This document contains the tables of evidence for the various PICO's.

Taskforce	Topic	Table	Title of PICO
I	Chemoprevention	1	Is there a role for chemoprevention (or even fundoplication) in the prevention of progression of BE?
II	Screening and casefinding	2	Is there an indication for screening for BE?
		3	Is there an indication for case finding for BE?
		4 a–d	What is the role of nonendoscopic or non-inhospital tools (**a** nonendoscopic cell collection devices, **b** breath tests, **c** ultraslim transnasal endoscopy, and **d** videcapsule endoscopy)?
		5	Is there an indication for screening of relatives of 'familiar BE-cancer'?
III	BE surveillance	6	What is the role of chromoendoscopy and virtual chromoendoscopy in the detection and characterization of BE lesions?
		7	What is the role of artificial intelligence in detection and characterization of neoplastic lesions in BE?
		8	What is the impact of inspection time on lesion detaction, and what is the role of dedicated endoscopy lists in Barrett's surveillance?
		9	Should BE surveillance endoscopies be performed under sedation in all BE patients?
		10	What should be the age-limit / stopping criteria for surveillance?
		11	Are the suggested intervals (stratified to BE length) still valid?
IV	Pathology sampling / risk stratification	12	Is there a role for nonendoscopic devices (Cytosponge, EsoCheck, EsophaCap), breath analysers and blood biomarkers for risk stratification of BE?
		13	What is the validity of the Seattle biopsy protocol in BE surveillance?
		14	What is the additional value of p53 in the histopathological assessment of BE biopsies and resection specimen?
		15	Is there a role for WATS 3D brushing as an alternative or as an adjunct to random BE biopsies?
		16	What is the role of biomarkers (e.g. Cernostics) either or not in conjunction to brushing?
		17	Is tumor budding a relevant histopathological parameter in the assessment of endoscopic resection specimen?
V	Treatment	18	How should HGD in BE be treated?
		19	Should we ablate all patients with (confirmed) LGD or is surveillance a valid alternative?
		20	Should we ablate all patients after endoscopic resection of lesion, or is close follow-up also an option?
		21	What is the position of new ablation techniques?
		22	How should low-risk T1a BE cancer be treated?
		23	What is the optimal endoscopic resection technique for BE lesions? ESD vs EMR
		24	How should low-risk T1b BE cancer be treated? Endoscopically, or is there still a role for surgery?
		25	What is the optimal treatment algorithm for high-risk T1a BE cancer?
		26	What is the optimal treatment algorithm for high-risk T1b BE cancer?
VI	Management after treatment	27	What is the most optimal medical therapy after EET?
		28	Should endoscopic follow-up after EET be centralized?
		29	Biopsy protocol after EET: When and where to take biopsies (SCJ, neosquamous epithelium)?
		30	What is the definition of recurrence after EET?
		31	What is the meaning of IM in random biopsies of normal-looking Z-line after EET?
		32	What are the most optimal intervals for post EET surveillance?
		33	When to stop posteradication surveillance?

Table 1	**Chemoprevention**											
Author (year)	**Methods**			**Population**				**Intervention**			**Outcomes**			
	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	Remarks	Notes
Jankowski, 2018	RCT	Yes	Yes	2557	>18 y/o	New and existing BE defined as at least 1 cm of histologically proven columnar-lined oesophagus	Patients with BE were randomised 1:1:1:1 using a computer-generated schedule held in a central trials unit to receive high-dose (40 mg twice-daily) or low-dose (20 mg once-daily) PPI, with or without aspirin (300 mg per day in the UK, 325 mg per day in Canada) for at least 8 years, in an unblinded manner		8.9 years (IQR 8.2–9.8)	all-cause mortality (time ratio)	High-dose PPI was superior to low-dose PPI (TR 1.27, 95% CI 1.01–1.58, p=0.038). Aspirin was not significantly better than no aspirin (TR 1.24, 0.98–1.57, p=0.068). If patients using non-steroidal anti-inflammatory drugs were censored at the time of first use, aspirin was significantly better than no aspirin (TR 1.29, 1.01–1.66, p=0.043; n=2236). Combining highdose PPI with aspirin had the strongest effect compared with low-dose PPI without aspirin (TR 1.59, 1.14–2.23, p=0.0068). NNT=34 for PPI vs NNT=43 for AAS	AspECT trial Lancet	
Chen, 2021	Systematic Review and Meta-analysis	N.a.	N.a.	155769	?	Articles can be regarded as eligible if they met the following criteria. (1) The research type was a cohort or case-control design. (2) The subjects of the study were patients with Barrett's esophagus from any region, country or race. (3) The outcome index was the incidence of EAC or HGD. (4)	PPI	Without PPI	N.a.	progression to HGD/EAC (OR)	OR=0.47, 95% CI 0.32 - 0.71	high heterogenity between studies; possible confounders (over-counter medication);dosis of PPI not reported; dysplasia at baseline not reported.	
Krishnamoorthi, 2018	Systematic Review and Meta-analysis	N.a.	N.a.	74943	?	(1) the study population consisted of BE subjects with NDBE and/or LGD; (2) reported progression to HGD and/or EAC as an outcome; (3) reported factos associated with progression from NDBE/LGD to HGD/EAC as na outcome; and (4) reported measures of association between risk factors and progression as hazard's ratio, relative risk, or OR with 95% CI	PPI/NSAIDs/statin	without PPI/NSAIDs/statin	N.a.	progression to HGD/EAC (OR)	OR=0.43, 95% CI 0.26–0.70	possible confounders; small sample size of some studies	

Singh, 2014	Systematic Review and Meta-analysis	N.a.	N.a.	2813	?	We included RCTs or observational studies that met the inclusion criteria: evaluated and clearly defined exposure to PPIs or H2RAs; reported OAC and/or BOHGD risk in patients with established BO; and reported HR, RR or OR, or provided data for their calculation.	With acid suppressive medications (PPI or H2RA)	without	N.a.	progression to HGD/EAC (OR)	PPI use was associated with a 71% reduction in risk of OAC and/or BO-HGD in patients with BO (adjusted OR 0.29; 95% CI 0.12 to 0.79). NNT 147 HR2A use was not associated with significant reduction of risk	Population: BE (85% non-dysplastic); high heterogeneity; possible confounders; all observational studies, with some not reporting dysplasia at baseline
Li, 2020	Systematic Review and Meta-analysis	N.a.	N.a.	7053	?	Studies that reported HR, RR or OR, or provided data for their calculation.	PPI	Without PPI	N.a.	progression to HGD/EAC (OR)	OR=0.61, 95% CI 0.29-1.26	high heterogenity; dosis of PPI not reported; publication bias not analysed
Hu, 2017	Systematic Review and Meta-analysis	N.a.	N.a.	5712	> 18 years	research object being patients with BE, study compared differences in incidence of HGD or EAC after taking PPIs vs. not	PPI	Without PPI	N.a.	progression to HGD/EAC (OR)	OR=0.43, 95% CI 0.17-1.08	Population: BE (included dysplasia on baseline); high heterogeneity between studies; possible confounders
Zang, 2014	Systematic Review and Meta-analysis	N.a.	N.a.	5446	?	(1) evaluated exposure to any type of COX inhibitors; (2) the primary outcome was clearly defined as EAC or high-grade	COX inhibitors or AAS	without COX inhibitors or AAS	N.a.	progression to HGD/EAC (RR)	COX inhibitors: RR=0.64, 95% CI 0.53-0.77 AAS: RR=0.63, 95% CI 0.43-0.94	Possible confounders
Heath, 2007	RCT	Yes	?	100	> 18 years	dysplasia (HGD); (3) patients included should be definitely diagnosed as having BE in the past or at present; (4) provided	Patients were randomly assigned to treatment with 200 mg of celecoxib or placebo, both administered orally twice daily, and then stratified by grade of dysplasia.		2 yrs	proportion of biopsy samples exhibiting dysplasia (IQR)	After 48 weeks of treatment, no difference was observed in the median change in the proportion of biopsy samples with dysplasia or cancer between treatment groups in either the low-grade (median change with celecoxib = − 0.09, interquartile range [IQR] = − 0.32 to 0.14 and with placebo = − 0.07, IQR = − 0.26 to 0.12; P = .64) or high-grade (median change with celecoxib = 0.12, IQR = − 0.31 to 0.55, and with placebo = 0.02, IQR = − 0.24 to 0.28; P = .88) stratum.	Possible confounders such as over-the-counter medication; small sample size; short time of follow-up

PICO search string:	(Barrett's esophagus[Title/Abstract] OR Barrett[Title/Abstract] OR columnar-lined esophagus[Title/Abstract]) AND (Prophylaxis[Title/Abstract] OR Prevention[Title/Abstract] OR Chemoprevention[Title/Abstract] OR Proton-pump inhibitor[Title/Abstract] OR Aspirin[Title/Abstract] OR Statin[Title/Abstract] OR Metformin[Title/Abstract] OR Ursodeoxycholic acid[Title/Abstract]) AND (Cancer[Title/Abstract] OR Adenocarcinoma[Title/Abstract] OR Dysplasia[Title/Abstract] OR Neoplasia[Title/Abstract] OR Survival*[Title/Abstract] OR Death[Title/Abstract] OR Mortality[Title/Abstract])

Table 2 — **Is there an indication for screening for BE**

Author (year)	Methods			Population				Screening tool	Intervention			Outcomes		Remarks	Notes
	Design	Randimsation / blinding		N	Age	Inclusion criteria			Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest		
Curtius (2021)	Markov models	na	na	na	na	na		endoscopy	Multi- state Markov models have most commonly been used to microsimulate large cohorts of individuals in silico from birth through clinical states of carcinogenesis. Model rates are inferred by fitting to cancer incidence and mortality data, or fitting to screening trial data explicitly. Then with a calibrated model, screening processes are built into microsimulations to compare strategies based on beneficial outcomes from interventions and to quantify associated costs such as numbers of expected surveillance exams. Two main validation datasets that are relevant for our optimization problem: Clinical Outcomes Research Initiative AND prospective, randomized control trial (BEST3) BE prevalence data for patients who were screened with sensitive technology Cytosponge-TFF3	na	na	optimal age to start screening, costs, number of endoscopies	For w = 1 (equal weighting of positive screen and safeguarding from cancer before screening), the optimal screening times were 64 years old for all males, 58 years old for males with GERD symptoms, 69 years old for all females, and 64 years old for females with GERD symptoms. if we screened x = 400,000 GERD men from the 1950 birth cohort, the number of futile endoscopies due to screening at optimal age 58 were estimated to be 24,000 less than expected with current practice for screening at 50. Assuming an EGD costs $745 USD, this implies a savings of $17.9 million for this birth cohort alone when choosing the optimal age to start screening recommendation versus current practice. For both screening ages 50 and 58 in GERD men, there was less than 0.4% risk of screening too late. If we screened the roughly x = 400, 000 US men with GERD from the 1950 birth cohort at-risk at those specific ages, the overall costs associated with screening at optimal age 58 were estimated to be 64,000 patient-years less spent on futile surveillance for this birth cohort alone than what would be expected for screening at age 50. And the number of futile endoscopies due to screening at optimal age 58 were estimated to be 24,000 less than expected with current practice for screening at 50.		
Singer (2021)	cost-effectiveness analysis	na	na	na	na	na		Upper endoscopy	WATS3D+Seattle protocol The reference case was a 60-year-old white male with gastroesophageal reflux disease. Cost was measured in 2019 US$, and the incremental cost-effectiveness ratio (ICER) was measured in $/QALY using thresholds for cost-effectiveness of $100,000/QALY and $150,000/QALY Effectiveness was measured by the number needed to screen to avert one cancer and one cancer-related death, and quality-adjusted life years (QALYs).	Seattle protocol only	na	Number needed to screen, cost, incremental cost-effectiveness ratio (ICER)	NNS: Between 320 and 337 people would need to be screened with WATS3D in addition to FB to avert one additional cancer, and 328–367 people to avert one cancer-related death. Screening with WATS3D costs an additional $1219 and produced an additional 0.017 QALYs, for an ICER of $71,395/QALY. All one-way sensitivity analyses resulted in ICERs under $84,000/QALY.		
Peters (2020)	Questionnaire	na	na	375	61.7	1000 subjects aged 50 to 75 years who were randomly selected from the municipal registry in the Netherlands.		upper endoscopy, transnasal endoscopy, capsule sponge and breath testing	Discrete choice experiment questionnaire was sent by postal mail. Each subject answered 12 discrete choice questions of 2 hypothetical screening tests comprising 5 attributes: EAC-related mortality risk reduction, procedure-related pain and discomfort, screening location, test specificity, and costs.	na	na	secreening acceptance discrete choice experiment (DCE) quantitative insights	Comparing different screening tests, uptake was 63% for upper endoscopy screening, 77% for transnasal endoscopy, 76% for capsule sponge testing, and 70% for breath testing, with a test specificity of 80%. The average expected uptake of EAC screening was 62.8% (95% confidence interval [CI] 61.1-64.5). A total of 17 (4.5%) respondents consequently chose never to be tested. On the contrary, 200 (53.3%) respondents consistently selected screening. In a subgroup analysis for of the remaining 158 (42.1%) respondents who did not consistently select screening, the effect of the attribute mortality reduction was much smaller compared with the group as a whole. Test specificity was the most important item for EAC screening, accounting for 27.2% of decision making, followed by pain and discomfort (importance score of 26.8%), and mortality reduction (importance score of 24.6%).	patients were randomly selected from the municipal registry in the Netherlands	

Study	Type			N	Age	Population/Inclusion	Screening tool	Methods	Comparator	Duration	Outcome	Results
Hammad (2019)	cohort	na	na	182	67.3	all patients diagnosed with EAC between 02/2005 and 09/2017	Esophagogastroduo denoscopy	We retrieved all cases of esophageal cancer from the institution's cancer registry, and then selected only EAC patients for a structured manual review of their entire VA electronic medical records though the Computerized Patient Record System. From endoscopy and pathology reports for the period from the patient's first VA encounter to their EAC diagnosis date, we collected information on the presence of BE (yes, no) and dysplasia status. We defined "proper BE surveillance" if a patient underwent appropriate surveillance EGD for BE based on findings from previous endoscopies (e.g. within 3–5 years after diagnosis of BE with no dysplasia). We defined missed opportunities for screening and surveillance of BE as: 1) failure to screen for BE in high risk patients, defined as male patients aged >50 years with GERD and at least one of obese, former or current smokers, or white ethnicity, 2) failure to diagnose BE in those that had a prior endoscopy, and 3) improper BE surveillance and treatment.	na	not stated	BE diagnosis previous to EAC BE surveillance Subgroup analysis by context of EAC (on BE sureveillance or not)	Only 45 patients (24.7%) had EGD at any time prior to the cancer diagnosing EGD, of whom 29 (15.9% of all EAC cases) had an established BE diagnosis. For the 29 patients with prior established BE, 22 (75.8%) were diagnosed with EAC as a result of surveillance EGD. Patients with prior established BE were more likely to be diagnosed at 0 or I stage (p<0.001) and managed with endoscopic or surgical modalities (p<0.001) than patients without prior BE.
Gupta (2014)	Population-Based Survey	na	na	136	63	A population-based survey was undertaken in 2008 to randomly sample a community of adults aged 50 years and older	VCE or uTNE or sEGD	The survey consisted of questions that assessed awareness of BE and its premalignant potential, willingness to participate in screening for BE/EAC, baseline acceptance of screening tests, preferences regarding optimal method of screening for BE, influence of a personal or family history of cancer on choosing to be screened, potential barriers to participation in population screening for BE/EAC, and demographic information. VCE or uTNE or sEGD	na	na	interest in screening screening tool preferences	72 % were interested in screening A history of undergoing screening tests and GI symptoms were predictive of interest. Unsedated techniques were preferred by 64 % (VCE: 56 % and uTNE: 8 %) versus sEGD (36 %)
Agha (2021)	cohort	na	na	108		All adults (>18 years of age) who underwent screening or routine surveillance for BE with both WATS3D and traditional cold FB between January 2015 and January 2019	WATS3D	After careful screening using white light and narrow-band imaging (NBI), FB were obtained every 1-2 cm in 4-quadrants along the length of the BE segment, followed by WATS3D brushings of the BE segment. WATS3D brushings were done after FB in all cases.	systematic 4-quadrant forceps biopsies	4 years	BE detection Agreement between tools in BE detection and exclusion	FB and WATS3D detected 62 (57.4%) and 83 (76%) cases of BE, respectively. The absolute difference of 21 cases (18.6%) of BE was attributed to the addition of WATS3D (BE missed by FB). The number needed to test with WATS3D was 5. Overall agreement by kappa statistic was 0.74
Sami (2021)	cost-effectivene ss analysis	na	na	500,000	na	na	sedated endoscopy, transnasal endoscopy, swallowable esophageal cell collection devices with biomarkers, and exhaled volatile organic compounds	Markov modeling 2 scenarios/contexts: (i) White men with chronic GERD (GERD-based); (ii) GERD-independent (all races, men and women), BE prevalence 1.6%; and (iii) GERD-independent, BE prevalence 5%. Screening tools: sedated endoscopy (sEGD), transnasal endoscopy, swallowable esophageal cell collection devices with biomarkers, and exhaled volatile organic compounds Willingness to pay threshold of $100,000 per qualityadjusted life year (QALY)	no screening	40 years	Incremental cost-effectiveness ratios (ICERs)	Swallowable esophageal cell collection devices with biomarkers were cost effective (<$35,000/QALY) and were the optimal screening tests in all scenarios. Exhaled volatile organic compounds had the highest ICERs in all scenarios. ICERs were low (<$25,000/QALY) for all tests in the GERD-based scenario The ICER for sEGD, SoS, hTNE, eVOC, and mTNE crossed the WTP threshold at BE prevalence values below 1.4%, 0.4%, 0.6%, 0.7%, and 0.2%, respectively
Gerson (2009)	cohort	na	na	126	42 - 56	Assymptomatic women scheduled for routine screening colonoscopy (for colorectal cancer) and women undergoing endoscopic examination before bariatric surgery	endosedated endoscopy	Outpatients at Stanford University and Palo Alto VA Health Care System. Biopsies of the esophagogastric junction in the setting of suspected BE, and completion of symptom and health-related quality of life questionnaires to ensure that subjects were asymptomatic.	no comparison	na	Identification of BE	BE in 8 (6%) of 126 subjects, including 3 (5%) of 61 of the women in the colorectal cancer screening cohort and 5 (8%) of 65 of the women in the pre–bariatric surgery cohort
Moriarty (2018)	RCT	Randomiz ed	unknwon	209	50 years of age or older	residents 50 years of age or older, without a known diagnosis of BE or EAC, who had previously completed a GI symptom questionnaire) were recruited	unsedated transnasal endoscopy (uTNE), sedated endoscopy (sEGD) and mobile research van (muTNE)	Patients were called 1 and 30 days after screening to assess loss of work (due to the screening procedure) and medical care sought after procedure. Direct medical costs were extracted from billing claims databases. Indirect costs (loss of work for subject and caregiver) were estimated using patient reported data. Screening tools: unsedated transnasal endoscopy (uTNE), sedated endoscopy (sEGD) and mobile research van (muTNE)	unsedated transnasal endoscopy (uTNE), sedated endoscopy (sEGD) and mobile research van (muTNE)	30 days	direct medical costs and indirect cost	30 day direct medical costs and indirect costs were significantly higher in the sEGD than the huTNE and muTNE groups. Total costs (direct medical + indirect costs) were also significantly higher in the sEGD than in the uTNE group. The muTNE group had significantly lower costs than the huTNE group.

Study	Type			N	Age	Inclusion criteria / Population	Test	Methods	Comparison		Outcome	Results
Qumseya (2021)	SR and MA	na	na	680	?	We included studies on patients who had undergone esophagogastroduodenoscopy (EGD) after SG.	na	We searched the major search engines ending in July 2020. The primary outcome was the prevalence of BE in patients who had undergone SG. We assessed heterogeneity using I2 and Q statistics.	na	na	Prevalence of BE in patients who had undergone SG.	The pooled prevalence of BE was 11.6% (95% confidence interval [CI], 8.1%-16.4%). The prevalence of BE was assessed in the general or low-risk populations and in those individuals with risk factors such as GERD, family history of BE or EAC, age >50 years, obesity, and male gender. When controlling for population Western versus non-Western, mean age, and gender distribution, there was a linear relationship between the number of risk factors and the risk of BE. As the number of risk factors increased, the risk of BE increased by 1.2% per additional risk factor
Marques de Sa (2020)	SR and MA	na	na	?	?	Inclusion criteria were original full-text articles published up to September 2018 and addressing BE in general and the GERD population, that met the globally accepted criteria for BE	na	Three databases were searched. Subgroup, sensitivity, and meta-regression analyses were conducted and pooled prevalence was comput.	na	na	Prevalence in general and GERD population	The pooled prevalence for general population: 0.96% (95% confidence interval: 0.75–1.18). The pooled prevalence for GERD was 6.72% (95% confidence interval: 3.61–9.83)
Stephanie (2022)	SR and MA	na	na	5137	?	Inclusion criteria: (i) original articles addressing prevalence and/or progression of BE from non-dysplastic BE or LGD to HGD or EAC in a general or GERD population, (ii) articles that defined BE as salmoncolored mucosa in the tubular esophagus with length ≥1 cm and histologic confirmation of IM, and (iii) reported outcomes in both women and men.	na	A comprehensive search was conducted using PubMed, Scopus, and Google Scholar. Studies were included that reported prevalence rates of BE or progression rates to neoplastic disease stratified by gender.	na	na	BE prevalence by gender and risk of progression by gender	The rate of BE in women was 1.29% ([95% CI: 0.76–2.19], I2=91%) compared to men at 4.66% ([95% CI: 3.31–6.53], I2=89%); OR: 0.33 ([95% CI: 0.27–0.42], I2=0%). The rate of annual progression of Barrett's to high-grade dysplasia or adenocarcinoma was 0.62% ([95% CI: 0.22–1.75]) in women compared to 1.54% ([95% CI: 0.83–2.81], I2=96%) in men; OR: 0.44 ([95% CI: 0.30–0.65], I2=22%).
Sharma (2017)	cohort	na	na	1378	62.27	The inclusion criteria for our population comprised Chinese patients over the age of 50 years with one or more of the following additional criteria: history of dyspepsia for at least 4 weeks; family history of gastric cancer; and any medical condition for which an upper gastrointestinal (GI) endoscopic investigation was warranted.	endoscopy	Patients were prospectively enrolled in this study. The questionnaire utilized to assess risk factors and symptoms was designed by two experts.	A separate group of subjects who underwent an esophagogastroduodenoscopy (EGD) during the same period for various indications, with no prior history of GERD, and who were found to have a normal looking esophagus at endoscopy, with minimal gastric pathology were recruited as controls.	?	risk factors for intestinal metaplasia	Age (odds ratio 1.081, 95% confidence interval 1.022–1.143) and male gender (odds ratio 4.808, 95% confidence interval 1.727–13.33) proved significant demographic factors for the presence of intestinal metaplasia
Gupta (2011)	cost-effectiveness analysis	na	na	na	na	50 y/o undergoing screening colonoscopy	endoscopy	No screening 18.08 $480 Reference EGD Screening, EET for HGD 18.08 $933 Dominated Ext EGD Scr & Surv of NDBE, EET for HGD 18.08 $961 $95,590/QAL	na	na	Optimal strategy	Optimal strategy: Scr & Surv of NDBE, EET for HGD
Benaglia (2013)	cost-effectiveness analysis	na	na	na	na	50 y/o man with GERD	endoscopy or Cytosponge	No screening 17.96 $132 Reference Cytosponge Scr, EGD Surv, EET for HGD 17.98 $373 $15,724/QALY EGD Scr, EGD Surv, EET for HGD 17.98 $431 Dominated	na	na	Optimal strategy	Optimal strategy: Cytosponge Scr, EGD Surv, EET for HGD
Heberle (2017)	cost-effectiveness analysis	na	na	na	na	60 y/o man with GERD	endoscopy or Cytosponge	No screening 15.08 $762 Reference Cytosponge Scr, EGD Surv, EET for HGD 15.10 $1485 $33,057/QALY EGD Scr, EGD Surv, EET for HGD 15.10 $2090 $330,361/QALY	na	na	Optimal strategy	Optimal strategy: Cytosponge Scr, EGD Surv, EET for HGD
Honing (2018)	cost-effectiveness analysis	na	na	na	na	50 y/o white man with GERD	mTNE, endoscopy or uTNE	No screening 18.43 $1436 Reference mTNE Scr, EGD Surv, EET for LGD 18.47 $2425 $29,446/QALY uTNE Scr, EGD Surv, EET for LGD 18.47 $2494 Dominated EGD Scr, EGD Surv, EET for LGD 18.47 $2957 Dominated	na	na	Optimal strategy	Optimal strategy: mTNE Scr, EGD Surv, EET for LGD

Author (year)	Study type			N	Age	Population	Intervention	Description	Comparator		Outcome measures	Results	Limitations
Baldwin-Hunter (2019)	cohort	na	na	2931	62.15	Patients ages 50–75 who underwent a first-time esophagogastroduodenoscopy	endoscopy	The first was the development of an a priori set of criteria, modified from clinical guide-lines, to identify patients likely to have BE. The purpose of this approach was to create a simple, highly efficient checklist which could be used to rapidly target patients for BE screening. The second approach was to use logistic regression modeling to identify EHR-based factors which predict BE. This information could then be used to generate a risk score based on the association of individual factors with BE.	na	?	risk factors for BE	Subjects who met screening criteria were more likely to have BE (3.3% vs. 1.1%, p = 0.001), and the criteria predicted BE with an AUROC of 0.65 (95% CI 0.59–0.71). A score based on logistic regression modeling included gastroesophageal reflux disease, sex, body mass index, and ever-smoker status and identified BE subjects with an AUROC of 0.71 (95% CI 0.64–0.77).	
Rubenstein (2007)	Markov model	na	na	na	na	50-year-old white men with GERD symptoms who were followed until age 80 years or death	endoscopy and capsule	A natural history or unscreened strategy was modeled for reference purposes as well as 2 screening strategies, conventional endoscopy and esophageal capsule endoscopy (ECE). In the ECE strategy, a positive test result would lead to a follow-up EGD. All patients confirmed to have BE by biopsy would undergo surveillance with conventional endoscopy for the remainder of the analysis. Although the risks of complications from EGD were included in the model, ECE was assumed to be without risk. The accuracy of EGD was modeled to be the same, regardless of whether it was an initial screening EGD or for the follow-up of a positive ECE study	na	na	indirect costs such as lost productivity from work	EGD screening prevented 60% of cancer deaths at a cost of $11,254 per quality adjusted life year gained compared with no screening. ECE prevented 53% of cancer deaths and provided 9 fewer quality-adjusted days at greater cost than EGD. ICER of $11,254 per QALY gained compared with no screening. The conventional endoscopy strategy was superior to ECE in every outcome analyzed, including unadjusted life years, quality-adjusted life years, cost, and cancer deaths prevented. The primary difference between the 2 strategies was the accuracy of the initial diagnostic procedure to detect BE. Although endoscopy was modeled to have superior diagnostic accuracy, because the prevalence of BE in this hypothetical cohort of patients is low (10%) and the percentage of patients who develop esophageal cancer is even lower, it is not surprising that the differences in outcomes between the 2 strategies were minimal.	limitation: exclusion of any possible adverse events from the ECE procedure
Gerson (2007)	Markov model	na	na	na	na	cohort of 50-year old men with chronic GERD for the presence of BE	endoscopy and capsule	GOAL: We compared the base-case strategy of no screening for BE to 2 competing screening strategies: (1) ECE followed by upper endoscopy (EGD) if BE were suspected or if there was poor visualization on the ECE; and (2) standard sedated EGD with biopsy The ability of an EGD to detect BE improved from 85% for an initial screening study to 95% when it was performed after a positive ECE, presumably because the knowledge of a prior positive study affected the endoscopy This study used a simple decision tree model design. patients with either a positive or inconclusive (inadequate visualization of the gastroesophageal junction) ECE study were modeled to undergo a follow-up EGD	na	na	number of BE cases detected, cost and endoscopic complication	Initial EGD was more expensive but more effective compared with the no screening strategy Conventional endoscopy was superior to ECE when using BE detection as an end point and was also less costly: Assuming a theoretical cohort of 10,000 patients with GERD, initial EGD cost $1988 and was associated with 18.54 life-years compared with $2392 and 18.36 life-years for the ECE arm and $901 and 18.30 life-years for the no screening arm. The incremental cost-effectiveness ratio of screening with EGD compared with the no screening arm was $4530 per life-year gained.	limitation: exclusion of any possible adverse events from the ECE procedure
Elsheaita (2020)	cohort	na	na	100		Inclusion criteria were clinical diagnosis of GERD, presence of endoscopic esophagitis and at least one of the following criteria: 1. Chronic GERD (more than 5 years). 2. Frequent GERD (weekly or more). 3. Two or more risk factors for BE or EAC: Old age >50 years, central obesity (waist circumference >102cm in males and >88cm in females), current or past history of smoking and family history of BE or EAC.	endoscopy	NBI	Seattle protocol only	na	Sensitivity, specificity, negative predictive value and positive predictive value	Sensitivity, specificity, negative predictive value and positive predictive value for Seattle protocol were 58.8%, 100%, 92.2%, 100% vs 76.5%, 100%, 95.4%, 100% respectively for narrow band imaging. A mean of 7.73 samples/patient was taken in Seattle protocol vs 3.42 samples in narrow band imaging (P<.001). A mean of 8.63 minutes was consumed in Seattle protocol vs 2.65 minutes in narrow band imaging (P<.001).	

Study	Design	Randomization	Blinding	N	Age	Population	Index	Methods	Comparator	?	Outcomes	Results	Limitations
Rubenstein (2020)	cohort	na	na	1241	58.3	Patients, ages 40–79 y, presenting either for their first esophagogastroduodenoscopy (EGD) or their first endoscopic therapy of early neoplastic BE, from April 2015 through June 2018	scores for identifying patients to be screened	We calculated risk scores for 6 previously published tools (the Gerson, Locke, Thrift, Michigan BE pREdiction Tools [M-BERET], Nord-Trøndelag Health Study [HUNT], and Kunzmann tools). We also investigated the accuracy of frequency and duration of gastroesophageal reflux disease (GERD), using data from a randomly selected 50% of patients undergoing their first EGD. We compared the ability of all these tools to discriminate patients with BE and early neoplasia from patients without BE, using findings from endoscopy as the reference standard. Gerson: 7-item questionnaire that queries severity of GERD and upper abdominal symptoms, plus race and ethnicity HUNT: age, sex, GERD symptoms in the prior 12 months, body mass index, and smoking Kunzmann: age, sex, body mass index, smoking, and esophageal conditions (GERD, Barrett's esophagus, hiatal hernia, esophageal stricture, fundoplication, or acid reducing medications). M-BERET: age, waist-to-hip circumference ratio, and an 8-item questionnaire querying typical frequency of heartburn or regurgitation without the use of acid reducing medications and a crude estimate of pack-years of cigarette smoking	presence of GERD	?	AUROC	All of the tools were more accurate in identifying patients with BE than the frequency and duration of GERD (AUROC for GERD, 0.579 vs range for other tools, 0.660–0.695), and predicted risk correlated well with observed risk (calibration). The AUROCs of the HUNT tool (0.796), the M-BERET (0.773), and the Kunzmann tool (0.763) were comparable in discriminating between patients with early neoplasia (n=94) vs no BE.	
Sami (2019)	cohort	na	na	200	57.9	Consecutive adult patients referred for clinical C-EGD with and without histologically confirmed BE of any length were invited to participate	endoscopy and capsule	All subjects underwent the 2 procedures on the same day, performed by blinded endoscopists. Patients completed preference and validated tolerability (10-point visual analogue scale [VAS]) questionnaires within 14 days of the procedures.	na	?	preferences, sensitivity, specificity of test	178 (89%) completed both procedures (11% failed EG Scan due to the inability to intubate the nasopharynx). A higher proportion of patients preferred the EG Scan (54.2%) vs the C-EGD (16.7%) (P<.001) EG Scan had a higher VAS score (7.2) vs the C-EGD (6.4) (P=.0004). The EG Scan identified any length BE with a sensitivity value of 0.90 (95% CI, 0.83–0.96) and a specificity value of 0.91 (95% CI, 0.82–0.96). The EG Scan missed 9 cases of BE, 2 of those were LSBE.	
Gerson (2004)	Markov model	na	na	na	na	cohort of 50-year-old men with chronic symptoms of gastroesophageal reflux.	endoscopy	We compared costs and life expectancy for a cohort of 50-year-old male patients with heartburn undergoing endoscopic screening for the presence of Barrett's esophagus compared with a cohort of patients with reflux who would not undergo endoscopic screening. Standard sedated videoendoscopy with biopsy was performed for all patients in the screening arm.	no screening and surveillance	na	incremental cost-effectiveness ratio; sensitivity analysis	The cost-effectiveness ratio of screening and surveillance compared with no screening was $12,140/life-year gained. Parameters associated with the greatest variation in incremental cost-effectiveness ratios included age at time of initial endoscopic screening, cost of endoscopy, and resectability of esophageal cancer. The screening with surveillance of patients with nondysplastic Barrett's esophagus was no longer cost-effective if the prevalence of Barrett's esophagus was <1%, with a cost-effectiveness ratio of $356,500/life-year gained. At a prevalence rate of 5%, the cost-effectiveness ratio was $23,730, and at a prevalence of 20%, the cost-effectiveness ratio was $7630.	
Shariff (2012)	RCT	Randomized	Blinded (endoscopist and pathologist)	82	60	patients with BE or those referred for diagnostic assessment	Unsedated transnasal endoscopy (TNE) and standard endoscopy (SE)	All patients underwent TNE followed by SE or the reverse. Spielberger State-Trait Anxiety Inventory short-form questionnaires, a visual analogue scale, and a single question addressing preference for endoscopy type were administered.	Unsedated transnasal endoscopy (TNE) and standard endoscopy (SE)	na	Diagnostic accuracy and tolerability	Tsensitivity and The sensitivity and specificity of TNE was 0.98 and 1.00, respectively he mean(+/- standard deviation) post-endoscopy Spielberger State-Trait Anxiety Inventory short-form score for TNE (30.0 1.10 standard error of the mean [SEM]) was lower than that for SE (30.7 1.29 SEM), (P .054). The visual analogue scale scores were no different (P <0.07). The majority of patients (59%) expressed a preference for TNE.	Limitations: This is a small study, with limited generalizability, a high prevalence of patients with BE, differential drop-out between the two procedures, and use of sedation.
Chak (2014)	RCT	Randomized	blinded (only pathologist)	184	58.9	veterans with or without GERD symptoms	Transnasal esophagoscopy (TNE) and Capsule esophagoscopy (ECE).	Eligible participants who agreed to screening were first randomized to either TNE or ECE in a 1:1 fashion using a computer-generated permuted-block randomization with variable blocks of 4 to 6.	Transnasal esophagoscopy (TNE) and Capsule esophagoscopy (ECE).	na	Acceptance of TNE and ECE and tolerability of TNE and ECE and effectiveness of BE screening.	Eleven (12.6%) patients randomized to TNE crossed the minimal clinically important threshold for overall procedure tolerability as opposed to none randomized to ECE (p = 0.001). Effectiveness of BE screening was not significantly different in both procedures (TNE vs. ECE = 3.2% vs. 5.4%, p = 0.47).	

Sami (2015)	RCT	Randomiz ed	blinded (only pathologist)	209	70	Subjects, ≥ 50 years of age who previously completed validated gastrointestinal symptom questionnaires	standad endoscopy (sEGD), or uTNE in a mobile research van (muTNE), or uTNE in a hospital outpatient endoscopy suite (huTNE)	Subjects, ≥ 50 years of age who previously completed validated gastrointestinal symptom questionnaires were randomized (stratified by age, sex and reflux symptoms) to one of three screening techniques (either sEGD, or uTNE in a mobile research van (muTNE), or uTNE in a hospital outpatient endoscopy suite (huTNE)) and invited to participate.	standad endoscopy (sEGD), or uTNE in a mobile research van (muTNE), or uTNE in a hospital outpatient endoscopy suite (huTNE)	na	participation rate	Complete evaluation of the esophagus was similar using muTNE (99%), huTNE (96%) and sEGD (100%) techniques (p=0.08). Mean recovery times (minutes) were longer for sEGD (67.3) compared to muTNE (15.5) and huTNE (18.5) (p<0.001).. Participation rates were numerically higher in the unsedated arms of muTNE (47.5%) and huTNE (45.7%) compared to the sEGD arm (40.7%), but were not statistically different (p=0.27).		
Gupta (2011)	cost-utility analysis	na	na	na	na	50-year-old patients already undergoing screening colonoscopy	endoscopy	Comparison of two strategies: performing and not performing a screening upper endoscopy at the time of screening colonoscopy.	no screening and surveillance	na	Incremental cost-effectiveness ratio (ICER)	One-time screening for the general population at the age of 50 for upper GI cancers required $115,664 per quality-adjusted life year (QALY) compared with no screening or surveillance. A strategy of screening and surveillance for Barrett's esophagus required only $95,559 per QALY saved		
Ross-Innes (2015)	case-control	na	na	1,110	55	In total, 1,110 individuals comprising 463 controls with dyspepsia and reflux symptoms and 647 BE cases swallowed a Cytosponge prior to endoscopy.	cytosponge	We conducted a case–control study to determine the sensitivity and specificity of the Cytosponge-TFF3 test for the detection of BE compared with endoscopy and biopsy as the reference standard	endoscopy	na	Diagnostic accuracy and tolerability	In all, 1,042 (93.9%) patients successfully swallowed the Cytosponge, and no serious adverse events were attributed to the device. The Cytosponge was rated favorably, using a visual analogue scale, compared with endoscopy (p < 0.001), and patients who were not sedated for endoscopy were more likely to rate the Cytosponge higher than endoscopy (Mann-Whitney test, p < 0.001). The overall sensitivity of the test was 79.9%(95% CI 76.4%–83.0%), increasing to 87.2% (95% CI 83.0%–90.6%) for patients with 3 cmof circumferential BE, known to confer a higher cancer risk The specificity for diagnosing BE was 92.4% (95% CI 89.5%–94.7%).		

PICO search string:	(((Screening[Title/Abstract]) AND (Barrett's Esophagus[Title/Abstract])) AND (Barrett's esophagus[Title/Abstract] OR adenocarcinoma[Title/Abstract] OR cancer[Title/Abstract] OR carcinoma[Title/Abstract] OR mortality[Title/Abstract] OR survival[Title/Abstract] OR death[Title/Abstract] OR costs[Title/Abstract] OR cost-effectiveness[Title/Abstract])) NOT (review[Publication Type])

| | | Table 3 | | | **Is there an indication for casefinding in BE** | | | | | |

Author (year)	Methods				Population		Intervention			Outcomes		Remarks
	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	
Sawas, 2021	Retrospective single-center cohort study & prospective multicenter cohort	N.a.	n=663 + n=645	>18 y/o	Patients with EAC.	Patients who do meet the ACG/BSG criteria for BE screening	Patients who do not meet the ACG/BSG criteria for BE screening	N.a.	proportion of preventable EAC defined as patients who met the ACG/BSG criteria for BE screening.	A significant proportion, up to 45%, of patients with EAC did not meet the ACG/BSG criteria for BE screening. Furthermore, a large percentage of patients with the non-BE/IM phenotype who had a poorer survival did not meet screening criteria. Reflux symptoms was found to be an unreliable symptom for screening because they were absent in half of the patients in the Mayo cohort at diagnosis.	Amrican journal of gastroeneterology
Sami, 2021	Markov simulation model	N.a.	N.a.	50 y/o (simulation)	Markov modeling was performed in 3 scenarios in 50 years old individuals: (i) White men with chronic GERD (GERD-based); (ii) GERD-independent (all races, men and women), BE prevalence 1.6%; and (iii) GERD-independent, BE prevalence 5%.	The simulation compared multiple screening strategies with no screening: sedated endoscopy (sEGD), transnasal endoscopy, swallowable esophageal cell collection devices with biomarkers, and exhaled volatile organic compounds.	No screening.	40 years (simulation)	The mean costs and quality-adjusted life years (QALYs) were calculated for each screening approach. These values were used to calculate incremental cost-effectiveness ratios (ICER).The **primary** outcome was to compare the ICER values (cost effectiveness) of the 6 screening tests with no screening to identify whether BE screening could be cost effective. The **secondary** outcome was to compare the ICER values of the 6 screening tests with each other to identify the optimal screening strategy defined as the one providing the most QALYs at a cost less than the willingness to pay (WTP) threshold of $100,000 per QALY.	Screening for BE in a GERD-independent manner (in 50-year-old men and women regardless of race or the presence of GERD symptoms) may be cost effective compared with no screening, particularly when using newer nonendoscopic tests. In the GERD-independent setting, mTNE, Cytosponge + TFF3, and SoS tests were all less costly and more effective than other screening tests. The Cytosponge + TFF3 test was the optimal choice at the lowest prevalence (1.6%), whereas the SoS test was the optimal test in both higher prevalence settings (5%, GERD-independent and 8% GERD-based).	Amrican journal of gastroenterology
Rubenstein, 2021	Systematic review	N.a.	N.a.	50 y/o white man with GERD (simulation)	Searching in PubMed from 1/2016 to 2/2020 with the following Medical Subject Headings: [Barrett esophagus or esophageal neoplasms] and [Health Care Costs or "Costs and Analysis" or Cost-Benefit Analysis or CEA.mp]. Eligible studies were published in English and presented results of a CEA comparing strategies of screening with outcomes expressed as quality-adjusted life-years (QALYs).	N.a.	N.a.	N.a.	Cost-effectiveness analysis comparing strategies of screening.	6 comparative models were found about BE screening. Screening for Barrett's esophagus and esophageal adenocarcinoma in men with symptoms of gastroesophageal reflux is a cost-effective strategy. Screening modalities that do not require sedation (eg, cytosponge, transnasal endoscopy) may be more cost-effective than screening with standard, sedated endoscopy.	Gastrointestinal Endoscopy Clinics of North America
Nguyen, 2021	retrospective analysis of a prospective cross-sectional study	N.a.	n=513	50-80 y/o	Study eligibility was based on the following: (1) age between 50 and 80 years; (2) no previous or current gastroesophageal surgery or cancer; (3) no active lung, colon, or breast cancer; (4) no current use of anticoagulants, which would be a relative contraindication for mucosal biopsy sampling during endoscopy; (5) no significant liver disease indicated by platelet count <70,000, ascites, or known gastroesophageal varices; and (6) no history of major stroke or mental condition that would limit the ability to answer questions.	Individuals who were asymptomatic (for GERD)	Entire study population	N.a.	We calculated the prevalence of BE as a proportion in the entire study population and the subgroup of individuals who were asymptomatic (for GERD). We compared cases versus control subjects in the overall population, cases versus control subjects among asymptomatic only, and symptomatic cases versus asymptomatic cases with respect to several sociodemographic and clinical factors. We examined the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive and negative likelihood ratios, area under the receiver operating characteristic curve (AUROC), and 95% CIs of diagnosing BE using the screening guidelines of 5 gastroenterological societies	Our data show that conditioning BE screening guidelines on the presence of GERD symptoms (or the lack of conditioning in the case of the AGA) comes at a heavy price as far as the performance of these screening guidelines. BE occurred in a large proportion without frequent GERD symptoms (56.8% of all BE cases in our study), which would lead to an under-diagnosis of BE if guidelines recommend that only those with frequent GERD symptoms are referred for screening. We found screening all those with frequent GERD symptoms and those without GERD symptoms but with 5 or more risk factors would render the best possible performance for the entire 2-pronged approach (sensitivity of 81.8%, specificity of 51.2%, and AUROC of .66). Additionally, it seems that with regard to specificity and PPV in this study population, the most important risk factors after GERD are non-Hispanic white race, followed by obesity, and then male sex or age ≥50 years. However, in a meta-analysis of over 13,400 obese patients undergoing preoperative EGD, the prevalence of BE was only .9%, and so obesity alone may not be a major risk factor for BE.	Gastrointestinal Endoscopy

Author, Year	Study type		n	Age	Population / Inclusion criteria	Index	Comparator		Methods / Outcome	Results / Conclusion	Journal
Rubenstein , 2020	Prospective cohort study	N.a.	n=1241	40-79 y/o	Individuals, aged 40–79, presenting for their first EGD and those presenting for their first endoscopic management of suspected early neoplastic Barrett's esophagus at either the University of Michigan (UM) or the Ann Arbor Veterans Affairs Medical Center (AAVA) from April 2015 through June 2018.	6 tools for predicting Barrett's esophagus and discriminating early neoplasia.	GERD alone for predicting Barrett's esophagus and discriminating early neoplasia.	N.a.	Tools were compared for discrimination using the AuROC and the optimal threshold for each tool was identified at the maximum of the Youden's index (the point on the receiver operating characteristic curve where sensitivity + specificity – 1 is maximized). We also calculated the Integrated Discrimination Improvement, which is similar to the Net Reclassification Improvement but integrates over all possible thresholds.	All of the tools were more accurate than GERD symptoms alone for predicting Barrett's esophagus. For discriminating early neoplasia, the HUNT tool, M-BERET, and Kunzmann tool were particularly accurate.	Gastroenterology
Hamel, 2020	Systematic review	N.a.	N.a.	>18 y/o	Adults (≥ 18 years old)a with chronic gastroesophageal reflux disease (GERD)b with or without other risk factorsc for esophageal adenocarcinoma (EAC).	Screening +	Screening -	N.a.	In adults (≥ 18 years) with chronic gastroesophageal reflux disease (GERD)a with or without other risk factorsb, what is the effectiveness (benefits and harms) of screening for esophageal adenocarcinoma (EAC) and precancerous conditions (Barrett's Esophagus (BE) and low- and high-grade dysplasia)? What are the effects in relevant subgroup populations? 1b If there is evidence of effectivenessc, what is the optimal time to initiate and to end screening, and what is the optimal screening interval (includes single and multiple tests and ongoing 'surveillance')? 2 In adults with chronic GERD with or without other risk factors,b who have been offered, received, or allocated to receive screening for EAC and precancerous conditions (BE and low- and high-grade dysplasia), how do they weigh the benefits and harms of screening, and what factors contribute to these preferences and to their decisions to undergo screening?	Ten studies evaluated the effectiveness of screening. One retrospective study reported no difference in long-term survival (approximately 6 to 12 years) between those who had a prior esophagogastroduodenoscopy and those who had not (adjusted HR 0.93, 95% confidence interval (CI) 0.58–1.50). Though there may be higher odds of a stage 1 diagnosis than a more advanced diagnosis (stage 2–4) if an EGD had been performed in the previous 5 years (OR 2.27, 95% CI 1.00–7.67). Seven studies compared different screening modalities, and showed little difference between modalities. Three studies reported on patients' unwillingness to be screened (e.g. due to anxiety, fear of gagging). The evidence on the effectiveness (benefits and harms) of screening for EAC and precancerous conditions (BE and dysplasia) is sparse and is of very low certainty, making it difficult to conclude whether or not people with chronic GERD should be screened for EAC and precancerous lesions.	Systematic reviews
Sharma, 2017	Prospective cohort study	N.a.	n=1378	>50 y/o	The inclusion criteria for our population comprised Chinese patients over the age of 50 years with one or more of the following additional criteria: history of dyspepsia for at least 4 weeks; family history of gastric cancer; and any medical condition for which an upper gastrointestinal (GI) endoscopic investigation was warranted.	N.a.	N.a.	N.a.	The primary outcome measure was the determination of the risk factors and symptoms for patients at risk of BE, namely IM, in an Asian setting.	We conclude that an increasing age and male gender are risk factors for BE in an Asian population, and should be targeted for a screening EGD.	JHG open
Desilets, 2014	Retrospective single-center cohort	N.a.	?	?	We examined the symptomology in newly diagnosed patients with BE in a tertiary care center from 1999 to 2008 with a focus on gender and segment length of BE	?	?	N.a.	?	Full article not available!Screening guidelines for BE focus on white males over age 50 with a history of chronic gastroesophageal reflux disease/heartburn. Future research is warranted to see if screening for BE should be expanded in select male populations under 50 with gastroesophageal reflux disease, who also have other recognized risk factors for the progression of BE to EAC, like smoking and obesity.	J Men's Health
Ward, 2006	Retrospective single-center cohort	N.a.	?	?	All patients referred for outpatient colonoscopy were eligible if they were at least 65 yr old and had not previously undergone esophagoscopy.	?	?	N.a.	?	Full article is not available! BE is common in unscreened male and female patients at least 65 yr of age who are referred for colonoscopy. Men were more likely than women to have BE although it occurred in both sexes. Reflux symptoms were fairly common but a poor predictor of BE.	American Journal of Gastroenterology
Nietert, 2003		N.a.						N.a.			
Inadomi, 2003		N.a.						N.a.			
Macdonald, 1997		N.a.						N.a.			
Langdong, 1997		N.a.						N.a.			

Table 4a								What is the role of non-endoscopic or non-inhospital tools for screening for Barrett's oesophagus (focusing on non-endoscopic cell collection devices - Cytosponge, EsoCheck, EsophaCap, E-nose, blood biomarkers)		

Author (year)	Methods		Population		Intervention			Outcomes		Remarks
	Design	Randimsation / blinding	N	Age	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	
Cytosponge										
Shaheen 2022 (Dig Dis Sci)	Case-control	No	No	191	Adults	Cytosponge and TFF3	Endoscopy	n/a	Diagnosis of BO	Sensitivity and specificity for BO	Of 191 patients, 99.5% successfully swallowed the device. Overall sample adequacy was 91% (171/188), with 84% (158/188) high quality. The detachment rate was 2/190 (1%). Overall sensitivity, specificity, and accuracy of the assay with TFF3 staining were 76%, 77%, and 76%. Sensitivity, specificity, and accuracy for ≥ 3 cm BE were 86%, 77%, and 82%. Asked if willing to repeat the procedure, 93% would, and 65% indicated a preference for the device over EGD. CONCLUSIONS: This study demonstrated a high rate of sample adequacy and promising acceptability of this non-endoscopic sampling device in a US population. Diagnostic characteristics suggest that non-endoscopic assessment of BE deserves further development as an alternative to endoscopy.
Gehrung 2021 (Nat Medicine)	Case-control/Machine Learning	No	No	2331 patients	Adults (no age range specified)	Machine Learning approach on Cytosponge specimens	Endoscopy	n/a	Diagnosis of BO through machine learning approach compared to endoscopy and biopsies	Diagnosis of BO and reduction (%) in workload of pathologist	By substituting manual review with automated review in low-priority classes, we can reduce pathologist workload by 57% while matching the diagnostic performance of experienced pathologists.
Fitzgerald 2020 (Lancet)	RCT	Yes	Yes	13514 (6531 to controls, 6983 to intervention group)	Age 50 years or older	Cytosponge + TFF3	Endoscopy	1 year	Diangosis of BO, Dysplasia, or OAC	Diangosis of BO, Dysplasia, or OAC	
Chettouh 2018 (Gut)	Case-control	No	No	Pilot cohort of 30 (20 cases and 10 controls). Validation cohort 278 (149 cases, 129 controls)	Not specified, but adults	Cytosponge + methylated biomarkers	Endoscopy	n/a	Diangosis of BO, Dysplasia, or OAC	Sensitivity and specificity for NDBE	Four methylation markers (TFPI2, TRIST1, ZNF345 and ZNF569) can detect Barrett's oesophagus when applied to Cytosponge samples.
Ross-Innes 2015 (Plos Medicine)	Case-control	No	No	1110 (463 controls and 647 cases of BE)	Not specified, but adults	Cytosponge + TFF3	Endoscopy	n/a	Main outcome measures Sensitivity and specificity estimates for detecting Barrett's oesophagus compared with gastroscopy	sensitivity and specificity for BO	The overall sensitivity of the test was 79.9% (95% CI 76.4%–83.0%), increasing to 87.2% (95% CI 83.0%–90.6%) for patients with ≥3 cm of circumferential BE, known to confer a higher cancer risk. The sensitivity increased to 89.7% (95% CI 82.3%–94.8%) in 107 patients who swallowed the device twice during the study course. There was no loss of sensitivity in patients with dysplasia. The specificity for diagnosing BE was 92.4% (95% CI 89.5%–94.7%).
Kadri 2010 (BMJ)	Cohort	No	No	504 patients	50-70 years	Cytosponge + TFF3	Endoscopy	n/a	Main outcome measures Sensitivity and specificity estimates for detecting Barrett's oesophagus compared with gastroscopy as the ideal method, and patient anxiety (short form Spielberger state trait anxiety inventory, impact of events scale) and acceptability (visual analogue scale) of the test.	sensitivity and specificity for BO	Compared with gastroscopy the sensitivity and specificity of the test was 73.3% (95% confidence interval 44.9% to 92.2%) and 93.8% (91.3% to 95.8%) for 1 cm or more circumferential length and 90.0% (55.5% to 99.7%) and 93.5% (90.9% to 95.5%) for clinically relevant segments of 2 cm or more. Most participants (355/496, 82%, 95% confidence interval 78.9% to 85.1%) reported low levels of anxiety before the test, and scores remained within normal limits at follow-up. Less than 4.5% (2.8% to 6.1%) of participants reported psychological distress a week after the procedure.

Study	Design			Sample size	Age	Test	Comparator		Outcome	Outcome measure	Findings
Xu 2019 (J Clin Path)	Case-control	No	No	258 (56 cases and 202 controls)	>20 years	Serology and Cytosponge	Endoscopy	n/a	Diagnosis of BO	Diagnosis of BO	Serology + Cy

EsophaCap

Study	Design			Sample size	Age	Test	Comparator		Outcome	Outcome measure	Findings
Iyer 2021 (GIE)	Case-control	No	No	Test set : 89 (60 BE and 29 controls		Esophacap and 5 Methylated DNA marker panel	Endoscopy	n/a	Diagnosis of BO	Sensitivity and specificty for BO	Sensitivity of the 5 MDM panel for BE diagnosis was 93% at 90% specificity in the training set and 93% at 93% specificity in the test set. Areas under the receiver operating characteristic curves were .96 and .97 in the training and test sets, respectively.
Iyer 2020 (Am J Gastro)	Case-control	No	No	201 (112 Cases, 89 controls)	Not specified, but adults	EsophaCap + Methylation Biomarkers	Endoscopy	n/a	Diagnosis of BO	Diagnosis of BO	The cross-validated sensitivity and specificity of a 5 MDM random forest model were 92% (95% confidence interval 85%–96%) and 94% (95% confidence interval 87%–98%), respectively. M
Iyer 2018 (Am J Gastro)	Case-control	No	No	40 (20 BE and 20 controls)	Not specified, but adults	EsophaCap + Methylation Biomarkers	Endoscopy	n/a	Diagnosis of BO	AUC	2-marker panel (VAV3 + ZNF682) yielded excellent BE discrimination (AUC = 1).
Zhou 2019 (Clin Exp Gastroenterol)	Case-control	No	No	28 patients testing cohort, 136 patients validation	Adults	EsophaCap and MUC2 immunohisto-chemistry	Endoscopy	n/a	Diagnosis of IM or dysplasia, and non-IM	Sensitivity and specificity for IM, dysplasia or EAC	Using the biopsy as our goldstandard to screen IM, dysplasia and EAC by combined cytology and MUC2 IHC, thesensitivity and specificity were 68% and 91%, respectively
Wang 2019 (Clin Can Research)	Prospective Cohort	No	No	80 (52 training, 28 test)	>18 years	EsophaCap + Methylation Biomarkers	Endoscopy	n/a	Diagnosis of BO	Diagnosis of BO	EsophaCap

Cost-effectiveness Studies for non-endsocopic devices

Study	Design			Sample size	Age	Test	Comparator		Outcome	Outcome measure	Findings
Swart 2021 (EClinicalMedicine)	Cost-utility analysis	No	No	Modelling	Adults 50yo on acid suppressant therapy	Cytosponge	Endoscopy	n/a	ICER and QALY of Cytosponge screening. Number of cases of incident sympotmatic oesophageal adenocarcinoma	ICER and QALY	At a willingness-to-pay threshold of £20,000 per QALY, the probability that Cytosponge-TFF3 was cost-effective was over 90%. Interpretation: Using data from a pragmatic randomised trial, one-off Cytosponge-TFF3 screen is cost-effective relative to usual care for patients with gastro-esophageal reflux disease, despite relatively low uptake and an older population in this trial setting than previously modelled. Improving Cytosponge-TFF3 uptake and targeting younger patients is likely to further improve cost-effectiveness.
Sami 2021 (Am J Gastro)	Cost-effectiveness analysis	No	No	Modelling	50 yo men white men with GERD, 50 yo men and women with BE prevalence 1.6%, and 50yo GERD independent with BE prevalence 5%	Swallowable cell collection devices, exhaled VOC, sedated endoscopy, transnasal endoscopy	Endoscopy	n/a	ICER and QALY	ICER and QALY	In both GERD-independent scenarios, most non-sEGD BE screening tests were cost effective. Swallowable esophageal cell collection devices with biomarkers were cost effective (<$35,000/QALY) and were the optimal screening tests in all scenarios. Exhaled volatile organic compounds had the highest ICERs in all scenarios. ICERs were low (<$25,000/QALY) for all tests in the GERD-based scenario, and all non-sEGD tests dominated no screening. ICERs were sensitive to BE prevalence and test costs. DISCUSSION: Minimally invasive nonendoscopic tests may make GERD-independent BE screening cost effective. Participation rates for these strategies need to be studied.
Heberle 2017 (Clin Gastro Hepatol)	Cost-effectiveness analysis	No	No	Modelling	Males age 60	Cytosponge	Endoscopy	n/a	ICER and QALY of Cytosponge screening. Number of cases of incident sympotmatic oesophageal adenocarcinoma	ICER for Cytosponge screening vs no screening	CONCLUSIONS: In a comparative modeling analysis of screening strategies for BE in patients with GERD, we found Cytosponge screening with endoscopic confirmation to be a cost-effective strategy. The greatest benefit was achieved by endoscopic screening, but with an unfavorable cost margin.
Benaglia 2013 (Gastroenterology)	Cost-effectiveness analysis	No	No	Modelling	Men 50 years or older	Cytosponge (acceptance rate 45%)	Endoscopy (acceptance rate 23%)	n/a	ICER and QALY of Cytosponge screening. Number of cases of incident sympotmatic oesophageal adenocarcinoma	ICER and QALY of Cytosponge screening. Number of cases of incident sympotmatic oesophageal adenocarcinoma	CONCLUSIONS: In a microsimulation model, screening 50-year-old men with symptoms of gastroesophageal reflux disease by Cytosponge is cost effective and would reduce mortality from esophageal adenocarcinoma compared with no screening.

Volatile organic compounds

Study	Study type			Sample size	Age	Method	Setting		Purpose	Outcome	Results	Notes
Peters 2019 (Gut)	Case-control	No	No	402 (141 BO, 140 GORD, 132 control)	Adults (no age range specified)	Electric Nose (Volatile Organic Compounds)	Endoscopy	n/a	Diagnosis of BO	Sensitivity and specificity for BO	This electronic nose was able to distinguish between patients with and without BO with good diagnostic accuracy (sensitivity 91% specificity 74%) and seemed to be independent of proton pump inhibitor use, the presence of hiatal hernia, and reflux. This technique may enable an efficient, well-tolerated, and sensitive and specific screening method to select high-risk individuals to undergo upper endoscopy.	
Chan 2017 (Gastroenterology)	Cross-sectional	No	No	122 (66 BE and 56 no BE)	Adults (no age range specified)	Electric Nose (Volatile Organic Compounds)	Endoscopy	n/a	Diagnosis of BO	Sensitivity and specificity for BO	Sensitivity of BO was 82%, specificity was 80% and accuracy was 81%	
Bhatt 2016 (GIE)	Case control	No	No	39 (20 OAC and 19 GERD)	Adults	Electric Nose (Volatile Organic Compounds)	Endoscopy	n/a	Diagnosis of OAC	Sensitivity and specificity for OAC	Results: The headspaces from 39 plasma samples (20 EAC, 19 GERD) were analyzed. The levels of 9 VOCs (acetonitrile, acrylonitrile, carbon disulfide, isoprene, 1-heptene, 3-methylhexane, [E]-2-nonene, hydrogen sulfide, and triethylamine) were significantly altered in EAC patients compared with GERD patients. A multivariable logistic regression analysis was performed to build a model for the prediction of EAC. The model identified patients with EAC with an area under the curve of 0.83 (95% confidence interval, 0.67-0.98). Conclusions: Plasma VOCs may be useful in diagnosing EAC. Larger studies are needed to confirm our pilot study observations.	

EsoCheck

Study	Study type			Sample size	Age	Method	Setting		Purpose	Outcome	Results	Notes
Moinova 2018 (Sci Transl Med)	Case-control	No	No	Not specified or pre-defined	Not specified, but adults	EsoCheck + methylated biomarkers	Endoscopy	n/a	Diangosis of BO, Dysplasia, or OAC	Sensitivity for NDBE, LGD, HGD, EAC		

MicroRNAs

Study	Study type			Sample size	Age	Method	Setting		Purpose	Outcome	Results	Notes
Inokuchi-2021 (J Clin Medicine)	Systematic Review	No	No	6 studies available for detection of BE, 14 studies available for detection of OAC	Adults	microRNA (measured using Nanostring miRNA analysis platform)	Endoscopy	n/a	Diagnosis of BO and OAC	Sensitivity, specificity and AUC for diangosis of BO and OAC	Results: microRNA-15a-5p, -195a-5p (Bansal et al), -194-5p, -451a (Bus et al), -143-3p, 215-5p, -194-5p (Cabbi et al 2016) and -130a-3p (Wang et al 2019) were upregulated, and -136-5p (Bust et al 2014) and -320e and -199a-3p (Pavlov et al 2018) were downregulated in BE diagnosis. For a diagnosis of OAC, miR-21-5p, -25-3p and -93-5p appeared in more than 1 publication and miR-21-5p was the most commonly observed aberrant miRNA in human cancer and is thought to promote metastasis of OAC by regulating cancer cell apoptosis and could be an independent prognostic biomarker for disease specific survival.	Pavlov 2018 included in SR
Craig 2020 (Clin Transl Gastro)	Case-control	No	No	31 patients (4 Normal, 8 GERD,7 BE, 5 LGD, 5 OAC)	Adults (no age range specified)	blood microRNAs	Endoscopy	n/a	AUC for discriminating normal, GERD, BE, LGD, OAC	AUC for discriminating normal, GERD, BE, LGD, OAC	Our data provide an miRNA signature of normal, precancerous, and cancerous tissue that may stratify patients at risk of progressing to EAC. We found that serum miRNAs have a limited ability to distinguish between disease states, thus limiting their potential utility in early disease detection.	
Wang 2018 (Onco Target Ther)	Case-control	No	No	Validation 130 (30 controls, 60 BE and 40 OAC)	Adults	microRNA	Endoscpy	n/a	Diagnosis of BO and OAC	Sensitivity, specificity and AUC for diangosis of BO and OAC	We found an increase in serum miR130a in low-grade and high-grade dysplasia BE patients compared to individuals with metaplasia. We also observed that miR130a expression levels increased gradually from early-stage (I, II) to advanced-stage (III, IV) EAC patients. Conclusion: Our preliminary results provide evidence that circulating miR130a is correlated with the development of BE and EAC.	
Li 2018 (Gastro)	Case-control	No	No	64 (38 BO cases, 26 controls)	No specified, but adults	Cytosponge + microRNAs	Endoscopy	n/a	Diagnosis of BO	Diagnosis of BO	Expression level of MIR192, MIR196a, MIR199a, combined that of trefoil factor 3, identified patients with BE with an AUC of 0.93, 93.1% sensitivity, and 93.7% specificity. Expression of MIR194 is increased in BE samples via epigenetic mechanisms that might be involved in BE pathogenesis.	
Bus 2016 (J Gastroenterology)	Case-control	No	No	22 (6 controls, 8 BE, 8 OAC)	Not specified, but adults	Serum microRNAs	Endoscopy	na	miRFA	Sensitivity and specificity for BO	The most informative diagnostic panel to distinguish controls from BE was the combination of miRNA-95-3p, -136-5p, -194-5p, and -451a showing the highest AUC 0.832 (95 % CI 0.698–0.967) with a sensitivity of 78.4 % (95 % CI 61.8–90.2) and specificity of 85.7 % (95 % CI 57.2–98.2) . A combination of three or more miRNAs was found to have a good diagnostic performance in discriminating BE from controls (AUC: 0.832), EAC from controls (AUC: 0.846), and BE from EAC (AUC: 0.797)	
Mallick 2016 (Dig Dis Sci)	Systematic Review	No	No	11 studies	Not specified	blood microRNAs	Endoscopy	n/a	miRNA	Correlation between miRNA expression and diagnosis of BO	MicroRNAs miR-192, -194, -203, -205, and -215 are promising tissue biomarkers for diagnosing BE.	

Study	Design			N	Population	Biomarker	Comparator		Index test	Outcome	Findings
Correia 2020 (Sci rep)	Case-control	No	No	246 (112 BE and 134 age-sex maatched controls)	Not specified, but only adults	Blood biomarker for BMP	Endoscopy	n/a	levels of BMP 2,4,5	levels of BMP 2,4,5	Concentration levels of BMP2, BMP4, and BMP5 were elevated in BE patients, with BMP2 and BMP5 significantly increased. BMP5 remained significant after multivariate analysis and was associated with an increased risk for BE with an OR of 1.49 (p value 0.01). Per log (pg/mL) of BMP5, the odds of having BE increased by 50%. Future optimization and validation studies might be needed to prove its utility as a non-invasive method for the detection of BE in high-risk populations and screening programs.
Xie 2020 (Cancer Epidemiol Biomarkers)	Meta-analysis	No	No	19 studies	Not specified	Blood based circulating levels of inflammatory and metabolic biomarkers	Endosopy	n/a	Assocaitions between blood biomarker and a diagnosis of BO/OAC	Assocaitions between blood biomarker and a diangosis of BO/OAC	Higher circulating levels of leptin, glucose, insulin, C-reactive protein, interleukin 6 and soluble tumor necrosis factor receptor 2 may be associated with an increased risk of EAC or BE.
Chu-2020 (J Cancer)	Case-control	No	No	355 (148 GV, 59 OAC and 148 health controls)	Adults	L1-cell Adhesion Molecule (serum biomarker)	Endoscopy	n/a	Diagnosis of OAC	Sensitivity, specificity and AUC for diangosis of OAC	Results: The concentrations of serum L1CAM were significantly lower in GC and EJA than those in healthy controls (P<0.001). Detection of L1CAM provided a sensitivity of 83.1%, a specificity of 62.2%, and an area under the curve (AUC) of 0.769 (95% CI: 0.715-0.823) in diagnosing GC, and a sensitivity of 66.1%, a specificity of 62.2%, and an AUC of 0.672 (95% CI: 0.590-0.755) in diagnosing EJA. Similar results were observed in the diagnosis of early-stage GC (0.681 (95%CI: 0.596-0.766)) and early-stage EJA (0.674 (95%CI: 0.528-0.820)). Analysis of clinical data showed that the levels of L1CAM were significantly associated with lymph node metastasis in GC (P<0.05). Conclusions:Our study showed that serum L1CAM might be a diagnostic biomarker for GC and EJA.
Maddalo 2017 (J Clinical Gastro)	Case-control	No	No	213 (53 BE, 53 OAC, 107 control)	Not specified, but adults	Serum Squamous Cellular Carcinoma Antigen	Endoscopy	n/a	Diangosis of BO or OAC	Sensitivity and specificity for BO or OAC	The cutoff for SCCA-IgM, as calculated by the ROC curves between patients affected by either BE or EAC and GERD patients was 56.6 AU/mL with a 91.5% sensitivity, 75.4% specificity, a positive predictive value (PPV) of 85.8%, a negative predictive value (NPV)
Zaidi 2014 (Cancer)	Case-control	No	No	Serum Discovert dataset 32 (20 GERD, 12 OAC). Serum Validation dataset 67 (36 GERD and 31 locaregional OAC)	Adults	Blood biomarkers	Endoscopy	n/a	Diagnosis of OAC	Senstiivity and specificity of biomarkers for OAC	By using serum data, a Bayesian rule-learning predictive model with 4 biomarkers (Myeloperoxidase, Protein S100-A9, Biglycan, Annexin) was developedto accurately classify disease class; the cross-validation results for the merged data set yielded accuracy of 87% and an area underthe receiver operating characteristic curve of 93%.CONCLUSIONS:Serum biomarkers hold significant promise for the early, noninva-sive detection of EAC.
Zhai 2012 (Neoplasia)	Case-control	No	No	36 (8 OAC, 8 mathced serum DNA, 10 BE, 10 healthy contorls)	Not specified, but adults	Blood methylation profile (Genome-wide DNA Methylation Profiling of cell-free Serum DNA)	Endsocopy	n/a	DNA Methylation profile in tissue and cfDNA	Correlation between DNA methylation profile in biopsies tissue and cdDNA from serum	We found that cfDNAm profiles were highly correlated to DNAm profiles in matched tumor tissue DNA (r = 0.92) in patients with EA.

PICO search string:	Medline: 265 papers screened "Early Detection of Cancer"[Mesh] OR "early detection"[tw] OR "Mass Screening"[Mesh] OR Screen*[tw] Filters: Humans, English AND "DNA Methylation"[Mesh] OR "Genetic Markers"[Mesh] OR "Biomarkers"[Mesh] OR "Electronic Nose"[Mesh] OR "Breath Tests"[Mesh] OR Cytosponge[tw] OR EsophaCap[tw] OR EsoCheck[tw] OR "non-endoscopic"[tw] OR nonendoscopic[tw] OR "MicroRNAs"[Mesh] OR "Volatile Organic Compounds"[Mesh] OR "sponge-on-string"[tw] OR "Methylation"[Mesh] OR "blood biomarkers"[tw] OR "E-nose"[tw] AND "Barrett Esophagus/diagnosis"[Mesh] OR "Barrett's oesophagus"[tw] OR "Gastroesophageal Reflux"[Mesh] OR "low grade dysplasia"[tw] OR "high grade dysplasia"[tw] OR "esophageal adenocarcinoma"[tw] OR "Adenocarcinoma Of Esophagus" [Supplementary Concept]

Manual inclusion of 3 papers - Chan et al 2017 Gastroenterology and Iyer et al 2021 Gastrointestinal Endoscopy and Zhou et al 2019 (Clin Exp Gastroenterol)

OVID: 330 papers screened, duplicated papers avoided and futher 7 papers included from OVID/Embase. Total 32 papers included from Medline and Embase
Ovid Search strategy
Embase <1974 to 2022 March 16>

1	Barrett esophagus/di [Diagnosis]	3648
2	barretts oesophagus.mp.	2610
3	esophageal adenocarcinoma/di [Diagnosis]	1611
4	oesophageal adenocarcinoma.mp.	1852
5	gastroesophageal reflux/di, dm [Diagnosis, Disease Man	8323
6	gastro-oesophageal reflux disease.mp.	3221
7	gastrointestinal dysplasia/di [Diagnosis]	315
8	low grade dysplasia.mp.	4741
9	high grade dysplasia.mp.	10315
10	intramucosal carcinoma.mp.	767
11	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10	28924
12	early detection.mp.	106278
13	screening/ or screening.mp.	1198516
14	cancer screening.mp. or cancer screening/	102665
15	mass screening.mp. or mass screening/	63280
16	12 or 13 or 14 or 15	1277716
17	dna methylation.mp. or DNA methylation/	103543
18	biomarker.mp. or biological marker/	495921
19	breath test.mp. or breath analysis/	22540
20	volatile organic compounds.mp. or volatile organic com	25186
21	Cytosponge.mp. or esophageal cell sampling device/	152
22	EsophaCap.mp.	20
23	EsoCheck.mp.	17
24	Electric nose.mp. or electronic nose/	1486
25	non-endoscopic.mp.	409
26	nonendoscopic.mp.	236
27	circulating microRNA/ec [Endogenous Compound]	1170
28	microRNA.mp.	204327
29	sponge-on-string.mp.	4
30	blood biomarker.mp.	1200
31	17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 o	803500
32	11 and 16 and 31	367
33	limit 32 to (human and english language)	330

Table 4b		**what is the role of volatile organic compounds Breath tests in BE screening**								
Author (year)	**Methods**		**Population**			**Intervention**			**Outcomes**	
	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	Remarks
Sacheen Kumar 2015	prospective study	nil	210	51-73	Eligible patients referred for UGI endoscopy	Exhaled breath samples were collected when patients were attending for one of the following investigations—staging laparoscopy and esophago-gastro-duodenoscopy (EGD), UGI endoscopic ultrasound, or EGD only.	All study participants had undergone upper gastrointestinal endoscopy on the day of breath sampling.		(1) to identify and quantify exhaled breath VOCs using SIFT-MS in patients with esophageal or gastric adenocarcinoma and compare them to noncancer controls and (2) to construct a VOC-based risk prediction model to distinguish patients with esophageal and gastric adenocarcinoma from non-cancer controls.	Receiver operating characteristic analysis and a diagnostic risk prediction model were used to assess the discriminatory accuracy of the identified VOCs. The 8 most significant predictors for adeno- carcinoma identified from stepwise logistic regression were decanal, nonanal, phenol, ethyl phenol, methyl phenol, hexanoic acid, hepta- nal, and butyric acid.	Twelve VOCs—pentanoic acid, hexanoic acid, phenol, methyl phenol, ethyl phenol, butanal, pentanal, hexanal, heptanal, octanal, nonanal, and decanal—were present at significantly higher concentrations (P < 0.05) in the cancer groups than in the noncancer controls
Yonne Peters 2019	prospective study	nil	513			Adult patients undergoing a clinically indicated upper endoscopy were invited to provide a 5min breath sample using an electronic nose					We showed as a proof of concept that it is possible to non-invasively detect the presence of BO by VOC breath analysis using an electronic nose device in patients with and without GORD with a sensitivity of 91% and a specificity of 74%.

PICO search string:

Search: (barrett's esophagus) AND (organic volatile compounds)
("barrett s oesophagus"[All Fields] OR "barrett esophagus"[MeSH Terms] OR ("barrett"[All Fields] AND "esophagus"[All Fields]) OR "barrett esophagus"[All Fields] OR ("barrett s"[All Fields] AND "esophagus"[All Fields]) OR "barrett s esophagus"[All Fields]) AND (("organic"[All Fields] OR "organically"[All Fields] OR "organics"[All Fields]) AND ("volatile"[All Fields] OR "volatiles"[All Fields] OR "volatilities"[All Fields] OR "volatilization"[MeSH Terms] OR "volatilization"[All Fields] OR "volatility"[All Fields] OR "volatilizations"[All Fields] OR "volatilize"[All Fields] OR "volatilized"[All Fields] OR "volatilizes"[All Fields] OR "volatilizing"[All Fields]) AND ("compound"[All Fields] OR "compound s"[All Fields] OR "compounds"[All Fields]))
Translations
barrett's esophagus: "barrett's oesophagus"[All Fields] OR "barrett esophagus"[MeSH Terms] OR ("barrett"[All Fields] AND "esophagus"[All Fields]) OR "barrett esophagus"[All Fields] OR ("barrett's"[All Fields] AND "esophagus"[All Fields]) OR "barrett's esophagus"[All Fields]
organic: "organic"[All Fields] OR "organically"[All Fields] OR "organics"[All Fields]
volatile: "volatile"[All Fields] OR "volatiles"[All Fields] OR "volatilities"[All Fields] OR "volatilization"[MeSH Terms] OR "volatilization"[All Fields] OR "volatility"[All Fields] OR "volatilizations"[All Fields] OR "volatilize"[All Fields] OR "volatilized"[All Fields] OR "volatilizes"[All Fields] OR "volatilizing"[All Fields]
compounds: "compound"[All Fields] OR "compound's"[All Fields] OR "compounds"[All Fields]

Table 4c	**Role of ultraslim transnasal endoscopy compared to a standard endoscopy in the management of Barrett's oesophagus**										
Author (year)	Methods		Population				Intervention			Outcomes		Remarks
	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	
Mohammed.K el al (2016)	Prospective,randomised cross-over single and tertiary ref centres	randomised, cross over	25	>18y	standard diagnosis,BO >2cm,	Transnasal endoSheath Endoscope	Standard endoscopy		sensitivity and specificity for diagnnosis of BO with IM,Optical quality of TEE with SE,pt tolerability	TEE accurate diagnosis of BP	
Shariff (2012)	Prospective , randomised, cross over	randomised cross over	95	>18y	BO>2cm,	TNE	standard endoscopy	35d	sen 98% and sp 100%, PPV 100% for endos diag of BE, IM sen 91.2 %, spec 100% PPV 100%	no difference in the endoscopic detection of BE (P □ .89)	
Hiroyuki Osawa et al (2009)	prospective comparison	comparative , non randomised	72	>18y	consecutive BE diagnosis pts using Conv TNE,	knownBE patients were observed for TNE vs FICE TNE	TNE vs FICE TNE				ni bx obtained ,The transnasal FICE system enables clear visualization of palisade vessels and provides better con- trasting images of the demarcation between the BE mucosa and the gastric mucosa, and thus contributes to easier diagnosis of endoscopic BE.
Sarmed S Sami (2015)	prospective Randomised comparative	randomised cross over	209	>50y	Prev completed validated GI sx questionnnaire were then randomised to either of teh three	huTNE vs mu TNE	sEGD		participation rate, complete evaluatio of oesophagus,mean recovery times , willingness to undergo procedure again	participation rate, To compare the clinical effectiveness of these techniques in a population based randomized trial, successful intubation, rate of complete evaluation	participation higher in non sedated-statistically not sig, pts with freq heartburn were more likely to participate,Complete evaluation of the esophagus was similar using (mobile van unsedated)muTNE (99%), (hospital outpatient unsedated)huTNE (96%) and (standard)sEGD (100%) techniques (p=0.08).biopsy acquisition was lower in muTNE ,
Gupta M (2014)	Prospective , mixed mode survey	non randomised, survey based	136	>50y	136 responded, 72% interested in screening,unsedated tech preferred by 64%,sEGD 36%	136 responded, 72% interested in screening,unsedated (VCE and uTNE) tech preferred by 64%,sEGD 36%	sEGD		We aimed to assess these attitudes via a survey.	71 % of responders agreeing to be screened, majority of adults were willing to undergo screening for BE/EAC, with a preference for unsedated techniques.	
Anne F Peery (2012)	Multi-center, prospective, cross-sectional study.	non randomised, office based ,two tertiary centres,self selected	422	40-85y	between the ages of 40 and 85, regardless of GERD symptomatology.	Examinations were performed using the one-wheel TNE-5000 endoscope system,Two endoscopists , in office settings	nil		procedure yield, 2) Completeness of examination, 3) Procedure length, 4) Adverse events and complications, 5) Choking, gagging, pain or anxiety during the examination, and 6) Overall tolerability	unsedated transnasal esophagoscopy is a safe and well-tolerated method to screen for esophageal disease in a primary care population, affordable option for general .	
Christopher H Belvins (2018)		prospective, randomized controlled trial	201	50-87Y	Consenting community patients without known BE were randomly assigned to receive huTNE, muTNE, or sEGD, followed by a telephone administered preference and tolerability assessment instrument 24 hours after study procedures. Patient preference was measured by the waiting trade-off method.	randomly assigned to receive huTNE, muTNE, or sEGD, followed by a telephone administered preference and tolerability assessment instrument 24 hours after study procedures.	sEGD		Although tolerability scores were superior for sEGD (P<0.001) compared with uTNE, scores for uTNE examinations were acceptable.	Patient preference is comparable between sEGD and uTNE for diagnostic examinations conducted in an endoscopy suite or in a mobile setting. Given acceptable tolerability, uTNE may be a viable alternative to sEGD for BE screening.	This suggests that patient preference may not be a significant factor when utilizing uTNE as an alternative to traditional sedated EGD for BE screening. When combined with lower direct and indirect costs and comparable clinical effectiveness, the rationale for the use of unsedated minimally invasive tools for BE screening is further strengthened.
Amitabh Chak	Randomized block study design with allocation concealment.	Outpatient clinic setting at a Veteran Affairs (VA) medical center.	184	45-85y	without a prior EGD in the past 10 years, and with no contraindications to ECE or TNE (history of recurrent epistaxis), with or without symptoms of GERD were eligible to participate.	To compare acceptance and tolerability of 2 novel unsedated office-based endoscopic screening techniques.	TNE vs ECE		(1) Transnasal esophagoscopy (TNE); (2) Capsule esophagoscopy (ECE).	(1) Acceptance of TNE and ECE; (2) Tolerability of TNE and ECE and effectiveness of BE screening.	General veteran population may not reflect the screening population for BE.
Joseph Y Chang et al (2011)	A Prospective Randomized Pilot Study	Mayo clinic	127	50y	A medical chart review of patients (N=2957) from the established Olmsted County cohort was used to identify the subset older than 50 years who had not undergone sEGD in the past 5 years	TNE and Oesophageal capsule	sEGD		The goals of this study were to demonstrate feasibility of population-based screening for BE and to obtain preliminary estimates of participation rates for noninvasive screening techniques (uTNE, VCE) in comparison with sEGD so that sample size for a larger study could be obtained.	Unsedated techniques may be acceptable, feasible, and safe alternatives to sEGD to screen for BE in the community.	

