LARYNGOLOGY

Validation of the European Laryngological Society classification of glottic vascular changes as seen by narrow band imaging in the optical biopsy setting

Francesco Missale[1,2] · Stefano Taboni[3,4] · Andrea Luigi Camillo Carobbio[1,5] · Francesco Mazzola[6] · Giulia Berretti[7] · Andrea Iandelli[1,5] · Marco Fragale[1,5] · Francesco Mora[1,5] · Alberto Paderno[7] · Francesca Del Bon[7] · Giampiero Parrinello[1] · Alberto Deganello[7] · Cesare Piazza[7] · Giorgio Peretti[1,5]

Abstract

Purpose In 2016, the European Laryngological Society (ELS) proposed a classification for vascular changes occurring in glottic lesions as visible by narrow band imaging (NBI), based on the dichotomic distinction between longitudinal vessels (not suspicious) and perpendicular ones (suspicious). The aim of our study was to validate this classification assessing the interobserver agreement and diagnostic test performance in detecting the final histopathology.

Methods A retrospective study was carried out by reviewing clinical charts, preoperative videos, and final pathologic diagnosis of patients submitted to transoral microsurgery for laryngeal lesions in two Italian referral centers. In each institution, two physicians, independently re-assessed each case applying the ELS classification.

Results The cohort was composed of 707 patients. The pathologic report showed benign lesions in 208 (29.5%) cases, papillomatosis in 34 (4.8%), squamous intraepithelial neoplasia (SIN) up to carcinoma in situ in 200 (28.2%), and squamous cell carcinoma (SCC) in 265 (37.5%). The interobserver agreement was extremely high in both institutions ($k = 0.954$, $p < 0.001$ and $k = 0.880$, $p < 0.001$). Considering the diagnostic performance for identification of at least SIN or SCC, the sensitivity was 0.804 and 0.902, the specificity 0.793 and 0.581, the positive predictive value 0.882 and 0.564, and the negative predictive value 0.678 and 0.908, respectively.

Conclusion The ELS classification for NBI vascular changes of glottic lesions is a highly reliable tool whose systematic use allows a better diagnostic evaluation of suspicious laryngeal lesions, reliably distinguishing benign ones from those with a diagnosis of papillomatosis, SIN or SCC, thus paving the way towards confirmation of the optical biopsy concept.

Keywords Narrow band imaging · European Laryngological Society classification · Optical biopsy · Laryngeal cancer · Vascular changes · Endoscopy

Francesco Missale and Stefano Taboni have equally contributed to the manuscript.

Cesare Piazza and Giorgio Peretti have equally contributed as senior authors.

✉ Andrea Luigi Camillo Carobbio
a.carobbio@studenti.unibs.it

1 IRCCS Ospedale Policlinico San Martino, Genoa, Italy

2 Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy

3 Section of Otorhinolaryngology, Head and Neck Surgery, Azienda Ospedaliera di Padova, University of Padua, Padua, Italy

4 Department of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy

5 Department of Surgical Sciences and Integrated Diagnostics (DISC), University of Genoa, Genoa, Italy

6 Department of Otolaryngology, Head and Neck Surgery, IRCCS Regina Elena National Cancer Institute, Rome, Italy

7 Unit of Otorhinolaryngology, Head and Neck Surgery, ASST Spedali Civili di Brescia, University of Brescia, Brescia, Italy

Introduction

Early detection and diagnosis of laryngeal squamous cell carcinoma (SCC) are crucially involved not only in reducing mortality, but also to optimize therapeutic approaches aimed at achieving the best organ and functional preservation [1, 2]. Fortunately, glottic SCC, the most common laryngeal tumor localization, is more frequently detected at an earlier stage than tumors originating in other subsites of the head and neck due to its early (albeit highly non-specific) symptoms [3]. Laryngeal SCC examination is usually performed by flexible (video)endoscopy under white light (WL) and relies on the analysis of superficial characteristics (size, color, location, single or multifocal appearance) and visible morphological features (smoothness, irregularity, keratinization, ulceration, submucosal growth), per se non-pathognomonic and possibly overlapping each other in malignant and benign pathologies, especially when diagnosed at early stages. This implies the frequent need to obtain an incisional biopsy before deciding on the therapeutic approach, with an increase in costs, anesthetic risks, and potential undue damage to the vocal cords.

Computed tomography and magnetic resonance imaging definitively play a major role in diagnosis of more advanced diseases, providing information about the involvement of laryngeal structures and lateral neck lymph nodes, while, on the other hand, they fall shortly in identifying and characterizing superficial mucosal lesions. By contrast, the judicious use of high definition and better contrasted videoendoscopic images now offer staggering details in evaluation of epithelial and superficial vascular patterns. Moreover, clinicians can increasingly benefit from novel optical diagnostic methods, providing information even closer to those obtained by formal histopathological examination, thus differentiating between normal mucosa and discrete lesions and, among the latter, between those with benign versus malignant behaviors [4]. In this context, narrow band imaging (NBI) is a well-established bioendoscopic technique using filtered wavelengths to enhance microvascular alterations associated with preneoplastic and neoplastic transformation of the upper aerodigestive tract (UADT) mucosa [5–8]. Since its first introduction in the late 1990s, the use of NBI has considerably upgraded physicians' ability for non-invasive detection and delineation of suspicious mucosal lesions, and is thus beneficial in the diagnosis of a variety of benign and malignant lesions [9]. However, the need for a common language to be shared among clinicians to describe NBI-enhanced vascular patterns led to the proposal of different classifications during the last decades [10–12]. In 2016, the Working Committee on Endoscopic Laryngeal Imaging of the European Laryngological Society (ELS) published a new proposal for a simplified (dichotomic) description

