Article published online: 2022-04-14

Surveillance of Barrett's esophagus using wide-area transepithelial sampling: systematic review and meta-analysis

OPEN ACCESS

Authors

Bashar Qumseya[1], Aymen Bukannan[2], Robyn Rosasco[3], Xiuli Liu[4], Amira Qumseya[5]

Institutions

1 Division of Gastroenterology, Hepatology and Nutrition, University of Florida, Gainesville, Florida, United States
2 Archbold Gastroenterology Group, Thomasville, Georgia, United States
3 Florida A & M University, 334 Palmer Ave, Tallahassee, Florida, United States
4 Department of Pathology, Immunology and Laboratory Medicine, University of Florida, Gainesville, Florida, United States
5 Department of Public Health, University of Florida, Gainesville, Florida, United States

submitted 9.6.2021
accepted after revision 10.11.2021

Bibliography
Endosc Int Open 2022; 10: E394–E402
DOI 10.1055/a-1783-9015
ISSN 2364-3722

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Corresponding author
Bashar Qumseya, MD, MPH, FASGE, Associate Professor, Division of Gastroenterology, Hepatology and Nutrition, University of Florida, P.O. Box 100214, 1329 SW 16th St., Suite 5251, Gainesville, FL 32610-0214, United States
Fax: +1-352-627-4761
bashar.qumseya@medicine.ufl.edu

Supplementary material is available under https://doi.org/10.1055/a-1783-9015

ABSTRACT

Background and study aims Wide-area transepithelial sampling (WATS) is an emerging technique that may increase dysplasia detection in Barrett's esophagus (BE). We conducted a systematic review and meta-analysis of patients who underwent surveillance for BE assessing the additional yield of WATS to forceps biopsy (FB).

Methods We searched Pubmed, Embase, Web of science, and the Cochrane library, ending in January 2021. The primary outcomes of interest were the relative and absolute increase in dysplasia detection when adding WATS to FB. Heterogeneity was assessed using I^2 and Q statistic. Publication bias was assessed using funnel plots and classic fail-safe test.

Results A total of seven studies were included totaling 2,816 patients. FB identified 158 dysplasia cases, whereas WATS resulted in an additional 114 cases. The pooled risk ratio (RR) of all dysplasia detection was 1.7 (1.43–2.03), $P<0.001$, $I^2=0$. For high-grade dysplasia (HGD), the pooled RR was 1.88 (1.28–2.77), $P=0.001$, $I^2=33\%$. The yield of WATS was dependent on the prevalence of dysplasia in the study population. Among studies with high rates of dysplasia, the absolute increase in dysplasia detection (risk difference, RD) was 13% (8%-18%, $P<0.0001$, number needed to treat [NNT] = 8). The pooled RD in HGD was 9% (2%-16%), $P<0.001$, NNT = 11. For studies with a low prevalence of dysplasia, RD for all dysplasia was 2% (1%-3%), $P=0.001$, NNT = 50. For HGD, the RD was 0.6% (0.2%-1.3%), $P=0.019$, NNT = 166.

Conclusions In populations with a high prevalence of dysplasia, adding WATS to FB results in a significant increase in dysplasia detection.

Introduction

Barrett's esophagus (BE) is a premalignant condition characterized by the development of specialized intestinal metaplasia that replaces squamous epithelium of the columnar esophagus [1]. BE is one of the most important risk factors for esophageal adenocarcinoma (EAC) [2], a disease with increasing incidence in Western countries [3]. Development of adenocarcinoma from BE appears to go through a cascade of steps starting with non-dysplastic BE (NDBE), low-grade dysplasia (LGD), high-grade dysplasia (HGD), intramucosal adenocarcinoma (IMC), and finally invasive AC [4, 5]. To stem the increasing incidence of EAC, we need to improve our ability to detect BE (screening) and our ability to detect dysplasia (surveillance). The American Society of Gastrointestinal Endoscopy (ASGE) recently published guidelines on screening and surveillance for BE, and made a conditional recommendation to include wide-area transepithelial sampling with computer-assisted 3D analysis ($WATS^{3D}$), hereafter termed WATS, to improve dysplasia detection in BE [6]. WATS offers a novel approach to increase dysplasia detection by combining an abrasive brush to allow sampling of larger areas of the suspected BE segment and molecular diagnostics [7]. The rate at which WATS can increase dysplasia detection varies greatly by study. WATS may also help to improve the yield of BE diagnosis at the time of screening [8]. Therefore, we aimed to conduct a systematic review and meta-analysis to assess the increased yield of dysplasia detection in patients with BE.

Methods

Study selection

This study was based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines using a protocol developed by the study team a priori. Inclusion criteria were as follows: 1) clinical trials, prospective, or retrospective studies; 2) meeting abstracts from the last 3 years; 3) studies that assessed the diagnostic yield of BE or dysplasia in patients undergoing EGD; 4) available results for WATS and forceps biopsy (FB); and 5) clear definition of dysplasia. Studies were excluded if they were: 1) case reports or case series; 2) poor quality; 3) English language full text was not available; 4) indefinite for dysplasia (IND) or crypt dysplasia (CD) could not be separated from other dysplasia; or 5) deemed as outliers with reported effect estimate more than eight times the expected rate.

