosal folds [19]. Esophagitis was classified according to the LA-classification [20]. CLE was measured endoscopically according to the Prague C/M-classification [15]. The gastroesophageal junction was video recorded both from above and below. One investigator (BW) classified the Z-line appearance according to the ZAP-classification [14], using the video recordings, blinded to the presence of any symptoms or clinical data. To test the reliability of the ZAP assessment, the inter-rater reliability of the ZAP classifications was evaluated by an endoscopist who did not participate in the original endoscopies or evaluations (PTS), who independently rated 40 subject videos. The agreement between endoscopists was 88% (35 out of 40); expected agreement was 40%, kappa = 0.79, $p < .0001$.

Definition of CLE

In order to determine which endoscopic alterations to the squamocolumnar junction can be attributed to a previous history of GER, three different definitions of CLE were used:

1. **ZAP I**: An irregular Z-line with a suspicion of tongue-like protrusions. Until recently an irregular Z-line was sufficient to be diagnosed with BE [5].
2. **ZAP II/III**: The presence of distinct, obvious tongues of metaplastic columnar epithelium. The definition has previously been linked to GER [11–13]. Since the discussion on long-segment, and short-segment BE, is no longer relevant, we have chosen to analyze ZAP II and III together.
3. **CLE $\geq 1$ cm**: The Prague C/M-classification with a minimum length of 1 cm (total length of metaplastic tongues and circular metaplasia). A 1 cm limit of the CLE has been proposed as cutoff for BE diagnosis [6].

Definition of GER

GER was defined as the occurrence of heartburn or acid regurgitation, in any of the three surveys from 1988, 1989 and 1995. In order to be classified as not having GER, the individuals needed to have negated heartburn and acid regurgitation at one of the occasions in the 80s, and in the survey from 1995. Since BE is known to be associated with a longstanding history of GER, recent symptoms (2011–2012) were not evaluated.

Inclusion criteria

Participants fulfilling all of the following criteria were included in the present study: (1) Data regarding regurgitation and heartburn from the questionnaires from 1988, 1989 and 1995; (2) Endoscopic data regarding Prague-classification from the endoscopy in 2012; (3) A video recording from the endoscopy in 2012 where the Z-line was possible to classify [13]. Since the discussion on long-segment, and short-segment BE, is no longer relevant, we have chosen to analyze ZAP II and III together.

1. **ZAP I**: An irregular Z-line with a suspicion of tongue-like protrusions. Until recently an irregular Z-line was sufficient to be diagnosed with BE [5].
2. **ZAP II/III**: The presence of distinct, obvious tongues of metaplastic columnar epithelium. The definition has previously been linked to GER [11–13]. Since the discussion on long-segment, and short-segment BE, is no longer relevant, we have chosen to analyze ZAP II and III together.
3. **CLE $\geq 1$ cm**: The Prague C/M-classification with a minimum length of 1 cm (total length of metaplastic tongues and circular metaplasia). A 1 cm limit of the CLE has been proposed as cutoff for BE diagnosis [6].

1. **ZAP I**: An irregular Z-line with a suspicion of tongue-like protrusions. Until recently an irregular Z-line was sufficient to be diagnosed with BE [5].
2. **ZAP II/III**: The presence of distinct, obvious tongues of metaplastic columnar epithelium. The definition has previously been linked to GER [11–13]. Since the discussion on long-segment, and short-segment BE, is no longer relevant, we have chosen to analyze ZAP II and III together.
3. **CLE $\geq 1$ cm**: The Prague C/M-classification with a minimum length of 1 cm (total length of metaplastic tongues and circular metaplasia). A 1 cm limit of the CLE has been proposed as cutoff for BE diagnosis [6].