of vocal fold vascular changes as seen under NBI [12]. In this system, the authors distinguished between normal and pathologic vascular patterns of the vocal folds. The latter, in turn, were divided into longitudinal and perpendicular vascular changes. Longitudinal vessels characterize benign lesions, while perpendicular ones (i.e. dot-like intrapapillary capillary loops [IPCL], enlarged and worm-like vessels) are considered signs of benign neoplasms (such as papillomatosis), squamous intraepithelial neoplasia (SIN), or frankly malignant lesions.

The present study aims to assess the performance of the ELS classification of vascular changes in a broad multicenter cohort, testing its interobserver agreement as primary endpoint, and analyzing its accuracy in predicting the final pathological results in an optical biopsy setting, i.e. by evaluating the diagnostic accuracy of NBI by comparing it with the final histopathologic diagnosis obtained after complete removal of the glottic lesion.

Methods

Study population

A retrospective study was carried out enrolling patients treated at the Departments of Otorhinolaryngology—Head and Neck Surgery of the Universities of Genoa (Center A; from January 2012 to December 2016) and Brescia (Center B; from January 2015 to December 2018), Italy.

All patients enrolled were affected by laryngeal lesions; a pre-treatment videoendoscopic evaluation with both WL and NBI was performed in the office as well as in the operatory theater, and the records were saved in his/her medical chart; the laryngeal lesion was treated by a transoral microsurgical approach by either cold instrumentation and carbon dioxide (CO_2) laser; postoperative histopathologic assessment was obtained to classify the resected tissue as benign, dysplastic or malignant. Histopathological diagnosis was performed according to the WHO classification system [13].

Clinical diagnostic work-up

All patients were preoperatively evaluated by high-definition television (HDTV)-WL and HDTV-NBI through a videorhinolaryngoscope ENF-VQ or ENF-VH coupled to an Evis Exera II HDTV camera connected to an Evis Exera II CLV-180B light source (Olympus Medical System Corp., Tokyo, Japan). Just before surgery, in the operating room, with patient under general anesthesia, intraoperative HDTV-WL and HDTV-NBI rigid endoscopy with 0° and 70° telescopes (Karl Storz, Tuttlingen, Germany) was also systematically performed. On the basis of this diagnostic work-up, laryngeal lesions were subsequently removed by either a

phonomicrosurgical approach (in case of benign lesions) or excisional biopsy (in case of papillomatosis, SIN, carcinoma in situ [CIS] or invasive SCC) by type I–III cordectomies according to the ELS classification of cordectomies [14].