Search strategy & data extraction

The search strategy was designed by the study team with the help of an expert librarian (RR). Previous searches on this topic were used to inform our strategy. Databases searched included MEDLINE (Ovid), Web of Science, Embase, Cochrane Library and CENTRAL, and World Health Organization International Clinical Trials Registry Platform (WHO ICTRP) from inception. The last update of the search was on January 8, 2021. Details of our search strategy are listed in **Appendix 1**. Citations were saved as an EndNote library (Thompson Reuters, Carlsbad, California, United States) then improved into Covidence (covidence.org). Duplicates were removed in EndNote and Covidence as well. Studies were screened by title and abstracts by two reviewers (BQ, AB). Conflicts were resolved by consensus. We extracted data on author, publication type (abstract vs. manuscript), study design, definition of dysplasia, number of patients with BE, number of patients with dysplasia on WATS and on FB, and basic patient demographics like age, gender, rate, and BE length if available.

Outcomes of interest

The primary outcome of interest was the increased yield of dysplasia detection on WATS compared to FB. The primary effect estimates, which refer to the summary estimates of choice, were the relative and the absolute increase in dysplasia detection. The relative increase was defined as a risk ratio (RR) = proportion of dysplasia detection on combined WATS with FB divided by the proportion of patients with dysplasia on FB only. The absolute increase in dysplasia detection was defined as the risk difference (RD) = proportion of dysplasia detection on combined WATS with FB minus the proportion of patients with dysplasia on FB only. We hypothesized that this rate is highly dependent on the rate of dysplasia in the study population. Therefore, we planned a priori to use a meta-regression to stratify the results to control for the rate of dysplasia. The rate of dysplasia was defined as the proportion of patients with dysplasia found on WATS and FB out of the total number of patients with BE who had surveillance endoscopy. CD and IND were excluded from all analyses because of high interobserver variability among pathologists. Some pathologists may also consider CD as a type of IND.

Because the definition of dysplasia varied by study and the influenced the effect estimate in each study, we standardized the definition of dysplasia to ensure that we were comparing studies fairly. We analyzed studies that reported rates of dysplasia defined as HGD/AC, LGD, or both. The reference standard in most studies included random biopsies using the Seattle protocol. Advanced imaging modalities, including chromoendoscopy (CE), were used in some studies as part of the reference. Finally, as a secondary outcome, we also assessed the rate of BE detection among patients who were undergoing screening for BE.

We hypothesized several sources of heterogeneity a priori. These included:

1. Degree of dysplasia in population
2. Variation in dysplasia definition
3. Variation in study design
4. Variation in publication type
5. Variation in WATS or FB protocols among centers.

To control for expected heterogeneity, we planned several sensitivity analyses a priori that included study design (prospective vs. retrospective), publication type (manuscript vs. abstract), and rate of dysplasia in the population (for RD).

Quality assessment

For quality assessment of individual studies, we used the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS 2) [9]. Quality assessment was only performed for the six included manuscripts. Abstracts lack sufficient information to accurately assess their quality. The final results were reported in a tabular form and assessed two domains: risk of bias and applicability. For each, answer choices were "yes," "no," or "unclear."

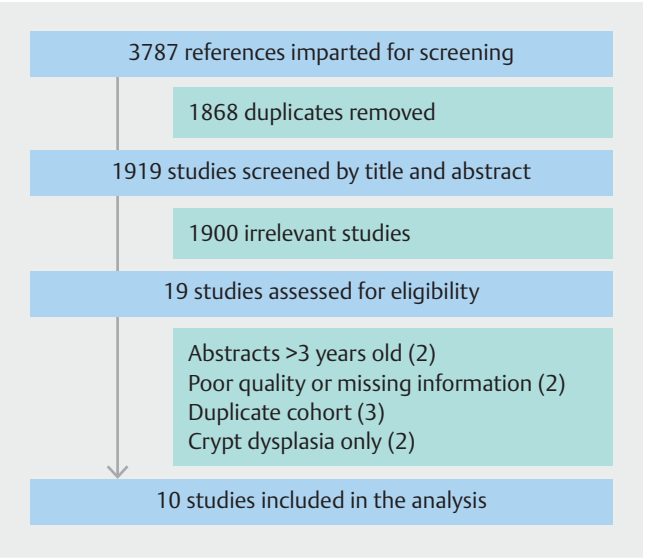
Statistical analysis

Because fixed effects models assume that the true effect size is the same in all studies, and given the heterogeneity in study designs and populations, we made the a priori decision to use random-effect modeling for all results. The primary effect estimates were the additional yield of all dysplasia, and the additional yield of HGD/AC, detected when adding WATS to FB. These was reported as the RR and RD with 95 %. We suspected a heterogeneity in RD based on the rate of dysplasia. Therefore, we planned a meta-regression to assess the effect of the rate of dysplasia on the absolute increase in dysplasia detection. We used β-coefficient to assess the degree of change in dysplasia detection based on the change in rate of dysplasia. Because the definition of high vs. low rate of dysplasia is subjective, we used the regression line to estimate a point at which the rate of dysplasia can be divided into two categories: high and low. This point was defined as the inflection point in the regression line. R^2 analog was used to assess the proportion of total between-study variance, which is explained by the regression model. We used Forest plots to show magnitude and direction of effect estimates. We used the I^2 to assess heterogeneity. This was defined as low (I^2 < 50 %), moderate (I^2 = 51 %–75 %), and high (I^2 > 75 %). To assess for publication bias, we used funnel plots and fail-safe test. We used CMA V3 (Biostat, Inc., Englewood, New Jersey, United States) for all statistical analyses.