Table 4d			**Role of video capsule endoscopy in the management of Barrett's oesophagus**								
Author (year)	**Methods**			**Population**			**Intervention**			**Outcomes**			**Remarks**

Author (year)	Design	Randimisation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	Remarks
Otto S Lin (2007)	Prospective and blinded, with no adjudication.	Prospective and blinded, with no adjudication.	96		patients with upcoming endoscopies for the indications of chronic gastroesophageal reflux symptoms (screening group) or Barrett's esophagus (surveillance group) in the Virginia Mason Medical Center endoscopy unit schedule.	ECE followed by EGD in each subject.			Sensitivity, specificity, and positive and negative predictive values of ECE for Barrett's esophagus by using EGD results, with histologic confirmation as the criterion standard. The secondary end points were adverse effects and complications from ECE and EGD.	ECE identified 14 of 21 patients with true Barrett's esophagus (sensitivity, 67%) and 58 of 69 patients without Barrett's esophagus (specificity, 84%). The positive predictive value was 22% and the negative predictive value was 98% for Barrett's esophagus	To assess the accuracy of ECE for the diagnosis of Barrett's esophagus.
Joseph Y Chang et al (2011)	A Prospective Randomized Pilot Study	Mayo clinic	127	50y	A medical chart review of patients (N=2957) from the established Olmsted County cohort was used to identify the subset older than 50 years who had not undergone sEGD in the past 5 years	TNE and Oesophageal capsule	sEGD		The goals of this study were to demonstrate feasibility of population-based screening for BE and to obtain preliminary estimates of participation rates for noninvasive screening techniques (uTNE, VCE) in comparison with sEGD so that sample size for a larger study could be obtained.	Unsedated techniques may be acceptable, feasible, and safe alternatives to sEGD to screen for BE in the community.	BE was identified in 3 patients and esophagitis in 8.
Amitabh Chak(2014)	Randomized block study design with allocation concealment.	Outpatient clinic setting at a Veteran Affairs (VA) medical center.	184	45-85y	without a prior EGD in the past 10 years, and with no contraindications to ECE or TNE (history of recurrent epistaxis), with or without symptoms of GERD were eligible to participate.	To compare acceptance and tolerability of 2 novel unsedated office-based endoscopic screening techniques.	TNE vs ECE		(1) Transnasal esophagoscopy (TNE); (2) Capsule esophagoscopy (ECE).	(1) Acceptance of TNE and ECE; (2) Tolerability of TNE and ECE and effectiveness of BE screening.	General veteran population may not reflect the screening population for BE.
Eliakim et al (2005)	prospective multicentre international	doubleblinded	Seven medical centers enrolled 109 consecutive patients	more than 18	histo of BE or chronic GERD sx and no prior UGIE and undergoing endoscopy	Ninety-three (88%) patients wereendoscoped for GERD symptomsand 13 (12%) for surveillanceof Barrett esophagus.Upper endoscopy using conscious sedation with meper-idine and midazolam was performed on the same day followingthe ECE.	ECE vs sEGD		In patients diagnosed with "Barrett esophagus" basedon the gold standard, the ECE demonstrated a sensitivity of97% and an NPVof 97%. Specificity and PPVs were estimatedas 99% for both parameters .Intention-to-treat (ITT)analysis for the diagnosis of "Barrett esophagus" demon-strated a sensitivity of 92% and NPV of 96%. The ITTspecificity and PPV for the diagnosis of "Barrett esophagus"were 99% and 97%, respectively.	ECE using the Given Esophageal Capsule is a simple, safe, and accurate method for the diagnosis of esophageal mucosal disease and is well tolerated by patients. Future generations of esophageal capsules with higher frame speed are in trials, and ECE using the esophageal capsule may become the primary diagnostic modality for the evaluation of patients with esophageal disease.	
Sabina Beg (2005)	This was a prospective single-blinded, diagnostic cohort study performed at Nottingham University Hospita	single blinded	Fifty patients were recruited into the study, of whom 47 (94%) completed the MACE procedure and 50 (100%) completed the EGD.	39-83y	Patients who were scheduled to undergo an EGD as part of the surveillance of known BE or EV were invited to participate in this study as cases.	Both diagnostic tests were undertaken on the same day by 2 independent operators blinded to the findings of the other modality. The MACE procedure was performed first to avoid any artefact caused by endoscope trauma or biopsy acquisition during the EGD.	MACE and EGD		This study was designed as a feasibility study, which will then inform power calculations for future trials.	MACE is both technically feasible and safe. Accuracy in the diagnosis of esophageal lesions compared with EGD was reasonable.There was a difference in technical success of 6% (94% with MACE versus 100% with EGD).Despite a rapid transit through the esophagus in some cases, the gastroesophageal junction was visualized in 100% and the z line in 91.5%. Patient comfort scores demonstrated a preference for MACE over EGD.	
P Sharma et al (2008)	blinded prospective	blinded	A total of 100 patients who met the inclusion and exclusion criteria (56 patients at the Veterans Affairs Medical Center, Kansas City, MO and 44 at The Southern Arizona Veterans Affairs Health Care System, Tucson, AZ) were prospectively enrolled in the study.		Subjects with reflux symptoms occurring at least twice a week, as documented by a validated GERD questionnaire (GERQ) (20), were included in the study. Enrollment of patients was also based on the following inclusion criteria: age 21 yr or older, ability to provide written informed consent, and willingness to undergo upper endoscopy and ECE.	ECE followed by EGD in each subject.			The primary aim of this blinded, prospective study was to assess the diagnostic accuracy of ECE for BE in patients presenting with GERD symptoms or for surveillance of BE. The secondary aims were to evaluate the diagnostic accuracy of ECE for erosive esophagitis and hiatal hernia, and assess the safety and adverse event profile of ECE.	ECE identified 41 of 53 patients with suspected BE (Fig. 2) based on the gold standard (findings at upper endoscopy). Thus ECE demonstrated a sensitivity and specificity of 77% and 85%, respectively. The PPV and NPV were 87% and 74%, respectively .If histologically confirmed BE on EGD was used as the reference, the sensitivity, specificity, PPV, and NPV of ECE to diagnose BE were 78%, 75%, 74%, and 79%, respectively. The sensitivity, specificity, PPV, and NPV of ECE for diagnosing BE in patients undergoing EGD for GERD symptoms, using upper endoscopy as the reference, were 67%, 87%, 60%, and 90%, respectively . Similarly,the sensitivity, specificity, PPV, and NPV in patients undergoing surveillance for BE were 79%, 78%, 94%, and 44%, respectively.The sensitivity and specificity of ECE for diagnosing LSBE were 83% and 85% and those for SSBE were 74% and 85%, respectively.	

Bhardwaj A (2009)	metanalysis	Eight studies were prospective blinded and one was retrospective blinded.	Nine studies comprising a total of 618 patients met the inclusion criteria.	8 studies meeting the following criteria were included: (i) blinded studies (prospective or retrospective) evaluating the diagnostic accuracy of ECE for BE; (ii) studies in which patients underwent EGD (with or without esophageal biopsy) and ECE during the study period; (iii) studies that used EGD or histologically conﬁrmed intestinal metaplasia (IM) as the reference standard; (iv) studies in which data were available or obtainable to determine the sensitivity and specificity of ECE for the diagnosis of BE; (v) abstracts or full-text articles in any language published in the last 5 years; although the search yielded no eligible studies in languages other than English.	To evaluate the diagnostic accuracy of esophageal capsule endoscopy (ECE) for Barrett ' s esophagus (BE) in patients with gastroesophageal reﬂux disease (GERD).				The pooled sensitivity and speciﬁcity of ECE for the diagnosis of BE for all studies were 77 and 86 % respectively. The pooled sensitivity and speciﬁcity of ECE for the diagnosis of BE using esophagogastroduodenoscopy (EGD) as the reference standard were 78 and 90 % , respectively; using histologically conﬁrmed intestinal metaplasia (IM) as the reference standard pooled sensitivity and speciﬁcity were 78 and 73 % , respectively.: Capsule endoscopy of esophagus has a moderate sensitivity and speciﬁcity for the diagnosis of BE in patients with GERD. The EGD remains the modality of choice for evaluation of suspected BE.	

	Table 5			Familial BE							
Author (year)	**Methods**			**Population**			**Intervention**			**Outcomes**	
	Design	Randimsation / blinding		N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	**Remarks**
Chak (2002)	Cross-sectional observational study	NA	NA	164	NA	58 patients with BE/EAC and 106 controls with GERD	comparison of family history of BE/EAC between these two groups	Review of medical records. No active intervention	NA	Percentage of confirmed family history of BE/EAC among cases and controls	higher prevalence of family history among cases than controls (24% vs. 5%). After multivariable analysis, patients with BE/EAC were 12 times more l ikely to have a positive family history compared to symptomatic GERD controls	
Romero (2002)	case-control study	not randomized	not blinded	100 cases and 100 controls	47 yo cases; 55 yo controls	100 fi rst-degree relatives (FDR) of patients with BE/EAC and GERD, and 100 controls with GERD who reported no family history of BE/EAC	100 cases and 100 controls were offered screening EGD	presence or absence of BE/EAC at endoscopy/patholog y	NA	Percentage of confirmed family history of BE/EAC among cases and controls	no s ignificant difference in the prevalence of BE/EAC among patients with family history of BE/EAC and those without (8% vs 5%, respectively). One patient (1%) among participants with positive family history had EAC versus none among controls.	
Chak (2004)	Cross-sectional observational study	NA	NA	screening endoscopies performed on 62 fi rst-degree relatives	44.6 yo for "isolated" relatives and 45.3 yo for "familial" relatives	relatives of probands with BE/EAC who never had endoscopy before	fi rst-degree relatives of patients with BE/EAC were offered endoscopy to screen for BE/EAC. A distinction was made between FDR of probands with known familial BE/EAC and FDR of probands with "i solated" BE/EAC (i .e. with no known family history of BE/EAC).	presence or absence of BE/EAC at endoscopy/patholog y	NA	percentage of BE/EAC at endoscopy/patholog y	Overall, 21% (13/62) of all FDR of patients with BE/EAC had positive findings, with most of these (11/13) found in FDR of probands with familial BE/EAC as opposed to probands with "i solated" BE/EAC. Only one case (1.6%) of EAC was found among 62 participants who had never had endoscopy before, and this was among the FDR of probands with family history of BE/EAC. While new members with positive BE/EAC were added to already known familial pedigrees, no new pedigrees were identified after offering endoscopy to FDR of patients with "i solated" BE/EAC	
Chak (2006)	Cross-sectional observational study	NA	NA	411 probands with BE/EAC	62.6	probands with BE/EAC	self-reported family history and medical review of relatives of those who reported positive family history	NA	NA	percentage of confirmed positive family history	Familial Barrett's esophagus was definitively determined in the case of 30 (7.3%) probands comprising17 of 276 (6.2%) with Barrett's esophagus, 11 of 116 (9.5%) with adenocarcinoma of the esophagus, and 2 of 21 (9.5%) with adenocarcinoma of the gastroesophageal junction.	
Juhasz (2011)	s ingle-center, prospective s ingle-arm study	not randomized	not blinded	48	44.4 yo	fi rst-degrre relatives of probands with HGD-BE/EAC	s ingle-center, prospective s ingle-arm study offering screening endoscopy to FDR of patients with HGD-BE/EAC	no comparison group	NA	BE/EAC at endoscopy	among 47 FDR who underwent endoscopy, 13 (27.7%) were positive for BE and none for EAC	
Ash (2011)	retrospective s ingle-center study	NA	NA	979 BE patients	58.4 (familial BE), 63.8 (non familial BE)	a cohort of BE patients	retrospctive review of medical records and interview of potential fi rst-degree relatives to confirm history of BE/EAC	comparison of sporadic BE and familial BE	NA	percentage of familial BE. Comparison of familial BE vs sporadic BE.	familial BE was seen in 5.8% of cases (35 cases). FBE probands were younger (mean, 58.4 vs. 63.8; p = 0.02) and had a s ignificant association with less-advanced neoplasia (adjusted OR 0.41, 95% CI 0.19–0.90). No difference with regard to BE length or hernia s i ze.	

Author (year)	Study type	Randomized	Blinded	Sample	Age	Population	Intervention	Comparison	Indication for RFA	Outcome	Results
Chak (2012)	Cross-sectional observational study	NA	NA	A total of 1397 individuals affected with either BE or EAC, of which 1146 were probands and 251 were relatives	NA	BE/EAC with or without family history	Comparison of age of diagnosis of EAC in non-familial EAC (i .e. probands with EAC with no known family history of BE/EAC), duplex families (i .e. probands with EAC and only one relative positive for BE/EAC) and multiplex families (i .e. probands with EAC and at least two relatives positive for BE/EAC)	Review of medical records. No active intervention	NA	Age of diagnosis of EAC	Median age of cancer diagnosis was s ignificantly younger in multiplex compared to duplex and non-familial kindreds (57 vs. 62 vs. 63 yrs, respectively, p = 0.045). Members of multiplex families developed EAC at an earlier age compared to non-familial EAC cases at multivariable mixed models (p = 0.019).
Mussetto (2013)	s ingle-center, prospective s ingle-arm study	not randomized	not blinded	18	52	patients with GERD and family history of BE/EAC	capsule endoscopy and standard endoscopy in patients with GERD and family history of BE/EAC	no comparison group	NA	BE/EAC at endoscopy	8/18 (44%) patients with GERD and family history of BE/EAC had a confirmed presence of Barrett's esophagus, and none had EAC
Verbeek (2014)	Cross-sectional observational study	NA	NA	603 probands with BE/EAC	64	probands with BE/EAC	self-reported family history and confirmation of positive family history with pathology from a centralized national dataset	NA	NA	percentage of confirmed positive family history	familial BE/EAC was definitive for 39 cases (7% of overall cases overall; 10% of fi rst-degree relatives)
Kharazmi (2018)	epidemiological observational study	NA	NA	13,325 esophageal cancers (any type)	66 yo at diagnosis (familial cases), 69 yo at diagnosis (sporadic cases)	histologically confirmed esophageal cancer	comparison between patients with or without family history of esophageal cancer (both overall, and by subtype)	comparison between patients with or without family history of esophageal cancer (both overall, and by subtype)	NA	Standardized incidence ratios (SIRs) of EAC	Esophageal adenocarcinoma in a fi rst- degree relative was associated with an increased familial ri sk of the same subtype of esophageal tumor (SIR 3.6, 95% CI 2.5–5.1).
Tofani (2019)	retrospective s ingle-center study of patients undergoing serial radiofrequency ablations (RFA)	NA	NA	301 patients who underwent RFA	62.1 yo	all patients who underwent RFA for BE/early EAC	retrospective review of all patients who underwent RFA for BE/early EAC	comparison between patients with or without family history of BE/EAC	as cl inically indicated for RFA	Age of diagnosis of EAC. Progression towards HGD/EAC.	Patients with family history of BE/EAC had an earlier age of diagnosis of EAC (51.5 vs 69 years) and higher ri sk of progression towards HGD/EAC (21.1% vs 8.7%). At multivariable analysis, there was an overall 5.5-fol d increased ri sk of progression towards HGD/EAC in patients with positive family history of BE/EAC (OR 5.55, 95%CI 1.47–20.0).
Qumseya (2019)	Meta-analysis	NA	NA	227 patients with family history of BE/EAC	NA	BE/EAC or family history for BE/EAC	Self-reported family history but unconfirmed with medical records, sefl-reported family history confirmed with medical records/pathology, screening endoscopy.	Many different protocols in the studies included (some did not have a control group). The meta- analysis attempts to compare patients with and without family history of BE/EAC	NA	BE/HGD-EAC	The pooled prevalence of BE in patients with family history positive for BE/EAC was 23.4% (95% CI, 13.7%-37.2%)
Rubenstein (2020)	Cross-sectional observational study	NA	NA	822 screened for CRC and were offered also upper endoscopy	58.5 yo (non-BE patients); 61 yo (BE patients)	anyone who went for a colonoscopy for screening for CRC	EGD offered to anyone who went for a screening colonoscopy	no comparison group	NA	percentage of BE/EAC at endoscopy/patholog y; odds ratio of family history of risk fo BE/EAC	Of the 822 men screened for CRC who underwent upper endoscopy, 70 were newly diagnosed with BE (8.5%). BE was associated with family histories of esophageal cancer (OR 2.63; 95%CI 1.07- 6.47)
Peters (2021)	multicenter case-control study	NA	NA	480 BE patients and 420 controls without BE who had a total of 6393 fi rst-degree relatives	65.8 yo cases; 62.5 yo controls	probands with BE/EAC and controls without BE/EAC	self-reported family history and confirmation of positive family history with pathology from a centralized national dataset	NA	NA	percentage of confirmed positive family history	pathologically confirmed positive family history was s ignificantly higher in BE/EAC patients compared to controls (6.5% versus 0.9%; p < 0.001)

Peters (2021)	Systematic review and meta-analysis	NA	NA	1,623 BE and 998 EAC patients	NA	BE/EAC or family history for BE/EAC	Self-reported family history but unconfirmed with medical records, sefl-reported family history confirmed with medical records/pathology, screening endoscopy	Many different protocols in the studies included (some did not have a control group). The meta- analysis attempts to compare patients with and without family history of BE/EAC	NA	BE/HGD-EAC	Pooled prevalence of a positive family history of BE/EAC among patients with BE was 8.8% (95%CI 5.5, 13.8). Pooled prevalence of a positive family history of BE/EAC among patients with EAC was 4.4% (95%CI 2.2, 8.7). Patients with family history positive for BE/EAC were three times more likely to have Barrett's esophagus (relative risk [RR] 3.26; 95%CI 1.43, 7.4) and two times more likely to have EAC (RR 2.19; 95%CI 1.14, 4.21)	
Glamour (2022)	Cross-sectional observational study	NA	NA	A total of 500 individuals affected with either HGD-BE or EAC	NA	HGD-BE/EAC with or without family history	Comparison of age of diagnosis of EAC in non-familial EAC (i .e. probands with EAC with no known family history of BE/EAC), duplex families (i .e. probands with EAC and only one relative positive for BE/EAC) and multiplex families (i .e. probands with EAC and at least two relatives positive for BE/EAC)	Review of medical records. No active intervention	NA	Age of diagnosis of EAC	There was a statistically significant difference for age of diagnosis for individuals in the multiplex families compared to the non-familial and duplex families (56.0 versus 64.3, 63.5; p = 0.049)	

| Table 6 | | | **Role of advanced imaging in detection and characterization** | | | | | | | | | | |

Author (year)	Methods			Population			Advance imaging	Intervention				Outcomes		Remarks	Notes
	Design	Randimsation / blinding		N	Age	Inclusion criteria		Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest			
Smith (2019)	cross-sectional	NA	NA	894	65	clinically indicated upper endoscopy	Volumetric laser endomicroscopy	1) All patients underwent WLE followed by VLE. 2) When a lesion was identified on VLE, the investigator would triangulate the location of the lesion by recording the distance and clockface registered with the WLE orientation. This information then was used to guide the investigator to acquire the tissue using WLE. 3) Investigators were given a questionnaire about if VLE guided tissue sampling or therapeutic decisions and if VLE identified suspicious areas not seen on WLE or other advanced imaging modalities.	NA	NA	=	Suspicion lesion detection by VLE Neoplasia detection alone by VLE (not identified by WLE nor random biopsies) VLE-guided tissue acquisition alone Diagnostic yield improvement with VLE	neoplasia detection alone=20 (2%) VLE-guided tissue acquisition=26 (3%) DY improvement= 55% (27 neoplasia found on VLE alone/49 patients found on WLE)	
Beg (2017)	cohort	NA	NA	1022		all patients with established BE confirmed histologically by the presence of IM	NBI and AAC	1) all patients with established BE were recorded in a prospectively maintained database. 2) use of AAC and NBI between 2011-2014	1) all patients with established BE were recorded in a prospectively maintained database. 2) use of WLE alone between 2007-2010	2007 - 2014	=	Dysplasia detection	Chromoendoscopy:28/170 (16.5%) vs WLE: 26/206 (12.6%), p=0.3	
Latos (2019)	cohort	NA	NA	24	38.7	not stated	autofluorescence endoscopy	1) None had any obvious mucosal abnormalities in WLE. 2) Biopsies randomly taken according to Seattle Protocol followed by AFE measurements taken at sites according to Seattle protocol	=	10 years		odds ratio fo dysplasia detection	6.443 (CI 2.8–15; p<0.0001)	
Everson (2018)	cohort	NA	NA	41	not stated	Patients attending 1 of 3 European referral centers for surveillance of at least C1M2 BE, were enrolled between February 2016 and October 2017.	iScan Optical Enhancement system	1) All examinations were recorded in HD-WLE and iScan OE before biopsy or endoscopic resection of suspicious areas. 2) In the majority of cases, tissue was resected by EMR or alternatively were sampled by forceps biopsy in accordance with Seattle protocol. 3) A total of 262 images were included for analysis (130 HD-WLE and 132 OE, mean and median 3 images per patient, range 1-5). 4) One of 2 experts assessed each image and delineated areas that represented histology-confirmed dysplastic 5) A second group of 7 experts and 7 trainee endoscopists were asked to individually assess each image. To simulate the selection of a site to target biopsy in the clinical setting, endoscopists were instructed to place a single marker on each image over the area that they believed was most likely to yield a biopsy with BE neoplasia	=	1 year	0	sensitivity, specificty, PPV, NPV, accuracy of HD-WLE and OE and Odds ratio (95% CI)	Sensitivity: HD-WLE 69% (600/868) VS OE 78% (739/938); OR 2.03 (1.57-2.63) Specificity: WLE 70% (668/952) VS OE 81% (694/854); OR 2.10 (1.64-2.70) PPV 68% (600/884) VS OE 82% (739/899); OR 2.14 (1.70-2.69) NPV: WLE 71% (668/936) VS OE 78% (694/893); OR 1.41 (1.13-1.74) Accuracy: WLE 70% (1268/1820) VS OE 80% (1433/1792): OR 1.84 (1.56-2.18)	
Verna (2014)	cohort	NA	NA	54	63	All patients scheduled for endoscopic examination from January2010 to April 2012 with a first diagnosis of BE were included inthe study.	iScan	1) After a baseline endoscopic examination, 3% acetic acid wassprayed onto the columnar lined oesophagus. 2) Then, combining the useof zooming, SE, CE and TE, an accurate evaluation was performedlooking for any irregularity or altered pattern of the metaplasticmucosa. All these areas were biopsied and labelled as "targetedbiopsies".		not stated		dysplasia detection	The dysplasia detection rate with targeted biopsies was 5.5%, while with random biopsies was 16.6%.	

Study	Type			N	Age	Inclusion criteria	Intervention	Description				Outcome	Results	Notes
Xiong (2017)	SR and meta-analysis	NA	NA	251	65-69	Studies were included if they met the following criteria:(1) evaluated the accuracy and/or additional detection rate(ADR) of NBI and CLE for the diagnosis of HGD/EAC inpatients with Barrett's esophagus; (2) provided effectivecomparison groups (NBI vs. CLE); (3) reported sufficient datato construct diagnostic 2 × 2 contingency tables of NBI andCLE at a per-patient and/or per-lesion level (a lesion wasdefined as a biopsy specimen or a biopsy location); (4) usedhistological biopsy as the standard diagnostic criterion forBE neoplasia.	NBI and confocal laser endomicroscopy	Five studies, involving a total of 251 patients, reported within-patient comparisons of NBI and CLE in thedetection of esophageal neoplasia, including HGD/EAC inBarrett's esophagus. Four studies were conductedin a single center [28—31]. All studies included were pub-lished after 2010. Four eligible studies, [29—32] involving213 patients and 2,272 lesions, reported the diagnostic accu-racy of NBI and CLE in the detection of esophageal neoplasia,including HGD and EAC in BE at the per-lesion level.	=	not stated	=	pooled additional detection rate and diagnostic accuracy,	Pooled ADR of CLE compared to NBI: 0.19 (0.05 - 0.33) Pooled ADR of CLE 19.3% (95% CI: 0.05—0.33, I2 = 74.6%) Pooled sensitivity: NBI 62.8% (95% CI: 0.56—0.69, I2 = 94.6%) VS CLE 72.3% (95% CI: 0.66—0.78, I2 = 89.3%) Pooled specificitY: NBI 85.3% (95% CI: 0.84—0.87, I2 = 92.1%) vs CLE 83.8% (95% CI: 0.82—0.85, I2 = 96.8%)	
Kato (2017)	cohort	NA	NA	214	59	Two experienced endoscopists (K.G. and M.K.) retrieved and selected still images of magnification NBI endoscopy (M-NBI) from patients with BE who were endoscopically diagnosed according to the Japanese classification	NBI	Two hundred and forty-eight HM-NBI images from macroscopically normal areas in patients with BE were retrieved from endoscopy databases and randomized for review by four endoscopists (two experts, two nonexperts). According to protocols from previous studies, we applied the simplified binary classification of NBI surface patterns to our study. The simplified NBI classification was created by integrating surface patterns of the three major NBI classifications (regular and irregular patterns of mucosal and vascular patterns).	=	not stated	=	inter- and intra-observer agreement of the interpretation of NBI surface patterns and the predicted histology (dysplasia vs. non-dysplasia)	The overall inter-observer agreements for mucosal pattern = 0.73) The overall inter-observer agreements for vascular pattern= 0.71 The overall inter-observer agreements for for predicting dysplastic histology = 0.80 The overall intra-observer agreement for mucosal = 0.84 The overall intra-observer agreement for vascular patterns = 0.86 The overall intra-observer agreement in predicting dysplastic histology = 0.89 The mean accuracy in predicting dysplastic histology was 95 % (experts: 96.8 %, non-experts: 93.1 %).	
Wolfsen (2015)	cohort	NA	NA	100	66.4	Study patients had either known or suspected BE and were referred for upper endoscopy. These included patients undergoing diagnostic upper endoscopy for the evaluation of either suspected or confirmed BE and follow-up of prior endotherapy for BE.	Volumetric laser endomicroscopy	Once the VLE imaging procedure was complete, the physician removed the Nvision VLE Optical Probe with balloon from the working channel of the endoscope and proceeded with further endoscopic examination and routine biopsy and/or EMR.	=	not stated	=	neoplasia detection	The final pathologic diagnoses of the patients studied were adenocarcinoma (4 patients), high-grade dysplasia (10 patients), low-grade dysplasia (11 patients), indefinite (5 patients), intestinal metaplasia (29 patients), and normal squamous cells (18 patients).	
de Groof (2020)	cohort	NA	NA	20	67	Patients were eligible for inclusion when referred for endoscopic workup of early Barrett's neoplasia (n Z 10), defined as histologically confirmed high-grade dysplasia or EAC, or referred for surveillance endoscopy of NDBE (n Z 10), defined as the histologically confirmed absence of dysplasia in previous endoscopies.	Deep learning algorithm detection		=	not stated	=	Accuracy, sensitivity, and specificity	Accuracy=90%,, sensitivity=91%,, and specificity=89% 9 of 10 neoplastic patients were correctly diagnosed. 9 of 10 NDBE patients were correctly diagnosed.	
Curvers (2011)	Multicenter, randomized, crossover study	randomized	? not blinded ?	99	63	?	Autofluorescence imaging	BE patients with confirmed low-grade intraepithelial neoplasia (LGIN) underwent both ETMI and SVE in random order (interval 6-16 weeks). During ETMI, BE was inspected with high-resolution endoscopy followed by autofluorescence imaging (AFI). All visible lesions were then inspected with narrow-band imaging. During ETMI and SVE, visible lesions were sampled followed by 4-quadrant random biopsies every 2 cm.		?		histological yield	TMI had a significantly higher targeted histological yield because of additional detection of 22 lesions with LGIN/high-grade intraepithelial neoplasia (HGIN)/carcinoma (Ca) by AFI. There was no significant difference in the overall histological yield (targeted+random) between ETMI and SVE. HGIN/Ca was diagnosed only by random biopsies in 6 of 24 patients and 7 of 24 patients, with ETMI and SVE, respectively.	abstract only; no access to the main article

Study	Type	Randomized	Blinding	N	Age	Patients	Technique	Methods	=	Follow-up	=	Outcome	Results
Curvers (2008)	cohort	NA	NA	14	73	Patients were eligible for this study when sched- uled for either a surveillance endoscopy of a known BE or referred for workup of high-grade intraepithelial neopla- sia (HGIN) or early cancer (EC).	NBI, indigo carmine and AAC	1) a number of randomly selected areas (upon the discretion of the endoscopist) were stabilized using the distal attachment cap and magnified for visualization of the mucosal morphology using WLE, NBI, ICC (0.5% indigo carmine), and AAC (2% acetic acid), respectively. 2) s. Seven endoscopists with no specific exper- tise in Barrett's esophagus or advanced imaging tech- niques and 5 international experts in this field evaluated these 22 areas for overall image quality, mucosal image quality, and vascular image quality	=	not stated	=	Yield of identifying HGIN/EC (%) and interobserver agreement	Yield of identifying HGIN/EC: WLE=86% vs WLE+NBI=84 VS WLE+ICC=70 VS WLE-AAC=83% Interobserver agreement: WLE=0.53 (0.50–0.57) VS WLE+NBI=0.39 (0.35–0.43) VS WLE+ICC=0.46 (0.42–0.50) VS WLE-AAC=0.42 (0.38–0.46)
Gross (2017)	Clinical trial	not randomiz ed	2nd pathologist was blinded	4203	59	All patients over age 18 who were scheduled for upper gastrointestinal endoscopy to evaluate symptoms of gastroesophageal reflux and suspected BE were enrolled.	WATS	Immediately after obtaining WATS specimens, investigators performed four-quadrant FB of the esophagus at 1 cm to 2 cm intervals and tested any endoscopically visible mucosal abnormality. All the FB samples were obtained after WATS sampling	=	not stated	=	Increased detection of LGD	Of the 33 patients diagnosed with LGD by WATS (only four had a history of BE and the average suspected length of BE in these patients was 1.4cm), 23 had negative FB results. High-grade dysplasia (HGD)/EAC was not detected by FB in this primarily screening population, although WATS detected a single case of EAC.
Kara (2005)	Clinical trial	randomiz ed	blinded	28	66	Patients were eligible for the study if they were referred to our department for the first time for work−up of recently diagnosed but endoscopically inconspicuous high−grade dysplasia or early cancer	NBI and indigo carmine	1)Patients with Barrett's esophagus underwent HRE−ICC and HRE−NBI (separated by 6±8 weeks) in a randomized sequence. 2)The 2 procedures were performed by 2 different endoscopists, who were blinded to the findings of the other examination.		not stated		sensitivity Lesions detected with chromoendoscopy or narrow−band imaging.	sensitivity for HGD/EC: HRE−ICC=93% VS HRE−NBI=86% VSHRE=79% (p=1.0) Lesions additionally detected with chromoendoscopy or narrow−band imaging: HRE−ICC=0/14 (0 %) HRE−NBI=0/14 (0 %) (p=1.0)
Shah (2018)	cohort	NA	The expert endomicrosco pist was blinded to endoscopic images, real-time pCLE interpretation, and pathology interpretations.	66	66	Patients referred for surveillance endoscopy for BE	Confocal laser endomicroscopy	1) Consecutive patients referred for surveillance endoscopy for BE underwent HD-WLE and NBI gastroscope. Areas suspicious for dysplasia were identified on NBI 2) Following visual examination, pCLE examination was performed using a 2.5mm gastroflex ultra-high definition probe		not stated		Sensitivity, Specificity, NPV, PPV for LGD and HGD	HGD: Sensitivity 67% (CI 9 – 99 %) Specificity 94% (85 – 98 %) NPV 33% (12 – 63%) PPV 98% (92 – 100%)
Jayasekera (2012)	cross-sectional	NA	NA	50	66	Patients included in this cross-sectional study had been referred for endoscopic evaluation and treatment of dysplastic Barrett's esophagus, which had been previously diagnosed by their referring physician.	Confocal laser endomicroscopy + NBI	A prediction of likely histology was made for each mucosal point (four-quadrant every 1cm and any mucosal point (four-quadrant every 1cm and any visible mucosal abnormality), first with HD-WLE, followed by NBI, and finally CLE. Biopsies were taken at all of these points.		NA		Sensitivity, specificity, and accuracy	HD−WLE: sens 79.1 %, spe 83.1 %, and accuracy 82.8 % NBI: sens 89.0 %, spe 80.1 % and accuracy 81.4 % CLE: sens 75.7 %, spe 80.0 %, and accuracy 79.9 %
Ormeci (2007)	cohort	NA	NA	109	62	Patients older than 18 years with an indication for eso-phagogastroduodenoscopy were selected for this study.	ICC	The patients for these groups were selected at the conventional endoscopy, and then chromoendoscopy was performed. The esophagus was stained with methylene blue, after which six biopsies were taken from stained and unstained areas.				Sensitivity andspecificity	dysplasia: sens 0.68 (0.46–0.85) and spe 0.77 (0.67–0.84) EAC: sens 0.95 (0.75–0.99) and spe 1.00 (0.95–1.00) (~WLE, p>0.05) ICC provided a full diagnosis of intestinal metaplasia and esophageal carcinoma by 6 biopsies instead of 12.
Singh (2008)	cohort	NA	NA	109	62	Patients known to have Barrett's esophagus and un−dergoing surveillance endoscopy or referred from other centers for workup of recently diagnosed dysplasia	NBI	1) mucosal patterns visualized in Barrett's esophagus on NBI−Z were classified into four easily distinguishable types: A, round pits with regular microvasculature; B, villous/ridge pits with regular microvasculature; C, absent pits with regular microvasculature; D, distorted pits with irregular microvasculature. 2) The NBI−Z grading was compared with the final histopathological diagnosis.		1 year		PPV and NPV accuracy inter- and intra-observer agreement	The PPV and NPV for type A pattern (columnar mucosa without intestinal metaplasia) were 100% and 97% respectively; for types B and C (intestinal metaplasia) they were 88% and 91% respectively, and for type D (high−grade dysplasia) 81% and 99% respectively. The mean accuracy of prediction of the pit patterns by the non- NBI−experts endoscopists was 83.9% (428/510); in the expert group it was 90.0% The mean k values of inter- and intra-observer agreement were 0.71 and 0.87 in the non−expert group; 0.78 and 0.91 in the expert group.

Author (year)	Study type			N	Age	Inclusion	Imaging	Methods		Outcome	Results
Sharma (2016)	cohort	NA	NA	97	65.7	Adult patients undergoing surveillance or endoscopic treatment for BE at 4 institutions in the United States and Europe after respective permissions from the Institution Review Boards of each institution. Patients included in the study were required to be able to tolerate oral proton pump inhibitors and to discontinue the use of aspirin, nonsteroidal anti-inflammatory drugs, or clopidogrel 7 days before and after all endoscopic ablation procedures.	NBI	1) The BING working group, composed of NBI experts from the United States, Europe, and Japan, met to develop a validated, consensus-driven NBI classification system for identifying dysplasia and cancer in BE. The group reviewed 60 NBI images of nondysplastic BE, highgrade dysplasia, and esophageal adenocarcinoma to characterize mucosal and vascular patterns visible by NBI; these features were used to develop the BING criteria. 2) We then recruited adult patients undergoing surveillance or endoscopic treatment for BE at 4 institutions in the United States and Europe, obtaining high-quality NBI images and performing histologic analysis of biopsies. 3) Experts individually reviewed 50 NBI images to validate the BING criteria, and then evaluated 120 additional NBI images (not previously viewed) to determine whether the criteria accurately predicted the histology results	not stated	Sensitivity, specificity, and accuracy	accuracy=85%, sensitivity=80%, specificity=88%, positive predictive value=81%, negative predictive value=88%
Bertani (2013)	cohort	NA	NA	100	59-60	This single-center study was conducted in patients who had previously received a diagnosis of BE without dysplasia and had accepted to undergo our BE surveillance program.	Confocal laser endomicroscopy	Patients were then assigned by a computer-generated randomization list to undergo HD-WLE only or pCLE in addition to HDWLE evaluation. In both groups, any visible mucosal abnormality suspicious for neoplasia was identified and biopsied; then, random four-quadrant biopsies at 1-cm interval were performed according to the Seattle biopsy protocol.	not stated	sensitivity, specificity, positive and negative predictive value	In the HD-WLE group, areas suspicious for neoplasia were not observed and dysplasia was diagnosed in 5/50 (10 %) patients (one with HGD). In the pCLE group, areas suspicious for neoplasia were observed by pCLE in 21/50 (42 %) patients; dysplasia was confirmed in 14 cases (28 %) (two with HGD). The dysplasia detection rate was significantly higher in the pCLE group than in the HD-WLE group (P = 0.04). The sensitivity, specificity, positive and negative predictive values of pCLE for dysplasia were 100, 83, 67, and 100 %.
Boerwinkel (2013)	Review of databases	NA	NA	374	65	not stated	Autofluorescence imaging	The original databases of 5 prospectively conducted trials on AFI were retrieved and assessed for patient demographics, endoscopic data, and pathology records. Data from the 5 databases were pooled into a single database	not stated	additional neoplasia detection	In 33 patients, AFI detected additional HGINs or IMCs next to lesions detected by primary white-light endoscopy
Sharma (2012)	Clinical trial	randomized	blinded	123	61	Patients over the age of 18 undergoing screening or surveillance for BO were prospectively enrolled at three tertiary referral centres	NBI	1) This was a multicentre, randomised, crossover trial comparing HD-WLE and NBI for the detection of IM and neoplasia in patients with BO. 2) Patients meeting the inclusion criteria were randomised to undergo endoscopy with either HD-WLE or NBI on day 1. Patients were randomised in a 1:1 ratio using a computergenerated list of random numbers and administered by study coordinators in sealed opaque envelopes that were opened after patient enrolment and immediately before the first study procedure. 3) The same patient then returned for the alternative procedure performed by a different endoscopist within 3-8 weeks. 4) During HD-WLE, four quadrant biopsies every 2 cm, together with targeted biopsies of visible lesions (Seattle protocol), were obtained. During NBI examination, mucosal and vascular patterns were noted and targeted biopsies were obtained.	not stated	dysplasia detection	Both HD-WLE and NBI detected 104/113 (92%) patients with IM, but NBI required fewer biopsies per patient (3.6 vs 7.6, p<0.0001). NBI detected a higher proportion of areas with dysplasia (30% vs 21%, p=0.01).
Kara (2006)	cohort	NA	NA	63	65	Patients were eligible for the study if they were scheduled for endoscopic examination because of a BE, with or without a previous history of dysplasia.	NBI	1) After routine inspection of the BE when using WLE, a number of small areas within the Barrett's segment were stabilized by the distal attachment cap and then magnified for visualization of the mucosal morphology. 2) For each investigated area, still images were taken with WLE and with NBI, and were saved on a computer, followed by 1 or 2 biopsies. 3) Early in the study, the images and biopsies of 15 areas were reviewed by the study's expert pathologist and 2 endoscopists in an unblinded manner.	not stated	Sensitivity, Specificity, NPV, PPV for HGD	A sensitivity of 94%, a specificity of 76%, a positive predictive value of 64%, and a negative predictive value of 98% for HGIN.

Study	Design			N	Age	Population	Intervention	Protocol		Outcome	Results	Notes
Bratlie (2015)	clinical trial	randomized	blinded			The study population was recruited from Sahlgrenska University Hospital, a tertiary referral high-volume endoscopy center for patients with suspected or histologically verified BE in the western region of Sweden. From November 2009 until November 2012, all patients between the ages of 20 and 80 years with histologically verified SIM or macroscopically suspected BE were considered for enrollment.	Multiple-band imaging	According to the study-protocol, visible lesions were commented on and biopsied seperately. SDWLE then continued with random 4-quadrant biopsies from the GEJ and every 2 cm of the extent of the BE during retraction. The HDMEMBI investigation started as in standard endoscopy with description of the relevant positions. Then, after switching MBI mode on and magnification of the image, the investigator thoroughly scanned the BE segment in search of irregularities in microvasculature and pit-pattern appearance and only grasping for lesions suspected of having dysplasia		not stated	neoplasia detection	There was no significant difference between groups in diagnostic yield for LGD (14 in HDMEMBI vs 13 in SDWLE) or HGD (4 HGDs were found: 3 using HDMEMBI and 1 using SDWLE) Significantly fewer biopsies were collected during the HDMEMBI procedure (P<.001).	
Kara (2006)	Cross-sectional	NA	NA	20	66	Patients with BE were included if they were scheduled for endoscopy because of (1) workup for a recently diagnosed (suspected) HGIN that, according to the referring physician, was endoscopically inconspicuous, or (2) follow-up after endoscopic therapy for HGIN.	Autofluorescence imaging	Twenty patients with BE with suspected or endoscopically treated HGIN were investigated with 2 prototype imaging systems: AFI (inspection with high-resolution videoendoscopy and autofluorescence imaging for detection of lesions) and NBI (for detailed inspection of mucosal and vascular patterns of identified lesions). Lesions were sampled for histopathologic evaluation		not stated	(1) the positive predictive value (PPV) of AFI alone and of AFI-NBI for detecting HGIN (2) the reduction of the total number of false-positive lesions (ie, lesions suspicious with AFI but without HGIN in corresponding biopsy specimens) because of the use of NBI.	All of the 28 lesions with HGIN were identified with AFI. Forty-seven suspicious lesions were detected with AFI: 28 contained HGIN (60%) and 19 were false positive (40%). With NBI, 25 of the true-positive lesions had definitely suspicious patterns, and 3 had dubiously suspicious patterns. Of the 19 false positives, 14 were not suspicious on NBI. The false-positive rate, therefore, was reduced from 40% to 10%. All of the 14 patients with HGIN were identified by AFI-NBI (sensitivity 100%).	
Curvers (2010)	cohort	NA	NA	87	67	All patients with BE referred to the participating centers for work-up of endoscopically inconspicuous HGD/Ca were eligible.	Endoscopic tri-modal imaging incorporates high-resolution endoscopy, autofluorescence imaging, and narrow band imaging.	The ETMI system consists of a high-resolution white-light endoscope with optical zoom (magnification 100 ; XGIF-Q240/260FZ; Olympus Inc, Tokyo, Japan) equipped with an AFI and NBI mode.				ETMI had a significantly higher targeted yield compared with SVE The yield of targeted biopsies of ETMI was significantly inferior to the overall yield of SVE. Detailed inspection with NBI reduced the false positive rate of HRE AFI from 71% to 48% but misclassified 17% of HGD/Ca lesions as not suspicious.	all advance imaging analysed together
Tholoor (2014)	cohort	NA	NA	655	66	Patients 18 years of age and olderwith adiagnosis ofBarrett's esophagus undergoing surveillance gastroscopy	Acetic acid	2 different groups of Barrett's esophagus surveillance: the standardized random biopsy protocol (SBP) and AAC protocol. Standardized: Barrett's mucosa was first assessed in detail with white light. Any visible lesions underwent biopsy separately. Four-quadrant biopsy specimens were then taken every 2 cm starting at the gastroesophageal junction and moving upward in a systematic fashion to the squamocolumnar junction. AA: Barrett's esophagus was first assessed with white-light examination. After the acetic acid spray, the following features were assessed endoscopically: (1) surface pattern (regular, irregular), (2) vascular pattern (regular or irregular), (3) focal loss of acetowhitening reaction (normal or abnormal). Abnormal areas identified after acetic acid spray were sampled in a targeted fashion and sent in a separate cassette. This was followed by 3 nontargeted biopsy specimens, which included 1 specimen from the gastroesophageal junction, 1 from the middle of Barrett's segment, and 1 from the squamocolumnar junction		not stated	gain of neoplasia detection number of biopsies	Gain of eoplasia detection AA 0.13 vs Seatle protocol 0.02, P< 0.0001 The number of biopsies required to detect 1 neoplasia was 15 times lower in the AAC cohort (40 biopsies) than in the SBP cohort (604 biopsies).	

Study	Type			N	Age	Description	Technique	Methods		Outcome	Outcomes	Results	Notes
Vázquez-Iglesiasa (2007)	cohort	NA	NA	100	53	One hundred upper endoscopies performed on 100 patients under follow-up were included in this study.	Acetic acid	After an initial endoscopic examination, 3% AA was sprayed onto the CLE, starting at 2 cm proximal to the squamocolumnar junction. Biopsy specimens were obtained by the standard followup protocol. Specimens were also taken from rough or irregular areas when present.		not stated	Sensitivity, Specificity, NPV, PPV for neoplasia detection	The percentage of dysplasia and adenocarcinoma in biopsy specimens was significantly higher in patients with rough or irregular areas (86.7%) than in those with normal uniform reticulum (0%) (P < 0.001). Sensitivity=100% (95% CI 71.7–100%). Specificity= 97.7% (91.2–99.6%) Positive predictive value=86% (58.4–97.7%) Predictive negative value=100% (94.6–100%).	
Raphael (2019)	cohort	NA	NA	138	65.2	Patients included in this study had a known diagnosis of BE, were older than 18 years, were undergoing scheduled surveillance.	Wide-Area Transepithelial Sampling	consecutive patients who underwent an examination that consists of high-definition white light endoscopy (HDWLE), narrowband imaging (NBI), volumetric laser endomicroscopy (VLE), and Seattle protocol (SP) biopsies (collectively termed HDWLE-NBI-VLE-SP examination). Raised lesions were removed by endoscopic resection. Areas suspicious for dysplasia on NBI and VLE were biopsied. This was followed by random biopsies and WATS-3D brush biopsies.		not stated	neoplasia detection	When restricting the analysis to LGD and higher, 21 dysplastic cases (15% of the total cases) were identified by HDWLE-NBI-VLE-SP, while WATS-3D found 4 additional new cases (3 with LGD and 1 with highgrade dysplasia) for an added yield of 19% (54/21, 95% confidence interval 0.6%–45.7%).	
Zhang (2012)	Meta-analysis	NA	NA	502	61 - 66	not stated	NBI	The following criteria were used to include published studies: (i) prospective study where NBI was used to assess BE, (ii) compared this with histology as the gold standard, (iii) description of typical mucosal pit patterns and vascular patterns, (iv) availability of adequate data, and (v) published as a full article or abstract.		not stated	Sensitivity and specificity for neoplasia Diagnostic odds ratio	Pooled sensitivity 0.91 (95% CI = 0.75–0.98) Pooled specificity 0.95 (95% CI = 0.91–0.97) Pooled diagnostic OR=66.65 (95% CI 6.52–680.83)	
Pascarenco (2016)	cohort	NA	NA	84				Every patient underwent a WLSE with random biopsies and after 4–6 weeks, a NBI examination was performed.				NBI detected significant more patients with low grade dysplasia (LGD) (7.1% vs. 0%; p = 0.03).	abstract only; no access to the main article
4	Clinical trial	randomized	NA	200	66	Inclusion criteria 1. aged 18 years or above 2. biopsy (histologically) proven Barrett's metaplasia 3. at least 2-cm length of Barrett's esophagus 4. informed consent	Acetic acid	1) We developed a training module using the PREDICT classification for recognition of Barrett's neoplasia with acetic acid 2) : Cross-over study, with each patient undergoing two gastroscopies 6-8 weeks apart, acting as their own control and randomised to either Seattle protocol gastroscopy (non-targeted mapping biopsies) or acetic acid assisted gastroscopy (targeted biopsies - Portsmouth protocol) first			neoplasia detection number of biopsies cost savings	Similar neoplasia detection: High grade dysplasia and cancer were detected with both protocols. Five low grade dysplasias were detected (two with acetic acid, four with nontargeted biopsies; one lesion was detected with both techniques). The paired analysis demonstrated a 6.5-fold decrease in the number of biopsies per pathology found by Portsmouth protocol as compared to the Seattle protocol, and a 9.6-fold decrease when restricted to high risk neoplasia (HGD and cancer). The total histopathology costs of Seattle protocol for the cohort was £125,987. In contrast, the cost for Portsmouth Protocol targeted biopsies was just £13,311, representing a 9.5x difference	ABBA study
Hajelssedig (2021)	Meta-analysis	NA	NA	493	57.4-66	not stated	NBI	Inclusion criteria are as follows: 1 prospective randomized clinical trials and controlled observational cohort studies; 2 studies that evaluated dysplasia or EAC in BE as primary or secondary outcomes; and 3 studies that compared NBI against WLE random biopsy or dye-based/virtual chromoendoscopy.		not stated	Sensitivity, Specificity, NPV, PPV for dysplasia	Sensitivity of 76% (95%CI:0.61–0.91) Specificity of 99% (95% CI:0.99–1.00) Positive predictive value of 97% (95% CI: 0.96–0.99) Negative predictive value of 84% (95% CI: 0.69–0.99)	

Study	Study type			N	Age	Population	Technique	Methods		Outcomes	Results
Furneri (2019)	Cost-consequence study	NA	NA	NA	NA	NA	NBI	A combined decision tree / Markov model approach was adopted to undertake cost-conse- quence, budget impact, and cost-effectiveness analyses of NBI with target biopsies. The decision tree approach was used to model the diagnostic and surveillance phases. Outputs from a literature review (conducted in May 2017) were used to identify relevant clinical inputs for the model using PICO query: Q1) What is the epidemiological burden and prognosis associated with BE?; Q2) What is the key evidence (from clinical trials, observational studies, reviews, guidelines) associated with NBI or other alternative techniques in the diagnosis and treatment of BE?; Q3) What is the compar- ative sensitivity, specificity, treatment effectiveness of NBI vs alternative techniques?; Q4) What are the main economic implications of BE management with NBI (vs alternative tech- niques) in the diagnosis / treatment of BE? The model estimated total costs (i.e. staff and overheads; histopathology; adverse events; capital equipment) and clinical implications of monitoring a cohort of patients with known/suspected BE, on an annual basis. In the simulation, BE patients (N = 161,657 at Year 1; estimated annual increase: +20%) entered the model every year and underwent esophageal endoscopy.	NA	cost reduction adverse events reduction	cost reduction of £458.0 mln Reduction of biopsies also determined savings for avoided adverse events (-£21.1 mln).
Kandiah (2017)	cohort	NA	NA	NA	NA		Acetic acid	The study was conducted in four phases: phase 1—development of component descriptive criteria; phase 2—development of a classification system; phase 3—validation of the classification system by endoscopists; and phase 4—validation of the classification system by non-endoscopists.	NA	Sensitivity, Specificity, NPV, PPV for dysplasia	In phase 3, the application of PREDICT (Portsmouth acetic acid classification) by endoscopists improved the sensitivity and negative predictive value (NPV) from 79.3% and 80.2% to 98.1% and 97.4%, respectively (p<0.001). But no difference in Specificity (76.2 VS 68.1 (186/273), p=0.03) nor in PPV (75.8 vs73.6, p=0.53). In phase 4, the application of PREDICT by non-endoscopists improved the sensitivity and NPV from 69.6% and 75.5% to 95.9% and 96.0%, respectively (p<0.001). But no difference in Specificity (86.7 vs 93.3, p=0.13) nor in PPV (84.0 vs 93.6 , p=0.05).
Sharma (2011)	Clinical trial	randomized	pathologists were blinded	101	65.1	Consecutive patients undergoing BE surveillance and/or referred for BE-associated neoplasia (HGD/EC) evaluation and treatment were prospectively enrolled in this trial at 5 hospitals	confocal laser endomicroscopy	All patients were examined by HD-WLE, narrow-band imaging (NBI), and pCLE, and the findings were recorded before biopsy samples were obtained. The order of HDWLE and NBI was randomized and performed by 2 independent, blinded endoscopists. All suspicious lesions on HD-WLE or NBI and 4-quadrant random locations were documented. These locations were examined by pCLE, and a presumptive diagnosis of benign or neoplastic (HGD/EC) tissue was made in real time.		Sensitivity and specificity	The sensitivity and specificity for HD-WLE were 34.2% and 92.7%, respectively, compared with 68.3% and 87.8%, respectively, for HD-WLE or pCLE (P = .002 and P < .001, respectively). HD-WLE or NBI or pCLE was more sensitive (P = .01) and less specific (P = .02) compared with HDWLE or NBI, and it had similar sensitivity (P = .25) and specificity (P = .05) compared with HD-WLE or pCLE.
Canto (2014)	Clinical trial	randomized	pathologists were blinded	192	not stated	consecutive adult patients undergoing outpatient endoscopy for either routine surveillance of BE (surveillance group) or suspected or biopsy-proven unlocalized BE-associated HGD and/or early intramucosal ECA (neoplasia group) referred for confirmation of diagnosis and/or endoscopic therapy	confocal laser endomicroscopy	All study patients had careful examination of the upper GI tract with a HDWLE upper endoscope. HDWLE: After examination with HDWLE, endoscopic diagnoses were recorded in real time based on the appearance of the BE mucosa. HDWLE'eCLE: immediately after HDWLE, fluorescein-aided eCLE13,14 imaging was performed with the endomicroscope	not stated	number of mucosal biopsies diagnostic yield for neoplasia sensitivity and specificity	HDWLE+eCLE+TB led to a lower number of mucosal biopsies (reduction in median biopsy number was from 6 to 3 in the neoplasia group (p=0.0001) and 3 to 1 in the surveillance group (p<0.0001)), higher diagnostic yield for neoplasia (34% vs. 7%, p<.0001) The addition of eCLE to HDWLE increased the sensitivity for neoplasia detection to 96% from 40% (p<.0001) without significant reduction in specificity.

Author (year)	Study type	Randomization	Blinding	N	Age	Inclusion criteria	Technique	Methods		Outcome	Results
Lim (2008)	Clinical trial	randomized	2nd endoscopist was blinded	30	67	All surveillance patients with a diagnosis of dysplasia in Barrett's esophagus from 2001 to 2003 and below 80 years of age were identified from computerized histopathology records.	methylene blue - MBDB	Patients were assigned to undergo either MBDB or RB as the first examination by using a table of random numbers. One endoscopist carried out all the MBDB to maintain visual consistency. The second endoscopist was blinded to the histology from the first endoscopy.	not stated	dysplasia detection	MBDB and RB were concordant in 3 patients with LGD and 2 patients with HGD. MBDB missed 7 patients with LGD and 4 patients with HGD. Overall, dysplasia was identified in 17 of 18 patients by RB and in 9 of 18 by MBDB.
Ragunath (2003)	Clinical trial	randomized	blinded to previous histology			Adults patients undergoing endoscopy	methylene blue - MBDB	Patients were assigned to undergo either MBDB or RB as the first examination by using a table of random numbers. One endoscopist carried out all the MBDB to maintain visual consistency. The second endoscopist was blinded to the histology from the first endoscopy.	not stated	Sensitivity and specificity dysplasia detection	Sensitivity 49% (95% CI 38-61%) Specificity 85% (95% CI 82-68%) There were no significant differences in the diagnostis of dysplasia and carcinoma between MBDB 12% and random biopsy 10%
Jin (2015)	cohort	NA	NA	50	29-78	The inclusion criteria were as follows: 1) Patients with gastroscopic biopsy for suspicious BE and being followed-up thereafter; 2) positive pathological confirmation, while conventional endoscopy was negative or suspicious; 3) patients subjected to endoscopic BE treatment (argon plasma coagulation) and followed up.	autofluorescence endoscopy	The suspicious lesions were rinsed with saline, and then the AFI was conducted to observe the color changes of suspicious lesions. The observation plane depth was adjusted to make the image clear. The AFI images were acquired and stored for the fluorescence diagnosis of suspicious lesions. Then, NBI with zooming function was performed to observe the shape and structure of gastric mucosal microvessels and pits	not stated	dysplasia detection	The false-positive rates decreased from 40.5% of AFI to 9.5% (7/74) of NBI-AFI (p<0.05). The positive predictive value of AFI in BE HGIN was 59.5% (44/74), while that of AFI-NBI combination was 84.8% (39/46; p<0.05).
Réaud (2006)	cohort	NA	NA	28	59.7	They presented with either proven or suspected BE and had an indication for upper gastrointestinal endoscopy for confirmation and/or monitoring.	Acetic acid	Twenty-eight patients were studied with magnifying chromoendoscopy. Endoscopy biopsies were performed on one or several zones of BE chosen random.	not stated	Sensitivity and specificity	sensibility and specificity were respectively 95.5% and 42.9%. Among the six biopsies that showed high-grade dysplasia, three were suspected
Fortun (2006)	cohort	NA	NA	64	62	Consecutive unselected patients referred to a specialist BE clinic with an endoscopic diagnosis of BE were invited to undergo EME for confirmation of diagnosis and formation of a surveillance.	Acetic acid	A routine upper gastrointestinal endoscopy was first performed. Following this, 10–20 mL of 3% acetic acid was sprayed on the lower oesophagus and Barrett's epithelium	not stated	The primary endpoint was to determine the diagnostic yield of SIM, dysplasia and adenocarcinoma by EME technique, using the descriptive statistics. As a secondary endpoint, we compared the findings with EME and previous random biopsy	Fifteen patients (24%) had a histological upgrade with enhanced magnification endoscopy: There was a high detection rate of specialized intestinal metaplasia even in short segment Barrett's oesophagus (74%), and additionally, there were two cancers, one with 2-cm Barrett's oesophagus and one ultra-short (1 cm). inter- and intra-observer agreement in assessing the pit patterns were 0.571 (0.041) and 0.709 (0.038)
Ngamrueng phong (2009)	Meta-analysis	NA	NA	450	not stated	not stated	methylene blue - MBDB	For a study to be included in this meta-analysis, predefined criteria had to be met. MB and RB should have been performed on each patient successively in included studies so that the diagnostic yields of the 2 tests were able to be compared under comparable conditions	not stated	increase in diagnostic yield	There was no significant IY with MB over RB for detection of dysplasia (IY 9%; 95% CI, -1% to 20%; 9 studies, n Z 450), and high-grade dysplasia and/or early cancer (IY 5%; 95% CI, -1% to 10%; 8 studies, n Z 405).
Nogales (2017)	cohort	NA	NA	100	58.4	From September 2014 to April 2015, gastroscopy images from 100 consecutive patients with previous or suspected diagnosis of BE were considered for inclusion in our study.	Acetic acid	Eight observers (4 staff endoscopists and 4 trainee endoscopists) evaluated 100 images selected from an anonymized bank of 470 photographs using the BING classification. Observers were to assign their individual assessment of the mucosal and vascular pattern, and pre-diction for dysplasia. Accuracy for dysplasia prediction and intra/interobserver agreement was calculated.	NA	Accuracy, sensitivity, and specificity	accuracy= 81.1%, sensitivity=48.4%, specificity=91%, positive predictive value=61.4 and negative predictive value=85.5%,

Study	Type	Randomization	Blinding	N	Age	Inclusion criteria	Intervention	Methods			Outcomes	Results
Pohl (2007)	clinical trial	randomized	2nd endoscopist was blinded	57	61.4	Patients were eligible if they were referred to our department for the first time for work−up of possible high grade intraepithelial neoplasia (HGIN) or early cancer suspected at the referring center. However, only patients with discrete mucosal alterations or macroscopically occult lesions were selected for this study. Patients were also included if they had been endoscopically treated for HGIN/early cancer previously	Acetic acid vs FICE	Upper gastrointestinal endoscopy was performed twice in each patient, with an interval of 4±6 weeks between procedures. The first procedure was randomly allocated to be either high resolution endoscopy with CAA or high resolution endoscopy with CVC. The CVC technology takes an ordinary endoscopic image from the video processor and arithmetically processes the reflected photons to reconstitute virtual images by increasing the relative intensity of narrowed blue light to a maximum and decreasing the intensity of narrowed red and green light to a minimum. The most recent development is computed virtual chromoendoscopy (CVC) was "Fujinon intelligent chromoendoscopy" (FICE).		not stated	Sensitivity, Specificity, NPV, PPV for dysplasia	I24/57 patients, 30 lesions with HGIN/ early cancer were detected. The sensitivity of targeted biopsies for HGIN/early cancer on a 'per lesion' basis was 87% (26/30) for both CAA and CVC. The positive predictive value was 39% (26/ 66) for CAA and 37% (26/70) for CVC. In the 'per patient' analysis, sensitivity was 83% (20/24) and 92% (22/24) for CAA and CVC, respectively (P = 0.617). Stepwise random four−quadrant biopsies identified only one patient with HGIN/ early cancer that was missed by both, CAA and CVC
Longcroft– Wheaton (2010)	cohort	NA	NA	119	65	All patients undergoing AA dye spray for evaluation of Barrett's esophagus between 2005 and 2008 were recorded prospectively on a computer database.	Acetic acid	Patients were examined with white light gastroscopy and visible abnormalities were identified. Acetic acid (2.5%) dye spray was used to identify potentially neoplastic areas and biopsy samples were collected from these, followed by quadrantic biopsies at 2 cm intervals of the remaining Barrett's mucosa.		not stated	Sensitivity and specificity Improvement in diagnostic detection	Sensitivity=95.5% and specificity=80% for the detection of neoplasia. Correlation between lesions predicted to be neoplasias by acetic acid and those diagnosed by histological analysis (r = 0.98). Significant improvement in the detection of neoplasia using acetic acid compared with white light endoscopy (P = .001).
Bhatti (2021)	Meta-analysis	NA	NA	224	not stated	not stated	autofluorescence endoscopy	All peer-reviewed and preprint original articles that reported the sensitivity and specificity of artificial intelligence-based models on white light endoscopic imaging as an index test against the standard criterion of histologically proven early oesophageal cancer in the background of BE reported as per-patient analysis were considered for inclusion			Sensitivity and specificity	Pooled sensitivity= 0.90 (95% CI, 0.83-0.944) and specificity=0.86 (95% CI, 0.781-0.91) The area under the curve for all the available evidence was 0.88.
Bertani (2013)	cohort	NA	NA	100	58.8-59.7	Patients who receive a diagnosis of BE (esophageal intestinal metaplasia) without dysplasia undergo repeated endoscopic examination after 3 years; patients with lowgrade dysplasia undergo repeated endoscopic examination every 6 months whereas those with high-grade dysplasia undergo endoscopic eradication.	Confocal laser endomicroscopy	Fifty of 100 patients underwent pCLE in addition to HD-WLE. Four-quadrant biopsy specimens according to the Seattle biopsy protocol were obtained in all patients to ensure standard-of-care. One-hundred consecutive patients, all fulfilling inclusion criteria and participating in our BE surveillance program and no one meeting exclusion criteria, accepted to be randomly assigned to HD-WLE evaluation only (HD-WLE group) (50 cases) or to added pCLE evaluation (pCLE group) (50 cases).		not stated	Sensitivity, Specificity, NPV, PPV for dysplasia Dysplasia detection	The sensitivity, specificity, positive and negative predictive values of pCLE for dysplasia were 100, 83, 67, and 100 %, respectively. The dysplasia detection rate was significantly higher in the pCLE group than in the HD-WLE group (P = 0.04).

PICO search string:	((Columnar lined esophagus OR Barrett's Oesophagus OR Intestinal metaplasia esophagus OR Barrett's Esophagus) AND (Advanced endoscopic imaging OR Gastroscopy OR esophagoscopy OR High definition endoscopy OR chromoendoscopy OR Narrow band imaging OR Blue light imaging OR Linked colour imaging OR i-scan OR Acetic acid OR Vinegar OR Methylene Blue OR Inidigo Carmine OR White light endoscopy OR Volumetric laser endoscopy)) AND (neoplasia OR Adenocarcinoma OR Mortality OR Low-grade dysplasia OR High-grade dysplasia OR Cost effectiveness)

	Table 7		**AI in BE**					

Author (year)	Methods		Population				Intervention	Outcomes	
	Design	Randimsation / blinding	N	imaging method	Inclusion criteria		Protocol details	Outcome measures (+ definitions)	Outcomes of interest
Ali° (2021)+AA11:J20	Prospective (videos)	No	131	HD-WLE ; NBI	Dataset 1: surveillance endoscopy 1st evaluation (n=68); Dataset 2: surveillance endoscopy no ttt(n=24). Dataset 3: before and after ttt (n=39)		Comparison between WLE-CAD system and expert (n=2) assessment. Evaluation of surface (phantom)	Scoring of BE length (C&M classification) and surface (cm2)	Marginal relative errors (8% and 7%) on C&M scores , 98.4% accuracy regarding surface of BE
De Groof (2019)	Pilot (ex vivo)	No	60	HD-WLE	Dataset 1: 40 Dysplastic BE patients. Dataset 2: 20 NDBE patients		Comparison between AI and expert assessment (n=6)	Diagnostic performances to detect dysplasia and delineation of suspicious areas (localization score preferred biopsy location a.k.a red-flag indication score)	Accuracy, sensitivity and specificity for detection were 92, 95 and 85%, respectively. The system localized and red- flagged the soft spot in 100% and 90% of cases, respectively
De Groof (2020)	Prospective (real-time videos)	No	20	HD-WLE	10 patients with HGD/Ca vs 10 NDBE patients		Comparison between AI and expert assessment (n=4)	diagnostic performance to detect dysplasia per level and per patient	Se 91%, Sp 89%, accuracy 90% per level
Hussein (2022)	Prospective (videos)	No	119	HD-WLE ; i-scan	Training set: 148,936 video frames from 31 dysplastic patients, 31 NDBE, two normal esophagus; Validation set : 25,161 images from 11 patient videos		Comparison between AI and expert assessment (n=6)	Diagnostic performances to detect dysplasia	The indirectly supervised CNN achieved a per image sensitivity in the test set of 91%, specificity 79%, area under receiver operator curve of 93% to detect dysplasia. Per-lesion sensitivity was 100%. AI outperformed 6 experts
Ebigbo (2019)	Pilot (still images)	No		HD-WLE ; NBI	248 images from 2 centers		Comparison between AI and expert assessment (n=13)	Diagnostic performances to detect dysplasia	HD-WLE: Se 92 and 97%, Sp 88 and 100%; NBI : Se 94%, Sp 80%; AI outperformed 13 experts
Ebigbo (2020)	Pilot (videos)	No		HD-WLE	Dataset 1 (training): 129 images; dataset 2 (validation) 62 mages from 14 patients			Diagnostic performances to detect dysplasia	Se 84%; Sp 100%; Accuracy 90%
Ebigbo (2021)	Retrospective	No	116	HD-WLE	230 still images		Comparison between AI and hstiopathology	Prediction of deep sm invasion (T1a vs T1b)	Se 77%, Sp 64%, accuracy 71%
Hashimoto (2020)	Retrospective (videos)	No	100	HD-WLE ; NBI; standard focus; near focus	Dataset 1 (training): 1374 images; dataset 2 (validation) 458 mages from 14 patients		Comparison between AI and expert assessment (n=2)	Diagnostic performances to detect dysplasia	Se 96.4%; Sp 94.2%; Accuracy 95.4%
Pan (2021)	Retrospective	No	187	HD-WLE	Training set: 354 images from 150 patients; Test set: 89 images from 37 patients		Comparison between AI and expert assessment (n=2)	Identification of BE evaluated by intersection over union (IOU)	IOU were 0.56 (GEJ) and 0.82 (SCJ)

Table 8						**role of inspection time and dedicated lists in barrett's surveillance**				
Author (year)	**Methods**			**Population**		**Intervention**		**Outcomes**			
	Design	**Randimsation / blinding**	**N**	**Age**	**Inclusion criteria**	**Protocol (details) of Intervention**	**Protocol (details) of Comparison**	**Follow-up**	**Outcome measures (+ definitions)**	**Outcomes of interest**	**Remarks**	**Notes**
Neil Gupta(2012)	Post hoc analysis of data obtained from a clinical trial.	single multicentre prospective / nil	112	65.5	referral for BE surveillance or evaluation/treatment of BE-associated neoplasia and the ability to provide informed consent.	Coordinators prospectively recorded the time spent inspecting the BE mucosa with a stopwatch.	historic data	Between November 2008 and September 2009,	Endoscopically suspicious lesions, high-grade dysplasia (HGD)/esophageal adenocarcinoma (EAC).	To evaluate the impact of Barrett's inspection time (BIT) on yield of surveillance.	As the total BIT increased, there was a greater propor- tion of patients found to have an endoscopically suspi- cious lesion (P .001) and a greater proportion found to have HGD/EAC . Endoscopists who av- eraged a BIT of longer than 1 minute per centimeter of BE had a higher endoscopically suspicious lesion detection rate, which translated into a trend toward a higher HGD/ EAC detection rate.	Longer inspection time is associated with increased detection of high-grade dysplasia and esophageal adenocarcinoma in Barrett's oesophagus
Joanne Ooi (2017)	This was a prospective 1-year multicenter study incorporating a historical 5-year cohort. The aim of this study was to assess the effect of dedicated BE surveillance lists on dysplasia detection rate (DDR).	nil / nil	729	62	A total of 729 patients with BE underwent surveil- lance endoscopy between 2007 and 2012.	dedicated endoscopy lists	historical data	dedicated lists information		This study demonstrated that a group of train- ed endoscopists undertaking Barrett's surveillance on dedi- cated lists had significantly higher DDR than a nonspecialist cohort. These findings support the introduction of dedica- ted Barrett's surveillance lists.	The overall detection rate of dysplasia/EAC after central pathology review was 18 % (26 /142) in Group A and 8 % (45 / 587) in Group B (P < 0.001).	Dedicated Barrett's surveillance sessions managed by trained endoscopists improve dysplasia detection rate
Linda. S (2022)	Endoscopists participated in one-hour education on recommended performance measures and endoscopic detection of pre-malignant pathologies. A controlled before-after study was performed, measuring adherence to guidelines and rates of detection of pathology in control and intervention groups.	This was a single-center controlled before-after study. / nil	2719	60 control, 58 interventi on	Consecutive patients (aged 18 years and over) referred for diagnostic gastroscopy were included.	Endoscopist education intervention. Interactive face-to-face 1-hour presentation covering recommendations from BSG and ESGE guidelines, and information on endoscopic screening, diagnosis and management of BE,SD and GIM, educational material, laminated info in endoscoy suites. The pre-education phase was between July 2018 and July 2019. Education 6 session was delivered in November 2019. The post-education phase was between November 2019 and March 2020, and March 2021 and October 2021.	Pre education data	post eductaion data	quality of current endoscopic practice vs BSG and ESGE guidelines	The aims of this study were to evaluate the quality of current Australian UGI endoscopic practice based on BSG and ESGE guidelines and the impact of an endoscopist educational intervention on the quality of diagnostic gastroscopy.	A simple endoscopist education session enhanced quality of UGI endoscopy by improving adherence to BSG and ESGE recommendations and increased detection of clinically significant pathology. A minimum inspection time of 7 minutes was associated with increased diagnostic yield and may be a feasible quality indicator for clinical practice Endoscopist education positively impacted all aspects of endoscopists practice .Longer (□7 minutes) procedures were associated with detection of BE	Quality upper gastrointestinal endoscopy: a prospective cohort study on impact of endoscopist education
James Britton 2018	Patients under- went their surveillance endoscopy on a dedicated BO list or a non-dedicated endoscopy list. This routing process was not randomised or influenced by the study team and occurred purely due to endoscopy capacity and patient availability on dates they were offered their test. We prospectively collected data against the BSG dataset for endoscopy reporting (see online supple- mentary material) while also recording the number of biopsies taken, histology results and appropriate- ness of surveillance intervals. Prospective surveillance data were then compared with each patient's previous surveillance endoscopy. Data are expressed as mean ±SDand percentiles unless otherwise stated. Fisher's exact test was used for comparison of means. A p value of <0.05 was taken to show statistical significance.	nil	326		All patients with BO surveillance between January 2016 and July 2017 at a single NHS district general hospital in the UK were included	single operater with sp interest in BO doing dedicated list	generic service lists	18 months	January 2016 and July 2017.	This study aimed to assess the quality of current surveillance delivery compared with a dedicated service	Histology results from the dedicated and non-dedicated list cohorts revealed similar rates of intestinal metaplasia (79.8% vs 73.1%, p=0.12) and dysplasia/oesophageal adenocarcinoma (4.3% vs 2.6%, p=0.41). The dedicated BO endoscopy list achieved signif- icantly greater adherence to the BSG guideline for endoscopy reporting when compared with both the non-dedicated and retrospective cohorts. In time, a dedicated service may provide a more stable transition to future guidelines, for example, easier adaptation to individual risk stratification models and advanced endoscopic techniques.	Dedicated service improves the accuracy of Barrett's oesophagus surveillance: a prospective comparative cohort study . The post-BSG guideline era of BO surveillance remains suboptimal in this UK hospital setting. A dedicated service appears to improve the accuracy and consistency of surveillance care, although the clinical significance of this remains to be determined.

| | Table 9 | | | | | sedation in detection | | | | | | | |

Author (year)	Methods			Population			Intervention			Outcomes		Remarks	Notes
	Design	Randimsation / blinding		N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest		
Hyun Jik Lee (2015)	retrospective study evaluated 28914 patients older than 20 years who underwent EGD at our insti- tution between January 2011 and December 2011.	nil	nil	18546	>20	patients who had diagnostic EGD:	sedated endoscopy	non sedated endoscopy	12 months		After statistical adjustment for age, sex, and body mass index, minimal change esophagitis, and hiatal hernia were significantly less frequently observed in the sedated EGD group [odds ratio (OR), 0.651; 95% confidence interval (CI), 0.586 to 0.722 and OR, 0.699; 95% CI, 0.564 to 0.866]. Nevertheless, there was no significant difference in other findings at the gastro- esophageal junction, such as reflux esophagitis with Los Angeles classification A, B, C, and D or Barrett's esophagus, between the two groups. Similarly, there were no differences in early gastric cancer, advanced gastric cancer, and gastric ulcer occurrence.	Sedation can impede the detection of minimal change esophagitis and hiatal hernia, but does not influence detec- tion of reflux esophagitis of definite severity and Barrett's esophagus.	
S Subramaniam	A retrospective analysis of all patients who underwent surveillance endoscopy for BE over a 5 year period (2009–2013) in a large district general hospital in North London were identi- fied using the audit tool on Unisoft Endoscopy reporting soft- ware.	nil	nil	181 endoscopies for BE surveillance were performed over 5 years.		patients who underwent surveillance endoscopy for BE over a 5 year period (2009–2013)	73 sedated	71 usedated			The mean LOT for sedated compared with unse- dated endoscopies was 12.47 min and 10.36 min respectively (p = 0.05, confidence interval= –4.23, 0.01). The average number of biopsies in sedated patients was 3.87 and 3.85 in the unse- dated (p = 0.47). The regression was a poor fit (R2 adjusted = –0.00033) and the overall relationship not significant: F (2, 141) = 0.976, p = 0.38. P values for sedation (p = 0.96) and length of BO (p = 0.16) did not achieve significance either.	In our study of patients undergoing endoscopy for BE surveillance, the LOT of endoscopic procedure was greater in patients receiving sedation than unsedated patients. The length of BE or the use of sedation did not have a significant effect on the number of biopsies taken. Sedation use did not affect num- ber of biopsies obtained and therefore may not increase dysplasia detection. We conclude that surveillance for BE patients can be performed without sedation.	DOES USE OF SEDATION AFFECT THE SPEED AT WHICH ENDOSCOPY IS PERFORMED AND NUMBER OF BIOPSIES OBTAINED IN BARRETT'S OESOPHAGUS?