Clinical evaluation applying the ELS classification

Clinical records of the study population, including demographic features and information on previous treatments in terms of laryngeal surgery, head and neck radiotherapy, or other treatments before the index transoral microsurgical procedure were retrieved from the hospital databases. Two independent physicians from each institution with at least a 3-year-experience in the use of NBI, blinded to the final histopathologic result, retrospectively and independently reviewed the intraoperative videoendoscopic recordings. Applying the ELS classification for laryngeal vascular changes [15], each case was categorized as suspicious for malignancy (presence of perpendicular vascular abnormalities as shown in Fig. 1) or non-suspicious
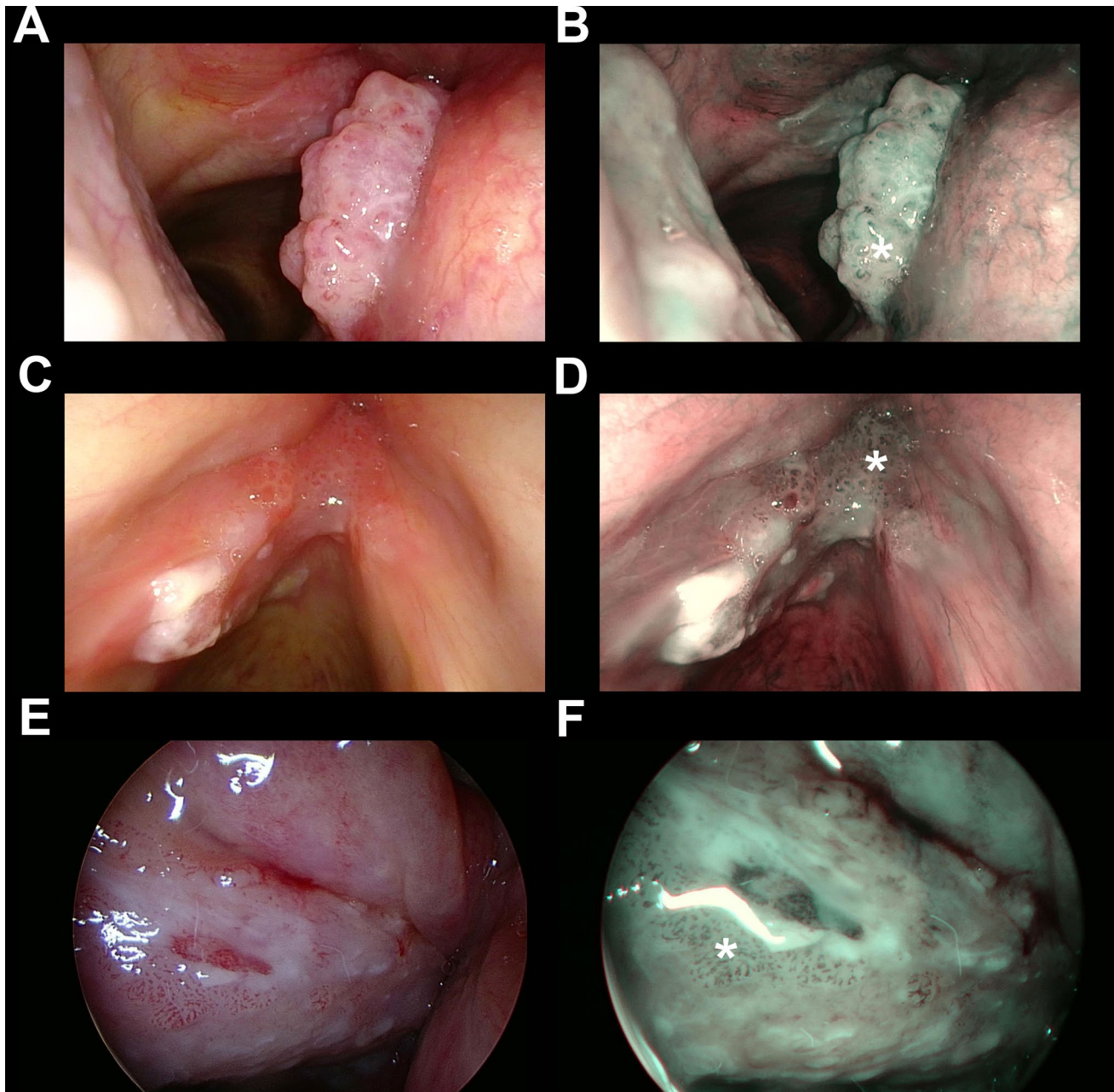