1. **ZAP I**: An irregular Z-line with a suspicion of tongue-like protrusions. Until recently an irregular Z-line was sufficient to be diagnosed with BE [5].
2. **ZAP II/III**: The presence of distinct, obvious tongues of metaplastic columnar epithelium. The definition has previously been linked to GER [11–13]. Since the discussion on long-segment, and short-segment BE, is no longer relevant, we have chosen to analyze ZAP II and III together.
3. **CLE $\geq 1$ cm**: The Prague C/M-classification with a minimum length of 1 cm (total length of metaplastic tongues and circular metaplasia). A 1 cm limit of the CLE has been proposed as cutoff for BE diagnosis [6].

1. **ZAP I**: An irregular Z-line with a suspicion of tongue-like protrusions. Until recently an irregular Z-line was sufficient to be diagnosed with BE [5].
2. **ZAP II/III**: The presence of distinct, obvious tongues of metaplastic columnar epithelium. The definition has previously been linked to GER [11–13]. Since the discussion on long-segment, and short-segment BE, is no longer relevant, we have chosen to analyze ZAP II and III together.
3. **CLE $\geq 1$ cm**: The Prague C/M-classification with a minimum length of 1 cm (total length of metaplastic tongues and circular metaplasia). A 1 cm limit of the CLE has been proposed as cutoff for BE diagnosis [6].

1. **ZAP I**: An irregular Z-line with a suspicion of tongue-like protrusions. Until recently an irregular Z-line was sufficient to be diagnosed with BE [5].
2. **ZAP II/III**: The presence of distinct, obvious tongues of metaplastic columnar epithelium. The definition has previously been linked to GER [11–13]. Since the discussion on long-segment, and short-segment BE, is no longer relevant, we have chosen to analyze ZAP II and III together.
3. **CLE $\geq 1$ cm**: The Prague C/M-classification with a minimum length of 1 cm (total length of metaplastic tongues and circular metaplasia). A 1 cm limit of the CLE has been proposed as cutoff for BE diagnosis [6].

1. **ZAP I**: An irregular Z-line with a suspicion of tongue-like protrusions. Until recently an irregular Z-line was sufficient to be diagnosed with BE [5].
2. **ZAP II/III**: The presence of distinct, obvious tongues of metaplastic columnar epithelium. The definition has previously been linked to GER [11–13]. Since the discussion on long-segment, and short-segment BE, is no longer relevant, we have chosen to analyze ZAP II and III together.
3. **CLE $\geq 1$ cm**: The Prague C/M-classification with a minimum length of 1 cm (total length of metaplastic tongues and circular metaplasia). A 1 cm limit of the CLE has been proposed as cutoff for BE diagnosis [6].

1. **ZAP I**: An irregular Z-line with a suspicion of tongue-like protrusions. Until recently an irregular Z-line was sufficient to be diagnosed with BE [5].
2. **ZAP II/III**: The presence of distinct, obvious tongues of metaplastic columnar epithelium. The definition has previously been linked to GER [11–13]. Since the discussion on long-segment, and short-segment BE, is no longer relevant, we have chosen to analyze ZAP II and III together.
3. **CLE $\geq 1$ cm**: The Prague C/M-classification with a minimum length of 1 cm (total length of metaplastic tongues and circular metaplasia). A 1 cm limit of the CLE has been proposed as cutoff for BE diagnosis [6].

1. **ZAP I**: An irregular Z-line with a suspicion of tongue-like protrusions. Until recently an irregular Z-line was sufficient to be diagnosed with BE [5].
2. **ZAP II/III**: The presence of distinct, obvious tongues of metaplastic columnar epithelium. The definition has previously been linked to GER [11–13]. Since the discussion on long-segment, and short-segment BE, is no longer relevant, we have chosen to analyze ZAP II and III together.
3. **CLE $\geq 1$ cm**: The Prague C/M-classification with a minimum length of 1 cm (total length of metaplastic tongues and circular metaplasia). A 1 cm limit of the CLE has been proposed as cutoff for BE diagnosis [6].