Results

Our searches resulted in a total of 3,787 studies. Of these, 1,868 were removed as duplicates and 1,919 were screened by title and abstract. Among those, 1,900 were excluded and 19 were assessed for inclusion. Ten studies were included in the final analyses (▶ **Fig. 1**).

These included six published manuscripts [7, 8, 10–13] and four meeting abstracts [14–17]. Study design included two RCTs [7, 14], four multicenter studies [8, 11–13, 17], and four retrospective cohort studies [10, 15–17]. In seven of the studies [7, 8, 10–12, 15, 16], dysplasia reported included LGD and HGD/AC. In six studies [7, 10, 13–16], dysplasia was reported as HGD/AC. These formed the primary cohort of our study. In patients with BE, the mean length of BE ranged from 1.2 cm to 4.6 cm. In most studies, the majority of patients were men. BE was defined as detection of intestinal metaplasia on FB or WATS. Most studies included patients undergoing surveillance for BE. One study excluded patients with nodules [12]. Further study and patient characteristics are presented in ▶ **Table 1** and ▶ **Table 2**.



▶ **Fig. 1** Flow chart of study inclusion.

Yield of WATS in dysplasia detection

A total of seven studies [7, 8, 10–12, 15, 16] were identified which reported rates of dysplasia detection in WATS compared to FB. The seven studies totaled 2,816 patients. Of those, 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) (▶ **Fig. 2a**). 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 I^2 = 0 and Q = 2.45.

Six studies [7, 10,13–16] 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, I^2 = 33 %, Q = 7.49 (▶ **Fig. 2b**. Thus, the additional yield of HGD/AC was 88 % [28 %–177 %].

In four studies [7, 8, 10, 16] 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, I^2 = 0, Q = 1.78, ▶ **Fig. 2c**. Thus, the additional yield of LGD was 50 % (14 %-99 %).

Effect of prevalence of dysplasia on the yield of dysplasia

As hypothesized, the absolute increase in dysplasia detection varied based on the prevalence of dysplasia in the underlying population. In a meta-regression, the rate of dysplasia in the population was significantly associated with the absolute increase in dysplasia (β = 0.32 [95 %CI.17–.46], P < 0.001, I^2 = 0 %) (▶ **Fig. 3a**). This would suggest for each 10 % increase in the prevalence of dysplasia, there is a 3.2 % increase in absolute dysplasia detection in WATS. The R^2 analog = 1 indicates that this model accounted for 100 % of between-study variability.

▶ **Table 1** Patient and study characteristics of included studies.

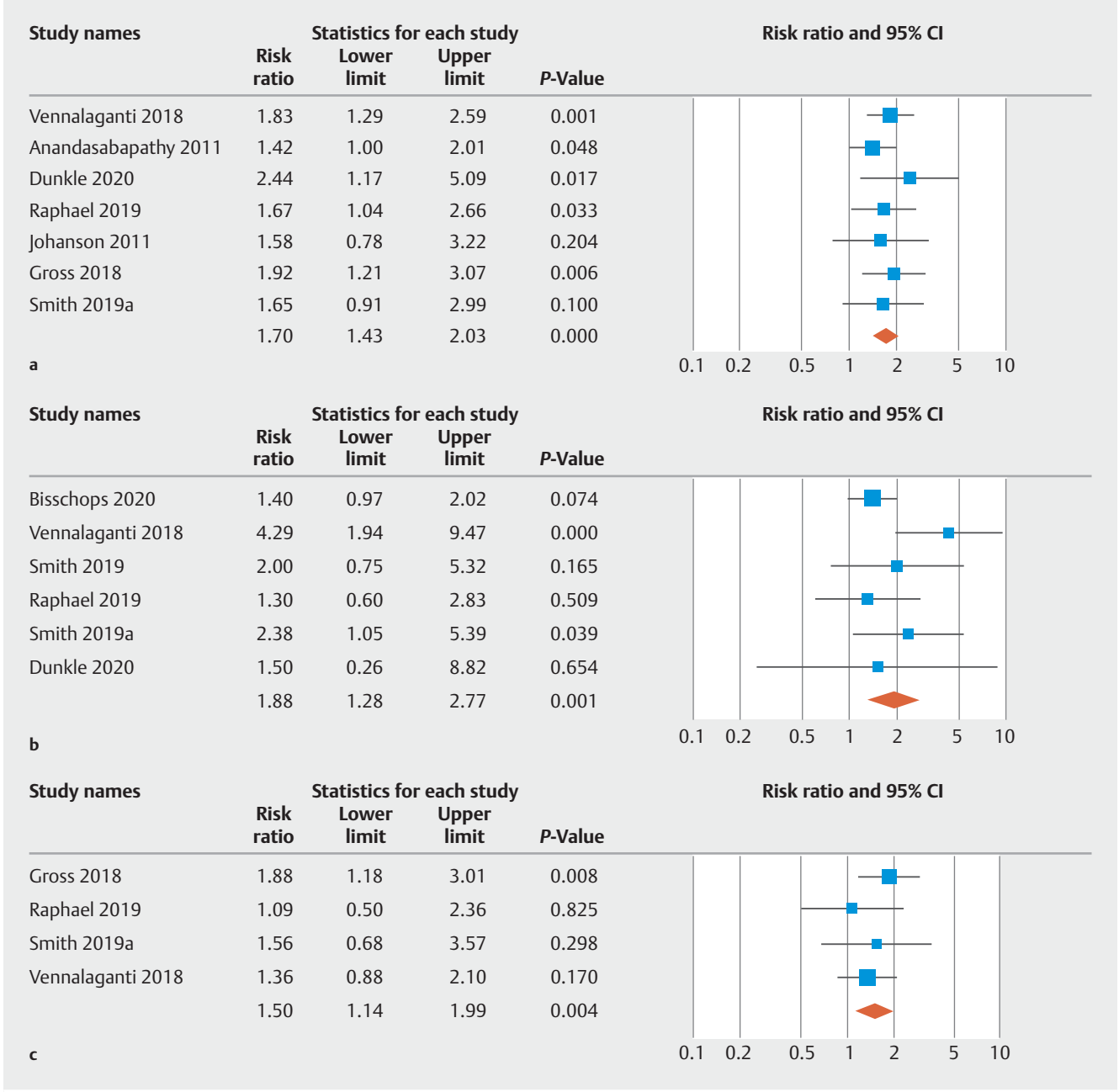
Study	Publication type	Study type	Patient population	Male %	Mean age or range (yr)	% White	BE length (cm)	Rate of dysplasia
Vennalaganti 2018	Manuscript	RCT-crossover	Surveillance pre- or post-ablation	76 %	63.4	95 %	4	0.400
Anandasabapathy 2011	Manuscript	Multicenter trial	BE with dysplasia, excluding nodules	82 %	65	84 %	4.6	0.358
Johanson 2011	Manuscript	Multicenter trial	Surveillance pre- or post-ablation	54 %	18–90	/	2.5	.0.49
Gross 2018	Manuscript	Multicenter trial	Screening & surveillance	43 %	59	/	NR	0.046
Raphael 2019	Manuscript	Retrospective	Surveillance pre- or post-ablation	73 %	65.2	/	3	0.330
Dunkle 2020	Abstract	Retrospective	Surveillance BE	/	/	/	/	0.185
Smith 2019a	Abstract	Retrospective	Post ablation	64 %	67	/	/	0.035
Bisschops 2020	Abstracts	RCT-crossover	Post EMR	84 %	68.4	/	/	/
Smith 2019	Manuscript	Multicenter trial	Screening & surveillance	39 %	56	/	1.2	/
Srinivasan 2019	Abstract	Retrospective		/	/	/	/	/