Fig. 1 Endoscopic picture of three representative cases of SCC (**a**–**d**) or CIS (**e**, **f**) correctly identified as suspicious by the presence of perpendicular vascular abnormalities (* in all panels) evaluating the NBI endoscopic appearance (**b**, **d**, **f**) and applying the ELS classification
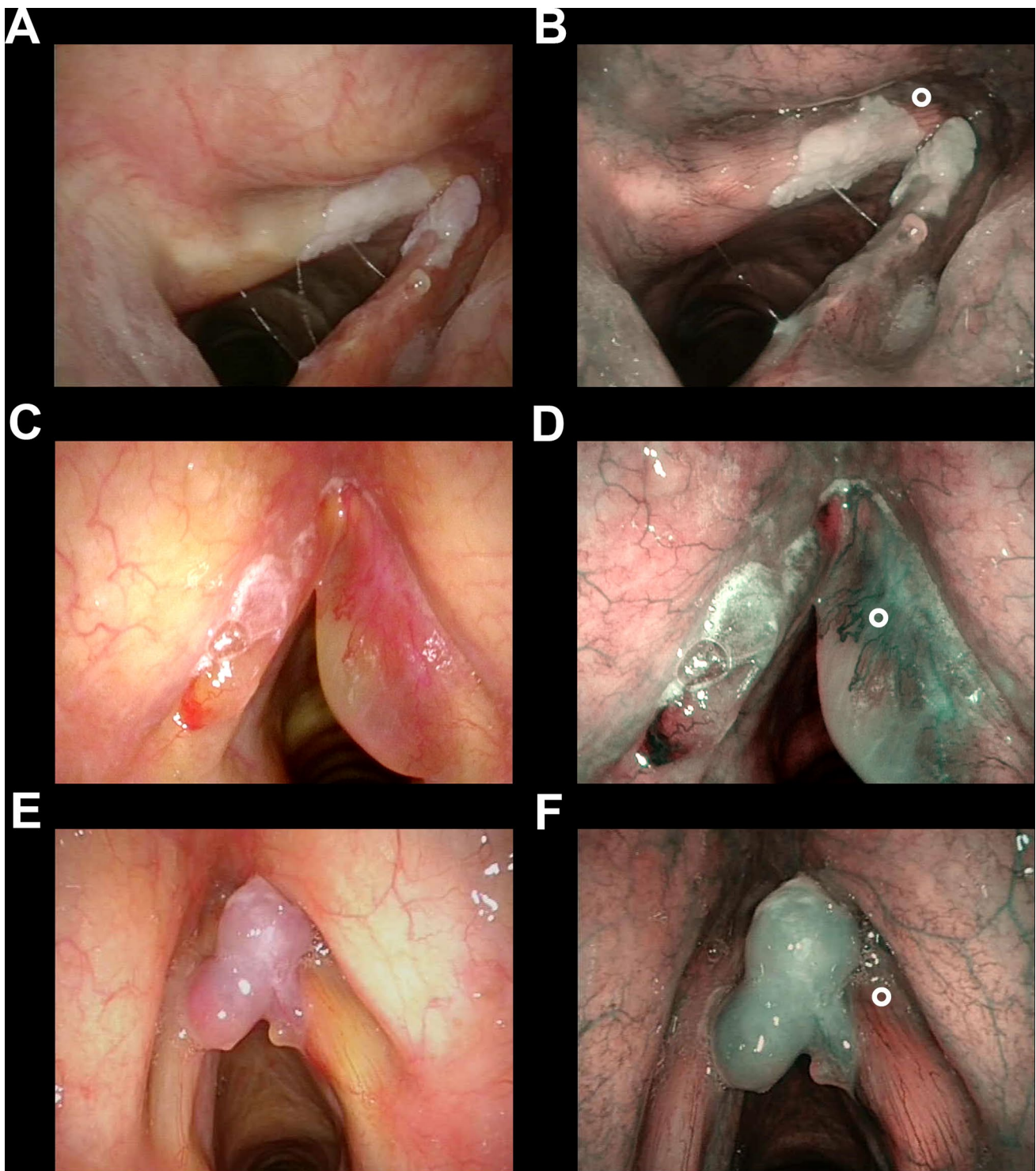
Fig. 2 Endoscopic picture of three representative cases of benign glottic lesions: keratosis without atypia (**a**, **b**), Reinke's edema (**c**, **d**) and polyp (**e**, **f**) correctly identified as benign lesions without identifying any perpendicular vascular abnormalities evaluating the NBI endoscopic appearance (**b**, **d**, **f**) and applying the ELS classification. The ° in all panels points to non-suspicious longitudinal vascular abnormalities that can be observed inside the lesion (**d**) of at its boundary (**b**, **f**)
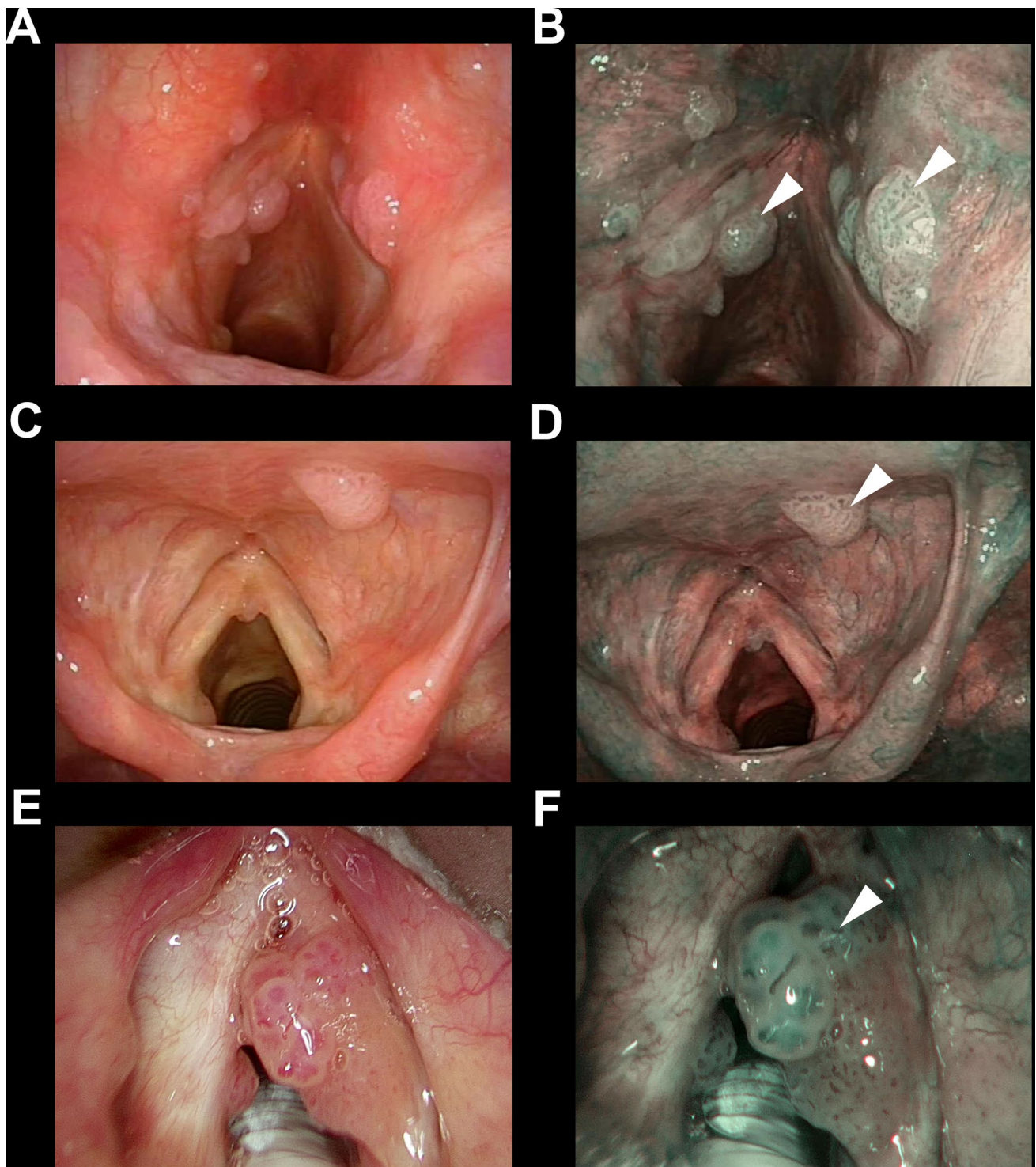
Fig. 3 Endoscopic pictures of three representative cases (**a**, **b**; **c**, **d**; and **e**, **f**) of recurrent laryngeal papillomatosis correctly identified detecting wide angle IPCLs (arrowheads in all panels) evaluating the NBI endoscopic appearance (**b**, **d**, **f**)