Table 10				age limit for surveillance						

Author (year)	Methods			Population			Intervention			Outcomes		Remarks
	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	
Omidvari, 2021	3 independently developed simulation models of EAC screening and surveillance that are part of the Cancer Intervention and Surveillance Modeling Network (CISNET) of the National Cancer Institute.	N.a.	N.a.	For the base case, we simulated 200 cohorts of US patients diagnosed with NDBE, and followed them until death or age 100 years. Each cohort was defined by a unique combination of starting age (66–90 years), sex (man or woman), and comorbidity level (none, mild, moderate, or severe) (simulation). We used the following 3 models: (1) Microsimulation Screening Analysis model for esophageal adenocarcinoma (MISCAN-EAC) from Erasmus MC University Medical Center Rotterdam and the University of Utah; (2) Esophageal AdenoCarcinoma Model (EACMo) from the Columbia University Medical Center and Massachusetts General Hospital; and (3) Multistage Clonal Expansion for EAC model from the Fred Hutchinson Cancer Research Center (MSCE-EAC).	N.a.	For each cohort, we simulated an additional surveillance at the current age, or no further surveillance. For example, a 70-year-old patient with NDBE with a mild comorbidity level either did or did not receive 1 more surveillance at age 70.	N.a.	N.a.	Using the average results of the 3 models for every cohort, we calculated the number of EAC cases, EAC deaths, life years (LYs), and quality-adjusted life years (QALYs) with and without 1 more surveillance. To estimate the total costs, we calculated the cost of cancer care, surveillance endoscopies, EETs, RFA touch-ups, and treatment of complications (ie, bleeding, perforation, and stricture) from a third-party payer perspective. Subsequently, we calculated incremental costs and QALYs gained from 1 additional endoscopic surveillance at the current age versus not performing surveillance at that age, using the average results. The incremental cost-effectiveness ratio (ICER) of performing a last surveillance was calculated for all 25 potential stopping ages (66–90 years), and the age with the highest ICER just less than the willingness-to-pay (WTP) threshold of $100,000 per QALY gained was considered the optimal age of last surveillance.	We found that for men with NDBE without comorbidity, the optimal age for last surveillance is 81 years, whereas it may be up to 8 years earlier for those with comorbidity. For women, we found that without comorbidity, the optimal age for last surveillance of patients with NDBE is 75 years but can be up to 6 years earlier if patients have comorbidities.	Gastroenterology

Table 11	surveillance intervals

Author (year)	Methods		Population					Intervention			Outcomes		Remarks
	Design	Randimisation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest		Remarks
Vissapragada, 2021	Systematic review	N.a.	N.a.	N.a.	Included studies were cost-effectiveness analyses of EGD surveillance of patients with BE, undertaken from 1975 to 2020, with the specified outcome of QALY or DALY. Studies were limited to those that used NDBE as the starting population and containing at least one strategy varying the interval frequency of EGD surveillance or limiting surveillance to a higher risk NDBE population.	Altered surveillance strategies, reducing low-value care.	Current Surveillance strategies.	N.a.	Incremental cost-effectiveness ratio (ICER) values were extracted from tables. Costs per QALY were calculated using converted dollar costs (converted costs/QALY). Results were not combined for meta-analysis due to heterogeneity in model design and parameters between studies, rather they were graphically represented to allow qualitative comparison.	EGD surveillance at current guideline-recommended intervals is not cost-effective. No surveillance is also not cost-effective at thresholds above $35,000/QALY (converted 2018 value). Modifying surveillance programs by eliminating patients at low risk of progression or increasing the surveillance intervals reduces the number of endoscopies performed and can be cost-effective strategies. Several studies provide evidence of cost-effectiveness with reduced frequency of endoscopy in male (Provenzale, Inadomi, Omidvari et al.) and female cohorts (Omidvari et al.) [27, 28, 31]. Gordon, Lindblad, Das, and Hao present evidence of cost-effectiveness in risk-stratified EGD surveillance programs, but this needs to be further evaluated to ensure appropriate selection and follow up [17, 26, 35, 36].		Surgical Endoscopy
Copidilly, 2018	Systematic Review and Meta-analysis	N.a.	N.a.	N.a.	Articles included in the systematic review and meta-analysis had to report at least 1 of our 4 outcomes of interest, which included (1) the primary outcome of reduction in EAC-related mortality, and/or secondary end points, including (2) all-cause mortality, (3) EAC stage at the time of diagnosis, and (4) receipt of surgical resection/esophagectomy. Additionally, the diagnosis of BE must have occurred at least 6 months before the diagnosis of EAC among patients in the surveillance group.	No surveillance	Current Surveillance strategies.	N.a.	Our primary outcome of interest was comparison of EAC-related mortality between groups after the development of EAC. We had 3 secondary outcomes: (1) all-cause (overall) mortality, (2) comparison of early-stage EAC detection rates among groups, and (3) comparison of surgical resection rates among groups. In addition, we also performed numerous sensitivity analyses to better understand the effect of BE surveillance on mortality (described in further detail in the "Addressing Heterogeneity" section).	A single case-control study did not show any association between surveillance and EAC-related mortality. A meta-analysis of 4 cohort studies found that lower EAC-related and all-cause mortality were associated with regular surveillance (relative risk, 0.60; 95% CI, 0.50–0.71; hazard ratio, 0.75; 95% CI, 0.59–0.94). Meta-analysis of 12 cohort studies showed lower EAC-related and all-cause mortality among patients with surveillance-detected EAC vs symptom-detected EAC (relative risk, 0.73; 95% CI, 0.57–0.94; hazard ratio, 0.59; 95% CI, 0.45–0.76). Lead- and length-time bias adjustment substantially attenuated/eliminated the observed benefits. Surveillance was associated with detection of EAC at earlier stages.		Gastroenterology
Thota, 2017	prospective multicenter cohort study	N.a.	n=1791	>18 y/o	Inclusion criteria for this study were as follows: (1) patients with columnar-lined mucosa in the distal esophagus of any length at endoscopy and the presence of IM with no evidence of dysplasia or cancer documented on histology, and (2) an endoscopic follow-up period of ≥ 1 year from the time of initial diagnosis. The study group of interest for this analysis consisted of patients with irregular Z line with IM (ie, esophageal columnar length <1 cm).	patients with a < 1 cm segment of columnar metapl asia	patients with >1 cm segment of columnar metapla sia	Median FU of 5.9 years	Unadjusted progression rates to HGD or EAC between patients with irregular Z line and IM were compared with those with BE ≥1 cm in length using Fisher's exact test because of a zero event rate.	In a prospective, multicenter cohort study, we found that patients with irregular Z line do not develop HGD or esophageal cancer within 5 years after index endoscopy.This supports recent guidelines that patients with columnar segments <1 cm should not be diagnosed as BE and that ongoing surveillance may not be indicated in these patients. Our findings have significant implications in cost reduction and tailoring surveillance intervals depending on individualized risk in patients with BE.		Gastroenterology
Matsuhashi, 2017	multicenter prospective cohort study	N.a.	n=209	>18 y/o	Between June 2011 and August 2015, 215 consecutive BE patients with an M value of 3 cm or longer were prospectively enrolled in this study.	N.a.	N.a.	FU of 1-4 years	The main outcome in the present study was to establish the incidence of EAC in Japanese patients with BE over 3 cm in its maximal length. Patients in whom EAC was diagnosed in the first endoscopic examination underwent subsequent treatment, and their prognosis was observed. Patients without EAC in the initial EGD were instructed to undergo EGD every year, and the incidence of EAC (number of cases per patient-year) was estimated.	In Japanese cohort, the incidence rate of novel EAC in patients with BE with a length of 3 cm or longer in Prague M criteria was estimated at 1.2% per year, which was similar to the values in reports from Western countries. All the cancers detected in the follow-up period were early cancers, while most of those detected at the initial endoscopic examination were at advanced stage.		Journal of Gastroenterology and Hepatology (Australia)

Gatenby, 2016	Description of United Kingdom Barrett's Oesophagus Registry	N.a.	n=1136	>18 y/o	Patients who had been registered with the United Kingdom Barrett's Oesophagus Registry from 9 centers who did not have prevalent adenocarcinoma (diagnosed at index endoscopy or within one year of the index endoscopy) and who had a minimum of one year of follow-up were included in the study cohort.	The covariates examined were segment length, previous biopsy findings, age at surveillance, duration of surveillance, year of surveillance and gender.	N.a.	Mean follow-up was 5.70 years and total follow-up was 6474 patient-years.	The three outcome measures were (1) development of any grade of dysplasia; (2) development of high-grade dysplasia or adenocarcinoma; and (3) development of adenocarcinoma.	There was no clear evidence of any change in overall adenocarcinoma or dysplasia incidence throughout the cohort period nor for the risk of development of all grades of dysplasia to change with increasing duration of follow-up. The risk of development of high-grade dysplasia and adenocarcinoma tended to increase with increased duration of surveillance, but more strikingly: there was a relationship between older age at surveillance and higher rate of detection of high-grade dysplasia and adenocarcinoma.The results of this study demonstrate that age at surveillance is an important factor for high-grade dysplasia and adenocarcinoma development and should be incorporated (with segment length and previous biopsy findings) into risk assessment in Barrett's oesophagus surveillance.	World Journal of Gastroenterology
Kastelein, 2015	Multicentre prospective cohort study	N.a.	n=714	>18 y/o	Consecutive patients were included presenting with known or newly diagnosed BO of at least 2 cm, without a history of HGD or OAC. The diagnosis was confirmed by the presence of intestinal metaplasia.	N.a.	N.a.	Median duration of 6 years and a total of 3992 person-years of follow-up.	We used a multistate Markov model to calculate progression rates from no dysplasia (ND) to low-grade dysplasia (LGD), high-grade dysplasia (HGD) and OAC. Progression rates were incorporated in a decision-analytic model, including costs and quality of life data. We evaluated different surveillance intervals for ND and LGD, endoscopic mucosal resection (EMR), radiofrequency ablation (RFA) and oesophagectomy for HGD or early OAC and oesophagectomy for advanced OAC. The incremental cost-effectiveness ratio (ICER) was calculated in costs per quality-adjusted life-year (QALY).	Based on a Dutch healthcare perspective and assuming a willingness-to-pay threshold of €35.000 per QALY, surveillance with EMR and RFA for HGD or early OAC, and oesophagectomy for advanced OAC is cost-effective every 5 years for ND and every 3 years for LGD.	Gut, BMJ
Yang, 2015	Cost-Effectiveness Analysis	N.a.	N.a.	N.a.	The possibilities of terminals and survival payoff are retrieved from literature search in PubMed through the search strategy as, "('barrett's oesophagus'[All Fields] OR 'barrett esophagus'[MeSH Terms] OR ('barrett'[All Fields] AND 'esophagus'[All Fields]) OR 'barrett esophagus'[All Fields] OR ('barrett's'[All Fields] AND 'esophagus'[All Fields]) OR 'barrett's esophagus'[All Fields]) AND (esophageal[All Fields] AND ('adenocarcinoma'[MeSH Terms] OR 'adenocarcinoma'[All Fields]))." The studies on Western population, which reported the prevalence of BE, incidence of EAC among patients with BE or subpopulation without BE, and the EAC-related survival outcome of AS, IAS, and NS groups among patients with BE, as well as subpopu lation without BE, were eligible for possibility retrieval.	Inadequate BE surveillance	Adequate BE surveillance	N.a.	Patients with BE were classified as adequate surveillance (AS), inadequate surveillance (IAS), and no surveil lance groups. Direct cost of endoscopy per person-year was estimated from diagnosis of BE to before diagnosis of EAC in the whole-population model, whereas the payoff was 2-year disease-specific survival rate of EAC	The results show that adequate endoscopic surveillance for patients with BE is able to be more cost-effective than IAS. Adequate surveillance for patients with BE had lower cost-effectiveness ratio (CER) than that of inadequate surveillance group, as well as lower incremental cost-effectiveness ratio (6116s/% vs 118,347s/%). Prolonging the surveillance years could decrease the yearly cost in whole population and also relevant CERs, despite increased total cost. Increasing the proportion of participants in AS group could improve the survival benefit. However, regarding optimal cost-effectiveness, further studies are still required to identify a high-risk subpopulation out of BE patients for endoscopic surveillance.	Medicine (United States)
Roberts, 2010	Singlecenter prospective cohort study	N.a.	n=302	>18 y/o	The cut off for inclusion into the study was July 2001 by which point patients must have had at least one surveillance endoscopy after index endoscopy. This permitted a minimum of 5 years follow-up for all patients. Surveillance consisted of scheduled annual endoscopy with four quadrant biopsies every 2 cm.	Patients with HGD, carcinoma-in-situ or invasive OA diagnosed during BE surveillance	Patients with HGD, carcinoma-in-situ or invasive OA diagnosed at initial endoscopy during the same period acted as a comparative group (prevalent lesions).	The surveillance group underwent a total of 652 years and 2 months of surveillance during which 923 endoscopies were performed. The average duration of surveillance was 25.9 months (range 9–63).	Primary endpoint was the 5-year survival of individuals who developed HGD, carcinoma-in-situ or invasive OA. All-cause mortality was used as the measure of survival. By combining data with costs associated with both the surveillance group and the prevalent lesion group, it is possible to infer cost-effectiveness of our surveillance programme.	Annual endoscopic surveillance seems clinically effective and financially acceptable. Whether annual surveillance confers benefit over less-frequent surveillance programmes is debatable. The cost per life years saved by this Barrett's surveillance programme is similar to the established practice of surveillance after treatment of colorectal adenocarcinoma.	Eur J Gastroenterol Hepatol

Macdonald, 2000	Observational study	N.a.	n=409	>18 y/o	Patients were regarded as having Barrett's mucosa only if endoscopy showed an area of abnormality ≥30 mm in length and biopsy samples confirmed the presence of columnar metaplasia. This definition excludes patients who are now considered to have "short segment" Barrett's oesophagus, but reflects the practice of assessment at the beginning of the study.	N.a.	N.a.	The average period of surveillance for the cohort was 4.4 years.	We monitored subsequent follow up and surveillance of patients through hospital records. In addition, we obtained details of subsequent illness and deaths for all patients who had Barrett's oesophagus from the notes of other local hospitals, general practitioners, and mortality records. Patients recruited into our surveillance programme since 1994 are not included in this report.	Of the 409 patients who had Barrett's oesophagus diagnosed, 143 were suitable for surveillance and participated for an average of 4.4 years. Although five of these patients developed cancer, the programme identified only one patient, who died postoperatively. In conclusion, the value of our surveillance programme seems extremely limited, and use of a more intensive biopsy protocol would not have altered clinical outcome.	BMJ
Weston, 1997											Full text not available.American Journal of Gastroenterology

Table 12			Non-endoscopic devices (Cytosponge, EsoCheck, EsophaCap), breath analysers and blood biomarkers for risk stratification of BE								
Author (year)	Methods		Population				Intervention			Outcomes	
	Design	Randimsation / blinding	N	Age	Inclusion criteria		Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	Remarks

Author (year)	Design	Rand.	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	Remarks
Shah-2018-(Mol Cell Proteomics) - Evaluation of Serum Glycoprotein Biomarker Candidates for Detection of Esophageal Adenocarcinoma and Surveillance of Barrett's Esophagus	Case-control	No No	Biomarker discovery set was 252 serum samples from 242 patients were analysed (Normal - 43, BE - 65, BE-LGD - 39, BE-HGD - 35, and EAC - 60, at baseline). Biomarker validation set was 49 serum samples from USA were also analysed (14 NSE, 13 BE, 3 BE-ID (Barrett's mucosa which is indefinite for dysplasia), 2 BE-LGD, 7 BE-HGD, and 10 EAC patients. analyzed)	Adults	Any patient with BE and dysplasia form the PROBE-NET (progression of BE to cancer network) study from 4 Australian states. Validation set comprises of 49 patients from a USA hospita (Ochsner Heatlh System)l who had previous treatment for BE were also recruited	Blood based biomarkers	Endoscopy and biopsies	n/a	Ability of biomarker panel to discriminate BE which does not require treatment (NDBE/LGD) from BE that require treatment (HGD/IMC)	The sensitivity and specificity of the biomarker panel for discriminating BE with HGD/IMC from NDBE/LGD in the validation set (Oschener - USA)	Complement C9 (C9), gelsolin (GSN), serum paraoxonase/arylesterase 1 (PON1) and serum paraoxonase/lactonase 3 (PON3) were validated as diagnostic glycoprotein biomarkers in lectin pull-down samples for EAC across both cohorts. A panel of 10 serum glycoprotein biomarker candidates discriminated BE patients not requiring intervention (BE low grade dysplasia) from those requiring intervention (BE with high grade dysplasia (BEHGD) or EAC) with an AUROC value of 0.93. Tissue expression of C9 was found to be induced in BE, dysplastic BE and EAC. In longitudinal samples from subjects that have progressed toward EAC, levels of serum C9 were significantly (p < 0.05) increased with disease progression in EPHA (erythroagglutinin from Phaseolus vulgaris) and NPL (Narcissus pseudonarcissus lectin) pull-down samples. The results confirm alteration of complement pathway glycoproteins during BE-EAC pathogenesis.
Maddalo 2017 (J Clinical Gastro) Squamous Cellular Carcinoma Antigen Serum Determination as a Biomarker of Barrett Esophagus and Esophageal Cancer: A Phase III Study	Case-control	No No	213 (53 BE, 53 OAC, 107 control)	Not specified, but adults	Patients recruited were: 1) patients with previously diagnosed BE undergoing routine follow-up endoscopy: 71 samples, 18 of which, deriving from patients with additional known synchronous or recent malignancies, were excluded from the analysis (53 samples); 2) EAC: 53 samples; these patients were arbitrarily subgrouped in "early stage" (stages 0-1-2, ie, those with direct surgical indication) and "advanced stage" (stages 3-4, ie, those who underwent neoadjuvant treatment before surgery or did not undergo surgery at all); 3) controls: 107 samples including 42 blood donors and 65 consecutive patients with GERD but with endoscopy and histology excluding a diagnosis of BE (GERD).	Serum Squamous Cellular Carcinoma Antigen	Endoscopy	n/a	Diangosis of BO or OAC	Sensitivity and specificity for BO or OAC	The cutoff for SCCA-IgM, as calculated by the ROC curves between patients affected by either BE or EAC and GERD patients was 56.6 AU/mL with a 91.5% sensitivity, 75.4% specificity, a positive predictive value (PPV) of 85.8%, a negative predictive value (NPV)
Risques-2007-(Cancer Epidemiol Biomarker Prev) Leucocyte telomere length predicts cancer risk in Barrett's esophagus	Prospective cohort	No No	300 patients with BE	Not specified, but adults	Patients were enrolled in the Seattle Barrett's Esophagus Research Program, a dynamic cohort study that began in 1983. Three hundred participants were eligible, as defined by the diagnosis of specialized intestinal metaplasia in esophageal biopsies, with no history of esophageal malignancy, with at least one follow-up endoscopy, and with a baseline blood sample available. Baseline was defined as the first endoscopy between January 5, 1995, and December 2, 1999. We have included all esophageal adenocarcinomas that developed subsequent to the baselineevaluation so that accurate risk stratification models can be developed based on findings at a single baseline endoscopy (18). INCLUDED THOSE WHO DEVELOPED OAC WITHIN 4 months	Leucocyte telemere length measured by quantitative PCR	Endoscopy and biopsies	Mean 5.8 years	Hazard Ratio of leucocyte telomere length for risk of OAC	Hazard Ratio of leucocyte telemere length for risk of OAC	Results: Shorter telomeres were associated with increased esophageal adenocarcinoma risk (age-adjusted HR between top and bottomquartiles of telomere length, 3.45; 95% confidence interval, 1.35-8.78; P = 0.009). This association was still significant when individually or simultaneously adjusted for age, gender, nonsteroidal anti-inflammatory drug (NSAID) use, cigarette smoking, and waist-to-hip ratio (HR, 4.18; 95% confidence interval, 1.60-10.94; P = 0.004). The relationship between telomere length and cancer risk was particularly strong among NSAID nonusers, ever smokers, and patients with low waist-to-hip ratio. Conclusion: Leukocyte telomere length predicts risk of esophageal adenocarcinoma in patients with Barrett's esophagus independently of smoking, obesity, and NSAID use. These results show the ability of leukocyte telomere length to predict the risk of future cancer and suggest that it might also have predictive value in other cancers arising in a setting of chronic inflammation.
Rumiato-2017-(Transl Res) Detection of genetic alterations in cfDNA as a possible strategy to monitor the neoplastic progression of Barrett's esophagus	Case-series	No No	40 BE patients. Of this, 19 patients also longgitudinally studied and cfDNA samples collected and 18 blood samples among dysplastic patients were also collected	Adults	Forty BE patients among those referred to the reference Endoscopic Unit of the Veneto Institute of Oncology IOV-IRCCS, Padova, Italy, between 2012 and 2015 were included in this study. Inclusion criteria consisted of a BE diagnosis and the availability of clinicopathological data; exclusion criteria were the presence of an invasive EAC, neoplasms in other sites, or major comorbidity, at the enrollment. Patients with a metaplastic esophagus who had a previous diagnosis of dysplasia were included in the BM category if at least 1 year had passed since endoscopic intervention; until then they were included in the posttreatment category. All biopsy specimens were assessed by a pathologist with elective experience and confirmed by a second opinion; in the presence of discordance, the cases were jointly reconsidered. Blood samples were collected at the first endoscopy visit, and sequential samples were also gathered from some patients at follow-up endoscopy visits (total 72 samples).	Plasma cell free DNA to assess for loss of heterozygosity among dysplastic and non-dysplastic BE	Endoscopy and biopsies	Median 24 months	The fractional allelic loss index (FAL) between NDBE and dysplastic BE (FAL is calculated for each cfDNA sample by dividing the number of LOH-positive markers by the total number of informative (heterozygous) loci.	FAL index in discriminating NDBE from LGD/HGD and between LGD/HGD post treatment	Here, we report that, in the cfDNA of dysplastic BE patients, the frequency of genetic alterations is statistically higher than that of metaplastic BE patients (P 5 0.005). Interestingly, after endoscopic treatment, the alteration frequency dropped, suggesting that cfDNA can also be used to monitor curative effects. Among the used markers, those that map nearby TP53 gene were the most discriminant between metaplastic and dysplastic BE. Furthermore, longitudinal follow-up cases showed that genetic alterations can be found in cfDNA before the appearance of a detectable lesion. Altogether, our data suggest that the use of liquid biopsy could become a minimally invasive diagnostic tool to implement BE patient monitoring. (Translational Research 2017;190:16–24)

Study	Design			Sample	Age	Methods	Biomarker	Reference standard	Follow-up	Primary endpoint		Results
Pilonis-2022- (Lancet Oncology) Use of a Cytosponge biomarker panel to prioritise endoscopic Barrett's oesophagus surveillance: a crosssectional study followed by a real-world prospective pilot	Cross-sectional and prospective cohort	No	No	557 training, 334 validation and 223 prospective validation	Adults >18 years	We first conducted a retrospective, multicentre, crosssectional study in patients with known Barrett's oesophagus who had a Cytosponge test followed by endoscopy. We included all available consecutive patients older than 18 years with a confirmed diagnosis of Barrett's oesophagus (with intestinal metaplasia confirmed by TFF3 and a minimum Barrett's segment length of 1 cm [tongues or circumferential by the Prague C and M criteria] who were having endoscopic surveillance as part of the BEST2 (ISRCTN12730505) and BEST3 (ISRCTN68382401) clinical trials.15,16 All patients in these trials received Cytosponge and confirmatory endoscopy from July 7, 2011, to April 1, 2019. Risk stratification was a stated secondary aim of these trials. Patients were recruited from across hospitals that were geographically dispersed across England. Eligible participants were split into training (n=557) and validation (n=334) cohorts on the basis of date of recruitment (training cohort 2011–13, validation cohort 2013 onwards).	Cytosponge	Endoscopy and biopsies	n/a	Endoscopic biopsy diagnosis of high-grade dysplasia or cancer was the primary endpoint.	Endoscopic biopsy diagnosis of high-grade dysplasia or cancer was the primary endpoint.	Findings The prevalence of high-grade dysplasia or cancer determined by the current gold standard of endoscopic biopsy was 17% (92 of 557 patients) in the training cohort and 10% (35 of 344) in the validation cohort. From the new biomarker analysis, three risk groups were identified: high risk, defined as atypia or p53 overexpression or both on Cytosponge; moderate risk, defined by the presence of a clinical risk factor (age, sex, and segment length); and low risk, defined as Cytosponge-negative and no clinical risk factors. The risk of high-grade dysplasia or intramucosal cancer in the high-risk group was 52% (68 of 132 patients) in the training cohort and 41% (31 of 75) in the validation cohort, compared with 2% (five of 210) and 1% (two of 185) in the low-risk group, respectively. In the real-world setting, Cytosponge results prospectively identified 39 (17%) of 223 patients as high risk (atypia or p53 overexpression, or both) requiring endoscopy, among whom the positive predictive value was 31% (12 of 39 patients) for high-grade dysplasia or intramucosal cancer and 44% (17 of 39) for any grade of dysplasia. Interpretation Cytosponge atypia, p53 overexpression, and clinical risk factors (age, sex, and segment length) could be used to prioritise patients for endoscopy. Further investigation could validate their use in clinical practice and lead to a substantial reduction in endoscopy procedures compared with current surveillance pathways.
Peleg-2011 (Endoscopy)- Neutrophil-to-lymphocyte ratio and risk of neoplastic progression in patients with Barrett's esophagus	Retrospective cohort	No	No	324 (240 NDBE, 3 IND, 69 LGD, 12 HGD/EAC)(13 progressed to HGD or higher)	Adults >18 years	All patients older than 18 years old with endoscopic and histologic diagnosis of BE from January 2013 to June 2018 were part of the study. Data was prospectively collected and retrospectively analyzed. Patients who had prior history of EAC or had any type of malignancy during presentation excluded from the statistical analysis. Patients with BE segment shorter than one cm in length (also known as ultra-short BE or irregular Z line), patients with concomitant HIV or other active infections and patients with any hematologic malignancy were also excluded.	Blood Biomarker (Neutrophil-to-lymphocyte ratio)	Endoscopy and biopsies	Mean 3.7 years	The primary end point of the study was progression to neoplasia which defined as development of HGD/EAC at least one year after index endoscopy in patients with NDBE or LGD on index endoscopy.	Diagnosis of HGD/OAC at least 1 year after index endoscopy with NDBE or LGD on index	Results: 324 patients were part of the final cohort, 241 (74.4%) were males, with mean age of 62.3 years. Thirteen patients demonstrated histologic progression to neoplasia over a mean follow up of 3.7 years (annual progression risk of 1.0% per year). The AUC of NLR for progression to high grade dysplasia (HGD) or EAC was 0.88 (95% CI 0.83-0.96), and baseline NLR was associated with 3-fold increase of progression to HGD and EAC during follow up (HR 3.2, 95% CI 1.5-5.8, p<0.001). Notably, in a subgroup analysis of patients with non-dysplastic BE (NDBE) at presentation, NLR was also a risk factor of histologic progression (HR 2.4, 95% CI 1.7-3.4, p<0.001). Conclusion: NLR predicts histologic progression in patients with BE. Patients with NDBE and NLR above 2.4 can be considered for specific surveillance programs with shorter intervals between sessions
Campos-2020-(J Gastrointestinal Surgery) Neutrophil-Lymphocyte Ratio as a Marker of Progressionfrom Non-Dysplastic Barrett's Esophagus to EsophagealAdenocarcinoma: a Cross-Sectional Retrospective Study	Retrospective cross-sectional	No	No	113 (72 NDBE, 11 Dysplastic, 30 EAC)		Medical records of all patients undergoing upper endoscopy between January 2013 and September 2017 that had columnar epithelium visualized in the distal esophagus were retrospectively analyzed. All endoscopic examinations that fulfilled criteria were included, even though belonging to the same patient in a different period of the surveillance. Exclusion criteria were absence of confirmed intestinal metaplasia on histology (goblet cells on Alcian- Blue staining), immunosuppression by drugs or chronic diseases, active or recent (< 6 months) infectious disease, history of cancer or any hematological or autoimmune disease, and previous surgery over the gastrointestinal tract, except fundoplication for GERD. To be considered eligible, cases also must have had a complete blood count (CBC) collected between the period of 6 months before and 6 months after endoscopy and necessarily outside of a context of clinical emergency (for example, on the emergency room, for any reason) or invasive procedures (such as surgery or esophageal dilation). Those criteria were assessed through a thorough exam of electronic medical records. In the presence of more than one eligible CBC, the one closest to the day of the endoscopy was selected. If an included patient had any endoscopy performed before 2013, those exams were also analyzed for inclusion according to the aforementioned criteria. Patients with EAC were selected from hospital discharge diagnostic records between January 2005 and December 2017 if they had the following codes, according to the international classification of diseases (ICD-10; C15: malignant neoplasm of the esophagus): C15.2 (abdominal esophagus), C15.5 (lower third), C15.8 (overlapping sites), and C15.9 (unspecified). Squamous cell carcinoma and gastroesophageal junction (GEJ) tumor types 2 or 3 of Siewert classification14 were excluded. Same exclusion criteria used for BE cases were applied, except the need of confirmed intestinal metaplasia.	Blood Biomarker (Neutrophil to lymphocyte ratio)	Endoscopy and biopsies	n/a	The primary endpoint was the correlation of an increasing NLR with advancing stages of BE progression. Subsequently, the same analysis was performed after subdivision of EAC patients.	The primary endpoint was the correlation of an increasing NLR with advancing stages of BE progression. Subsequently, the same analysis was performed after subdivision of EAC patients.	Results NLR progressively increased across groups (NDBE, 1.92 ± 0.7; DBE, 2.92 ± 1.1; EAC 4.54 ± 2.9), with a significant correlation between its increasing value and the presence of dysplasia or neoplasia (r = 0.53, p < 0.001). NLR > 2.27 was able to diagnose EAC with 80% sensitivity and 71% specificity (area under the curve = 0.8). Conclusion NLR correlates with advancing stages of BE progression, a finding that reinforces the role of immune imbalance in EAC carcinogenesis and suggests a possible use of this marker for risk stratification on surveillance strategies.

PICO search string:	"Barrett Esophagus"[Mesh] OR "Barrett's oesophagus"[tw] OR "low grade dysplasia"[tw] OR "high grade dysplasia"[tw] OR "Barrett's Esophagus related neoplasia"[tw] OR "Barrett's Oesophagus related neoplasia"[tw] OR "esophageal adenocarcinoma"[tw] OR "Adenocarcinoma Of Esophagus" [Supplementary Concept] OR "intramucosal adenocarcinoma"[tw] AND "DNA Methylation"[Mesh] OR "Genetic Markers"[Mesh] OR "Biomarkers"[Mesh] OR "Electronic Nose"[Mesh] OR "Breath Tests"[Mesh] OR Cytosponge[tw] OR EsophaCap[tw] OR EsoCheck[tw] OR "non-endoscopic"[tw] OR nonendoscopic[tw] OR "MicroRNAs"[Mesh] OR "Volatile Organic Compounds"[Mesh] OR "sponge-on-string"[tw] OR "Methylation"[Mesh] OR "blood biomarkers"[tw] OR "E-nose"[tw] AND "risk stratification"[tw] OR surveillance[tw] OR "Disease Progression"[Mesh]

2 papers manually included: Peleg et al 2021 Endoscopy, and Campos et al 2020 (J Gastrointestinal Surgery)

#1 AND #2 AND #3 Filters: **Humans, English**   463 papers; after screening 9 were included and 2 additional papers were manually searched. After detailed review, 7 papers included in data extraction sheet

| | Table 13 | | | **validity of Seattle protocol** | | | | | | | |

Author (year)	Methods			Population			Intervention			Outcomes		Remarks
	Design	Randimsation / blinding		N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	Remarks
Vithayathil (2022)	Crossover RCT	Seattle vs AFI directed pCLE with random biopsies	To referring histology	134	67	BE at least C2 or M3 with no visible lesions	HRWLE with Seattle protocol biopsies (standard arm)	endoscopy with AFI-directed pCLE and targeted biopsies for molecular biomarkers (experimental arm).	no, cross-sectional	Sensitivity for dysplasia all grades 80.0%; 95% CI, 63.1–91.6 ; sensitivity for HGD/cancer 80.0%; 95% CI, 63.1–91.6; Sensitivity for dysplasia all grades (overall histology) 51.9%; 95% CI, 37.8–65.7 ; sensitivity for HGD/cancer (overall histology) 43.3%; 95% CI, 25.5–62.6	Dysplasia detection (all grades), HGD/cancer. Separate analysis for overall histology, which included histology from biopsies within 12 months from the first trial endoscopy	there wasn't any "no surveillance arm"
Sharma (2013)	Crossover RCT	Seattle vs NBI targeted	To other arm results	123	61	patients under Barrett's surveillance	HRWLE with Seattle protocol biopsies (standard arm)	NBI with targeted biopsies	no, cross-sectional	WLE Sensitivity for dysplasia all grades 58% ; sensitivity for HGD/cancer 64; NBI Sensitivity for dysplasia all grades 70.9%; sensitivity for HGD/cancer 78.5%	Dysplasia detection (all grades), HGD/cancer.	
Wolfsen (2008)	Prospective tandem	Seattle vs NBI targeted	To other arm results	65		patients referred with dysplasia in Barrett's	HRWLE with Seattle protocol biopsies (standard arm)	NBI with targeted biopsies	no, cross-sectional	SP sensitivity for dysplasia 43%; NBI sensitivity 57%	Dysplasia detection (all grades),	
Nachiappan (2019)	Retrospective chort	N/A		222	67	BE with dysplasia diagnosis (prevalent or incident)	random, quadrantic biopsies (Seattle)	Targeted only	no, cross-sectional	Proportion of all grade diagnosed by Seattle vs targeted (73% vs 27%); Proportion of HGD/caner diagnosed by Seattle vs targeted (61% vs 39%)	Dysplasia detection (proportion of all grades of dysplasia diagnosis by Seattle or targeted); HGD/cancer detection (proportion of HGD/cancer diagnosed by Seattle or targeted);	there wasn't any "no surveillance arm"
Lee, 2018	Prospective comparative cohort	N/A		143	55-65 (4 different groups)	BE at least 1 cm who received NBI targeted biopsies	Repeat OGD with Seattle protocol biopsies	Index OGD with targeted biopsies by NBI only	no, cross-sectional	1 case of dysplasia all grade diagnoses by targeted and 5 diagnosed by Seattle.	Dysplasia detection (number of dysplasia diagnoses by Seattle or targeted);	segments of metaplastic columnar epithelium < 1 cm excluded
Alshelleh, 2018	Retrospective chort	N/A		141	65	BE with no visible lesions	OGD with Seattle protocol	OGD with random biopsies only	no, cross-sectional	overall dysplasia yield for each group was 5.7% in the RB group , 19.6% in the SP group; overall HGD/IMC yield for each group was 1.1% in the RB group , 3.8% in the SP group	Dysplasia detection (number of dysplasia diagnoses by Seattle protocol (SP) or random non Seattle (RB);	Risk of bias: the SP had significantly longer BE segment. Other arms not included here had VLE diagnosis.
Longcroft-Wheaton (2020)	Crossover RCT	AA chromo with targeted bx vs Seattle	To other arm results	174	66	patients under Barrett's surveillance with no history of neoplasia	Seattle protocol gastroscopy (non-targeted mapping biopsies)	acetic acid assisted gastroscopy (targeted biopsies)		overall dysplasia yield for each group was 3.4% in the SP group , 2.2% in the AA group; overall HGD/IMC yield for each group was 1.1% in the SP group , 1.1% in the AA group	Dysplasia detection (number of dysplasia diagnoses by Seattle protocol (SP) oracetic acid (AA));	

Abela (2008)	Retrospective chort	N/A		362		patients under Barrett's surveillance with no history of neoplasia	OGD with Seattle protocol (group A)	OGD with random non systematic biopsies only (group B)	Prevalence of low-grade dysplasia (per patient): 18.9% versus 1.6% (P << 0.001). Prevalence of high-grade dysplasia: 2.8% versus 0% (P = 0.03). Incidence of low-grade dysplasia: 2.2% versus 6.6% (NS). Incidence of high-grade dysplasia: 2.8% versus 0% (P = 0.03).	Dysplasia detection (number of dysplasia diagnoses by Seattle or targeted);
Roberts 2010	Retrospective cohort	N/A		376		Patient with diagnosis of Barrett's undergoing (302) or not surviellance (74) 1994 and 2001	OGD with Seattle protocol (group A)	Prevalent OAC in the same period	5y mortality was 50% in group A and 97% in group B	5 year suvival
Theron 2016	Retrospective cohort	NA		431	55 group A, 58 group B	Patient with diagnosis of Barrett's undergoing (247) or failing to attend (184) surviellance between 1982 and 2007	OGD with Seattle protocol every 2 years (group A n=274)	diagnosis of BE byt failed to attend surveillance (n=184)	Group A: 5y Dis spec mortality 52% within cancer group 4.4% overall; Group B: 5y DSS 100% within cancer group 4.9% overall. All cause of mortality: group a 35%, group B 25%	DSS; overall mortality	patient with HGD/OAC diagniss within 1 year of index endsocopy were excluded. Data not adjusted
PICO search string:	Barrett's esophagus OR Barrett esophagus OR Barrett's metaplasia AND White-light endoscopy OR mapping biopsies OR seattle protocol OR Endoscopic monitoring OR endoscopic surveillance High-grade dysplasia OR Dysplasia OR cancer OR stage of cancer OR overall survival OR disease specific survival OR mortality OR death (English, Humans) - 2323 hits - 10 included after abstract and/or full paper assessment									

Table 14	**What is the additional value of p53 in the histopathological assessment of BE biopsies and resection specimen?**								
Author (year)	Methods		Population			Intervention			Outcomes		Remarks
	Design	Randomisation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	
Skacel 2002	Interobserver study	N/A	16 patients, 3 raters	N/A	N/A	N/A	N/A	N/A	Interobserver consensus, progression risk	A correlation with clinical progression was seen for p53 positivity (p = 0.017; log-rank test), and for either p53 positivity or complete agreement among three GI pathologists on LGD diagnosis (p = 0.014; log-rank test). The p53 staining demonstrated 88% sensitivity and 75% specificity for progression of LGD to HGD/CA. Adding complete interobserver agreement on LGD among three experienced GI pathologists to p53 positivity resulted in improved sensitivity with no change in specificity (100% and 75%, respectively).	These studies assess the added benefit of P53 IHC on diagnostic consensus
Kaye 2009	Interobserver study	N/A	186 patients, 5 raters	N/A	N/A	N/A	N/A	N/A	Interobserver variation, correlation with outcome, diagnostic accuracy	Addition of P53 increased concordance expressed as kappa from 0.45 to 0.52	
Kaye 2016	Interobserver study	N/A	72 cases, 10 raters	N/A	N/A	N/A	N/A	N/A	Interobserver consensus	Average reproducibility (kappa) for P53 assessment of 0.6. Inclusion of P53 increased interobserver concordance from 0.47 to 0.55.	
Wel 2018	Interobserver study	N/A	60 cases, 10 raters	N/A	N/A	N/A	N/A	N/A	Outcomes were: (i) proportion of 'indefinite for dysplasia' (IND) diagnoses; (ii) interobserver agreement; and (iii) diagnostic accuracy as compared with a consensus 'gold standard' diagnosis defined at an earlier stage by five core expert BO pathologists after their assessment of this case set.	Addition of p53 IHC decreased the mean proportion of IND diagnoses from 10 of 60 to eight of 60 (P = 0.071). Mean interobserver agreement increased significantly from 0.45 to 0.57 (P = 0.0021). The mean diagnostic accuracy increased significantly from 72% to 82% (P = 0.0072) after p53 IHC addition.	
Januszewicz 2022	Interobserver study	N/A	216 cases, 4 raters	N/A	N/A	N/A	N/A	N/A	diagnostic concordance with and without P53, proportion of cases reclassified as indefinite for dysplasia	Use of p53-IHC led to a >40% reduction in indefinite for dysplasia diagnoses (P < 0.001) and increased interobserver concordance for all BE grades (Kappa= 0.46 (NDBE), 0.26 (BE-IND), 0.49 (LGD), 0.35 (HGD/IMC).	
Wel 2020	Interobserver study	N/A	55 cases, 51 raters	N/A	N/A	N/A	N/A	N/A	diagnostic concordance with and without P53, major diagnostic error, demographic features in MVA	Addition of P53 reduced major diagnostic error (8.8% to 8.3%) and whilst Working in a non-teaching hospital was associated with increased odds of major diagnostic error (OR 1.76, 95% CI 1.15 to 2.69), this was neutralised when pathologists viewed p53 labelled slides.	
Redston 2022	Retrospective cohort analysis and prospective validation	N/A	1449	not given	Any	N/A	N/A	not given	Risk of progression	Abnormal P53 predicts risk of progression amongst all histologic classes (p<0.001) although too few NDBE cases progressed to evaluate value in this context.	These studies assess the added benefit of P53 IHC on disease progression risk
Kastelein 2013	case-control on prospectively collected cohort	Participants blinded to outcome	635	60 (median)	All patients with newly diagnosed BE with IM in 9 hospitals between Nov 2003 and Dec 2004	N/A	N/A	6.6 yrs		Aberrant P53 increased with increasing grades of dysplasia (38% in LGD, 100% in EAC), 11% of non-dysplastic BE bxs also showed aberrant P53. Aberrant P53 in normal BE was associated with an increased risk of neoplastic progression (RR 4.3); aberrant P53 and synchronous BE strongly predict risk of progression (RRR 12.2).	

Tokuyama 2020	Retrospective cohort analysis	N/A	970	60	Any BE with IM	P53 IHC	No P53 IHC	6	Prediction of progression risk	In the subgroup of patients diagnosed with IND (n = 109), abnormal p53 expression was associated with a fourfold increase (1.2– 13.3, P = 0.023) in risk of HGD/ OAC relative to untested patients diagnosed with IND, independent of other risk factors.	
Hadjinicolaou 2020	Prospective follow up study	N/A	127	65.6	Any BE with IM at least C2M4	N/A	N/A	4.6	Prediction of progression risk	Aberrant P53 predicts progression within 12mo RR 6.0	
PICO search string:		Barrett's esophagus AND P53 AND Consensus OR Concordance, searched reference lists for additional references									

Table 15						Role of WATS3D in BE surveillance					

Author (year)	Methods		Population			Intervention			Outcomes		Remarks
	Design	Randimisation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	
Fatima 2022	retrospective crosse sectional study	NA	78	68	post endoscopic eradiation therapy for LGD, HGD, IMCa of EAC; all complete eradicatoin of IM on 2 endoscopies with mapping with 3-6 months interval: Patients with visible abnormalities were excluded	WATS 3D after Seattle protocol in previously treated area	Seattle biopsy protocol in previously treated area every 1 cm	NA	detection of IM and dysplasia recurrence after EET CR-IM	NDBE was detected in 9 patients with SPB but missed on WATS-3D and in 21 patients with WATS-3D whose SPB results were negative. Similarly, LGD was detected in 1 patient with SPB that was missed on WATS-3D, whereas WATS-3D detected LGD in 6 patients missed with SPB; the absolute increase in yield was 26.9% (95%CI 16.6737.18%), NNT 3.71 (95%CI 2.69-6.00) when WATS3D was combined with biopsies	retrospective
Qumseya 2022	Systematic review and meta-analysis	NA	2816 in 7 studies			WATS 3D	Seattle biopsy protocol	NA	detection of dysplasia, stratefied for studies reporting only HGD/CA and those also reporting LGD seperately	A total of seven studies were identified which reported rates of dysplasia detection in WATS compared to FB. The seven studies totaled 2,816 patients.FB alone identified 158 patients with dysplasia, whereas adding WATS resulted in a total of 272 cases of dysplasia (114 additional cases of dysplasia due to WATS). In the seven studies, on random- effect modeling, the pooled RR was 1.7 (95% confidence interval [CI] 1.43–2.03, P < 0.001) . This means that adding WATS to FB resulted in a relative increase in dysplasia detection of 70% (43%-103 %). There was no evidence of heterogeneity with I2=0 and Q=2.45. Six studies reported on the additional yield of HGD/AC (separate from LGD). There were 3,821 patients, of whom, 68 had HGD/AC on FB. WATS increased that number to 126 patients. Therefore, the pooled RR was 1.88 (95 %CI 1.28– 2.77), P = 0.001, I2=33%, Q=7.49 Thus, the additional yield of HGD/AC was 88% [28%–177 %]. In four studies that reported the additional yield of LGD, there were 2,155 patients, of whom, 74 had LGD on FB and 113 had LGD on WATS with FB. The pooled RR was 1.5 (95% CI 1.14–1.99), P = 0.004, I2=0, Q= 1.78,. Thus, the additional yield of LGD was 50% (14%-99%).	The clinical advantage for the studies reporting additional ylied for HGD/CA needs to be put in perspective. In the study by Bisschops et al most of patients with WATS3d+/FB- for HHGD/CA had at least LGD on FB, which would be an indication for referral. The meta-analysis clearly showed that the effect is associated with the prevalence of dysplasia in the study. We need to balance this against costs. THE METANALYSIS CONTAINS A REFERENCE THAT CANNOT BE FOUND IN MEDLINE OR ON THE WEBSITE OF MODERN PATHOLOGY
Demeester 2021	RCT multicenter	computer-generated set of stickers that were sequentially numbered and randomly designated for wide-area transepithelial sampling (WATS) or forceps biopsy sampling (FB) and sent to each participating center.	1002 ; 818 without history of known IM with 204 having a CLE endoscopic ally ; 184 with known IM or prior to ablation	57	pt for routine upper GI endoscopy or knwon dysplasia or surveillance	WATS 3D	Seattle biopsy protocol	NA	detection of IM and dysplasia secondary compliance with protocol	The overall frequency of finding IM was 21% and was similar with FB (19.6%) and WATS (22.7%, P=.2). There was no difference in detection of dysplasia between FB and WATS. In patients with no history of IM, WATS found significantly more IM compared with FB when a columnar-lined esophagus (CLE) was present (32.4% with WATS vs 15.2% with FB, P<.004). In 184 patients with known BE, FB and WATS found IM with similar frequency (38.5% FB vs 41.9% WATS, P= .6) with no difference in short- or long-segment BE. The likelihood of noncompliance was significantly higher in the FB group compared with the WATS group (27.5% vs 7%, respectively; P < .01).	includes also IM of the GEJ , however only 6 patients had also IM in the stomach// WATS after targeted biopsies was allowed ! // Study bias : in 350 patients with BE both modalities were used

Study	Study type			N	Age	Setting/Population	Intervention	Comparator		Outcome	Results	Comments
Codipilly 2022	Systematic review and meta-analysis	NA		3206 in 7 studies		prospective studies in humans	WATS 3D	Seattle protocol biopsy	NA	Primary incremental yield of WATS3D over FB for the diagnosis of dysplasia, defined as new cases of DBE identified by WATS3D alone but negative for dysplasia by FB. A composite outcome incorporating indefinite for dysplasia (IND), LGD, HGD, and EAC was used as the endpoint for this outcome. Crypt dysplasia was excluded as an outcome in our analysis. Secondary outcomes were the incremental yield of WATS3D in the diagnosis of HGD/EAC, miss rate of WATS3D (FB þ WATS3D – results), reconfirmation of WATS3D dysplasia results on subsequent FB samples when the initial result was positive on WATS3D but negative on FB, additional time taken for acquiring WATS3D samples, adverse events from WATS3D, and progression to HGD/EAC and mortality (both EAC-specific and all-cause) for patients diagnosed with WATS3D-only dysplasia compared with FB alone.	Diagnosis of dysplasia for FB = 15,9% (95% CI 5,4-30,5%)-, WATS 3D incremental yield of 7,2% (95% CI 3,9%-11,5%) Meta-analysis of 6 studies demonstrated that FB diagnosed HGD/EAC in 2.3% of patients, whereas the io-incremental yield with WATS3D was 2.1% (95% confidence interval, .4%-5.3%; I2Z92.7%). Notably,WATS3D was negative in 62.5% of cases where FB identified dysplasia. Two studies reported reconfirmation ofWATS3D dysplasia with FB histology in only 20 patients.	WATS3D increases dysplasia detection; however, the clinical significance of this increased dysplasia detection remains uncertain. Data from endoscopic follow-up to ascertain FB histology in patients with dysplasia based solely on WATS3D are needed to determine the optimal clinical application and significance of WATS3Donly dyslasia
Agha 2021	retrospective observational study	NA		108	63.5	screening or surveillance of BE with both methods	Wats 3D + FB	FB alone	NA	Detection of BE	FB and WATS3D detected 62 (57.4%) and 83 (76%) cases of BE, respectively. The absolute difference of 21 cases (18.6%) of BE was attributed to the addition of WATS3D. The number needed to test with WATS3D was 5.	retrospective; no data on endoscopic apperance of BE in additional detections
Kumar 2020	meta-analysis and systematic review	NA		20932 endoscopies in 11 studies	NA	screening endoscopies with patients with known BE				incremental yield of detection for IM and dysplasia	total of 6643 lesions (intestinal metaplasia) were identified when WATS was used with FB versus 3310 with FB alone; an additional 3333 lesions identified. The absolute increase in detection was 16% (measured as RD −0.16, 95% CI 0.10 to 0.22, p<0.00001) and a 1.62 times relative increase in detection rates of BE (measured as RR −1.62, 95% CI 1.28 to 2.05, p<0.0001) with a number needed to test (NNT) of 6.1 when WATS was used in adjunct with FB. Out of the 19 950 screening and surveillance endoscopies done in dysplasia naïve patients across nine studies, 533 lesions (dysplasia) were identified with WATS combined with FB vs 213 lesions with FB alone. With 320 more lesions identified, it translates to a marginal 2% but significant absolute increase (measured as RD 0.02, 95% CI 0.01 to 0.03, p=0.001) in additional diagnostic yield from WATS. There is a 2.05 times relative increase in the detection rate of ED (measured as RR 2.05, 95% CI 1.42 to 2.98, p=0.0001) yielding an NNT of 50 patients.	7 studies of high quality 4 of fair quality; 60 % attribution by 1 study (smith et al 2019)
Singer 2020	cost effective model	NA		NA	60	60 year old male undergoing surveillance				cost effectiveness with 100,000 and 150,000 $ per qaly as threshold	Between 320 and 337 people would need to be screened with WATS3D in addition to FB to avert one additional cancer, and 328–367 people to avert one cancer-related death. Screening with WATS3D costs an additional $1219 and produced an additional 0.017 QALYs, for an ICER of $71,395/QALY. All one-way sensitivity analyses resulted in ICERs under $84,000/QALY.	Probably not applicable to Europpean standards
Raphael 2019	observational cohort study	NA		138	53.2	Patients with and without dysplasia and potentially pre and postablation therapy	WATS3D as adjunct to HDWLE, NBI and VLE targeted biopsies and Seattle protocol	HDWLE, NBI and VLE targeted biopsies and Seattle protocol		detection of dysplasia	25% dysplasia after HDWLE, NBI and VLE targeted biopsies and Seattle protocol. WATS3D identified 12 new dysplasias: +34,3% (crypyt dysplasia and 3 LGD)	Not clear if visible lesions were sampled, not an exclusion criterion. High prevalence of dysplasia 50/138. What is the significance of cryt dysplasia on WATS 3D ??
Smith 2019	Multicenter prospective cohort	Not randomized. Alternating between WATS or FB first	no control of adherence to order	12899		Patients referred for screening or surveillance of BE in 21 sites, 58 endoscopists	Wats 3D as adjunct to FB	Seattle FB alone		detection of dysplasia = primary endpoint ; detection of BEa = sec endpoint	DYSPLASIA : FB : 88/12899 dysplasia ; WATS additionally detected 213 cases: 1,65% increase, relatively 242%. NNT 61// BE DETECTION : FB 1684, WATS3D additionjal 2570 cases +19,9% detection absolut 13,1-->33%); relative 153%	Community based. No control for effective performance of biopsies., Most extra detection of dysplasia was LGD on WATS3D
Gross 2018	multicenter prospective cohort	not randomized. WATS followed by FB	no blinding	4203	59	Patients referred for screening or surveillance of BE in 25 sites (USA)	Wats 3D as adjunct to FB	Seattle FB alone		detection of dysplasia and BE	BE : 594 by FB alone, 493 additional cases with WATS, 83% oincrease. LGD : 26 detected by FB, 23 additional by WATS , increasing 88,5%.	Again no endoscopic data available on what a BE was. The example included in the paper shows in fact brushing of a normal Z-line and GEJ.

Vennalaga nti 2018	RCT multicenter (16 centers USA)	Randomiz ed FB first or WATS 3D first	160	63.4	BE surveillance patients or referred for endoscopic treatment for BE (1-10 cm length)	Wats 3D as adjunct to FB	Seattle FB alone		detection of HGD/EAC	23 additional detections by WATS 3D 23/160, absolute increase 14,4%, relative 428,6%.	not clear if visible lesions were sampled, not an exclusion criterion. High prevalence of dysplasia, including visible lesions. Of the 23 additional cases 12 were LGD and 91,3% had a previous diagnosis of dysplasia on FB. 1 case HGD/EAC by FB was LGD on WATS3D.
Bisschops 2020/22	RCT multicenter (Europe)	Randomiz ed FB first or WATS 3D first	172	68	BE patients with known LGD/HGD or EAC after removal of all visible lesions	WATS 3D	Seattle FB alone		The primary endpoint of this study was the concordance/disconcordance between the detection of Barrett's associated HGD and/or EAC using WATS brushing and random FB. Secondary endpoints were (i) the adjunctive value of WATS, defined as the absolute and relative increased detection when WATS is added to FB; (ii) the relative sensitivity, defined as the total number of HGD/EAC diagnosed with WATS divided by any HGD/EAC diagnosis using either of the techniques; (iii) the effect of the order of WATS and FB on the outcome; (iv) the time needed for both procedures; and (v) related complications.	WATS detected 39/172 (23%) cases with HGD/EAC, and FB detected 33/172 (19%) HGD/EAC cases. A total of 21 patients (12%) were diagnosed with HGD/EAC on both modalities; WATS detected an additional 18 patients (10%) with HGD/EAC that were not detected with FB (i.e. WATS+/FB-). In contrast, FB detected 12 (7%) additional patients with HGD/EAC that were not diagnosed with WATS (i.e. WATS-/FB+). Also no significant difference in the PP analysis excluding patients with insufficient samples. WATS as adjunct : absolute increase was 10,4 % 18/172 cases . Relative increase 60% NNT 10.	Works as adjunct, but not as standalone.
Shaheen 2022	retrospective cohort study of the clinical database of commercial utilization of the WATS3D technology (CDx Diagnostics, Suffern, NY, USA) in the United States from 2013 to 2019	not randomize d	4545	62-68	4374 NDBE, 128 crypt dysplasia and 43 LGD baseline WATS3D diagnosis.				primary outcome was the crude progression rate, defined as the proportion of patients per patient-year who demonstrated progression on forceps biopsy sampling to either HGD or EAC, stratified by baseline WATS3D histologic grade.	Progression to HGD/EAC NDBE : 0,08 % per patient year / Crypt dysplasia : 1,42 % per patient year (p<0,01) // LGD : 5,79 % per patient year (p<0,05 compared to crypt dysplasia)	Limitations : Only 55% of the WATS diagnosis had forceps diagnosis available at baseline. No expert review pathology because of retrospective large database. Small number of progressors. No data on BE length or other risk factors

	Table 16		What is the role of biomarkers either or not in conjunction to biopsy/brushing?

Author (year)	Methods		Population			Intervention			Outcomes		Remarks

	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	Remarks
p53											
Snyder (2019)	meta-analysis	NA	NA	8 case-control studies comprising 1435 patients and 7 cohort studies comprising 582 patients	Mean age was 62.5 years (range 60.0–67.3),	Included studies that (1) followed subjects with NDBE, IND, or LGD for progression to HGD or EAC, (2) assessed p53 immunohistochemistry in BE biopsies taken prior to progression, (3) specified follow-up duration, and (4) reported the number of subjects with neoplastic progression.	NA	NA	followed for median 4.4 years (range 1.9–6.7 years)	Progression to HGD/EAC. Separate meta-analyses of case-control studies and cohort studies.	In the case-control study meta-analysis of the risk of neoplasia with aberrant p53 expression, the fixed- and random-effect estimates of average effect size with aberrant p53 expression were OR 3.84, $p < .001$ (95% CI 2.79–5.27) and OR 5.95, $p < .001$ (95% CI 2.68–13.22), respectively. In the cohort study meta-analysis, the fixed- and random-effect estimates of average effect size were RR = 17.31, $p < .001$ (95% CI 9.35–32.08) and RR = 14.25, $p < .001$ (95% CI 6.76–30.02), respectively. Separate meta-analyses of case-control and cohort studies of BE patients who had baseline biopsies with p53 immunostaining revealed consistent, strong, and significant associations between aberrant p53 immunostaining and progression to HGD/EAC. These findings support the use of p53 immunostaining as an adjunct to routine clinical diagnosis for dysplasia in BE patients.	p53. Meta-analysis that included studies until August 2017. For that reason, with regard to p53, I have summarized below here only original studies published after 08/2017 (with a couple of exceptions of studies that have been quoted often). Below there are other 2 meta-analyses on p53 as well
Altaf (2017)	meta-analysis	NA	NA	NA (different studies included for different biomarkers)	NA (different studies included for different biomarkers)	Included studie: BE with or without HGD/EAC, comparing biomarkers (p53, Ki-67, *p16*, DNA content abnormalities) in predicting neoplastic progression of BE to HGD/EAC. Only studies published in the English.	NA	NA	NA (different studies included for different biomarkers)	Progression to of BE to HGD/EAC. Comparing biomarkers (p53, Ki-67, p16, DNA content abnormalities) in predicting neoplastic progression of BE to HGD/EAC.	102 clinical studies were included in the study. Mutation of p53 had the following findings: diagnostic odds ratio (DOR) 10.91, sensitivity 47%, specificity 92%, positive likelihood ratio (PLR) 4.71, negative likelihood ratio (NLR) 0.65, area under the curve (AUC) 0.79. Loss of p53 had the following findings: DOR 16.16, sensitivity 31%, specificity 98%, PLR 6.66, NLR 0.41, AUC 0.923. Both mutations and loss of p53 were found to be superior to other p53 abnormalities such as loss of heterozygosity (LOH) and overexpression. Ki-67 had DOR 5.54, sensitivity 82%, specificity 48%, PLR 1.59, NLR 0.42, AUC 0.761. Aneuploidy had DOR 12.08, sensitivity 53%, specificity 87%, PLR 4.26, NLR 0.42, AUC 0.846. Tetraploidy had DOR 5.87, sensitivity 46%, specificity 85%, PLR 3.47, NLR 0.65, AUC 0.793. Loss of Y chromosome had DOR 9.23, sensitivity 68%, specificity 80%, PLR 2.67, NLR 0.49, AUC 0.807. p16 aberrations (hypermethylation, LOH, mutation and loss) failed to demonstrate any advantage over the other biomarkers studied.	p53, p16, Ki-67 and DNA content abnormalities (aneuploidy, tetraploidy, loss of Y chromosome). Meta-analysis. Included studies until March 2016.
Janmaat (2017)	meta-analysis	NA	NA	16 different IHC biomarkers were studied in 36 studies. These studies included 425 cases and 1835 controls	NA (different studies included for different biomarkers)	criteria: (1) association between IHC biomarker expression on FFPE material and risk of neoplastic progression was assessed; (2) a cohort or case-control study design; (3) patients with known or newly diagnosed BE with or without LGD at baseline; (4) patients defined as cases had to have progressed to either HGD or EAC during follow-up; (5) mean follow-up of at least one year from the time of initial BE diagnosis; (6) the possibility to extract an OR.	NA	NA	follow up time was 51 months versus 59 months for cases versus controls, respectively	Progression to of BE to HGD/EAC. Comparing biomarkers (p53, aspergillus oryzae lectin (AOL), Cyclin A, Cyclin D and alpha-methylacyl-CoA racemas) in predicting neoplastic progression of BE to HGD/EAC.	Meta-analyses were possible for p53, AOL, Cyclin A, Cyclin D, and alpha-methylacyl-CoA racemase (AMACR), which were studied 13, 2, 4, 3, and 2 times respectively. Aberrant p53 expression was significantly associated with an increased risk of neoplastic progression with an OR of 3.18 (95% CI 1.68 to 6.03). The overall OR, for aberrant p53 IHC on neoplastic progression, after stratification for histology, was 3.86 (95% CI 2.03 to 7.33).This association was confirmed for both non-dysplastic BE (6.12; 95% CI 2.99 to 12.52) and BE with LGD (OR 8.64; 95% CI 3.62 to 20.62). Another promising biomarker to predict neoplastic progression was AOL, with an OR of 3.04 (95% CI 2.05 to 4.49). Cyclin A showed a tendency towards increased risk (OR 1.90; 95% CI 0.85 to 4.22). Cyclin D did not show a significant association (OR 1.01; 95% CI 0.14 to 7.03). Alpha-methylacyl-CoA racemase had an OR 4.07 (95% CI 0.66 to 25.12). The following IHC biomarkers were investigated only once up to this meta-analysis: β-catenin, CD1a, COX2, HER2, Ki67, Lewis, Mcm2, Sialyl Lewis, SOX2, and WGA	p53, aspergillus oryzae lectin (AOL), Cyclin A, Cyclin D and alpha-methylacyl-CoA racemase. Meta-analysis. Included studies until 09/2016

Study	Type			N	Age	Inclusion	Intervention	Comparison	Follow-up	Aim	Results	Notes
Kastelein (2013)	case-control	NA	NA	Multicenter study; 635 patients with BE (≥2 cm) and no HGD/EAC followed up in time (49 cases who developed HGD/EAC and 586 controls who did not)	Cases: 60 yo (53–69); Controls: 65 yo (55–70)	Multicenter study; 635 patients with BE (≥2 cm) and no HGD/EAC followed up in time according to 2008 ACG guidelines, who had paraffin material suitable for immunohistochemistry	No intervention, just regular clinical follow up according to ACG 2008 guideline. Evaluation of >12,000 biopsies and comparison of p53 expression among cases (those who developed HGD/EAC) and controls (those who did not develop HGD/EAC at follow up).	No intervention, just regular clinical follow up according to ACG 2008 guideline. Evaluation of > 12,000 biopsies and comparison of p53 expression among cases (those who developed HGD/EAC) and controls (those who did not develop HGD/EAC at follow up).	Patients followed for a median 6.6 years, and median 4 endoscopies.	comparison of p53 expression among cases and controls	Normal BO without dysplasia was seen in 1085 (73%) biopsy series, LGD in 347 (23%), HGD in 35 (3%) and OAC in 14 (1%). Aberrant p53 expression was more common in biopsy series of cases (49%) than in biopsy series of controls (14%). Multivariable analysis included age, gender, BO length and oesophagitis. Interobserver agreement for p53 expression was good (κ=0.79; 95% CI 0.75 to 0.83). P53 overexpression was associated with an increased risk of neoplastic progression at multivariable analysis (RR 5.6, 95%CI 3.1, 10.3). Loss of p53 expression was associated with a greater increased risk of neoplastic progression (RR 14, 95%CI 5.3, 37.2). The positive predictive value for neoplastic progression increased from 15% with histological diagnosis of LGD to 33% with LGD and concurrent aberrant p53 expression. Aberrant p53 protein expression appears to be a more powerful predictor of neoplastic progression than histological diagnosis of LGD. Only 11% of patients are diagnosed with LGD and aberrant p53 expression, but this subgroup has a much higher risk of neoplastic progression.	p53. this is one of the most cited papers on the subject of p53 and progression
Davelaar (2015)	Prospective Follow-Up Study	NA	NA	116 (80 IM, 13 IND, 7 LGD, 6 HGD, and 10 EAC cases)	64 (±13) years	Only patients with proven intestinal metaplasia (IM) in biopsies were included. At study entry the brush cytology specimens were taken prior to the biopsies during endoscopy. Brushes of normal squamous epithelium were taken at least 3 cm above the BE segment for control purposes.	Evaluated if immunohistochemistry (IHC) on tissue specimen and DNA fluorescent in situ hybridization (FISH) on brushing are complementary tools to assess TP53 abnormalities and tested their prognostic value for progression to HGD/EAC in a long-term prospective follow-up of a BE cohort	Utility of IHC alone, FISH alone and their combination in predicting progression to HGD/EAC. Regular clinical follow up according to the ACG 2008 guideelines. One normal squamous, four BE and four EAC cell lines were used to validate the patient data.	71 (IQR 47–84) months. 505 patient-years of endoscopic surveillance.	Evaluated if immunohistochemistry (IHC) on tissue specimen and DNA fluorescent in situ hybridization (FISH) on brushing are complementary tools to assess TP53 abnormalities and tested their prognostic value for progression to HGD/EAC in a long-term prospective follow-up of a BE cohort.	The frequency of IHC and FISH TP53 abnormalities increased significantly with increasing histological stage ($P<0.001$). The number of patients positive for TP53 abnormalities was significantly higher in the progressors than in nonprogressors for both IHC (63.6% vs. 7.5%) and FISH (36.4% vs. 7.5%; all cases p<0.05). Also when IHC and FISH were combined, this result was observed. Multivariate analysis showed that IHC (hazard ratio: 17, 95% CI: 3.2–96, P50.001) and FISH (hazard ratio: 7.3, 95% CI: 1.3–41, P50.02) were both independent significant predictors of progression. Combining TP53 FISH with TP53 IHC positivity was associated with an OR of 25.5 (95% CI: 4.90–133, P<0.001) and had a sensitivity and specificity of 81.8% and 85%, respectively for predicting progression. Similar results were observed for progression to only HGD and only EAC. A large number of patients was detected by only one technique (from a total of 40 with TP53 abnormalities, only 9 tested positive for both FISH and IHC).Combining the techniques in assessing TP53 abnormalities leads to an increased detection rate of TP53 aberrations (detected TP53 abnormalities in 100% of patients with LGD, HGD, and EAC) and thus and improved accuracy for predicting BE progression. There was little overlap between the abnormalities found by IHC and FISH at cell lines, thus alterations missed by one method may be picked up by the other, and vice-verse. Hence these two methods may be complementary.	p53. Comparison of FISH and IHC to assess p53
di Pietro (2015)	cross-sectional study	NA	NA	157 (training cohort), 46 (validation cohort)	66.4 (training cohort), 68.7 (validation cohort)	Age ≥18 yrs, BE with a length of at least C ≥ 2 or C<2M≥4 according to the Prague classification with or without visible lesions.	Biopsies were performed either according to the standard Seattle protocol or under the guidance of autofluorescence imaging (AFI)	Comparison of the accuracy of a panel of molecular biomarkers on AFI-directed biopsies with conventional Seattle protocol biopsies for HGD and EAC. Secondary aims: (i) assessment of diagnostic accuracy for the biomarkers for any grade of dysplasia and (ii) validation of a large panel of biomarkers in an independent prospective study by an independent laboratory.	NA	Accuracy of biomarkers in identifying HGD/EAC. Panel of biomarkers assessed: Aneuploidy/tetraploidy; 9p and 17p loss of heterozygosity; RUNX3, HPP1 and p16 methylation; p53 and cyclin A immunohistochemistry. The statistical analysis consisted of three stages: (1) per-biopsy analysis (correlation between biomarkers and histological outcome in individual targeted areas); (2) per-patient analysis (correlation between overall biomarker result and overall histological outcome in individual patients); and (3) comparison between AFI+ versus AFI− areas (comparative analysis of biomarker diagnostic accuracy for dysplasia in biopsies from AFI+ vs AFI− areas).	All of the biomarkers associated with the presence of confirmed dysplasia, with the exception of 9p LOH (p16). At stricter analysis for HGD/EAC, aneuploidy, p53 immunohistochemistry and cyclin A had the strongest association with dysplasia in the per-biopsy analysis and, as a panel, had an area under the receiver operating characteristic curve of 0.97 (95% CI 0.95 to 0.99) for diagnosing HGD/EC. The diagnostic accuracy for HGD/EC of the three-biomarker panel from AFI+ areas was superior to AFI− areas (p<0.001). Average AUCs in all six databases were significantly higher for the threebiomarker panel assessed on AFI+ areas compared with AFI− areas. The biomarker panel had a sensitivity and a specificity of 95.8% (95% CI 76.9% to 99.8%) and 88.6% (95% CI 79.7% to 94.1%), respectively, for a diagnosis of HGD/EC. By comparison, the Seattle protocol had similar sensitivities, namely, 95.8% (95% CI 76.9% to 99.8%) for a diagnosis of HGD/EC (p=1.0 when compared with three-biomarker panel). Importantly, using this novel approach 2.8 biopsies per patient were taken on average compared with 12.8 for the standard biopsy protocol (4.5 fold reduction; p<0.001). At validation cohort, the panel had a sensitivity and a specificity of 100% and 85% (95% CI 98.9% to 95.0%), respectively, for a diagnosis of HGD/EAC.	Panel of biomarkers assessed: Aneuploidy/tetraploidy; 9p and 17p loss of heterozygosity; RUNX3, HPP1 and p16 methylation; p53 and cyclin A immunohistochemistry
Younes (2017)	retrospective	NA	NA	275 BE patients with no HGD/EAC	mean 62, median 63, range 21-91	275 BE patients with no HGD/EAC at baseline	regular clinical follow up; none received ablative therapy during follow up	comparisons of baseline p53 among progressors and non-progressors	mean 41 months, median 43 months, range 3-112 months	Role of p53 (measured by ICH) in the progression to HGD/EAC.	Of the 275 patients, 227 had initial biopsies completely negative for p53 and of these one (0.4%) progressed to HGD/EAC; none of 24 (0%) patients with scattered positive cells and none of 4 (0%) of patients with multifocal scattered positive cells progressed. By contrast, 5 of 16 (31.25%) patients with aggregates of positive cells and 3 of 4 (75%) of those with multifocal aggregates of positive cells progressed to HGD/EAC. Kaplan-Meier analysis with log rank statistics showed the difference in progression rate between the five groups to be highly significant (p<0.0001). There were 20 cases with p53 protein overexpression on initial biopsies of which 8 (40%) progressed to HGD/EAC. This is in sharp contrast to the 255 cases with initial biopsies negative for p53 protein accumulation of which only one (0.3%) progressed to HGD/EAC (p <0.0001).	p53
Stachler (2018)	retrospective case–control study	NA	NA	97 (24 cases-progressors and 73 controls - non progressors)	cases 67.5 vs. controls 61.6 years	Patients with biopsy-proven Barrett's esophagus with NDBE, IND or LGD. If someone developed HGD/EAC within 12 months, they were excluded.	Routine clinical follow up-surveillance.	Comparison of progressors vs. non-progressors	mean follow-up (controls 6.7 vs. cases 3.3 years, P<0.001)	Porgression to HGD/EAC. From each patient, we selected a single tissue sample obtained more than 1 year before progression (cases) or more than 2 years before the end of follow up (controls). Pathogenic mutations, gene copy numbers, and ploidy were compared between samples from progressors and non-progressors.	TP53 mutations were detected in 46% of samples from progressors and 5% of non-progressors. In this case–control sample set, TP53 mutations in BE tissues increased the adjusted risk of progression 13.8-fold (95% CI, 3.2–61.0) (P<.001). These results were confirmed in a separate validation set of 16 NDBE who progressed and 28 NDBE who did not progress. We did not observe significant differences in ploidy or copy number profile between groups. We identified 147 pathogenic mutations in 57 distinct genes—the average number of pathogenic mutations was higher in samples from progressors (2.5) than non-progressors (1.2) (P<.001). TP53 and other somatic mutations were recurrently detected in samples with limited copy number changes (aneuploidy). Beyond TP53, several known GEA tumor suppressors and oncogenes (ARID1B, APC, ERBB2, RB1, RUNX1, LARP4B, and BIRC5) had more frequent mutations in progressors. Among these, ARID1B, APC, and ERBB2 were significantly enriched in the progressors (P<0.05).	p53; also ARID1B, APC, and ERBB2

Duits (2018)	nested case-control study	NA	NA	260 (130 cases-progressos vs 130 controls non-progressors)	Progressors:59.9 ± 9.6; non-progressors 59.2 ± 9.7	Patients with biopsy-proven Barrett's esophagus with NDBE, IND or LGD. If someone developed HGD/EAC within 24 months, they were excluded.	Routine clinical follow up-surveillance.	Comparison of progressors vs. non-progressors	progressors 3.7 vs. non-progressors 4.8 years; P < 0.001	Porgression to HGD/EAC. We assessed abnormal DNA content, p53, Cyclin A, and Aspergillus oryzae lectin (AOL) in FFPE sections.	LGD, p53 and AOL were independently associated with neoplastic progression. In the multivariate analysis, expert LGD had a 36-fold increased odds of progression (OR, 35.7; 95% CI, 1.4–920.8). Abnormal p53 expression (OR, 4.1; 95% CI, 1.4–12.4) and abnormal AOL expression in three epithelial compartments had a 4-fold increased odds of progression (OR, 4.3; 95% CI, 0.7–26.3). Cyclin A did not predict progression, and DNA ploidy analysis by image cytometry was unsuccessful in the majority of cases, both were excluded from the multivariate analysis. The multivariable biomarker model had an area under the receiver operating characteristic curve of 0.73.	AOL, and p53, cyclin A, DNA ploidy. Utilized the Amsterdam-based ReBus nested case-control cohort, a multicenter prospective cohort study
Tokuyama (2020)	retrospective	NA	NA	78 patient with both histology and p53 vs 892 patients with histology alone	mean (range) 63.2 (36.5–92.6) for those with both histology and p53 and 57.5 (16.3–88.2) histology only	BE with non-dysplastic oesophagus (NDBO), indefinite dysplasia (IND) or LGD, no prior diagnosis of HGD or OAC and no prior ablation or oesophageal surgery.	Routine clinical follow up-surveillance.	Comparison of progressors vs. non-progressors	NA	role of p53 in the progression to HGD/EAC.	Almost half (46.9%) of patients with abnormal p53 expression were diagnosed with HGD or OAC within 5 years, compared to 5.9% with wild-type p53, and 7.6% of patients not tested (P <0.0001). However, this difference was heavily influenced by other risk factors, including dysplasia grade, in multivariate analyses. In the subgroup of patients diagnosed with IND (n = 109), abnormal p53 expression was associated with a fourfold increase (1.2–13.3, P = 0.023) in risk of HGD/EAC relative to untested patients diagnosed with IND, independent of other risk factors. Selective use of p53 IHC in conjunction with routine histology modestly improved risk stratification by identifying patients with IND at higher risk of a subsequent diagnosis of HGD/EAC.	p53
Hadjinicolaou (2020)	Prospective multicenter study	NA	NA	127 (42 progressors, 85 non-progressors)	median age of 65.6 years (IQR, 13.7 yrs)	Age ≥18 yrs, BE with a length of at least C ≥ 2 or C<2M≥4 according to the Prague classification with or without visible lesions. Patients had targeted biopsies dureing endoscopy with autofluorescence imaging (AFI).	The primary endpoint of this study was progression from NDBO/ID to any grade of dysplasia. The two secondary endpoints were a) progression from NDBO/ID to HGD/OAC, and b) any histologic progression i.e. NDBO/ID to LGD, NDBO/ID to HGD, and LGD to HGD.	Comparions of biomarkers between progressors and non-progressors. This was a validation cohort from the cross-sectional study described above (di Pietro 2015)	Median 4.6 yrs (IQR, 4.3 yrs). Progressors: 1.2 yrs (IQR, 2.7 yrs) until progression	Evaluation of previously defined 9-molecular biomarker panel (in the study of di Pietro 2015 above) with progression in BE patients. Histological progression was defined as transition from a NDBO or indefinite for dysplasia (ID) to any dysplasia, or if low-grade dysplasia already present, to a higher grade of dysplasia or cancer. Biomarkers: p53 and cyclin A were analysed by immunohistochemistry (IHC); aneuploidy and G2/tetraploidy, were analysed by flow cytometry; p16, RUNX3 and HPP1 hypermethylation was analysed by quantitative methylation-specific PCR (Methylight); and LOH at 9p and 17p loci was analysed by the use of microsatellite markers. Snap frozen biopsies in DMSO were used for aneuploidy, G2 tetraploidy, LOH markers and methylation assays.	Amongst progressors, there were 12 (28.6%) that progressed from NDBO/ID to LGD, 16 (38.1%) that progressed from NDBO/ID to HGD/OAC and 14 (33.3%) that progressed from LGD to HGD/OAC. Of the nine molecular biomarkers, at multivariable analysis p53 and aneuploidy were the only significant predictors of any progression. However, when we excluded patients with progression within 12 months of follow up (prevalent dysplasia), only aneuploidy retained statistical significance. The presence of positive aneuploidy at index endoscopy led to a 6.6-fold higher risk of dysplastic progression over no progression (95% CI: 1.8 24.8, p = 0.005). However the sensitivity of the test to predict progression was low (32%, 95% CI: 16, 52%). ROC analysis showed that a clinical model using patient age and BO length (AUC=0.55; CI: 0.45,0.66) was outperformed in the prediction of any histologic progression by a molecular biomarker model comprising of aneuploidy and p53 with a cut-off of one positive biomarker out of two (AUC=0.68; CI: 0.59,0.77). ROC analysis showed that a model with aneuploidy as the only predictor of dysplastic progression outperformed the clinical model (AUC=0.63; CI: 0.54 0.72). p53 appeared to correlate more with short-term progression. Patients with aberrant p53 expression at index endoscopy had an odds ratio of 6.0 (95% CI: 3.1, 11.2, p = 0.007) of missed dysplasia on endoscopic biopsies.	p53, cyclin A; aneuploidy and G2/tetraploidy; p16, RUNX3 and HPP1 hypermethylation; LOH at 9p and 17p loci; aneuploidy, G2 tetraploidy
Helminen (2022)	retrospective case-control	NA	NA	45 cases with HGD/EAC (24 were progressive from LGD and 21 progressive from metaplasia) and 92 controls (45 non-progressive LGD and 52 non-progressive metaplasia)	NA for the whole cohort - has Table 1 with different ages for each case-control group	Patients with ≥1 EGD >6 months before the diagnosis of HGD/EAC. Controls: non-progressive BE with or without LGD confirmed with follow-up EGDs performed at least 5 years after the initial diagnosis, matched by age (± 5 years) and sex to the cases.	Routine clinical follow up-surveillance.	Comparison of progressors vs. non-progressors	NA for the whole cohort - has Table 1 with different FU times for each case-control group	Porgression to HGD/EAC. Expression of p53, Ki67 and toll-like receptor 5 (TLR5) between progressors and non-progressors	p53 associated with a high risk of progression (OR (6.7, 95% 1.8–24.6)). The previously suggested markers Ki67 and TLR5 were not associated with disease progression.	p53, Ki67 and toll-like receptor 5 (TLR5). Based on Northern and Central Finland patients

Study	Study design			Cohort	Age	Population	Follow-up	Comparison	FU time	Aim	Results	Conclusion
Redston (2022)	retrospective and prospective	NA	NA	Retrospective cohort: 561 patients (249 non-progressors and 312 progressors). Prospective validation cohort: 1487 (646 non-dysplastic-NDBE, 389 indeterminate-BE-IND, 414 low grade BE-LGD)	NA for the whole cohort - has Table 1 with different ages for each group	Patient with BE and NDBE, BE-IND and BE-LGD followed up in time	Routine clinical follow up-surveillance.	Commparison of p53 among progressors (patients having a baseline diagnosis of NDBE, BE-IND, or BE-LGD, followed by a diagnosis of BE-HGD or EAC) and nonprogressors (patients having a baseline diagnosis of NDBE, BE-IND, or BE-LGD with ≥3 years of follow-up without progression confirmed by at least 1 additional EGD with biopsies)	NA for the whole cohort - has Table 1 with different FU times for each group	utility of p53 in predicting progression to HGD/EAC	Abnormal p53 IHC highly correlated with TP53 mutation status (90.6% agreement) and was strongly associated with neoplastic progression in the retrospective cohorts, regardless of histologic diagnosis (P < .001). In patients who progressed to advanced disease, p53-ABNL in baseline endoscopies was 49.7%, 90.0%, and 94.2% in NDBE, BE-IND, and BE-LGD, respectively. These numbers were dramatically lower in nonprogressing patients, with 1.7%, 15.4%, 45.4% positive in NDBE, BE-IND,and BE-LGD, respectively (P < .00001 for all). In our NDBE case-control testing cohort, p53-ABNL in the baseline endoscopy had a sensitivity of 50.8% and specificity of 98.3% for progression, with an odds ratio (OR) of 58 (95% confidence interval [CI], 17.9–188.5; P < .0001). In our abnormal histology testing cohort, p53-ABNL in the baseline endoscopy with a diagnosis of IND was associated with a sensitivity of 90.0% and specificity of 84.6% for progression with an OR of 49.5 (95% CI, 10.0–245.0; P < .0001). p53-ABNL in the baseline endoscopy with a diagnosis of LGD was associated with sensitivity of 94.2% and specificity of 54.6% for progression with an OR of 17.8 (95% CI, 6.4–49.5; P < .0001). In the retrospective cohort, abnormal p53 was associated with a hazard ratio (HR) of 5.03 (95% CI 3.88–6.5) for the whole cohort and a HR 5.27 (95% CI 3.93–7.07) for patients with exclusively nondysplastic disease before progression, adjusted for age, sex, and histologic diagnosis. The prevalence of p53-ABNL was calculated for sequential time points before progression and was found to be stable over time. In contrast, a diagnosis of BE-IND or BE-LGD occurred closer to progression with the prevalence of histologic abnormalities (BE-IND or BE-LGD) in all progressor biopsies falling steadily between 1 and 3 years before progression. Abnormalities in p53 IHC were present at a higher frequency than abnormalities in morphologic diagnosis at all time points, and this difference was most striking at time points more than 2 years before progression. In the prospective validation cohort, p53 IHC predicted progression among NDBE, BE-IND and LGD (P < .001). In the entire cohort, p53-ABNL corresponded to an HR of 12.51 (95% CI, 7.984–19.61; P < .0001). Performing a subanalysis based on baseline diagnosis showed an HR of 3.29 (95% CI, 2.05–5.29; P = .0002) for patients with baseline BE-LGD and an HR of 5.10 (95% CI, 2.11–12.28; P < .0001) for patients with baseline BE-IND, and HR of 11.83 (95% CI of 0.15–919.5; P < .0001). Under all conditions of having the index endoscopy or only surveillance endoscopies and under all histologic diagnoses, p53 IHC was able to stratify progressors from nonprogressors.	The largest available study on p53 and progression
Pinto (2022)	cross-sectional study	NA	NA	BE (n = 19) and EAC (n = 145) samples	NA	NDBE and EAC samples	investigation of prevalence of mutations in BE/EAC	investigation of prevalence of mutations in BE/EAC	NA	Investigation of the prevalence of core genetic (TP53 mutations and microsatellite instability (MSI) status) and epigenetic (DNA promoter hypermethylation of APC, CDKN2A, MGMT, TIMP3 and MLH1) modifications in a cohort of non-dysplastic BE and EAC samples.	Overall, none of the BE harbored TP53 mutations, whereas 30 out of 108 (28%) EAC samples carried mutations. None of the BE lesions and seven out of 108 tumors (6%) showed MSI. The promoter DNA methylation status of four genes (APC, CDKN2A, MGMT and TIMP3) was evaluated. For each gene, the promoter methylation frequency was significantly higher in BE or tumor samples compared to the tumor adjacent normal counterpart (p<0.05 for all). APC, CDKN2A, MGMT and TIMP3 promoter hypermethylation is frequently seen in both BE and EAC (21–89%), as well as in a subset of adjacent normal samples (up to 12%). Overall, 16% BE and 7% EAC samples showed hypermethylation of all four genes simultaneously.	Prevalence of core genetic (TP53 mutations and microsatellite instability (MSI) status) and epigenetic (DNA promoter hypermethylation of APC, CDKN2A, MGMT, TIMP3 and MLH1)
Roumans (2022)	multicenter cohort study	NA	BA	631 (3276 endoscopies were performed)	median age of 60 years. (IQR 53–69).	Consecutive BE patients from 15 Dutch hospitals. Histologically confirmed intestinal metaplasia in biopsies obtained from columnar lined epithelium in the esophagus, BE≥2 cm, and absence of a history of HGD/EAC. To exclude prevalent cases of neoplasia at baseline, only BE with ≥6 months of FU in the study without HGD/EAC development were selected for this analysis.	Routine clinical follow up-surveillance.	Comparison of progressors vs. non-progressors	median FU time was 6.8 years (IQR 4.9–9.8). Overall, 4475 person-years were observed	Progression to HGD/EAC. Usefulness of dynamic use of dysplasia and p53 (overexpression or loss of expression) and SOX2 (loss of expression) immunohistochemistry to predict outcome and therefore inform on follow up strategies for each individual patient, rather than having fixed strategies for all. A multivariate joint model was used in which longitudinal data of the risk of aberrant measurements of a biomarker are combined with the risk of neoplastic progression. Secondly, these models were combined with a time-varying Cox proportional hazards model in the framework of a multivariate joint model, to estimate the dynamic risk of neoplastic progression. Both the Cox model and the multivariate joint model were adjusted for baseline values age and gender, and time-varying dichotomous covariates BE length and esophagitis. The risk of neoplastic progression was estimated based on the longitudinal evolution of LGD, p53, and SOX2. For the results of the biomarkers of the multivariate joint model the value (representing the current risk of neoplastic progression, based on all previous measurements) and the accumulated effect of this biomarker (representing the overall risk of neoplastic progression, based on the history of the measurements of biomarkers) are reported.	The risk of aberrant expression of p53 increased in time (OR 1.17, p<0.01), with older age (OR 1.82, p<0.01), and male gender (OR 0.23, p<0.01 (ref. male)), but also with a long segment BE (OR 2.77, p<0.01). Loss of SOX2 expression was not influenced by any of these factors. An increased risk of developing HGD/EAC during surveillance was associated with aberrant expression of p53 (HR value 1.26, p<0.01) or SOX2 (HR value 1.43, p<0.01). The results of LGD (HR value 1.02, p = 0.78 & HR accumulated effect 1.02, p = 0.12) were not statistically significant. Consequently, there may be an association with an increased risk of neoplastic progression and LGD, but the association with aberrant expression of p53 or SOX2 is considerably larger. This indicates that multiple successive aberrant measurements of p53 or SOX2 increase the risk of developing HGD/EAC. If the model will be used in an online application, it will include all demographic and clinical variables of that individual patient, as well as the longitudinal evolutions of histological diagnosis, p53, and SOX2 to estimate the neoplastic progression risk. The risk estimations can be updated at every surveillance endoscopy, based on new additional measurements of histological diagnosis and immunohistochemistry. This results in dynamic risk estimations for each patient, according to its individual patient characteristics. These risks for the development of HGD/EAC gradually evolve from low risk to high risk. Each risk will have its consequence in surveillance, for example for low risk the interval can be chosen to remain long (e.g. three years), for medium risk the interval could be shortened (e.g. towards one year), and for high risk patients endoscopic eradication therapy (EET) may be applied. Optimism-adjusted estimates of the AUC were between 0.80 and 0.88 at different time points, indicating good performance of the model.	p53, SOX2. This is the first 'proof-of-principle' study to explore a "dynamic" approach to risk assessment, and therefore personalize BE surveillance/management.