BE, Barrett's esophagus; RCT, randomized controlled trial; NR, not recorded; EMR, endoscopic mucosal resection.

▶ **Table 2** Increased detection of dysplasia and Barrett's esophagus in each of the included studies.

Study	Dysp type	# BE	WATS + FB	FB only	Dysp type	# BE	WATS + FB	FB only	# BE patients	WATS + FB	FB only	# Screened	BE WATS + FB	BE FB alone
Vennalaganti 2018	All	160	64	35	HGD/AC	160	30	7	160	38	28	/	/	/
Anandasabapathy 2011	All	151	54	38	NA	/	/	/	/	/	/	/	/	/
Johanson 2011	All	391	19	12	NA	/	/	/		/	/	792	243	142
Gross 2018	All	1,087	50	26	NA	/	/	/	1,087	49	26	4,203	1,087	594
Raphael 2019	All	106	35	21	HGD/AC	106	13	10	106	12	11	/	/	/
Dunkle 2020	All	119	22	9	HGD/AC	119	3	2	802	14	9	/	/	/
Smith 2019a	All	802	28	17	HGD/AC	802	19	8	/	/	/	/	/	/
Bisschops 2020	/	/	/	/	HGD/AC	147	49	35	/	/		/	/	/
Smith 2019	/	/	/	/	HGD/AC	802	19	8	/	/	/	11,093	4,254	1,684
Srinivasan 2019	/	/	/	/	/	/	/	/	/	/	/	108	82	62

Dysp, dysplasia; BE, Barrett's esophagus; FB, forceps biopsies; HGD, high-grade dysplasia; AC, adenocarcinoma; NA, not available

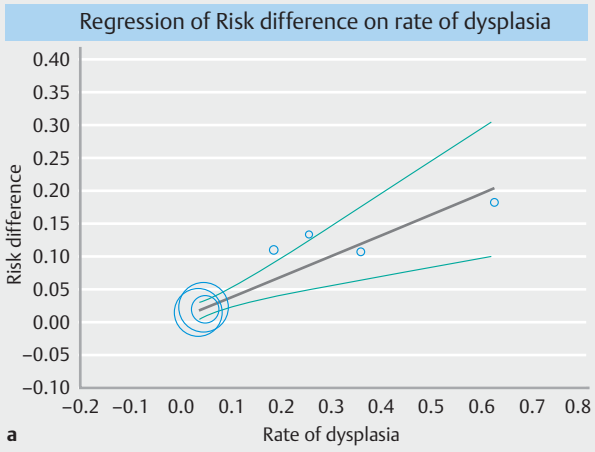


▶ **Fig. 2** Forest plot of **a** the relative increase (risk ratio, RR) of all dysplasia detection in seven included studies; **b** the relative increase of high-grade dysplasia (HGD) detection in six included studies; and **c** the relative increase of low-grade dysplasia (LGD) detection in four included studies.