(undetectable perpendicular vascular changes or longitudinal ones as shown in Fig. 2). In case of interobserver disagreement, consensus was reached by direct comparison between the examiners. The identification of features of respiratory papillomatosis (i.e., wide angle IPCL) was also considered as a secondary endpoint (Fig. 3).

Table 1 Definition of interrater agreement qualitative scores according to Altman et al. [15, 16] and Clinical Utility indexes grading according to Mitchell [18]

Agreement classification		Clinical Utility Index classification	
κ	Strength of agreement	CU+Ve or CU−Ve	Grading
<0.21	Poor	<0.49	Poor utility
0.21–0.40	Fair	0.49–0.63	Satisfactory utility
0.41–0.60	Moderate	0.64–0.80	Good utility
0.61–0.80	Good	0.81–1.00	Excellent utility
0.81–1.00	Very good		

CU + Ve positive clinical utility index, *CU−Ve* negative clinical utility index

Statistical analysis

Clinical data were reported as absolute and relative frequencies. The reliability of the ELS classification was assessed for each independent cohort measuring the Cohen's k statistic and the strength of agreement graded according to Altman et al. [16, 17], as reported in Table 1. Considering the final consensus of the evaluation, we assessed the performance of the diagnostic test for prediction of the final pathologic result (at least SIN1, up to SCC) in an optical biopsy setting. For better understanding of the clinical utility of applying the classification and detecting papillomatosis-like features, the Clinical Utility (CU) indexes were also derived, taking into account the measures of occurrence (sensitivity or specificity) together with the possibility of discrimination (positive [PPV] or negative predictive values [NPV]), and their qualitative grading were judged accordingly [18, 19] (Table 1). The Positive Clinical Utility Index (CU + Ve) is defined as sensitivity*PPV, and a high CU + Ve results should characterize "case finding" tests. By contrast, good Negative Clinical Utility Index (CU−Ve), defined as specificity*NPV, should be ideal for "screening" tests [18, 19].

In all analyses, a two-tailed *p* value < 0.05 was considered significant. GraphPad Prism (San Diego, CA, USA), Stata (version 13.0, College Station, Texas, USA) and R (version 3.6.2) were used for statistical analysis and rendering graphs.

Results

Clinical data

A total of 707 patients met enrolment criteria: 434 (61.3%) had been evaluated and treated at the center A, and 273 (38.7%) at the center B. Five-hundred and fifty six (78.6%) were males and 151 (21.4%) females, with a mean age of 61.8 years (range 18–91). Four-hundred and seventy-eight

Table 2 Clinical features of the cohort

Variables	All		Longitudinal vessels		Perpendicular vessels	
	n	%	*n*	%	*n*	%
All	707	100.0	283	40.0	424	60.0
Gender						
Male	556	78.6	184	26.0	372	52.6
Female	151	21.4	99	14.0	52	7.4
Previous treatments°						
No	478	67.6	221	31.3	257	36.4
Previous surgery	174	24.6	45	6.4	129	18.2
Previous RT	17	2.4	4	0.6	13	1.8
Previous biopsy	54	7.6	8	1.1	46	6.5
Histology						
Benign	208	29.4	184	26.0	24	3.4
Papillomatosis without atypia	34	4.8	0	0.0	34	4.8
SIN1	46	6.5	24	3.4	22	3.1
SIN2	61	8.6	26	3.7	35	5.0
SIN3/CIS	93	13.2	15	2.1	78	11.0
SCC	265	37.5	26	3.7	239	33.8