Study	Col2	Col3	Col4	Col5	Col6	Col7	Col8	Col9	Col10	Aim	Results	Conclusion
Vithayathil (2022)	Multicenter randomized crossover study	Block-randomized using computer generated randomization in blocks of 4	Endoscopists could not be blinded to the intervention arm but were blinded to the endoscopy and histology results of the pretrial endoscopy and other study arm.	154 recruited and randomized; 134 completed both arms of the study	67.3 (38.0–89.0)	≥18 years, BE >C2 and/or M3 on pretrial endoscopy referred for surveillance of nondysplastic BE (NDBE) or assessment of flat dysplasia	AFI-directed pCLE and targeted biopsies for molecular biomarkers (experimental arm).	Patients randomized to receive either high-resolution white-light endoscopy (HRWLE) with Seattle protocol biopsies (standard arm) or endoscopy with AFI-directed pCLE and targeted biopsies for molecular biomarkers (experimental arm). Patients crossed over to the other arm after 6 to 12 weeks.	NA	The primary outcome was the diagnostic accuracy for dysplasia of AFI-guided pCLE using the trial histology as the gold standard. Secondary outcomes included the following: (1) diagnostic accuracy of AFI-guided pCLE for dysplasia with reference to the overall histology, which included biopsy specimens taken within 12 months before enrollment; (2) added diagnostic value of molecular biomarkers; (3) time to perform the endoscopy; and (4) patient-reported experience related to experimental and standard endoscopy. A 3-biomarker panel including cyclin A, p53 (assessed with immunohistochemistry), and aneuploidy (assessed with image cytometry) was used.	Within the experimental arm, pCLE had a higher sensitivity than HRWLE for HGD/IMC (P =.046) and all grades of dysplasia (P = .01), but lower specificity (HGD/IMC, P = .01; all grades of dysplasia, P = .02). In per-patient analysis, there was no difference in the sensitivity of pCLE for dysplasia compared with HRWLE with the Seattle protocol (76.5%; 95% CI, 50.1–93.2 vs 76.5%; 95% CI, 50.1–93.2, respectively; P = 1.00 for HGD/IMC; 74.3%; 95% CI, 56.7–87.5 vs 80.0%; 95% CI, 63.1–91.6, respectively; P = .48, for all grades of dysplasia). The use of AFI-targeted pCLE led to 2.1 optical biopsies per patient on average compared with 12.3 tissue biopsy specimens taken in the Seattle protocol. Standard endoscopy missed 28 cases of dysplasia (miss rate, 51.9%), 11 of which were detected by experimental endoscopy. Experimental endoscopy missed 20 dysplastic cases (miss rate, 37%), of which 5 were diagnosed correctly by standard endoscopy. In the overall histology analysis, AFI-guided pCLE had a higher sensitivity for HGD/IMC than Seattle protocol biopsies (73.3%; 95% CI, 54.1–87.7 vs 43.3%; 95% CI, 25.5–62.6, respectively; P = .02). The difference in sensitivity for all grades of dysplasia was not statistically significant (63.0%; 95% CI, 48.7–75.7 vs 51.9%; 95% CI, 37.8–65.7, respectively; P = .13). In the per-patient analysis the sensitivity and specificity for dysplasia of individual biomarkers were 48.6% and 93.9% for p53, 47.1% and 69.4% for cyclin A, and 40.0% and 88.5% for aneuploidy, respectively. At multivariable analysis, p53, aneuploidy, and optical dysplasia correlated significantly with a diagnosis of dysplasia. A panel comprising these 3 biomarkers showed an area under the receiver operating curve of 0.83 (95% CI, 0.76–0.91) for a diagnosis of any grade of dysplasia and 0.88 (95% CI, 0.78–0.97) for a diagnosis of HGD/IMC. Using a threshold of 1 positive biomarker, this panel had a higher sensitivity than the Seattle protocol in detecting dysplasia in the overall histology analysis (81.5% vs 51.9%; P < .001) The difference was not statistically significant in the trial histology analysis (91.4% vs 80.0%; P = .16).	3-biomarker panel including cyclin A, p53, and aneuploidy. The addition of molecular biomarkers could improve diagnostic accuracy

<p style="color:red; text-align:center">the studies above show the value of p53 in predicting progression to HGD/EAC - the studies below show its value in diagnosing dysplasia</p>

Study	Col2	Col3	Col4	Col5	Col6	Col7	Col8	Col9	Col10	Aim	Results	Conclusion
Kaye (2009)	retrospective	NA	NA	173	NA	BE with or without dysplasia	H&E alone vs addition of p53 staining	H&E alone vs addition of p53 staining	NA	Use of haematoxylin and eosin (H&E) alone using the Vienna classification and assessed the p53 staining using a qualitative system to determine dysplasia in BE	Weighted kappa scores between pairs of pathologists showed substantial agreement and improved after p53 immunohistochemistry. The Vienna classification is useful and reproducible in BE. p53 immunohistochemistry assists in diagnosis in difficult cases and predicts progression.	p53 is useful for defining dysplasia in BE and increases IOA, and predicts progression
Kaye (2016)	retrospective	NA	NA	72 cases encompassing the full spectrum of BE	NA	72 cases encompassing the full spectrum of BE	H&E alone vs addition of p53 staining	H&E alone vs addition of p53 staining	NA	Use of haematoxylin and eosin (H&E) alone using the Vienna classification and assessed the p53 staining using a qualitative system to determine dysplasia in BE	For the four-tier Vienna system, the average unweighted kappa was 0.30. Weighted kappa values varied from 0.27 to 0.69 with an average of 0.47. When grouped into definite dysplasia versus no definite dysplasia the average kappa was 0.55, but the kappa for low-grade dysplasia (LGD) versus high grade dysplasia (HGD) was only 0.31. For p53, using the three recognized patterns, the unweighted kappa was 0.6 (confidence interval 0.58–0.63). When cases were evaluated with both H&E and p53 the average kappa was 0.61 for definite dysplasia versus the rest. p53 immunohistochemistry interpretation is more reliable than dysplasia diagnosis, even with limited training. As it is predictive of prognosis and improves diagnostic reproducibility, it is suitable for routine use by pathologists as an adjunct to dysplasia diagnosis.	p53 is useful for defining dysplasia in BE and increases IOA
Kinra (2018)	retrospective	NA	NA	consecutive cases of BE (n=59), over a period of 4 years	NA	59 cases encompassing the full spectrum of BE	H&E alone vs addition of Alpha-Methyl Acyl-CoA Racemase (AMACR), p53, CyclinD1, β-catenin, H2AX and M30 immunohistochemical (IHC) stains	H&E alone vs addition of Alpha-Methyl Acyl-CoA Racemase (AMACR), p53, CyclinD1, β-catenin, H2AX and M30 immunohistochemical (IHC) stains	NA	Use of haematoxylin and eosin (H&E) alone using the Vienna classification and assessment of Alpha-Methyl Acyl-CoA Racemase (AMACR), p53, CyclinD1, β-catenin, H2AX and M30 immunohistochemical (IHC) stains to determine dysplasia in BE	Among the IHC stains performed, p53, β-catenin, H2AX and M30 stains were significantly useful to differentiate between IFD and LGD (P values: 0.04, 0.004, 0.05 & 0.04, respectively). AMACR and β-catenin stains though were up-regulated in HGD/adenocarcinomas than in other categories, their expression were not statistically different between the IFD and LGDs. Using a combined panel of IHC stains seems helpful in detection of dysplasia in BE, especially to differentiate the IFD and LGD changes in BE.	a combined panel of IHC stains seems helpful in detection of dysplasia in BE
van der Wel (2018)	retrospective	NA	NA	66 H&E slides (20 LGD, 20 HGD and 20 NDBE)	NA	BE with or without dysplasia	H&E alone vs addition of p53 staining	H&E alone vs addition of p53 staining	NA	Assess the added value of p53 immunohistochemistry (IHC) for the homogeneity within a group of dedicated gastrointestinal (GI) pathologists. 66 H&E slides (20 LGD, 20 HGD and 20 NDBE) were digitalised and independently assessed twice in random order by 10 dedicated GI pathologists. After a 'wash-out' period, cases were reassessed with the addition of a corresponding p53 IHC slide.	Addition of p53 IHC decreased the mean proportion of IND diagnoses from 10 of 60 to eight of 60 (P = 0.071). Mean interobserver agreement increased significantly from 0.45 to 0.57 (P = 0.0021). The mean diagnostic accuracy increased significantly from 72% to 82% (P = 0.0072) after p53 IHC addition. Addition of p53 IHC significantly improves the histological assessment of BO biopsies, even within a group of dedicated GI pathologists. It decreases the proportion of IND diagnoses, and increases interobserver agreement and diagnostic accuracy.	p53, improves diagnosis over histology alone
Toon (2019)	retrospective	NA	NA	28 sections from 23 patients	NA	BE with or without dysplasia	description of p53 profiles in BE with or without dysplasia	just descriptive	NA	description of p53 profiles in BE with or without dysplasia	In non-neoplastic epithelium, normal p53 staining was weak, heterogenous and localised to the crypts. In dysplastic epithelium, p53 overexpression was seen which was of moderate to strong intensity in either a crypt predominant location or diffuse involving crypt and surface epithelium. The crypt predominant pattern was observed more commonly in low grade dysplasia while the diffuse pattern was more commonly seen in high grade dysplasia. In a minority of cases, there was complete loss of p53 staining in dysplastic epithelium and contiguous neoplasia (null phenotype).	p53. Simply description of p53 profiles in BE with or without dysplasia

Study	Type			N	Age	Population	Index	Comparator	FU	Aim	Results	Conclusion
van der Wel (2020)	retrospective	NA	NA	51 international pathologist, each assessed 55 digitised BE biopsies before and after viewing matched p53 labelling.	NA	BE with or without dysplasia	H&E alone vs addition of p53 staining	H&E alone vs addition of p53 staining	NA	Aim: assess BE concordance rates and pathologist features predictive of diagnostic discordance.	Excellent concordance (>70%) for NDBE and HGD, and intermediate concordance for LGD (42%) and IND (23%). Major diagnostic errors were found in 248 diagnoses (8.8%), which reduced to 232 (8.3%) after viewing p53 labelled slides. At least 5 years of professional experience was protective against major diagnostic error for H&E slide review (OR 0.48, 95% CI 0.31 to 0.74). Working in a non-teaching hospital was associated with increased odds of major diagnostic error (OR 1.76, 95% CI 1.15 to2.69); however, this effect was neutralised if pathologists in these settings viewed cases with additional p53 labelled slides (OR 1.44, 95% CI 0.92 to 2.28). As expected, routine use of p53 labelled slides was associated with reduced odds of major diagnostic error. Excellent diagnostic agreement when reporting NDBO, LGD and HGD on H&E-stained slides alone (84.4%, 65.3% and 78.3%, respectively), rising to 89.4% when LGD and HGD diagnoses were combined. After revealing the matching p53 labelled slide for the 55 cases, agreement further improved to 85.9% for ND, 72.7% for LGD and 76.7% for HGD, rising to 91.9% when LGD and HGD were combined. Addition of matched p53 labelled slides improved diagnostic concordance, with small but clinically meaningful improvements seen in the diagnostic concordance between participating pathologists for NDBO reference diagnosis cases (83.8% vs 78.8% on H&E slide) and LGD/HGD combined reference diagnosis cases (79.3% vs 77.5% on H&E slide). In addition to this, p53 labelled slides also had a small but beneficial impact on reducing the number of major overinterpretations and underinterpretations (8.3%, 232 of 2805 diagnoses), representing 0.5% fewer overall major misinterpretations compared with H&E-stained slide diagnosis alone.	p53, improves diagnosis over histology alone
Januszewicz (2022)	multicenter retrospective cohort study	NA	NA	216 BE-IND specimens from 185 patients	mean (±SD) 64.8 (±11.6)	Inclusion criteria were: *i.* age 18 years or older *ii.* endoscopic evidence of BE ≥1 cm in length; *iii.* Presence of intestinal metaplasia (IM) on biopsies; *iv.* at least one biopsy from the study endoscopy was reported as BE-IND.	evaluation of p53 to help with evaluation of dysplasia in patients defined as "indeterminate for dysplasia"	Routine clinical follow up-surveillance.	median FU time was 5.3 years (IQR 1.3 – 9.2 years).	assess the utility of p53 immunohistochemistry (p53-IHC) in assessing BE-IND specimens. Comparison of the rate of changed diagnosis and the interobserver agreement (IOA) for all BE grades before and after p53-IHC.	Over half of the cases were reclassified to a non-dysplastic BE (NDBE), while 5.6% of cases in Group A and 7.4% in Group B were reclassified to definite dysplasia (pathologists were divided into two groups, to assure that each sample was assessed independently by a pathologist in one group and a pathologist in another group). Use of p53-IHC led to a >40% reduction in BE-IND diagnoses (P<.001), and increased IOA for all BE grades (κ=0.46 [NDBE], 0.26 [BE-IND], 0.49 [LGD], 0.35 [HGD/IMC]). An aberrant p53-IHC pattern significantly increased the likelihood of reclassifying BE-IND to definite dysplasia (odds ratio 44.3, 95%CI:18.8-113.0).	use of p53 for the definition of BE with indefinite for dysplasia
BarreGEN												
Ellsworth (2012)	retrospective	NA	NA	271 patients	NA (given in groups of 10 year each in Table 1, but not an overall measure)	patients with BE (NDBE, IND, LGD, HGD)	evaluation of mutational load (ML)	just descriptive	NA	Evaluation of mutational load NDBE, IND, LGD and HGD. Mutational load (ML) defined by loss of heterozygosity (LOH) using a panel of 16 LOH mutational markers associated with common tumor suppressor genes relevant to BE. The panel contained markers at the following 10 chromosomal loci (associated genes in parenthesis): 1p (CMM1, L-myc), 3p (VHL, HoGG1), 5q (MCC, APC), 9p (CDKN2A), 10q (PTEN, MXI1), 17p (TP53), 17q (NME1), 18q (DCC), 21q (TFF1 and PSEN2) and 22q (NF2). The presence or absence of mutations and the clonality of each mutation were determined for each marker	ML correlated with the histological classification with increasingly severe histology having higher ML. Three levels of mutation load (no ML, low ML, and high ML) were defined. NDBE and IND had mostly no/low, whereas HGD had consistently high ML. In HGD targets a relatively high proportion of DNA (>75%) was found with these mutations (high clonality). Mutations found in non-dysplastic histological classifications (NDBE, IND) were typically low clonality. There were less mutations detected in targets histologically classified as normal squamous epithelium and epithelium containing columnar cells that were not intestinalized (columnar, non-Barrett's epithelium), and importantly, there were no high clonality mutations found in these microdissected targets. The highest frequency of mutations was seen at 17p (TP53), with mutations present in 14/16 (88%) HGD targets, 27/39 (69%) LGD targets, and 49/138 (36%) "indefinite for dysplasia" targets. 9p (CDKN2A) was also more frequently mutated than other loci with 7/16 (44%) HGD targets, 20/39 (51%) LGD targets, and 45/138 (33%) "indefinite for dysplasia" targets displaying mutations. Mutational load was positively correlated to histological classification, with the number and clonality of mutations increasing with increasingly severe histological classification.	Mutational load (ML) defined by loss of heterozygosity (LOH) using a panel of 16 LOH in 1p (CMM1, L-myc), 3p (VHL, HoGG1), 5q (MCC, APC), 9p (CDKN2A), 10q (PTEN, MXI1), 17p (TP53), 17q (NME1), 18q (DCC), 21q (TFF1 and PSEN2) and 22q (NF2)
Khara (2014)	retrospective	NA	NA	415 patients histologically known to have BE	NA	patients with BE (NDBE, IND, LGD, HGD)	ML measured the presence and clonality of LOH mutations and the presence of MSI at each genomic locus examined. Low clonality (50–75% of the DNA containing LOH), and high clonality (>75 % of the DNA containing LOH)	just descriptive	NA	Evaluation of mutational load NDBE, IND, LGD and HGD. Mutational load (ML) defined by loss of heterozygosity (LOH) using a panel of 16 LOH mutational markers associated with common tumor suppressor genes relevant to BE. The panel contained markers at the following 10 chromosomal loci (associated genes in parenthesis): 1p (CMM1, L-myc), 3p (VHL, HoGG1), 5q (MCC, APC), 9p (CDKN2A), 10q (PTEN, MXI1), 17p (TP53), 17q (NME1), 18q (DCC), 21q (TFF1 and PSEN2) and 22q (NF2). The presence or absence of mutations and the clonality of each mutation were determined for each marker. Microsatellite instability (MSI) around these tumor suppressor genes was also included in the assessment of ML. High-clonality LOH mutations were assigned a value of 1, low-clonality mutations were assigned a value of 0.5, and MSI 0.75 at the first loci, and 0.5 for additional loci. These values were summed to the ML. Three levels of ML: (1) "No ML" contained microdissected targets that lacked mutations. (2) "Low ML" contained targets that had mutations, but the level of ML in this category was below the top 5th percentile of IM targets that had the highest ML. (3) "High ML" contained microdissected targets with an ML similar to those targets in the top 5th percentile of IM targets with the highest ML.	Low clonality LOH and MSI mutations were detected in similar or higher abundance than high clonality LOH mutations in less severe histological classifications (IM, IND, LGD), suggesting these mutations may occur prior to the appearance of more advanced histological stages of BE. Increasing ML correlated with increasingly severe histology. By contrast, proportions of targets that lacked mutations decreased with increasingly severe histology. Importantly, LOH and MSI mutations at all loci were detected with less advanced stages of BE histology (IM, IND, LGD) but were found more frequently with more advanced stages of BE histology (HGD, EAC). When a weighted value for MSI mutations was included in the assessment of ML, the positive correlation between ML and increasingly severe histology was slightly improved (correlation coefficient=0.69, p<0.0001). Importantly, the addition of MSI to the assessment of ML helped to better discriminate the difference in ML between less advanced (IM, IND, LGD) and more advanced (HGD, EAC) histological classifications of BE. The average difference in ML between less advanced histology (IM, IND, LGD) and more advance histology (HGD, EAC) was statistically higher when both LOH and MSI was considered as compared to only LOH (2.54 vs. 2.21, p=0.02). Our results demonstrate that histology-guided assessment for ML provided an objective measure of genomic instability amongst BE histological classifications. Increasing ML correlated with increasingly severe BE histology. The addition of MSI characterization at each of the ten genomic loci to the assessment of ML slightly increased this correlation.	Mutational load (ML) defined by loss of heterozygosity (LOH) using a panel of 16 LOH in 1p (CMM1, L-myc), 3p (VHL, HoGG1), 5q (MCC, APC), 9p (CDKN2A), 10q (PTEN, MXI1), 17p (TP53), 17q (NME1), 18q (DCC), 21q (TFF1 and PSEN2) and 22q (NF2). Microsatellite instability (MSI) around these tumor suppressor genes was also included in the assessment of ML.

Study	Type			Patients	Age	Description	Follow-up	Comparison	Follow-up duration	Outcome	Results	Notes
Eluri (2015)	case–control study	NA	NA	69 patients (46 controls and 23 cases)	62.5 (controls; non-progressors) 63.9 (cases; progressors)	Cases had NDBE or LGD at baseline and developed HGD/EAC ≥1 year later. Controls were matched 2:1, had NDBE or LGD and no progression at follow-up.	Routine clinical follow up-surveillance.	Comparison of progressors vs. non-progressors	mean 4 years	Prediction of progression to HGD/EAC with the use of mutational load (ML) defined as in the studies above by LOH and MSI	Mean ML in pre-progression biopsies was higher in cases (2.21) than in controls (0.42; *P* <0.0001). The ML of progressors ranged from 0.5 to 6.75 and the ML of nonprogressors ranged from 0.00 to 2.75. No case had an ML of 0 in their pre-progression tissue, compared with 25/46 (54%) of controls. Sensitivity was 100% at ML ≥0.5 and specificity was 96% at ML ≥1.5. Accuracy was highest at 89.9% (95%CI 80.2–95.8)for ML ≥1. ROC curves for ML ≥1 demonstrated an area under the curve (AUC) of 0.95 (95% CI 0.89–1.0). All 10 loci showed a markedly higher rate of mutation in cases compared with controls. The most frequent locus in the control group was 9p. In the cases, the most frequent loci were 9p, 17p, and 5q. ML in pre-progression BE tissue predicts progression to HGD or EAC.	ML defined as above
Eluri (2018)	nested case-control study	NA	NA	159 patients (101 controls-nonprogressors and 58 cases-progressors)	58.9 ± 8.1	Cases had baseline NDBE and developed HGD/EAC ≥2 year later. Controls were matched 2:1, had NDBE and no progression at follow-up.	Routine clinical follow up-surveillance.	Comparison of progressors vs. non-progressors	mean 4.0 ± 1.6 years	Prediction of progression to HGD/EAC with the use of mutational load (ML) defined as in the studies above by LOH and MSI. However, crude DNA lysates were used for ML determination in this study.	There was no difference in mean ML in pre-progression tissue in progressors and non-progressors (ML = 0.73 ± 0.69 vs. ML = 0.74 ± 0.61, P = 0.93). ROC curves showed poor discrimination of ML in predicting progression with AUC of 0.50 at ML ≥ 1. AUC did not vary with different ML cut-points. OR for MLML ≥ 1 was 0.77 (0.33–1.83). The utility of the to stratify BE patients for risk of progression was not confirmed in this study. The etiology for discrepancies between this and prior studies showing high predictiveness is likely due to the use of crude lysates in this study, but this requires further investigation.	this is a validation study on ML (defined as above); similar methods as Eluri 2015 - discrepancy of results, possible from different methods used
Das (2016)	cost-effectiveness study - Markov decision modeling and simulation	NA	NA	NA	mean 50 yo	Hypothetical cohort of men with mean age of 50 year old and diagnosed with NDBE at baseline EGD follwed by ACG guidelines 2011	simulation study	strategy I, natural history without surveillance; strategy II, surveillance per current guidelines; strategy III, ablation for all patients; strategy IV, risk stratification with use of a biomarker panel to assess genomic instability (i. e., mutational load [ML]: no ML: minimal surveillance; low ML: standard surveillance, high ML: ablation).	a calculated period of 174,853 person-years	Comparison of different strategies, including a strategy guided by biomarkers. Mutational load (ML) include loss of heterogeneity (LOH) in 17p (TP53), 9p (CDKN2A), 1p (CMM1, L-myc), 3p (VHL, HoGG1), 5q (MCC, APC), 10q (PTEN, MXI1), 17q (NME1), 18q (DCC), 21q (TFF1, PSEN2), and 22q (NF2) genomic loci, and presence of microsatellite instability (MSI) at these loci. ML was assessed with BarreGen and PathFinderTG in esophageal biopsy specimens from patients with NDBE	Strategy IV provided the best values for quality-adjusted life years (QALYs), ICER, and INHB in comparison with strategies II and III. Results were robust in sensitivity analysis. In a Monte Carlo analysis, the relative risk for the development of cancer in the patients managed with strategy IV was decreased. Critical determinants of strategy IV cost-effectiveness were the complete response rate, cost of ablation, and surveillance interval in patients with no ML. The use of ML to stratify patients with NDBE by risk was the most cost-effective strategy for preventive EAC treatment.	Mutational load (ML) was assessed with BarreGen and PathFinderTG in esophageal biopsy specimens from patients with NDBE (LOH) in 17p (TP53), 9p (CDKN2A), 1p (CMM1, L-myc), 3p (VHL, HoGG1), 5q (MCC, APC), 10q (PTEN, MXI1), 17q (NME1), 18q (DCC), 21q (TFF1, PSEN2), and 22q (NF2))
Trindade (2019)	single-center, retrospective pilot study	NA	NA	28 (8 progressors, 20 non-progressors)	mean 63.75 yo	Patient with BE and IND	Routine clinical follow up-surveillance.	Comparison of progressors vs. non-progressors	NA (different follow ups mentioned in Table 1, but not a global one)	Association of ML (defined as above) with progresson from IND to LGD/HGD/EAC	Out of 8 patients who progressed, 7 had some level of genomic instability detected in their IND biopsy (ML ≥0.5). Ten of the 20 (50%) who did not progress had no ML level. At an ML cut-off above 1.5, the risk of progression to high-grade dysplasia was 33% vs 0% (p=0.005), with a sensitivity of 100% and a specificity of 85%.	ML defined as above. Small, retrospective study

Study	Type			Patients		Description				Outcome	Results	Notes
Chamil Codipilly (2022)	meta-analysis	NA	NA	18,842 patients in 7 studies (of which 3206 were BE and of which 302 were dysplastic BE)		Studies that compare use of both biopsy forceps and WATS3D vs biopsy alone	NA	NA	NA	Primary outcome: the incremental yield of WATS3D-detected dysplasia (defined as a composite of IND, LGD, HGD and EAC) over forcep biopsy. Secondary outcomes: incremental yields of HGD/EAC and rate of reconfirmation of WATS3D dysplasia on subsequent forcep biopsy when the initial result was positive on WATS3D but negative on forcep biopsy, additional time taken for acquiring WATS3D samples, adverse events from WATS3D, and progression to HGD/EAC and mortality.	Forceps biopsy diagnosed dysplasia in 15.9% (95%CI, 5.4%-30.5%) of cases, whereas the incremental yield with WATS3D was 7.2% (95% CI 3.9%-11.5%). Meta-analysis of 6 studies demonstrated that forceps biopsy diagnosed HGD/EAC in 2.3% (95% CI, .6%-5.1%) of patients, whereas the incremental yield with WATS3D was 2.1% (95% CI .4%-5.3%). These results must be interpreted within the context of most studies lacking reconfirmation of WATS3D-only dysplasia by subsequent endoscopic forceps biopsy. Notably, WATS3D was negative in 62.5% of cases where forceps biopsy identified dysplasia. Hence, its role likely remains that of an adjunctive technique to endoscopic surveillance with forceps biopsy at this time.Two studies reported reconfirmation of WATS3D dysplasia with forceps biopsy histology in only 20 patients in whom WATS3D was positive but biopsies were negative and who underwent new EGDs with biopsies (of these 4/9 in one study and 5/11 in another were confirmed). Procedural time (only 1 study): the addition of WATS3D brushings increased procedure time by an average of 4 minutes and 31 seconds, in addition to the 6 minutes and 55 seconds required for obtaining Seattle protocol biopsy samples. No adverse events were recorded in either the FB or WATS3D groups. WATS3D increases dysplasia detection; however, the clinical significance of this increased dysplasia detection remains uncertain.	Meta-analysis from 2000 to 2020, included 7 studies. CDx Diagnostics provided funding support for 5 of 7 studies. Since WATS3d is only about dysplasia and not properly biomarkers, I will only include the meta-analyses here.
Qumseya (2019)	ASGE guideline 2019 where brief meta-analysis was performed	NA	NA	6 studies, 6271 patients with BE		Studies that compare use of both biopsy forceps and WATS3D vs biopsy alone	NA	NA	NA	performance of WATS3D	WLE with random biopsy sampling detected 125 cases of dysplasia. The performance of WATS resulted in identification of 137 additional cases missed by WLE with random biopsy sampling. The relative increase in dysplasia detection was 48% (95% CI, 34%-60%%). For patients with history of dysplasia, the relative increase in dysplasia detection using WATS was 47% (95% CI, 32%-61%). The absolute increase in dysplasia detection using WATS was 10.6% (95% CI, 1.5%-19.8%). The relative increase in LGD detection was 21% (95% CI, 24%-40%). For studies reporting all patients with or without a history of dysplasia, referred to as all-comers, the relative increase in dysplasia detection was 52% (95% CI, 21%-82%), whereas the absolute increase in dysplasia detection was 2% (95% CI, 1.5%-2.5%). Four of the 6 studies reported the yield of WATS in detection of LGD. When the random effects model was used, the absolute increase in LGD detection was 1.8% (95% CI, 1.4%-2.3%).	Barrett's ASGE guideline 2019 where brief meta-analysis on WATS3D was performed. All the aforementioned studies were funded by the manufacturer (CDx Diagnostics, Suffern, NY).

Study	Type			N		Inclusion/Design	Methods	Comparison	Surveillance	Outcome	Results	Notes
Suresh Kumar (2020)	meta-analysis	NA	NA	20 392 endoscopies across 11 studies	NA	Studies that compare use of both biopsy forceps and WATS3D vs biopsy alone	NA	NA	NA	The primary outcome of the study was to analyse the incremental yield in the detection rates of BE and ED by WATS when used as an adjunct to FB. This was done by calculating the absolute and relative increase in detection rates of BE and ED, respectively, while using WATS with FB as compared with using FB alone. The relative increase in reduction rate is defined as the relative odds of detecting BE and ED by using WATS in conjunction to FB versus FB alone; while the absolute increase in detection rate is defined as the absolute difference of detection rates between WATS with FB and FB alone.	There was an absolute increase in detection of 16% (95% CI 0.10% to 0.22%, p<0.00001) when WATS was used with forceps biopsy. A relative increase of 1.62 times was seen in detection rates of BE (95% CI 1.28 to 2.05, p<0.0001) with the number needed to test (NNT) of 6.1 patients. For BE with dysplasia, a 2% absolute increase (95% CI 0.01 to 0.03, p=0.001) in additional diagnostic yield from WATS. A relative increase of 2.05 times was seen in the detection rate of ED (95% CI 1.42 to 2.98, p=0.0001) yielding an NNT of 50 patients.	Meta-analysis until 04/2020.

Study	Type			N		Inclusion/Design	Methods	Comparison	Surveillance	Outcome	Results	Notes
Prichard (2015)	cross-sectional study	NA	NA	39 biopsies (22 with non-dysplaastic BE and 17 with HGD)	NA	biopsies from patient with NDBE and HGD	NA	NA	NA	the first study that identified the biomarkers' images that were included in the TissueCypher and that distinguish NDBE from HGD	Multiple image analysis features derived from epithelial and stromal biomarkers, including immune biomarkers and morphology, showed significant differences between HGD and NDBE	development of TissueCypher as an adjunct to histology - description of the technology
Critchley-Thorne (2016)	Multicenter nested case–control study	NA	NA	366 (cases - those who progressed to HGD/EAC ≥1 year: 79; controls- those who did not progress ≥1 year: 287)	NA	Inclusion criteria: BE cases with a diagnosis of non-dysplasia (ND), indefinite for dysplasia (IND) or low-grade dysplasia (LGD). Exclusion criteria: were history of HGD/EAC, diagnosis of HGD/EAC in less than 1 year, insufficient tissue quality as assessed by a pathologist, and preparation of tissue with Bouin's fixative or methylene blue.	Develop and validate a test that predicts progression of BE to HGD/EAC based upon quantification of epithelial and stromal variables in baseline biopsies. Biopsies from cases and controls were randomly assigned to training or validation sets. Immunofluorescence analyses were performed for 14 biomarkers and quantitative biomarker and morphometric features were analyzed. Prognostic features were selected in the training set and combined into classifiers. The top-performing classifier was assessed in the validation set.	Comparison of biomarker and histology feature classifiers between progressors and non-progressors. A 3-tier, 15-feature classifier was selected from the process described in column "I" in the training set and tested in the validation set. The classifier stratified patients into low-, intermediate- and high-risk classes. The following candidate panel of 14 protein biomarkers was selected and examined in the study: K20, Ki-67, β-catenin, p16INK4a, AMACR, p53, HER2/neu, CDX-2, CD68, NF-κB-p65, COX-2, HIF-1α, CD45RO, and CD1a.	Surveillance time of 5.6 years	Progression of BE to HGD/EAC.	Cut-points were established for a low risk class (risk score of 0 to less than 5.5), intermediate risk class (risk score of 5.5 to less than 6.4), or high risk class (risk score of 6.4 to 10). In the training set, HRs were 4.19 (95% C.I. 1.52, 11.57) for intermediate- vs. low-risk and 14.73 (95% C.I. 6.55, 33.16) for high- vs. low-risk. At multivariable analysis, these results were significant when ajusted for pathologist's type (generalist and subspecialist), segment length, age, sex and percent cells overexpressing p53. In the validation set, HRs were 2.45 (95% C.I. 0.99, 6.07) for the comparison of the intermediate-risk versus low-risk group and 9.42 (95% C.I. 4.61, 19.24), for high-risk versus low-risk (p<0.0001 for log-rank and score tests). At multivariable analysis, these were independent from pathology analysis, segment length, age, sex, or p53 overexpression. The 15-feature classifier was the top performing risk prediction model with AUROC 0.772. The probability of progression to HGD/EAC by 5 years increased continuously as the 15-feature risk score increased. Prevalence-adjusted NPV and PPV were 0.98 and 0.26 using reported progression rates. Thus, the 3-tier classifier identifies patients at very low risk of progression within 5 years, as demonstrated by the prevalence-adjusted NPV of 0.98 in the validation set. The classifier also identifies patients at very high risk of progression, with prevalence-adjusted PPV estimated at 0.26, demonstrating high predictive performance considering the very low frequency of progression in BE.	TissueCypher. Many of the authors of this study hold equity ownership or stock options in Cernostics, Inc., the commercial entity that developed the proprietary TissueCypher technology used in this study. The study was partially funded/supported by industry.
Critchley-Thorne (2017)	Multicenter nested case–control study	NA	NA	175 (30 cases - patients with baseline ND, IND and LGD who progressed to HGD/EAC; 145 controls - patients who did not progress)	NA	Cases: patients with baseline ND, IND and LGD who progressed to HGD/EAC. Controls: patients who did not progress to HGD/EAC. Inclusion criteria were availability of tissue blocks and clinicopathologic data, and confirmation of intestinal metaplasia by a GI subspecialist.	Determine if the 3-tier classifier developed in the study by Critchley-Thorne 2016, could detect abnormalities indicative of a field effect in non-dysplastic (ND), indefinite for dysplasia (IND) or low-grade dysplasia (LGD) biopsies from BE patients with prevalent HGD/EAC.	In this study, we tested the hypothesis that the patients in the predicted high-risk class (according to the 3-tier classifier) have significantly higher risk for presence of prevalent HGD/EAC than patients in the predicted low-risk class	5.6 years	Determine if the 3-tier classifier developed in the study by Critchley-Thorne 2016, could detect abnormalities indicative of a field effect in non-dysplastic (ND), indefinite for dysplasia (IND) or low-grade dysplasia (LGD) biopsies from BE patients with prevalent HGD/EAC. Cases were patients who had HGD/EAC on repeat endoscopy in <1 year or had prior history of treated HGD/EAC, returned to ND, IND or LGD and had HGD/EAC on repeat endoscopy. The non-progressor controls did not show HGD/EAC on repeat endoscopy and had median HGD/EAC-free surveillance time of 5.6 years	The 15-feature risk classifier had the capability to distinguish prevalent HGD/EAC from non-progressors with AUROC of 0.893, whereas the % cells overexpressing p53 had AUROC 0.594. The 15-feature classifier could stratify patients with significantly different risks for prevalent HGD/EAC; ORs were 46.0 (95% C.I. 14.86–169, p<0.0001) for the comparison of the high-risk versus low-risk group and 7.67 (95% C.I. 2.24–28.14, p=0.001) for intermediate-risk versus low-risk group. The classifier identified both non-dysplastic and LGD biopsies from prevalent cases as high-risk. The probability of diagnosis of HGD/EAC on repeat endoscopy increased continuously as the 15-feature risk score increased. At multivariable analysis, this effect was independent of the type of pathologist (generalist or specialist), and in some cases the classifier identified prevalent HGD even when pathology was ND/LGD.	TissueCypher. Many of the authors of this study hold equity ownership or stock options in Cernostics, Inc., the commercial entity that developed the proprietary TissueCypher technology used in this study. The study was partially funded/supported by industry.

Study	Study design		Blinding	Sample size	Age	Population	Method	Comparison	Time period	Objective	Results	Conflict of interest
Hao (2019)	cost-effectiveness study - Markov decision modeling and simulation	NA	NA	Hypothetical cohort of 10,000 individuals with BE and NDBE, IND or LGD was used	Hypothetical cohort was 69.1 yrs old	Hypothetical cohort of 10,000 individuals with BE and NDBE, IND or LGD was used	Cost-effectiveness Markov decision model was constructed in Excel using Palisade DecisionTools Suite (Palisade Corporation, Ithaca, New York) for the disease progression of patients with BE and their surveillance and treatment protocol over a 5-year time frame.	Comparison between a TissueCypher-guided management that stratified patients with BE in different risk levels for progression towards HGD/EAC *versus* standard of care (SOC)	Simulation for a 5-year period	Compare cost and quality-adjusted life-years (QALYs) from the perspective of a US health insurer with care delivered by an integrated health system.	Base-case model results for a 5-year period comparing the Assay-directed care to the standard of care (SOC) estimated an incremental cost-effectiveness ratio (ICER) of $52,483/QALY in 2012 US dollars. Assay-directed care increased the use of endoscopic treatments by 58.4%, which reduced the progression to HGD, EAC and reduced EAC-related deaths by 51.7%, 47.1%, and 37.6%, respectively, over the 5-year period. A surveillance interval of 5 years in BE patients scored low-risk by the Assay, independent of pathologic diagnosis (NDBE, IND, LGD), resulted in a 16.6% reduction in endoscopies. Targeting of endoscopic therapies to patients scored high-risk by the Assay increased the number of endoscopic treatments by 58.4%, which resulted in reducing the incidence of HGD, EAC and EAC-related deaths by 51.7%, 47.1%, and 37.6% over 5 years, respectively. Sensitivity analysis indicated that the probability of the Assay being cost-effective compared to the SOC was 57.3% at the 100,000/QALY acceptability threshold. While the Assay strategy was estimated to add cost during the initial 3 years of adoption, it was estimated to lower future costs and improve outcomes due to reduced surveillance in low-risk patients, and early treatment in high-risk patients over a 5-year period.	The authors of this study hold equity ownership or stock options in Cernostics, Inc., the commercial entity that developed the proprietary TissueCypher technology used in this study. The study was partially funded/supported by industry.
Davison (2020)	A single-blinded, case-control study	NA	Single blinding (biopsies were tested in a blinded way)	268 overall (58 cases and 210 controls)	NA	Patients with BE and biopsies showing non dysplasia (ND), indefinite for dysplasia (IND) or low-grade-dysplasia (LGD). Cases (incident progressors) were diagnosed with HGD/EAC ≥1 year after baseline. Controls (non-progressors) with no progression for >5yrs. Cases were matched to controls based on age, sex, BE length and pathologic diagnosis.	Regular clinical follow up. The study was single-blinded. The risk prediction test was run on each specimen without knowledge of clinicopathologic information, and results were reported to an outside statistician.	Comparison between the results of the TissueCypher technology in cases (progressors) vs. controls (non-progressors). Same 3-tier risk classifier developed in the Critchley-Thorne 2016 study was used, to evaluate the 5-year risk of progression to HGD/EAC	Cases: follow up for median 2.7 years until HDG/EAC. Controls: median follow up of 7 years.	Progression of BE to HGD/EAC.	Sensitivity and specificity of the test at 5 years for 3-tier classification (low-, intermediate-, high-risk) were 29% and 86%, respectively, and 40% and 86%, respectively for 2-tier classification (Low-, intermediate/high-risk combined). The prevalence-adjusted PPV was 23%, indicating that 23% of patients who score high-risk with the test in the BE surveillance population would progress to HGD/EAC within 5 years. The prevalence-adjusted NPV was 96.4%. Patients who scored high-risk were 4.7x (95% C.I. 2.5-8.8) more likely to progress within 5 years than patients who scored low-risk (P<0.0001). The risk prediction test demonstrated superior risk stratification than p53 alone in the full set of patients (HR = 4.7 (95% C.I. 2.5-8.8), P<0.0001 for the risk prediction test versus HR = 1.6 (95% C.I. 0.8-3.5), P= 0.1923 for p53 alone. Patients with expert pathologist-confirmed ND who scored high-risk were at 5.1-fold (95% C.I. 2.1-12.3) increased risk of progression compared to ND patients who scored low-risk (p=0.0003). The adjusted PPV for the test in expert pathologist-confirmed ND was 26%, indicating that 26% of ND patients who score high-risk will progress to HGD/EAC within 5 years. Patients with expert pathologist-confirmed LGD were at 3.8-fold (95% C.I. 1.9-7.6) increased risk for progression compared to patients with ND, and adjusted PPV for confirmed LGD was 21.8%. Thus, the test identifies a subset of ND patients who are at higher risk of malignant progression than patients with expert-confirmed LGD. At multivariable analysis, the high-risk class and male sex provided predictive power that was independent of pathologic diagnosis, age, segment length and hiatal hernia. Biopsies taken with different protocols and different levels of the esophagus, still provided similar overall results.	TissueCypher. Many of the authors of this study hold equity ownership or stock options in Cernostics, Inc., the commercial entity that developed the proprietary TissueCypher technology used in this study. The study was partially funded/supported by industry.
Frei (2020)	Blinded, nested case-control cohort	NA	The risk prediction assay was run on each specimen in a blinded manner without knowledge of the clinical outcome (nonprogressor or progressor) or other clinicopathologic information.	76 pt with ND-BE (38 progressed HGD/EAC and 38 did not), matched on age, sex, and BE length)	Mean age 63±9 years	The Amsterdam ReBus cohort consists of BE patients who progressed toHGD/EAC (progressors) during endoscopic surveillance and those who never showed progression (nonprogressors) during endoscopic follow-up.	Evaluation of the TissueCypher performance if spatial samples (all random biopsy levels from the baseline endoscopy) and temporal samples (all available previous endoscopies back to 10 years before progression) are added to a single biopsy. The risk prediction assay was run on each biopsy specimen.	Single-biopsy vs multiple biopsies (spatial or temporal) evaluation.	Median 3.2 years (IQR 2.3–4.3) for until HGD/EAC diagnosis for progressors and 6.1 years (IQR 5.5–7.2) for non-progressors	Evaluation of the TissueCypher performance if spatial samples (all random biopsy levels from the baseline endoscopy) and temporal samples (all available previous endoscopies back to 10 years before progression) are added to a single biopsy. The risk prediction assay was run on each biopsy specimen.	A high-risk score was associated with a prevalence-adjusted annual progression rate of 6.9%. Patients who scored high risk were 3.23 (95% Cl 1.6–6.5; P = 0.0032) more likely to progress to HGD/EAC than patients who scored low risk. Sensitivity and specificity of the assay at 5 years were 30.4% and 95.0%, respectively. Evaluation of the highest scoring of all additional spatial biopsy levels from the BL endoscopy significantly increased the detection rate of progressors by 63.5% (from 30.4% to 49.8%; P 5 0.016). Specificity remained 95.0% irrespective of the number of tested levels. When evaluating all BL spatial biopsy levels, patients who scored high risk were 5.53(95% Cl 2.7–11.4; P<0.0001) more likely to progress than patients who scored low risk. The prevalence-adjusted positive (PPV) and negative (NPV) values for prediction of progression within 5 years were 34.6% and 97.7%, respectively, indicating an annual progression risk of 6.9% in NDBE patients scoring high risk. Temporal analysis of endoscopies from different times led to an additional, although not statistically significant increase of the detection rate by 37.6% (from 49.8% to 68.5%). Analysis of the highest scoring biopsy from all available BL and pre-BL endoscopies showed that patients who scored high risk were 7.0-fold (95% Cl 3.3–14.8; P <0.0001) increased risk of progression within 5 years vs patients who scored low risk	TissueCypher. Some of the authors of this study hold equity ownership or stock options in Cernostics, Inc., the commercial entity that developed the proprietary TissueCypher technology used in this study.
Diehl (2021)	cross-sectional study	NA	NA	60	65.2 ± 11.8	Patients with BE in regular clinical follow up	When providers ordered TissueCypher they were sent questionnaire on planned treatment/follow up program before receiving the TissueCypher results. Then, after receiving the results, they were sent again questinnaires on treatment/follow up program.	Pre- and post-result management program	NA	Two physicians recorded their management indications before and after receiving the TissueCypher results. Evaluation of the change in management that TissueCypher results brought about.	TissueCypher results impacted 55.0% of management decisions. In 13 (21.7%) patients, the test upstaged the management approach, resulting in endoscopic eradication therapy (EET) or shorter surveillance interval. The test downstaged the management approach in 20 (33.4%) patients, leading to surveillance rather than EET. The management plan was not impacted for 25 patients (42.4%) of whom 18 scored TissueCypher low-risk, three scored intermediate-risk, and four scored high-risk.	TissueCypher. Some of the authors of this study hold equity ownership or stock options in Cernostics, Inc., the commercial entity that developed the proprietary TissueCypher technology used in this study.

Study	Design		Risk assay	N	Age	Patient population	Evaluation	Comparison	Follow-up / Progressers	Aims	Results	Biomarker / Notes
Frei 2021	A blinded, retrospective cohort study	Patients derived from the screening cohort of a RCT of SURveillance vs RFA for BE-LGD patients	The risk prediction assay was run on each specimen in a blinded manner without knowledge of the clinical outcome (nonprogressor or progressor) or other clinicopathologic information.	155 (34 progressers, 121 non-progressors)	mean age 61 ± 10 years	Patients derived from the screening cohort of a RCT of SURveillance vs RFA for BE-LGD patients. Presence of LGD in the community at baseline.	Evaluation by 3 expert pathologists, and evaluation with the TissueCypher assay	Comparison of progressors and non-progressors among patients with LGD. Evaluation of cases where downstaging from expert pathologists to ND-BE is confirmed or not by TissueCypher assay.	Progressers (median time 2.4 years until HGD/EAC; IQR 1.0–5.1) and non-progressors (median time 7.9 years; IQR 5.9–10.3)	Investigate whether this TissueCypher assay can risk stratify BE patients with communitybased diagnosis of LGD, and to compare the predictive performance of this assay to 3 expert GI pathologists providing histologic diagnoses. Prediction of progression of LGD-BE to HGD/EAC.	Patients who scored intermediate/high risk with the assay were 6.7 (95% CI 3.2–13.8) more likely to progress to HGD/EAC than patients who scored low risk. Patients diagnosed as IND/LGD by the 3 pathologists were 4.3 (95% CI 2.0–9.3), 5.9 (95% CI 2.7–12.9), or 6.6 (95% CI 3.1–13.8), respectively, more likely to progress than patients who were downstaged to ND. The risk prediction assay sensitivity was 68% vs 76% for the 3 pathologists, and specificity was 79% vs 64%–77.0% for the pathologists. The PPV of the risk prediction assay was 38.4% for prediction of progression within 5 years, and 56.9% for prediction of progression within 10 years. By contrast, the PPVs associated with the 3 pathologists' diagnostic classes of IND/LGD were 29.2%, 32.2%, and 39.2% for progression within 5 years, and 40.7, 50.7%, and 51.9% within 10 years. The NPV of the assay was 92.5%, compared with 93.0%, 93.5%, and 94.2% for the ND groups according to the 3 pathologists. The assay detected 50%–56% of progressors that were downstaged to NDBE by the pathologists3), whereas the pathologists detected 33.3%–50.0% of the progressors that scored low risk. The PPV of the assay in patients who were downgraded to ND ranged from 16.5% to 20.3% depending on the pathologist. This indicates that patients who are downstaged to ND on expert review but score intermediate/high risk with TissueCypher progress at a rate of 3.3%–4.1% per year. The assay stratified BE patients with a community-based diagnosis of LGD with overall predictive accuracy comparable with 3 expert pathologists. The assay provided objective risk stratification, whereas there was significant variability between the 3 pathologists who agreed on only 51.7% of cases in this study. Furthermore, the assay identified approximately half of the progressors that the expert pathologists downstaged to NDBE.	TissueCypher. Some of the authors of this study hold equity ownership or stock options in Cernostics, Inc., the commercial entity that developed the proprietary TissueCypher technology used in this study.
Iyer (2022)	A Pooled Analysis of International Multicenter Studies	NA	NA	552 patients with baseline NDBE (n = 472), IND (n = 32), or LGD (n = 48)	61.5 (11.0)	Studies that included patients with BE but without HGD/EAC and evaluated with TissueCypher for progression to HGD/EAC.	Pooled analysis of five different studies. Deidentified data from these studies were shared with the authors with the permission of all investigators.	Comparison of progressors (n=152) and non-progressors (n=400)	NA	Aims: (1) assessing the incremental utility of the TissueCypher test in predicting progression in BE (NDBE, indefinite for dysplasia [IND], or LGD) patients with over clinical variables alone; and (2) assessing the incremental utility of the test in predicting progression in those with baseline NDBE over clinical variables. Secondary aims included: (1) assessing the incremental utility of the TissueCypher test in predicting both prevalent and incident HGD/EAC over 5 years; and (2) developing a BE progression risk score combining demographic, clinical, and TissueCypher test scores.	Of the 152 progressors (overall), 58 (38%) were in the high-risk class, 25 (16.4%) were in the intermediate risk class, and the remaining 69 (45%) were in the low risk class. A high-risk test class independently predicted increased risk of progression to high-grade dysplasia/adenocarcinoma (odds ratio, 6.0; 95% confidence interval, 2.9–12.0), along with expert confirmed low-grade dysplasia (odds ratio, 2.9; 95% confidence interval, 1.2–7.2). Model prediction of progression with the TissueCypher risk class incorporated was significantly superior than without, in the whole cohort (c-statistic 0.75 vs 0.68; P < .0001) and the nondysplastic BE subset (c-statistic 0.72 vs 0.63; P < .0001). Sensitivity and specificity of the high risk TissueCypher class were 38% and 94%, respectively. Sensitivity of a high-/intermediate-risk class in the entire cohort was 0.55, with a specificity of 0.82. A model was developed to output predicted probability of progression within 5 years.This risk calculator can predict the probability of progression for each patient, based on their baseline characteristics (it will need to be validated).	TissueCypher. Partially funded by Cernostics, Inc., the commercial entity that developed the proprietary TissueCypher technology used in this study. Five studies were used in this pooled analysis (each summarized above): Critchley-Thorne 2016, Critchley-Thorne 2017, Davison 2020, Frei 2020 and Frei 2021

DNA abnormalities

Study	Design			N	Age	Patient population	Evaluation	Comparison	Follow-up	Aims	Results	Biomarker / Notes
Altaf (2017)	meta-analysis	NA	NA	NA (different studies included for different biomarkers)	NA (different studies included for different biomarkers)	Included studie: BE with or without HGD/EAC, comparing biomarkers (p53, Ki-67, p16, DNA content abnormalities) in predicting neoplastic progression of BE to HGD/EAC. Only studies published in the English.	NA	NA	NA (different studies included for different biomarkers)	Progression to of BE to HGD/EAC. Comparing biomarkers (p53, Ki-67, p16, DNA content abnormalities) in predicting neoplastic progression of BE to HGD/EAC.	102 clinical studies were included in the study. Mutation of p53 had the following findings: diagnostic odds ratio (DOR) 10.91, sensitivity 47%, specificity 92%, positive likelihood ratio (PLR) 4.71, negative likelihood ratio (NLR) 0.65, area under the curve (AUC) 0.79. Loss of p53 had the following findings: DOR 16.16, sensitivity 31%, specificity 98%, PLR 6.66, NLR 0.41, AUC 0.923. Both mutations and loss of p53 were found to be superior to other p53 abnormalities such as loss of heterozygosity (LOH) and overexpression. Ki-67 had DOR 5.54, sensitivity 82%, specificity 48%, PLR 1.59, NLR 0.42, AUC 0.761. Aneuploidy had DOR 12.08, sensitivity 53%, specificity 87%, PLR 4.26, NLR 0.42, AUC 0.846. Tetraploidy had DOR 5.87, sensitivity 46%, specificity 85%, PLR 3.47, NLR 0.65, AUC 0.793. Loss of Y chromosome had DOR 9.23, sensitivity 68%, specificity 80%, PLR 2.67, NLR 0.49, AUC 0.807. p16 aberrations (hypermethylation, LOH, mutation and loss) failed to demonstrate any advantage over the other biomarkers studied.	p53, p16, Ki-67 and DNA content abnormalities (aneuploidy, tetraploidy, loss of Y chromosome). Meta-analysis. Included studies until March 2016.
Rabinovitch (2001)	prospective cohort study	NA	NA	307	NA	307 patients who had baseline biopsies (and histological and flow cytometric DNA content and S-phase assessment) between 07/1983-06/1998, as well as at least one follow-up evaluation	No specific intervention. Just regular clinical follow up. Evaluation of baseline genetic characteristics and progression towards EAC in the future	No specific intervention. Just regular clinical follow up. Evaluation of baseline genetic characteristics and progression towards EAC in the future	Total patient-years of follow-up in this study were 1338. The mean follow-up time from baseline to cancer or last endoscopy was 56 months (median 40, range 0.5 to 174).	Presence of DNA-content abnormality (aneuploidy, tetraploidy [4N]), and S-phase assessment at baseline biopsy in predicting the future risk of EAC.	42 (13.7%) developed cancer during follow-up. HGD, aneuploidy, or increased 4N (tetraploidy) fractions were detected at the baseline endoscopy in 90% of patients who progressed to cancer (38/42) and in 100% of patients who progressed to cancer within 5 yr (34/34). Tetraploidy increased risk of cancer in all different cut-offs considered (6%, 6-15% or >15%). The presence of aneuploidy with <2.7N had low risk of progression. The presence of elevated 4N fractions alone (RR = 11, CI = 5.1–25) or elevated 4N fractions and aneuploidy together (RR = 20, CI = 9.0–44) was more predictive of cancer outcome than were the presence of aneuploidy alone (RR = 4.4, CI = 1.4–14). The presence of either aneuploidy or elevated 4N fraction assumes greater predictive strength in the subset of patients with negative, indefinite, or low-grade dysplasia (RR = 25, CI = 6.5–98). Of 307 patients, 137 had an S-phase fraction of >5.5% at baseline endoscopy. The incidence of cancer 3 yr and 5 yr after such an elevated S-phase was 17% (CI = 12–25) and 21% (CI =15–30), respectively. The RR for patients with this elevated S phase compared to those without was 2.3 (CI = 1.2–4.4, p = 0.02). However, it was not significant when HGD was accounted for. Flow cytometry can be an adjunt to biopsies and helps define low- and high-risk subsets of patients without HGD.	DNA-content abnormality. Part of The Seattle Barrett's Esophagus Study
Famhy (2004)	retrospective	NA	NA	40 (21 with HGD/EAC and 19 with no dysplasia)	NA	BE patients with HGD/EAC or without dysplasia	regular clinical follow up/treatment	HGD/EAC vs NDBE	at least 5 years (for NDBE)	FISH to determine if there are specific genetic changes in Barrett's esophagus with associated high-grade dysplasia/intramucosal adenocarcinoma compared to those without dysplasia. Centromeric enumeration probes (CEP) for chromosomes 6, 7, 11, and 12, and locus-specific probes (LSI) for 9p21 (p16 gene), and 17p13.1 (p53 gene) loci along with their corresponding CEP (9 and 17, respectively) were used in this study. A positive FISH result was defined as the presence of cells with >2 CEP signals or with a loss of the LSI signals relative to their corresponding CEP.	p53 locus loss and/or aneusomy of chromosomes 6, 7, 11, and 12 abnormalities could be detected by FISH from 95% of HGD/EAC with a specificity of 100%. When combining the results of both FISH probe sets, aneusomy of chromosomes 6, 7, 11, and 12 or a p53 loss were identified in 20/21 (95%) of the HGD/carcinoma cases, including all five cases with cytologic changes classified as indefinite for dysplasia. The sensitivity and specificity for the detection of HGD/carcinoma using the combination of the above probes was 95 and 100%, respectively (95% CI: 74–99.8% and 79.1–100%; Loss of the p16 locus was seen commonly in patients both with (86%) and without (47%) dysplasia/carcinoma.	FISH for Centromeric enumeration probes (CEP) for chromosomes 6, 7, 11, and 12, and locus-specific probes (LSI) for 9p21 (p16 gene), and 17p13.1 (p53 gene) loci along with their corresponding CEP (9 and 17, respectively) can distinguish between HGD/EAC and NDBE

Study	Design			Sample	Age	Population	Intervention	Comparison	Follow up	Objective	Results	Conclusion/Notes
Maley (2006)	prospective cohort study	NA	NA	268 patients with BE followed in time	NA for all cohorts; <40 (6.8%), 40-49 (18%), 50-59 (25%), 60-69 (24%), 70-79 (22%), ≥80 (4%)	268 patients with BE with at least one follow up endoscopy	No specific intervention. Just regular clinical follow up endoscopy	No specific intervention. Just regular clinical follow up. Evaluation of baseline genetic characteristics and progression towards EAC in the future	Follow up for an average of 4.4 years (range 0.1–8.4) for a total of 1,179 patient years between 01/1995 and 08/2003	The role of number of clones, clonal diversity (measured by the Shannon diversity index) and genetic divergence (number of loci showing differences by loss of heterogeneity [LOH] divided by the number of heterozygous in normal loci) in the progression from BE to EAC. These were measured at baseline, and risk of progression was assessed in time. The Shannon index and the number of clones are strongly correlated, and one can substitute for the other in a multivariate model	The upper quartiles of the number of clones, Shannon index and genetic divergence based on LOH were strongly predictive of increased progression to EAC, showing that neoplasms with greater clonal diversity at baseline were more likely to progress to cancer than neoplasms with lesser clonal diversity. Controlling for TP53 LOH, tetraploidy and aneuploidy, which are known to be associated with genomic instability, we found that number of clones, divergence and Shannon index (all based on LOH), individually predict cancer outcome. The combination of factors that best predicted progression included the number of clones, genetic divergence, TP53 LOH and ploidy lesions.	Number of clones, clonal diversity and genetic divergence. Part of The Seattle Barrett's Esophagus Study. Clinically, assessment of clonal diversity may be a unified method to identify high-risk patients for early detection as well as warning of possible variants that may be resistant to cancer prevention interventions.
Galipeau (2007)	prospective cohort study	NA	NA	243	NA	BE patients with or without dysplasia	regular clinical follow up	comparison of progressors and non-progressors	A total of 17,139 patient-months with a mean of 71 mo (median 80.5 mo, range 2.3–130.8 mo)	TP53 and CDKN2A (p16) alterations (methylation/mutation), 17p LOH and 9p LOH, tetraploidy and aneuploidy were evaluated at baseline biopsies and patients followed up in time to observe progression	At 10 y, each molecular and DNA content abnormality, when analyzed alone (univariate) in a patient at baseline, made a significant contribution to prediction of EA risk: 17p LOH (10-y RR = 10.6; 95% CI 5.2–21.3, p, 0.001), 9p LOH (10-y RR =2.6; 95% CI 1.1–6.0, p= 0.03), p53 (RR 7.3 (3.7–14.3), tetraploidy RR 8.8 (4.3–17.7), aneuploidy RR 8.5 (4.3–17.0), but not but not CDKN2A mutation (10-y RR 1.8; 95% CI 0.8–4.1, p=0.13) and CDKN2A methylation (RR 2.1; 95% CI 0.8–4.1, p = 0.09). At final multivariable model, the following risks were seen: 17p LOH 5.4 (2.5–12.0) ,0.001; Tetraploidy 2.9 (1.4–5.9) ,0.001; Aneuploidy 3.4 (1.6–7.1) 0.001; 9p LOH 2.4 (1.0–5.5) 0.045. Compared to no abnormality present, the following relative risks (RR) were observed: one abnormality RR 1.8 (0.48–6.87), p . 0.38; two abnormalities RR 9.0 (2.4–33.3), p , 0.001; three abnormalities RR 38.7 (10.8–138.5), p , 0.001	TP53 and CDKN2A (p16) alterations (methylation/mutation), 17p LOH and 9p LOH, tetraploidy and aneuploidy. Part of the Seattle Barrett's Esophagus Study cohort
Fritcher (2008)	retrospective	NA	NA	92 (of whom, 84 had HGD/EAC, 7 had LGD and 1 NDBE)	mean age of 64.4 years (range, 34-87 years).	Patient with BE in surveillance	regular clinical follow up	regular clinical follow up	mean 267 days (range, 0-1304 days)	Comparison of sensitivity and specificity of conventional cytology, DNA ploidy analysis with digital image analysis (DIA), and fluorescence in situ hybridization (FISH) for the detection of dysplasia (FISH probes to probes to 8q24 (C-MYC), 9p21 (P16), 17q12 (HER2), and 20q13)	FISH was more sensitive (P < .05) than cytology and DIA for low-grade dysplasia, HGD, and EA.Sensitivities: for LGD: cytology, DIA, and FISH 5%, 5%, and 50% respectively; for HGD: 32%, 45%, and 82%, respectively; for EAC: 45%, 45%, and 100%, respectively. Specificity (on patients with only benign squamous mucosa: 93%, 86%, and 100% (P = .22). There was a significant difference between FISH categories (negative, 9p21 loss, gain of a single locus, and polysomy) for progression to HGD/EA (P < .001).	Comparison of sensitivity and specificity of conventional cytology, DNA ploidy analysis with digital image analysis (DIA), and fluorescence in situ hybridization (FISH) for the detection of dysplasia (FISH probes to probes to 8q24 (C-MYC), 9p21 (P16), 17q12 (HER2), and 20q13)
Sikkema (2009)	nested case-control study	NA	NA	54 patients (27 cases-progressors and 27 controls-non-progressos) matched on age, sex and follow up duration	cases: 58.8 years (range 36.6-76.2); controls 56.2 (range 29.6, 74.2)	BE who progressed to HGD/EAC and BE who did not progress to HGD/EAC	regular clinical follow up	Progressors to HGD/EAC vs. non-progressors. Regular clinical follow up	cases: 6.9 years (range 0.4, 16.3) and controls 7.9 (range 1.9-18.9)	Presence of LGD, DNA ploidy status (flow cytometry) and Ki67 and p53 expression (immunochemistry) to compare progressors towards HGD/EAC and non-progressors.	At univariate analysis, LGD (HR 3.6, 95%CI 1.6, 8.1), aneuploidy (HR 3.6; 95%CI 1.6, 8.1) strong (>20%) Ki67 expression (HR 5.2, 95%CI 1.5, 17.6) and moderate (>15%) p53 expression (HR 6.5, 95%CI 2.5, 17.1) increased the risk of progression towards HGD/EAC. At multivariable analysis, that included LGD as covariate, only moderate p53 (HR 5.4, 95%CI 2.0, 14.5) remained statistically significant. Strong Ki67 showed only a trend towards significance (20-50%: HR 2.2; 95%CI 0.9, 5.1 and >50% HR 3.2; 95%CI 0.9, 11.2) Aneuploidy was no longer a predictor (HR 2.3, 95%CI 0.8, 6.3).	p53, aneuploid and Ki67. Only p53 was strong at multivariable analysis, whereas Ki67 ws weak, and
Borovicka (2009)	cross-sectional multicenter	NA	NA	164 patients with 239 endoscopies	62.5 years (± 12.9)	patients with BE with or without dysplasia	regular clinical follow up	compare HGD/EAC with all other BE patients	2.6 years (± 3.3)	Addition of brush cytology processed by digital image cytometry (DICM) to biopsies to further increase the detection of patients with Barrett's esophagus who are at risk of neoplasia. DICM results (Aneuploidy/intermediate pattern vs. diploidy) to distinguish between HGD/EAC and all other BE patients	An intermediate DICM result carried a relative risk (RR) of 12 (95%CI 7.4–19.7) and aneuploidy a RR of 27 (95%CI 13.5–64.5) for high-grade dysplasia/adenocarcinoma. Adding DICM to the standard biopsy protocol, a pathological cytometry result (aneuploid or intermediate) was found in 25 of 239 endoscopies (11 %; 18 patients) with low-risk histology (no high-grade dysplasia or adenocarcinoma). During follow-up of 14 of these 18 patients, histological deterioration was seen in 3 (21%). Sensitivity for detecting histologically confirmed adenocarcinoma or HGD was 52% (95 %CI 33%–70%); the specificity with respect to a histological finding of SIM or no SIM was 98% (95 %CI 94%–99%). The positive predictive value of a nondiploid DICM was 80% (95 %CI 56%–94%), while the negative predictive value was 92% (95 %CI 87%–95%). The likelihood ratio of obtaining an aneuploid pattern from a patient with adenocarcinoma or HGD than from patients with SIM or no SIM was 13 (95 %CI 6.2–25).	Aneuploidy/intermediate pattern vs. diploidy to distinguish between HGD/EAC and all other BE patients
Merlo (2010)	prospective cohort study	NA	NA	239 patients with BE	63 yo	Patients with BE who had baseline (starting at 01/1995) information on 9p LOH, 17p LOH, ploidy, microsatellite shifts and sequence mutations in CDKN2A and TP53, and who were followed prospectively with clinical follow-up	No specific intervention. Just regular clinical follow up. Evaluation of baseline genetic characteristics and progression towards EAC in the future	No specific intervention. Just regular clinical follow up. Evaluation of baseline genetic characteristics and progression towards EAC in the future	Patients were followed up for an average of 5.2 years (range = 0.10 – 7.5 years)	Clonal diversity and risk of progression to EAC. Major diversity measurement methods, including genetic divergence and entropy based measures were calculated with microsatellite shifts and loss of heterozygosity, DNA content tetraploidy and aneuploidy, methylation and sequence mutations.	All diversity measures were strong and highly significant predictors of progression and the type of alterations evaluated had little effect on the predictive value of most of the diversity measures. Diversity is a strong predictor of progression from BE to EAC regardless of whether divergence, number of clones, Shannon index, Simpson index or other q values are used. Diversity is also a robust predictor of progression across all types of genetic alterations, whether LOH alone, selected alterations, neutral alterations, or all genetic alterations were used to define a clone. The consistency of our results with respect to the type of diversity measure and alterations used to define clones suggests that diversity measures are robust biomarkers for risk stratification.	Clonal diversity. Part of The Seattle Barrett's Esophagus Study. Same group as Maley 2006 paper.

Study	Study type			N	Age	Population	Intervention	Comparison/Aim	Follow-up	Outcome	Results	Biomarkers
Bird-Lieberman (2012)	nested case-control study	NA	NA	380 (89 cases - progressors and 291 controls - non progressors)	Cases: 63.8±11.9. Controls: 63.8±11.3	Cases were BE patients from the NIBR who developed HGD/EAC ≥6 months after their initial BE diagnosis.Each case was matched to up to 5 controls who had BE but had not developed HGD/EAC by the study censor date. All had to be diagnosed with pathology.	Routine clinical follow up-surveillance.	Progressors to HGD/EAC vs. non-progressors. Regular clinical follow up	mean 6.7 years (±3.3 y)	Progression of BE to HGD/EAC. The following biomarkers were evaluated: abnormal DNA ploidy, p53, and cyclin A expression, levels of sialyl Lewisa, Lewisx, and Aspergillus oryzae lectin [AOL] and binding of wheat germ agglutinin, evaluate on formalin-fixed paraffin-embedded tissue	A panel comprising low-grade dysplasia, abnormal DNA ploidy, and AOL most accurately identified progressors and nonprogressors. For each 1-point increase in risk score, BE patients with LGD are almost at 4-fold increased odds of progressing to develop EAC (OR, 3.90; 95% CI, 2.39 – 6.37), whereas BE patients without LGD have a 3-fold increased odds of progression (OR, 3.31; 95% CI, 1.81– 6.05). C-statistic for the reduced model of LGD, abnormal DNA ploidy, and AOL was 0.73.	LGD, abnormal DNA ploidy, and AOL. Used the population-based Northern Ireland BE Register (NIBR)
di Pietro (2015)	cross-sectional study	NA	NA	157 (training cohort), 46 (validation cohort)	66.4 (training cohort), 68.7 (validation cohort)	Age ≥18 yrs, BE with a length of at least C ≥ 2 or C<2M≥4 according to the Prague classification with or without visible lesions.	Biopsies were performed either according to the standard Seattle protocol or under the guidance of autofluorescence imaging (AFI)	Comparison of the accuracy of a panel of molecular biomarkers on AFI-directed biopsies with conventional Seattle protocol biopsies for HGD and EAC. Secondary aims: (i) assessment of diagnostic accuracy for the biomarkers for any grade of dysplasia and (ii) validation of a large panel of biomarkers in an independent prospective study by an independent laboratory.	NA	Accuracy of biomarkers in identifying HGD/EAC. Panel of biomarkers assessed: Aneuploidy/tetraploidy; 9p and 17p loss of heterozygosity; RUNX3, HPP1 and p16 methylation; p53 and cyclin A immunohistochemistry. The statistical analysis consisted of three stages: (1) per-biopsy analysis (correlation between biomarkers and histological outcome in individual targeted areas); (2) per-patient analysis (correlation between overall biomarker result and overall histological outcome in individual patients); and (3) comparison between AFI+ versus AFI− areas (comparative analysis of biomarker diagnostic accuracy for dysplasia in biopsies from AFI+ vs AFI− areas).	All of the biomarkers associated with the presence of confirmed dysplasia, with the exception of 9p LOH (p16). At stricter analysis for HGD/EAC, aneuploidy, p53 immunohistochemistry and cyclin A had the strongest association with dysplasia in the per-biopsy analysis and, as a panel, had an area under the receiver operating characteristic curve of 0.97 (95% CI 0.95 to 0.99) for diagnosing HGD/EC. The diagnostic accuracy for HGD/EC of the three-biomarker panel from AFI+ areas was superior to AFI− areas (p<0.001). Average AUCs in all six databases were significantly higher for the threebiomarker panel assessed on AFI+ areas compared with AFI− areas. The biomarker panel had a sensitivity and a specificity of 95.8% (95% CI 76.9% to 99.8%) and 88.6% (95% CI 79.7% to 94.1%), respectively, for a diagnosis of HGD/EAC. By comparison, the Seattle protocol had similar sensitivities, namely, 95.8% (95% CI 76.9% to 99.8%) for a diagnosis of HGD/EC (p=1.0 when compared with three-biomarker panel). Importantly, using this novel approach 2.8 biopsies per patient were taken on average compared with 12.8 for the standard biopsy protocol (4.5 fold reduction; p<0.001). At validation cohort, the panel had a sensitivity and a specificity of 100% and 85% (95% CI 98.9% to 95.0%), respectively, for a diagnosis of HGD/EAC.	Panel of biomarkers assessed: Aneuploidy/tetraploidy; 9p and 17p loss of heterozygosity; RUNX3, HPP1 and p16 methylation; p53 and cyclin A immunohistochemistry
Allan (2015)	cross-sectional	NA	NA	20	NA	Biopsies of BE from dysplasia and non-dysplatic areas	NA	comparison of dysplastic and non-dysplatic areas in each individual	NA	FISH probes to 9p12 (CDKN2A; p16), 17q11.2-12 (HER2), 8q24.12-13 (CMYC) and 20q13.2 (ZNF217) to compare dysplastic and non-dysplatic areas in each individual	HER2, CMYC and ZNF217 showed significant (all P < 0.0005) increases in copy number in dysplastic mucosa; CDKN2A had an insignificant (P = 0.852) decrease when compared to non-dysplastic mucosa. While aneusomy was strongly predictive of dysplasia, eusomy did not rule it out. Odds ratio for detecting dysplasia: CDKN2A 0.991 (95%CI 0.953–1.031, p= 0.658); HER2 1.359 (95%CI 1.126–1.641,p=0.001); CMYC 1.414 (95%CI 1.145–1.745,p=0.001); ZNF217 1.318 (95%CI 1.111–1.563,p=0.002). HER2 has the greatest area under the curve (AUC) at 0.942 (95%CI 0.874–1), suggesting strong predictive power, followed closely by CMYC at 0.941 (95%CI 0.861–1) and ZNF217 with 0.895 (95%CI 0.796–0.994). CDKN2A (p16) has the smallest area at 0.538 (95%CI 0.349–0.726), suggesting almost no predictive power. While definite aneusomy strongly predicts dysplasia, non-detection of aneusomy with this particular FISH panel does not allow one to conclude that dysplasia is absent, or that the case is necessarily low risk.	Aneusomy detected by FISH to 9p12 (CDKN2A; p16), 17q11.2-12 (HER2), 8q24.12-13 (CMYC) and 20q13.2 (ZNF217)
Timmer (2016)	prospective cohort study	NA	NA	428	60 yrs	Barrett's esophagus without dysplasia (NDBE) at baseline, and length >1 cm, who had a baseline brushing. Patients with histological progression within 6 months were excluded from the final analysis.	Evaluation of six molecular markers by DNA fluorescence in situ hybridisation on brush cytology specimens to study the risk of progression to HGD/EAC	Progressors to HGD/EAC vs. non-progressors. Regular clinical follow up	2019 patient-years (median 45 months per patient; IQR 35–72 months)	Risk of progression to HGD/EAC based on baseline molecular markers. These were measured by DNA FISH using locus-specific probes to p16 (CDKN2A), p53 (TP53), Her-2/neu (ERBB2), 20q, and MYC, and centromeric probes for chromosomes 7 and 17 (Abbott Molecular) to detect aneusomy (aneusomy detected by the centromeric probes for CEP7 and CEP17 was used as a surrogate marker to assess DNA ploidy changes).	An 'Abnormal Marker Count' that counted abnormalities in p16, MYC gain and aneusomy, significantly improved risk prediction beyond using just age and BE length. In multivariate analysis, these three factors (age, Barrett's length, and the markers p16, MYC, and aneusomy) identified a high-risk group with an 8.7-fold (95% CI, 2.6 to 29.8) increased hazard ratio compared with the low-risk group, with an area under the curve of 0.76 (95% CI, 0.66 to 0.86). The specificity of the model was 0.54 meaning that 46% of the non-progressors would be incorrectly classified as high-risk. The PPV of the model was 9%, but its NPV was 99% (meaning that 99% of the patients would be safely classified as low-risk and not progress to high-grade dysplasia or cancer during follow-up.) increased risk with each biomarker: Her-2 gain: 1.02 (0.73-1.44), p16 loss: 1.07 (02-1.12), MYC gain: 1.01 (1.00-1.02), 20q: 1.0 (0.99-1.01), aneusomy: 1.23 (1.06 to 1.43), Abnormal Marker Count: 1.91 (1.29 to 2.81)	p16, MYC, her-2/neu, 20q, Aneusomy (chromosomes 7 & 17). Martinez 2016, Tinner 2016 and Hoefnagel 2020 use the same dataset, only with different focus and different follow up times
Brankley (2016)	Retrospective cohort study of BE patients who underwent endoscopic surveillance/treatment between 04/2003 and 10/2010	NA	NA	245 patients with bening squamous epithelium, BE, BE with LGD and BE with HGD at baseline	mean (standard deviation [SD] age was 66 years (10), ranging from 35 to 90 years	high-risk patients, having a history of biopsy-confirmed HGD without EA	Evaluation of polysomy in patients with HGD but not EAC.	Comparison was to observe the rates of progression to EAC between patients with HGD with polysomy and those without	median 3.6 years (IQR 2–5 years)	Role of polysomy on the risk of progression to EAC. Polysomy was assessed by FISH probes targeting 8q24 (MYC), 9p21 (CDKN2A), 17q12 (ERBB2), and 20q13 (ZNF217). Polysomy was defined as multiple chromosomal gains (displaying ≥ 3 signals for ≥ 2 probes); specimens containing ≥ 4 cells exhibiting polysomy were considered polysomic.	At baseline, 93 (38.0%) had a polysomic FISH result and 152 (62.0%) had a non-polysomic FISH result. Patients with a non-polysomic result had an estimated risk of 1.4% (95% CI: 0% to 3.3%) of developing EA within 3 years (unchanged between years 1–3), as opposed to polysomic patients' risk of 10.2% (95% CI: 3.9% to 16.6%) within 1 year, 14.2% (95% CI: 6.7% to 21.7%) within 2 years, and 20.4% (95% CI: 11.3% to 29.5%) within 3 years. In this high-risk patient cohort, in whom all patients had a prior history of HGD, the negative predictive value of a non-polysomic FISH result is estimated to be 98.6% at 2 years after brushing collection. Further, the positive predictive value of a polysomic FISH result is estimated to be 14.2% within 2 years. Thus, patients with a polysomic FISH result had a significantly higher risk of developing EAC within 2 years (14.2%) compared with patients with a non-polysomic FISH result (1.4%, P < 0.001). There was no association between a polysomic FISH and a non-polysomic FISH result and the stage of cancer (P = 0.57). There was an increased risk of progressing to EA as the severity of the simultaneous histologic diagnosis increased with 2-year risk estimates of 0%, 2.9%, 3.8%, and 10.4% for absence of benign squamous epithelium, IM, LGD, and HGD, respectively. Patients with simultaneous HGD were more likely to develop EAC than those that did not have HGD on biopsy (HR = 4.5, 95% CI: 1.8–11.5, 2-year risk estimates 10.4% vs. 2.6%, P = 0.0005). When considering all polysomy and baseline HGD combinations together, those with both had the highest risk of EA compared with those with neither (HR = 11.8, 95% CI: 3.4–40.9)	polysomy

Study	Design			Sample	Age	Inclusion	Follow-up	Comparison	Follow-up time	Aim / Biomarkers	Results	Biomarker
Poneros (2017)	retrospective	NA	NA	192 specimens (46 non-BE, 42 NDBE, 23 IDBE, 10 LGD, 29 HGD, and 42 EAC)	NA	specimen of different BE grades of dysplasia	NA	NA	NA	FISH analysis using probes 8q24 (MYC), 9p21 (CDKN2A), 17q12 (ERBB2), and 20q13 (ZNF217) to study the progression of BE to HGD/EAC	Polysomy had the highest discriminatory power for most comparisons, with an AUC of .83 in discriminating EA/HGD from all other diagnoses. If at least 10% cells in a specimen were positive for polysomy, we were able to obtain a sensitivity and specificity of 88 and 75% for EA versus rest and a sensitivity and specificity of 80 and 88% for EA/HGD versus rest, respectively. When LGD and IGD were added, specificity increased to 96% but sensitivity decreased to 74%. If at least 10% cells in a specimen were positive for polysomy, we were able to obtain a sensitivity and specificity of 88 and 75% for EA versus rest and a sensitivity and specificity of 80 and 88% for EA/HGD versus rest, respectively. When LGD and IGD were added, specificity increased to 96% but sensitivity decreased to 74%.	polysomy
Choi (2018)	retrospective	NA	NA	80 FFPE BO samples with high-grade dysplasia (HGD), 38 LGD, 21 IND and 14 negative for dysplasia (ND).	64 years (range 35–87)	patients with BE with or without dysplasia	regular clinical follow up	evaluate HGD/EAC vs others, and evaluate utility of DNA alterations for progressions	33 months (range 0–170)	DNA flow cytometry as a diagnostic marker of dysplasia and facilitate risk stratification of low-grade dysplasia (LGD) and indefinite for dysplasia (IND) patients	DNA content abnormality was identified in 76 HGD (95%), 8 LGD (21.1%), 2 IND (9.5%) and 0 ND samples. As a diagnostic marker of HGD, sensitivity was 95%, specificity 85%, 90% positive predictive value (PPV) and 92% negative predictive value (NPV). One-year, 4-year, 5-year and 12-year detection rates of HGD or EAC for LGD patients with DNA content abnormality were 85.4% (p<0.001), 85.4% (p=0.003), 100% (p<0.001) and 100% (p<0.001), respectively, whereas LGD patients in the setting of normal DNA content had 1-year, 4-year, 5-year and 12-year detection rates of 13.5%, 30.8%, 30.8% and 30.8%, respectively. All IND patients with DNA content abnormality were subsequently found to have HGD or OAC within 2 years (p=0.001), whereas 1-year, 2-year and 13-year detection rates of HGD or OAC in the setting of normal DNA content remained stable at 5.9%. For patients with DNA content abnormality detected at baseline LGD or IND, the univariate HRs for subsequent detection of HGD or EAC were 7.0 (95%CI 2.184 to 24.073) and 20.0 (95% CI 1.872 to 436.409), respectively (p =<0.001).	aneuploidy/tetraploidy as a biomarker of presence and progression towards HGD
Duits (2018)	nested case-control study	NA	NA	260 (130 cases-progressos vs 130 controls non-progressors)	Progressors:59.9 ± 9.6; non-progressors 59.2 ± 9.7	Patients with biopsy-proven Barrett's esophagus with NDBE, IND or LGD. If someone developed HGD/EAC within 24 months, they were excluded.	Routine clinical follow up-surveillance.	Comparison of progressors vs. non-progressors	progressors 3.7 vs. non-progressors 4.8 years; P < 0.001	Porgression to HGD/EAC. We assessed abnormal DNA content, p53, Cyclin A, and Aspergillus oryzae lectin (AOL) in FFPE sections.	LGD, p53 and AOL were independently associated with neoplastic progression. In the multivariate analysis, expert LGD had a 36-fold increased odds of progression (OR, 35.7; 95% CI, 1.4–920.8). Abnormal p53 expression (OR, 4.1; 95% CI, 1.4–12.4) and abnormal AOL expression in three epithelial compartments had a 4-fold increased odds of progression (OR, 4.3; 95% CI, 0.7–26.3). Cyclin A did not predict progression, and DNA ploidy analysis by image cytometry was unsuccessful in the majority of cases, both were excluded from the multivariate analysis. The multivariable biomarker model had an area under the receiver operating characteristic curve (AUC) of 0.73.	AOL, and p53, cyclin A, DNA ploidy. Utilized the Amsterdam-based ReBus nested case-control cohort, a multicenter prospective cohort study
Stachler (2018)	retrospective case–control study	NA	NA	97 (24 cases-progressors and 73 controls - non progressors)	cases 67.5 vs. controls 61.6 years	Patients with biopsy-proven Barrett's esophagus with NDBE, IND or LGD. If someone developed HGD/EAC within 12 months, they were excluded.	Routine clinical follow up-surveillance.	Comparison of progressors vs. non-progressors	mean follow-up (controls 6.7 vs. cases 3.3 years, P<0.001)	Porgression to HGD/EAC. From each patient, we selected a single tissue sample obtained more than 1 year before progression (cases) or more than 2 years before the end of follow up (controls). Pathogenic mutations, gene copy numbers, and ploidy were compared between samples from progressors and non-progressors.	TP53 mutations were detected in 46% of samples from progressors and 5% of non-progressors. In this case–control sample set, TP53 mutations in BE tissues increased the adjusted risk of progression 13.8-fold (95% CI, 3.2–61.0) (P<.001). These results were confirmed in a separate validation set of 16 NDBE who progressed and 28 NDBE who did not progress. We did not observe significant differences in ploidy or copy number profile between groups. We identified 147 pathogenic mutations in 57 distinct genes—the average number of pathogenic mutations was higher in samples from progressors (2.5) than non-progressors (1.2) (P<.001). TP53 and other somatic mutations were recurrently detected in samples with limited copy number changes (aneuploidy). Beyond TP53, several known GEA tumor suppressors and oncogenes (ARID1B, APC, ERBB2, RB1, RUNX1, LARP4B, and BIRC5) had more frequent mutations in progressors. Among these, ARID1B, APC, and ERBB2 were significantly enriched in the progressors (P<0.05).	no difference in ploidy between progressors and non-progressors; TP53 abnormalities, ARID1B mutation, APC mutatio, ERBB2 mutation
Hadjinicolaou (2020)	Prospective multicenter study	NA	NA	127 (42 progressors, 85 non-progressors)	median age of 65.6 years (IQR, 13.7 yrs)	Age ≥18 yrs, BE with a length of at least C ≥ 2 or C<2M≥4 according to the Prague classification with or without visible lesions. Patients had targeted biopsies during endoscopy with autofluorescence imaging (AFI).	The primary endpoint of this study was progression from NDBO/ID to any grade of dysplasia. The two secondary endpoints were a) progression from NDBO/ID to HGD/OAC, and b) any histologic progression i.e. NDBO/ID to LGD, NDBO/ID to HGD, and LGD to HGD.	Comparisons of biomarkers between progressors and non-progressors. This was a validation cohort from the cross-sectional study described above (di Pietro 2015)	Median 4.6 yrs (IQR, 4.3 yrs). Progressors: 1.2 yrs (IQR, 2.7 yrs) until progression	Evaluation of previously defined 9-molecular biomarker panel (in the study of di Pietro 2015 above) with progression in BE patients. Histological progression was defined as transition from a NDBO or indefinite for dysplasia (ID) to any dysplasia, or if low-grade dysplasia already present, to a higher grade of dysplasia or cancer. Biomarkers: p53 and cyclin A were analysed by immunohistochemistry (IHC); aneuploidy and G2/tetraploidy, were analysed by flow cytometry; p16, RUNX3 and HPP1 hypermethylation was analysed by quantitative methylation-specific PCR (Methylight); and LOH at 9p and 17p loci was analysed by the use of microsatellite markers. Snap frozen biopsies in DMSO were used for aneuploidy, G2 tetraploidy, LOH markers and methylation assays.	Amongst progressors, there were 12 (28.6%) that progressed from NDBO/ID to LGD, 16 (38.1%) that progressed from NDBO/ID to HGD/OAC and 14 (33.3%) that progressed from LGD to HGD/OAC. Of the nine molecular biomarkers, at multivariable analysis p53 and aneuploidy were the only significant predictors of any progression. However, when we excluded patients with progression within 12 months of follow up (prevalent dysplasia), only aneuploidy retained statistical significance. The presence of positive aneuploidy at index endoscopy led to a 6.6-fold higher risk of dysplastic progression over no progression (95% CI: 1.8 24.8, p = 0.005). However the sensitivity of the test to predict progression was low (32%, 95% CI: 16, 52%). ROC analysis showed that a clinical model using patient age and BO length (AUC=0.55; CI: 0.45,0.66) was outperformed in the prediction of any histologic progression by a molecular biomarker model comprising of aneuploidy and p53 with a cut-off of one positive biomarker out of two (AUC=0.68; CI: 0.59,0.77). ROC analysis showed that a model with aneuploidy as the only predictor of dysplastic progression outperformed the clinical model (AUC=0.63; CI: 0.54 0.72). p53 appeared to correlate more with short-term progression. Patients with aberrant p53 expression at index endoscopy had an odds ratio of 6.0 (95% CI: 3.1, 11.2, p = 0.007) of missed dysplasia on endoscopic biopsies.	p53, cyclin A;; aneuploidy and G2/tetraploidy; p16, RUNX3 and HPP1 hypermethylation; LOH at 9p and 17p loci; aneuploidy, G2 tetraploidy