Based on the regression line, we found that a 10 % dysplasia rate may be a good fit to separate studies with low versus high rates of dysplasia.

Among studies with high rates of dysplasia, on random-effect modeling, the pooled absolute increase in dysplasia detection was 13 % (RD.13 [.08–.18], $P<0.0001$) with no evidence of heterogeneity ($I^2=0$, $Q=5.17$) (▶ **Fig. 3b**). Based on this finding, the number needed to treat (NNT) was eight (95 %CI: 5.6–12.5). Therefore, in high-risk populations, we need to add WATS to FB in eight patients to detect one additional case of dysplasia. Similarly, the pooled absolute increase in HGD/AC detection in this population was 9 % (RD.09 [.02–.16], $P<0.001$, $I^2=54\%$, NNT for HGD = 11 (95 %CI: 6.3–50) (▶ **Fig. 3c**).

For studies with a low prevalence of dysplasia (<10 %), on random-effect modeling, the pooled absolute increase in dysplasia detection was only 2 % (RD.02 [.001–.03], $P=0.001$) with no evidence of heterogeneity ($I^2=0$, $Q=3.4$) (▶ **Fig. 3b**). Similarly, the pooled absolute increase in HGD detection in this population was .6 % (RD.06 [.002–.013], $P=0.019$, $I^2=32\%$, NNT for HGD/AC = 166 [95 % CI: 76.9– 500]) (▶ **Fig. 3c**].



▶ **Fig. 3a** Meta-regression of the absolute increase (risk difference, RD) of dysplasia detection based on the rate of dysplasia in the population.

WATS as a replacement for FB

We found five studies [8, 10–12, 15] that assessed the number of patients with dysplasia on WATS compared to the number of cases of dysplasia with FB. These studies had a total of 2,126 patients. Considering each modality separately, dysplasia was detected in 106 patients using FB and 103 patients using WATS. There was no significant difference in dysplasia detection between the two modalities (RR.96 [95 %CI:.69–1.35], $P = 0.816$, $I^2 = 36\,\%$, $Q = 6.2$) (▶ **Fig. 3d**). While WATS identified cases that were missed by FB, WATS also missed cased of dysplasia that were detected on FB (**Supplementary Table 1**). This indicates that using WATS to replace FB would not result in an increase in dysplasia detection, as the additional cases of dysplasia picked up on WATS may be offset by the number cases missed on WATS but detected on FB.

Screening for BE

As a secondary outcome, we reviewed studies that assessed screening for BE and found four such studies [8, 11, 13, 17]. They totaled 16,196 patients, in whom FB detected 2,482 patients with BE, whereas the addition of WATS increased this number to 5,666. The indications for screening varied by study and within study. There was very significant heterogeneity ($I^2 = 97\,\%$); therefore, the estimate was not pooled. The relative increase in BE detection ranged from 32 % to 250 % (**Supplementary Fig. 1**). Furthermore, it is unclear how many of these may have been done for an irregular z-line rather than BE of at least 1 cm using the current definition of BE.

Risk bias and quality assessment

We tried to assess and control for bias in several ways. First, to assess for publication bias, we used a funnel plot. There was no evidence of publication bias (**Supplementary Fig. 2a**). We recognized that the results were limited due to the low number of studies; therefore, we further assessed publication bias using the classic fail-safe test. This showed that we need 56 null studies to change the *P* value to non-significance. When one study was removed at a time, we noted that none of the included studies had an overall influence on the pooled effect estimate (**Supplementary Fig. 2b**).

In addition, we assessed the quality of each study using QUADAS 2. This showed no major concerns with the quality of the included studies (**Supplementary Table 2**). One abstract by Elden et al. [18] was removed from the primary analysis for two reasons. The study population was not clearly defined and the study was identified as an outlier based on our a priori criteria. In this study by Elden et al., the relative risk of dysplasia when adding WATS to FB was very high at 9.4 (95 %CI 3.77–23.44), thus exceeding the a priori threshold for exclusion as an outlier. In a sensitivity analysis, when Elden et al. was included, the pooled additional yield of all dysplasia was 95 % (95 %CI 47 %–155 %), $I^2 = 54\,\%$, Q = 15.4 (**Supplementary Fig. 2c**). Therefore, adding this outlier study does not change the overall direction or conclusion of our study.

Finally, we conducted sensitivity analyses based on study design and publication type.

In sensitivity analyses, study design (prospective vs. retrospective) and publication type (abstract vs. manuscript) did not affect the RR of dysplasia detection (RR = 1.67 [95 %CI 1.35–2.05] for prospective studies, RR = 1.79 [95 %CI 1.29–2.09] for retrospective studies, $P = 0.712$) (**Supplementary Fig. 2d**; RR = 1.93 [1.21–3.06] for abstracts, and RR = 1.67 [1.38–2.02] for manuscripts, $P = 0.57$) (**Supplementary Fig. 2e**).