The sum of rows is 723 since 16 patients had two different previous treatments

RT radiotherapy, *SIN* squamous intraepithelial neoplasia, *CIS* in situ carcinoma, *SCC* squamous cell carcinoma
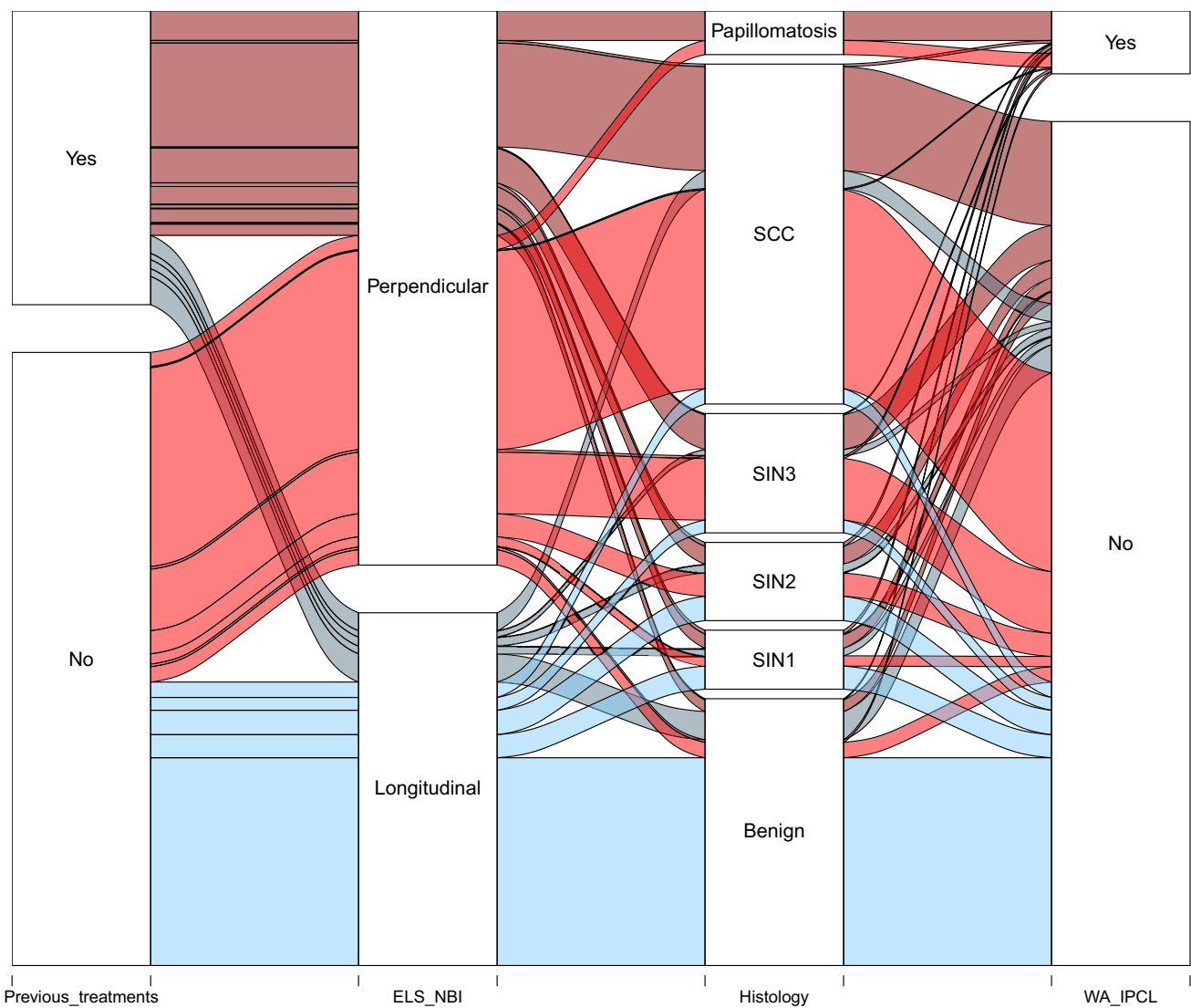
Fig. 4 Alluvial chart showing frequency distribution of previous treatment, ELS classification results, histology, and presence of wide angle IPCL features (WA IPCL). Color code according to different matching of previous treatments and ELS classification results

(67.6%) patients were submitted to endoscopic evaluation and surgical procedures without previous treatments, whereas 174 (24.6%) had been already surgically treated, 17 (2.4%) received head and neck radiotherapy, and 54 (7.6%) had been previously biopsied elsewhere. The final pathologic report was consistent with a benign lesion in 208

Table 3 Agreement analysis by Cohen's *k* test

Institution	Type of lesion	*N*	%	Agreement (%)	*κ*	95% CI (*κ*)	*p*
University of Genoa (Center A)	All	434	100	97.7	0.954	0.86–1.0	<0.0001
	Untreated	284	65.4	97.5	0.949	0.833–1.0	<0.0001
	Previous biopsy/surgery/RT	150	34.6	98.0	0.945	0.785–1.0	<0.0001
University of Brescia (Center B)	All	273	100	94.9	0.872	0.754–0.991	<0.0001
	Untreated	194	71.1	94.3	0.864	0.723–1.0	<0.0001
	Previous biopsy/surgery/RT	79	28.9	96.2	0.894	0.674–1.0	<0.0001

(29.5%) cases, papillomatosis without atypia in 34 (4.8%), mild SIN (SIN1) in 46 (6.5%), moderate SIN (SIN2) in 61 (8.6%), severe SIN (SIN3) or CIS in 93 (13.1%), and SCC in 265 (37.5%). Full details are summarized in Table 2, Fig. 4.