Study	Study design			Sample	Age	Population	Intervention	Comparison	Follow-up	Outcome	Results	Theme
Killcoyne (2020)	retrospective demographically matched case-control study	NA	NA	777 biopsies, from 88 patients with BE. Independent validation cohort of 76 and 248 patients	NA	patients with BE in surveillance	regular clinical follow up	compare progressors to non-progressors	NA	Shallow whole genome sequencing on 777 biopsies, sampled from 88 patients in BE surveillance over a period of up to 15 years to study genome-wide CN instability (Copy number (CN) alterations) as a marker for risk of progression to HGD or intramucosal cancer (IMC). These findings are validated on two independent cohorts of 76 and 248 patients.	Risk classifications was based on the enrichment of samples from progressor or non-progressor patients to maximize the sensitivity of classes: 'low' (Pr≤0.3; sensitivity=0.87, specificity=0.65), 'moderate' (0.3>Pr<0.5), or 'high' (Pr≥0.5, sensitivity=0.72, specificity=0.82). Aggregating predictions either per-endoscopy (mean or max sample predictions) or per-patient (mean or max predictions excluding HGD/IMC samples) did not measurably increase the prognostic accuracy, suggesting that a single sample (e.g. pooled 4-quadrant biopsy) may be sufficient for prediction which could be ideal for clinical application. The model was then used to predict and classify risks per-sample for the validation cohort (76 patients, 213 samples). 78/142 (55%) samples from non-progressor patients were classified as low risk, and 55/71 (77%) of samples from patients who progressed were classified as high risk. As in the discovery cohort, high risk classification of progressor patient samples was largely independent of histopathology. Similarly, when we used our model to classify the historical Seattle study patient dataset (n=248, samples=1273 SNP array) we again find that samples from progressors are classified as high risk regardless of pathology. Most progressive patient samples are classed as high risk throughout the disease history, while non-progressive patient samples are consistently low risk. For patients that progress, 50% (8/16) of endoscopies had at least one sample classified as high risk 8 or more years prior to transformation. Cases which lack early CN patterns of progression acquired these over the following years, leading to 78% (18/23) of endoscopies with at least one high risk sample one to two years prior to HGD/IMC diagnosis. With each increment in the number of patients the predictive accuracy of the model increased, reaching a (cross validated) AUC of 0.89 (specificity=0.83, sensitivity=0.82) when combining all discovery and validation patients.	evaluation of copy number (CN) instability as a marker of progression to HGD/EAC
Douville (2021)	cross-sectional study	NA	NA	79 (training set) and 268 (validation set)	NA	Brushng obtained from patients without BE, with non-dysplastic BE (NDBE), low-grade dysplasia (LGD), high grade dysplasia (HGD), or adenocarcinoma (EAC)	Assessment of aneuploidy from a single esophageal brushing that widely sampled the esophagus in these different types of patients (training and validation cohorts)	NA	NA	Presence of aneuploidy. To assess aneuploidy, we employed RealSeqS, a technique that uses a single primer pair to interrogate ~350,000 genome-spanning regions and identify specific chromosome arm alterations. A classifier to distinguish NDBE from EAC was trained on results from 79 patients. An independent validation cohort of 268 subjects was used to test the classifier at distinguishing patients at successive phases of BE progression.	Aneuploidy progression was associated with gains of 1q, 12p, and 20q and losses on 9p and 17p. The entire chromosome 8q was often gained in NDBE, whereas focal gain of 8q24 was identified only when there was dysplasia. Based on the Global Aneuploidy Score (GAS) score and this panel of six specific chromosomal alterations, we developed a simple decision tree classifier, termed BAD (Barrett's Aneuploidy Decision), for distinguishing stages of BE progression. BAD sorted samples into three categories. Not-BAD cases had GAS≤0.6, indicating relative non-aneuploidy. Maybe-BAD cases had GAS>0.6 but none of the six specific chromosome alterations, possibly indicating a greater potential risk of progression. Very-BAD cases had GAS>0.6 and losses of 9p or 20q, gains of 1q, 12p, or 20q, or a focal gain of 8q24. The BAD classification system, which used both specific chromosome changes plus GAS scores, outperformed GAS scores alone. In particular, compared to the aneuploidy classification of GAS, the Very-BAD classification markedly improved both the specificity for rejecting NDBE and the positive predictive value (PPV) for identifying HGD plus EAC cases. These results were confirmed at validation set: the AUCs were 0.86 and 0.87 in the Training and Validation Sets, respectively. samples from the patients with NDBE in the Validation Set exhibited the same bimodal distribution as observed in the Training Set, with the GAS threshold of 0.60 again cleanly separating Validation Set NDBE into two populations, 36.6% as aneuploid and 63.4% as non-aneuploid. When the BAD classifier was applied to the Validation Set, it again was more accurate than GAS alone. Specifically, 96.4% of EAC and 67.9% of HGD were classified as Very-BAD. In contrast, only 7.3% of NDBE were classified as Very-BAD. Among Validation Set NDBE cases, 7.3% were classified as Very-BAD, 29.3% as Maybe-BAD, and 63.4% as Not-BAD. In the Validation Set, as in the Training Set, the BAD classification system outperformed GAS scores alone. Tthe total number of chromosome arms lost or gained steadily increased during disease progression (6 for aneuploid NDBE, 10 for aneuploid LGD, 17 for aneuploid HGD, and 22 for aneuploid EAC). The BAD classification of DNA from esophageal brushings is highly correlated with histopathologic classification of the same patients.	aneuploidy; use of RealSeqS technique
Bowman (2021)	retrospective	NA	NA	82 formalin-fixed paraffin-embedded samples from the 45 patients, including 78 HGD/IMC, 2 LGD, and 2 IND. Eight non-dysplastic BE samples were used as controls.	67 years (range 42–89)	patients with IDN, LGD or HGD/IMC who underwent endoscopic therapy	standard clinical practice according to ACG 2016 guidelines	standard clinical practice according to ACG 2016 guidelines	mean 16 months (range 1 month-9 years)	Predictors of persistent or recurrence of HGD/EAC after endoscopic resection. DNA flow cytometry was one of the predictors	Sixty (73%) of the 82 specimens showed abnormal DNA content (aneuploidy or elevated 4N fraction). These were all specimens with HGD/IMC (representing 77% of that group). Of these, 42 (70%) were associated with subsequent development of persistent/recurrent HGD/IMC (n = 41) or esophageal adenocarcinoma (EAC; n = 1). In contrast, only 6 (27%, all HGD/IMC) of the 22 remaining samples (all with normal DNA content) were associated with persistent/recurrent HGD/IMC. Abnormal DNA content (HR = 6.0, 95%CI 1.8–19.6; p = 0.003) and treatment with EMR alone (HR = 2.7, 95% CI 1.0–7.0; p = 0.047) remained as significant risk factors for treatment failure in a multivariate analysis. DNA abnormality also serves as a diagnostic marker of HGD/IMC with an estimated sensitivity of 77%.	aneuploidy/tetraploidy may predict prro response to therapy

Study	Study type	Randomization	Blinding	N	Age	Inclusion criteria	Intervention	Comparator		Outcomes	Results	Biomarker/Conclusion
Vithayathil (2022)	Multicenter randomized crossover study	Block-randomized using computer generated randomization in blocks of 4	Endoscopists could not be blinded to the intervention arm but were blinded to the endoscopy and histology results of the pretrial endoscopy and other study arm.	154 recruited and randomized; 134 completed both arms of the study	67.3 (38.0–89.0)	≥18 years, BE >C2 and/or M3 on pretrial endoscopy referred for surveillance of nondysplastic BE (NDBE) or assessment of flat dysplasia	AFI-directed pCLE and targeted biopsies for molecular biomarkers (experimental arm).	Patients randomized to receive either HRWLE with Seattle protocol biopsies (standard arm) or endoscopy with AFI-directed pCLE and targeted biopsies for molecular biomarkers (experimental arm). Patients crossed over to the other arm after 6 to 12 weeks.	NA	The primary outcome was the diagnostic accuracy for dysplasia of AFI-guided pCLE using the trial histology as the gold standard. Secondary outcomes included the following: (1) diagnostic accuracy of AFI-guided pCLE for dysplasia with reference to the overall histology, which included biopsy specimens taken within 12 months before enrollment; (2) added diagnostic value of molecular biomarkers; (3) time to perform the endoscopy; and (4) patient-reported experience related to experimental and standard endoscopy. A 3-biomarker panel including cyclin A, p53 (assessed with immunohistochemistry), and aneuploidy (assessed with image cytometry) was used.	Within the experimental arm, pCLE had a higher sensitivity than HRWLE for HGD/IMC (P =.046) and all grades of dysplasia (P = .01), but lower specificity (HGD/IMC, P = .01; all grades of dysplasia, P = .02). In per-patient analysis, there was no difference in the sensitivity of pCLE compared with HRWLE with the Seattle protocol (76.5%; 95% CI, 50.1–93.2 vs 76.5%; 95% CI, 50.1–93.2, respectively; P = 1.00 for HGD/IMC; 74.3%; 95% CI, 56.7–87.5 vs 80.0%; 95% CI, 63.1–91.6, respectively; P = .48, for all grades of dysplasia). The use of AFI-targeted pCLE led to 2.1 optical biopsies per patient on average compared with 12.3 tissue biopsy specimens taken in the Seattle protocol. Standard endoscopy missed 28 cases of dysplasia (miss rate, 51.9%), 11 of which were detected by experimental endoscopy. Experimental endoscopy missed 20 dysplastic cases (miss rate, 37%), of which 5 were diagnosed correctly by standard endoscopy. In the overall histology analysis, AFI-guided pCLE had a higher sensitivity for HGD/IMC than Seattle protocol biopsies (73.3%; 95% CI, 54.1–87.7 vs 43.3%; 95% CI, 25.5–62.6, respectively; P = .02). The difference in sensitivity for all grades of dysplasia was not statistically significant (63.0%; 95% CI, 48.7–75.7 vs 51.9%; 95% CI, 37.8–65.7, respectively; P = .13). In the per-patient analysis the sensitivity and specificity for dysplasia of individual biomarkers were 48.6% and 93.9% for p53, 47.1% and 69.4% for cyclin A, and 40.0% and 88.5% for aneuploidy, respectively. At multivariable analysis, p53, aneuploidy, and optical dysplasia correlated significantly with a diagnosis of dysplasia. A panel comprising these 3 biomarkers showed an area under the receiver operating curve of 0.83 (95% CI, 0.76–0.91) for a diagnosis of any grade of dysplasia and 0.88 (95% CI, 0.78–0.97) for a diagnosis of HGD/IMC. Using a threshold of 1 positive biomarker, this panel had a higher sensitivity than the Seattle protocol in detecting dysplasia in the overall histology analysis (81.5% vs 51.9%; P < .001) The difference was not statistically significant in the trial histology analysis (91.4% vs 80.0%; P = .16).	3-biomarker panel including cyclin A, p53, and aneuploidy. The addition of molecular biomarkers could improve diagnostic accuracy

Others: these are either other types of biomarkers that are not as established as the ones above, or combinations of biomarkers

Mucins

Study	Study type	Randomization	Blinding	N	Age	Inclusion criteria	Intervention	Comparator		Outcomes	Results	Biomarker/Conclusion
Niv (2018)	meta-analysis	NA	NA	1020 patients and 2102 controls		Observational studies describing mucin expression in esophageal normal mucosa, premalignant, and malignant states were included. Only studies that clearly included data on esophageal lesions compared with normal esophageal tissue. In all the studies, biopsies were taken because of GERD and for surveillance of BE	NA	NA	NA	Difference in mucin expression between normal esophagus, BE with or without dysplasia, and EAC.	In the total 5 studies meta-analyzed (45 sets of data), mucin expression was significantly higher in esophageal lesions than in normal squamous esophageal mucosa (OR, 5.456; 95% CI, 1.883-15.807; P=0.002). ORs for mucin expression in Barrett's mucosa without dysplasia, LGD, HGD, and EAC (each in the fixed model) compared with normal controls were 5.297 (95% CI, 2.277-12.321, P<0.0001; 2.671 (0.889-8.028), P=0.080; 5.787 (2.427-13.801), P<0.0001 and 14.696 (7.003-30.842), P<0.0001, respectively. ORs for total expression in esophageal lesions of MUC1, MUC2, MUC3, MUC 4, MUC5AC, MUC6, and MUC 16 compared with normal esophageal controls were 0.001(95% CI, 0.000-0.003), P<0.0001; 87.177 (24.533-309.784), P<0.0001; 42.566 (11.005-164.634), P<0.0001; 0.010 (0.003-0.039), P<0.0001; 68.73 (21.826-216.489), P<0.0001; 38.928 (12.700-119.323), P<0.0001; 13.650 (7.292-25.551), P<0.0001, respectively. There is a gradient of mucin expression and complexity in esophageal premalignant to malignant lesions, lower in Barrett's mucosa with low grade dysplasia (LGD), increased in high grade dysplasia (HGD), and highest in esophageal adenocarcinoma (EAC). MUC2, MUC3, MUC5AC, and MUC6 expression was higher in EAC than HGD, and higher in HGD than in LGD mucosa. The opposite was found for MUC1 and MUC4.	Mucins expressed on biopsies of normal esophagus, BE with or without dysplasia, and EAC (MUC1, MUC2, MUC3, MUC4, MUC5AC, MUC6, MUC7, MUC16). Meta-analysis. Included studies until October 30, 2016
Butt (2017)	cross-sectional study	NA	NA	NA	NA	biopsies of Barrett's esophagus with or without dysplasia/EAC	NA	NA	NA	Mucin glycoprotein 1 (MUC1) expression in BE and EAC using gene set enrichment analysis (GSEA); development of a therapy (photoimmunotherapy) targeting MUC1	MUC1 was present in 21% and 30% of significantly enriched pathways comparing BE and EAC to squamous epithelium, respectively. MUC1 gene expression was 2.3 times compared to normal squamous epithelium (p=<0.001) and 2.2 times higher in EAC (p=0.03). MUC1 immunohistochemical expression increased progressively during progression to LGD, HGD and EA and followed tumor invasion.	Mucin glycoprotein 1 (MUC1); development of a therapy (photoimmunotherapy) targeting MUC1
Ozcan (2018)	retrospective	NA	NA	72 pt with biopsies from squamous epithelium, or BE with or without dysplasia/EAC	NA	biopsies from squamous epithelium, or BE with or without dysplasia/EAC	evaluation of mucinous markers in biopsies from squamous epithelium/BE with or without dysplasia/EAC	evaluation of mucinous markers in biopsies from squamous epithelium/BE with or without dysplasia/EAC	NA	evaluation of mucinous markers in biopsies from squamous epithelium/BE with or without dysplasia/EAC	Multilayer squamous epithelium showed only MUC-1 positivity in the EAC group (p<0.05) while MUC-2 and MUC-5AC staining could not be detected in this group. MUC2 and MUC5AC expressions were prevalent in the surface columnar lined esophageal epithelium (P < 0.05 for both). Strong and diffused membranous or cytoplasmic staining of CK7 was observed at squamous, ductal, surface columnar and/or glandular epithelium. c-p27 staining was diffused and moderate in the cellular membranes observed in all groups except for esophageal epithelial metaplasia without intestinal metaplasia. Additionally, weakly focal cytoplasmic staining in squamous epithelium of p27 in EAC was detected.	MUC-1, MUC-2, MUC-5AC, CK7, and cytoplasmic p27. this is mostly a descriptive study

cyclin A

Janmaat (2017)	meta-analysis	NA	NA	16 different IHC biomarkers were studied in 36 studies. These studies included 425 cases and 1835 controls	NA (different studies included for different biomarkers)	criteria: (1) association between IHC biomarker expression on FFPE material and risk of neoplastic progression was assessed; (2) a cohort or case-control study design; (3) patients with known or newly diagnosed BE with or without LGD at baseline; (4) patients defined as cases had to have progressed to either HGD or EAC during follow-up; (5) mean follow-up of at least one year from the time of initial BE diagnosis; (6) the possibility to extract an OR.	NA	NA	follow up time was 51 months versus 59 months for cases versus controls, respectively	Progression to of BE to HGD/EAC. Comparing biomarkers (p53, aspergillus oryzae lectin (AOL), Cyclin A, Cyclin D and alpha-methylacyl-CoA recemas) in predicting neoplastic progression of BE to HGD/EAC.	Meta-analyses were possible for p53, AOL, Cyclin A, Cyclin D, and alpha-methylacyl-CoA racemase (AMACR), which were studied 13, 2, 4, 3, and 2 times respectively. Aberrant p53 expression was significantly associated with an increased risk of neoplastic progression with an OR of 3.18 (95% CI 1.68 to 6.03). The overall OR, for aberrant p53 IHC on neoplastic progression, after stratification for histology, was 3.86 (95% CI 2.03 to 7.33).This association was confirmed for both non-dysplastic BE (6.12; 95% CI 2.99 to 12.52) and BE with LGD (OR 8.64; 95% CI 3.62 to 20.62). Another promising biomarker to predict neoplastic progression was AOL, with an OR of 3.04 (95% CI 2.05 to 4.49). Cyclin A showed a tendency towards increased risk (OR 1.90; 95% CI 0.85 to 4.22). Cyclin D did not show a significant association (OR 1.01; 95% CI 0.14 to 7.03). Alpha-methylacyl-CoA racemase had an OR 4.07 (95% CI 0.66 to 25.12). The following IHC biomarkers were investigated only once up to this meta-analysis: β-catenin, CD1a, COX2, HER2, Ki67, Lewis, Mcm2, Sialyl Lewis, SOX2, and WGA	p53, aspergillus oryzae lectin (AOL), Cyclin A, Cyclin D and alpha-methylacyl-CoA racemase. Meta-analysis. Included studies until 09/2016
Lao-Sirieix (2007)	case-control study	NA	NA	48 (16 cases, 32 controls)	64.4±0.6	Patients with biopsy-proven Barrett's esophagus	Routine clinical follow up-surveillance.	Progressors to HGD/EAC vs. non-progressors. Regular clinical follow up	6.5 years for cases, 5.5 years for controls	Progression from BE to HGD/EAC. Cyclin A was measured from brushing	In the case-control cohort, patients with biopsies expressing cyclin A at the surface were more likely to progress to adenocarcinoma than those who did not (odds ratio, 7.5; 95% confidence interval, 1.8-30.7). The sensitivity and specificity of cyclin A expression in brushings for the detection of high-grade dysplasia and cancer patients were 97.8% and 58.7%, respectively. The associated NPV was 97.4%. In other words,if a brushing sample is negative for cyclin A,it is 97.4% likely that the patient does not have high-grade dysplasia or adenocarcinoma. To differentiate between dysplastic and nondysplastic samples,the sensitivity and specificity are 88.3% and 64.2%,respectively, with a NPV of 81.6%	cyclin A
di Pietro (2015)	cross-sectional study	NA	NA	157 (training cohort), 46 (validation cohort)	66.4 (training cohort), 68.7 (validation cohort)	Age ≥18 yrs, BE with a length of at least C ≥ 2 or C<2M≥4 according to the Prague classification with or without visible lesions.	Biopsies were performed either according to the standard Seattle protocol or under the guidance of autofluorescence imaging (AFI)	Comparison of the accuracy of a panel of molecular biomarkers on AFI-directed biopsies with conventional Seattle protocol biopsies for HGD and EAC. Secondary aims: (i) assessment of diagnostic accuracy for the biomarkers for any grade of dysplasia and (ii) validation of a large panel of biomarkers in an independent prospective study by an independent laboratory.	NA	Accuracy of biomarkers in identifying HGD/EAC. Panel of biomarkers assessed: Aneuploidy/tetraploidy; 9p and 17p loss of heterozygosity; RUNX3, HPP1 and p16 methylation; p53 and cyclin A immunohistochemistry. The statistical analysis consisted of three stages: (1) per-biopsy analysis (correlation between biomarkers and histological outcome in individual targeted areas); (2) per-patient analysis (correlation between overall biomarker result and overall histological outcome in individual patients); and (3) comparison between AFI+ versus AFI− areas (comparative analysis of biomarker diagnostic accuracy for dysplasia in biopsies from AFI+ vs AFI− areas).	All of the biomarkers associated with the presence of confirmed dysplasia, with the exception of 9p LOH (p16). At stricter analysis for HGD/EAC, aneuploidy, p53 immunohistochemistry and cyclin A had the strongest association with dysplasia in the per-biopsy analysis and, as a panel, had an area under the receiver operating characteristic curve of 0.97 (95% CI 0.95 to 0.99) for diagnosing HGD/EC. The diagnostic accuracy for HGD/EC of the three-biomarker panel from AFI+ areas was superior to AFI− areas (p<0.001). Average AUCs in all six databases were significantly higher for the threebiomarker panel assessed on AFI+ areas compared with AFI− areas. The biomarker panel had a sensitivity and a specificity of 95.8% (95% CI 76.9% to 99.8%) and 88.6% (95% CI 79.7% to 94.1%), respectively, for a diagnosis of HGD/EAC. By comparison, the Seattle protocol had similar sensitivities, namely, 95.8% (95% CI 76.9% to 99.8%) for a diagnosis of HGD/EC (p=1.0 when compared with three-biomarker panel). Importantly, using this novel approach 2.8 biopsies per patient were taken on average compared with 12.8 for the standard biopsy protocol (4.5 fold reduction; p<0.001). At validation cohort, the panel had a sensitivity and a specificity of 100% and 85% (95% CI 98.9% to 95.0%), respectively, for a diagnosis of HGD/EAC.	Panel of biomarkers assessed: Aneuploidy/tetraploidy; 9p and 17p loss of heterozygosity; RUNX3, HPP1 and p16 methylation; p53 and cyclin A immunohistochemistry
van Olphen (2016)	case-control study within a prospective cohort	NA	NA	625 (50 cases-progressors vs. 575 controls-non progressors)	median age of 60 years (interquartile range (IQR) 53–69))	Patients with biopsy-proven Barrett's esophagus and no HGD/EAC within 9 months	Routine clinical follow up-surveillance.	Progressors to HGD/EAC vs. non-progressors. Regular clinical follow up	median duration of 6.7 years (IQR 5.0–7.4). Progressors: median follow-up of 3.2 years (IQR 1.9–5.3).	Progression from BE to HGD/EAC. Cyclin A expression was determined by immunohistochemistry in biopsies of 625 patients; these results were combined with the histological diagnosis and our previous p53, AMACR, and SOX2 data in loglinear regression models. Differences in discriminatory ability were quantified as changes in area under the ROC curve (AUC) for predicting neoplastic progression.	Cyclin A surface positivity significantly increased throughout the metaplasia–dysplasia–carcinoma sequences and was seen in 10% (107/1050) of biopsy series without dysplasia, 33% (109/335) in LGD, and 69% (34/50) in HGD/EAC (p<0.001). Positive cyclin A expression was associated with an increased risk of neoplastic progression (adjusted relative risk (RRa) 2.4; 95% CI: 1.7–3.4), and was particularly seen in biopsy series with LGD (adjusted RR of 5.8; 95% CI: 3.7–9.0). In per-biopsy analysis, cyclin A had an AUC of 0.59 (95% CI: 0.54–0.64) for predicting neoplastic progression with a sensitivity of 32%, a specificity of 86%, a PPV of 21%, and a NPV of 92%. Increases in AUC were substantial for P53 (+0.05), smaller for SOX2 (+0.014), minor for cyclin A (+0.003), and none for AMARC (0.00). There is a challenge in interpretating cyclin A immunohistochemistry with a moderate interobserver agreement with a kappa value of 0.46. Area under the curve (AUC) for predicting neoplastic progression was calculated (pathological diagnosis grade of dysplasia AUC of 0.62 (95% CI: 0.58–0.68), pathological diagnosis+p53 and SOX2 immunohistochemistry AUC of 0.72 (95% CI: 0.67–0.77) and pathological diagnosis+p53, SOX2, and cyclin A immunohistochemistry AUC of 0.72 (95% CI: 0.67–0.77)).	LGD, SOX3, Cyclin A and P53
Duits (2018)	nested case-control study	NA	NA	260 (130 cases-progressos vs 130 controls non-progressors)	Progressors:59.9 ± 9.6; non-progressors 59.2 ± 9.7	Patients with biopsy-proven Barrett's esophagus with NDBE, IND or LGD. If someone developed HGD/EAC within 24 months, they were excluded.	Routine clinical follow up-surveillance.	Comparison of progressors vs. non-progressors	progressors 3.7 vs. non-progressors 4.8 years; P < 0.001	Porgression to HGD/EAC. We assessed abnormal DNA content, p53, Cyclin A, and Aspergillus oryzae lectin (AOL) in FFPE sections.	LGD, p53 and AOL were independently associated with neoplastic progression. In the multivariate analysis, expert LGD had a 36-fold increased odds of progression (OR, 35.7; 95% CI, 1.4–920.8). Abnormal p53 expression (OR, 4.1; 95% CI, 1.4–12.4) and abnormal AOL expression in three epithelial compartments had a 4-fold increased odds of progression (OR, 4.3; 95% CI, 0.7–26.3). Cyclin A did not predict progression, and DNA ploidy analysis by image cytometry was unsuccessful in the majority of cases, both were excluded from the multivariate analysis. The multivariable biomarker model had an area under the receiver operating characteristic curve of 0.73.	AOL, and p53, cyclin A, DNA ploidy. Utilized the Amsterdam-based ReBus nested case-control cohort, a multicenter prospective cohort study

Study	Design	Randomization	Blinding	Population/N	Age	Inclusion criteria	Primary endpoint	Secondary endpoint	Follow-up	Objective	Results	Biomarkers
Hadjinicolaou (2020)	Prospective multicenter study	NA	NA	127 (42 progressors, 85 non-progressors)	median age of 65.6 years (IQR, 13.7 yrs)	Age ≥18 yrs, BE with a length of at least C ≥ 2 or C<2M≥4 according to the Prague classification with or without visible lesions. Patients had targeted biopsies dureing endoscopy with autofluorescence imaging (AFI).	The primary endpoint of this study was progression from NDBO/ID to any grade of dysplasia. The two secondary endpoints were a) progression from NDBO/ID to HGD/OAC, and b) any histologic progression i.e. NDBO/ID to LGD, NDBO/ID to HGD, and LGD to HGD.	Comparisons of biomarkers between progressors and non-progressors. This was a validation cohort from the cross-sectional study described above (di Pietro 2015)	Median 4.6 yrs (IQR, 4.3 yrs). Progressors: 1.2 yrs (IQR, 2.7 yrs) until progression	Evaluation of previously defined 9-molecular biomarker panel (in the study of di Pietro 2015 above) with progression in BE patients. Histological progression was defined as transition from a NDBO or indefinite for dysplasia (ID) to any dysplasia, or if low-grade dysplasia already present, to a higher grade of dysplasia or cancer. Biomarkers: p53 and cyclin A were analysed by immunohistochemistry (IHC); aneuploidy and G2/tetraploidy, were analysed by flow cytometry; p16, RUNX3 and HPP1 hypermethylation was analysed by quantitative methylation-specific PCR (Methylight); and LOH at 9p and 17p loci was analysed by the use of microsatellite markers. Snap frozen biopsies in DMSO were used for aneuploidy, G2 tetraploidy, LOH markers and methylation assays.	Amongst progressors, there were 12 (28.6%) that progressed from NDBO/ID to LGD, 16 (38.1%) that progressed from NDBO/ID to HGD/OAC and 14 (33.3%) that progressed from LGD to HGD/OAC. Of the nine molecular biomarkers, at multivariable analysis p53 and aneuploidy were the only significant predictors of any progression. However, when we excluded patients with progression within 12 months of follow up (prevalent dysplasia), only aneuploidy retained statistical significance. The presence of positive aneuploidy at index endoscopy led to a 6.6-fold higher risk of dysplastic progression over no progression (95% CI: 1.8 24.8, p = 0.005). However the sensitivity of the test to predict progression was low (32%, 95% CI: 16, 52%). ROC analysis showed that a clinical model using patient age and BO length (AUC=0.55; CI: 0.45,0.66) was outperformed in the prediction of any histologic progression by a molecular biomarker model comprising of aneuploidy and p53 with a cut-off of one positive biomarker out of two (AUC=0.68; CI: 0.59,0.77). ROC analysis showed that a model with aneuploidy as the only predictor of dysplastic progression outperformed the clinical model (AUC=0.63; CI: 0.54 0.72). p53 appeared to correlate more with short-term progression. Patients with aberrant p53 expression at index endoscopy had an odds ratio of 6.0 (95% CI: 3.1, 11.2, p = 0.007) of missed dysplasia on endoscopic biopsies. Cyclin A was a negative confounder of p53	p53, cyclin A;; aneuploidy and G2/tetraploidy; p16, RUNX3 and HPP1 hypermethylation; LOH at 9p and 17p loci; aneuploidy, G2 tetraploidy
Vithayathil (2022)	Multicenter randomized crossover study	Block-randomized using computer generated randomization in blocks of 4	Endoscopists could not be blinded to the intervention arm but were blinded to the endoscopy and histology results of the pretrial endoscopy and other study arm.	154 recruited and randomized; 134 completed both arms of the study	67.3 (38.0–89.0)	≥18 years, BE >C2 and/or M3 on pretrial endoscopy referred for surveillance of nondysplastic BE (NDBE) or assessment of flat dysplasia	AFI-directed pCLE and targeted biopsies for molecular biomarkers (experimental arm).	Patients randomized to receive either high-resolution white-light endoscopy (HRWLE) with Seattle protocol biopsies (standard arm) or endoscopy with AFI-directed pCLE and targeted biopsies for molecular biomarkers (experimental arm). Patients crossed over to the other arm after 6 to 12 weeks.	NA	The primary outcome was the diagnostic accuracy for dysplasia of AFI-guided pCLE using the trial histology as the gold standard. Secondary outcomes included the following: (1) diagnostic accuracy of AFI-guided pCLE for dysplasia with reference to the overall histology, which included biopsy specimens taken within 12 months before enrollment; (2) added diagnostic value of molecular biomarkers; (3) time to perform the endoscopy; and (4) patient-reported experience related to experimental and standard endoscopy. A 3-biomarker panel including cyclin A, p53 (assessed with immunohistochemistry), and aneuploidy (assessed with image cytometry) was used.	Within the experimental arm, pCLE had a higher sensitivity than HRWLE for HGD/IMC (P =.046) and all grades of dysplasia (P = .01), but lower specificity (HGD/IMC, P = .01; all grades of dysplasia, P = .02). In per-patient analysis, there was no difference in the sensitivity of pCLE for dysplasia compared with HRWLE with the Seattle protocol (76.5%; 95% CI, 50.1–93.2 vs 76.5%; 95% CI, 50.1–93.2, respectively; P = 1.00 for HGD/IMC; 74.3%; 95% CI, 56.7–87.5 vs 80.0%; 95% CI, 63.1–91.6, respectively; P = .48, for all grades of dysplasia). The use of AFI-targeted pCLE led to 2.1 optical biopsies per patient on average compared with 12.3 tissue biopsy specimens taken in the Seattle protocol. Standard endoscopy missed 28 cases of dysplasia (miss rate, 51.9%), 11 of which were detected by experimental endoscopy. Experimental endoscopy missed 20 dysplastic cases (miss rate, 37%), of which 5 were diagnosed correctly by standard endoscopy. In the overall histology analysis, AFI-guided pCLE had a higher sensitivity for HGD/IMC than Seattle protocol biopsies (73.3%; 95% CI, 54.1–87.7 vs 43.3%; 95% CI, 25.5–62.6, respectively; P = .02). The difference in sensitivity for all grades of dysplasia was not statistically significant (63.0%; 95% CI, 48.7–75.7 vs 51.9%; 95% CI, 37.8–65.7, respectively; P = .13). In the per-patient analysis the sensitivity and specificity for dysplasia of individual biomarkers were 48.6% and 93.9% for p53, 47.1% and 69.4% for cyclin A, and 40.0% and 88.5% for aneuploidy, respectively. At multivariable analysis, p53, aneuploidy, and optical dysplasia correlated significantly with a diagnosis of dysplasia. A panel comprising these 3 biomarkers showed an area under the receiver operating curve of 0.83 (95% CI, 0.76–0.91) for a diagnosis of any grade of dysplasia and 0.88 (95% CI, 0.78–0.97) for a diagnosis of HGD/IMC. Using a threshold of 1 positive biomarker, this panel had a higher sensitivity than the Seattle protocol in detecting dysplasia in the overall histology analysis (81.5% vs 51.9%; P < .001) The difference was not statistically significant in the trial histology analysis (91.4% vs 80.0%; P = .16).	3-biomarker panel including cyclin A, p53, and aneuploidy. The addition of molecular biomarkers could improve diagnostic accuracy

p16

Study	Design	Randomization	Blinding	Population/N	Age	Inclusion criteria	Primary endpoint	Secondary endpoint	Follow-up	Objective	Results	Biomarkers
Xu (2013)	systematic review and meta-analysis	NA	NA	41 articles (39 reporting DNA methylation in tissue samples and 7 reporting DNA methylation in blood samples; 11 for EAC, 29 for SCC and 1 for both)	NA	any article about esophageal cancer, both EAC and SCC	in English and Chinese	evaluation of frequency of DNA methylation for p16 in esophageal cancer	NA	p16 in esophageal cancer	The summarized frequency of DNA methylation of p16 gene was 0.53(95% CI: 0.44–0.61) in esophageal cancer tissues. The frequency of DNA methylation of p16 gene varied with the tumor differentiation grades (well differentiated: 0.37; moderately differentiated: 0.61; poorly differentiated: 0.63) (Fig. 3). No significant difference in the frequency of DNA methylation was observed between EAC and ESCC (0.55 vs. 0.52) For EAC: frequency was 0.55 (95%CI 0.42–0.68)	p16	
Altaf (2017)	meta-analysis	NA	NA	NA (different studies included for different biomarkers)	NA (different studies included for different biomarkers)	Included studie: BE with or without HGD/EAC, comparing biomarkers (p53, Ki-67, p16, DNA content abnormalities) in predicting neoplastic progression of BE to HGD/EAC. Only studies published in the English.	NA	NA	NA (different studies included for different biomarkers)	Progression to of BE to HGD/EAC. Comparing biomarkers (p53, Ki-67, p16, DNA content abnormalities) in predicting neoplastic progression of BE to HGD/EAC.	102 clinical studies were included in the study. Mutation of p53 had the following findings: diagnostic odds ratio (DOR) 10.91, sensitivity 47%, specificity 92%, positive likelihood ratio (PLR) 4.71, negative likelihood ratio (NLR) 0.65, area under the curve (AUC) 0.79. Loss of p53 had the following findings: DOR 16.16, sensitivity 31%, specificity 98%, PLR 6.66, NLR 0.41, AUC 0.923. Both mutations and loss of p53 were found to be superior to other p53 abnormalities such as loss of heterozygosity (LOH) and overexpression. Ki-67 had DOR 5.54, sensitivity 82%, specificity 48%, PLR 1.59, NLR 0.42, AUC 0.761. Aneuploidy had DOR 12.08, sensitivity 53%, specificity 87%, PLR 4.26, NLR 0.42, AUC 0.846. Tetraploidy had DOR 5.87, sensitivity 46%, specificity 85%, PLR 3.47, NLR 0.65, AUC 0.793. Loss of Y chromosome had DOR 9.23, sensitivity 68%, specificity 80%, PLR 2.67, NLR 0.49, AUC 0.807. p16 aberrations (hypermethylation, LOH, mutation and loss) failed to demonstrate any advantage over the other biomarkers studied.	p53, p16, Ki-67 and DNA content abnormalities (aneuploidy, tetraploidy, loss of Y chromosome). Meta-analysis. Included studies until March 2016.

Author (Year)	Study type			N		Inclusion		Comparison	Follow-up	Aim	Results	Conclusion
Nieto (2018)	systematic review and meta-analysis	NA	NA	14 studies (404 patients)	NA	Studies which reported on progression from NDBE to BE with HGD or EAC in the same patient cohort. Relevant epigenetic markers are DNA methylation, histone modification, chromatin remodeling and micro and non-coding RNAs.	NA	NA	NA	epigenetic biomarkers for the progression of Barrett's sophagus to HGD/EAC	14 studies met the inclusion criteria. 42 epigenetic markers were identified, and 5 studies developed models aiming to predict progression to OADC. The evidence from this systematic review is suggestive of a role for p16 as an individual epigenetic biomarker in predicting progression from BE to EAC. Prognostic models incorporating this and other markers also suggest a role for p16 in combination with HPP1, RUNX3 and/or clinical markers.	This started as a meta-analysis, but the results were too heterogeneous, the biomarkers too numerous and the studies too small to be used in a quantitative meta-analysis. So the authors did only a systematic review of available data.
Famhy (2004)	retrospective	NA	NA	40 (21 with HGD/EAC and 19 with no dysplasia)	NA	BE patients with HGD/EAC or without dysplasia	regular clinical follow up/treatment	HGD/EAC vs NDBE	at least 5 years (for NDBE)	FISH to determine if there are specific genetic changes in Barrett's esophagus with associated high-grade dysplasia/intramucosal adenocarcinoma compared to those without dysplasia. Centromeric enumeration probes (CEP) for chromosomes 6, 7, 11, and 12, and locus-specific probes (LSI) for 9p21 (p16 gene), and 17p13.1 (p53 gene) loci along with their corresponding CEP (9 and 17, respectively) were used in this study. A positive FISH result was defined as the presence of cells with >2 CEP signals or with a loss of the LSI signals relative to their corresponding CEP.	p53 locus loss and/or aneusomy of chromosomes 6, 7, 11, and 12 abnormalities could be detected by FISH from 95% of HGD/EAC with a specificity of 100%. When combining the results of both FISH probe sets, aneusomy of chromosomes 6, 7, 11, and 12 or a p53 loss were identified in 20/21 (95%) of the HGD/carcinoma cases, including all five cases with cytologic changes classified as indefinite for dysplasia. The sensitivity and specificity for the detection of HGD/carcinoma using the combination of the above probes was 95 and 100%, respectively (95% CI: 74–99.8% and 79.1–100%; Loss of the p16 locus was seen commonly in patients both with (86%) and without (47%) dysplasia/carcinoma.	FISH for Centromeric enumeration probes (CEP) for chromosomes 6, 7, 11, and 12, and locus-specific probes (LSI) for 9p21 (p16 gene), and 17p13.1 (p53 gene) loci along with their corresponding CEP (9 and 17, respectively) can distinguish between HGD/EAC and NDBE
Galipeau (2007)	prospective cohort study	NA	NA	243	NA	BE patients with or without dysplasia	regular clinical follow up	comparison of progressors and non-progressors	A total of 17,139 patient-months with a mean of 71 mo (median 80.5 mo, range 2.3–130.8 mo)	TP53 and CDKN2A (p16) alterations (methylation/mutation), 17p LOH and 9p LOH, tetraploidy and aneuploidy were evaluated at baseline biopsies and patients followed up in time to observe progression	At 10 y, each molecular and DNA content abnormality, when analyzed alone (univariate) in a patient at baseline, made a significant contribution to prediction of EA risk: 17p LOH (10-y RR = 10.6; 95% CI 5.2–21.3, p, 0.001), 9p LOH (10-y RR =2.6; 95% CI 1.1–6.0, p= 0.03), p53 (RR 7.3 (3.7–14.3), tetraploidy RR 8.8 (4.3–17.7), aneuploidy RR 8.5 (4.3–17.0), but not but not CDKN2A mutation (10-y RR 1.8; 95% CI 0.8–4.1, p=0.13) and CDKN2A methylation (RR 2.1; 95% CI 0.8–4.1, p = 0.09). At final multivariable model, the following risks were seen: 17p LOH 5.4 (2.5–12.0) ,0.001; Tetraploidy 2.9 (1.4–5.9) ,0.001; Aneuploidy 3.4 (1.6–7.1) 0.001; 9p LOH 2.4 (1.0–5.5) 0.045. Compared to no abnormality present, the following relative risks (RR) were observed: one abnormality RR 1.8 (0.48–6.87), p . 0.38; two abnormalities RR 9.0 (2.4–33.3), p , 0.001; three abnormalities RR 38.7 (10.8–138.5), p , 0.001	TP53 and CDKN2A (p16) alterations (methylation/mutation), 17p LOH and 9p LOH, tetraploidy and aneuploidy. Part of the Seattle Barrett's Esophagus Study cohort
Fritcher (2008)	retrospective	NA	NA	92 (of whom, 84 had HGD/EAC, 7 had LGD and 1 NDBE)	mean age of 64.4 years (range, 34-87 years).	Patient with BE in surveillance	regular clinical follow up	regular clinical follow up	mean 267 days (range, 0-1304 days)	Comparison of sensitivity and specificity of conventional cytology, DNA ploidy analysis with digital image analysis (DIA), and fluorescence in situ hybridization (FISH) for the detection of dysplasia (FISH probes to probes to 8q24 (C-MYC), 9p21 (P16), 17q12 (HER2), and 20q13)	FISH was more sensitive (P < .05) than cytology and DIA for low-grade dysplasia, HGD, and EA.Sensitivities: for LGD: cytology, DIA, and FISH 5%, 5%, and 50% respectively; for HGD: 32%, 45%, and 82%, respectively; for EAC: 45%, 45%, and 100%, respectively. Specificity (on patients with only benign squamous mucosa: 93%, 86%, and 100% (P = .22). There was a significant difference between FISH categories (negative, 9p21 loss, gain of a single locus, and polysomy) for progression to HGD/EA (P < .001).	Comparison of sensitivity and specificity of conventional cytology, DNA ploidy analysis with digital image analysis (DIA), and fluorescence in situ hybridization (FISH) for the detection of dysplasia (FISH probes to probes to 8q24 (C-MYC), 9p21 (P16), 17q12 (HER2), and 20q13)
Jin (2009)	retrospective, multicenter, double-blinded validation study of 8 BE progression prediction methylation biomarkers	NA	NA	195 BE biopsies (145 NPs and 50 Ps)	NA	Patients with biopsy-proven Barrett's esophagus with or without dysplasia	Routine clinical follow up-surveillance.	Routine clinical follow up-surveillance.	NA	Promoter methylation levels of 8 genes (p16, HPP1, RUNX3, CDH13, TAC1, NELL1, AKAP12 and SST) to distinguish progressors from non-progressors	Areas under the ROC curve (AUCs) were high in the 2-, 4-year and combined data models (0.843, 0.829 and 0.840; p<0.001, p<0.001 and p<0.001, respectively). In addition, even after rigorous overfitting correction, the incremental AUCs contributed by panels based on the 8 markers plus age vs. age alone were substantial (Δ-AUC = 0.152, 0.114 and 0.118, respectively) in all three models. Patients were classified as LR with a threshold that corresponded to 90% true positives and 43% false-positives; the HR group was defined using a threshold that yielded 43% true-positives and 10% false-positives. Assuming a cumulative progression rate to HGD and/or EAC of 7.5% over 5 years(19), the corresponding negative predictive value relating to our LR threshold was 98.7% (i.e., progression risk in the LR group was 1.3%) and the positive predictive value relating to HR was 27% (i.e., progression risk in the HR group was 27%).	Promoter methylation levels of 8 genes (p16, HPP1, RUNX3, CDH13, TAC1, NELL1, AKAP12 and SST)

Study	Study type			N	Age	Inclusion criteria	Method	Comparison	Follow-up	Aim / Panel	Results	Biomarkers
di Pietro (2015)	cross-sectional study	NA	NA	157 (training cohort, 46 (validation cohort)	66.4 (training cohort), 68.7 (validation cohort)	Age ≥18 yrs, BE with a length of at least C ≥ 2 or C<2M≥4 according to the Prague classification with or without visible lesions.	Biopsies were performed either according to the standard Seattle protocol or under the guidance of autofluorescence imaging (AFI)	Comparison of the accuracy of a panel of molecular biomarkers on AFI-directed biopsies with conventional Seattle protocol biopsies for HGD and EAC. Secondary aims: (i) assessment of diagnostic accuracy for the biomarkers for any grade of dysplasia and (ii) validation of a large panel of biomarkers in an independent prospective study by an independent laboratory.	NA	Accuracy of biomarkers in identifying HGD/EAC. Panel of biomarkers assessed: Aneuploidy/tetraploidy; 9p and 17p loss of heterozygosity; RUNX3, HPP1 and p16 methylation; p53 and cyclin A immunohistochemistry. The statistical analysis consisted of three stages: (1) per-biopsy analysis (correlation between biomarkers and histological outcome in individual targeted areas); (2) per-patient analysis (correlation between overall biomarker result and overall histological outcome in individual patients); and (3) comparison between AFI+ versus AFI− areas (comparative analysis of biomarker diagnostic accuracy for dysplasia in biopsies from AFI+ vs AFI− areas).	All of the biomarkers associated with the presence of confirmed dysplasia, with the exception of 9p LOH (p16). At stricter analysis for HGD/EAC, aneuploidy, p53 immunohistochemistry and cyclin A had the strongest association with dysplasia in the per-biopsy analysis and, as a panel, had an area under the receiver operating characteristic curve of 0.97 (95% CI 0.95 to 0.99) for diagnosing HGD/EC. The diagnostic accuracy for HGD/EC of the three-biomarker panel from AFI+ areas was superior to AFI− areas (p<0.001). Average AUCs in all six databases were significantly higher for the threebiomarker panel assessed on AFI+ areas compared with AFI− areas. The biomarker panel had a sensitivity and a specificity of 95.8% (95% CI 76.9% to 99.8%) and 88.6% (95% CI 79.7% to 94.1%), respectively, for a diagnosis of HGD/EAC. By comparison, the Seattle protocol had similar sensitivities, namely, 95.8% (95% CI 76.9% to 99.8%) for a diagnosis of HGD/EC (p=1.0 when compared with three-biomarker panel). Importantly, using this novel approach 2.8 biopsies per patient were taken on average compared with 12.8 for the standard biopsy protocol (4.5 fold reduction; p<0.001). At validation cohort, the panel had a sensitivity and a specificity of 100% and 85% (95% CI 98.9% to 95.0%), respectively, for a diagnosis of HGD/EAC.	Panel of biomarkers assessed: Aneuploidy/tetraploidy; 9p and 17p loss of heterozygosity; RUNX3, HPP1 and p16 methylation; p53 and cyclin A immunohistochemistry
Allan (2015)	cross-sectional	NA	NA	20	NA	Biopsies of BE from dysplasia and non-dysplatic areas	NA	comparison of dysplastic and non-dysplatic areas in each individual	NA	FISH probes to 9p12 (CDKN2A; p16), 17q11.2-12 (HER2), 8q24.12-13 (CMYC) and 20q13.2 (ZNF217) to compare dysplastic and non-dysplatic areas in each individual	HER2, CMYC and ZNF217 showed significant (all P < 0.0005) increases in copy number in dysplastic mucosa; CDKN2A had an insignificant (P = 0.852) decrease when compared to non-dysplastic mucosa. While aneusomy was strongly predictive of dysplasia, eusomy did not rule it out. Odds ratio for detecting dysplasia: CDKN2A 0.991 (95%CI 0.953–1.031, p= 0.658); HER2 1.359 (95%CI 1.126–1.641,p=0.001); CMYC 1.414 (95%CI 1.145–1.745,p=0.001); ZNF217 1.318 (95%CI 1.111–1.563,p=0.002). HER2 has the greatest area under the curve (AUC) at 0.942 (95%CI 0.874–1), suggesting strong predictive power, followed closely by CMYC at 0.941 (95%CI 0.861–1) and ZNF217 with 0.895 (95%CI 0.796–0.994). CDKN2A (p16) has the smallest area at 0.538 (95%CI 0.349–0.726), suggesting almost no predictive power. While definite aneusomy strongly predicts dysplasia, non-detection of aneusomy with this particular FISH panel does not allow one to conclude that dysplasia is absent, or that the case is necessarily low risk.	Aneusomy detected by FISH to 9p12 (CDKN2A; p16), 17q11.2-12 (HER2), 8q24.12-13 (CMYC) and 20q13.2 (ZNF217)
Timmer (2016)	prospective cohort study	NA	NA	428	60 yrs	Barrett's esophagus without dysplasia (NDBE) at baseline, and length >1 cm, who had a baseline brushing. Patients with histological progression within 6 months were excluded from the final analysis.	Evaluation of six molecular markers by DNA fluorescence in situ hybridisation on brush cytology specimens to study the risk of progression to HGD/EAC	Progressors to HGD/EAC vs. non-progressors. Regular clinical follow up	2019 patient-years (median 45 months per patient; IQR 35–72 months)	Risk of progression to HGD/EAC based on baseline molecular markers. These were measured by DNA FISH using locus-specific probes to p16 (CDKN2A), p53 (TP53), Her-2/neu (ERBB2), 20q, and MYC, and centromeric probes for chromosomes 7 and 17 (Abbott Molecular) to detect aneusomy (aneusomy detected by the centromeric probes for CEP7 and CEP17 was used as a surrogate marker to assess DNA ploidy changes).	An 'Abnormal Marker Count' that counted abnormalities in p16, MYC gain and aneusomy, significantly improved risk prediction beyond using just age and BE length. In multivariate analysis, these three factors (age, Barrett's length, and the markers p16, MYC, and aneusomy) identified a high-risk group with an 8.7-fold (95% CI, 2.6 to 29.8) increased hazard ratio compared with the low-risk group, with an area under the curve of 0.76 (95% CI, 0.66 to 0.86). The specificity of the model was 0.54 meaning that 46% of the non-progressors would be incorrectly classified as high-risk. The PPV of the model was 9%, but its NPV was 99% (meaning that 99% of the patients would be safely classified as low-risk and not progress to high-grade dysplasia or cancer during follow-up.) increased risk with each biomarker: Her-2 gain: 1.02 (0.73-1.44), p16 loss: 1.07 (02-1.12), MYC gain: 1.01 (1.00-1.02), 20q: 1.0 (0.99-1.01), aneusomy: 1.23 (1.06 to 1.43), Abnormal Marker Count: 1.91 (1.29 to 2.81)	p16, MYC, her-2/neu, 20q, Aneusomy (chromosomes 7 & 17). Martinez 2016, Tinner 2016 and Hoefnagel 2020 use the same dataset, only with different focus and different follow up times
Martinez (2016)	prospective cohort study	NA	NA	320	58.9±11.7	Barrett's esophagus without dysplasia (NDBE) at baseline.	Multicolour fluorescence in situ hybridization (FISH) data to assess the genetic clonal diversity at single-cell resolution in NDBE patients.	Progressors to HGD/EAC vs. non-progressors. Regular clinical follow up	Median of 43 months (range 11–130 months). Progressors follow up until HGD/EAC diagnosis (median 34 months)	Study the dynamics of clonal evolution, including changes in the clonal diversity, as well as the frequency and rate of clonal expansions in patients with NDBE. A minimum of 50 cells per sample were scored for abnormalities by FISH at seven markers including CEP7, CEP17, p53, p16, Her-2/neu, 20q and MYC.	The loss of one p16 allele (hemizygous loss of p16) was the most frequently observed alteration overall with 51% of the patients (n=163). p53 loss was found in only 7.5% of patients (n=24) and relative p53 locus loss in 10.6% of patients (n=34). Genetic diversity correlates with the risk of progression to cancer, independently of the choice of diversity statistics used. Although clonal expansions in our patients were generally rare, we provide quantification of the frequency and rate of clonal expansions in a human neoplasm. We only observed one significant clonal expansion every 36.8 patient years of follow-up, and in those cases, the clones grew at an average of 1.58 cm2 per year. Importantly, our data show that measures of clonal diversity are more prognostic than 'traditional biomarkers' that are based on the detection of particular individual genetic abnormalities. Most strikingly, we demonstrate that several of the single-probe diversity measures (MYC and CEP 7) were the best predictors in the multivariate analysis, and conversely that p16-abnormalities are poor prognosticators due to the initial expansion and then contraction of those clones. We show that the level of genetic diversity is invariant over time , suggesting an absence of strong selection in the evolution in NDBE and consequently that progression risk is predetermined by the invariant baseline level of diversity. Used two probe sets: risk of progression: Probe set 1, HR: 4.0 (95% CI 1.6 – 10.0) Probe set 2, HR: 4.0 (1.7 – 9.8)	clonal diversity: CEP7, CEP17, p53, p16, Her-2/neu, 20q and MYC. Martinez 2016, Tinner 2016 and Hoefnagel 2020 use the same dataset, only with different focus and different follow up times
Hadjinicolaou (2020)	Prospective multicenter study	NA	NA	127 (42 progressors, 85 non-progressors)	median age of 65.6 years (IQR, 13.7 yrs)	Age ≥18 yrs, BE with a length of at least C ≥ 2 or C<2M≥4 according to the Prague classification with or without visible lesions. Patients had targeted biopsies during endoscopy with autofluorescence imaging (AFI).	The primary endpoint of this study was progression from NDBO/ID to any grade of dysplasia. The two secondary endpoints were a) progression from NDBO/ID to HGD/OAC, and b) any histologic progression i.e. NDBO/ID to LGD, NDBO/ID to HGD, and LGD to HGD.	Comparisons of biomarkers between progressors and non-progressors. This was a validation cohort from the cross-sectional study described above (di Pietro 2015)	Median 4.6 yrs (IQR, 4.3 yrs). Progressors: 1.2 yrs (IQR, 2.7 yrs) until progression	Evaluation of previously defined 9-molecular biomarker panel (in the study of di Pietro 2015 above) with progression in BE patients. Histological progression was defined as transition from a NDBO or indefinite for dysplasia (ID) to any dysplasia, or if low-grade dysplasia already present, to a higher grade of dysplasia or cancer. Biomarkers: p53 and cyclin A were analysed by immunohistochemistry (IHC); aneuploidy and G2/tetraploidy, were analysed by flow cytometry; p16, RUNX3 and HPP1 hypermethylation was analysed by quantitative methylation-specific PCR (Methylight); and LOH at 9p and 17p loci was analysed by the use of microsatellite markers. Snap frozen biopsies in DMSO were used for aneuploidy, G2 tetraploidy, LOH markers and methylation assays.	Amongst progressors, there were 12 (28.6%) that progressed from NDBD/ID to LGD, 16 (38.1%) that progressed from NDBO/ID to HGD/OAC and 14 (33.3%) that progressed from LGD to HGD/OAC. Of the nine molecular biomarkers, at multivariable analysis p53 and aneuploidy were the only significant predictors of any progression. However, when we excluded patients with progression within 12 months of follow up (prevalent dysplasia), only aneuploidy retained statistical significance. The presence of positive aneuploidy at index endoscopy led to a 6.6-fold higher risk of dysplastic progression over no progression (95% CI: 1.8 24.8, p = 0.005). However the sensitivity of the test to predict progression was low (32%, 95% CI: 16, 52%). ROC analysis showed that a clinical model using patient age and BO length (AUC=0.55; CI: 0.45,0.66) was outperformed in the prediction of any histologic progression by a molecular biomarker model comprising of aneuploidy and p53 with a cut-off of one positive biomarker out of two (AUC=0.68; CI: 0.59,0.77). ROC analysis showed that a model with aneuploidy as the only predictor of dysplastic progression outperformed the clinical model (AUC=0.63; CI: 0.54 0.72). p53 appeared to correlate more with short-term progression. Patients with aberrant p53 expression at index endoscopy had an odds ratio of 6.0 (95% CI: 3.1, 11.2, p = 0.007) of missed dysplasia on endoscopic biopsies.	p53, cyclin A;; aneuploidy and G2/tetraploidy; p16, RUNX3 and HPP1 hypermethylation; LOH at 9p and 17p loci; aneuploidy, G2 tetraploidy

Study	Design			Sample	Age	Inclusion				Biomarker/Aim	Results	Conclusion
Chueca (2020)	cross-sectional investigative	NA	NA	77 paraffin-embedded human esophageal samples from 55 patients (15 samples of squamous epithelium, 36 NDBE, 3 IND, 24 LGD, 4 HGD and 12 adenocarcinoma)	median age 62 years (IQR 58–65)	only biopsies	only investigation of p16 methylation in different biopsies ranging from normal epithelium to EAC	only investigation of p16 methylation in different biopsies ranging from normal epithelium to EAC	NA	only investigation of p16 methylation in different biopsies ranging from normal epithelium to EAC	There were increasing rates of p16 methylation as histological lesion appears and progresses.Squamous epithelium showed the lowest methylation rates: 6% (IQR 5–11) vs. 11%(7–39.50) in negative/indefinite for dysplasia, p<0.01; 10.60%(6–24) in low-grade dysplasia, p<0.05; and 44.50%(9–66.75) in high-grade dysplasia/adenocarcinoma, p<0.01. This latter group also exhibited higher methylation rates than Barrett's epithelium with and without low-grade dysplasia (p<0.05).	p16 methylation
Yousaf (2020)	retrospective casse-control	NA	NA	22 patient for evaluation of biomarkers to discriminate between dysplastic and non-dysplastic, and 134 (170 slides) patients for evaluation of progression	NA for the whole cohort - has different ages for different groups involved	patient with BE of all different histologies, ranginf from non-dysplastic to EAC	regular clinical follow up	comparison between dysplastic and non-dysplastic BE, and then comparison between progressors and non-progressors	NA for the whole cohort; different FU for different groups involved	Evaluation of markers involved in the cell cycle (cyclin D1 [CyD1], Ki-67, P16), cell-cell interaction, and cell differentiation (β-catenin, SATB2, CD44, OCT4) and senescence (γ-H2AX) with ICH to differentiate between dysplastic and non-dysplastic BE, and risk of progression	Difference between dysplastic and non-dysplastic BE: Significant differences (P < .05) between the two groups were found in the surface compartment for Ki-67, γ-H2AX, CD44, and CyD1; in the neck compartment for Ki-67, γ-H2AX, and CD44; and in the base only for γ-H2AX. No significant differences were noted in the expression of P16, β-catenin, SATB2, or OCT4 in any compartments. Ki-67 expression showed the largest difference in expression and smallest P value (P < .001) for identifying dysplasia. At less than 5% expression, surface Ki-67 showed sensitivity of 100%, specificity of 31%, PPV of 69%, and NPV of 100%. At a cutoff level of more than 5%, PPV increased to 91% and NPV declined to 82%; at a cutoff of more than 50%, PPV remained at 91% but NPV declined to 74%. Reevaluation without and with ancillary surface Ki-67 improved Cohen κ correlation between individual pathologists among themselves and with the consensus diagnosis from moderate (overall κ = 0.55; range, 0.40-0.73) to substantial (overall κ = 0.77; range, 0.68-0.95) for discriminating dysplastic from ND lesions. In aggregate, sensitivity of Ki-67 plus histology vs histology alone was 88% vs 64%, respectively; specificity was 67% vs 77%, PPV was 69% vs 69%, and NPV was 88% vs 73%. Progression: The odds ratio for progression between surface Ki-67-positive and Ki-67-negative cases was 15.3 (95% CI, 9.6-24.7). The Pearson correlation coefficient for progression was moderate with Ki-67 (r = 0.56) and without Ki-67 (r = 0.42).	ki67 was the best to discriminate NDBE from dysplastic BE. P16 was not able to discriminate. Study of markers involved in the cell cycle (cyclin D1 [CyD1], Ki-67, P16), cell-cell interaction, and cell differentiation (β-catenin, SATB2, CD44, OCT4) and senescence (γ-H2AX)

Other - mixed

Study	Design			Sample	Age	Inclusion				Biomarker/Aim	Results	Conclusion
Wang (2014)	systematice review and meta-analysis	NA	NA	558 patients (140 Barrett's) from 9 eligible studies. Both SCC and EAC	NA	any article about esophageal cancer, both EAC and SCC	in English and Chinese	evaluation of influence of RUNX3 methylation in esophageal cancer	NA	RUNX3 methylation and association with esophageal cancer	RUNX3 methylation was significantly higher in esophageal cancer than in normal squamous mucosa from the proximal resection margin or esophageal benign lesions (OR = 2.85, CI = 2.01–4.05, P=0.00001). RUNX3 methylation was significantly higher in esophageal adenocarcinoma (EAC) than Barrett's esophagus (OR = 0.35, CI = 0.20–0.59, P=0.0001). RUNX3 methylation was significantly higher in ESCC/EAC than in normal squamous mucosa from the proximal resection margin or esophageal benign lesions. The pooled OR from 6 studies including 347 esophageal cancers and 246 normal squamous mucosa showed OR= 2.85, CI = 2.01–4.05, P=0.00001.	RUNX3 methylation. However, both SCC and EAC were included. These were not differentiated in the analysis
Mallick (2016)	systematice review and meta-analysis	NA	NA	11 studies	NA	studies in English that included microRNA and BE with or without dysplasia	studies in English that included microRNA and BE with or without dysplasia	role of microRNAs to distinguish between BE and non-BE and between non-dysplastic BE and dysplastic BE	NA	role of microRNAs to distinguish between BE and non-BE and between non-dysplastic BE and dysplastic BE	Increased miR-192, -194, and -215, and reduced miR-203 and -205 expression in BE compared to normal. Elevated miR-192, -194, and -215, and diminished miR-203 and -205 levels were also noted for comparisons of HGD or EAC against normal. In contrast, a consistent microRNA expression difference was absent for the comparisons of HGD or EAC against BE. MicroRNAs miR-192, -194, -203, -205, and -215 are promising tissue biomarkers for diagnosing BE.	microRNA to distinguish between BE and non-BE
Nieto (2018)	systematice review and meta-analysis	NA	NA	14 studies (404 patients)	NA	Studies which reported on progression from NDBE to BE with HGD or EAC in the same patient cohort. Relevant epigenetic markers are DNA methylation, histone modification, chromatin remodeling and micro and non-coding RNAs.	NA	NA	NA	epigenetic biomarkers for the progression of Barrett's sophagus to HGD/EAC	14 studies met the inclusion criteria. 42 epigenetic markers were identified, and 5 studies developed models aiming to predict progression to OADC. The evidence from this systematic review is suggestive of a role for p16 as an individual epigenetic biomarker in predicting progression from BO to OADC. Prognostic models incorporating this and other markers also suggest a role for p16 in combination with HPP1, RUNX3 and/or clinical markers.	This started as a meta-analysis, but the results were too heterogeneous, the biomarkers too numerous and the studies too small to be used in a quantitative meta-analysis. So the authors did only a systematic review of available data.

Study	Design			N	Age	Inclusion criteria	Surveillance	Comparison	Follow up	Outcome/Biomarkers	Results	Biomarkers/Notes
Janmaat (2017)	meta-analysis	NA	NA	16 different IHC biomarkers were studied in 36 studies. These studies included 425 cases and 1835 controls	NA (different studies included for different biomarkers)	criteria: (1) association between IHC biomarker expression on FFPE material and risk of neoplastic progression was assessed; (2) a cohort or case-control study design; (3) patients with known or newly diagnosed BE with or without LGD at baseline; (4) patients defined as cases had to have progressed to either HGD or EAC during follow-up; (5) mean follow-up of at least one year from the time of initial BE diagnosis; (6) the possibility to extract an OR.	NA	NA	follow up time was 51 months versus 59 months for cases versus controls, respectively	Progression to of BE to HGD/EAC. Comparing biomarkers (p53, aspergillus oryzae lectin (AOL), Cyclin A, Cyclin D and alpha-methylacyl-CoA racemas) in predicting neoplastic progression of BE to HGD/EAC.	Meta-analyses were possible for p53, AOL, Cyclin A, Cyclin D, and alpha-methylacyl-CoA racemase (AMACR), which were studied 13, 2, 4, 3, and 2 times respectively. Aberrant p53 expression was significantly associated with an increased risk of neoplastic progression with an OR of 3.18 (95% CI 1.68 to 6.03). The overall OR, for aberrant p53 IHC on neoplastic progression, after stratification for histology, was 3.86 (95% CI 2.03 to 7.33).This association was confirmed for both non-dysplastic BE (6.12; 95% CI 2.99 to 12.52) and BE with LGD (OR 8.64; 95% CI 3.62 to 20.62). Another promising biomarker to predict neoplastic progression was AOL, with an OR of 3.04 (95% CI 2.05 to 4.49). Cyclin A showed a tendency towards increased risk (OR 1.90; 95% CI 0.85 to 4.22). Cyclin D did not show a significant association (OR 1.01; 95% CI 0.14 to 7.03). Alpha-methylacyl-CoA racemase had an OR 4.07 (95% CI 0.66 to 25.12). The following IHC biomarkers were investigated only once up to this meta-analysis: β-catenin, CD1a, COX2, HER2, Ki67, Lewis, Mcm2, Sialyl Lewis, SOX2, and WGA	p53, aspergillus oryzae lectin (AOL), Cyclin A, Cyclin D and alpha-methylacyl-CoA racemase. Meta-analysis. Included studies until 09/2016
Murray (2006)	nested case-control study	NA	NA	210 (35 cases and 175 controls)		Cases were BE patients from the NIBR who developed HGD/EAC ≥6 months after their initial BE diagnosis.Each case was matched to up to 5 controls who had BE but had not developed HGD/EAC by the study censor date. All had to be diagnosed with pathology.	Routine clinical follow up-surveillance according to ACG guidelines	Progressors to HGD/EAC vs. non-progressors. Regular clinical follow up	Mean period of follow up in the whole cohort was 3.7 years (range 0–8.0); it was 2.3 years (range 0.5–7.4) in the cases and 3.9 years (range 1.1–7.8) in controls	Progression from BE to HGD/EAC. Biopsies from the time of diagnosis of BE were stained immunohistochemically for TP53, cyclin D1, cyclooxygenase 2 (COX-2), and b-catenin proteins	The odds of diffuse or intense TP53 staining were substantially elevated in biopsies from patients who developed oesophageal adenocarcinoma compared with controls (odds ratio (OR) 11.7 (95% confidence interval (CI) 1.93, 71.4)). No increased risk was associated with focal TP53 staining, defined as ,10% cells positive for this marker. When diffuse or intense staining was compared with none or focal staining, the OR (95% CI) for patients who developed definite OA showing TP53 staining was 9.28 (1.78, 48.3). This difference was also present when all cases were considered (OR 8.42 (95% CI 2.37, 30.0). Despite the association with TP53 staining, only 32.4% of cases had an initial biopsy showing diffuse/intense TP53 staining. There were no significant associations between cyclin D1, COX-2, or b-catenin staining and case control status. The OR for positive staining for both TP53 and COX-2 was markedly increased in cases compared with controls (OR 27.3 (95% CI 2.89, 257.0)) although only 15% of cases had positive staining for both markers.	Used the population-based Northern Ireland BE Register (NIBR). Only p53 and COX-2 were associated with progression, but not cyclin D1 or b-catenin proteins
Galipeau (2007)	prospective cohort study	NA	NA	243	NA	BE patients with or without dysplasia	regular clinical follow up	comparison of progressors and non-progressors	A total of 17,139 patient-months with a mean of 71 mo (median 80.5 mo, range 2.3–130.8 mo)	TP53 and CDKN2A (p16) alterations (methylation/mutation), 17p LOH and 9p LOH, tetraploidy and aneuploidy were evaluated at baseline biopsies and patients followed up in time to observe progression	At 10 y, each molecular and DNA content abnormality, when analyzed alone (univariate) in a patient at baseline, made a significant contribution to prediction of EA risk: 17p LOH (10-y RR = 10.6; 95% CI 5.2–21.3, p, 0.001), 9p LOH (10-y RR =2.6; 95% CI 1.1–6.0, p= 0.03), p53 (RR 7.3 (3.7–14.3), tetraploidy RR 8.8 (4.3–17.7), aneuploidy RR 8.5 (4.3–17.0), but not but not CDKN2A mutation (10-y RR 1.8; 95% CI 0.8–4.1, p=0.13) and CDKN2A methylation (RR 2.1; 95% CI 0.8–4.1, p = 0.09). At final multivariable model, the following risks were seen: 17p LOH 5.4 (2.5–12.0) ,0.001; Tetraploidy 2.9 (1.4–5.9) ,0.001; Aneuploidy 3.4 (1.6–7.1) 0.001; 9p LOH 2.4 (1.0–5.5) 0.045. Compared to no abnormality present, the following relative risks (RR) were observed: one abnormality RR 1.8 (0.48–6.87), p . 0.38; two abnormalities RR 9.0 (2.4–33.3), p , 0.001; three abnormalities RR 38.7 (10.8–138.5), p , 0.001	TP53 and CDKN2A (p16) alterations (methylation/mutation), 17p LOH and 9p LOH, tetraploidy and aneuploidy. Part of the Seattle Barrett's Esophagus Study cohort
Fritcher (2008)	retrospective	NA	NA	92 (of whom, 84 had HGD/EAC, 7 had LGD and 1 NDBE)	mean age of 64.4 years (range, 34-87 years).	Patient with BE in surveillance	regular clinical follow up	regular clinical follow up	mean 267 days (range, 0-1304 days)	Comparison of sensitivity and specificity of conventional cytology, DNA ploidy analysis with digital image analysis (DIA), and fluorescence in situ hybridization (FISH) for the detection of dysplasia (FISH probes to probes to 8q24 (C-MYC), 9p21 (P16), 17q12 (HER2), and 20q13)	FISH was more sensitive (P < .05) than cytology and DIA for low-grade dysplasia, HGD, and EA.Sensitivities: for LGD: cytology, DIA, and FISH 5%, 5%, and 50% respectively; for HGD: 32%, 45%, and 82%, respectively; for EAC: 45%, 45%, and 100%, respectively. Specificity (on patients with only benign squamous mucosa: 93%, 86%, and 100% (P = .22). There was a significant difference between FISH categories (negative, 9p21 loss, gain of a single locus, and polysomy) for progression to HGD/EA (P < .001).	Comparison of sensitivity and specificity of conventional cytology, DNA ploidy analysis with digital image analysis (DIA), and fluorescence in situ hybridization (FISH) for the detection of dysplasia (FISH probes to probes to 8q24 (C-MYC), 9p21 (P16), 17q12 (HER2), and 20q13)
Jin (2009)	retrospective, multicenter, double-blinded validation study of 8 BE progression prediction methylation biomarkers	NA	NA	195 BE biopsies (145 NPs and 50 Ps)	NA	Patients with biopsy-proven Barrett's esophagus with or without dysplasia	Routine clinical follow up-surveillance.	Routine clinical follow up-surveillance.	NA	Promoter methylation levels of 8 genes (p16, HPP1, RUNX3, CDH13, TAC1, NELL1, AKAP12 and SST) to distinguish progressors from non-progressors	Areas under the ROC curve (AUCs) were high in the 2-, 4-year and combined data models (0.843, 0.829 and 0.840; p<0.001, p<0.001 and p<0.001, respectively). In addition, even after rigorous overfitting correction, the incremental AUCs contributed by panels based on the 8 markers plus age vs. age alone were substantial (Δ-AUC = 0.152, 0.114 and 0.118, respectively) in all three models. Patients were classified as LR with a threshold that corresponded to 90% true positives and 43% false-positives; the HR group was defined using a threshold that yielded 43% true-positives and 10% false-positives. Assuming a cumulative progression rate to HGD and/or EAC of 7.5% over 5 years(19), the corresponding negative predictive value relating to our LR threshold was 98.7% (i.e., progression risk in the LR group was 1.3%) and the positive predictive value relating to HR was 27% (i.e., progression risk in the HR group was 27%).	Promoter methylation levels of 8 genes (p16, HPP1, RUNX3, CDH13, TAC1, NELL1, AKAP12 and SST)

Study	Study type			Patients	Age	Population	Intervention	Comparison	Follow-up	Aim	Results	Biomarker/Comments
Rygiel (2008)	prospective cohort study	NA	NA	99 patients with BE with different stages of dysplasia/EAC	Median age was 62 (range, 31-87) years	patients with BE with different stages of dysplasia/EAC	Routine clinical follow up-surveillance (brush cytology performed)	Routine clinical follow up-surveillance.	NA	Investigation of 7p12 (EGFR), 8q24 (c-myc), and 20q13 with FISH for the progression of BE to HGD/EAC	Gains (3-4 copies) of chromosome 17, 8q24 (c-myc), and 20q.13 loci were found in the low frequencies in NDBE. Their frequencies increased with the stage of dysplasia and reached a high incidence in EAC. Amplification (>4 copies) of at least 1 of the loci was observed in 14% of HGD and increased to 50% in EAC (P = 0.015). The most frequently amplified locus was c-myc (18%), followed by 20q13 (13%) and EGFR (11%) in the HGD/EAC cases. High amplification levels (>10 copies) of the loci were more frequent in EAC (72%) compared with HGD (20%; P = 0.049). Gains of the loci might be of value as prognostic markers because they are already present in NDBE and may precede the later event of the amplification as observed in HGD/EAC.	7p12 (EGFR), 8q24 (c-myc), and 20q13
Rossi (2009)	retrospective study	NA	NA	21	63 years (range 37–84).	Patients with biopsy-proven Barrett's esophagus with or without dysplasia	Routine clinical follow up-surveillance.	Comparison of progressors vs. non-progressors	median follow-up of 36 months (range 12–120)	Progression, defined as any worsening of histology in time. Measurement of HER2 status was studied by immunohistochemistry and fluorescence in situ hybridization (FISH) on paraffin-embedded tissue	At univariate analysis, HER2 overexpression/amplification (P=0.038) and the existence of LGD or HGD at diagnosis (P= 0.038)were the only variables significantly associated with the occurrence of progression.	only univariable analysis, small study, progression defined as any type of progression, not only HGD/EAC
Bird-Lieberman (2012)	nested case-control study	NA	NA	380 (89 cases - progressors and 291 controls - non progressors	Cases: 63.8±11.9. Controls: 63.8±11.3	Cases were BE patients from the NIBR who developed HGD/EAC ≥6 months after their initial BE diagnosis.Each case was matched to up to 5 controls who had BE but had not developed HGD/EAC by the study censor date. All had to be diagnosed with pathology.	Routine clinical follow up-surveillance.	Progressors to HGD/EAC vs. non-progressors. Regular clinical follow up	mean 6.7 years (±3.3 y),	Progression of BE to HGD/EAC. The following biomarkers were evaluated: abnormal DNA ploidy, p53, and cyclin A expression, levels of sialyl Lewisa, Lewisx, and Aspergillus oryzae lectin [AOL] and binding of wheat germ agglutinin, evaluate on formalin-fixed paraffin-embedded tissue	A panel comprising low-grade dysplasia, abnormal DNA ploidy, and AOL most accurately identified progressors and nonprogressors. For each 1-point increase in risk score, BE patients with LGD are almost at 4-fold increased odds of progressing to develop EAC (OR, 3.90; 95% CI, 2.39 – 6.37), whereas BE patients without LGD have a 3-fold increased odds of progression (OR, 3.31; 95% CI, 1.81– 6.05). C-statistic for the reduced model of LGD, abnormal DNA ploidy, and AOL was 0.73.	LGD, abnormal DNA ploidy, and AOL. Used the population-based Northern Ireland BE Register (NIBR)
Alvi (2013)	2 cohorts: retrospective and prospective	NA	NA	retrospective cohort (60 BE, 36 dysplastic and 90 EAC) and a prospective multicenter study (98 BE patients, including 28 dysplastic and 9 early EAC)	NA	BE with or without dysplasia	regular clinical follow up	DNA methylation to diagnose presence of dysplasia	NA	DNA methylation to diagnose presence of dysplasia and distinguish dysplastic vs non-dysplastic BE	23% genes demonstrated statistically significant changes in methylation in EAC vs. BE (Wilcoxon P<0.05). Six out of seven genes successfully validatedboth internally internally (Pearson's P<0.01) and externally (ANOVA P<0.001). These were SLC22A18 (tumor suppressing subtransferable candidate 5, a paternally imprinted gene), PIGR (polymeric immunoglobulin receptor), GJA12 (gap junction protein, gamma 2), RIN2 (Ras and Rab interactor 2), RGN (senescence marker protein-30, X-linked gene), TCEAL7 (transcription elongation factor A - like 7, X-linked gene). For SLC22A18, PIGR, TCEAL7 and RIN2 genes it was a gradual increase, whereas for RGN the biggest change in methylation occurred at the onset of dysplasia and for GJA12 this occurred between dysplasia and EAC. Individually GJA12 (AUC=0.973) was best able to distinguish between dysplasia/EAC and non-dysplastic BE followed by PIGR (AUC=0.963), SLC22A18 (AUC=0.954), RIN2 (0.922), RGN (AUC=0.865) but only in males and lastly TCEAL7 (AUC=0.788). The greatest AUC of 0.988 (P<0.01) was obtained using the four gene combination (SLC22A18 + PIGR + GJA12 + RIN2) which had a sensitivity of 94% and a specificity of 97%. This methylation panel was able to stratify patients from the prospective cohort into three risk groups based on the number of genes methylated (low risk: <2 genes, intermediate: 2 and high: >2). The data demonstrated that the risk of both dysplasia and EAC increased with the number of genes methylated (Figure 6b). 17.6% of the cases in the <2 gene methylated group were dysplastic (low grade dysplasia only). In the group with 2 genes methylated the proportion of dysplastic cases increased to 42.3% including 11.5% high grade dysplasia/EAC. In the group with >2 genes methylated 14.5% of cases had LGD and 32.8% had HGD/EAC (combined cases of dysplasia and EAC: 47.3%).	DNA methylation (SLC22A18 + PIGR + GJA12 + RIN2 genes) for the identification of dysplasia
Revilla-Nuin (2013)	prospective cohort study	NA	NA	5 patients for identification of miRNAs and 24 patients for validation	NA	Patient with BE, with or without dysplasia	regular clinical follow up	comparison of progressors to non-progressors	Progressors: median: 4.6 years; range: 3–6 years. Non-progressors: median: 7.1 years; range: 5–10 years.	A quantitative reverse transcription PCR (qRT-PCR)-based study to study the use of micro-RNAs (miRNA) from biopsies to predict progression to HGD/EAC	Only miR-192, 194, 196a, and 196b showed a significantly higher expression in BE samples from patients with progression to EAC compared with those who did not progress to EAC. ROC analysis for each of these 4 miRNAs differentially expressed in BE samples from BE-no-EAC and BE-EAC groups resulted in AUC values between 0.60 and 0.80. At the cutoff value of 8.14 for miR-196a, the sensitivity was 71% and the specificity was 70.6%. For miR-194 at the cutoff value of 176.1, the sensitivity was 71.4% and the specificity was 64.7%. At the cutoff value of 139.9 for miR-192, the sensitivity was 85.7% and the specificity was 50%. And finally, for miR-196b at the cut-off value of 3.08, the sensitivity was 71.43% and the specificity was 68.75%.	use of micro-RNAs (miRNA) from biopsies to predict progression to HGD/EAC
Murao (2015)	case-control study	NA	NA	59 (9 cases with EAC, and 50 sex- and age-matched controls with BE)	70.9 years (SD 15.1) and 67.4 (SD 8.2)	EAC and BE patients	no intervention, just analysis of biopsies to identify biomarkers	comparison of BE and EAC or BE mucosa associated with EAC	NA	Oligonucleotide microarray analysis to identify candidate genes differentially associated with EAC and BE, and validation in a separate cohort.	Out of numerous genes identified, CD55 (Decayaccelerating factor: DAF) appeared to be differentially associated with with BE and EAC. The median CD55 expression levels in cancer lesions (0.030) and BE of the EAC group (P = 0.038) were higher than those in BE of the controls. However, expressions of the other genes were not significantly different by pairwise multiple comparison procedures. CD55-positive staining was observed in the surface of the cancer lesions most strongly and in the BE lesions surrounding the cancer lesions, while there was no CD55-positive staining lesion in the adjacent esophageal or gastric mucosa. ROC curve analyses revealed CD55 mRNA relative expression levels in brushing samples differentiating BE in the EAC group from BE in the control group with the area under the ROC curve (AUC) of 0.837 (95% CI:0.641–1.000, P = 0.002). The most suitable cut-off point of CD55 was 0.02, which enabled the estimation of the sensitivity and specificity for CD55 to be 87.5% and 74.0%, respectively (OR 19.9, 95% C.I. 2.2–177.8, P = 0.007).	CD55