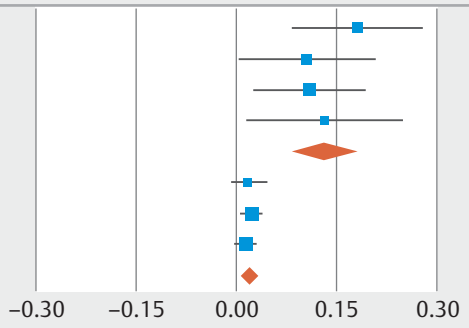
Discussion

Clinical implications

In this systematic review and meta-analysis, we report that adding WATS to FB results in an increase in dysplasia detection, including HGD/AC. In populations with high rates of dysplasia, the absolute increase in dysplasia was high (9 % for HGD and 13 % for all dysplasia) with a low NNT (11 for HGD and 8 all dysplasia). However, the rate was much smaller in studies in which patients had low rates of dysplasia (0.6 % for HGD and 2 % for all dysplasia) and the NNT was high (166 for HGD and 50 all dysplasia). WATS did not perform better than FB as a stand-alone modality to replace FB.

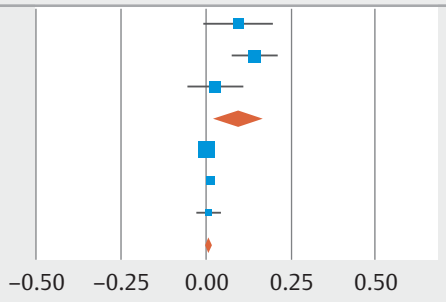
Improving dysplasia detection in BE patients is of great importance [19]. As described above, the main finding of our study is that the absolute increase in dysplasia detection by WATS varied considerably based on the rate of dysplasia in the underlying population. Among low-risk populations, such as those with no history of dysplasia, and who form the majority of BE patients in practice, our data indicate that there is a high NNT. While no studies, to our knowledge, have addressed cost-effectiveness of WATS in BE surveillance, the high NNT may be cost-prohibitive. Therefore, further studies are needed before WATS can be routinely used for such patient populations. On the other hand, our results support the use of WATS in high-risk populations, given the low NNT both to increase detection of HGD and all dysplasia (LGD and HGD). High-risk patients may

Group by Rate	Study name	Risk difference	Lower limit	Upper limit	*P*-Value
High	Vennalaganti 2018	0.18	0.08	0.28	0.000
High	Anandasabapathy 2011	0.11	0.00	0.21	0.044
High	Dunkle 2020	0.11	0.02	0.19	0.011
High	Raphael 2019	0.13	0.01	0.25	0.027
High		0.13	0.08	0.18	0.000
Low	Johanson 2011	0.02	−0.01	0.05	0.199
Low	Gross 2018	0.02	0.01	0.04	0.005
Low	Smith 2019a	0.01	−0.00	0.03	0.096
Low		0.02	0.01	0.03	0.001



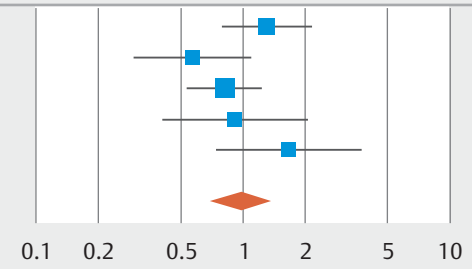
b

	Study name	Risk difference	Lower limit	Upper limit	*P*-Value
High	Bisschops 2020	0.10	−0.01	0.20	0.069
High	Vennalaganti 2018	0.14	0.08	0.21	0.000
High	Raphael 2019	0.03	−0.06	0.11	0.507
High		0.09	0.02	0.16	0.012
Low	Smith 2019	0.00	−0.00	0.01	0.156
Low	Smith 2019a	0.01	0.00	0.03	0.033
Low	Dunkle 2019	0.01	−0.03	0.04	0.651
Low		0.01	−0.00	0.01	0.157



c

	Study name	Risk ratio	Lower limit	Upper limit	*P*-Value
	Gross 2018	1.31	0.79	2.16	0.297
	Raphael 2019	0.57	0.30	1.10	0.095
	Anandasabapathy 2011	0.82	0.54	1.24	0.339
	Johanson 2011	0.92	0.41	2.05	0.832
	Dunkle 2019	1.67	0.74	3.76	0.219
		0.96	0.69	1.34	0.816



d

▶ **Fig. 3** **b** Forest plot of absolute increase (RD) in dysplasia detection stratified by rate of dysplasia (high vs. low) among patients with Barrett's esophagus. c Forest plot of absolute increase (RD) in HGD/AC detection stratified by rate of dysplasia (high vs. low) among patients with Barrett's esophagus. **d** Forest plot of the relative risk (RR) of dysplasia detection on WATS alone compared to FB alone.

include those with a history of dysplasia or history of endoscopic eradication therapy. To our knowledge, our study is the first and largest to highlight this difference, which is clinically relevant.

Another important question we aimed to address is whether WATS can be used to replace the Seattle protocol, which can be time-consuming and resource-intensive. The goal of many advances in the dysplasia detection has been to forgo the Seattle protocol. In the case of chromoendoscopy, a previous meta-analysis showed that targeted biopsies combined with Seattle protocol biopsies produced the highest yield of dysplasia [20]. In our analysis, we showed that WATS led to increased dysplasia detection by finding cases of dysplasia missed by FB and WATS also missed many cases of dysplasia identified by FB. Our results indicate that WATS should not be used alone and favors recent ASGE guidelines [6], which recommend adding WATS to the Seattle protocol. If dysplasia is found on WATS but not FB, we have limited data on how various centers deal with this. However, this was not the focus of this systematic review and may require further studies to address.