ELS classification interobserver reliability

Cohen's k statistic was used to assess the agreement between judgment of each lesion by two independent raters in each institution applying the ELS classification. According to the criteria by Altman et al. [16, 17], reported in Table 1, for the entire cohort the result was satisfactory and showed a very good agreement between observers at both the center A ($\kappa = 0.954$; 95% confidence interval [CI] 0.86–1.0, $p < 0.0001$) and the center B ($\kappa = 0.872$; 95% CI 0.754–0.991, $p < 0.0001$) (Table 3). The agreement was consistent and significant ($p < 0.0001$) for both Institutions, as well as for untreated ($\kappa = 0.949$; 95% CI 0.833–1.0 and $\kappa = 0.864$; 95% CI 0.723–1.0, respectively) and previously treated patients ($\kappa = 0.945$; 95% CI 0.785–1.0 and $\kappa = 0.894$; 95% CI 0.674–1.0, respectively) (Table 3).

Diagnostic performance

Considering the final score in the entire cohort (24 cases with initial disagreement were resolved between the examiners), performance of the diagnostic test was assessed investigating the detection of at least SIN1-SCC (Table 4). The best sensitivity and NPV were obtained for detection of SCC (0.90 and 0.91, respectively) and, accordingly, the best specificity and PPV for diagnosis of at least SIN1 (0.79

and 0.88, respectively). Considering previous treatments as a potential source of bias, for untreated patients the ELS classification reached the best performance with sensitivity and NPV for detection of SCC of 0.93 and 0.95, respectively, and specificity and PPV for diagnosis of at least SIN1 of 0.88 and 0.91, respectively. In previously treated patients, the performance of endoscopic evaluation was still satisfactory in terms of sensitivity (from 0.82 to 0.86), while it was poorer in terms of specificity (from 0.34 to 0.49), NPV (from 0.46 to 0.76), and PPV (from 0.45 to 0.84) (Table 4, Fig. 5).

The measurement of the CU indexes confirmed this observation with a good CU + Ve and CU−Ve for all outcomes except one in untreated patients, whereas no more than satisfactory or even poorer results were obtained for most outcomes in previously treated or biopsied patients, as shown in Table 5.

Diagnostic performance in respiratory papillomatosis

Among perpendicular vascular changes, the ELS classification well defines the vascular pattern of recurrent respiratory papillomatosis lesions, characterized by vessel loops with wide angle turning point, embedded in a three-dimensional warty structure [15]. We tested the identification of these features by NBI in our cohort, confirming their value for correct identification of this disease with a sensitivity of 1.0 (95% CI 0.90–1.0), specificity of 0.98 (95% CI 0.96–0.99), PPV of 0.69 (95% CI 0.55–0.82), and NPV of 1.0 (95% CI 0.99–1.0), as shown in Table 4. Moreover, the measure of CU indexes confirmed the excellent performance of NBI

Table 4 Diagnostic test results

	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
All				
At least SIN1	0.80 (0.77–0.84)	0.79 (0.74–0.84)	0.88 (0.85–0.91)	0.68 (0.62–0.73)
At least SIN2	0.84 (0.80–0.87)	0.75 (0.70–0.80)	0.83 (0.79–0.87)	0.76 (0.71–0.81)
At least SIN3	0.89 (0.85–0.92)	0.69 (0.64–0.74)	0.75 (0.70–0.79)	0.86 (0.81–0.89)
SCC	0.90 (0.86–0.94)	0.58 (0.53–0.63)	0.56 (0.52–0.61)	0.91 (0.87–0.94)
Untreated				
At least SIN1	0.80 (0.75–0.84)	0.88 (0.83–0.93)	0.91 (0.87–0.95)	0.74 (0.67–0.79)
At least SIN2	0.85 (0.80–0.89)	0.86 (0.81–0.90)	0.88 (0.84–0.92)	0.82 (0.76–0.87)
At least SIN3	0.90 (0.86–0.94)	0.81 (0.75–0.86)	0.81 (0.76–0.86)	0.90 (0.86–0.94)
SCC	0.93 (0.88–0.96)	0.68 (0.63–0.74)	0.61 (0.55–0.67)	0.95 (0.91–0.97)
Previous biopsy/surgery/RT				
At least SIN1	0.82 (0.75–0.87)	0.49 (0.35–0.63)	0.84 (0.77–0.89)	0.46 (0.33–0.59)
At least SIN2	0.83 (0.76–0.89)	0.44 (0.33–0.56)	0.75 (0.68–0.82)	0.56 (0.42–0.69)
At least SIN3	0.85 (0.78–0.91)	0.40 (0.31–0.51)	0.65 (0.58–0.72)	0.68 (0.54–0.79)
SCC	0.86 (0.77–0.92)	0.34 (0.26–0.43)	0.45 (0.41–0.57)	0.76 (0.63–0.86)
Papillomatosis detection	1.00 (0.90–1.00)	0.98 (0.96–0.99)	0.69 (0.55–0.82)	1.00 (0.99–1.00)