Study	Study type			N	Age	Inclusion	Method	Comparison	Follow-up	Aim	Results	Biomarkers
van Olphen (2015)	case-control study within a multicentre prospective cohort	NA	NA	12,000 biopsies from 635 patients (584 controls-nonprogressors, and 51 cases-progressors)	Controls: median 60 years (IQR 53–69); cases: 65 years (54–71)	Newly diagnosed BE of ≥2 cm. Patients with a history of HGD/EAC ≤9 months were excluded.	regular clinical FU according to ACG 2008 guidelines	comparison of progressors to non-progressors	Controls: median 6.5 years (IQR 5.2–7.2), cases: 3.3 (1.9–5.3)	Role of SOX2, with or without aberrant p53, in predicting progression to HGD/EAC	NDBE showed homogeneous nuclear staining for SOX2; SOX2 was progressively lost in dysplastic BE (Loss of SOX2 in only 2% of biopsy series without dysplasia, in contrast to 28% in LGD and 67% in HGD/EAC; loss of SOX2 expression was more common in biopsy series of cases (25%) than in biopsy series of controls (7%)). Loss of SOX2 expression was associated with an increased risk of neoplastic progression in BE patients after adjusting for gender, age, BE length, and esophagitis (adjusted RR 4.8; 95% CI 3.2–7.0). The sensitivity of aberrant SOX2 expression for predicting neoplastic progression was 25% with a specificity of 94%. Aberrant expression of either SOX2 or p53 was more common in biopsy series of cases than in biopsy series of controls, in both nondysplastic BE (22% vs. 7%) and in LGD (22% vs. 9%).The PPV for neoplastic progression increased from 16% with LGD alone to 56% with concurrent loss of SOX2 and aberrant p53 expression. Aberrant SOX2 or p53 expression in nondysplastic BE was associated with an increased risk of neoplastic progression with an adjusted RR of 5.3 (95% CI 3.3–8.3); the risk was even higher with concurrent LGD (adjusted RR 7.3; 95% CI 4.6–11.5), but aberrant expression of both SOX2 and p53 in BE with LGD was associated with the highest risk of neoplastic progression (adjusted RR 18.5; 95% CI 11.1–31.2).	SOX2, p53
di Pietro (2015)	cross-sectional study	NA	NA	157 (training cohort), 46 (validation cohort)	66.4 (training cohort), 68.7 (validation cohort)	Age ≥18 yrs, BE with a length of at least C ≥ 2 or C<2M≥4 according to the Prague classification with or without visible lesions.	Biopsies were performed either according to the standard Seattle protocol or under the guidance of autofluorescence imaging (AFI)	Comparison of the accuracy of a panel of molecular biomarkers on AFI-directed biopsies with conventional Seattle protocol biopsies for HGD and EAC. Secondary aims: (i) assessment of diagnostic accuracy for the biomarkers for any grade of dysplasia and (ii) validation of a large panel of biomarkers in an independent prospective study by an independent laboratory.	NA	Accuracy of biomarkers in identifying HGD/EAC. Panel of biomarkers assessed: Aneuploidy/tetraploidy; 9p and 17p loss of heterozygosity; RUNX3, HPP1 and p16 methylation; p53 and cyclin A immunohistochemistry. The statistical analysis consisted of three stages: (1) per-biopsy analysis (correlation between biomarkers and histological outcome in individual targeted areas); (2) per-patient analysis (correlation between overall biomarker result and overall histological outcome in individual patients); and (3) comparison between AFI+ versus AFI− areas (comparative analysis of biomarker diagnostic accuracy for dysplasia in biopsies from AFI+ vs AFI− areas).	All of the biomarkers associated with the presence of confirmed dysplasia, with the exception of 9p LOH (p16). At stricter analysis for HGD/EAC, aneuploidy, p53 immunohistochemistry and cyclin A had the strongest association with dysplasia in the per-biopsy analysis and, as a panel, had an area under the receiver operating characteristic curve of 0.97 (95% CI 0.95 to 0.99) for diagnosing HGD/EC. The diagnostic accuracy for HGD/EC of the three-biomarker panel from AFI+ areas was superior to AFI− areas (p<0.001). Average AUCs in all six databases were significantly higher for the threebiomarker panel assessed on AFI+ areas compared with AFI− areas. The biomarker panel had a sensitivity and a specificity of 95.8% (95% CI 76.9% to 99.8%) and 88.6% (95% CI 79.7% to 94.1%), respectively, for a diagnosis of HGD/EAC. By comparison, the Seattle protocol had similar sensitivities, namely, 95.8% (95% CI 76.9% to 99.8%) for a diagnosis of HGD/EC (p=1.0 when compared with three-biomarker panel). Importantly, using this novel approach 2.8 biopsies per patient were taken on average compared with 12.8 for the standard biopsy protocol (4.5 fold reduction; p<0.001). At validation cohort, the panel had a sensitivity and a specificity of 100% and 85% (95% CI 98.9% to 95.0%), respectively, for a diagnosis of HGD/EAC.	Panel of biomarkers assessed: Aneuploidy/tetraploidy; 9p and 17p loss of heterozygosity; RUNX3, HPP1 and p16 methylation; p53 and cyclin A immunohistochemistry
van Olphen (2016)	case-control study within a prospective cohort	NA	NA	625 (50 cases-progressors vs. 575 controls-non progressors)	median age of 60 years (interquartile range (IQR) 53–69))	Patients with biopsy-proven Barrett's esophagus and no HGD/EAC within 9 months	Routine clinical follow up-surveillance.	Progressors to HGD/EAC. Regular clinical follow up	median duration of 6.7 years (IQR 5.0–7.4). Progressors: median follow-up of 3.2 years (IQR 1.9–5.3).	Progression from BE to HGD/EAC. Cyclin A expression was determined by immunohistochemistry in biopsies of 625 patients; these results were combined with the histological diagnosis and our previous p53, AMACR, and SOX2 data in loglinear regression models. Differences in discriminatory ability were quantified as changes in area under the ROC curve (AUC) for predicting neoplastic progression.	Cyclin A surface positivity significantly increased throughout the metaplasia–dysplasia–carcinoma sequences and was seen in 10% (107/1050) of biopsy series without dysplasia, 33% (109/335) in LGD, and 69% (34/50) in HGD/EAC (p<0.001). Positive cyclin A expression was associated with an increased risk of neoplastic progression (adjusted relative risk (RRa) 2.4; 95% CI: 1.7–3.4), and was particularly seen in biopsy series with LGD (adjusted RR of 5.8; 95% CI: 3.7–9.0). In per-biopsy analysis, cyclin A had an AUC of 0.59 (95% CI: 0.54–0.64) for predicting neoplastic progression with a sensitivity of 32%, a specificity of 86%, a PPV of 21%, and a NPV of 92%. Increases in AUC were substantial for P53 (+0.05), smaller for SOX2 (+0.014), minor for cyclin A (+0.003), and none for AMARC (0.00). There is a challenge in interpretating cyclin A immunohistochemistry with a moderate interobserver agreement with a kappa value of 0.46.	LGD, SOX3, Cyclin A and P53
Timmer (2016)	prospective cohort study	NA	NA	428	60 yrs	Barrett's esophagus without dysplasia (NDBE) at baseline, and length >1 cm, who had a baseline brushing. Patients with histological progression within 6 months were excluded from the final analysis.	Evaluation of six molecular markers by DNA FISH using locus-specific probes to p16 (CDKN2A), p53 (TP53), Her-2/neu (ERBB2), 20q, and MYC, and centromeric probes for chromosomes 7 and 17 (Abbott Molecular) to detect aneusomy (aneusomy detected by the centromeric probes for CEP7 and CEP17 was used as a surrogate marker to assess DNA ploidy changes).	Progressors to HGD/EAC vs. non-progressors. Regular clinical follow up	2019 patient-years (median 45 months per patient; IQR 35–72 months)	Risk of progression to HGD/EAC based on baseline molecular markers. These were measured by DNA FISH on brush cytology specimens to study the risk of progression to HGD/EAC	An 'Abnormal Marker Count' that counted abnormalities in p16, MYC gain and aneusomy, significantly improved risk prediction beyond using just age and BE length. In multivariate analysis, these three factors (age, Barrett's length, and the markers p16, MYC, and aneusomy) identified a high-risk group with an 8.7-fold (95% CI, 2.6 to 29.8) increased hazard ratio compared with the low-risk group, with an area under the curve of 0.76 (95% CI, 0.66 to 0.86). The specificity of the model was 0.54 (meaning that 46% of the non-progressors would be incorrectly classified as high-risk). The PPV of the model was 9%, but its NPV was 99% (meaning that 99% of the patients would be safely classified as low-risk and not progress to high-grade dysplasia or cancer during follow-up.) increased risk with each biomarker: Her-2 gain: 1.02 (0.73-1.44), p16 loss: 1.07 (02-1.12), MYC gain: 1.01 (1.00-1.02), 20q: 1.0 (0.99-1.01), aneusomy: 1.23 (1.06 to 1.43), Abnormal Marker Count: 1.91 (1.29 to 2.81)	p16, MYC, her-2/neu, 20q, Aneusomy (chromosomes 7 & 17). Martinez 2016, Tinner 2016 and Hoefnagel 2020 use the same dataset, only with different focus and different follow up times
Martinez (2016)	prospective cohort study	NA	NA	320	58.9±11.7	Barrett's esophagus without dysplasia (NDBE) at baseline.	Multicolour fluorescence in situ hybridization (FISH) data to assess the genetic clonal diversity at single-cell resolution in NDBE patients.	Progressors to HGD/EAC vs. non-progressors. Regular clinical follow up	Median of 43 months (range 11–130 months). Progressors follow up until HGD/EAC diagnosis (median 34 months)	Study the dynamics of clonal evolution, including changes in the clonal diversity, as well as the frequency and rate of clonal expansions in patients with NDBE. A minimum of 50 cells per sample were scored for abnormalities by FISH at seven markers including CEP7, CEP17, p53, p16, Her-2/neu, 20q and MYC.	The loss of one p16 allele (hemizygous loss of p16) was the most frequently observed alteration overall with 51% of the patients (n=163). p53 loss was found in only 7.5% of patients (n=24) and relative p53 locus loss in 10.6% of patients (n=34). Genetic diversity correlates with the risk of progression to cancer, independently of the choice of diversity statistics used. Although clonal expansions in our patients were generally rare, we provide quantification of the frequency and rate of clonal expansions in a human neoplasm. We only observed one significant clonal expansion every 36.8 patient years of follow-up, and in those cases, the clones grew at an average of 1.58 cm2 per year. Importantly, our data show that measures of clonal diversity are more prognostic than 'traditional biomarkers' that are based on the detection of particular individual genetic abnormalities. Most strikingly, we demonstrate that several of the single-probe diversity measures (MYC and CEP 7) were the best predictors in the multivariate analysis, and conversely that p16-abnormalities are poor prognosticators due to the initial expansion and then contraction of those clones.We show that the level of genetic diversity is invariant over time , suggesting an absence of strong selection in the evolution in NDBE and consequently that progression risk is predetermined by the invariant baseline level of diversity. Used two probe sets: risk of progression: Probe set 1, HR: 4.0 (95% CI 1.6 – 10.0) Probe set 2, HR: 4.0 (1.7 – 9.8)	clonal diversity: CEP7, CEP17, p53, p16, Her-2/neu, 20q and MYC. Martinez 2016, Tinner 2016 and Hoefnagel 2020 use the same dataset, only with different focus and different follow up times

Study	Study type			Sample	Age	Patients	Method	Comparison	Follow-up	Outcome assessed	Results	Biomarker
Stachler (2018)	retrospective case–control study	NA	NA	97 (24 cases-progressors and 73 controls - non progressors)	cases 67.5 vs. controls 61.6 years	Patients with biopsy-proven Barrett's esophagus with NDBE, IND or LGD. If someone developed HGD/EAC within 12 months, they were excluded.	Routine clinical follow up-surveillance.	Comparison of progressors vs. non-progressors	mean follow-up (controls 6.7 vs. cases 3.3 years, P<0.001)	Porgression to HGD/EAC. From each patient, we selected a single tissue sample obtained more than 1 year before progression (cases) or more than 2 years before the end of follow up (controls). Pathogenic mutations, gene copy numbers, and ploidy were compared between samples from progressors and non-progressors.	TP53 mutations were detected in 46% of samples from progressors and 5% of non-progressors. In this case–control sample set, TP53 mutations in BE tissues increased the adjusted risk of progression 13.8-fold (95% CI, 3.2–61.0) (P<.001). These results were confirmed in a separate validation set of 16 NDBE who progressed and 28 NDBE who did not progress. We did not observe significant differences in ploidy or copy number profile between groups. We identified 147 pathogenic mutations in 57 distinct genes—the average number of pathogenic mutations was higher in samples from progressors (2.5) than non-progressors (1.2) (P<.001). TP53 and other somatic mutations were recurrently detected in samples with limited copy number changes (aneuploidy). Beyond TP53, several known GEA tumor suppressors and oncogenes (ARID1B, APC, ERBB2, RB1, RUNX1, LARP4B, and BIRC5) had more frequent mutations in progressors. Among these, ARID1B (OR 11.3 (1.7-75.9)), APC (NA), and ERBB2 (4.6 (0.4-59.8)), muttion burden (1.5 (1.1-2.1)) were significantly enriched in the progressors (P<0.05).	p53; ARID1B, APC, and ERBB2
Duits (2018)	nested case-control study	NA	NA	260 (130 cases-progressos vs 130 controls non-progressors)	Progressors:59.9 ± 9.6; non-progressors 59.2 ± 9.7	Patients with biopsy-proven Barrett's esophagus with NDBE, IND or LGD. If someone developed HGD/EAC within 24 months, they were excluded.	Routine clinical follow up-surveillance.	Comparison of progressors vs. non-progressors	progressors 3.7 vs. non-progressors 4.8 years; P < 0.001	Porgression to HGD/EAC. We assessed abnormal DNA content, p53, Cyclin A, and Aspergillus oryzae lectin (AOL) in FFPE sections.	LGD, p53 and AOL were independently associated with neoplastic progression. In the multivariate analysis, expert LGD had a 36-fold increased odds of progression (OR, 35.7; 95% CI, 1.4–920.8). Abnormal p53 expression (OR, 4.1; 95% CI, 1.4–12.4) and abnormal AOL expression in three epithelial compartments had a 4-fold increased odds of progression (OR, 4.3; 95% CI, 0.7–26.3). Cyclin A did not predict progression, and DNA ploidy analysis by image cytometry was unsuccessful in the majority of cases, both were excluded from the multivariate analysis. The multivariable biomarker model had an area under the receiver operating characteristic curve of 0.73.	AOL, and p53, cyclin A, DNA ploidy. Utilized the Amsterdam-based ReBus nested case-control cohort, a multicenter prospective cohort study
Moinova (2018)	cross-sectional explorative study	NA	NA	173 (training cohort; 62 control patients without BE, and 111 patients with BE/EAC at endoscopy) and 149 individuals (validation cohort; 30 control patients without BE, and 119 patients with BE/EAC at endoscopy)	NA	Patients with or without BE who had undergone endoscopy and had biopsies/brushing	First with Reduced Representation Bisulfite Sequencing (RRBS) on a set of 26 EAC biopsies and respective matched normal squamous biopsies, 15 biopsy or brushing samples of BE, and 5 EAC cell lines. After identifying patches of DNA that were associated with BE (at the CCNA1 alleles), they performed targeted resequencing of this differentially-methylated region with next-generation sequencing (NGS) based assay. Once identified the methylated CCNA1 they performed cytology brushings of the distal esophagus to evaluate it. Further validation of these biomarkers in cytology obtained in a balloon before endoscopy in 86 patients who did it	As explained in column "I"	NA	Tested CCNA1 DNA methylation as a BE biomarker in cytology brushings of the distal esophagus, in comparison with and in addition to the methylated vimectin gene (mVIM; a BE biomarker)	In esophageal biopsies, mCCNA1 detected in 81% percent of nondysplastic BE, 68% percent of BE with high-grade dysplasia, and 90% percent of EAC, but in only 1% of normal squamous samples. In the training set, cytology brushing, CCNA1 DNA methylation demonstrated an area under the curve (AUC)=0.95 for discriminating BE-related metaplasia and neoplasia cases versus normal individuals, performing identically to methylation of VIM DNA, an established BE biomarker (90.7% sensitivity and 98.4% specificity for BE/EAC). At validation cohort, AUC was 0.95, and similar sensitivity/specificity. Molecular cytology assays of distal esophageal brushings, by bisulfite-sequencing detection of the 2-marker panel of mVIM and mCCNA1 DNAs, detects BE and EAC with sensitivity 95% and specificity 91%. In balloon samples from 86 individuals, tests of CCNA1 plus VIM DNA methylation detected BE metaplasia with 90.3% sensitivity and 91.7% specificity, with AUC=0.92 for mCCNA1 and AUC =0.91 for mVIM. The combination of mVIM and mCCNA1 was more sensitive than either marker individually, even when the specificity of each individual marker was adjusted to match that of the combination.	CCNA1 DNA methylation
Li (2019)	cross-sectional explorative study	NA	NA	564 samples of 4 tissue types of esophagus (normal squamous esophagus (NSE), BE, EAC, and SCC). Also 996 samples from neck squamous cell carcinoma and stomach adenocarcinoma	NA		Identification of aberrant methylation in different esophageal hisktology (from normal to BE, EAC and SCC)	Identification of aberrant methylation in different esophageal hisktology (from normal to BE, EAC and SCC)	NA	Identification of aberrant methylation in different esophageal histology (from normal to BE, EAC and SCC). Also 996 samples from neck squamous cell carcinoma and stomach adenocarcinoma were included to reduce the "noise" from possible adjacent tumors. Traning set composed of 377 biopsies was then tested in the test set (187 biopsies), and then in a validation set (184 biopsies). Then, independent prognostic methylation markers were identified using multivariable Cox regression for EAC (n=79) and SCC (n=90).	Twelve CpG sites with frequency greater than or equal to 9 were selected to construct the diagnostic methylation classifier. The 12 CpG-based diagnostic classifier achieved total accuracy rate of 93.9% (95% CI 91.0%–96.1%) at multinomial logistic regression model and an AUC of 0.992. At the testing set, the total accuracy rate was 93.1% (95% CI: 88.4%-96.3%) and the AUC was 0.990. At the validation set, the diagnostic classifier could effectively predict group membership in 159 (86.4%, 95% CI: 80.6%-91.0%) of 184 samples, with an AUC of 0.978. For 12 CpG sites, the distribution of methylated levels in the validation set was consistent with those in the training and test sets. Only one CpG site was overlapped between the independent prognostic markers of EAC and those of SCC, revealing that two types of esophageal cancer had distinct sets of prognostic methylation markers. Prognostic methylation classifiers were constructed with 3 CpG sites for EAC and 2 CpG sites for ESCC. Patients were classified into high-risk group and low-risk group based on the median of the risk score of classifiers. The Kaplan-Meier survival curve showed a significant difference in survival time between the two groups (Log-rank P < 0.0001). The 3 CpG-based prognostic classifier for EAC (Hazard ratio [HR] = 5.164, 95%CI 2.199-12.130, p=0.02) was an independent risk factors by multivariate Cox regression adjusting clinical risk factors. Time-dependent ROC curve analysis indicated that the predictive performance of the prognostic methylation classifiers was superior to those of clinical risk factors. After combining the prognostic classifier and tumor stage, patients were divided into 4 risk levels of G1 (low-risk and early-stage), G2 (low-risk and advanced-stage), G3 (high-risk and early-stage), and G4 (high-risk and advanced-stage). Kaplan-Meier curves showed that patients in the different levels of risk stratification demonstrated significantly different prognoses (Log-rank P < 0.0001).	DNA methylation for discriminating and prognosis of of EAC/SCC

Study	Design			Sample	Age	Inclusion criteria	Follow-up	Comparison	Follow-up time	Marker evaluation	Results	Biomarkers
Porter (2020)	cross-sectional explorative study	NA	NA	58 EAC, 15 non-BE, 14 BE adjacent to EAC, 78 BE, 15 LGD, 14 BE adjacent to dysplasia	NA	NA	NA	NA	NA	exploration on biopsy specimen of subcellular epithelial protein (HMGB1, p53, RUNX3) expression, alongside expression of CD20, CD4, CD8 and Foxp3 to characterize stromal B lymphocyte, and helper, cytotoxic and regulatory T-lymphocyte cell infiltrate, assessed by immunohistochemistry	There was an increased intensity of nuclear HMGB1 in the background BO in those that had progressed to either dysplasia (71%) or cancer (67%) compared with BO from non-progressors (27%), p ≤ 0.017 and p = 0.024, respectively. In addition, patients who had progressed to cancer also expressed weaker epithelial cytoplasmic HMGB1 in their background BO (absent + weak intensity in 67%) compared with patients who did not have malignancy (absent + weak intensity in 24%), p = 0.015. Cytoplasmic expression of HMGB1 was similar in background BO whether dysplasia was present or not. Nuclear p53 was absent in the majority (80%) of normal epithelium and emerged in dysplastic BO (87% as moderate + strong expression), p < 0.001, as expected. Oesophageal adenocarcinoma expressed stronger nuclear p53 compared with normal mucosa (p = 0.002) and non-dysplastic BO (p < 0.001), and weaker nuclear p53 expression compared with dysplastic BO (p = 0.006). RUNX3 does not seem discriminatory between dyspastic/EAC and non-dyspolastic BE/normal tissue. Compared with normal epithelium, non-dysplastic BO is associated with reduced lymphocytic infiltration of CD20+ B-cells (p < 0.001), CD4+ T-cells (p < 0.001) and CD8+ T-cells (p < 0.001). In areas of dysplastic BO there is an increase of CD20+ B-cells (p = 0.003) and CD8+ T-cells (p = 0.012) and an increase in Foxp3+ Tregs (p < 0.001) compared with non-dysplastic BO. Individuals with BO who progressed to dysplasia demonstrated an immune cell infiltrate signature in background non-dysplastic BO characterised by increased CD20 + B-cells (p = 0.038) compared with non-dysplastic BO of non-progressors. Similarly, patients progressed to adenocarcinoma displayed increased CD20+ (p < 0.001), CD4+ (p = 0.003) and CD8+ (p = 0.014) lymphocytes in the background non-dysplastic BO compared with nonprogressors.	HMGB1, p53, RUNX3, lymphocyte populations
Hoefnagel (2020)	prospective cohort study	NA	NA	334 (220 from 6 community hospitals 114 from an academic center	Median 60.0 years (interquartile range (IQR) = 15.75	Barrett's esophagus without dysplasia (NDBE). Patients who received endoscopic treatment or who had dysplasia in the past or who progressed within 6 months from baseline were excluded. In addition, patient with inadequate number of cells for FISh analysis were also excluded	Routine clinical follow up-surveillance according to ACG guidelines	Progressors to HGD/EAC vs. non-progressors. Regular clinical follow up	median follow up time of 86 months (IQR 39.7) overall, median until progression 39.9 (IQR 55.3) for progressors, and for non-progressors median follow up was 90.2 (IQR 39.1)	Porgression to HGD/EAC, according to biomarkers' score. Biomarkers included: aberrations for chromosomes 7, 17, and structural abnormalities for c-MYC, CDKN2A, TP53, Her-2/neu and 20q assessed by DNA fluorescence in situ hybridization on brush cytology specimens, were used to determine marker scores and to perform clonal diversity measurements, and aneusomy. Several diversity scores were calculated including the Shannon and Simpson indices and normalized clone score.	At univariate analyses the diversity variable, namely the normalized clone score over markers CEP7, CEP17, 20q (20q13.2) and c-MYC (8q24.12) was the most significant predictor of progression. Similarly, increased clonal abundancy was found to be associated with an increased risk of progression. At multivariable analysis, the model, which included the diversity measure defined as the normalized clone score over the FISH markers CEP7, CEP17, 20q and c-MYC and the clinical variables of age and C-Barrett length, was superior to other models with regard to predicting progression. It had a median integrated Area Under the Curve of 0.88 (IQR 0.84–0.91). When choosing a relatively low risk score of 3.93 as a cut off to select for high and low risk patients, the model yields a sensitivity of 91% and a negative predictive value of 97% (95%CI 93%, 99%). This is highly beneficial for the safety of the at-risk patients, and of patients that would test negative for the test. The down side is that the low specificity (38%) leads to many cases which will falsely test positively. The annual progression risk to HGD/EAC in the low risk group was 0.31% versus 1.92% in the high risk group.	aberrations for chromosomes 7, 17, and structural abnormalities for c-MYC, CDKN2A, TP53, Her-2/neu and 20q . Martinez 2016, Tinner 2016 and Hoefnagel 2020 use the same dataset, only with different focus and different follow up times
Hadjinicolaou (2020)	Prospective multicenter study	NA	NA	127 (42 progressors, 85 non-progressors)	median age of 65.6 years (IQR, 13.7 yrs)	Age ≥18 yrs, BE with a length of at least C ≥ 2 or C<2M≥4 according to the Prague classification with or without visible lesions. Patients had targeted biopsies dureing endoscopy with autofluorescence imaging (AFI).	The primary endpoint of this study was progression from NDBO/ID to any grade of dysplasia. The two secondary endpoints were a) progression from NDBO/ID to HGD/OAC, and b) any histologic progression i.e. NDBO/ID to LGD, NDBO/ID to HGD, and LGD to HGD.	Comparions of biomarkers between progressors and non-progressors. This was a validation cohort from the cross-sectional study described above (di Pietro 2015)	Median 4.6 yrs (IQR, 4.3 yrs). Progressors: 1.2 yrs (IQR, 2.7 yrs) until progression	Evaluation of previously defined 9-molecular biomarker panel (in the study of di Pietro 2015 above) with progression in BE patients. Histological progression was defined as transition from a NDBO or indefinite for dysplasia (ID) to any dysplasia, or if low-grade dysplasia already present, to a higher grade of dysplasia or cancer. Biomarkers: p53 and cyclin A were analysed by immunohistochemistry (IHC); aneuploidy and G2/tetraploidy, were analysed by flow cytometry; p16, RUNX3 and HPP1 hypermethylation was analysed by quantitative methylation-specific PCR (Methylight); and LOH at 9p and 17p loci was analysed by the use of microsatellite markers. Snap frozen biopsies in DMSO were used for aneuploidy, G2 tetraploidy, LOH markers and methylation assays.	Amongst progressors, there were 12 (28.6%) that progressed from NDBO/ID to LGD, 16 (38.1%) that progressed from NDBO/ID to HGD/OAC and 14 (33.3%) that progressed from LGD to HGD/OAC. Of the nine molecular biomarkers, at multivariable analysis p53 and aneuploidy were the only significant predictors of any progression. However, when we excluded patients with progression within 12 months of follow up (prevalent dysplasia), only aneuploidy retained statistical significance. The presence of positive aneuploidy at index endoscopy led to a 6.6-fold higher risk of dysplastic progression over no progression (95% CI: 1.8 24.8, p = 0.005). However the sensitivity of the test to predict progression was low (32%, 95% CI: 16, 52%). ROC analysis showed that a clinical model using patient age and BO length (AUC=0.55; CI: 0.45,0.66) was outperformed in the prediction of any histologic progression by a molecular biomarker model comprising of aneuploidy and p53 with a cut-off of one positive biomarker out of two (AUC=0.68; CI: 0.59,0.77). ROC analysis showed that a model with aneuploidy as the only predictor of dysplastic progression outperformed the clinical model (AUC=0.63; CI: 0.54 0.72). p53 appeared to correlate more with short-term progression. Patients with aberrant p53 expression at index endoscopy had an odds ratio of 6.0 (95% CI: 3.1, 11.2, p = 0.007) of missed dysplasia on endoscopic biopsies.	p53, cyclin A;; aneuploidy and G2/tetraploidy; p16, RUNX3 and HPP1 hypermethylation; LOH at 9p and 17p loci; aneuploidy, G2 tetraploidy
Yousaf (2020)	retrospective casse-control	NA	NA	22 patient for evaluation of biomarkers to discriminate between dysplastic and non-dysplastic, and 134 patients for evaluation of progression	NA for the whole cohort - has different ages for different groups involved	patient with BE of all different histologies, ranginf from non-dysplastic to EAC	regular clinical follow up	comparison between dysplastic and non-dysplastic BE, and then comparison between progressors and non-progressors	NA for the whole cohort; different FU for different groups involved	Evaluation of markers involved in the cell cycle (cyclin D1 [CyD1], Ki-67, P16), cell-cell interaction, and cell differentiation (β-catenin, SATB2, CD44, OCT4) and senescence (γ-H2AX) with ICH to differentiate between dysplastic and non-dysplastic BE, and risk of progression	Difference between dysplastic and non-dysplastic BE: Significant differences (P < .05) between the two groups were found in the surface compartment for Ki-67, γ-H2AX, CD44, and CyD1; in the neck compartment for Ki-67, γ-H2AX, and CD44; and in the base only for γ-H2AX. No significant differences were noted in the expression of P16, β-catenin, SATB2, or OCT4 in any compartments. Ki-67 expression showed the largest difference in expression and smallest P value (P < .001) for identifying dysplasia. At less than 5% expression, surface Ki-67 showed sensitivity of 100%, specificity of 31%, PPV of 69%, and NPV of 100%. At a cutoff level of more than 5%, PPV increased to 91% and NPV declined to 82%; at a cutoff of more than 50%, PPV remained at 91% but NPV declined to 74%. Reevaluation without and with ancillary surface Ki-67 improved Cohen κ correlation between individual pathologists among themselves and with the consensus diagnosis from moderate (overall κ = 0.55; range, 0.40-0.73) to substantial (overall κ = 0.77; range, 0.68-0.95) for discriminating dysplastic from ND lesions. In aggregate, sensitivity of Ki-67 plus histology vs histology alone was 88% vs 64%, respectively; specificity was 67% vs 77%, PPV was 69% vs 69%, and NPV was 88% vs 73%. Progression: The odds ratio for progression between surface Ki-67-positive and Ki-67-negative cases was 15.3 (95% CI, 9.6-24.7). The Pearson correlation coefficient for progression was moderate with Ki-67 (r = 0.56) and without Ki-67 (r = 0.42).	ki67 was the best to discriminate NDBE from dysplastic BE. Study of markers involved in the cell cycle (cyclin D1 [CyD1], Ki-67, P16), cell-cell interaction, and cell differentiation (β-catenin, SATB2, CD44, OCT4) and senescence (γ-H2AX)

Study	Study type			Sample	Age	Cases/Controls	Intervention/exposure	Comparison	Comparison details		Objective	Results	Biomarkers/Conclusion
Goda (2021)	Exploratory cross-sectional case-control study	NA	NA	28 (training set), 53 (validation set)	Short BE: cases 68.1, controls 67.6. Long BE: cases 65.5, controls 73.9	Cases: patients with EAC from either long or short BE. Controls: patient with either long or short BE but not EAC. Exclusion criteria were history of endoscopic ablation of BE or gastrectomy, gastric cancer, or other malignant lesions, hemorrhagic diseases, cirrhosis, or renal failure.	Brushing of BE and EAC patients was performed and used to identify biomarkers. First a training set was used to identify candidate biomarker genes with microarray analysis, and then confirmed with RT-qPCR in a validation set.	No comparison, just exploratory study	NA		Identification of biomarkers associated with BE/EAC which could suggest progression from BE to EAC.	Training set: The following 16 genes were identified as significantly differentially expressed (p < 0.05) in BE samples among the four groups (controls vs. SSBE or LSBE, EAC patients vs. SSBE or LSBE): Decay accelerating factor: DAF (CD55); desmocollin 2 (DSC2); topoisomerase type IIα (TOP2A); matrix metalloproteinases (MMP-1, -3, -9, -10, and -12); mannose receptor C type 1 (MRC1); phosphoenolpyruvate carboxykinase 1 (PCK1); serpin peptidase inhibitor, clade B, member 7 (SERPINB7); thyrotropin-releasing hormone degrading enzyme (TRHDE); PDZ domain containing 1 (PDZK1); sodium channel epithelial 1 beta subunit (SCNN1B); brain abundant membrane attached signal protein 1 (BASP1); and cortical thymocyte-like protein (CTXL). Validation set: expression levels of CD55 (p = 0.05) and TOP2A (p = 0.021) in the cancer lesions of the SSBE EAC group were significantly higher than those in the control group. The expression levels of TRHDE (p = 0.034) and SERPINB7 (p = 0.037) in the non-cancer lesions of the SSBE EAC group were significantly higher than those in the control group. The expression levels of SCNN1B (p = 0.022) in the cancer lesions of the SSBE EAC group were significantly lower than those in the control group. The expression levels of MRC1 (p = 0.004) in the non-cancer lesions of the SSBE EAC group were significantly lower than those in the control group. The expression levels of MMP12 (p = 0.057) in the cancer lesions of the LSBE EAC group tended to be higher than those in the non-cancer lesions. Comparing the gene expression in the SSBE control group and the LSBE control group, the median expression levels of CD55 (p = 0.011), DSC2 (p = 0.037), PCK1 (p = 0.002), TRHDE (p = 0.001), PDZK1 (p < 0.005), and SERPINB7 (p = 0.005) in the LSBE group were significantly higher. It was found that the expression levels of BASP1 (p = 0.048), SCNN1B (p = 0.004), and MRC1 (p = 0.029) were also significantly lower.	Exploration of gene mutations that can lead to progression
Rickelt (2022)	retrospective	NA	NA	481 formalin-fixed, paraffin-embedded (FFPE) biopsies of BE and BE–related neoplasia from 321 patients undergoing endoscopy between 1991 and 2019 in 3 independent cohorts	NA (as a single value)	Patients with BE, with or without dysplasia	regular clinical follow up	regular clinical follow up	NA (as a single value)		Investigation of gene expression patterns of extracellular matrix (ECM) molecules in BE and BE–related neoplasia as biomarkers for the diagnosis of BE–related neoplasia and to predict neoplastic progression. Immunohistochemical expression of basement membrane (BM) marker agrin (AGRN) and p53 was analyzed in biopsies of BE–related neoplasia.	Differential gene-expression analysis revealed significant enrichment of ECM matrisome gene sets in dysplastic BE and EAC compared with controls. Loss of BM AGRN expression was observed in both BE–related dysplasia and EAC. The mean AGRN loss in BE glands was significantly higher in BE–related dysplasia/EAC compared with NDBE (P < 0.001; specificity =82.2% and sensitivity=96.4%). Loss of AGRN was significantly higher in NDBE samples from progressors compared with non-progressors (P < 0.001) and identified patients who progressed to advanced neoplasia with a specificity of 80.2% and sensitivity of 54.8%. Combination of AGRN loss and abnormal p53 staining identified progression to BE–related advanced neoplasia with a specificity and sensitivity of 86.5% and 58.7%.	Agrin, p53
Helminen (2022)	retrospective case-control	NA	NA	45 cases with HGD/EAC (24 were progressive from LGD and 21 progressive from metaplasia) and 92 controls (45 non-progressive LGD and 52 non-progressive metaplasia)	NA for the whole cohort - has Table 1 with different ages for each case-control group	Patients with ≥1 EGD >6 months before the diagnosis of HGD/EAC. Controls: non-progressive BE with or without LGD confirmed with follow-up EGDs performed at least 5 years after the initial diagnosis, matched by age (± 5 years) and sex to the cases.	Routine clinical follow up-surveillance.	Comparison of progressors vs. non-progressors	NA for the whole cohort - has Table 1 with different FU times for each case-control group		Porgression to HGD/EAC. Expression of p53, Ki67 and toll-like receptor 5 (TLR5) between progressors and non-progressors	p53 associated with a high risk of progression (OR (6.7, 95% 1.8–24.6)). The previously suggested markers Ki67 and TLR5 were not associated with disease progression.	p53, Ki67 and toll-like receptor 5 (TLR5). Based on Northern and Central Finland patients
Pinto (2022)	cross-sectional study	NA	NA	BE (n = 19) and EAC (n = 145) samples	NA	NDBE and EAC samples	investigation of prevalence of mutations in BE/EAC	investigation of prevalence of mutations in BE/EAC	NA		Investigation of the prevalence of core genetic (TP53 mutations and microsatellite instability (MSI) status) and epigenetic (DNA promoter hypermethylation of APC, CDKN2A, MGMT, TIMP3 and MLH1) modifications in a cohort of non-dysplastic BE and EAC samples.	Overall, none of the BE harbored TP53 mutations, whereas 30 out of 108 (28%) EAC samples carried mutations. None of the BE lesions and seven out of 108 tumors (6%) showed MSI. The promoter DNA methylation status of four genes (APC, CDKN2A, MGMT and TIMP3) was evaluated. For each gene, the promoter methylation frequency was significantly higher in BE or tumor samples compared to the tumor adjacent normal counterpart (p<0.05 for all). APC, CDKN2A, MGMT and TIMP3 promoter hypermethylation is frequently seen in both BE and EAC (21–89%), as well as in a subset of adjacent normal samples (up to 12%). Overall, 16% BE and 7% EAC samples showed hypermethylation of all four genes simultaneously.	Prevalence of core genetic (TP53 mutations and microsatellite instability (MSI) status) and epigenetic (DNA promoter hypermethylation of APC, CDKN2A, MGMT, TIMP3 and MLH1)

Study	Design			N	Age	Population	Method	Comparison	Follow-up	Outcome	Results	Conclusion
Roumans (2022)	multicenter cohort study	NA	BA	631 (3276 endoscopies were performed)	median age of 60 years (IQR 53–69).	Consecutive BE patients from 15 Dutch hospitals. Histologically confirmed intestinal metaplasia in biopsies obtained from columnar lined epithelium in the esophagus, BE≥2 cm, and absence of a history of HGD/EAC. To exclude prevalent cases of neoplasia at baseline, only BE with ≥6 months of FU in the study without HGD/EAC development were selected for this analysis.	Routine clinical follow up-surveillance.	Comparison of progressors vs. non-progressors	median FU time was 6.8 years (IQR 4.9–9.8). Overall, 4475 person-years were observed	Progression to HGD/EAC. Usefulness of dynamic use of dysplasia and p53 (overexpression or loss of expression) and SOX2 (loss of expression) immunohistochemistry to predict outcome and therefore inform on follow up strategies for each individual patient, rather than having fixed strategies for all. A multivariate joint model was used in which longitudinal data of the risk of aberrant measurements of a biomarker are combined with the risk of neoplastic progression. Secondly, these models were combined with a time-varying Cox proportional hazards model in the framework of a multivariate joint model, to estimate the dynamic risk of neoplastic progression. Both the Cox model and the multivariate joint model were adjusted for baseline values age and gender, and time-varying dichotomous covariates BE length and esophagitis. The risk of neoplastic progression was estimated based on the longitudinal evolution of LGD, p53, and SOX2. For the results of the biomarkers of the multivariate joint model the value (representing the current risk of neoplastic progression, based on all previous measurements) and the accumulated effect of this biomarker (representing the overall risk of neoplastic progression, based on the history of the measurements of biomarkers) are reported.	The risk of aberrant expression of p53 increased in time (OR 1.17, p<0.01), with older age (OR 1.82, p<0.01), and male gender (OR 0.23, p<0.01 (ref. male)), but also with a long segment BE (OR 2.77, p<0.01). Loss of SOX2 expression was not influenced by any of these factors. An increased risk of developing HGD/EAC during surveillance was associated with aberrant expression of p53 (HR value 1.26, p<0.01) or SOX2 (HR value 1.43, p<0.01). The results of LGD (HR value 1.02, p = 0.78 & HR accumulated effect 1.02, p = 0.12) were not statistically significant. Consequently, there may be an association with an increased risk of neoplastic progression and LGD, but the association with aberrant expression of p53 or SOX2 is considerably larger. This indicates that multiple successive aberrant measurements of p53 or SOX2 increase the risk of developing HGD/EAC. If the model will be used in an online application, it will include all demographic and clinical variables of that individual patient, as well as the longitudinal evolutions of histological diagnosis, p53, and SOX2 to estimate the neoplastic progression risk. The risk estimations can be updated at every surveillance endoscopy, based on new additional measurements of histological diagnosis and immunohistochemistry. This results in dynamic risk estimations for each patient, according to its individual patient characteristics. These risks for the development of HGD/EAC gradually evolve from low risk to high risk. Each risk will have its consequence in surveillance, for example for low risk the interval can be chosen to remain long (e.g. three years), for medium risk the interval could be shortened (e.g. towards one year), and for high risk patients endoscopic eradication therapy (EET) may be applied. Optimism-adjusted estimates of the AUC were between 0.80 and 0.88 at different time points, indicating good performance of the model.	p53, SOX2. This is the first 'proof-of-principle' study to explore a "dynamic" approach to risk assessment, and therefore personalize BE surveillance/management.

<mark>cost-effectiveness biomarkers</mark>

Study	Design			N	Age	Population	Method	Comparison	Follow-up	Outcome	Results	Conclusion
Rubestein (2005)	cost-effectiveness study - Markov decision modeling and simulation	NA	NA	NA	50 yo	50-year-old Caucasian men with gastro-oesophageal reflux, who were monitored until age 80.	simulation study	Strategies: (1): observation only, (2) current practice (dysplasia guided surveillance based on ACG 2002 guidelines), (3) surveillance every 3 months for patients with a positive biomarker (biomarker guidedsurveillance), and (4) oesophagectomy immediately for a positive biomarker (biomarker-guided oesophagectomy)	NA	The primary outcome was the threshold cost and performance characteristics needed for a biomarker to be more cost-effective than current practice.	Regardless of the cost, the biomarker needs to be at least 95% specific for biomarker-guided oesophagectomy to be cost-effective. For biomarker-guided surveillance to be cost-effective, a $100 biomarker could be 80% sensitive and specific. Biomarkers predicting the development of oesophageal adenocarcinoma would need to be fairly accurate and inexpensive to be cost-effective.	This is a study to identify the characteristics of a good biomarker that can be useful as a cost-effective method of BE surveillance/risk stratification
Gordon (2014)	cost-effectiveness study - Markov decision modeling and simulation	NA	NA	NA	hypothetical cohort based on BE surveillance program in Adelaide, Australia	Interventions: (1) No surveillance, (2) 2-yearly endoscopic surveillance of patients with NDBE and 6-monthly surveillance of patients with LGD (2005 UK guidelines), (3) a hypothetical strategy of biomarker-modified surveillance.	cost-effectiveness in different scenarios (nothing, following the 2005 U.K. British Society of Gastroenterology Guidelinesl or using biomarkers)	a 5-year decision-analytic model and traces patients from a diagnosis of HGD or adenocarcinoma	Interventions: (1) No surveillance, (2) 2-yearly endoscopic surveillance of patients with non-dysplastic BE and 6-monthly surveillance of patients with low-grade dysplasia, (3) a hypothetical strategy of biomarker-modified surveillance.	Compared with no surveillance, surveillance produced an estimated incremental cost per QALY ratio of $60,858. This was reduced to $38,307 when surveillance practice was modified by a hypothetical biomarker-based strategy. Sensitivity analyses indicated that the likelihood that surveillance alone was cost-effective compared with no surveillance was 16.0% and 60.6% if a hypothetical biomarker-based strategy was added to surveillance, at an acceptability threshold of $100,000 per QALY gained. Endoscopic surveillance of patients with non-dysplastic BE is unlikely to be cost-effective for the majority of patients and depends heavily on progression rates between dysplasia grades. However, strategies that modify surveillance according to cancer risk (e.g. through biomarkers) might be cost-effective, provided that high-risk individuals can be identified and prioritized for surveillance.	addition of biomarker(s) in risk assessment and follow up strategy may be cost-effective
Das (2016)	cost-effectiveness study - Markov decision modeling and simulation	NA	NA	NA	mean 50 yo	Hypothetical cohort of men with mean age of 50 year old and diagnosed with NDBE at baseline EGD follwed by ACG guidelines 2011	simulation study	strategy I, natural history without surveillance; strategy II, surveillance per current guidelines; strategy III, ablation for all patients; strategy IV, risk stratification with use of a biomarker panel to assess genomic instability (i. e., mutational load [ML]: no ML: minimal surveillance; low ML: standard surveillance, high ML: ablation).	a calculated period of 174,853 person-years	Comparison of different strategies, including a strategy guided by biomarkers. Mutational load (ML) include loss of heterogeneity (LOH) in 17p (TP53), 9p (CDKN2A), 1p (CMM1, L-myc), 3p (VHL, HoGG1), 5q (MCC, APC), 10q (PTEN, MXI1), 17q (NME1), 18q (DCC), 21q (TFF1, PSEN2), and 22q (NF2) genomic loci, and presence of microsatellite instability (MSI) at these loci. ML was assessed with BarreGen and PathFinderTG in esophageal biopsy specimens from patients with NDBE	Strategy IV provided the best values for quality-adjusted life years (QALYs), ICER, and INHB in comparison with strategies II and III. Results were robust in sensitivity analysis. In a Monte Carlo analysis, the relative risk for the development of cancer in the patients managed with strategy IV was decreased. Critical determinants of strategy IV cost-effectiveness were the complete response rate, cost of ablation, and surveillance interval in patients with no ML. The use of ML to stratify patients with NDBE by risk was the most cost-effective strategy for preventive EAC treatment.	Mutational load (ML) was assessed with BarreGen and PathFinderTG in esophageal biopsy specimens from patients with NDBE (LOH) in 17p (TP53), 9p (CDKN2A), 1p (CMM1, L-myc), 3p (VHL, HoGG1), 5q (MCC, APC), 10q (PTEN, MXI1), 17q (NME1), 18q (DCC), 21q (TFF1, PSEN2), and 22q (NF2))

Study	Type			Population	Age	Cohort	Model	Comparison	Time frame	Objective	Results	Conflicts/Funding
Hao (2019)	cost-effectiveness study - Markov decision modeling and simulation	NA	NA	Hypothetical cohort of 10,000 individuals with BE and NDBE, IND or LGD was used	Hypothetical cohort was 69.1 yrs old	Hypothetical cohort of 10,000 individuals with BE and NDBE, IND or LGD was used	Cost-effectiveness Markov decision model was constructed in Excel using Palisade DecisionTools Suite (Palisade Corporation, Ithaca, New York) for the disease progression of patients with BE and their surveillance and treatment protocol over a 5-year time frame.	Comparison between a TissueCypher-guided management that stratified patients with BE in different risk levels for progression towards HGD/EAC *versus* standard of care (SOC)	Simulation for a 5-year period	Compare cost and quality-adjusted life-years (QALYs) from the perspective of a US health insurer with care delivered by an integrated health system.	Base-case model results for a 5-year period comparing the Assay-directed care to the standard of care (SOC) estimated an incremental cost-effectiveness ratio (ICER) of $52,483/QALY in 2012 US dollars. Assay-directed care increased the use of endoscopic treatments by 58.4%, which reduced the progression to HGD, EAC and reduced EAC-related deaths by 51.7%, 47.1%, and 37.6%, respectively, over the 5-year period. A surveillance interval of 5 years in BE patients scored low-risk by the Assay, independent of pathologic diagnosis (NDBE, IND, LGD), resulted in a 16.6% reduction in endoscopies. Targeting of endoscopic therapies to patients scored high-risk by the Assay increased the number of endoscopic treatments by 58.4%, which resulted in reducing the incidence of HGD, EAC and EAC-related deaths by 51.7%, 47.1%, and 37.6% over 5 years, respectively. Sensitivity analysis indicated that the probability of the Assay being cost-effective compared to the SOC was 57.3% at the 100,000/QALY acceptability threshold. While the Assay strategy was estimated to add cost during the initial 3 years of adoption, it was estimated to lower future costs and improve outcomes due to reduced surveillance in low-risk patients, and early treatment in high-risk patients over a 5-year period.	The authors of this study hold equity ownership or stock options in Cernostics, Inc., the commercial entity that developed the proprietary TissueCypher technology used in this study. The study was partially funded/supported by industry.
Vissapragada (2021)	an overview of cost-effectiveness modalities	NA	NA	NA (evaluation of many different CE studies)	NA	cost-effectiveness studies	NA	NA	NA	cost-effectiveness in BE surveillance	Reducing the number of surveillance endoscopies performed by clinical or biomarker-based risk stratification was reported to have high chance of being cost-effective between $10,000/QALY and $45,000/QALY (converted 2018 USD value) willingness-to-pay threshold. When using a guidance from biomarkers, even with significant heterogeneity in model structure, inputs, and strategies, all studies applying a risk stratification EGD surveillance showed cost-effectiveness of their primary strategy. Roughly, the excluded proportions of patients were 85.5% for Gordon (2014), 77% for Das (2016), 77% for Hao (2019). In other words, more than a third of patients currently in guideline-recommended EGD surveillance programs may need to be excluded to make it cost-effective.	an overview of cost-effectiveness modalities

	Table 17		role of tumor budding							

Author (year)	Methods		Population			Intervention			Outcomes		Remarks
	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	Remarks
Dhingra-2021-Diagnostic Pathology	Retrospective	No	No	42 patients (from a cohort of 52), 8 with tumor budding	Mean age 68 years (range: 46–83)	Patients with ESD specimens showing BE invasive adenocarcinoma on final pathologic examination	Pathological staging of ESD specimens	No	No data	Presence of dysplasia, invasive adenocarcinoma, peritumoral inflammation, desmoplasia, lymphovascular and perineural invasion; tumor differentiation, depth of invasion, morphology, and budding; and margin status for dysplasia or carcinoma.BE, dysplasia, or carcinoma on follow-up endoscopic biopsies, pathology of follow-up esophagectomy, presence of metastasis. Tumor budding was assessed at the advancing tumor edge (peritumoral) and scored as 1 (low), 2 (inter- mediate), or 3 (high) on hematoxylin and eosin-stained sections	Peritumoral budding (PTB) present in 8 cases; budding was as low-grade in 1 case and intermediate- or high-grade in 7 cases. Tumor budding was preset in 44% of patients with sub- mucosal adenocarcinoma. No peritumoral tumor budding was observed in patients with intramucosal adenocarcinoma.	Tumor budding when present was associated with other high-risk features. No possible to conclude if PTB was independent risk factor for metastasis.Low n (n=8).
Brown M-2010-*Histopathology*	Two-center retrospective study	No	No	69 SCC, 287 Adenoc	Not available	transthoracic esophagogastrecto my with lymphadenectomy for esophageal or gastroesophageal cancer	Tumour budding at the invasive front was assessed histologically.	Inflammatory response at the invasive front were assessed histologically.	The 30-day mortality was 18/356 (5.1%). Median follow-up of the surviving 87 patients was 74 months (range 21–148 months). Median time to death in patients who died (n = 269) was 13 months (range 0–136 months).	Tumor budding has prognostic significance in many carcinomas and is defined as the presence of detached isolated single cells or small cell clusters up to 5 cells at the invasion front (peritumoral budding [PTB]) or within the tumor (intratumoral budding [ITB]).	Intraobserver agreement of scores in 90% of cases and a K of 0.58. The median number of tumour buds was four (range 0–50). Of the 356 cases, 184 had fewer than five tumour buds. TB was associated with poorly differentiated tumours (P = 0.0001), higher T stage (P = 0.0001), lymph node metastases (P = 0.0001), incomplete excision (P = 0.0001), higher overall TNM stage (P = 0.0001) and low inflammatory response (P = 0.0001). Not associated with patient age (P = 0.477), tumour type including adenocarcinoma and SCC (P = 0.089), tumour length (P = 0.967) or use of neoadjuvant chemotherapy (P = 0.052). Low-grade TB: overall survial 31 months; high-grade TB: OS 15 months (p < 0.0001) TB independent significant prognostic factor on multivariate analysis. Prognostic impact of TB independent of histologic type (SCC, p = 0.021; AC, p = 0.0001), age, N-stage, overall stage	No separate analysis between SCC and AC was performed.
Thies-2016-Human Pathology	Two-center retrospective study	No	No	162 patients (from a cohort of 200 cases)	NA	Primary resected EACs of all stages.	Standard hematoxylin and eosin (H&E) stains versus pancytokeratin for detection and quantification of TB	Pancytokeratin for detection and quantification of TB. Tumor staging; lymphatic, vascular, and perineural invasion; tumor grading.	NA	Tumor budding is defined as the presence of detached isolated single cells or small cell clusters up to 5 cells at the invasion front (peritumoral budding [PTB]) or within the tumor (intratumoral budding [ITB])	High PTB and ITB rates were seen in more advanced tumor categories (P < .001 each); tumors with lymph node metastases (P < .001/P = .002); and lymphatic, vascular, and perineural invasion and higher tumor grading (P < .001 each). Survival analysis showed an association with worse survival for high-grade ITB (P = .029) but not PTB (P = .385). However, in multivariate analysis, lymph node and resection status, but not ITB, were independent prognostic parameters.	Pancytokeratin staining was superior to hematoxylin and eosin staining for the detection of buds with substantial to excellent interobserver agreement and used for subsequent analysis, with intraclass correlation coefficient (ICC) of 0.93. Higher budding counts were identified by using the pancytokeratin stains compared to the H&E stains (P b .001; range, 1.54-fold for 1 HPF ITB to 1.7-fold for 10 HPFs both ITB and PTB).

Study	Design			N	Age	Tumor type	Aim/Assessment	Method	Follow-up	Tumor bud definition	Results	Conclusion
Landau-2014-Mod Pathol	Retrospective	No	No	210 surgically resected therapy-naïve esophageal adenocarci-nomas.194 were included in the survival analysis. 93 with TB	66 years (median) (interquartile range 60–74 years) a	Superficial (stage T1) esophageal adenocarcinomas	Extent of tumor budding (none, focal, and extensive)	depth of invasion (intramucosal versus submucosal), angiolymphatic invasion, tumor grade and tumor size.	44 months (median)	A tumor bud is defined as a detached cluster of fewer than 5 cells at the invasive front of a tumor. extensive tumor budding (tumor budding in Z3 20X microscopic fields). Tumor budding. None. Focal (1–2 budding fields). Extensive (>= 3 budding fields)	41% (24 out of 59) of tumors with extensive tumor buddingwere metastatic to regional lymph nodes, compared with 10% (12 out of 117) of tumors with no tumor budding, and 15% (5 out of 34) of tumors with focal tumor budding (Po0.001). When controlling for all pathologic risk factors in a multivariate analysis, extensive tumor budding remains an independent risk factor for lymph node metastasis with a 2.5-fold increase (95% CI ¼ 1.1–6.3, P ¼ 0.039) in the risk of nodal metastasis. Extensive tumor budding was associated with a 3.3-fold increased risk of death (95% confidence interval ¼ 1.5–7.4, P ¼ 0.004), after controlling for these other prognostic variables in the multivariate analysis. Extensive tumor budding was associated with a 3.2-fold increased risk of recurrence (95% confidence interval ¼ 1.4–7.0, P ¼ 0.005), independent of T and N stage in the multivariate analysis.	Tumor budding is na independent risk factor for nodal metastasis and survival in T1 ADC
Lohneis-2021-Virchows Archiv	Retrospective	No	blinded to clinical outcome	104 surgically resected therapy-naïve esophageal adenocarci-nomas	Mean age 68 years (range: 39–86)	Resected esophageal adenocarcinoma. Adenocarcinomas with signet ring cell morphology were excluded	1)To determined the interclass correlation coefficient (ICC) between 3 independent observers in assessment of TB. 2)To compare the use of H&E and cytokeratin stained slides using the ITBCC (International Tumor Budding Consensus Conference) methods. 3) Assess cytokeratin-, H&E-tumor budding, age, sex, pT stage, pN stage, and UICC on survival	H&E Vs cytokeratin for TB assessment	NA	1)The interclass correlation coefficient between the three observers was 0.831 (95% CI: 0.689–0.915), indicating good to very good interobserver correlation. 2)The mean count of tumor buds per HPF on H&E stained slides was 8.5 per high power field, the median 5 (range 0– 35). 50, 11, and 43 tumors could be assigned to budding groups Bd1, Bd2, and Bd3, respectively. The number of tu- mor buds per HPF in the cytokeratin stained cases was higher than observed with H&E staining. The mean count was 11.6 per high power field, the median 8 (range 0–54). Thirty-nine, 15, and 50 tumors were assigned to budding groups Bd1, Bd2, and Bd 3, respectively.A tumor bud was defined as a single tumor cell or (non-glandular) clusters of up to four tumor cells. Budding was grouped according to ITBCC into Bd 1 (0–4 buds), Bd 2 (5–9 buds), and Bd 3 (10 or more buds).	An increased tumor bud count was a signif- icant negative prognostic marker for OS in all studied patients (H&E: HR = 1.05 (95% CI 1.029–1.073), p < 0.001; cytokeratin: HR = 1.073 (95% CI 1.045–1.101), p < 0.001). Low budding tumors significantly associated with a longer OS. Median OS was 202.2/202.2 (H&E/ cytokeratin) months for Bd1 tumors, 85.8/ median not reached (H&E/cytokeratin) months for Bd2 tumors, and 19.6/11.3 (H&E/cytokeratin) months for Bd3 tumors (p = 0.002/ p < 0.001). In a multivariate survival analysis, including ITBCC bud- ding group determined by H&E, age, sex, pT stage, pN stage, and UICC stage, tumor budding (HR = 1.63 (95% CI 1.15– 2.31), p = 0.006), pT stage (HR = 2.13 (95% CI 1.10–4.13), p = 0.025) and UICC stage (HR = 1.92 (95% CI 1.14–3.24), p = 0.015) retained their prognostic status on OS. In a multi- variate analysis including budding group determined by cytokeratin, age, sex, pT stage, pN stage, and tumor grade, tumor budding was not an independent prognostic factor (HR = 1.20 (95% CI 0.83–1.74), p = 0.324).	Analysis of tumor buds according to the ITBCC in esophageal adenocarcinomas needs further verification by prospective studies in order to be considered a prognostic tool for survival prediction. The main benefit would be the easy and inexpensive implementation into routine diagnostic processes.
Nowak-2013-Lab Invest CONGRESS ABSTRACT	Retrospective	No	No	42 patients	Mean age 64.4 (range 46-87)	T1 EAC treated by primary surgical resection	Assessemnt of tumor size, grade, type, depth of invasion, lymphovascular invasion, lymphnode metastasis, presence and extent of TB	NA	NA	The presence of lymphnode metastasis was was associated with decreased time for recurrence (HR 7.75, p value 0.041). High tumor grade correlatedwith reduced overall survival (HR 3.06, p value 0.015) and with tumor recurrence (p value 0.012).	TB was a strong predictor of recurrence (HR 14.21, p 0.022). The presence of >=10 buds per 20x field was associated with 20 fold increase in risk of tumor recurrence (HR 19.99, p 0.007). The amount of TB was correlelated with nodal metastases (p< 0.001)	Abstract, scareced data, no multivariable analyses.
Lohneis-2022-Pathology - Research and Practice	Retrospective	No	No	253 of 273 patients included	Media 66 (range 39–86)	neoadjuvant treated oesophageal with a poor response (> 10% vital residual tumour) to neoadjuvant therapy.	Prognostic significance of tumour budding in 278 oesophageal adenocarcinomas with a poor response (> 10% vital tumour cells, minor regression) to neo-adjuvant therapy followed by surgery	NA	NA	Assessed the correlation of the clinicopathological pa- rameters patients age (≤65 and >65 years), sex (male and female), ypT classification (ypT1/2 and ypT3/4), nodal status (negative nodal status (ypN0) and positive nodal status (ypN1–3)), UICC tumour stage and grading of the tumour before neoadjuvant treatment (G1/2 and G3) with budding groups	Significant positive correlation (p < 0.05) between the budding group, ypN stage and UICC tumour stage. An increased tumour bud count was a significant negative prog-nostic marker for OS in all studied patients (HR = 1.039 (95% CI 1.012–1.066), p = 0.004). Low budding tumours associated with a significantly longer OS. Median OS was 26.19 months for Bd 1 tumours, 18.40 months for Bd 2 tumours and 14.98 months for Bd 3 tumours (p = 0.032) (Fig. 2B). In a multivariate survival analysis, including ITBCC budding group determined by H&E, age, sex, ypT stage, ypN stage and UICC stage, only tumour budding (HR = 1.63 (95% CI 1.15–2.31), p = 0.006), ypN stage (HR = 2.13 (95% CI 1.10 – 4.13), p = 0.025) and UICC stage (HR = 1.92 (95% CI 1.14 – 3.24), p = 0.015) retained their prognostic status on OS.	Some patients went through chemoradiotherapy (70%), others only through chemotherapy (30%)
PICO search string:			(adenocarcinoma) AND (esophagus) AND (budding); (barrett) AND (budding); (esophagus) AND (budding); 78 results. Manuscripts bibliography search: 1 more paper and one more abstract. March 2022									

Table 18	management of HGD					
TITLE	AUTHORS / JOURNAL	STUDY DESIGN	STUDY POPULATION	CONTROL POPULATION	OUTCOME	REMARKS
Radiofrequency ablation in Barrett's esophagus with dysplasia.	Shaheen NJ, Sharma P, Overholt BF, Wolfsen HC, Sampliner RE, Wang KK, et al N Engl J Med. 2009 May 28;360(22):2277-88. PMID: 19474425	RCT	120 Barrett's patients, 63 had HGD, 42 treated with RFA	21 sham control (Barrett's with HGD) not treated with RFA	CR-IM 73.8%, CR-D 81%; Among patients with high-grade dysplasia, 19.0% of those in the control group had progression to esophageal cancer, as compared with 2.4% of those in the ablation group (P=0.04).	
Late Recurrence of Barrett's Esophagus After Complete Eradication of Intestinal Metaplasia is Rare: Final Report From Ablation in Intestinal Metaplasia Containing Dysplasia Trial.	Cotton CC, Wolf WA, Overholt BF, Li N, Lightdale CJ, Wolfsen HC, Pasricha S, et al. Gastroenterology. 2017 Sep;153(3):681-688. PMID: 28579538	5-year FU of above mentioned RCT			Incidence of dysplasia recurrence was 7.3 per 100 personyears for baseline HGD	
Photodynamic therapy with porfimer sodium for ablation of high-grade dysplasia in Barrett's esophagus: international, partially blinded, randomized phase III trial.	Overholt BF, Lightdale CJ, Wang KK, Canto MI, Burdick S, Haggitt RC, et al. Gastrointest Endosc. 2005 Oct;62(4):488-98. PMID: 16185958	international, partially blinded, randomized phase III trial.	208 patients, 138 PDT	70 surveillance	eradication of HGD in 77% in the treatment group and 38% in the surveillance group; progression to cancer 13% in PDT group vs 28% in the surveillance group (p 0.006)	
Five-year efficacy and safety of photodynamic therapy with Photofrin in Barrett's high-grade dysplasia.	Overholt BF, Wang KK, Burdick JS, Lightdale CJ, Kimmey M, Nava HR, et al. Gastrointest Endosc. 2007 Sep;66(3):460-8. PMID: 17643436	Five-year efficacy and safety of aboven mentioned RCT			5 year FU study of RCT above: cancer progression 15% in PDT group, 29% in surveillance group. (P=.004)	
Efficacy and durability of radiofrequency ablation for Barrett's Esophagus: systematic review and meta-analysis.	Orman ES, Li N, Shaheen NJ Clin Gastroenterol Hepatol. 2013 Oct;11(10):1245-55. PMID: 23644385	Systematic review and meta-analysis	Total 3802 patients, 31% HGD; treated with RFA		CR-IM 68%, CR-D 85%, progression to EAC 0,4% annual risk	
Efficacy of Cryotherapy as a Primary Endoscopic Ablation Modality for Dysplastic Barrett's Esophagus and Early Esophageal Neoplasia: A Systematic Review and Meta-Analysis.	Raseen Tariq R, Enslin S, Hayat M, Kaul V. Cancer Control. 2020 Jan-Dec;27(1):1073274820976668. PMID: 33297725	Systematic review and meta-analysis	Total 405 patients, not clear how much HGD; treated with cryo-ablation		in the high-quality studies a pooled proportion of CE-D 91.3% and pooled proportion of CE-IM 71.6% (95% CI, 59.0-82.9, I 2 = 80.9%).	Recurrence rates only assessed in a small number of studies

Table 19					LGD treatment vs surveillance							

Author (year)	Methods			Population			Intervention			Outcomes		
	Design	Randi msation / blinding		N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow- up	Outcome measures (+ definitions)	Outcomes of interest	Remarks
Wolfson 2022 GIE	prospecti ve registry study 2008-2018	no	no	225 (19% of total cohort)	67.2	LGD confirmed by two expert pathologists	RFA	NA	7856 patient years, they report 10yrs, but median FU in months is not mentioned	CR-D: 2 endoscopies with biopsies without dysplasia; CR- IM: 2 endoscopies with biopsies without IM and BE tongues <3cm	CR-IM 65.2%; CR-D 87.2%; 0,7% of patients initially treated for LGD developed invasive cancer; crude incident range 0.52 per 100 patient years	treatment in expert center; patients who had treatment >24months were excluded from analysis
van Munster 2022 Gut	prospecti ve registry study 2008-2018	no	no	375 treatment cohort; 306 in durability cohort	65	LGD confirmed by expert pathologist	RFA	NA	median 43months from first treatment	Endoscopic eradication and no IM in biopsies from the cardia	outcomes not reported for different types of initial dysplasia	treatment in expert centers; all patients treated in the Netherlands included; no report on only LGD patients, therefore I think exclude for this PICO
Barret 2021 Gut	RCT	yes	no	82 (40 RFA and 42 surveillan ce)	62.3	LGD confirmed by 2 pathologists and by central reading by 1 pathologist	RFA (junction was always ablated, but not always with the focal device)	surveillance 12, 24 and 36 months after randomization	1 and 3 years	prevalence of LGD 3 years after randomization; secondary outcomes were rate of neoplastic progression at 1 and 3 years; LGD at 1 year; CR-D and CR-IM	Neoplastic progression in 13,5% in RFA group vs 26.2% in the surveillance group; CR-IM 35% in RFA group vs 0% in surveillance group; Prevalence LGD 34.3% in RFA group and 58.1% in surveillance group; complication at least one adverse event in 7/37 patients after RFA, mostly mild stricture, dysphagia or bleeding in 4,5%	Extremely low rates of CR-IM, 14 participating centers, so some centers may have treated <3 patients; also patients with a history of HGD/cancer in their Barrett's were included; in the RFA group 37/40 patients had RFA, only 33 had a second RFA session
Klair 2021 Dig Dis	Systemat ic review and meta-analysis	na	na									excluded, all 4 studies in this systematic review are included in this data extraction sheet
Canto 2020 AJG	prospecti ve multicent er trial	no	no	29 with LGD	65	BE with LGD, HGD or Ca, after central pathology review; treatment naive	cryoballoon ablation	NA	12 months from start treatment	CE-D at 12 months; CE-IM at 12 months; progression of LGD to HGD/Ca at 12 months; buried Barrett's; number of CBA treatments to achieve CE-D/IM at 12 months; complications	CE-D 76% (ITT) and 97% (PP); CE-IM 72% (ITT) and 91% (PP) > no significant differences in CE-D and CE-IM rates when stratified by dysplasia grade. No progressors in the LGD group. Serious adverse events in 1% of treatments. Stricture 12.5%.	ook biopten in de cardia genomen; APC touch-up voor kleine eilandjes Barrett
Wronska 2021 Endoscopy	RCT	yes	no	71	60/62	LGD confirmed by expert pathologist	APC	APC 90 vs 60W and PPI 40mg vs 120mg	6wks and 2yrs after APC treatment	complete ablation rate at 6 weeks; secondary adverse event rate, complete ablation rate after 2 years, recurrence rates	Ccomplete endoscopic and histologic ablation at 6 weeks after 6 weeks in 70%; complete endoscopic and histological response at 2 years confirmed in 62%; No progression to HGD or cancer in any of the patients.	
Phoa 2014 JAMA	RCT	yes	no	136	63	LGD confirmed by expert pathologist	RFA until endoscopic eradication	surveillance according to guideline	median FU 36 months, all patients followed at least 24 months.	neoplastic progression to HGD or cancer during 3yr FU; secondary outcomes were CR- D and CR-IM and adverse events	ablation reduced risk of progression by 25%; risk of progression to cancer alone by 7.4%; CR-D 92.6% and CR-IM in 88.2%; 12% strictures resolved with 1 dilation	only 1 diagnosis of LGD was required for inclusion; 28% of the control group had no dysplasia during FU
Phoa 2017 GIE	cost effectiven ess SURF	see Phoa 2014										Cost-effectiveness spin-off of the SURF-trial, at a willingness to pay 40,915 per prevented event of progression, RFA is efficient (> based on high eradication rates in the SURF trial)
Pouw 2020 GIE	Long term FU of RCT	yes	no	136		LGD confirmed by expert pathologist	RFA until endoscopic eradication	surveillance according to guideline	73months		Sustained clearance of BE in 91% and LGD in 96% of patients	RFA resulted in an absolute risk reduction of 32.4%; NNT 3.1; after end of RCT, surveillance patients could also be ablated
Westerveld 2020 EIO	Systemat ic review and meta-analysis						cryoballoon ablation					excluded, review
Omidvari 2020 CGH	simulatio n study										in men the optimal strategy was treatment of LGD after confirmation by repeat endoscopy ($ 53,044/QALY) ; for women the same	Cost-effectiveness

Study	Type			N	Age	Population	Intervention	Comparator	Follow-up	Outcome	Results	Remarks
Song 2020 Dis Es	single center retrospec tive cohort study	no	no	69	65.2	BE with confirmed LGD by an expert GI pathologist	na		surveillance	3.74 years		16/69 (23.2%) of patients developed HGD/EAC; for persistent LGD 6.44/100 rate compared to 2.61/100 rate in case of non-persisting LGD	
Rosmolen 2019 GIE	QoL from RCT	see Phoa 2014											QoL in ablated and surveyed patients was comparable, patients in the RFA group had less concerns and less threatening view of their disease (however, the other group knew of course that they were not treated)
Inadomi 2019 TGH	review												No relevant papers in this review, mix of strategies for NDBE and LGD
Pollit 2019 Cur Med Res Op	Cost- effectiven ess												I could not differentiate outcomes for LGD and HGD separately
Tan 2018 UEGJ	retrospec tive cohort study	no	no	21		LGD confirmed by at least 2 GI-pathologists	RFA				stratification on chance of reaching CR-D and CR-IM based on baseline histology	CR-D 95.2%, CR-IM 71.4%	
Pandey 2018 End	Systemat ic review and meta-analysis												review, no new studies identified; some studies have long- term FU in the mean while and those were included
Kahn 2018 Dis Es	retrospec tive single center cohort												
Qumseya 2017 Am J Gastr	Systemat ic review and meta-analysis												
Duits 2017 Gastro	cohort study	no	no	255	63	initial diagnosis LGD	revision by expert pathology panel	na		42months	progression to neoplasia	number of pathologists confirming LGD was strongly associated with progression to neoplasia; risk increased when all 3 agreed (OR 47%); repetitive LGD also increased risk (OR 9.3%); multifocal LGD was not significantly associated with progression	study indicating relevance of revision and of a repetitive diagnosis of LGD when assessing risk of progression and thus the possible prophylactic value or ablation
Small 2015 Gastro	retrospec tive cohort study	no	no	45 RFA, 125 surveillan ce		LGD confirmed by expert pathologist	RFA		surveillance	889 and 848 days	assess association between progression and RFA	annual rates of progression to HGD/Ca was 6.6% in the surveillance group and 0.77% in RFA group; risk of progression was significantly lower after RFA HR 0.06	
Jankowski 2015 Am J Gastr	consensus Delphi												excluded, Delphi process
Almond 2014 BJS	meta- analysis												excluded, since this contains all different ablation techniques and includes studies <2000
Curvers 2010 Am J Gastr	cohort study			147		LGD in community setting	expert pathology review 2 pathologists			51.1 months	progression to neoplasia	15% of LGD diagnosis were confirmed; cumulative risk of progression HGD/Ca 85% compared to 4.6% if patients were downstaged (incidence rate if LGD was confirmed was 13.4% per patient per year vs 0.49% if downstaged)	full tekst available in PhD-thesis
Sharma 2008 End	prospecti ve cohort	no	no	10	66.9yrs	BE 2-6 cm with LGD at at least 2 endoscopies confirmed by 2 pathologists	only circumferential RFA 1 or 2 times; after 1 yr focal ablation for visible BE	na		2yrs	CR-D and CR-IM at 2yr follow- up with all biopsies negative	CR-D 100%, CR-IM 90%; no strictures	
Shaheen 2009 NEJM	RCT	yes	yes	42 RFA vs 22 sham	66/64yrs	1x LGD confirmed by expert pathologist, no nodules	cRFA 2x12 , followed by fRFA 2x? At 2-4-9 months	sham, no ablation		6 and 12 months endoscopy with biopsies	proportion of patients with CR-D and CR-IM at 12 months; secondary proportion of patients with progression	eradication of all IM (also in HGD patients) ITT 77% vs 2% NNT 1.3; eradication LGD 90% vs 23% NNT 1.5; progression LGD to HGD 5% vs 14% and progresion LGD to cancer 0 vs 0	no severe adverse events

	Table 20				**Ablation of residual BE**								
Author (year)	**Methods**				**Population**			**Intervention**				**Outcomes**	
	Design	Randomisation / blinding	N	Age	Inclusion criteria	Technology	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest		**Remarks**
Peerally (2019)	Clinical trial	Randomized	not	76	69.7	Patient age 18 to 85 years, histology HGD or T1a cancer with a maximum depth of invasion on ER of T1m3, EUS and CT or positron emission tomography-CT scan negative for locally advanced or metastatic disease	RFA or APC	Randomization was 1:1 to RFA or APC All patients received high-dose (twice daily) proton pump inhibitors; ER was performed at entry if not done within the previous 6 months.	–	1 year	Dysplasia clearance; BE clearance; adberse events; costs	Dysplasia clearance at 12 months: RFA 79.4% and APC 83.8% (OR 0.7; 95% CI: 0.2-2.6); BE clearance: RFA 55.8%, and APC 48.3% (OR 1.4; 95% CI, 0.5-3.6). Buried BE: 6.1% in RFA and 13.3% in APC Adverse events (including stricture rate) were similar: RFA 3/36 [8.3%] and APC 3/37 [8.1%]) cost: RFA cost $27491 more per case than APC.		Short follow-up - long-term BE clearance?
Subramaniam (2018)	Cohort	NA	NA	91	72 - 73	All patients referred for consideration of ablation therapy of dysplastic Barrett's esophagus were eligible	RFA	Focal and circumferential ablation was performed after initial follow up endoscopy postresection. Patients proceeded straight to RFA in the absence of any visible lesions. EMR VS ESD VS RFA alone	–	2 years	Dysplasia clearance rate; IM clearance rate	Dysplasia clearance rate: 96.3% in the ESD group, 88.4% in the EMR group, and all patients in the RFA alone group. IM clearance rate: EMR in 81.4% vs. ESD in 85.2%, but higher in the RFA alone group (90.5%).		Short follow-up - long-term BE clearance?
Pouw (2018)	Clinical trial	Randomized	Blinded	79	64-66	Patients aged 18–85 years were eligible if there was presence of a treatment-naive short segment Barrett's oesophagus eligible for treatment with the Barrx 90 device, or in the case of residual Barrett's oesophagus after initial (one or two sessions) circumferential radiofrequency ablation treatment, endoscopic resection, or a combination thereof.	RFA	Patients were randomly assigned (1:1) to the new simplified regimen (3 × 12 J/cm², without cleaning) or the standard regimen (2 ablations with 15 J/cm²), with variable block sizes of four, six, and eight patients, stratified by participating hospital. Focal radiofrequency ablation was done every 3 months, up to a maximum of three treatments, until all Barrett's oesophagus was eradicated.		9 months	Dysplasia and IM regression Number of ablation treatments to achieve eradication Adverse events	Dysplasia and IM regression: 32 (74%, 95% CI 59–87) in simplified protocol VS 0 (83%, 95% CI 67–94) in the standard protocol (p=0·34). Median Barrett's oesophagus surface regression: 98% (IQR 95–100) in the simplified regimen group VS 100% (97–100) in the standard regimen group. The difference between medians was 2% (95% CI −0·562 to 3·162) A median of 2 (IQR 1–2) focal radiofrequency ablation treatment sessions was required to achieve complete eradication of all Barrett's epithelium for both regimens. Stenoses requiring dilatation: 9% of 43 patients in the simplified VS 11% of 36 in the standard regimen group. Post-procedural bleeding requiring endoscopy: 2% in the simplified ablation group VS 8% patients in the standard group.		Short follow-up - long-term BE clearance?
Frei (2018)	cohort	NA	NA	48	66	Patients were detected by the intern endoscopic resection registry. Additionally we scanned the database of the Institute of Pathology applying the keywords "dysplasia" and "Barrett." Matching patients with endoscopic treatment were enclosed to the registry.	RFA	Remaining Barrett's metaplastic tissue was burned according to an RFA treatment protocol. Either the Halo-360 system for circumferential or the Halo-90 system for sectorial ablation was used at 3-month intervals until full eradication was achieved or persistent metaplasia was not removable. Escape-EMR was performed if any visible lesions appeared or if persistent dysplastic tissue was resistant to RFA.	–	41 months	Dysplasia and cancer eradication and recurrence	Complete eradication: 33/35 in HGD/T1a and 11/13 T1b in adenocarcinoma. Recurrence: 0/35 in HGD/T1a and 1/13 in T1b.		No data on IM eradication
Schwameis, (2018)	Cohort	NA	NA	40	66	All patients who had undergone endoscopic therapy forHGDorIMCbetween 2001 and 2010 at the University of Southern California. These patients had formed the basis of a prior report comparing the early outcomes of endotherapy to esophagectomy	RFA	Endotherapy consisted of ER and/or mucosal ablation, typically with radiofrequency energy (Barrx), but in some cases with the cryoballoon (C2 Therapeutics). When any nodule or lesion was present in the columnar mucosa ER was used as the initial step to remove and stage the lesion. In some patients with SSBE, complete ER was done of the entire segment. Mucosal ablation was preferred for flat, non-nodular Barrett's, particularly when long segments were present.	–	82 months	Complete IM eradication IM recurrence Overal survival	Complete resolution of IM: in 30 patients (83%) at a median of 21 months. Recurrence of IM: In 18 patients (60%), CRIM was maintained, whereas 12 patients developed recurrence at a median of 14 months. Additional endotherapy (n ¼ 11) led to CRIM again in 10 patients (83%). There were no cancer deaths when CRIM was achieved. Overall survival with endotherapy was 73% at 5 years and 67% at 10 years.		No comparison group
Vliebergh (2018)	cohort	NA	NA	295	65	The inclusion criterion was all patients undergoing RFA for curative eradication of BE.	RFA	Between February 2008 and January 2017, data from 7 different expert centers were prospectively collected in the registry.	–	25 months	Complete IM and dysplasia eradication IM recurrence adverse events	CR-IM= 88% and CR-D=93 % Sustained remission=65% Recurrence of IM was most frequently observed (59/227 patients, 26%), followed by direct recurrence of HGD (14 patients, 6%), LGD (6 patients, 2.5 %) and adenocarcinoma (1 patient, 0.5 %). Immediate complications occurred in 4% of all procedures and late complications occurred in 9% of all procedures.		No comparison group
Matos (2019)	SR and meta-analysis	NA	NA	1950	not stated	not stated	RFA	These were comparative studies involving RFA using the Halo technique (BARXX Medical), either with or without the use of an endoscopic resection, in patients with BE. Studies were included regardless of randomization status.		not stated	Risk difference (RD) for dysplasia	There was a significant difference between the two groups [RD 0.35(0.15, 0.56)]. There was no significant difference between the two groups [RD 0.03 (0.00, 0.05)].		Observational studies