Whether WATS can increase dysplasia detection in patients who undergo surveillance using other advanced imaging is yet

to be decided. A study by Raphael et al. [21] compared the additional yield of WATS to high-definition white light endoscopy with CE and volumetric laser endomicroscopy. The authors noted that the WATS added yield was 19 %. However, the CI was high (0.6 %–45.7 %). While these limited data suggest that WATS is still beneficial even in cases where other advanced imaging modalities have been performed, given the low number of patients in the study, larger studies are need to confirm this trend. The lack of standardization of the "reference" procedure was evident in the studies included in our cohort. Based on the recent ASGE guidelines on BE, we argue that standard biopsies should include targeted biopsies based on CE, followed by Seattle protocol biopsies. Using such a reference standard is crucial for future studies and will help improve patient care.

As a secondary outcome, we reviewed four studies that assessed screening for BE [8, 11, 13, 17]. Although, in a total of 16,196 patients, FB detected 2,482 patients with BE while the addition of WATS increased this number to 5,666, there was very significant heterogeneity (I^2 = 97 %) with the relative increase in BE detection ranging from 32 % to 250 % in these studies. It is unclear how many of these may have been done for an irregular z-line rather than BE of at least 1 cm. Thus, the role of WATS in BE screening remains to be determined.

Strength and limitations

Based on our systematic review, we noticed that studies assessing the use of WATS in BE had several limitations, which include: the variable definition of dysplasia reported CD; indefinite for dysplasia (IND), LGD, and HGD/AC; varying indications for surveillance and inclusion criteria (surveillance post-EET, surveillance in NDBE); and the fact that most WATS studies were industry-sponsored. A previous meta-analysis [22] tried to synthesize WATS data but resulted in very high heterogeneity, making the results largely uninterpretable. Our study tried to control for the above-mentioned limitations in various ways.

The first limitation we tried to address is the heterogeneity in dysplasia definition. Some studies included LGD, while others excluded LGD. Similarly, some studies included IND with LGD as one category. Other studies also included CD in the dysplasia categories. Therefore, trying to analyze all studies together when the outcomes are not the same is inappropriate. This is a major hurdle in analyzing WATS data and has contributed to the significant heterogeneity reported in a previous meta-analysis with an I^2 of 97 % [22]. Therefore, for the primary analyses, we separated studies based on how dysplasia was reported. This standardized approach was proposed a priori and is essential for this kind of data synthesis.

The second limitation is the varying indication for surveillance among studies including various inclusion and exclusion criteria. Most studies included patients undergoing surveillance for BE. We recognized that the variation in rate of dysplasia in the study population was a major source of heterogeneity. This is not a factor when calculating relative values, but it is a major factor when looking at absolute effect estimates, i. e. RD. A study of patients who already had dysplasia and underwent radiofrequency ablation prior to surveillance would be expected to detect far more cases of dysplasia compared to a study of patients who had mostly NDBE. Our meta-regression indicated that the prevalence of dysplasia is a significant contributor to heterogeneity in calculation of absolute increase in dysplasia detection. In stratifying the data based on prevalence of dysplasia, we were able to address heterogeneity and calculate more accurate absolute measures and NNT.

The prevalence of dysplasia was not a factor in the calculation of RR. This finding was predicted and expected. Take the case of two hypothetical studies of 100 patients each. One study has low prevalence of dysplasia of 6 %. The other has high prevalence of dysplasia of 30 %. In the first study, FB detects three cases of dysplasia, while WATS detects an additional three cases. In the second study, FB detects 15 cases and WATS picks up another 15 cases of dysplasia. Note that the RR calculation for both studies is two. That is, WATS doubled the number of cases of dysplasia in both studies. However, the absolute increase is markedly different between the two studies. In the first study, the absolute increase is 3 %. In the second study, the absolute increase in dysplasia detection is 15 %. Thus, in absolute terms, the prevalence of dysplasia would be expected to contribute to heterogeneity in a way that should not occur in the relative ratio calculation. These findings can be interpreted to mean that WATS consistently increases dysplasia detection in all patient populations relative to FB. However, this increase may not be clinically relevant in low-prevalence populations.

Furthermore, our study team has no industry support, which removes some of the limitations of previous studies. On the other hand, there has been concern about the correlation between dysplasia detected on WATS compared to FB. The criteria used to define dysplasia for the formalin-fixed tissue, performed during WATS, is identical to routine pathology on biopsies.

Therefore, it is reasonable to assume that dysplasia on WATS should be treated similarly to dysplasia on FB. Further understanding such a difference, if one does exist, is beyond the scope of this study and should be the focus of future studies on this topic.

Finally, it is impossible to predict whether dysplasia detected on WATS and missed on FB would be detected on future FBs. None of the studies addressed this issue. However, we know that dysplasia can be missed even in tertiary centers, where most of these studies are done. So even in expert hands, dysplasia can be missed, especially in longer segments of BE. Therefore, WATS likely has a role in these patients. Yet, this requires future investigation.

Conclusions

WATS is associated with an increase in dysplasia detection that is most pronounced in surveillance of populations with a high prevalence of dysplasia. The clinical value of the increased detection rate of dysplasia for WATS requires further study because there are limited data about how various centers deal with these findings.

Competing interests

The authors declare that they have no conflict of interest.