1. **ZAP I**: An irregular Z-line with a suspicion of tongue-like protrusions. Until recently an irregular Z-line was sufficient to be diagnosed with BE [5].
2. **ZAP II/III**: The presence of distinct, obvious tongues of metaplastic columnar epithelium. The definition has previously been linked to GER [11–13]. Since the discussion on long-segment, and short-segment BE, is no longer relevant, we have chosen to analyze ZAP II and III together.
3. **CLE $\geq 1$ cm**: The Prague C/M-classification with a minimum length of 1 cm (total length of metaplastic tongues and circular metaplasia). A 1 cm limit of the CLE has been proposed as cutoff for BE diagnosis [6].
Histology

The presence of IM was determined from biopsies from the squamocolumnar junction. To eliminate any doubt concerning the origin of the IM, the intention was to obtain biopsies from the Z-line containing both squamous and columnar epithelium. The biopsy specimens were fixed in 4% phosphate-buffered formaldehyde then processed and embedded in paraffin wax. From each specimen, 4-μm sections were cut and stained with haematoxylin and eosin, and periodic acid-Schiff reagent. The highly experienced study pathologist (MV), an expert in esophageal pathology and blinded to all clinical data, examined the stained sections. IM was considered present if goblet cells were identified.

Statistics

Logistic regression was performed to identify which endoscopic alterations to the Z-line that was independently associated with a history of GER. The following alterations to the Z-line with binary cutoff’s were investigated as dependent variables:

1. ZAP I classification (ZAP0 vs. ZAPI).
2. ZAP II/III classification (ZAPII/I vs II/III).
3. The combined length of the C and M value according to the Prague C/M-classification (<1 cm vs. ≥1 cm).

These binary cutoffs were investigated with or without the presence of IM. Separate multivariable logistic models were performed with each group of alterations to the Z-line as the dependent variable and GER as the independent variable. Models were adjusted for body mass index, Helicobacter pylori infection and use of proton pump inhibitor, if display-
ing statistically significant univariate association with the endoscopic alteration. All models were adjusted for gender and age. Data analyses were performed in the Stata/IC software (StataCorp. 2013. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP, USA).

Ethics

The studies from 1988, 1989 and 1995 were approved by the Ethical Review Board of the Medical Faculty of Uppsala University (Dnr. 1988/103, 1989/220 and 1995/39). Approval for the 2011–2012 study was obtained from the Ethics Committee of Uppsala University (Dnr. 2010/443). All participants gave their informed consent.

Results

A total of 165 subjects were included in the study. Table 1 shows the demographics of the population. The biopsies from the squamocolumnar junction contained both squamous and columnar epithelium in 95% of the subjects. In this cohort, only eight subjects had a C value of at least 1 cm, and all of these also had a M component (2 C1M2, 2 C1M3, 1 C1M4, 1 C2M3, 1 C5M5 and 1 C7M10). Of the 165 subjects investigated, 99 had a ZAP I (of them 12% had IM), 26 had ZAP II/III (35% IM) and 40 had CLE ≥1 cm (21% IM).

Table 2 and 3 show the association between different endoscopic findings and previous symptoms of GER. Table 2

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR</th>
<th>95% Conf. interval</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZAPI vs. ZAP0</td>
<td>2.06</td>
<td>0.85–5.01</td>
<td>.110</td>
</tr>
<tr>
<td>ZAPIII vs. ZAP0/I</td>
<td>3.60</td>
<td>1.49–8.70</td>
<td>.004</td>
</tr>
<tr>
<td>The Prague C/M-classification</td>
<td>1.64</td>
<td>0.77–3.50</td>
<td>.199</td>
</tr>
</tbody>
</table>

Bold values indicates p-values < 0.05.

Figure 1. The study flow chart, illustrating the study population and the dropouts.