Study	Type			N	Age	Inclusion criteria	Intervention	Procedure details		Follow-up	Outcomes	Results	Comments
Pollit (2018)	cost-effectiveness analysis	NA	NA	NA	NA	NA	RFA	A hypothetical cohort of United Kingdom (UK) patients diagnosed with BE entered the model. At time t=0, patients who do not receive EET enter the model in one of three health states: NDBE, LGD or HGD. Thereafter, individuals either remain in their previous health state, move between health states (including progression to EAC) or die based on a series of transition probabilities. All patients who transition to the EAC state undergo surgery with esophagectomy, from which a small proportion will die or are cured. In the treatment group, patients with either confirmed LGD or HGD are initially managed by EET (EMR being used in a proportion of patients prior to RFA as per most recent published studies	NA	NA	Incremental Cost-Effectiveness Ratio	EET for patients with LGD and HGD arising in BE is cost-effective compared to endoscopic surveillance alone (lifetime ICER £3,006 per QALY gained). The results show that as the time horizon increases, the treatment becomes more cost-effective. The five year financial impact to the UK NHS of introducing EET is £7.1m.	
Krajciova (2019)	cohort	NA	NA	136	64	Patients older than 18 years with confirmed diagnosis of BE (visible at least 1 cm long segment of metaplastic mucosa with IM) were included.	RFA	All patients with macroscopically visible lesions underwent ER (one patient underwent endoscopic submucosal dissection- ESD) that allowed histopathological staging. If the staging did not show an indication for surgery, physicians continued with endoscopic treatment with RFA.	NA	27.5 mobths	complete remission of neoplasia (CR-N), complete remission of intestinal metaplasia (CR-IM), recurrence and safety	CR-IM and CR-N were achieved in 77.9% (95% CI 70.0-84.6%) and 98.5% (95% CI 94.8-99.8%). Among 30 patients without CR-IM, 22 (73%) did not have macroscopic signs of BE. Recurrent neoplasia was detected in 4.5% of patients (6/134) and 15% (16/106) experienced a recurrence of IM at the level of the neo-Z-line. Diagnosis of cancer was an independent risk factor for recurrent IM after RFA (OR 7.0, 95% CI 1.6-30.9, p<0.0005)	No comparison group
Kaul (2020)	cohort	NA	NA	57	68.5	Patients who underwent SCT with the indication of dysplastic BE or early adenocarcinoma of the esophagus at our academic tertiary care referral center from August 2008 to February 2019.	Cryo	43.9% (25/57) of patients underwent radiofrequency ablation (RFA) during the course of treatment (e.g. after initiating SCT). 33.3% of patients (19/57) were RFA failures prior to SCT. 68.4% (39/57) of patients underwent endoscopic resection (EMR) prior to SCT	NA	4.8 years	complete eradication of intestinal metaplasia (CE-IM) and complete eradication of dysplasia (CE-D).	CE-IM was achieved in 75% (39/52) of patients, and CE-D in 98.1% (51/52). The mean durability of CE-IM in these patients is 3.4 years.	No subgroup analyses (with or withouT prebious RFA)
Knabe (2022)	cohort	NA	NA	154	64	Consecutive patients with neoplastic BE undergoing ablation after curative endoscopic resection (89.6%) or primarily were included into this prospective trial in 9 European centers. Inclusion criteria: 1. Patients aged .18 and ,85 years with BE-N following curative endoscopic resection of visible lesions (histologically low risk, i.e., T1m, G1/2, L0 V0 R0 basal) with planned complete BE eradication or primary ablation for low-grade intraepithelial neoplasia or macroscopically invisible high-grade intraepithelial neoplasia 2. BE length C > 1 cm and < 10 cm	Hybrid APC	Endoscopic resection was performed according to institutional standards and had to be performed in case of visible lesions. Primary ablation of BE-N was only allowed with low grade dysplasia or HGD histology and strictly invisible lesions. Argon plasma coagulation was performed using the H-APC probe after prior injection with saline using the same method as described before. Up to 5 ablation sessions were allowed for complete eradication of BE (initial complete eradication of intestinal metaplasia by definition including BE-associated neoplasia, documented by 1 negative endoscopy with biopsies.	NA	2 years	CE-IM in intention-to-treat (ITT) and per-protocol (PP) samples at 2 years rate of recurrence-free cases (sustained CE-IM) documented by negative follow-up endoscopies with biopsies and immediate/delayed adverse events.	Initial CE-IM was achieved in 87.2% of 148 cases in the PP analysis (ITT 88.4%) initial BE-associated neoplasia was 98.0%. On 2-year follow-up of the 129 successfully treated cases, 70.8% (PP) or 65.9% (ITT) showed sustained CE-IM; recurrences were mostly endoscopy negative biopsy-proven BE epithelium and neoplasia in 3 cases. Adverse events on a patient basis were seen in 6.1% including a stricture rate of 3.9% requiring a few dilatation sessions.	No comparison group
Magee (2020)	cohort	NA	NA	112	67	1. Diagnosed with Barrett's Oesophagus with intramucosal cancer, high grade dysplasia or low grade dysplasia. 2. Visible lesions removed by endoscopic resection (ER) prior to RFA. 3. Treated with the new BarrxTM 360 Express catheter as index RFA treatment 4. 3-month follow	RFA	In the UK and Ireland, the treatment protocol for BE related neoplasia constitutes initial removal of visible neoplastic lesions via endoscopic resection. The protocol for RFA following this is shown in Fig. 1, with endoscopies planned at 3 monthly intervals with further RFA treatment given when there is visible Barrett's or Barrett's seen on biopsies. End of treatment (EoT) biopsies are then taken at 12 months to assess for the complete resolution of intestinal metaplasia (CR-IM) and complete resolution of dysplasia (CR-D).	NA	3 months	Regression of BE at 3 months, symptomatic stricture formation, resolution of intestinal metaplasia (CR-IM) and dysplasia (CR-D)	The mean reduction in Circumferential (C) length was 78% ± 36 and mean reduction in Maximal length (M) was 55% ± 36. 17 patients (15%) developed strictures requiring dilation. Stricture formation when the 12 J energy was used (p < 0.05). 47 patients had EoT biopsies, 40 (85%) had CR-D and 34 (76%) had CR-IM.	No comparison group Short follow-up - long-term BE clearance?
Munster (2022)	cohort	NA	NA	1386	65.5	We included all patients with BE containing early neoplasia who underwent endoscopic eradication therapy with at least one RFA treatment between 1 January 2008 and 31 December 2018	RFA	Visible lesions were removed with endoscopic resection. RFA was used to treat flat BE using the Barrx system. RFA was repeated every 3 months and was eventually followed by touch-up treatment using argon plasma coagulation or endoscopic resection for persisting BE islands of < 10mm and > 10mm, respectively. If a new nonflat neoplastic lesion was detected during one of the RFA treatments ("incident lesion"), additional endoscopic resection was performed. Upon complete endoscopic eradication of BE, random 4 quadrant biopsies were obtained < 5mm below the neosquamocolumnar junction for histological correlation. Patients with complete endoscopic eradication of BE and no dysplasia in the cardia biopsies were considered as CE-BE. Persisting intestinal metaplasia in cardia biopsies was also considered as CE-BE. Patients with persisting visible BE after RFA were classified as treatment failure. RFA was stopped if we anticipated that we would be unable to achieve CE-BE or if expected benefits of continued RFA were considered smaller than the risks.	NA	3 years	Complete esophageal healing.	In 134 patients with poor healing (10 %), additional time and acid suppression resulted in complete esophageal healing, and 67/134 (50%) had normal squamous regeneration with 97% CE-BE. Overall, 74 patients had poor squamous regeneration (5 %). Compared with patients with normal regeneration, patients with poor squamous regeneration had a higher risk for treatment failure (64% vs. 2 %, relative risk [RR] 27 [95% confidence interval [CI] 18–40])	No comparison group

Study	Design			N	Age	Inclusion/Population	Treatment		Follow-up	Outcome	Results	Comments
Desai (2021)	SR and meta-analysis	NA	NA	794	64.6	subjects with high-grade dysplasia and/or superficial adenocarcinoma who underwent BET (ablation ± endoscopic mucosal resection)	RFA	(a) subjects with high-grade dysplasia and/or superficial adenocarcinoma who underwent BET (ablation ± endoscopic mucosal resection); (b) BET completion by confirmation of complete eradication of neoplasia (CE-N) and intestinal metaplasia (CE-IM) with systematic sampling and (c) clearly defined follow-up (endoscopy and biopsy) protocol of ≥2 years thereafter for detection of recurrence. Pooled estimates of CE-N and CE-IM after BET completion and follow-up were analysed.	NA	>2 years	CE-IM and CE-N	Despite high efficacy of BET at therapy completion (CE-N: 95.9 [91.7-98.7]%; CE-IM: 90.9 [83-96.6]%), this declined (CE-N: 89 [73.4-98.2]%; CE-IM: 77.8 [65.6-88]%) over 3.4 years of follow-up	short follow-up - Only two studies reported a post-BET follow-up of >5 years (CE-IM 50 [41.5%-58.5]%).
Munster (2019)	Cohort	NA	NA	94	74	All patients who underwent ER for a neoplastic lesion between 2008 and 2018, without further ablation therapy.	No treatment	No additional ablation was performed for several reasons; in 73 patients (78%), the main argument was expected limited life expectancy.	NA	21 month	progression to HGD/EAC	17 patients (18%) developed HGD/EAC: all were curatively treated endoscopically.	No comparison group In 78%, there was a expected limited life expectancy Short follow-up
Oliphant (2014)	Cohort	NA	NA	72	73	Case series reviewed data from patients with biopsy proven Barrett's-associated HGD and/or IMC treated with therapeutic intent by ER with or without mucosal ablation between 2004 and 2011.	APC/RFA/photodynamic therapy	In addition to ER, 43% of patients were treated with argon plasma coagulation, 17% with radiofrequency ablation, and 11% with photodynamic therapy.	~	38 months	Disease progression and disease regression	8 (13%) patients with HGD at baseline and 0 (0%) with IMC progressed to invasive carcinoma. Dysplasia was completely eradicated (CE-D) with reversion to non-dysplastic Barrett's in 21 (29%) patients. 8 (13%) patients with HGD and 1 (1%) patient with intra-mucosal cancer regressed to low-grade dysplasia. Complete eradication of intestinal metaplasia (CE-IM) was achieved in 1 (1.4%) patient with a baseline diagnosis of HGD. Overall disease regression was achieved in 28 (44%) patients with HGD and 7 (38%) patients with IMC.	Short follow-up; no data on BE clearance?
Manner (2013)	Clinical trial	Randomized	not	63	63	Patients in whom complete remission from early Barrett's cancer or high grade intraepithelial neoplasia (HGIN) had been achieved following endoscopic resection	APC	Patients in whom complete remission from early Barrett's cancer or high grade intraepithelial neoplasia (HGIN) had been achieved following endoscopic resection and who fulfilled the inclusion criteria for the present study were randomly assigned to undergo APC ablation with concomitant PPI treatment with esomeprazole ("ablation group") or to undergo surveillance with PPI treatment ("surveillance group"). Randomization (1:1) was computer-based and was performed using a semi-deterministic minimization method	~	25-28 months	secondary lesions	For complete Barrett's ablation, a mean number of 4±1.6 APC sessions were required (range 2–7). The number of secondary lesions was 1 in the ablation group (3 %), and 11 in the surveillance group (36.7 %), leading to significantly higher recurrence-free survival for the patients undergoing ablation (P=0.005).	
Haidry (2015)	Cohort	NA	NA	508	68-69	Patients undergoing RFA/EMR for BE-related neoplasia from 2008 to 2013	RFA	Before RFA, visible lesions were removed by EMR. Thereafter, patients had RFA 3-monthly until all BE was ablated or cancer developed (endpoints). End of treatment biopsies were recommended at around 12 months from first RFA treatment or when endpoints were reached.	~	12 months	clearance of dysplasia (CR-D) and BE (CR-IM)	CR-D and CR-IM improved significantly between the former and later time periods, from 77% and 56% to 92% and 83%, respectively (p<0.0001). Rescue EMR after RFA decreased from 13% to 2% (p<0.0001). Progression to OAC at 12 months is not significantly different (3.6% vs 2.1%, p=0.51).	
Johnson (2015)	Cohort	NA	NA	49	61	Patients were included in the study if they had biopsy proven Barrett's esophagus with or without dysplasia or intramucosal adenocarcinoma.	fundoplication	The patients underwent RFA ± EMR followed by Nissen fundoplication. The application of EMR and RFA were not standardized in the study	~	24 months	clearance of dysplasia (CR-D) and BE (CR-IM)	Complete remission of intestinal metaplasia (CR-IM) was achieved in 26 (53 %) patients, complete remission of dysplasia (CR-D) in 16 (33 %) patients, and 7 (14 %) had persistent neoplastic Barrett's. After fundoplication, 18/26 (70 %) remained in CR-IM. An additional 10/16 CR-D achieved CR-IM and 4/7 with persistent dysplasia achieved CR-IM.	No comparison group. Short follow-up
Canto (2015)	Cohort	NA	NA	74	69	Patients with neoplastic Barrett's esophagus who had not undergone previous ablation (treatment-naïve group) or who had persistent or recurrent neoplasia despite previous treatment (rescue treatment group) were enrolled.	Cryo	EMR was performed at least 2 weeks prior to ablation therapy in order to remove visible lesions. The primary goal of cryotherapy treatment was the eradication of all neoplasia. The secondary goal was eradication of all intestinal metaplasia. The Food and Drug Administration-approved PolarW and system using compressed CO2 (cryogen) in a pressurized tank was used. Cryogen was delivered at 6–8L/min through a low-pressure flexible polyethylene catheter	NA	4.2 years	clearance of dysplasia (CR-D) and BE (CR-IM) recurrence adverse events	Between 2006 and 2013, 64 evaluable patients (20 treatment naïve, 44 rescue treatment) were treated. At 1 year, the overall complete response rates were 77% for cancer (10/13), 89% for dysplasia 57/64, 94% for HGD (60/64; 100% for treatment naïve, 91% for rescue treatment), and 55% for intestinal metaplasia (35/64). Long-term complete response for neoplasia with rescue therapy was 87% (56/64). Recurrent or new intestinal metaplasia was detected in 20/64 (31 %) after two negative follow-up procedures. Serious adverse events were noted in two patients (3 %). Post-cryotherapy pain occurred in four patients (6 %; only two needed analgesics).	No comparison group
Künzli (2014)	Cohort	NA	NA	83	68	Barrett's esophagus patients undergoing focal RFA using the simplified protocol in four tertiary referral centers were retrospectively included.	RFA	A prior endoscopic resection for visible lesions was allowed, as was a preceding circumferential RFA treatment with the HALO360 device. During each focal ablation, the gastroesophageal junction (GEJ) was ablated circumferentially in addition to Barrett's esophagus islands or tongues. Sessions continued at 8 to 12-week intervals until complete resolution of Barrett's esophagus.	NA	16 weeks	complete remission of dysplasia and of intestinal metaplasia and stenosis requiring dilation.	Complete remission of dysplasia in 78/83 (94 %) and of intestinal metaplasia in 72/83 (87 %). Stenosis requiring dilation developed in 9/83 (11 %), necessitating a median 2 dilation sessions (range 1–9), with ≥8 sessions in three patients.	No comparison group. Short follow-up
Natour (2018)	Cohort	NA	NA	35	65	Inclusion criteria were: 1) at least one previous endoscopic treatment for BE-related dysplasia or early neoplasia; 2) presence of islands or tongues of persistent BE mucosa ≤1cm not suspicious for dysplasia at follow-up endoscopy; 3) hot avulsion applied as the single treatment of these areas	Hot avulsion	All included patients had undergone hot avulsion for eradication of residual focal areas of BE that were ≤ 1 cm and not suspicious for dysplasia, following at least one previous endoscopic treatment for dysplasia or intramucosal cancer.	NA	17.4 months	complete eradication of residual focal BE	ll patients achieved complete eradication of residual focal BE. One of the patients required a second hot avulsion treatment	No comparison group. Short follow-up
Natour (2018)	Cohort	NA	NA	60		All patients with biopsy-proven BEHGD and IMC (T1a), who were treated endoscopically between 2007 and 2014, were prospectively analyzed	RFA	Treatment algorithms were determined by consensus opinion after presentation at gastrointestinal tumor board. Patients underwent EMR and/or RFA until eradication-of-dysplasia and complete remission of intestinal metaplasia (CRIM) was achieved. Patients were then enrolled in an endoscopic surveillance program.	NA	33 months	CE-IM and CE-N recurrence	Fifty-five (92%) patients achieved CE-N and 52(87%) CE-IM. Nine (15%) patients relapsed after CRIM with nondysplastic-BE (6), BE with low-grade dysplasia (1), and high-grade dysplasia (2).	No comparison group. Short follow-up

Study	Type			N	Age	Inclusion criteria	Intervention		Follow-up	Outcome	Results	Comments
Brow (2015)	Cohort	NA	NA	61	68.8	(i) One or more RFA treatments, (ii) age greater than 18, (iii) presence of endoscopic and biopsy-proven BE of equal to or greater than 1 cm. 61=Five patients had intra-mucosal cancer (8.2%), 38 HGD (62.3%), and 18 LGD (29.5%).	RFA	This is a retrospective cross-sectional study of patients who underwent at least one treatment with either focal and/or balloon RFA devices who were identified from two tertiary centers.	–	not stated	BE reduction	RFA with a focal device resulted in greater percentage reduction in BE length compared to the balloon system (73% vs. 39%, p<0.01). After adjusting for initial BE length, pre-treatment BE length, hernia status, prior endoscopic mucosal resection (EMR), prior RFA, and prior EMR/RFA sessions, RFA with a focal device at each session remained an independent predictor for a significant reduction in BE extent as compared to the balloon system. Those who underwent RFA with the focal device only required fewer RFA sessions to achieve CE-IM (mean 1.6 vs. 3.8 sessions, p<0.01). Only one patient in the balloon RFA group had esophageal stricture while none of the patients in the focal RFA group had a stricture	follow-up?
Manner (2015)	Cohort	NA	NA	60	62.4	Patients who had a residual BE segment of at least 1 cm after endoscopic resection of early Barrett's neoplasia underwent thermal ablation of BE by Hybrid-APC.	APC	Prior to thermal ablation, submucosal injection of sodium chloride 0.9 % was carried out. Check-up upper GI endoscopy was carried out 3 months after macroscopically complete ablation including biopsies from the neo-Z-line and the former BE segment, and recording of stricture formation.	NA	3 months	complete eradication of residual focal BE + stricture	Forty-eight out of 50 pt (96 %; ITT: 49/60, 82 %) achieved macroscopically complete remission after a median of 3.5 APC sessions [SD 2.4; range 1–10]. There was one treatment-related stricture (2 %).	No comparison group. Short follow-up
Koutsoumpas (2016)	Cohort	NA	NA	118	71	Patients undergoing esophageal EMR for early neoplasia in Barrett's esophagus from 2009 to 2014.	EMR	EMR was performed using a maximum of two band ligation mucosectomies per endoscopic session; thereafter, follow-up was 3-monthly and EMR was repeated as required for Barrett's eradication.	NA	24 months	complete eradication of residual focal BE + stricture	complete Barrett's excision in 85.0 %. One patient developed a stricture (1.1 %), one a delayed bleeding, and there were no perforations.	No comparison group. Short follow-up
Fujii-Lau (2017)	SR and meta-analysis	NA	NA	4355 patients (3213 treated with RFA, 567 with SRER, and 575 both modalities)	not stated	Patients who achieved CE-IM after endoscopic therapy of Barrett's esophagus with IM, dysplasia or early adenocarcinoma (EAC)	RFA vs EMR	Studies were included if they met the following strict criteria: (1) study design – randomized control trial, cohort studies, case series with at least 20 patients; (2) patient population – patients who achieved CE-IM after endoscopic therapy of Barrett's esophagus with IM, dysplasia or early adenocarcinoma (EAC); (3) intervention – primary endoscopic therapy with stepwise complete endoscopic resection (SRER) or radiofrequency ablation (RFA) with or without focal endoscopic mucosal resection(EMR); (4) outcome – reported number of patients with recurrent IM, dysplasia, or EAC on histology; and (5) mean followup of at least 1 year after the first endoscopy confirming complete eradication.	–	11 837.63 patient-years	recurrence	The pooled incidence of any recurrence was 7.5 (95 %CI 6.1 – 9.0)/100 PY with a pooled incidence of IM recurrence rate of 4.8 (95%CI 3.8 – 5.9)/100 PY, and dysplasia recurrence rate of 2.0 (95%CI 1.5 – 2.5)/ 100 PY. Compared to the SRER group, the RFA group had significantly higher overall [8.6 (6.7 – 10.5)/100 PY vs. 5.1 (3.1 – 7)/100 PY, P = 0.01] and IM recurrence rates [5.8 (4.3 – 7.3)/100 PY vs. 3.1 (1.7 – 4)/100 PY, P < 0.01] with no difference in recurrence rates of dysplasia.	
Gupta (2013)	Cohort	NA	NA	592	64	Inclusion criteria were: patients over the age of 18 with the presence of endoscopically (at least 1 cm endoscopically evident columnar mucosa in the tubular esophagus) and histologically-confirmed BE	APC/RFA/photodynamic therapy/cryotherapy	All subjects who achieved CRIM remained on endoscopic surveillance. The frequency of endoscopy performed in the surveillance program was dependent on the original histology of BE treated.	NA		Complete remission of intestinal metaplasia (CRIM) was defined as eradication of IM (in esophageal and gastro esophageal junction biopsies), documented by 2 consecutive endoscopies. Recurrence was defined as presence of IM or dysplasia after CRIM in surveillance biopsies.	Twenty-four months after CRIM, the incidence of recurrence was 33%; 22% of all recurrences observed were dysplastic BE.	No comparison group
Haidry (2013)	Cohort	NA	NA	335	68	All patients were referred for consideration of ablative management of dysplastic BE at a collaborating center.	RFA	Before patients could undergo HALO RFA, any visible lesions were removed by endoscopic resection. After the initial ablation with either the focal HALO90 or circumferential HALO360 catheter, patients were discharged home with highdose acid suppression	NA	19 montha	eradication	HGD was cleared from 86% of patients, all dysplasia from 81%, and BE from 62% at the 12-month time point, after a mean of 2.5 (range, 2 6) RFA procedures. Invasive cancer developed in 10 patients (3%) by the 12-month time point and disease had progressed in 17 patients (5.1%) after a median follow-up time of 19 months. Symptomatic strictures developed in 9% of patients and were treated by endoscopic dilatation.	Short follow-up - long-term BE clearance?
Trindade (2017)	Cohort	NA	NA	27	68	Consecutive patients with BE who received Liquid nitrogen cryotherapy following endoscopic mucosal resection (EMR) of intramucosal cancer.	Cryo	All patients underwent EMR of the nodular portion of BE in which histology showed the nodules were T1a stage cancer. Cryotherapy was subsequently performed at a different session, 6-8 weeks after the EMR, to allow the EMR site to heal.	NA	2 years	complete eradication of dysplasia (CE-D), intestinal metaplasia (CE-IM), and development of invasive cancer during follow up	22/27 patients (82%) achieved CE-D after cryotherapy; 19/27 patients (70%) achieved CE-IM; 1/27 patients (4%) developed invasive cancer	No comparison group Short follow-up - long-term BE clearance? Low number of patients
Phoa (2013)	Cohort	NA	NA	55	65	Patients were initially included if they had endoscopically visible BE with histology proven HGIN and/or early-stage cancer demonstrated on at least 2 separate endoscopies	RFA	Forty patients (72%) underwent endoscopic resection of visible abnormalities before the first RFA treatment. After a minimum interval of 6 weeks after endoscopic resection for endoscopically visible abnormalities, patients were treated with primary circumferential ablation using the HALO360 system. Subsequently, patients underwent a series of circumferential or focal RFA procedures	NA	5 Yyears	CE-IM and CE-N	After RFA treatment, CR-neoplasia/CR-IM was achieved in 54 of 55 (98%) patients. In total, 46 patients were followed for at least 5 years. In this cohort of patients, sustained CR-neoplasia and CR-IM were maintained in 43 of 46 patients (93%; 95% CI: 82.5 97.8).	No comparison group
Okoro (2012)	Cohort	NA	NA	90	60-67	Patients treated for BE associated with dysplasia or intramucosal cancer.	RFA	compared complications and histologic outcomes between patients who had EMR before RFA and those who received only RFA.	–	20.5 months	Stricture, CE-IM and CE-N	Stricture rates were 14% in group 1 and 9% in group 2 (odds ratio, 1.53; 95% CI, 0.26 –9.74). The rates of CR-IM were 43% in group 1 and 74% in group 2 (odds ratio, 0.33; 95% CI, 0.14 – 0.78). The rates of complete resolution of dysplasia were 76% in group 1 and 71% in group 2 (odds ratio, 1.28; 95% CI, 0.39 – 4.17). The adjusted odds ratio for CR-IM in group 1 (adjusting for age, segment length, and grade of dysplasia) was 0.50 (95% CI, 0.15–1.66).	Short follow-up - long-term BE clearance?
Herrero (2011)	Cohort	NA	NA	26	66	Consecutiive patients with BE 10 cm with early neoplasia	RFA	RFA with or without prior endoscopic resection (ER) for BE >10 cm containing neoplasia.		29 months	CR-IM and CR-neoplasia	CR-neoplasia and CR-IM were achieved in 83% (95% confidence interval [CI], 63%- 95%) and 79% (95% CI, 58%-93%), respectively. None of the patients had fatal or severe complications and 15% (95% CI, 4%-35%) had moderate complications.	short follow-up Low number of patients

Study	Design			N	Age	Inclusion criteria	Intervention	Comparison		Follow-up	Outcome	Results	Comments
Phoa (2015)	Cohort	NA	NA	132	65	Patients with BO≤12 cm with HGD and/or EC on 2 separate endoscopies were eligible for inclusio	RFA	Visible lesions (<2 cm length; <50% circumference) were removed with ER, followed by serial RFA every 3 months (max 5 sessions). Follow-up endoscopy was scheduled at 6 months after the first negative post-treatment endoscopic control and annually thereafter.	NA	27 months	complete eradication of neoplasia (CE-neo) and intestinal metaplasia (CE-IM)	After entry-ER in 119 patients (90%) and a median of 3 RFA (IQR 3–4) treatments, CE-neo was achieved in 121/132 (92%) and CE-IM in 115/132 patients (87%), per intention-to-treat analysis. Perprotocol analysis, CE-neo and CE-IM were achieved in 98% and 93%, respectively. After a median of 27 months following the first negative post-treatment endoscopic control, neoplasia and IM recurred in 4% and 8%, respectively.	No comparison group Short follow-up - long-term BE clearance?
Barret (2016)	Cohort	NA	NA	40	66	Patients with early Barrett's neoplasia and a visible lesion undergoing combined endoscopic resection and focal radiofrequency in a single session	RFA	Consecutive ablation procedures were performed every 8 to 12 weeks until complete endoscopic and histological eradication of dysplasia and intestinal metaplasia was reached. Endoscopic resection was performed using the multiband mucosectomy technique in 80% of cases and for radiofrequency ablation the Barrx90 catheter for focal ablation was used.	NA	19 months	complete eradication of BE + stricture	Stenoses occurred in 33% of cases, successfully managed with a median number of 2 dilations. In 43% of patients one single session treatment session resulted in complete histological remission of intestinal metaplasia.	No comparison group Short follow-up - long-term BE clearance?
Vilsteren (2015)	Clinical trial	Randomized	Pathologists were blinded	55	69	Patients were eligible if they met the following criteria: (1) age between 18 and 85 years; (2) BO length ≤5 cm; (3) HGD and/or EC in BO in specimens obtained at two separate endoscopies; (4) no signs of deep submucosal invasion, regional lymph node involvement or distant metastases on EUS and CT of thorax and abdomen (in the case of EC); (5) no prior endoscopic treatment of BO other than a single prior ER for staging; (6) in the case of a prior diagnostic ER, specimens with a negative deep resection margin, no deep submucosal invasion (≤T1sm2), no lymphatic/vascular invasive growth and no poorly or undifferentiated cancer (G3eG4);	Patients with BO <5 cm containing HGD/EC were randomised to SRER or ER/RFA. In all other patients with visible lesions, ER was performed after randomisation: in the SRER arm the lesion was removed together with the first 50% of the BO segment in the same session	–		24 months	complete eradication of BE + stricture	CR-neoplasia was achieved in 25/25 (100%) SRER and in 21/22 (96%) ER/RFA patients. CR-IM was achieved in 23 (92%) SRER and 21 (96%) ER/RFA patients. The stenosis rate was significantly higher in SRER (88%) versus ER/RFA (14%; p<0.001), resulting in more therapeutic sessions in SRER (6 vs 3; p<0.001) due to dilation	Short follow-up - long-term BE clearance?

PICO search string:	(((Barrett[Title/Abstract] OR Barrett's esophagus[Title/Abstract]) AND (residual[Title/Abstract] OR after Endoscopic Mucosal Resection[Title/Abstract] OR after Endoscopic Mucosal Dissection[Title/Abstract] OR mucosectomy[Title/Abstract] OR Endoscopic resection[Title/Abstract])) AND (Ablation[Title/Abstract] OR Radiofrequency[Title/Abstract] OR Argon Plasma[Title/Abstract] OR Cryoballom[Title/Abstract] OR Cryoablation[Title/Abstract])) AND (survival[Title/Abstract] OR dysplasia[Title/Abstract] OR adenocarcinoma[Title/Abstract] OR cancer[Title/Abstract] OR neoplasia[Title/Abstract])

	Methods			**Population**			**Intervention**			**Outcomes**		
Author (year)	**Design**	**Randimsation / blinding**		**N**	**Age**	**Inclusion criteria**	**Protocol (details) of Intervention**	**Protocol (details) of Comparison**	**Follow-up**	**Outcome measures (+ definitions)**	**Outcomes of interest**	**Remarks**

Table 21 — FRA versus alternative ablation methods

Author (year)	Design	Randimsation / blinding		N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	Remarks
Genere, 2022	retrospective	no	no	23/23	NA	persistent (<50% regression) or dysplasia after 2 RFA	continued RFA	spray cryotherapy	NA	CEIM, CED, esophageal strictures, number of treatment sessions	CEIM and CED 96% vs 83%, p = NS, strictures = 71 % vs 14%, p = 0,02; treatment sessions 12 vs 19, p<0,01	disputable inclusion criteria 71% stricture rates in the RFA group very unusual
Fasullo, 2021	retrospective	no	no	100/62	67	BE undergoing EET (residual BE after ER, flat HGD or flat LGD)	RFA	spray cryotherapy	> 12 months	CEIM, CED, esophageal strictures, number of treatment sessions, outcome of switched therapy, recurence of IM and dysplasia after CEIM, severe AE	CEIM 64% vs 66%, p = 0,78 ; CE D 81% vs 71 %, p = 0,14; n treatment = 3,5 vs 4,8, p = 0,004; no severe AE (ASGE) ; switched therapy : 15% vs 24%, p = 0,15, with CE IM 60 % vs 73 %, p = 0,44, and 3 vs 5,1 treatment sessions	
Thota, 2018	retrospective	no	no	73/81	66/70	BE undergoing EET	RFA (360 and 90)	spray cryotherapy	25 months	CEIM, CED, esophageal strictures, number of treatment sessions, recurrence rate, neoplastic progression rate, esophageal cancer related mortality	CE IM 67% vs 41%, p = 0,002, CE D 88% vs 79%, p = 0,15, treatment sessions 3 vs 3 , recurrence rate 11% vs 14 %, neoplastic progression rate 12,5% vs 12,5 %, esophageal cancer related mortality 1% vs 5% (p = 0,37)	C3M5 BE in both groups no data on AE worrying number of EAC related deaths in the cryo group
Solomon, 2019	prospective	no	no	59/35	NA	BE undergoing EET	Focal RFA	spray cryotherapy	1 month	post treatment pain (numeric pain scale), dysphagia	immediate pain score 1,18 vs 0,41, p = 0,001, 48h pain score 1,77 vs 0,76, p = 0,013, no difference at 3 weeks, and dysphagia similar in both groups at all time points (p = 0,4)	
Van Munster, 2018	prospective	no	yes (assessment of BE regression)	26/20	67	BE undergoing EET	Focal RFA, triple 12J/cm²	cryoballoon, focal ablation system, 10 s application	3 months	BE regression, 14 day scores for pain and dysphagia	BE regression 90% vs 88%, p = 0,6, AUC for pain, dysphagia, analgesics smaller for cryoablation, pain duration 4 vs 2 day (RFA vs cryo), analgesic use 4 vs 2 days, p<0,01	single blinding (BE regression assessment) , possible bias for pain evaluation
Peerally, 2019	prospective	yes	no	36/40	70	residual BE after ER	RFA, 4 sessions	APC, 4 sessions, 2L/mn, forced 60 W or pulsed 50 W	> 12 months	CE IM, CE D, AE, quality of life, costs, buried glands	CEIM 56%vs 48%, CED 79% vs 84%, 3,2 vs 3,4 treatment sessions, stricture 8,3% vs 8,1 %, Bleeding 2,8 % vs 5,4% Tolerance to treatment at D7 similar, QOL 6M and 12M similar, treatment duration30 vs 24 mn, cost 33170 USD vs 5678 USD per case, buried glands 6% vs 13%	Comparable efficacy and safety, with LSBE (mean C3-4M6), and much lower costs. Pilot trial only (not an adequately powered non inferiority trial)
Knabe, 2020 (ABSTRACT)	prospective	yes	no	48/55	65	residual BE after ER	RFA, 360 10J/cm² x 2 or 90 with triple 12J/cm²	hybrid APC, 60W, saline injection	12 months	CE IM, postprocedural pain intensity (0-10) and duration, strictures	CEIM 87% vs.91%, p = NS, post treatment pain score 4 vs 2, p<0,05, post treatment pain duration 6 vs 3 days, p<0,05, strictures 13% vs 2%, p<0,05	only an abstract with short term results, 58 patients analyzed, early study termination
Alshelleh, 2021 (ABSTRACT)	retrospective	no	no	26/25	NA	BE undergoing EET	focal cryoballoon ablation	spray cryotherapy	12 monts	CEIM, CED, stricture rate	CEIM 85 vs 80 % p = NS, CED 96% vs 96 %, p = NS, similar stricture rate	
Tariq, 2020	metanalysis	no	no	405	60-71	BE undergoing EET	spray LN, spray C02, focal cryoballoon	none	3-54 months	CEIM, CED, AE	CE IM = 72 %, CE D = 91%, AE = 12,2 %	heterogeneity in devices (spray with liquid nitrogen, spray with C02, balloon with nitrous oxyde)
Westerweld, 2020	metanalysis	no	no	272, 7 studies	66	BE undergoing EET	focal cryoballoon ablation 10 s	none	NA	CEIM, CED, AE	CE IM = 86 %, CE D = 94%, AE = 12,5 % among which stricture = 6%, mucosal laceration 0,7 %, perforation 0,4%, bleeding 0,4%, all successfully managed endoscopically	small retrospective studies
Hamade, 2019	metanalysis	no	no	282, 6 studies	65	BE undergoing EET	cryotherapy			CEIM, CED, , neoplastic progression rate, , recurrence rate of IM and dysplasia, AE	CE IM = 70 % CE D = 98 %, neoplastic progression rate = 4 %, recurrent IM = 19% per year, recurrent dysplasia = 10% per year, stricture rate = 5 %	heterogeneity in the cryoablation devices
Orman, 2013	metanalysis	no	no	3802, 18 studies	na	BE undergoing EET	RFA			CEIM, CED, neoplastic progression rate, , recurrence rate of IM and dysplasia, AE	CE IM = 78 %, CE- D= 91 %, neoplastic progression rate = 0,2% during ttt and 0,7% after CE-IM, recurrent IM = 13 %, stricture rate = 5%	
Pasricha, 2013	prospective	no	no	3728	62	BE undergoing EET	RFA		2,4 years	CEIM, recurrence rate	CE IM 85 %, recurrent IM = 20% at two years	1300 patients lost to FU, mean length of recurrence = 0,6 cm
Van Munster, 2021	prospective	no	no	1154	65	BE undergoing EET	RFA		43 months	CE IM CE D, stricture, recurrence rate	CE IM 94%, AE = 21%, esophageal stricture 15 %, bleeding 4 %, perforation 1%	

Wolfson, 2022	prospective	no	no	2535	67	BE undergoing EET	RFA		120 months	CEIM, CED, neoplastic progression rate, , recurrence rate of IM	CE IM = 63 %, CE D = 88%, neoplastic progression = 4,1 % at 10 years, recurrent IM = 19% at 8 years,	
Manner, 2014	prospective	yes	no	33/30	62	residual BE	APC	surveillance	28 months	CE IM CE D, stricture, recurrence rate, number of treatment session,	CE IM = CE D = 97 %, neoplastic progression rate = 3%, stricture rate = 9,1%, 4 treatment session	C5M6 BE
Knabe, 2022	prospective	no	no	154	64	residual BE	hybrid APC		24 months	CE IM CE D, stricture, recurrence rate, number of treatment session,	CE IM = 87 %, CE D = 98 %, AE = 6,1 %, stricture = 4%, bleeding = 4% , perforation = 0,6 %, recurrent BE = 10% (13/129) at two years, recurrent neoplasia 2% (3/129) at rwo years	

Table 22   Management of low-risk T1a

TITLE	AUTHORS / JOURNAL	STUDY DESIGN	STUDY POPULATION	CONTROL POPULATION	OUTCOME
Long-term efficacy and safety of endoscopic resection for patients with mucosal adenocarcinoma of the esophagus.	*Pech O, May A, Manner H, et al.* *Gastroenterology 2014; 146: 652–660.*	retrospective cohort study	1000 consecutive patients (mean age, 69.1 ± 10.7 years; 861 men) with mAC (481 with short-segment and 519 with long-segment Barrett's esophagus) who presented at a tertiary care center from October 1996 to September 2010		After a mean follow-up period of 56.6 ± 33.4 months, 963 patients (96.3%) had achieved a complete response; surgery was necessary in 12 patients (3.7%) after endoscopic therapy failed. Metachronous lesions or recurrence of cancer developed during the follow-up period in 140 patients (14.5%) but
Surgery versus radical endotherapies for early cancer and high-grade dysplasia in Barrett's oesophagus.	Bennett C, Green S, DeCaestecker J, Almond M, Barr H, Bhandari P, Ragunath K, Singh R, Jankowski J. Cochrane Database Syst Rev. 2020 May 22;5(5):CD007334. doi: 10.1002/14651858.CD007334.pub5. PMID: 32442322	Systematic review			no RCTs available
Endotherapy versus surgery for early neoplasia in Barrett's esophagus: a meta-analysis.	Wu J, Pan Y, Wang T, Gao D, Hu B. Gastrointest Endosc. 2014 Feb;79(2):233-241.e2. PMID 24079410	meta-analysis	This study included seven studies involving 870 patients, 510 treated endoscopically	360 patients treated with esophagectomy	no significant difference between endoscopic therapy and esophagectomy in the neoplasia remission rate (relative risk [RR] 0.96), overall survival rate at 1 year (RR 0.99), 3 years (RR 1.03), and 5 years (RR 1.00). Endoscopic therapy was associated with a higher dysplasia recurrence rate (RR 9.50) and fewer major adverse events (RR 0.38).

Table 23							EMR vs ESD for BE				

Author (year)	Methods		Population				Intervention			Outcomes		
	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	Remarks
Perez-2022-Endoscopy	Retrospective, Multicentric	No	No	150 EMRs/93 ESDs	71 (IQR 62–78) EMR 69 (63–75) ESD	BE-HGD or T1a EAC . Patients who underwent both cEMR and ESD or received surgery or chemoradiation before endoscopic therapy were excluded from the analysis.	EMR+ ablation	ESD+ablation	EMR vs ESD: media (interquartile range) 15.5 (6.75–30) Vs 8 (2– 18) months (p<0.001) RFA was performed following EMR 62.5 % and ESD 69.4 % (P = 0.40).	Repeat endoscopic treatment procedure, defined as a fol-low-up endoscopic resection procedure (EMR, ESD) or cryo-therapy needed to treat local recurrence. RFA ablation was not included. Efficacy outcomes. En bloc resection was defined when the target lesion was resected as a single specimen. R0 resection was the absence of the highest grade depth of invasion within the lesion (HGD or T1a), at the lateral or deep margin. Curative resection was an HGD/EAC R0 resection and histo- pathological confirmation of the absence of lymphovascular invasion, absence of poor grading (G3), and restriction to the mucosa and superficial submucosa (< 500 µm). Metachronous lesion, defined as detection of HGD/EAC on biopsies from a visible lesion at a location not in the proxi- mity of the post-endoscopic resection scar. Adverse events. Major early bleeding was defined as bleed- ing requiring premature termination of the procedure, transfusion of red blood cells, or not amenable to manage- ment by endoscopic intervention. Major delayed bleeding was defined as clinical signs of bleeding, with a hemoglobin drop > 2 g/dL. Perforation was defined as a visible defect in the esophageal wall that allowed visualization into the med-iastinum, or as signs of mediastinitis with post-intervention- al imaging proving extravasation of contrast medium. Stricture was defined as a post-treatment luminal narrowing that could not be traversed with a diagnostic upper gastrointes- tinal endoscope.	EMR Vs ESD: en bloc resection 43 Vs 89% (P<0.001); R0: 56 vs 73 (P = 0.01), perforation 0.7 vs 0% (P > 0.99), early bleeding 0.7 vs. 1 % (P > 0.99), delayed bleeding 3.3 vs. 2.1%; (P=0.71), stricture 10vs. 16% (P=0.16). EMR Vs ESD: Recurrent/residual disease 31.4vs 3.5 %; 48-month recurrence-free survival rate 68.6 vs. 96.5 %; (P < 0.001) Additional end treatment to treat recurrent/residual disease 24.2 vs. 3.5 % (P < 0.001).	Follow-up longer in the EMR group. Good ESD results even with low ESDs per center: Cleveland Clinic, 15 EMR and 26 ESD; University Hospitals, 79 EMR and 4 ESD; Uni- versity of Florida, 34 ESD; Mayo Clinic Arizona, 42 EMR and 13 ESD; University of Kentucky, 14 EMR; Johns Hopkins, 3 ESD; Brigham and Women's Hospital, 9 ESD; and University of Sao Paulo, 4 ESD.
Codipilly-2022-Clin Gastroen Hepat.	Prospectively maintained database Single center	No	No	537 patients: 456 cEMR, 81 ESD.	Mean (SD) EMR65.2 (9.8) ESD 68.6 (10.3)	Dysplastic BE or EAC Patients with surgery or chemoradiation before endoscopic therapy were excluded from the analysis.	EMR+ ablation	ESD+ablation	cEMR: follow-up period of 11.2 years (interquartile range [IQR], 6.5–15.7 y), ESD: follow-up period of 1.4 years (IQR, 0.8–2.0 y).	CRD: absence of dysplasia on biopsy specimens from the tubular esophagus and gastroesophageal junc- tion, during at least 1 surveillance endoscopy) CRIM: absence of intestinal metaplasia on biopsy specimens from the tubular esophagus and gastroesophageal junction, during at least 1 surveillance endoscopy	cEMR Vs ESD: Cumulative probabilities of CRD at 2 years were 75.8% and 85.6% in the cEMR and ESD groups, respectively (P < .01). Independent predictors of CRD were: ESD (hazard ratio [HR], 2.38; P < .01) and shorter BE segment length (HR, 1.11; P < .01). The cumulative probabilities of CRIM at 2 years were 59.3% and 50.6% in the cEMR and ESD groups, respectively (P > .05). The only independent predictor of CRIM was a shorter BE segment (HR, 1.16; P < .01).	BE patients with dysplasia or intramucosal adenocarcinoma undergoing ESD reach CRD at higher rates than those treated with cEMR, although CRIM rates at 2 years and complication rates were similar between the 2 groups.
Draganov-2021-Gastroeterology.	multicenter prospective study on ESD across 10 centers	No	No	692 ESDs, 181 esophagus, 171 BE/EAC								Dificult to assess detaildes data for BE because data are aggregated.
van Munster-2021-Endoscopy.	Retrospective, multicentric	No	No	138 ESDs (intention to treat analysis) ESD procedures completed n = 130 (per protocol analysis)	median age of 68 (61–73) years.	ESD for BE neoplasia	ESD			Procedure-related outcomes: (i) the en bloc resection rate, defined as the proportion of resections assessed as being en bloc at the end of the ESD procedure; (ii) the R0 resection rate, defined as the proportion of en bloc resections with lateral and vertical resection margins that were R0 (namely, free of cancer, or free of HGD if HGD was the worst diagnosis), strati- fied for depth of invasion; (iii) lesion histology; and (iv) proce- dure-related adverse events. Endpoints related to follow-up (i) the incidence of residual cancer ("persistent neoplasia") as detected during endoscopic follow-up or surgical resection after ESD, for R0 and R1 resections (see Table 1 s in Supplementary Material for all definitions); (ii) the incidence of recurrent lesions during fol- low-up; and (iii) the reliability of an endoscopic assessment for the presence of residual neoplasia after ESD.	126/ 130 (97%) en bloc resections. Among resections of HGD or T1a EAC lesions, 87% (95%CI 75%–92%) were both en bloc and R0; the corresponding value for T1b EAC lesions was 49% (36%–60%). R1 resections, 10/34 (29%)with residual cancer, all detected at first endoscopic follow-up. The remaining 24 patients (71 %) showed no residual neoplasia. Six of these patients underwent sur gery with no residual tumor; the remaining 18 underwent endoscopic follow-up during median 31 months with 1 local recurrence (annual recurrence rate 2%). Among R0 resec- tions, annual local recurrence rate during median 27 months was 0.5 %.	There were no procedure-related deaths. Post-procedural bleeding: 4/138; 2.9 % [0.7 %–5.8 %]. A small perforation occurred in 1 patient treated endoscopically 18 patients developed a stricture (18/138; 13% [7.2 %–18.8 %]) managed with a dila- tions (range 1–12). All the patients who developed esophageal strictures had undergone resection of > 50 % of the esophageal circumference.

Study	Design			N	Age	Indication	Intervention	Comparator	Follow-up	Aims/Outcomes	Results	Adverse events / Conclusions
Omae-2021-Endosc Int Open.	Retrospective, single center	No	Npo	90 ESDs: 22 were C-ESD (24 %) and 68 were WF-ESD (76 %)	median age 69 (IQR 63–74),	Dysplasic BE	C-ESD	WF-ESD. Wide field ESD: wider resection margins (5–10mm),	follow-up of 13.4 and 9.4 months in the C-ESD and WF-ESD cohorts, respectively.	The primary outcomes were the rates of en bloc, R0, curative resection, recurrence and procedure time. The secondary out- comes were adverse events, namely bleeding, perforation, stricture, and need for unplanned medical assistance after discharge. R0 or radical resection was defined as the presence of negative lateral (LM) and vertical margins (VM) for high-grade dys- plasia/cancer in the ESD specimens as described before. Non-radical resection corresponded to R1 resections. Curative resections were the R0 resections of well/moder- ately differentiated neoplasias that engaged the mucosa or su- perficial submucosa (Sm1) and had no lymphovascular involve- ment. Non-curative resections were resections in which one of those parameters were not fulfilled. The duration of ESD was defined as the time (in minutes) from the first incision until retrieval of the specimen. A velocity index of time (min) divided by the resected size (mm) was cal- culated in all ESDs. Local recurrence was defined as high-grade dysplasia/cancer detected within 2 cm of the previous ESD scar and a meta- chronous recurrence was defined as neoplasia located more than 2 cm away from the ESD scar, at any of the follow-up endoscopies, before ablation treatment. Complete remission of neoplasia was defined as the absence of suspicious lesions on gastroscopy with ME-NBI and chromoendoscopy with acetic acid and no histologic evidence of dysplasia/cancer applying the Seattle protocol on the follow-up endoscopies.	The en bloc resec-tion rate was 95 vs 100% (ns), the positive lateral margin rate was 23% vs 3% (P <0.01), the R0 rate was 73% vs 90%, and the curative resection rate was 59 % vs 76 % in the C-ESD and WF-ESD groups, respectively, (both P > 0.05). WF-ESD was associated with less post-operative strictures, 6% vs 27% (P= 0.01), with no local recurrence but no significantly reduced risk of metachronous recurrence (Hazard Ratio = 0.46, 95 % CI = 0.14–1.46), during a follow-up. In the WF-ESD cohort, depth of invasion (presence of deep submucosal cancer relative to mucosal cancer) was the only identifiable factor for non-radicality (OR 60.0, 95% confidence interval [CI] 10.0–361.0, P < 0.001). Lesion histology (presence of EAC relative to absence of adenocarcinoma) was the only fac- tor associated with non-curability (OR 21.0 95% CI 14.5–97.2, P < 0.001). Resection of more than 75 % of the circumferential lumen was the only identifiable factor for the presence of com- plications, namely strictures (OR 6.4 95 % CI 1.4–29.1 P < 0.02).	There was one case of bleeding and one case of perforation, both in the WF-ESD group and both were treated conservatively. Strictures were less frequent in the WF-ESD group, four of 68 (6 %) vs the C-ESD group, six of 22 (27 %) (P = 0.01). In large resections, that engaged > 75 % of the luminal circumference, strictures were more frequent in the C-ESD group (3/3, 100%) than in WF-ESD group 4/15 (27 %), P = 0.04. The later received oral steroids prophylactically
Podboy-2020-Digestive Endoscopy.	Retrospective	No	No	31 EMRs, 20 ESDs	(67.48 EMR vs 70.95 ESD)	BE with dysplasia/cancer	EMR	ESD	Years: 2.8 1.6 EMR, 1.4 1.1 ESD	Submucosal involvement was rated in quartiles across the entire length of each sample. Artifact was defined as any significant cautery effect or procedural hemorrhage that impaired interpreta- tion of the pathology specimen. This was subjectively graded from mild (minimally affecting pathologic interpre- tation) to severe (severely affecting ability to accurately interpret the specimen). R0 resection was defined as vertical and lateral resection margins free of carcinoma or high grade dysplasia (cases with only low grade dysplasia or no dysplasia in the specimen were excluded from the R0 calculation). An en bloc resection was defined as endoscopic removal of the entire visible lesion in one piece. Any discrepancy between the reviewing pathologists was resolved by consensus review.	ESD produced significantly larger specimens (mean length (mm) xmean width (mm) 27.5 9 17.9 for ESD vs 12.4 9 9.3 in EMR); with a smaller total number of specimens (1.4 ESD vs 2.9 EMR, all P values < 0.05). ESD produced more R0 resections (81.3% vs 20.0%), more en bloc resections (100% vs 51.6%), better oriented specimens (90.7% vs 66.0%), greater presence of submucosal glands (70.0% vs 38.7%), with a greater percentage of present submucosa (100% vs 74.2%) compared to EMR (all P values <0.05). There was no difference in the presence or severity of cautery artifact, or fragmentation of specimen between the two groups. EMR was associated with a significant number of more equivocal lateral margins (13/31, 41.9% vs 1/20, 5.0%, P < 0.05) and more equivocal vertical margins (13/ 31, 41.9% vs 0/20, 0%, P = 0.004). Definitive pathologic diagnosis not possible in 13/31 EMR specimens compared to 0/20 ESD specimens (P = 0.002).	EMR equivocal vertical margins were secondary to iatrogenic artifact (cautery or crush effect) of the specimen in 9/13 or secondary to limited submucosa in 4/13. Equivocal lateral margins were secondary to significant artifact in five and malorientation or fragmentation in eight. Of the 13 EMR specimens with equivocal pathology, 11 were noted to have 'at least intramucosal adenocarcinoma'. Four of the 11 patients chose to undergo elective esophagectomy with final surgical pathol- ogy demonstrating ≤T1a disease in 2, and ≥T1b disease in two. No significant different in adverse events
Yutaka Tomizawa-2020-End Int Open	multicenter retrospective	No	No	31 ESDs, 12 savage ESDs	Median age was 71 (IQR 55–79)	ESD for BE visible neoplasia.	ESD in naive BE	Savage ESD	mean 197 days	Savage ESD: ESD in recurrent BE- related neoplasia after prior endoscopic therapy.	Complete R0 resection was obtained in 75% in the salvage group and 80% in the non-salvage group (P = 1.00). In seven patients (22 %), the pre-ESD diagnosis was upgraded on post-ESD histopatholo- gy (1 low-grade dysplasia to high grade dysplasia [HGD], 4 HGD to early esophageal carcinoma (EAC), and 2 intramu- cosal EAC to invasive EAC). No perforations occurred in ei- ther group.Two late adverse events occurred, both in the salvage group (P = 0.133). Delayed bleeding occurred in a patient who had just resumed warfarin and stricture occurr- ed in a patient who had a circumferential resection requir- ing serial dilation and stent placement.	Small n. Salvage ESD isnsafe, achieving similar high rates of en-bloc resection and complete R0 resection as in treatment-naïve patients. During mean follow-up of 197 days, no cancer death or procedure-related deaths occurred.
Dennis Yang-2020-GIE	multicenter, prospective cohort study	No	No	205 BE ESD	median of 69 years ([IQR], 63-75)	ESD for BE visible neoplasia.	ESD			The primary aim was to assess the freequency of overall change of histopathologic diagnosis be- tween pre-ESD (biopsy sampling or EMR) and ESD pathologic specimens. The secondary aim was to identify potential factors associated with a change of histopatholog- ic diagnosis after ESD.	113 patients with a change of diagnosis, ESD resulted in upstaging in 77 (68.1%) and downstaging in 36 (31.9%) on the basis of baseline histopathology . On multivariate logistic regression, location at the distal esophagus or GEJ (OR, 2.1; 95% CI, 1.1-3.9; P = .02) and prior RFA (OR, 2.5; 95% CI, 1.2-5.5; P = .02) remained as factors associated with a change in histologic diagnosis after ESD.	

Dennis Yang-2018-GIE	meta-analysis	No	No	11 studies, 501 patients, 524 lesions		ESD in Early BE neoplasia				The pooled mean follow-up 22.9 months (95% CI, 17.5-28.3)	Early BE neoplasia was defined as either dysplastic BE (low- or high-grade dysplasia) or EAC based on preprocedural staging (ie, cross-sectional imaging, EUS, histopathology). Efficacy was determined based on the en bloc and R0 (complete) resection rates. En bloc resection was defined as exci- sion of the targeted lesion in a single specimen. R0 resection was defined as negative lateral and deep mar- gins for BE dysplasia and/or EAC in the ESD specimen. Secondary outcomes included curative resection rate and recurrence. Curative resection was defined when all the following were present: R0 resection, well to moderately differentiated histology, and absence of lym- phovascular invasion. In patients with curative resection, recurrence was defined as histologically confirmed BE neoplasia at the previous resection site on repeat endos- copies during the follow-up period. Adverse events included perforation, bleeding, and esophageal stricture formation. Immediate adverse events were defined as those occurring within 48 hours of the procedure and delayed referred to those occurring after 48 hours of the procedure.	Pooled estimate for en bloc resection was 92.9% (95% CI, 90.3%-95.2%). The pooled R0 and curative resection rates were 74.5% (95% CI, 66.3%-81.9%) and 64.9% (95% CI, 55.7%-73.6%), respectively. There was no association between R0 or curative resection rates and study setting (Asia vs West), length of BE, lesion characteristics, procedural time, or length of follow-up. The pooled estimates for perforation and bleeding were 1.5% (95% CI, .4%-3.0%) and 1.7% (95% CI, .6%-3.4%), respectively. Esophageal stricture rate was 11.6% (95% CI, .9%-29.6%). Incidence of recurrence after curative resection was .17% (95% CI, 0%-.3%) at a mean follow-up 22.9 months (95% CI, 17.5-28.3). The pooled incidence of recurrence, defined as histopathologic evidence of early BE neoplasia on follow-up after curative resection had been achieved, was .16% (95% CI, .1%-.3%; Cochran Q test P≥.51;I2 Z).	There were overall 11 studies, of which 8 were retrospective (group 1) and 3 were prospective observational (n Z 2) or a randomized controlled trial (n Z 1) (group 2). We performed subgroup meta-analysis on groups 1 and 2 separately. There was no statistically significant difference in the outcome measures in each group when compared with the pooled results from all 11 studies. Furthermore, we also compared the results be- tween groups 1 and 2 and did not find any statistically significant differences; however, this may be because of the limited studies in each group, particularly group 2.
D. Yang-2017-GIE.	Multicenter retrospective	No	No	46 patients	Median 69 years (range, 42-82 years).	ESD for BE-HGD or EAC	ESD			.	En bloc resection was defined as excision of the tar- geted lesion in a single specimen. R0 (complete) resection was defined as negative lateral and deep margins for HGD or cancer in the ESD specimens. R1 (incomplete) resection was defined as en bloc resection, but with microscopically positive lateral or deep margins for HGD or cancer. Cura- tive resection was defined when all the following were present: (1) negative lateral and deep margins, (2) well to moderately (G1-G2) differentiated histology, and (3) absence of lymphovascular invasion. Complete remission of neoplasia was defined as the absence of histologic evidence of HGD or EAC on biopsy specimens during follow-up evaluation. Early (<48 hours) and late (>48 hours) adverse events .	En bloc and curative resection rates were 96% (44/46) and 70% (32/46), respectively. Most lesions (11/20; 55%) diagnosed as BE-HGD on biopsy were upstaged to intramucosal or invasive EAC on post-ESD his- topathology. There were 4 early (<48 hours) adverse events (3 bleeding and 1 perforation), and all were treated endoscopically. Seven patients (15%) developed esophageal strictures that were managed endoscopically. Com- plete remission of BE neoplasia was found in 100% (32/32) of patients with curative resection	
Subramaniam-2017-GIE.	Multicenter retrospective	No	No	143 ESDs , 124 patients	mean 71.21 years.	BE neoplasia	ESD. 32 ESDs (22.4%) were performed in patients who had undergone previous endoscopic resection for Barrett's neoplasia. One patient required ESD for recurrence after esophagec- tomy and 2 others for recurrence after radical chemoradio- therapy for advanced cancer.			median follow-up time 21.6 months (interquartile range, 11.0-32.6)	En-bloc resection was defined as resection of the marked area in a single piece. R0 resection was defined as cancer-free deep and lateral margins. Dysplasia-free mar- gins were recorded concurrently. Sm1 lesions were defined as lesions with a depth of submucosal inva- sion <500 mm below the muscularis mucosae. Poor prog- nostic histologic features were defined as poor differentiation, presence of signet ring cells or LVI. Curative resection of cancer was defined as R0 resection of IMC in the absence of poor prognostic histologic features. Expanded curative resection of cancer criteria included R0 resection of both IMC and Sm1 cancers without poor prognostic histologic features. Local recurrence was defined as HGD/cancer noted within 2 cm of the previous ESD resection site. Neoplastic lesions identified >2 cm away from the ESD site were defined as metachronous.	The en-bloc resection rate was 90.8% and R0 resection rate 79% . Adverse event rate was 3.5% (1.4% bleeding, 0% perforation, and 2.1% stricture formation). The expanded curative resec tion rate was 65.8%. Multivariable logistic regression: submucosal cancer was identified as a significant factor affecting the R0 resection rate (odds ratio, .08; 95% confidence interval, .03-.22; P <.01). Follow-up: further neoplasia was seen in 16 cases. Seven of 16 were considered recurrent (within 2 cm of the ESD scar) and 9 of 16 as metachronous. Dis- counting the metachronous lesions, the true recurrence rate is 7 of 121 (5.8%). Four of 7 recurrent lesions and 8 of 9 metachronous lesions were treated by further endo- scopic resection.	
Abe-2019-GIE	Multicenter retrospective	No	No	Total 372: 51/321 (EMR/ESD). 204 SSBE, 34 LSBE	Median (range) 68 (30-90), y	histologically diagnosed adenocarcinoma, histologically diagnosed mucosal or SM cancer, and esophageal cancer (EA)or esophagogastric cancer (EAGJ) involving the esophagus.	EMR	ESD		The median follow-up period was 6.6 years with 5-year follow-up available in 77.7% of all patients.	Depth of invasion was classified as superficial muscularis mucosa (SMM), lamina propria mucosa (LPM), deep muscularis mucosa (DMM), SM 500 mm, or SM > 500 mm	Positive lateral margin in EMR compared with ESD (49.0% vs 7.5%, P < .01). There was no significant difference in positive deep margin between the 2 groups (13.7% vs 6.2%, P = .08). The en bloc resection rate and R0 resection rate in the EMR and ESD group were 60.8% and 99.1% and 49.0% and 87.9%, respectively (p < .01). 5-year cumulative incidences of local recurrence were 13% and .5% in the EMR and ESD group, respectively (P < .01). Metachronous AEGJ developed in 4 patients (1.3%) during a median follow-up period of 5.0 years, with 5-year follow-up avail- able in 50.0% of patients.	Only 34 had LSBE

Study	Design			Population	Age	Inclusion	Intervention	Comparator	Follow-up	Definitions	Results	Quality
Grischa Terheggen-2016-Gut	Prospective, RCT	Yes	No	ESD (n=20) or EMR (n=20). There were no losses and exclusions after randomisation.	Mean (SD): 64±12 (ESD) 65±11 (EMR)	BE with a focal lesion with HGIN or EAC, ≤3 cm were randomised to either ESD or EMR.	EMR	ESD	Mean 23.1±6.4 months	Complete resection: complete single piece (en bloc) resection of the targeted lesion plus histological confirmation of horizontal and vertical free margins (R0) for both EAC and HGIN. Curative resection of the targeted neoplastic area: histologically complete resection (R0 basal and lateral) of HGIN/mucosal EAC or EAC with low-risk superficial submucosal invasion. En-bloc resection, defined as resection of the targeted lesion including all coagulation markers in one piece in a single session, irrespective of basal and lateral tumour margins infiltrated or undetermined. CR from neoplasia and recurrences at FU, defined as histologically complete resection (R0) or incomplete resection (R1) of HGIN or EAC followed by at least one FU endos- copy including biopsies from the resection scar and mapping biopsies of the original BO segment, indicating no residual or metachronous area of HGIN or EAC. Recurrent or metachronous HGIN/EAC, defined as histolo- gically confirmed HGIN/EAC at the previous resection site or other areas of BO after previous CR-HGIN/EAC.	R0 was more frequent with ESD (10/17 vs 2/17, p=0.01). No difference in complete remission from neoplasia at 3 months (ESD 15/ 16 vs EMR 16/17, p=1.0). Recurrent EAC in one case in the ESD group. Elective surgery was performed in four and three cases after ESD and EMR, respectively (p=1.0). 2 perforations in ESD group and none for EMR (p=0.49).	Poor quality ESD: 10% of perforations in the ESD group, R0 only 58%. Really small n, in most of evaluated parameters p > 0.2 due to poor sample size
Pech-2014-Gastroenterology	Prospective database	No	No	1000 consecutive patients	mean 69.1. (+- 10.7) years	Mucosal adenocarcinoma of the esophagus (mAC). Patients with low- grade and high-grade dysplasia and submucosal or more advanced cancer were excluded.	EMR+ ablation		Mean 56.6 (+ - 33.4) months	Complete remission of HGD and cancer: R0 resection plus one normal endoscopic checkup examination. In R1 or Rx situations (Rx meaning that evaluation of the margin was not possible due to coagulation artifacts) on the lateral margin of the resected specimen, 2 consecutive endoscopies without evidence of resid- ual tumor were required to conclude that there was a complete response. Tumor-associated death: death caused by metastatic esophageal adenocarcinoma or metastatic adenocarcinoma from an unknown primary. Long-term complete remission rate: status of complete remission at time of writing, inde- pendently of previous recurrences. Major complications were defined as perforation and bleeding, with a decrease in the hemoglobin level of >=2 g/dL. Minor complications consisted of symptomatic esophageal strictures requiring dilation or bougienage.	963 patients (96.3%) with complete response; surgery needed in 12 patients (3.7%) after endoscopic therapy failed. Metachronous lesions or recurrence of cancer in 140 patients (14.5%) but endoscopic retreatment was successful in 115, resulting in a long-term complete remission rate of 93.8%. 2 died of BE cancer. The 10-year survival rate of ER was 75%. Major complications: in 15 patients (1.5%) managed conservatively.	

| PICO search string: | ((adenocarcinoma) AND (esophageal)) AND (submucosa): 172, ((adenocarcinoma) AND (esophagus)) AND T1b: 79, cancer AND (esophagus)) AND T1b: 122.   April 2022 |

| Table 24 | | | | | | low-risk T1b surgery vs endoscopy | | | | | |

	Methods			**Population**			**Intervention**			**Outcomes**		
Author (year)	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	Remarks
Dunn, 2022	retrospective	no	no	133/136 52/19 T1b	NA	Esophagectomy or ER for T1 to T3 EAC	esophagectomy	EET	NA	all cause mortality, disease specific mortality, 30 Day mortality, severe (CD 3 or more) AE, patient stay, hospital costs	all cause mortality = (HR 1.85, 95% CI 0.73, 4.72) disease specific mortality : (HR 1.10, 95% CI 0.26, 4.65), 30 D mortality = 0%, severe AE = 26 % vs 0,7%, p<0,05, hospital stay 14 vs 0, costs 16360 vs 8786 GBP	9 / 12 T1b sm 1 , outcomes not specifcally reported for this subgroup
Saunders, 2020	retrospective	no	no	96/93	69	Esophagectomy or EMR for T1 EAC	esophagectomy	EMR	5,5 years	rate of LN metastases, overall survival, disease free survival,	LN metastases = 2% in LR sm1, 15 % in HR sm1, OS = 65% vs 94% at 5 years, no p value… DFS : HR = NS in multivariate analysis between surg and emr	14/184 sm1 patients 0% LN metastases among the 9 LRsm1 patients
Qin, 2020	retrospective (propensity score analysis)	no	no	196/224	80	esophagectomy or endoscopic therapy for T1 EAC	esophagectomy	endoscopic therapy (various)	NA	5 year OS , cancer specific survival	5 year OS = 50,5% vs 40% p= 0,4, CSS = 70,3% vs 72,2%, p=0,6	heterogeneity in endoscopic therapies (ablative techniques included) 80% EAC, 51% T1b in the esophagectomy group vs 21% in the ER group proportion of low risk T1b EAC ?
Nelson, 2018	retrospective	no	no	49/23	64/71	esophagectomy or endoscopic therapy for T1b EAC	esophagectomy (ivor lewis> minimally invasive > transhiatal)	EMR	43	local recurrence, regional recurrence, distant recurrence, cancer related mortality, DFS	N+ on surgery = 17 % 3Y local recurrence = 0% vs 23 %, p<0,05, 3Y regional recurrence = 5,6 % vs 4,7%, p = 0,6, 3Y distant recurrence = 8,1 vs 13,2 %, p=0,4. CSS and DFS not specified	sm1 = 43% in the esophagectomy group vs 74 % in the EMR group, p = 0,01 - LVI 48 % vs 26 %, G3 52 % vs 30 % proportion of low risk T1b ?
Zeng, 2017	retrospective (propensity score analysis)	no	no	621/624	NA	esophagectomy or endoscopic therapy for T1 esophageal cancer	esophagectomy	endoscopic therapy (various)	92/45	OS, CSS,	for the subgroup of EAC : OS : HR = 1,3, 95%CI0,81-1,96, p = 0,3, CSS : HR = 0,5, 95% CI 0,27-0,88, p = 0,085.	75 % EAC - 15 % of T1b disease - proportion of low risk T1b EAC ? similar OS and CSS after esophagectomy or endoscopic therapy for early adenocarinoma
Chu, 2018	markow model	no	no	NA	75	Esophagectomy or EMR for T1b EAC	esophagectomy	EMR	until death	CSS, QALYn cost effectiveness	esophagectomy yielded more QALY than EMR 4,07 vs. 3,85, but not cost effective : incremental cost effectiveness ration 156981 USD	modelization study based on SEER data proportion of low risk T1b ?
Tian, 2011	retrospective	no	no	39/29	68	EMR of T1b EAC	esophagectomy	surveillance	NA	median survival (OS)	13 % N+ disease in the esophagectomy group - median survival = 49 vs 35 months, p = 0,09	comparable groups - no significant survival gain with esophagectomy after ER for T1b. No mention of DFS. proportion of low risk T1b EAC ???.
Kamarajah , 2020 (abs)	retrospective	no	no	946/3352	NA	Esophagectomy or EMR for T1b N0 EAC	esophagectomy	EMR	na	5 year OS	15 % N+ disease among operated patients ; 5 Y OS = 61 % vs 53 %, p = 0,3 ; Hr = 0,87, 95% CI 0,7-1,04, p= 0,1	proportion of low risk T1b EAC unknown, but T1b N+ excluded from this analysis
Ramai, 2020 (abs)	retrospective	no	no	367/84	67	Esophagectomy or EMR for T1b EAC	esophagectomy	EMR	na	CSS	CSS at 5 years 73% vs 74 %, p = NS	esophagectomy and ER associated with comparable long term outcomes in terms of CSS
Otaki, 2020	retrospective	no	no	68/73	64/73	Esophagectomy or EMR for T1b EAC	esophagectomy (transthoracic or transhiatal)	EMR	49/43 months	OS, DFS	5 y OS = 89% vs 59 %, p<0,01, 5 Y DFS = 92 % vs 69 %, p = 0,09	patients in the ER group older with more comorbidities. 26% recieved CRT in the endoscopy group. 57% had deep R1 margins … proportion of low risk T1b EAC ???
Datta, 2019 (abs)	retrospective	no	no	286/141	na	Esophagectomy or EMR for T1b EAC	esophagectomy (transthoracic or transhiatal)	EMR	55/42 months	OS, 30 and 90 day mortality	30 D mortality = 2,1% vs 0,8 %, p = 0,3 90 D mortality = 3,9 % vs 2,2 %,p = 0,4 OS = HR = 1,051 (95% CI 0,69-1,59, p = 0,8)	proportion of low risk T1b = ? Similar OS at 5 years after esophagectomy or ER, despite the unusual post ER mortality
Rosmolen, 2010	retrospective	no	no	27/64	3	esophagectomy or EET for Tis or T1 EAC	esophagectomy (tt ou th)	EET	23 months	QOL, fear of cancer recurrence, anxiety 12 to 60 months after treatment	QOL (SF 36 and EORTC QLQ C30) : similar results ; more eating problems (OR 18, p<0,001), GER symptoms (3,4, p = 0,05) in the surgery group - more fear of cancer recurrence i=(p = 0,003) in the ER group.	33/8 T1b EAC in total proportion of low risk T1b = ? long term Impaired qol after esophagectomy
Bennett, 2020	systematic review	no	no								"No data were available to permit an analysis of the relative benefits or harms of endotherapies compared with surgery"	no specific data on low risk T1b
Scholvinck, 2016	retrospective	no	no	69, 14 R0 LR tumor	70	EMR for T1b EAC	EMR	none	60 months	distant recurrence	0% recurrence in the R0 resected LR T1b (1/14 operated), 16 %for the R0 resected HR group (12/32 operated), 30 % for the R1 resected group (13/23 operated)	ER and follow-up is safe in LR T1b

Author, year	Design			N	Age	Intervention	Treatment	Comparator	FU	Outcomes	Results	Comments
Manner, 2015	retrospective	no	no	72 T1b sm1 EAC	62	EMR for T1b EAC	EMR	none	60 months	LN metastases,	LN metastases in 2% of LR T1b patients vs 9 % HR T1 patients, p = 0,24	First study to answer the question nb = mortality of eophagectomy was 3% nb 2 = only 8 patients operated in the LR group, - LN status assessed by sugery or FU, mixing LN metastases and regional recurrences
Manner, 2013	retrospective	no	no	66 LR T1b EAC	63	EMR for T1b EAC	EMR	none	47 months	LN metastases (= distant recurrence), 5 Y OS, AE,	LN metastase (distant recurrence) = 1,9 %, 5 Y OS = 84 %, major AE = 1,5 %	supports the safety of EMR for low risk T1b EAC
Nieuwenhuis, 2022	retrospective	no	no	55 LR T1b	76	EMR for T1 EAC	EMR/ ESD	none	30 months	LN metastases during FU	LN metastases rate = 2 % (1/55) at 30 months, annual risk = 0,7 %, vs 8 % for T1bHR (annual risk = 3%), and 20 % in the T1a HR group (annual risk 7%)	supports the safety of EMR for low risk T1b EAC ,and the importance of poor histoprognostic factors even in T1a lesions
Gotink, 2021	retrospective	no	no	248 T1b, 56 T1bsm1	66	T1b EAC treated by ER or esophagectomy	EMR/ESD/Esophagectomy	none	5,5 years	LN metastases, distant metastases	30% cumulative incidence of metastases, 5,9% for LR T1 <20 mm, 16,1% for LR T1 > 20 mm	20 % R1 - aggregation of positive LN at esophagectomy and regional recurrence diagnosed during FU
Frei, 2019	retrospective	no	no	13 T1b	66	EMR of T1 EAC	EMR	none	36 months	endoscopic treatment success, cancer related death	9 LR T1b, no local or distant recurrence	small sample size, short FU
graham, 2018	retrospective	no	no	60 T1b, 13 LR T1b	70	EMR of T1 EAC, followed or not by esophagectomy	EMR / Esophagectomy	nonne	41 months	LN metastases, rate	LN rate = 0 % in LR T1b, 21 % in the HR T1 b lesions	12/13 LR T1b followed up and 1/13 operated on - aggregation of positive LN at esophagectomy and regional recurrence diagnosed during FU
Sihag, 2021	retrospective	no	no	158 T1b	65	esophagectomyfor T1 EAC	esophagectomy	none	48 months	LN metastases, distant recurrence during FU	LN metastases = 18,4 % in T1b lesions, distant recurrence at 5 years : 17 % fot T1b vs 6% for T1a , p = 0,018	median interval to recurrence = 2 years no data on LR T1b
Lorenz, 2014	retrospective	no	no	126 T1b, 37 T1b sm1	63	esophagectomyfor T1 EAC	esophagectomy	none	64 months	LN metastases, distant recurrences during FU	LN metastases 8% in sm1, vs 27 % in sm 2 and 25 % in sm3- distant recurrences = 8 % in T1bsm1 vs 4% in sm2 and 21 % in sm3 - 5 Y OS = 83% for T1b sm1 vs 100% and 89 % fot T1am1-m3 and 68 %-70 % for sm2 and sm3, p = ns. Tumor recurence m1-sm1 = 5,7% vs sm2-sm3 = 22%, p = 0,021	only data on sm1 vs T1a or sm2/3, no specific data on LR T1b. However, the 8% recurrence rate after esophagectomy for T1b sm1 should be noted
Oetzùann von sochaczewski, 2020	retrospective	no	no	164 T1b, 42 T1bsm1	na	esophagectomyfor T1 EAC	esophagectomy	none	87 months	LN metastases at surgery, distant recurrence during FU	LN metastases 9 %, distant recurrence 9,5%	no data on LR T1b
Leers, 2011	retrospective	no	no	51 T1b, 19 T1bsm1	64	esophagectomy for T1 EAC	esophagectomy	none	50 months	LN metastases, 5 year OS	21% LN metastases in T1bsm1, 5 % (1/19) in low risk T1b, 5 year OS = 79 %	the only surgical serie with data on LR T1b (n= 1 patient)
Gamboa, 2016	retrospective	no	no	353 T1b	na	esophagectomy for T1 EAC	esophagectomy	none	na	LN metastases	positive LN at surgery in 9 % of T1b G1 and <2 cm vs25 % in other T1b lesions	no data on LVI, no specific data on T1b sm 1 , and on low risk T1b lesion
Newton, 2018	retrospective	no	no	415 T1b	64	esophagectomy for T1 EAC	esophagectomy	none	33 months	LN metastases, 5 Y OS	LN rate in G1 T1b <2 cm = 4,2% 5 year OS = 64%inT1b, T1b stage increased the risk of death at 5 years (HR = 1,8 (1,07-3,02, p= 0,03)	no clear LN rates for low risk T1b number of T1b sm 1 ?
Merkow, 2014	retrospective	no	no	2159 T1b		T1 esophageal cancer	esophagectomy	none	NA	LN metastases, distant recurrence during FU	positive LN at surgery in 17 %, 9 % in LR T1b <2cm	89 % of EAC number of T1b sm 1 ? no direct comparison possible between esophagectomy and ER for T1b, and even less low risk T1b lesions. Many missing information : LVI ? Distant recurrence ? R1/R0 ?
Westerterp, 2005	retrospective	no	no	66 T1b	65	esophagectomy for T1 EAC	esophagectomy	none	40 months	LN metastases, distant recurrence during FU	LN + : 0% - distant recurrence : 4%	25 t1b sm 1 - proportion of LR tumors ? (transhiatal esophagectomy

Table 25					management of high-risk T1a							
Author (year)	**Methods**		**Population**			**Intervention**			**Outcomes**		**Remarks**
	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	
Boys 2015	retrospective	no	19		Esophagectomy of T1 Barrett's cancer after ER		NA	NA	LN metastasis	No LN metastasis in T1a despite presence of LVI in 5.3%, poor differentiated AC in 21.1%	
Pech 2014	retrospective	no								54 patients with poorly differentiated AC: no LN metastasis	
Oetzmann von Sochaczewski C 2020	retroespective	no	53		Esophagectomy of T1 Barrett's cancer					12 patients with LVI: 2 LN metastasis	

Table 26				Management of high-risk T1b							
Author (year)	Methods			Population					Intervention		Outcomes			

Author (year)	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	Remarks
Gotink-2022-Endoscopy	Retrospective, national	No	No	248	65.6 medican(57.8–72.5 IQR)	T1b	Treatment with ER (33.1%) or surgery (66.9%). 59.8% of ER were followed by surgery		The median follow-up time was 5.5 years (IQR 4.9–7.7) in patients treated with endoscopic resection only and when < 12 lymph nodes were present in surgical resection specimens. The median follow-up time was 3.3 years (IQR 1.8–5.3) in patients treated with primary surgery with ≥ 12 lymph node dissections during surgery.	specimens (≥ 12 resected lymph nodes), or the development of metastases during follow-up.	5-year cumulative incidence of metastasis was 30.9%, and increased with submucosal invasion depth (subdistribution hazard ratio [SHR] 1.08, 95 %CI 1.02–1.14, for every increase of 500 μm), lymphovascular invasion (SHR 2.95, 95 %CI 1.95–4.45), and for larger tumors (SHR 1.23, 95 %CI 1.10–1.37, for every increase of 10 mm).	The described model demonstrated good discriminative ability (c-statistic 0.81, 95 %CI 0.75–0.86). Differentiation grade was only found to be significant in univariable analysis. The vast majority of metastasis occurred within the first year, which suggests that the metastatic spread may have already occurred at the time of resection, and that our ability to assess for early metastatic spread is limited
van de Ven-2021-UEG Journal	Retrospective, national	No	No	248	65.6 medican(57.8–72.5 IQR)	T1b	Treatment with ER (33.1%) or surgery (66.9%). 59.8% of ER were followed by surgery		The median follow-up time was 5.5 years (IQR 4.9–7.7) in patients treated with endoscopic resection only and when < 12 lymph nodes were present in surgical resection specimens. The median follow-up time was 3.3 years (IQR 1.8–5.3) in patients treated with primary surgery with ≥ 12 lymph node dissections during surgery.	specimens (≥ 12 resected lymph nodes), or the development of metastases during follow-up.	On multivariable analysis: Invasion depth (subdistribution hazard ratio [SHR] 1.08,95 %CI 1.02–1.14, for every increase of 500 μm), lymphovascular invasion (SHR 2.95, 95 %CI 1.95–4.45), and for larger tumors (SHR 1.23, 95 %CI 1.10–1.37, for every increase of 10 mm).	Inter-observer variability was excellent for LVI (κ = 0.88). Presence of only 1 LVI focus was not an independent predictor for metastases.
von Sochaczewski -2020-World J Surg	Restrospective, single center	No	No	217 patients (53 T1a and 164 T1b) treated by esophagectomy		EAC treate with esophagectomy	Esophagectomy in high-risk cases. In addition, suspicious lymph nodes on EUS failed ER, and repeated local tumor recurrences were also subject to surgical treatment.		Median FU 87 and 75 months for patients operated for T1a and T1b tumors	A high-risk: G3 grading, venous or lymphovascular invasion in the staging endoscopy.	Multivariate analysis, risk for metastatis: OR for tumor infiltration depth increased to 1.31 (95% CI 1.02–1.67; P = .033) per infiltration depth level, and theOR for lymphatic vessel invasion was 3.34 (95% CI 1.70–6.56; P < .001). Males also had a higher risk OR 4.60 (95% CI 1.03–20.6; P = .046). In comparison with sm1 tumors , the OR for any metastases increased to a similar level in both sm2 tumors with 3.44 (95% CI 1.00–11.9; P = 0.049) and sm3 tumors, whose OR was 3.88 (95% CI 1.25–12.04; P = 0.019) in multivariate analysis	The risk of metastases increased with the number of risk factors. If all three risk factors were present, 75% had lymph node metastases at surgery and 75% developed tumor recurrences during follow-up

Study	Design			N	Age	Population	Treatment	Follow-up	Outcome	Results	Comments
Otaki-2020-GIE.	Retrospective, multicentric	No	No	141	64.1 (59.8, 71.9) surgical cohort, 73.4 (65.8, 80.5) endoscopic cohort <.001	T1b EAC, Thise with neoadjuvant treatment before surgery or endoscopic resection were excluded.	48% esophagectomy, 52% endoscopically treated (EMR)	Median follow-up 49.4 months in esophagectomy and 43.4 in endoscopic cohorts	overall and cancer-free survival.	5 year overall survival rates in the surgical and endoscopic cohorts were 89% and 59%, respectively. 5-year cancer-free survival rates were 92% and 69% (P =.09). The presurgical EMR histology in patients with metastatic lymphadenopathy had cancer involving the deep margin in 8 (80%), lym- phovascular invasion (LVI) was positive in 3 (30%), and was poorly/undifferentiated cancer in 3 (30%). Presence of any high-risk histologic feature was associated with increased overall mortality in the entire cohort (hazard ratio, 2.20; 95% confidence interval, .78-6.17; P Z .014). When stratified into the surgical and endoscopic cohorts, this association was present only in the endoscopic cohort.In multivariable analysis, only esophagectomy (HR 0.22) and deep margin positivity (HR 2.46) were correlated with mortality.	In the EMR roup 57.5% had positive deep margin. Two patients (2.9%) died of postoperative adverse events within 90 days of surgery. Cause of death not available in 68% and 43% of patients in the endoscopic and surgical groups, respectively. Only 8 (17%) had documented esophageal cancer– related mortality. Esophagectomy was associated with improved overall but not cancer-free survival (comparing to EMR)
Mitchell-2020-Eur Surg	systematic review	No	No	16 studies, a total of 1382 T1 cases were included: T1b adenocarcinoma (849 patients)Subgroup analysis of T1b lesions available in 8 reports (365 patients).		T1a and T1b oesophageal adenocarcinoma, esophagectomy, lymph node yield for the study ≥15. Neoadjuvant therapy cases were excluded	esophagectomy	NA	Incidence of node positivity	T1b disease a rate of 22%of node positivity. For SM1, SM2 and SM3 was 19.9%, 20.2% and 32.9%, respectively. The rate SM3 is significantly higher than for SM1 or SM2 disease (p = 0.001) with no significant difference between SM1 and SM2 disease	there is no analysis of factors correlated with Incidence of node positivity
von Sochaczewski-2019-Diseases of esophagus	Restrospective, single center	No	No	217 patients (53 T1a and 164 T1b) treated by esophagectomy	mean 64 years in T1b (range 38–82)	EAC treate with esophagectomy		Median FU 75 months for patients operated for T1b tumors	Lymph node metastases	LNM in sm1 (4/42), sm2 (11/44) and sm3 tumors (18/78). Locoregional recurrences were rare in sm1 (1/42), and sm2 (2/44), but frequent in sm3 (12/78). Distant metastases: 4/42 in sm1, 4/44 in sm2, and 13/78 in sm3. Overall metastatic rates of 11.9% in sm1, 27.3% in sm2, and 32.1% in sm3 tumors. The 5-year overall survival rates were 90% in sm1, 66% in sm2, and 71% in sm3 tumors (P = .020) (Fig. 3A). The 5-year disease-specific survival rates were 97% in sm1, 91% in sm2, and 80% in sm3 tumors, respectively.	The 5-year overall survival of patients with sm2 tumors was comparable to that of patients with sm3 invasion depth (66% vs. 71%), although both the disease-specific survival and the recurrence-free survival were higher in patients with sm2 invasion depth (91% vs. 80% and 87% vs. 77%)
Ramay-2019-Ann Surg Oncol	systematic review	No	No	27 retrospective studies,2 prospective studies,1 meta-analysis		T1N0 Esophageal Adenocarcinoma					The data of different studies is not combined
Kunzli-2018-UEG Journal	Restrospective, single center	No	No	35 patients (17 low-risk, 18 high-risk)	68 (median)	Submucosal EAC treted with ER (EMR,ESD)	FU with endoscopy and EUS.	median months 23 (IQR 15–43). Median follow-up was 22 (IQR 15–47) months for low-risk patients, and 23 (IQR 15–39) months for high-risk patients.	Primary outcome: number of patients diagnosed with LNM; secondary outcomes:intraluminal recurrences.	None were diagnosed with LNM. Five (14%) patients developed a local intraluminal recurrence a median of 18 (IQR 11–21) months after baseline ER that were treated endoscopically. In none of the patients were distant metastases diag- nosed during follow-up.	2 pts FU discontinued: 1 patient died 11 months after initial ER of a non- tumor-related cause, 1 patient refused additional follow-up because of advanced age. Small series.