References

[1] Hvid-Jensen F, Pedersen L, Drewes M et al. Incidence of adenocarcinoma among patients with Barrett's esophagus. N Engl J Med 2011; 365: 1375–1383

[2] Shaheen NJ, Richter JE. Barrett's oesophagus. Lancet 2009; 373: 850–861

[3] Hur C, Miller M, Kong CY et al. Trends in esophageal adenocarcinoma incidence and mortality. Cancer 2013; 119: 1149–1158

[4] Qumseya BJ, Wani S, Gendy S et al. Disease progression in Barrett's Low-grade dysplasia with radiofrequency ablation compared with surveillance: systematic review and meta-analysis. Am J Gastroenterol 2017; 112: 849–865

[5] Wani S, Qumseya B et al. Standards of Practice Committee. Endoscopic eradication therapy for patients with Barrett's esophagus-associated dysplasia and intramucosal cancer. Gastrointest Endosc 2018; 87: 907–931 e9

[6] Qumseya B, Sultan S et al. American Society of Gastrointestinal Endoscopy Standards Of Practice Committee. ASGE guideline on screening and surveillance of Barrett's esophagus. Gastrointest Endosc 2019; 90: 335–359 e2

[7] Vennalaganti PR, Kaul V, Wang KK et al. Increased detection of Barrett's esophagus-associated neoplasia using wide-area trans-epithelial sampling: a multicenter, prospective, randomized trial. Gastrointest Endosc 2018; 87: 348–355

[8] Gross SA, Smith MS, Kaul V et al. Increased detection of Barrett's esophagus and esophageal dysplasia with adjunctive use of wide-area transepithelial sample with three-dimensional computer-assisted analysis (WATS). United Europ Gastroenterol J 2018; 6: 529–535

[9] Whiting PF, Rutjes AW, Wetwood ME et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011; 155: 529–536

[10] Raphael KL, Stewart M, Sejpal D et al. Adjunctive yield of wide-area trans-epithelial sampling with computer-assisted three-dimensional analysis in detection of dysplasia after advanced imaging and random biopsies in Barrett's esophagus. Am J Gastroenterol 2019; 114: S258–S259

[11] Johanson JF, Frakes J, Eisen D et al. Computer-assisted analysis of abrasive transepithelial brush biopsies increases the effectiveness of esophageal screening: a multicenter prospective clinical trial by the EndoCDx Collaborative Group. Dig Dis Sci 2011; 56: 767–772

[12] Anandasabapathy S, Sontag S, Graham DY et al. Computer-assisted brush-biopsy analysis for the detection of dysplasia in a high-risk Barrett's esophagus surveillance population. Dig Dis Sci 2011; 56: 761–766

[13] Smith MS, Ikonomi E, Bhuta R et al. Wide-area transepithelial sampling with computer-assisted 3-dimensional analysis (WATS) markedly improves detection of esophageal dysplasia and Barrett's esophagus: Analysis from a prospective multicenter community-based study. Dis Esophagus 2019; 32: doy099

[14] Bisschops R, Hairdry R, Messmann H et al. 243 Wide area transepithelial sample esophageal biopsy combined with computer assisted 3-dimensional tissue analysis (WATS3D) for detection of high grade dysplasia and adenocarcinoma in Barrett: European multicenter, prospective, randomized, tandem study. Gastrointest Endosc 2020; 91: AB23

[15] Dunkle A, Andersen M, Lisovsky M et al. Comparative analysis of wide-area transepithelial sampling (wats) versus endoscopic biopsy in diagnosing dysplasia in Barrett's esophagus. Modern Pathol 2020; 33: 1849

[16] Smith MS, Kaul V, Odze R. Wide area transepithelial sampling improves detection of intestinal metaplasia and dysplasia following endoscopic ablation of Barrett's esophagus. Gastroenterology 2019; 156: S284

[17] Srinivasan S, Agha YH, Hyder J et al. WATS3D vs. traditional forceps biopsy in screening of barrett's esophagus: a community hospital experience. Am J Gastroenterol 2019; 114: S212–S213

[18] Elden AC, Kumar P, Ghosh S et al. Evidence supporting a continued role for standard 4-quadrant biopsies along with wide area transepithelial sampling with computer-assisted 3d analysis (WATS3D) in the diagnosis and management of Barrett's esophagus: a large single-center experience. Gastrointest Endosc 2020; 91: AB402–AB403

[19] Sharma P, Savides TJ, Canto MI et al. The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on imaging in Barrett's esophagus. Gastrointest Endosc 2012; 76: 252–254

[20] Qumseya BJ, Wang H, Badie N et al. Advanced imaging technologies increase detection of dysplasia and neoplasia in patients with Barrett's esophagus: a meta-analysis and systematic review. Clin Gastroenterol Hepatol 2013; 11: 1562–70 e1-2

[21] Raphael KL, Stewart M, Sejpal DV et al. Adjunctive Yield of wide-area transepithelial sampling for dysplasia detection after advanced imaging and random biopsies in Barrett's esophagus. Clin Transl Gastroenterol 2019; 10: e00107

[22] Suresh KVC, Harne P, Patthipati VS et al. Wide-area transepithelial sampling in adjunct to forceps biopsy increases the absolute detection rates of Barrett's oesophagus and oesophageal dysplasia: a meta-analysis and systematic review. BMJ Open Gastroenterol 2020; 7: e000494