Table 1. Demographics of the investigated population (n = 165).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (year)</td>
<td>59 (38–79)</td>
<td>27 (18–44)</td>
<td>86 (51%)</td>
</tr>
<tr>
<td>BMI (mean, range)</td>
<td>23 (15%)</td>
<td>28 (17%)</td>
<td>40 (25%)</td>
</tr>
<tr>
<td>Intestinal metaplasia the Z-line</td>
<td>7 (4%)</td>
<td>57 (35%)</td>
<td></td>
</tr>
</tbody>
</table>
shows the association between previous symptoms of GER and the endoscopic appearance, while Table 3 also adds the prerequisite of IM to the endoscopic appearance. Table 2 shows that in this study distinct, obvious tongues of metaplastic columnar epithelium (ZAP II/III) were associated with previous GER, and that no association was found between ZAP I and previous GER, or between CLE $\geq 1\text{ cm}$, and previous GER. Table 3 shows that adding IM to the endoscopic findings did not strengthen the association with previous GER. With the addition of IM, no endoscopic appearance reached a statically significant association with previous symptoms of GER.

**Discussion**

This study is unique in having the possibility to investigate current endoscopic findings with more than 20-year-old, prospectively gathered data regarding symptoms. This makes it possible to explore which present endoscopic appearance of the Z-line are associated with symptoms of GER 20 years ago, something that is of great interest, since both BE and EAC are known to be linked to longstanding GER [20]. The finding of CLE without IM could thus be the result of sampling error, with biopsy sampling being performed too distal in the CLE segment. In this study, the biopsy sampling was made according to a strict protocol and with great caution. The fact that 95% of the biopsies from the Z-line contained both squamous and columnar epithelium, clearly demonstrates that we succeeded in minimizing sampling errors. However, adding IM to the endoscopic findings, did not strengthen the association with previous GER. This could indicate that including IM as a prerequisite for BE might not be necessary, as some expert consensus groups have suggested [6,7], but the present study is not definitive in this regard.

It has been shown that the inter-observer agreement regarding the endoscopic length measurement in the esophagus is poor even under ideal conditions [10]. The mean difference between true length and measured length has been shown to be more than 1 cm. Because of this measurement bias, all definitions that include length measurement in the esophagus will inevitably suffer from this inherited weakness. This is a known limit and one feasible way to bypass this problem is to use a qualitative description instead of a quantitative measurement. This is for instance correct when judging the anti-reflux barrier of the GEJ, where the qualitative Hill classification can be used instead of quantitative measurement of axial length of a hiatal hernia [25].

Regarding the definition of BE one way to avoid the endoscopic length measurement problem could be to use obvious tongues of metaplastic epithelium (ZAP II/III) as the criterion rather than length, something that is supported by the present study. One could argue that by using the presence of metaplastic tongues as an endoscopic criterion, patients with a C value but without M value, according to the Prague-classification, may not be recognized as having BE. In this study however, all subjects with a C value of at least 1 cm, also had tongues of metaplastic epithelium (M value).

Another weakness of the present study is that the ZAP classification was done by one investigator, while the Prague C/M measurements were made by the individual endoscopists. The reason for this was that at the time of designing the study, the plan was to only investigate which ZAP grades could be attributed to longstanding GER. Only after the implementation of the study did it come to our attention that there was a new proposal of a 1 cm limit of the CLE in order to be diagnosed with BE. Since it is not possible to evaluate the Prague C/M length measurements from the video clips, the only way to assess also this finding was to use the measurements done by the individual endoscopists. Even though the Prague C/M classification did not show any association with previous GER, it is of great importance in describing the extent of overt BE, especially as a tool to try...
to map changes over time, for example in the context of endoluminal therapies. In that context, a post hoc analysis of a cohort of BE-patients with a median columnar metaplasia of C3M5, has shown the agreement between endoscopist to be good, especially in segments longer than 5 cm [26].

In conclusion, this study indicates that a definitive endoscopic alteration of the Z-line that can be linked to long-standing GER is the presence of obvious tongues of metaplastic columnar epithelium (ZAP II/III). This could be of importance regarding endoscopic requirements for the diagnosis of BE.

Acknowledgement

The present study was supported by Olympus Solna Sverige AB, who supplied equipment for the realization of the study. The authors have no other affiliation with this company.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The present study was supported by Olympus Solna Sverige AB, who supplied equipment for the realization of the study.

ORCID

Bengt Wallner http://orcid.org/0000-0002-9234-7946

References