Study	Design			N / Patients	Age	Population	Intervention	Follow-up	Outcome	Results	Additional Findings
Graham-2018-World J Gastroenterol	Restrospective, single center	No	No	60 patients: 22 surgically managed (1 low-risk and 21 high-risk patients), 38 treated conservatively (12 low-risk and 26 high-risk)	70 (IQR 66-75)	T1b EAC managed with either surgery or ER (EMR).	EMR followed by surgery or conservative treatment	median 45 mo (IQR 32-72)	LNM	All 10 pts with LNM occurred in patients with high- risk lesions (21% of the total high-risk lesions). RF for LNM Sm1 vs Sm2/3/X (P = 0.035); no LVI Vs positive for LVI/X (P = 0.012); G1/2 Vs G3 (P = 0.488); and R0 vs R1 (P = 0.049). No statistically significant difference in tumor-related deaths and disease-specific survival between surgery or ER (P = 0.636) and (P = 0.376), respectively. In high-risk group 14/ 47 (30%) died during FU. Of these, 7 died of EAC. No significant difference for EAC-related deaths during the follow-up period when comparing treatment modality in high-risk patients	During the follow-up period 3/60 (5%) patients developed a local recurrence, 1 patient from the LR group (surgery) and 2 patients from the HR group (EMRs).All treated sucessfully with EMR
Nelson-2018-J Thoracic and Cardiov Surg		No	No	72 pts		T1b adenocarcinoma treated with EMR or esophagectomy. Excluded: EMRs with a positive deep margin were excluded from analysis, pre-treatment chemoradiation, previous treatment for their index esophageal cancer lesion, were medically unfit to receive any further chemoradiation or surgery, or were lost to follow-up.	23 EMR with esophageal preservation and 49 esophagectomy.	Median: 43.5 months esophagectomy and 45.1 months EMR	High-risk patients were those with LVI or larger, deeper, or poorly differentiated lesions. Depth of invasion, tumor grade, and presence of LVI were used as covariates to construct the propensity model.	EMR associated with an increased risk of local recurrence (P = .01), but not distant recurrence (P = .44). No difference in distant recurrence rate (P = .66). Low-risk patients with no recurrences or cancer-related deaths. High-risk patients showed a trend toward increased distant recurrence after EMR. Patients with high risk of nodal metastasis had a trend toward increased risk of distant recurrence with esophageal preservation (P = .066). After propensity matching, patients who underwent esophageal preservation continued to show an increased risk of local recurrence, although this difference was no longer statistically significant (P = .083). However, there continued to be no difference in the risk of regional or distant recurrence (P = .578 and P =.658, respectively).	Low n
Weksler-2017-Journal of Thoracic and Cardiovascular Surgery	Retrospective, national	No	No								
Newton-2017-Ann Surg Oncol	Retrospective, national	No	No	782 patients (512 T1b)	Mean 64.1 ± 9.6	Tis, T1a, or T1b EAC who had primary surgical resection and microscopic examination of at least 15 lymph nodes.	Surgery	32.9 months median follow-up	LNM, Survival	With T1b tumors, the LNM rate with LVI compared with no LVI was 43.3% versus 15.6% (OR 4.15, 95% CI 2.44–7.06, P \ 0.001). LNM was more likely with mod- erately differentiated (OR 3.04, 95% CI 1.03–8.99, P = 0.044) and poorly/undifferentiated tumors (OR 6.05, 95% CI 2.08–17.64, P = 0.001) than well-differentiated tumors. LNM rates with poorly/undifferentiated, moder- ately differentiated, and well-differentiated tumors were 33.5, 20.2, and 7.7%, respectively. With tumors C 2 cm, the LNM rate was 29.2% compared with 18.1% in patients with tumors \ 2 cm (OR 2.39, 95% CI 1.63–3.49, P \ 0.001).	Perioperative Mortality: 2.9% (30 days), 5.0% (90 days). For well-differentiated T1b tumors <2 cm without LVI, the LNM rate (4.2%) was lower than 90-day mortality (6.0%). On multivariable Cox regression, pT stage 1b and LNM were independent predictors of worse survival (P < 0.05). Patients with T1a tumors and poor differentiation or size >2 cm had a higher LNM rate than patients with T1b tumors without other high-risk tumor factors.

Study	Type			N	Age	Cancer/Inclusion	Treatment		Follow-up	Outcome	Results	Conclusion
Gamboa-2016-Cancer.	Retrospective, national	No	No	353 T1b	NA	EAC	Only patients with localized disease at the time of diagnosis who underwent surgical resection with lymph node dissection and pathologic exami- nation of lymph nodes were included		NA	LNM	T1b tumors: 26.5% with high-grade histology were found to have lymph node involvement compared with 14.6% with low-grade histology (P =.008). Patients with tumors measuring >2 cm in size, the incidence of lymph node positivity was 24.3%, compared with 12.1% in patients with tumors measuring <2 cm in size (P =.006). Among patients with low-grade tumors measuring <2 cm in size, the incidence of lymph node positivity was 8.6% (P 5.001)	There was no significant associa- tion observed between lymph node status and patient age, patient sex, patient race, geographic location, or location of the tumor within the esophagus
Ishihara-2016-J Gastroen	Retrospective, multicentric. Japan	No	No	154 SM EAC	64.5 (MEDIAN)	Esophageal or esophagogastric carcinoma. Exclusion criteria were: (1) chemotherapy or radiation before surgery, (2) chemotherapy or radiation before or after ER, (3) deeper invasion (SM invasion C500 lm) in the cardia than in the esophagus, and (4) any other primary invasive cancer that had been treated within the preceding 5 years.	ER or surgery		surgery: median (range) follow- up period of 9 (2–24) months. ER: > 5 years	Metastasis was considered positive if one of following criteria was fulfilled: (1) histologically confirmed metas- tasis in surgical specimen or (2) clinically confirmed metastasis during follow- up after surgery or ER.	Multivariate analysis in deep MM and SM cancers: lymphovascular involvement [odds ratio (OR) 6.20; 95 % confidence interval (CI) 3.12–12.32; p < 0.001], a poorly differentiated component (OR 3.69; 95 % CI 1.92–7.10; p < 0.001), and lesion size (OR 3.12; 95 % CI 1.63–5.97; p = 0.001) as independent risk factors for metastasis.	
Manner-2016-Diseases of Esopahgus	Restrospective, single center	No	No	38 sm2 lesions, 69 sm3. 23/38 pt with pT1b sm2 lesions and 39/69 pt with sm3 lesions fulfilled inclusion criteria.		EAC. LN rate was only evaluated in pt who had proven maximum invasion depth of sm2/sm3, and who in case of ER had a FU by EUS of at least 24 months.	ER or surgery		FU in the patients who underwent endoscopic therapy mean 41–43 months.	Rate of LN metastasis was analyzed. histologically low-risk (hisLR): G1-2, L0, V0; histologically high-risk (hisHR): ≥1 criterion not fulfilled; macroscopically low-risk (macLR): gross tumor type I-II, tumor size ≤2 cm; macroscopically high-risk (macHR): ≥1 criterion not fulfilled; combined low-risk (combLR): hisLR+macLR; combined high-risk (combHR): at least 1 risk factor.	The overall rate of LN metastasis was 21.7% in pT1b sm2 (5/23) and 35.9% (14/39) in sm3 lesions. In the pT1b sm2 group, rate of LN metastasis in the hisLR, hisHR, combLR, and combHR groups were 8.3% (1/12), 36.3% (4/11), 0% (0/5), and 27.8% (5/18). In the pT1b sm3 group, rate of LN metastasis in the hisLR, hisHR, combLR and combHR groups were 28.6% (2/7), 37.5% (12/32), 25% (1/4), and 37.1% (13/35).	30-day mortality of surgery was 1.7% (1/58 pt).In EAC with pT1b sm2/3 invasion, the frequency of LN metastasis depends on macroscopic and histological risk patterns. no treatment-related mortality of endoscopic therapy.

Study	Type			N patients	Age	Population	Intervention	Comparison	Follow-up	Definitions	Results	Comments
Schölvinck-2016-Sur Endosc.	Restrospective, 2 centers	No	No	69 pats [23 R1-resections and 46 R0-resection (14 R0-LR and 32 R0-HR)]	70 (64–76)	Submucosal EAC in the ER-specimen. Exclusion criteria were the presence of proven lymph node or distant metastasis , chemotherapy, radiotherapy after ER , lack of follow-up.	EAC patients undergoing surgery versus conservative therapy. 26 patients underwent surgical treatment (1 R0-LR, 12 R0-HR and 13 R1).		med- ian FU of 60 months.	R1 considered when tumor extended beyond in vertical margin. EACs classified as low risk (LR; submucosal invasion <500 nm, G1–G2, no LVI) or high risk (HR; deep submucosal invasion [500 nm, G3– G4 and/or LVI). Metastatic disease defined as LNM in surgical resection specimen and/or evidence of malignant disease during FU.	12 (17 %) patients developed LNM and/or distant metastasis. The rate per group was 0 % for the low-risk R0 group, 16 % for the high-risk R0-group and 30 % for the patients with an irradical ER (R1 group).None of the 14 R0-LR patients developed metastatic disease after a med- ian FU of 60 months. In the R0-HR group and R1 group, metastatic disease was diagnosed in 16 and 30 % of patients, respectively. Surgical patients tended to have a better overall survival than non-surgical patients (p = 0.09). Tumor-related deaths, however, were 12 % in both groups. There was no in-hospital mortality and no 30- and 90-day mortality. However, one HR patient (4 %) died because of a tra-cheoesophageal fistula (trachea to gastric tube) 50 months after surgery.	Small n Kaplan–Meyer analysis shows a trend of better overall survival after surgery than after conservative (endoscopic) therapy (log-rank test p = 0.09), based on all causes of death. Additional Kaplan–Meyer analysis shows that there was no difference in occurrence of tumor-related deaths between the two treatment strategies (p = 0.88).
Manner-2015-Surg Endosc.	Restrospective, single center	No	No	72	62.3 ± 9 year (range 44–81) LR. 62.8 ± 10 year (range 40–78) HR group.	EAC Sm1	LN metastasis was only evaluated in patients who had a proven maximum invasion depth of sm1 (ER and/or sur-gery), and who in case of ET had a follow-up (FU) by EUS of at least 24 months.		In endoscopically treated LR patients (37/49), mean EUS-FU was 60 ± 30 mo (range 25–146); in HR patients under-going ET (6/23), it was 63 ± 17 mo (46–86; p = 0.4).	low-risk (LR; G1–2, L0, V0) and high-risk lesions (HR; G3, L1, V1; C 1 risk factor)	49 patients had LR (68 %) and 23 HR lesions (32 %). The rate of LN metastasis was 2 % in the LR (1 patient) and 9 % in the HR group (2 patients; p = 0.24). Mortality of esophagectomy was 3 %. 3 deaths in the 49 LR patients (6 %). No tumor-related death. In the HR group, 2 deaths were observed (8.7 %). One tumor-related death (4.3 %). One patient died of a SIRS within 30 days after surgery. No significant difference (p = 0.65) number of deaths in the LR and HR group.	The rate of LN metastasis in pT1b sm1 early adenocarcinoma with histological LR pattern was lower than the mortality rate of esophagectomy.
Tian-2011-GIE	Restrospective, single center	No	No	68 patients	68.3 (mean)	T1b EAC	Surgery (39)	No surgery (29): 5 (17.2%) CRT, 15 (51.7%) endoscopic treatment, 5 (17.2%) both CRT and ET, 4 (13.8%) had neither			Among patients who underwent esophagectomy, 13 (33.3%) had LNM, and the mortality rate was 50.0% and 11.1% for those with and without LNM, respectively (P <.01).Overall mortality not significantly different for patients with and without esophagectomy (25.6% and 37.9%, respec-tively, P =. .28). LNM found in esophagectomy specimens had worse survival duration (hazard ratio 6.8, P = .02). Other factors not associated.	After adjustment for age, sex, and Charlson comorbidity index, esophagectomy was not associated with mortality rate in logistic regression or survival duration in Cox pro- portional hazard analysis
Herrero-2010-Endoscopy	Restrospective, 2 centers			82 patients: 57 m3, 12 with sm1, 13 with sm2/3 cancers.	70 (59 – 78)	ER EAC if the endoscopic resection specimen showed m3 or submucosal cancer.	ER		Median 26 months (interquartile range [IQR] 14 – 41).	Infiltration into the muscularis mucosae was defined as m3. Submucosal invasion was classified as sm1 (≤ 500 μm) or sm2/3 (> 500 μm).	13 were poorly differentiated and 5 had lymphovascular invasion. After initial ER, seven patients underwent surgery and 75 endoscopic therapy. No lymph node metastases were found in 158 lymph nodes of the esophagect-omy specimens. None of the endoscopically treated patients were diagnosed with lymph node metastasis.	
Gockel-2011-Expert Review Gastro & Hepat	Systematic review	No	No	7645 patients (SCC and EAC)		Submucosal esophageal cancer (EAC and SCC)				Lymph node (N), lymphatic (L) and vascular (V) invasion	EAC: LNM, 6% in sm1, 23% in sm2 and 58% in sm3; The L+:N+ ratio remained constant from sm1 to sm3 tumor depths (sm1: 1.5; sm2: 1.17; sm3: 1.31), suggesting that lymphatic channel invasion may be a predictor of lymph node metastasis in EAC; No vascular invasion was found in sm2 and sm3 patients with ADC.	

PICO search string: ((adenocarcinoma) AND (esophageal)) AND (submucosa): 172, ((adenocarcinoma) AND (esophagus)) AND T1b: 79, cancer AND (esophagus)) AND T1b: 122.   April 2022

Table 27	**What is the most optimal medical therapy after EET (long-term PPI?)**									
Author (year)	Methods		Population			Intervention			Outcomes		
	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	Remarks
Phoa (2013)	4 Cohort studies, prospectives	no	46	65 + 10	CR at last endoscopy after RFA	Esomeprazole 40 mg twice a day H2-receptor antagonist at bedtime and sucralfate suspension after every meal for 2 weeks after each therapeutic endoscopy		5 years	CR-neoplasia: sustained complete histological remission of HGD/early-stage cancer CR-IM: sustained complete histological remission of IM	CR-neoplasia: 93% (95%CI: 82.5-97.8) CR-IM: 93% (95% CI: 82.5-97.8)	
Phoa (2016)	Cohort study, Prospective	no	115 with CR-IM 121 with CR-neoplasia	65 + 14	CR at last endoscopy after RFA	High-dose PPI twice daily H2-receptor antagonist at bedtime and sucralfate suspension after every meal for 2 weeks after each therapeutic endoscopy		2 years	CR-neoplasia: sustained complete histological remission of HGD/early-stage cancer CR-IM: sustained complete histological remission of IM	CR-IM: 92% (95% CI: 86-96) CR-neoplasia: 96% (95%CI: 90-98)	
Ramay (2017)	Cohort study, Retrospective	no	30 with CE-IM 46 with CE-neoplasia 41 patients were eligible for 5-y analysis (26 with inital CE-IM)	61.9 + 8.6	CR at last endoscopy after cryotherapy	PPI twice daily		3 years 5 years (mean follow up 69.7 + 13.8 mo)	durability of response, calculated as the percentage of patients maintaining eradication without treatment CE-HGD: durability of complete eradication of HGD CE-D: complete eradication of dysplasia CE-IM: complete eradication of IM	For the 30 with CE-IM, CE-HGD 97%, CE-D 93%, CE-IM 87%, two progressed to ADK For the 26 with CE-IM, CE-HGD 96%, CE-D 92%, CE-IM 81%	
Dulai (2013)	Comparative (long and ultra-long BE), retrospective		34 patients (17 ULSBE and 17 LSBE)	66 and 68 years	CR-IM and/or dysplasia at last endoscopy after RFA	After ablation, all patients were placed on twice-daily proton pump inhibitors until the BE was fully ablated as evidenced by follow-up surveillance biopsies. Once eradicated, patients were continued on standard acid suppression		3 years		CR-IM 65% at 3 years and 82% at 2 years	

Author (year)	Study type	Comparative/randomized	Population	Age (mean ± SD)	Definition	Intervention (PPI)	Control	Follow-up	Outcome	Results	Notes
Komanduri (2017)	Cohort study, prospective	no, but comparative	205 BE (ND, HGD, LGD, IMC) with completed EET and CR-IM	65 + 11.6	CR at last endoscopy after RFA (3 sessions)	During EET, Omeprazol 40 mg twice a day was utilized as the standard and minimal equivalen dose of PPI. 24 pH-impedance and high-resolution manometry in instances of: (i) poor symptom control; (ii) EE identifi ed aft er at least one session of EET; or (iii) inability to achieve CE-IM aft er 3 RFA sessions. Patients optimized on twice-daily PPI (confi rmed adherence with dosing ~30 min before AM and PM meals and no insurance coverage mandated dosage reductions) and abnormal refl ux testing were referred for anti-refl ux surgery. Aft er modifying the refl ux management strategy, patients underwent a subsequent round of RFA. Patients not achieving CE-IM aft er 6 ablative sessions were deemed non-responders and off ered alternative therapies such as cryotherapy.	64 historical controls treated with EET without any formal anti-reflux strategy	40 months	recurrencce of IM	recurrence of IM 4.8 vs 10.9% (historical control), p=0.04	
Guarner-Argente (2013)	Cohort study, retrospective	no	156 with CR-neoplasia and 137 with CR-IM	68 + 11	EMR, PDT, APC and RFA	standard or high-dose proton pump inhibitor therapy		ND	sustained CR-neoplasia and sustained CR-IM	sustained CR-neoplasia 153/156 and sustained CR-IM 135/137	
Cotton (2017)	AIM dysplasia trial (RCT RFA vs placebo)		73 patients with dysplastic EB and CR-IM after RFA completing 5 years of surveillance	ND	CR-IM and/or dysplasia at last endoscopy after RFA	esomeprazole 40 mg twice daily.		5 years		sustained CE-IM 90% and sustained CE-D 99%	
van Munster (2022)	Prospective Registry	no	1,154 patients in the durability cohort	64 + 9	patients with successful EET with RFA defined as CE-BE with at least 1-year FU	high-dose PPI twice daily		10 years (median 32 mo, 16-59, after the last treatment	Proportion of patients with sustaine	sustained success in 1,116; recurrent dysplasia in 38	
Konda (2014)	Prospective database	no	74 patients with HGD/cancer and CR-IM after completing EET with EMR	ND	patients with successful EET with EMR	high-dose PPI twice daily		median 33 mo	sustained CR-cancer, sustained CR-HGD and sustained CR-IM	sustained CR-cancer 100%, CR-HGD 100%, CR-dysplasia 95.9%, CR-IM 71.6%	
Fleisher (2010)	Prospective AIM-II	no	50 out of 60 patients with NDBE and CR-IM at 2.5 years of fu. 10 patients declined participation	54 + 11	CR-IM at endoscopy 2.5 years after RFA	esomeprazol 40 mr twice a day the month after RFA, esomeprazol 40 mg per day the first 2.5 years and at discretion of the investigator between 2.5 and 5 years		5 years	sustained CR-IM	sustaied CR-IM 92%	
Skrobic (2016)	Prospective	no, patients' preference	47 patients with CE-IM after RFA	47.3 + 10.8	CE-IM at the last endoscopy	PPI (type and dose NA), n=25	Nissen, n=22	2 years	recurrence of IM	PPI, n=5 (20%) at 6 mo (1), 1 year (2) and 2 years (2) and Nissen, n=2 (9%) at 2 years; p=0.423	No differences between PPI and Nissen groups in age, gender, BMI, mean of RFA procedures, BE histology, BE length, HH size

Author (year)	Study type	Rand.	Patients	Age	CE definition	PPI regimen	Antireflux	Follow-up	Outcome	Results	Conclusions
Sharma (2006)	prospective	no	35 patients with NDBE or LGD treated with APC (n=19) or MPEC (n=16)		not defined	High dose of PPI		2 years	not clear	CE-IM in 24/35 after two years of FU	
Kobayashi (2022)	Retrospective cohort using a prospectively maintained database	no	345 patients with HGD/EAC treated with EMR or RFA and CE-N	64 + 10.7	CE-N at the last endoscopy	Once or twice-daily PPI during and after		30.8 mo (15 - 54)	recurrence of advanced neoplasia	20 patients (5.8%)	CE-IM was associated with a significantly lower hazard of recurrence (HR 0.2, 95%CI 0.1, 0.6), whereas the number of endoscopic treatments to achieve CE-N was associated with a significantly higher hazard of recurrence (HR 1.1, 95% CI 1.0, 1.2)
Akiyama (2012)	Retrospective	no	45 patients with pH monitoring before RFA and no need of adjustement of PPI doses or Nissen		NA	PPI at single, double or triple dosage		3 months after last treatment	NA	Reduction of BE surface area, 37 (82%) more than a 90% reduction, complete elimination (100%) in16 (36%). Normal to mild EAE and smaller hiatal hernia size were the independent factors associated with CE (OR = 6.78 [95 % CI 1.09–42.03], p = 0.04; OR 0.45 (95 % CI 0.23–0.86), p = 0.02)	Depite having PPI and/or funduplication or both, only 25/45who repeated pH monitoring (56%) after EER achieved normalization of 24-h intra-esophageal pH, while 7 (16%) had mild, 8 (18%) had moderate, and 5 (11%) had severe EAE
Basu (2002)	Retrospective from local database	no	50 patients with BE treated with APC, 34 with successful ablation and 16 with persistent BE		successful ablation if 90% or more BE was eliminated after several sessions of APC but also patients with no response were included in the analysis at 1 year	omeprazol 20 mg twice daily or lansoprazol 30 mg daily ad increase up to ompeprazol 60 mg daily if abnomal pH measurement before APC. Same doses after APC		1 year	recurrence of BE one year after APC (and lenght)	At 1 year, 17 patients had BE (including the initial 16) and 8 of them had reduced PPI medication. Only 4 patients reduced the dose in the responder group, p=0.01	Patients who reduced PPI dose to omeprazol 20 mg once daily or less (I guess that during FU but it is not clear) had a significantly greater incidence of BE recurrence (but the initial no responders are also included and it seems that only 1 has a real recurrence)
Ferraris (2007)	Retrospective from prospective database	no, but comparative	94/96 with BE treated with APC	57.1 (21-79)	CE-IM at the last endoscopy and confirmed by biopsies 3 months later	omeprazol 40 mg/day, n=50, same dosage during FU	Lap fundoplication before RFA, n=46, no PPI during FU	3 years	recurrence of IM	18% of recurrence	Lap funduplication was the only independent factor of sustained absence of BE during FU
Madisch (2005)	Retrospective	no, but comparative	66/70 NDBE successfuly treated with APC	55 + 12	CE-IM after APC	omeprazol 20 or 40 mg daily, n=33	Lap fundoplication during FU, n=22, and 8 of them also PPI	51 mo (9-85)	relapse of IM	13/66 (19.7%) of recurrence, 8 of them confirmed histologically	no predictive factors identified
O'Connell (2010)	Retrospective	no, but comparative	47 patients		patients treated with RFA and one endoscopy at least 12 mo after	PPI twice daily, n=28	Lap and open Nissen fundoplication before, during or after RFA, n= 19	1 year	persistent or recurrent BE	8 patients with persistent or recurrent BE (1 in Nissen group, 5.3% and 7 in PPI group, 25%; p=0.03)	Long-segment BE predisposes patients to recurrent BE after EERF ablation without funduplication (median 10 cm, 6-12, vs 3 cm, 2-5; p=0.03
Gupta (2013)	Retrospective from prospective database	no	229 with D-BE or ADC treated with RFA		CRIM: complete erradication of IM	PPI once or twice daily		median 22 mo	survival free of recurrence	recurrence by 1 year was 20% and 33% by two years	No factors associated with recurrence

Table 28						Centralization of FU after EET						
Author (year)	**Methods**			**Population**				**Intervention**			**Outcomes**	
	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest		**Remarks**
Wolf 2015	prospective	no	no	4982	62	patients with RFA for BE	RFA, EGD with biopsies in academic or community hospitals	none	2,7 years	EAC, death by EAC	100 (2%) EAC , 9 (0,2%) died from EAC EAC incidence was not impacted by treatment in an academic vs a community center (13.9 per 1,000 person-years vs 5.7 per 1,000 person-years, RR adjusted for baseline histologic grade 0.92 [95% CI 0.62-1.38] , and EAC mortality was not impacted by treatment in a community vs. academic setting		(n= 113 centers, vs 35 tertiary referral), supports the possibility of treatment and FU in community centers
Pasricha 2014	prospective	no	no	1634	62	patients with BE, CEIM by RFA ,and follow up data	RFA, EGD with biopsies in academic or community hospitals	none	2.4 years	BE recurrence	BE recurrence in 20 % of the patients same recurrence rate in community vs academic/referal centers **OR =1.04 (0.78 – 1.38)**		33% treated at a tertiary referral center, 67% at community centers (n= 113 centers, vs 35 tertiary referral), supports the possibility of treatment and FU in community centers
Cameron, 2014	prospective	no	no	69	69	BE with dysplasia on biopsies	EGD for BE in community hospital	EGD for BE in expert center	none	mucosal lesion or EAC detection rate	lesion detection 94% vs 42 %, p<0,001 56% increased EAC detection rate (10/28)		lack of proficiency for lesion detection in community centers
Roumans, 2020	metaanalysis	no	no	14002	NA	EGD for BE	EGD for BE in community hospital	EGD for BE in university hospital	none	adherence to FU and sampling guidelines	OR = 23 (7,16-73,6) for adherence to suveillance guidelines in university vs community hospital		lack of adherence to follow up guidelines for BE in community hospitals
Lyday 2009	retrospective	no	no	429	59	patients with RFA for BE	RFA, EGD with biopsies in community hospitals	RFA	20 months	safety, long term efficacy	safety cohort = 429, long term efficacy (biopsy available 1 year after CRIM), n= 137, CRD 100%, CRIM 77%		the **68 % dropout rate between the safety and the efficacy cohort is worrying**, vs 1-25% in referral centers (van munster et al, gut 2021, wolfson et al, GIE 2022)
Noorrdzij 2021	retrospective	no	no	170	67	BE with dysplasia on biopsies	Endoscopic workup of BE in an expert center	not applicable	none	rate of visible lesions missed at the referring center	36 missed visible lesions. 15% of the lesions associated with EAC missed at the referring center. Final pathology upstaged in 36% of the cases		lack of proficiency for lesion detection in community centers
Nieuwenhuis, 2022	retrospective	no	no	248	69	BE with confirmed LGD	EGD for BE in community center	EGD for BE in expert center	3 months	upstaging with HDG or EAC	57/248(23%) HGD or EAC		lack of proficiency for lesion detection in community centers

Author, year	Design	col3	col4	N	Age	Population	Intervention	Comparator	FU time	Outcome	Results	Comment
Abrams, 2009	retrospective	no	no	10958	62	BE with surveillance EGD in a community hospital	EGD for BE in community center	no	none	adherence to FU and sampling guidelines	adherence to guidelines 51,2%	lack of adherence to follow up guidelines for BE in community hospitals
Tsoi, 2021	retrospective	no	no	75	72	BE with confirmed LGD	EGD for BE in community center	EGD for BE in expert center	none	upstaging with HGD or EAC	20/75 (27%) HGD or EAC visible lesion detected in 52 %vs 12 % in referral vs community hospital, p = 0,029	lack of proficiency for lesion detection in community centers
Pohl, 2008	retrospective	no	no	1317	59	EGD for BE	EGD for BE in community hospital	EGD for BE in expert center	none	lesion detection rate, quality of surveillance EGD	lesion detection 1,3 vs 0,8 %, p=ns fewer biopsies in CH vs EC (2,5 vs 4,1, p<0,001)n fewer complete EGD documentation (25,1 vs 58 %, p<0,001)	lack of proficiency for lesion detection in community centers, poor quality of EGD in Community centers
Scholvinck, 2016	retrospective	no	no	197	66	EGD with HGD or EAC	EGD for BE in community hospital	EGD for BE in expert center	none	lesion detection rate	lesion detection 87% vs 60%, p<0,001 repeat endoscopy in CH in case of HGD/EAC <50 %	lack of proficiency for lesion detection in community centers
Rayner Hartley, 2016	retrospective	no	no	77	65	BE with dysplasia on biopsies	EGD for BE in community hospital	EGD for BE in expert center	none	lesion detection rate	18,4% lesion upstaging in the referral center	lack of proficiency for lesion detection in community centers
Curvers 2011	prospective	yes	no	99	63	BE with LGD	Endoscopic trimodal imaging	white light endoscopy	none	overall histological yield	no improvement in dysplasia detection HGD or EAC was diagnosed on random biopsies in 6/24 vs 7/24 patients	the results of AFI in expert centers (Curvers et al, Gastroenterology 2010; Cuvers et al, Gut 2008) were not reproduced in community centers
Pouw 2011	prospective	yes	no	84	70	BE with visible lesion	MBM	cap EMR	none	safety , efficacy, duration, costs	superiority of MBM over ER cap in terms of duration, costs.	50/84 cases performed in community centers. No direct comparison between experts and community centers.
Wani, 2019	retrospective	no	no	58709	61	EGD for BE	EGD for BE	none	none	adherence to FU and sampling guidelines	adherence to guidelines 83%	excluded because no direct comparison between community and expert centers
Klaver 2021	prospective	no	no	985	57	BE and without HGD or EAC, and follow-up in one of 6 community centers around Amsterdam.	Follow-up EGD every 3 years or 6-12 months fot LGD, on a dedicated program by one endoscopist, under the supervision of a specialized nurse	not applicable	7.9 years	progression to HGD or EAC	progression rate 0.78% per year 5 patients developped advanced (t>1) disease. 2 patients had surgery for T1 N0 disease. overall 8 progressors/67 (12%) were not amanable to endoscopic treatment, and treated by surgery or CRT	low progression rates in community centers - but expert nurses and pathologists involved.

Table 29			**biopsy protocol after EET**								
Author (year)	**Methods**		**Population**			**Intervention**			**Outcomes**		**Remarks**
	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	
Omar 2019	retrospective databank analysis	NA	50		Recurrence of IM after EET				Endoscopic location of recurrence relative to the SCJ	Recurrence was withing 2cm of SCJ or visible in 49/50 cases Non visibble recurrences in 17 patients (35%)	
Gupta M 2013	retrospective analysis	NA			Recurrence of IM and dysplasia after EET				Recurrence after EET	37/229 (16%) patients had recurrence 17 had recurrence at the SCJ (13 IM, 3 HGD, 1 LGD) 18 had recurrrence in the esophageal body (15 IM, 1 LGD, 1 HGD, 1 Ca) 2 had recurrence at the SCJ and esophageal body	
Cotton 2015	retrospective cohort study	NA	198		CRIM after RFA			3 years	Recurrence after EET	16.2% had recurrence All recurrences >1cm proximal to the GEJ were visible (HD-WL, NBI)	
Guthikond a 2016	retrospective cohort study	NA			CRIM after RFA for dysplastic Barrett				Recurrence after EET	272 patients with CRIM; 52 (24%) recurrence of IM 33 recurrences in the esophagus 17 at the cardia, 2 in esophagus and cardia 27 dysplastic recurrences (8/10 (80%) of recurrences in esophagus visible but only 3/17 (23.5%) recurrences at the cardia visible)	
Van Munster 2022	retrospective cohort study	NA	1154		Recurrence after EET				Dysplastic recurrence after EET	38 patients had dysplastic recurrence (LGD (n=14), HGD (n=7), or EAC (n=17)) All recurrences were visible	
Belghazi 2017	retrospective cohort study	NA	73		Recurrence after stepwise redical ER				Recurrence after EET	1 recurrence of T1b cancer, 4 dysplastic recurrences, 12 IM recurrences (all visible), 5 patients with buried IM, 27 patients with IM at cardia (not visible)	
van Munster 2022 (GUT)	retrospective cohort study									random biopsies of normal appearing cardia (5mm distal neo-Z-Linie) in 2733 FU endoscopies in 1121 patiens: Cardia random biopsies: IM found in 14% (95%CI 12% to 16%); Persisting IM n=78; recurrent IM n=72; IM was reproduced during a median of 3 FU endoscopies in 33%; LGD found in 0.81% (9/1121); reproduced on FU endoscopis in 75%; no progrssion to HGD or EAC	

											3 patients developed LGD (2 non visible in cardia, 1 visible in a newly developed island); no progression to HGD or EAC		
											Biopsies from neo-squamous epithelium: Buried IM found in 2.7% of patients; never reprodueced; progression to LGD, HGD EAC 0%		
											Cardia biosies: no difference in recurrence rate of dysplasia in the phase where no routine cardia biopasies were performed compared to the phaase before with routing cardia biopsies		
											Biopsies of NSE: Outcomes before 2013 (annual recurrence risk 1.3 (95% CI 0.5 to 2.1)) did not differ significantly from those after 2013 (annual recurrence risk 1.0 (95% CI 0.6 to 1.3)) (p 0.56)		

Table 30			Definition of a recurrence after EET								
Author (year)	**Methods**		**Population**			**Intervention**			**Outcomes**		
	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest (=definition of recurrence)	Remarks
van Munster et al. (2022)	Retrospective cohort	NA	1154	-	-	ER + RFA	-	43 months	-	3 grades of recurrence detected in the oesophagus or cardia after achieving CE-IM: (1) LGD in normal appearing cardia; (2) recurrent BE with dysplasia/EAC; (3) advances EAC that exceeded boundaries for curative endoscopic treatment	Clinical relevance of endoscopically normal cardia with LGD was considered negligible
Knabe et al. (2022)	prospective multicenter	NA	154	-	-	(ER +) Hybrid APC	-	24 months	The primary outcome: rate of complete BE eradication (initial CE-IM including CE-N) determined by 1 negative follow-up endoscopy with negative biopsies. Patients with negative endoscopy but positive biopsy went back to ablation (if within allowed number of sessions). Follow-up started only after reaching the primary outcome.

The secondary outcomes: rate of patients without recurrence at 2 years (sustained CE-IM including CE-N), vice versa recurrence rates of IM and neoplasia, number of ablation sessions, and AEs (AE and SAE, immediate and late) of ablation therapy including measures (e.g., dilatation of strictures). | Endoscopically suspected BE with confirmed by positive biopsy or biopsy alone positive for IM | Recurrence rate 34.1% (ITT analysis) or 29.2% (PP analysis) |
Wronska et al. (2021)	Prospective RCT	randomisation	71	-	BE with flat LGD	APC (90W or 60W) with omeprazole 120 mg or 40 mg			The primary outcome was the complete ablation rate at 6 weeks after APC treatment. Secondary outcomes were: 1) adverse event rate during APC treatment and within the 6-week post-treatment period; 2) complete ablation rate 2 years after APC treatment and at the end of follow-up; 3) recurrence rate, defined as the percentage of patients with complete ablation in whom endoscopic and histologic evidence of BE was found during follow-up.	BE recurrence was defined as endoscopic and histologic detection of Barrett's mucosa in a patient with prior complete ablation	BE recurrence 7%; easily treated with additional APC
Pouw et al. (2020)	Retrospective cohort (data end SURF in may 2013 until 2017)	RFA vs surveillance	136		BE with confirmed LGD	RFA	Surveillance		The primary outcome was rate of progression to HGD and/or cancer in patients randomized to RFA and to endoscopic surveillance. Progression was defined as a diagnosis of HGD or cancer in biopsy specimens or ER specimens, assessed by an expert pathologist. Secondary outcomes: (1) Recurrence of BE, dysplasia, and focal IM distal to a normal-appearing neosquamocolumnar junction among patients who achieved complete clearance of IM with LGD (CIM) by RFA treatment. Recurrence of BE was defined as endoscopically visible Barrett's mucosa, also when no biopsy specimens were obtained to confirm the presence of IM. IM found in biopsy specimens obtained from a normal-appearing neosquamocolumnar junction was not considered a recurrence of BE. CIM was defined as a single endoscopy without endoscopic evidence of Barrett's mucosa and without IM or dysplasia in biopsy specimens obtained just distal to the neosquamocolumnar junction after RFA treatment; (2) Regression of LGD in the surveillance group without any ablative treatment, defined as no more LGD in biopsy specimens obtained according to the Seattle protocol at any follow-up endoscopy after randomization in the SURF study; (3) the outcomes of patients with progression to HGD/cancer at any point during the SURF study	Recurrence of BE was defined as endoscopically visible Barrett's mucosa, also when no biopsy specimens were obtained to confirm the presence of IM. IM found in biopsy specimens obtained from a normal-appearing neosquamocolumnar junction was not considered a recurrence of BE.	
Wani et al. (2020)	Prospective multicenter study (2013-2018)	NA	807		Patients who had endoscopic evidence of BE and accompying biopsies with IM and who received EET	EET consisting of EMR and RFA	NA	2317 person years	The primary study outcome was the overall rate of intestinal metaplasia and dysplasia recurrence in BE patients achieving CE-IM. The secondary outcomes were (1) recurrence rates stratified by baseline histology, (2) recurrence rate changes over time, (3) postrecurrence histology and clinical outcomes, and (4) intestinal metaplasia and dysplasia recurrence predictors in BE patients who underwent EET and achieved CE-IM.	Recurrence of intestinal metaplasia or dysplasia was defined as histologic evidence of intestinal metaplasia or dysplasia on biopsies or EMR specimens taken from the esophagus or SCJ after CE-IM was achieved in the presence or absence of endoscopically visible BE	Recurrence of IM: 15%; Recurrence of dysplasia: 5.1% after median 2317 person-years.
Schwameis et al. (2020)	Retrospective chart review		40		Patients who had endoscopic therapy from 2001 to 2010 in a single center (1 phycsician)	EET consisting of ER and ablation (RFA / cryo)		82 months	The aim of this study was to evaluate the workload associated with endotherapy, the frequency and type of recurrence, long-term QOL, and late oncologic outcomes. CRIM was defined as 2 sequential endoscopies with biopsies showing no residual columnar-lined esophagus (CLE) or islands of columnar mucosa, and biopsies showing no IM, including at the gastroesophageal junction (GEJ). Short-segment Barrett's esophagus (SSBE) was defined as length of CLE <3 cm with IM, whereas patients with CLE ≥3 cm and IM were considered to have long-segment Barrett's esophagus (LSBE)	No clear defitiion stated; as derived from tekst: recurrence is presence of IM, dysplasia or cancer after CR-IM was achieved.	

Study	Design		N	Population	Intervention	Intervention 2	Follow-up	Endpoints	Recurrence definition	Results
Krajciova et al. (2019)	Retrospective analysis of prospective registry	NA	136	Patients treated with RFA for Barrett;s esophagus related neoplasia for LGD, HGD or EAC (after ER or ESD) in the Czech Republic between 2009-2016.	EET consisting of ER and ablation with RFA		27,5 months	The primary endpoints were assessment of CR-N, CR-IM and recurrences of both IM and neoplasia. We also assessed risk factors for recurrence of IEN/cancer and of IM and safety parameters	Recurrence of intestinal metaplasia. Recurrence of neoplasia.	4.5% recurrence of neoplasia; 15% recurrence of IM
Omar et al. (2019)	Retrospective study (2006-2015)		549 (50 in final analysis)	Patients in a large database who underwent EET and achieved complete eradication of IM from 2005-2015.	EET consisting of ER and ablation		12 months	The primary outcome of the study was the endoscopic location of recurrence of IM and/or dysplasia relative to the SCJ. Secondary outcomes were whether recurrences were visible or nonvisible in nature, the timing of recurrence after CE-IM, and histology of recurrence.	Recurrence of IM or neoplasia was defined by the detection of IM or neoplasia on surveillance biopsy specimens in tubular esophagus or at the GEJ after achieving CE-IM. Consistent with prior reports, the discovery of IM or neoplasia from the cardia was considered GEJ recurrence for the purposes of this study. Recurrences were further stratified into visibleand nonvisible recurrences depending on whether IM or neoplasia was discovered at the site of visible BE or only via random sampling, respectively.	13,8% recurrence (21% dysplasia, 11.5% HGD/EAC)
Sami et al. (2019)	Retrospective		594	Patients with histologically confirmed BE with or without dysplasia who underwent RFA (CR-IM achieved between 204 and 2017)	EET consisting of ER and RFA		2,8 years		Recurrence was defined as the histological presence of IM with or without dysplasia on biopsy specimens taken from either the tubular oesophagus or the GOJ or both after CRIM was achieved	Annual incidence rate of all recurrences: 9.6%l dysplasic 2.8% and HGD/cancer 1.6%
Vliebergh et al. (2019)	Prospective multiccenter registry		342	All pts undergoing RFA for curative eradication of BE (2008-2017)	EET consisting of ER and RFA		2,4 years	The primary outcome parameters included CR-IM and CR-D (endoscopically, and absence of dysplasia or IM under the neo-Z-line). Secondary outcomes were durability of CR-IM and CR-D and safety. Safety outcomes included immediate and late adverse events. Bleeding was considered clinically significant if it required hospitalization, blood transfusion, or an additional endoscopic intervention. A stenosis was defined as narrowing of the esophagus with symptomatic dysphagia requiring dilation.	No clear definition stated; recurrence of IM, HGD, LGD and EAC reported.	12% recurrence of neoplasia after long-term FU of 2.4 years (26% recurrence IM, of which 48% single finding; 6% recurrence HGD, 2.5% recurrence LGD, 0.5% recurrence EAC)
Frei et al. (2019)	Retrospective study		48	Patients with histologically confirmed diagnosis of dysplasia treated with endoscopic resection or ablation from January 2004 until December 2014	EET consisting of ER and RFA for HGD/T1a	EET consisting of ER and RFA for T1b	41 months	Primary endpoint(s): Complete eradication of dysplasia and neoplasia (CE-neo): Proportion of patients with complete histological eradication of dysplasia (sampling of neo-squamous epithelium and neo-z-line) after endoscopic treatment, defined as 2 histology reports at 3 months intervals negative for dysplasia in the treatment group HGD/T1a or T1b. Secondary endpoint(s): - Recurrence-free survival after CE-neo. - Total number of therapeutic interventions per patient. - Adverse events following endoscopic intervention with EMR and RFA, defined as mild (hospital admission, bleeding without transfusion, stenosis), moderate (< 2 units blood transfusion, repeated endoscopic intervention, hospital admission > 4 days), severe (intensive care unit [ICU], need for surgery) and fatal (death related to intervention).	No clear definition stated; recurrence of dysplasia/neoplasia. Also reported: no systemic recurrences.	11% recurrence of neoplasia
Tan et al. (2019)	Retrospective cohort study		430	Veteran patients with BE and with >1 RFA treatment	At least 1 RFA before Feb 2016.		2,7 yrs	The primary outcome was recurrence of BE with or without neoplasia after successful treatment (ie, CEIM) of BE with RFA. CEIM was defined by endoscopy and biopsy after RFA demonstrating both absence of columnar epithelium endoscopically and intestinal metaplasia on esophageal biopsies.	Histopathology reports with intestinal metaplasia, dysplasia, or neoplasia from targeted biopsies during follow-up endoscopy; endoscopic recurrence of columnar epithelium was not considered a recurrence, because of possibility of abnormal epithelium post-RFA. BE recurrence within 6 months of CEIM was also not considered because it likely represents lack of true BE eradication.	
Cotton et al. (2019)	Prediction model using retrospective data			US RFA registry (2004-2013) & UK NHR registry (2007-2015) : all pts who achieved CE-IM and entered endoscopic surveillance	Model for the incidence of neoplastic recurrence after CE-IM.			CEIM = defined as 1 posttreatment endoscopy showing no histologic or endoscopic evidenc	Neoplastic recurrence = the first finding on histologic examination of LGD, HGD or EAC in the esophagus or cardia after CE-IM confirmed in a single endoscopy	

Study	Design		N		Patients	Treatment		Follow-up	Outcomes	Recurrence definition	Incidence
Belghazi et al. (2018)	Prospective study		73		Pts treated with SRER for BE<5 cm with HGD or early cancer, and who had reached complete eradication of IM (CE-IM) and neoplasia (CE-neo)	SRER (with possible additional APC)		76 months	Primary outcomes: recurrence of neoplasia (HGD/EC), recurrence of dysplasia (including indefinite for dysplasia) and recurrence of endoscopically visible BE. Secondary outcomes: buried Barrett's glands, IM in biopsy specimens obtained from normal-appearing neo-SQJ, need for re-treatment, and sustained CE-IM / CE-neo at last FU endoscopy	Different types of recurrences: (1) recurrence of HGD and/or cancer, found in any biopsy or ER specimen obtained form the tubular esophagus or from an irregular eno-SCJ during FU; (2) recurrence of dysplasia define as any dysplasia (cancer, HGD, LGD or indifinite for dysplasia) found in any biopsy or ER speimen obtained for the tubular esophagus or from an irregula neo-SCJ during FU; (3) recurrence of endoscopically visible Barrett's mucosa, defined as any endoscopically visible columnar epithelium in the tubular esophagus (an irregular neo-SCJ was considered to be endsocopically visible Barrett;s mucosa) with or without histologic confirmation of IM. CAVE: IM foundin a normal appearing neo-SCJ was **not** considered a recurrence of BE.	Annual incidence for HGD/cancer: 0,22% per patient-year. Annual incidence for recurrence of dysplasia: 0.87 per patient-year; Annaul incidence rate for recurrence BE: 2.62% per patient-year.
Cotton et al. (2017)	Prospective study		110		Pts with endoscopic evidence of non-nodular dysplastic BE<8 cm in length who achieved CE-IM at 2 years after initial treatment	(ER +) RFA		401 person-years	CEIM = a single endoscopy visit wihotu endoscopic evidence of BE and with biopsies negative for intestinal metaplasia or dysplasia.	(1) Any recurrence = the first recurrence with IM, dysplasia or adenocarcinoma. Dysplastic recurrence = the first recurrence with dysplasia or adenocarcinoma.	(1) Incidence rate of any recurrence was 10.8 per 100 person-years; (2) Incidence rate of dysplastic recurrence was 5.2 per 100 person-years
Ramay et al. (2017)	Retrospective study of prospectively collected data		50			Cryotherapy with liquid nitrogren spray			Primary outcomes consisted of rates of CE-HGD, CE-D, and CE-IM at the 3-year and 5-year biopsy sessions, allowing for interval touch-up ablation. Secondary outcome measures included durability of response, incidence rates, and location of recurrent intestinal metaplasia and dysplasia, response to retreatment, and disease progression; CE-IM was defined as the absence of endoscopic evidence of BE and the absence of intestinal metaplasia on biopsy specimens of both the neosquamous esophageal epithelium and the biopsied area below the neosquamocolunnar junction at 1 endoscopic session. CE-D was defined as the absence of low-grade dysplasia (LGD) and HGD endoscopically and on biopsy of both areas.	Recurrence was defined as histologic evidence of intestinal metaplasia, dysplasia, or neoplasia on endoscopic biopsy during the surveillance period	
Godat (2017)	retrospective cohort (single center)								The main goal of this study was to evaluate the efficiency and safety of endoscopic management in case of relapse of neoplastic BE. The secondary objective was to determine predictive factors that could involve the failure of endoscopic treatment for relapse, reduce the global survival or the disease-free survival. The disease-free survival was defined as the time from achieved remission until second relapse.	Recurrence = histological presence of HGD or superficial EAC at least 6 months after the end of successful initial endotherapy	
Guthikonda et al. (2017)	retrospective cohort study		218			RFA			CE-IM was defined as complete histological and endoscopic remission of IM after a single endoscopy with biopsies taken throughout the prior extent of BE. CE-D was defined as complete histological and endoscopic remission of dysplasia after a single endoscopy with biopsies taken throughout the prior extent of BE. Recurrence was defined as dysplastic or non-dysplastic IM in the esophagus or any dysplasia in the cardia on histology subsequent to CE-IM	Recurrence was defined as any presence of IM or dysplasia in the tubular esophagus or dysplasia in the gastric cardia subsequent to CE-IM; Because non-dysplastic IM of the cardia is a common finding in subjects with chronic gastroesophageal reflux disease (14), a finding of non-dysplastic IM of the cardia alone was not considered disease recurrence	
Le Page et al. (2016)	Prospectove study		50			ER + RFA	Surgery			No clear definition stated, but from text: Recurrence of HGD/Invasive cancer	
Phoa et al. (2016)	Prospective study		132			ER + RFA		27	Complete eradication of neoplasia (CE-neo), defined as absence of HGD and EC in all biopsies obtained at the first endoscopy with complete endoscopic clearance of BO or from residual BO after the maximum number of endoscopic treatment sessions had been performed. Complete eradica tion of IM (CE-IM), defined as absence of IM, in all oesophageal biopsies obtained at the first endoscopy with complete endoscopic clearance of BO.	No clear definitions stated; recurrence of HGD/mucosal cancer and recurrence of IM described in results.	
David et al. (2015)	Retrospective observational cohort study		342			RFA or EMR+RFA or Ps-PDT		14,2 months		BE recurrence defined as the subsequent detection of specialized columnar mucosa in biopsy samples from the target mucosa in the distal esophagus	
Canto et al. (2015)	Retrospective cohort		64			Cryoablation (CO2)		4,2 yrs		Recurrence of Barrett's esophagus was defined as any esophageal biopsy with intestinal metaplasia after two sets of negative post-cryotherapy biopsies (at 3 and 6 months).	
Small et al. (2015)	Retrospective cohort		256			EMR + ablation		5 yrs	Primary outcomes in each group (HGD or IMC) were the proportion with CE-IM after multimodal endoscopic therapy, defined as a single endoscopy with biopsy specimens demonstrating no IM, recurrence of dysplasia or neoplasia after CE-IM was achieved, and recurrence of IM defined as detection of IM on biopsy specimens from a single surveillance endoscopy after eradication any time during the follow-up period. A secondary outcome was to determine recurrence of dysplasia/neoplasia in patients who had eradication of HGD/IMC but persistent IM despite endoscopic therapy.	No clear definition stated; recurrence for dysplasia and neoplasia, and recurrence of IM described in results.	

Konda et al. (2014)	Prospective database		86			EMR		33 months	The primary outcome was treatment efficacy as determined by complete eradication of BE, and associated neoplasia as determined by the combination of endoscopy, histology, and, in the cases of cancer, EUS, without evidence of malignant lymphadenopathy. Secondary outcomes included accuracy of diagnosis, safety, and durability. Complete remission was defined as showing no pathologic evidence of cancer, HGD, or residual intestinal metaplasia on the most recent endoscopy and histology in surveillance.	Recurrence was defined as disease that was encountered during surveillance.
Anders et al. (2014)	Retrospective study		90		Pts who had undergone successful resection of neoplastic BO (defiend as at least 2 enfative endoscopies wit biopsies showing no neoplasia or BO) and at least 3 yrs of FU.	EMR		64.8	The primary outcome parameters were the rates of recurrence of neoplasia (low-grade and high-grade intraepithelial neoplasia (LGIN/HGIN) and cancer) and of Barrett's epithelium without neoplasia following successful complete removal of BO and neoplasia (defined as two negative follow-up endoscopies with biopsies that were negative for BO and neoplasia).	Recurrence of neoplasia: LGD, HGD and cancer
Gosian et al. (2013)	Retrospectice study (single center)		32			Cryo with liquid nitrogen			The primary outcomes were complete eradication of HGD (CE-HGD) and complete eradication of IM (CE-IM) at the 2-year biopsy session, allowing for interval touch-up ablation. The secondary outcome measures included the durability of response and response to retreatment. Durability of response was calculated as the number of patients maintaining CE-HGD and CE-IM without retreatment. Response to retreatment was described as the ability to ablate HGD or IM with additional cryotherapy. In both analyses, time 0 was considered the last cryotherapy session before the start of the surveillance period. Disease progression was defined as any patient found to have esophageal adenocarcinoma at any follow-up.	Recur rence was defined as histologic evidence of IM or neoplasia seen during the surveillance period
Phoa et al. (2013)	Prospective registry		54			ER + RFA			Primary outcome parameters were sustained complete histological remission of HGIN/early-stage cancer (CR-neoplasia) and sustained complete histological remission of IM (CR-IM). Secondary outcome parameters were presence of IM in biopsies obtained <5 mm distal to neo-SCJ (gastric cardia); presence of buried Barrett's glands in neosquamous epithelium biopsies/endoscopic resection specimens; and presence of abnormalities on EUS.	No definition of recurrence described; study only reports the recurrence of neoplasia or IM.
Guarner-Argente et al. (2013)	Retrospective cohort		166			ER with PdT / RFA or APC			Primary endpoints included (1) complete eradication of all HGD and/or intramucosal carcinoma: negative biopsies for HGD and intramucosal carcinoma (complete eradication of dysplasia and neoplasia) and (2) complete eradication of all BE, with absence of intestinal metaplasia in all biopsies, including those obtained from neosquamous mucosa and immediately distal to the neosquamocolumnar junction (complete eradication of intestinal metaplasia).	Recurrence of intestinal metaplasia, dysplasia, or cancer during the follow-up (biopsies with intestinal metaplasia, low-grade dysplasia [LGD], HGD, or cancer after the primary endpoints had been achieved at least in one follow-up endoscopy)
Dulai et al. (2013)	Retrospective evaluation of prospectively collected data		72		Long BE vs ulta-long BE	ER + RFA		34 / 45 months		No definition of recurrence described; study only reports the recurrence of neoplasia or IM.
Halsey et al (2011)	Single-center retrospective study of prospectively collecte data		36			Cryoabaltion using liquid nitrogen spray		24 months	The primary outcome measure was the location and histology of recurrent disease. The secondary outcome measures included the number of treatment sessions for complete response, time interval for complete response, time interval between complete response and recurrence, final treatment success, and duration of follow-up after complete response.	Recurrence was defined as the presence of intestinal metaplasia with or without neoplasia on endoscopic biopsy during the surveillance period.
Alvarez etl al. (2011)	prospective trial				BE > 10 cm containing neoplasia	ER + RFA			Complete remission, defined as endoscopic resolution of BE and no intestinal metaplasia (CR-IM) or neoplasia (CR-neoplasia) in biopsy specimens.	Eecurrence of neoplasia during follow-up; recurrence of BE during follow-up (either endoscopic or histological)
Pouw et al (2010)	Multicenter prospective cohort study		24		BE < 12 cm containing HGD / EAC	EMR + RFA		22 months	The primary end points were histology-based (biopsy specimens obtained 2 months after the last therapeutic intervention). A complete response was defined as all biopsies negative for IM (complete response [CR]-IM) and neoplasia (CR-neoplasia), reported separately. Secondary end points were as follows: disease progression, adverse events, and durability of CR-IM and CR-neoplasia at the last biopsy available.	No definition of recurrence described; study only reports no recurrences of neoplasia or IM.
Gondrie et al. (2008)	Prospective trial		12		Pts with HGD or EAC diagnosed at 2 separate endoscopies	EMR + RFA		14		No definition of recurrence described; study only reports no recurrences of neoplasia or IM.

	Table 31		**What is the meaning of IM in random biopsies of a normal-looking z-line after EET?**								
Author (year)	**Methods**			**Population**			**Intervention**			**Outcomes**		**Remarks**
	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest	Remarks
Hirota 1999	Cross-sectional	N/A	889	62	N/A	N/A	N/A		SIM on biopsy	5.6% GOJ-SIM in an all-comer endoscopy population	
Gupta 2013	Retrospective cohort study	N/A	229	64	N/A	N/A	N/A		SIM on GOJ bx	Of the 17 recurrences at the GEJ, 72% (13 patients) of recurrences were IM and 28% (4 patients) of recurrences were dysplastic BE (3 patients had HGD, and 1 patient had LGD). On entry histology, all 4 patients had HGD.	No information re endoscopic aspect of dysplastic recurrences
Orman 2013	Retrospective cohort study	N/A	112	not known	N/A	N/A	N/A		SIM on GOJ bx	Eight patients (7% of those with CE-IM) had recurrent disease after a median of 235 days (range 55-1124 days). Progression to IMC (n=1) or EAC (n=2) occurred in 3 of these 8 patients, all of whom had pre-ablation high-grade dysplasia (HGD). Five patients had recurrence of non-dysplastic BE (n=3), low-grade dysplasia (n=1), and HGD (n=1). During 155 patient-years of observation, recurrence occurred in 5.2%/year, and progression occurred in 1.9%/year.	Limited information re endoscopic aspect of dysplastic recurrences
Cotton 2015	Retrospective cohort study	N/A	198	69	N/A	N/A	N/A		SIM on GOJ bx	In a mean 3.0 years of follow-up, 32 (16.2%; 95%CI 11.0%-22.0%) patients had recurrence of disease, 13 had dysplastic recurrence, and 5 (2.5%; 95% CI, 0.3%-4.7%) of whom progressed beyond their worst histology before treatment.	Three of eight patients (2 with HGD and 1 with IMC) with high grade recurrences had no endoscopic abnormality
Cotton 2017	Retrospective cohort study	N/A	110	66	N/A	N/A	N/A		SIM on GOJ bx	After mean 3.6 years F/U, 35 of 110 (32%) patients had recurrence of BE or dysplasia, and 19 (17%) had dysplasia recurrence. The incidence rate of IM recurrence was 10.8 per 100 person-years overall (95% CI, 7.8–15.0); 8.3 per 100 person-years among patients with baseline low-grade dysplasia (95% CI, 4.9–14.0), and 13.5 per 100 person-years among patients with baseline high-grade dysplasia (95% CI 8.8–20.7). The incidence rate of dysplasia recurrence was 5.2 per 100 person-years overall (95% CI 3.3–8.2); 3.3 per 100 person-years among patients with baseline low-grade dysplasia (95% CI 1.5–7.2), and 7.3 per 100 person-years among patients with baseline high-grade dysplasia (95% CI 4.2–12.5).	Limited information re endoscopic aspect of dysplastic recurrences

Study	Type		N							Outcome	Comment
Desai 2017	Systematic review	N/A	774	not known	N/A	N/A	N/A		SIM on GOJ bx	Recurrence of EAC, dysplasia, and IM was 1.4%, 2.6%, and 16.1%, respectively, in the focal EMR and RFA group	Limited information re endoscopic aspect of dysplastic recurrences
Fuji 2017	Systematic review	N/A		not known	N/A	N/A	N/A		SIM on GOJ bx	pooled incidence of IM recurrence rate of 4.8 (95%CI 3.8 – 5.9)/100 PY, and dysplasia recurrence rate of 2.0 (95%CI 1.5 – 2.5)/100 PY.	Limited information re endoscopic aspect of dysplastic recurrences
Guthikonda 2017	Retrospective cohort study	N/A	218	70	N/A	N/A	N/A		SIM on GOJ bx	28 patients (13%) experienced any dysplastic recurrence	Most dysplastic recurrences were in the cardia, and the majority were not visible but detected on random biopsies.
Sawas 2018	Systematic review	N/A	1973		N/A	N/A	N/A		SIM on GOJ bx	Dysplasia detection IR after the first year was 1% (95% CI: 1–2%) and significantly higher in the first year compared to the years after (RR: 1.92 (95% CI: 1.32–2.8).	Limited information re endoscopic aspect of dysplastic recurrences
Omar 2019	Retrospective cohort study		443		N/A	N/A	N/A		SIM on GOJ bx	48 NDBE recuurrences and 20 neoplastic recurrences. Recurrence generally follows baseline histology. About 505 of recurrrences in first year not visible, only 10% thereafter.	Heterogenous definition incluudes NDBE
Sawas 2019	Systematic review		4410		N/A	N/A	N/A		SIM on GOJ bx	The pooled cumulative incidence of any dysplasia recurrence after achieving CRIM was 5% (95% CI, 3%-7%) and 12% (95% CI, 4%-23%) after achieving CR-D only.	Limited information re endoscopic aspect of dysplastic recurrences
Wani 202	Retrospective cohort study		807		N/A	N/A	N/A		SIM on GOJ bx	Intestinal metaplasia recurred in 121 patients (15%; IR, 5.2/100 person-years), and dysplasia recurred in 41 patients (5.1%; IR, 1.8/100 person-years). Recurrence peaked at 1.5 yrs with Bell curve behaviour	Limited information re endoscopic aspect of dysplastic recurrences
Solfisburg 2021	Retrospective cohort study		633		N/A	N/A	N/A		SIM on GOJ bx	Median follow-up was 47 months. Dysplasia recurrence was 2.2% per year. Recurrent GEJIM after endoscopic eradication of BE was not associated with an increased risk of subsequent dysplasia (log rank .07)	Limited information re endoscopic aspect of dysplastic recurrences
Van Munster	Retrospective cohort study		1154		N/A	N/A	N/A		SIM on GOJ bx	Random biopsies from the cardia were noted to contain LGD in 9 patients (9/11121; 0.8%). All other 24 recurrences were detected as visible lesions at the GOJ or in the tubular oesophagus. No statistically significant association was found between a finding of non-dysplastic IM in the cardia and the risk for recurrence (adjusted HR 0.5 (95% CI 0.2 to 1.7)). None of the patients with IM progressed to HGD or EAC.	
PICO search string:			"Barrett's esophagus" AND recurrence AND "intestinal metaplasia"								

Table 32				Most optimal surveillance intervals after EET						

Author (year)	Methods		Population			Intervention			Outcomes		
	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest (=most optimal surveillance interval)	Remarks
van Munster et al. (2022)	Retrospective cohort (durability cohort, 7 Barrett Expert Centers, 200-2018)	NA	1154	-	All patients with successful EET defined as complete endoscopic eradication of BE (CE-BE) with at least 1-year FU at the moment of data collection.	Treatment: ER + RFA ; FU-protocol: endoscopies every 3 months in the first year, followed by annual endoscopies in the next 2-5 yrs, and eventually followed by endoscopies every 2-3 year.	FU-protocol (after 2015): endoscopies annually in the first 5 year after treatment, and eventually followed by endoscopies every 2-3 year.	43 months (after baseline); 32 months after last treatment	Primary endpoint: Proportion of patients with sustained eradication of LGD, HGD and EAC during long-term endoscopic FU. A patient was considered a failure if recurrent LGD, HGD or EAC was detected in the oesophagus or cardia, or if lymph node or distant metastasis from EAC were found during FU. Failure was categorised into three groups according to the severity of recurrent disease: (a) LGD in a normal appearing cardia without recurrent BE; (b) recurrent BE with LGD/HGD/EAC amendable for curative endoscopic treatment; and (c) advanced EAC that exceeded boundaries for curative endoscopic treatment. Secondary durability endpoints: - Sustained eradication of HGD and EAC (recurrent LGD was considered as success). - Progression to advanced EAC that exceeded boundaries for curative endoscopic treatment. - Recurrence of non-dysplastic BE. - Diagnostic yield for FU endoscopies and random biopsies. - Association between frequent endoscopies in the first FU year and recurrence. - Association between IM in the cardia and recurrence. - Unrelated mortality rates and causes of death.	During FU: - Recurrence of LGD, HGD and EAC occured in 38/1154 pts (3%), which corresponds to an annual risk of 1% (95%CI 0.8-1.4). - 24/1154 had recurrence HGD/EAC (2%; annual risk 0.7 (95%CI 0.4-1.0)). - All recurrences were detected as endoscopic abnormalities (recurrent BE or visible lesion). Occurence median 31 months (19-43) after CE-BE; categorisation of recurrences: (1) LGD in normal appearing cardia (9/38, 24%, who had re-treatment or surveillance without progression) (2) recurrent BE with dysplasia/EAC (24/38, 63%, successful endoscopic re-treatment to achieve CE-BE again) and (3) advanced EAC that exceeded boundaries for curative endoscopic treatment (5/38, 13%, 4 had surgery of which 3 died of metastasised disease). - Diagnostic yield per FU-endoscopy for recurrent LGD/HGD/EAC was 1% (95% CI 0.7 - 1.3); for recurrence HGD/EAC 0.5 (95% CI 0.4-0.9). -Recurrent NDBE 9% (8-11), the majority of which islands; 1 progressed to LGD, 0 progressed to HGD/EAC during median 24 months of FU (annual risk for recurrent BE islangs 3% in year 1-2 and 1% in the years thereafter; the risk for recurrent BE tongues was 0.4% in the first 2 years and 1% in the years thereafter). - Regarding the different FU-intervals before and after 2015: no significant assosication was found between the frequency of FU in year 1 and dysplastic recurrence during the first 30 months (afjusted HR 1.6 (95% CI 0.6-4.1). Also during long-term FU no significant association was found between frequency FU in year 1 and progression to advanced neoplasia (adjusted HR 0.8 (95% CI 0.1-5.8).	
Cotton (2018)	Retrospective study to build prediction model (US 2004-2013 and UK 2007-2015)	NA	3478 pts in the surveillance cohort for the model		Patients that achieved CEIM and entered surveillance were included.			-	CE-IM: defined as one post-treatment endoscopy showing no histological or endoscopic evidence of intestinal metaplasia or associated neoplasia; endoscopic surveillance: defined as having at least one additional surveillance endoscopy with histologic sampling following CEIM. **Primary outcome:** neoplastic recurrence defined as the first finding on histologic examination of LGD, HGD or EAC in the esophagus or cardia after CE-IM confirmed in a single endoscopy. **Variables for the model:** age at first RFA treatment, sex, initial BE segment length in centimeters, presence of prior endoscopically resected nodular disease, performance of any endoscopic mucosal resection of nodular disease during endoscopic eradication treatment, and the patients' most severe histologic grade prior to CEIM	**Model:** annual rate of recurrence of neoplasia (LGD, HGD and EAC) was 0.19% (0,09-0,40) for patients with pre-CEIM NDBE/indeterminate for dysplasia, 1.98% (1,34-2,93) for patients with pre-CEIM LGD and 5.93% (4,77-7,36) in patients with pre-CEIM HGDIMC. In the higher risk groups, neoplastic recurrence occured at a higher rate in the first year but at a constant rate thereafter. Large majority of recurrences was amendable for endoscopic re-treatment. Proposed intervals: patients with pre-CEIM LGD 1 and 3 years after CE-IM; for patients with pre-CEIM HGD/IMC 0.25, 0.5, 1, 2, 3, 4, 5 years follow-up after CE-IM.	Recommendations beyond 5 years requires extrapolation beyond present data.
Wolfson (2022)	Prospective registry	NA	2535 pts, from which 1175 details analysis with relapses		Patients with LGD, HGD or IMC which was confirmed by 2 expert pathologists before EET started in Jan 2008 until Dec 2018.	RFA treatement	NA	7856 patient-years	**The primary aim** was to determine the 10-year cancer progression in patients undergoing EET for BE. **Secondary aims** included understanding the durability of complete remission of dysplasia (CR-D) and complete remission of intestinal metaplasia (CR-IM) together with rates of relapse from these states.	**Definitions:** CR-IM was defined as all biopsies clear of intestinal metaplasia at two consecutive endoscopies together with residual tongues of glandular mucosa measuring less than 3cm and the absence of dysplasia. Relapse from CR-D was defined as the biopsy proven recurrence of dysplasia at a single endoscopy. Relapse from CR-IM the same followed successful clearance of IM.	Rate of relapse from CR-IM was 4.2% in year 1, 10.1% in at 2 years and 18.7% at 8 years. In year 1: 3.2% LGD, 5.2% HGD, 2.8% IMC. Year 2: 9.8% LGd, 11.9% HGD, 11.4% IMC. Year 8: 11.3% LGD, 22.1% HGD, 14.7% IMC. 78.4% of all relapses relapses within the first 2 years after CRIM. Most relapses occur within 2 years. The authors state that the data suggests that continuing annual surveillance endoscopies beyond 2 years may offer little additional benefit whilst providing no evidence that the histological sybtype prior to EET should be used to guide surveillance guidelines.	
Recurrence characteristics											

Study	Type		N		Population	Treatment		Follow-up	Definitions / Aim	Results	Comments
Wani (2020)	Prospective study at 4 tertiary-care referral centers (2013-2018)	NA	807		Patients referred for EET with endoscopic evidence of BE, accompying biopsies with IM and who underwent EET.	Treatment: ER + RFA ; FU-protocol: for HGD/IMC every 3-6 months for years 1 and 2, then annually; for LGD every 6 months for 1 year, then annually.		2317 person-years (mean FU of 3.3 years) afer CE-IM	(1) CE-IM was defined as the absence of endoscopically visible BE and intestinal metaplasia on esophageal biopsies that included the entire pretreatment BE length and SCJ after a single endoscopy; (2) Complete eradication of dysplasia was defined as the absence of dysplasia on esophageal biopsies in the presence or absence of endoscopically visible BE after a single endoscopy; (3) Recurrence of intestinal metaplasia or dysplasia was defined as histologic evidence of intestinal metaplasia or dysplasia on biopsies or EMR specimens taken from the esophagus or SCJ after CE-IM was achieved in the presence or absence of endoscopically visible BE. **Primary outcome:** the overall rate of intestinal metaplasia and dysplasia recurrence in BE patients achieving CE-IM. The secondary outcomes were (1) recurrence rates stratified by baseline histology, (2) recurrence rate changes over time, (3) postrecurrence histology and clinical outcomes, and (4) intestinal metaplasia and dysplasia recurrence predictors in BE patients who underwent EET and achieved CE-IM	**Recurrence of IM:** 121 pts (15%) during FU of 2317 person years (baseline diagnosis LGD in 27/239 (11.3%) and baseline diagnosis HGD/EAC 55 of 507 (10.8$). IR of IM was 5.2 (95% CI 4.4-6.2) per 100 person years (for baseline LGD IR 2.8, for baseline HGD/EAC 6.7). **Recurrence of dysplasia:** 41 pts (5.1%); baseline LGD 6/239 (2.5%) and baseline HGD/EAC 35/507 (6.9%). IR for dysplasia 1.8 (95% CI 1.3-2.4) per 100 person years (for baseline LGD IR 0.6 (0,3-1,4) & for baseline HGD/EAC IR 2.9 (2,1-3,9). **Time to recurrence:** peak at 1.6 years after CE-IM; rate of recurrences was not constant and time to any recurrence was normally distributed	The authors raise the question if there is need for aggressive surveillanc in year 1 after achieving CE-IM.
Omar (2019)	Retrospective analysis multicenter database		443 (50 in final analysis)		Patients who underwent EET and achieved complete eradication of IM (2005-2015)	Treatment: ER +/- RFA. FU not described.		Median time until recurrence 12 months (range 1054)	CE-IM was defined as the absence of endoscopically visible BE plus the absence of IM on all surveillance biopsy samples in the tubular esophagus or at the GEJ on a single endoscopy after EET; Recurrence of IM or neoplasia was defined by the detection of IM or neoplasia on surveillance biopsy specimens in tubular esophagus or at the GEJ after achieving CE-IM. **Primary outcome:** The primary outcome of the study was the endoscopic location of recurrence of IM and/or dysplasia relative to the SCJ; Secondary outcomes were whether recurrences were visible or nonvisible in nature, the timing of recurrence after CE-IM, and histology of recurrence.	50 patients with a recurrence (overall recurrence rate of 13.8%), 33/55 (66%) visible abnormality vs 7/50 random biopsies. Most recurrences <2cm of the SCJ. Histology of recurrences: 74% NDBE (n=37), LGD 10% (n=5), HGD 10% (n=5), EAC 4% (n=2), indeterminate for dysplasia 2% (n=1). All novsiible recurrences were detecte within 2 years after CE-IM.	
Cotton (2017)	Retrospective analysis on long-term durability from prospective trial		110		All patient that achieved CE-IM at 2 years after the start of the AIM trial or after salvage RFA were included for this analysis	RFA; FU-protocol: LGD after 6 and 12 months, HGD after 3-6-9-12 months		401 person-years (mean 3.6 year per patient) after achieving CE-IM	The aim of this study was to assess the rate of recurrence of BE in prospectively followed patients who had presented with dysplasia and then achieved CEIM.	**Any recurrence:** 35/110 (32%) any recurrence over 324.5 person-years. IR of any recurrence qas 10.8 (7,8-15,0) per 100 person-years overall (IR LGD 8.3 (4,9-14,0) and HGD 13.5 (8,8-20,47). **Dysplatic recurrence:** 19 (17%) had dysplastic recurrence over 363.1 person-years; IR 5.2 (3,3-8,2) per 100 person-years overall (IR LGD 3.3 (1,5-7,2) and HGD 7.3 (4,2-12,5). **Recuerrnce over time:** greater probability of any recurrence or dyspalstic recurrence in the 1st year follwong CE-IM than in the follwong 4 years.	After 4 years of surveillance no recurrence of BE or dysplasia was identified in this study.
Guthikonda (2017)	Retrospective cohort		306		Patients treated with RFA for dysplastic BE			540.6 person-years	**Aim:** to describe the clinical outcomes associated with recurrence following CE-IM in BE treated with RFA; also analyze histological and endoscopic features of recurrence. **CE-IM** was defined as complete histological and endoscopic remission of IM aft er a single endoscopy with biopsies taken throughout the prior extent of BE; **CE-D** was defined as complete histological and endoscopic remission of dysplasia aft er a single endoscopy with biopsies taken throughout the prior extent of BE. **Recurrence** was defined as dysplastic or non-dysplastic IM in the esophagus or any dysplasia in the cardia on histology subsequent to CE-IM. Because non-dysplastic IM of the cardia is a common finding in subjects with chronicgastroesophageal reflux disease, a fi nding of non-dysplastic IM of the cardia alone was not considered disease recurrence.	**Recurrence:** 52/218 (24%) experienced recurrence of IM or Barrett's associated neoplasia over 540.6 person-years (IR 9.6%/year) in esophagus and cardia. Int he esophagus 36 had a recurrence, 12 NDBE (33%). 15 LGD (42%), HGD (14%) IMC 4 (11%; 30/52 (58%) achieved again CE-IM after re-treatment. 4 (1.8% of total, 7.7% of recurrences) progressed to invasive carcinoma (IR 0.65%/year). Mean time to recurrence was 1,88 years (SD 1.42) after CE-IM and mean n of endsocopie 2.3. REcurrences	
Study outcomes prospective trials											
Knabe (2022)	Prospective multicenter study					Hybrid APC				Not used for outcome of interest: only study results	
Phoa (2016)	Prospective multicenter study					ER + RFA				Not used for outcome of interest: only study results	
Konda (2014)	Retrospective cohort					ER				Not used for outcome of interest: only study results	
Gosian (2013)	Single-center retrospective study					Cryoablation with liquid nitrogen				Not used for outcome of interest: only study results	
Guarner-Argente (2013)	Retrospective cohort					Endoluminal therapy (ER + PDT / RFA / APC)				Not used for outcome of interest: only study results	
van Vilsteren (2011)	Prospective multicenter trial					SRER	ER + RFA			Not used for outcome of interest: only study results	
Pouw (2008)	Prospective trial					RFA				Not used for outcome of interest: only study results	
Anders (2014)	Retrospective anslysis					widespread ER				Not used for outcome of interest: only study results	

Table 33			**When to stop surveillance after EET (taken into account risk of recurrence and competing mortality)**							

Author (year)	Methods		Population				Intervention			Outcomes		Remarks
	Design	Randimsation / blinding	N	Age	Inclusion criteria	Protocol (details) of Intervention	Protocol (details) of Comparison	Follow-up	Outcome measures (+ definitions)	Outcomes of interest (=definition of recurrence)	Remarks
van Munster et al. (2022)	Retrospective cohort (durability cohort, 7 Barrett Expert Centers, 200-2018)	NA	1154	-	All patients with successful EET defined as complete endoscopic eradication of BE (CE-BE) with at least 1-year FU at the moment of data collection.	Treatment: ER + RFA ; FU-protocol: endoscopies every 3 months in the first year, followed by annual endoscopies in the next 2-5 yrs, and eventually followed by endoscopies every 2-3 year.	FU-protocol (after 2015): endoscopies annually in the first 5 year after treatment, and eventually followed by endoscopies every 2-3 year.	43 months (after baseline); 32 months after last treatment	Primary endpoint: Proportion of patients with sustained eradication of LGD, HGD and EAC during long-term endoscopic FU. A patient was considered a failure if recurrent LGD, HGD or EAC was detected in the oesophagus or cardia, or if lymph node or distant metastasis from EAC were found during FU. Failure was categorised into three groups according to the severity of recurrent disease: (a) LGD in a normal appearing cardia without recurrent BE; (b) recurrent BE with LGD/HGD/EAC amendable for curative endoscopic treatment; and (c) advanced EAC that exceeded boundaries for curative endoscopic treatment. Secondary durability endpoints: - Sustained eradication of HGD and EAC (recurrent LGD was considered as success). - Progression to advanced EAC that exceeded boundaries for curative endoscopic treatment. - Recurrence of non-dysplastic BE. - Diagnostic yield for FU endoscopies and random biopsies. - Association between frequent endoscopies in the first FU year and recurrence. - Association between IM in the cardia and recurrence. - Unrelated mortality rates and causes of death.	During FU: - Recurrence of LGD, HGD and EAC occured in 38/1154 pts (3%), which corresponds to an annual risk of 1% (95%CI 0.8-1.4). - 24/1154 had recurrence HGD/EAC (2%; annual risk 0.7 (95%CI 0.4-1.0)). - All recurrences were detected as endoscopic abnormalities (recurrent BE or visible lesion). Occurence median 31 months (19-43) after CE-BE. ;- Regarding the different FU-intervals before and after 2015: no significant assosication was found between the frequency of FU in year 1 and dysplastic recurrence during the first 30 months (afjusted HR 1.6 (95% CI 0.6-4.1). Diagnostic yield per endoscopy for the recurrence of HGD/EAC was 0.6% per endosocpy. All-cause mortality during FU (median 60 months after baseline and 49 months after the last treatment): 96 patients died , of which 92 (8%) due to unrelated causes and 4 due to metastasised EAC (0.3%).	
Wolfson (2022)	Prospective registry	NA	2535 pts, from which 1175 details analysis with relapses		Patients with LGD, HGD or IMC which was confirmed by 2 expert pathologists before EET started in Jan 2008 until Dec 2018.	RFA treatement	NA	7856 patient-years	The primary aim was to determine the 10-year cancer progression in patients undergoing EET for BE. Secondary aims included understanding the durability of complete remission of dysplasia (CR-D) and complete remission of intestinal metaplasia (CR-IM) together with rates of relapse from these states.	Definitions: CR-IM was defined as all biopsies clear of intestinal metaplasia at two consecutive endoscopies together with residual tongues of glandular mucosa measuring less than 3cm and the absence of dysplasia. Relapse from CR-D was defined as the biopsy proven recurrence of dysplasia at a single endoscopy. Relapse from CR-IM the same followed successful clearance of IM.	Rate of relapse from CR-IM was 4.2% in year 1, 10.1% in at 2 years and 18.7% at 8 years. In year 1: 3.2% LGD, 5.2% HGD, 2.8% IMC. Year 2: 9.8% LGd, 11.9% HGD, 11.4% IMC. Year 8: 11.3% LGD, 22.1% HGD, 14.7% IMC. 78.4% of all relapses relapses within the first 2 years after CRIM. Most relapses occur within 2 years. The authors state that the data suggests that continuing annual surveillance endoscopies beyond 2 years may offer little additional benefit whilst providing no evidence that the histological sybtype prior to EET should be used to guide surveillance guidelines.	
van Munster et al (2021)	Retrospective cohort		94	mean 74	ER for a neoplastic lesion without subsequent ablation				The first primary endpoint was progression to HGD/EAC in the remaining BE. For patients with remaining NDBE or LGD, detection of HGD/EAC was considered to be progres sion. For patients with persisting flat HGD, new EAC was progression as was a new visible lesion containing HGD. The second primary endpoint was all-cause mortality. This endpoint reflects whether the decision to prefer surveillance over ablation was justified for patients with expected limited life expectancy.Only patients who had ER mono based on age and/or comorbidity were included in the analysis of this second endpoint (n=73). Secondary endpoints included symptomatic EAC and/or EAC-related death and predictors for progression	During median 21 months of FU with a median of 4 endoscopies 0 patients progressed to advanced cancer; 17 patients progressed to HGD or LR-EAC (annual progression risk 8%); median time to progressin 26 months. 40% of the patients (29/73) died due to unrelated causes median 28 months after ER at a median age of 80 years (72-85%). endoscopic FU was stopped in 37/73 patients at median 20 months (5-59) after ER.	
Omidvari et al. (2021)	Model simulation				Surveillance for BE	Comparative modeling analysis				For men with no, mild, moderate and severe comorbidity, the optimal ages of last surveillance were 81, 80, 77 and 73 years respectively. For women, these ages were younger: 75, 73, 73 and 69 years.	