PPV positive predictive value, *NPV* negative predictive value, *CI* confidence interval, *SIN* squamous intraepithelial neoplasia, *SCC* squamous cell carcinoma
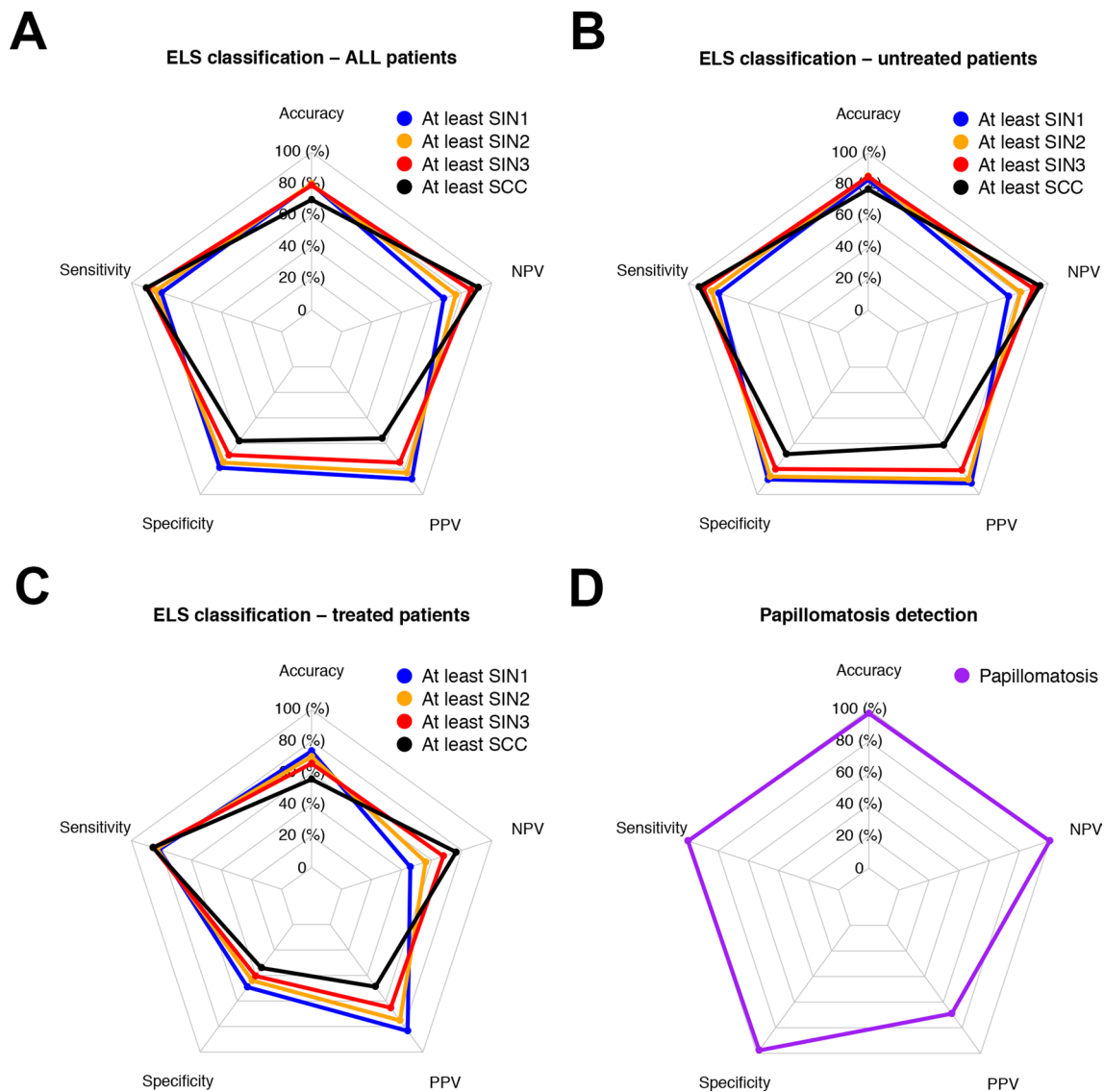
Fig. 5 Radar charts showing the diagnostic test applying the ELS classification for the detection of different histologic targets in the whole cohort (**a**), in the untreated group (**b**), and in the previously surgical or RT treated group (**c**). Diagnostic test results referred to the detection of wide angle IPCLs for the diagnosis of laryngeal papillomatosis (**d**)

with a CU–Ve of 0.98 (95% CI 0.97–0.99) and good performance in terms of CU + Ve of 0.69 (95% CI 0.56–0.83), as reported in Table 5.

Discussion

Among the several bioendoscopic techniques now available for routine evaluation of the UADT, NBI appears to be the most effective in evaluation of the larynx, hypopharynx, oral and oropharyngeal cavities [1, 20, 21]. The easy use of NBI and other bioendoscopic tools based on similar principles, which aims to enhance the vascular features of tissues (e.g. SPIES [22] or iSCAN [20]), is mainly due to

full integration of high-definition videoendoscopes, easily activated by pressing a button during in-office endoscopic examination or during pre- and intraoperative assessment. Interestingly, the superior in-depth evaluation of the bioendoscopic features of a given lesion may pave the way to the proof of concept of the optical biopsy, i.e. the capability to understand the nature of a given vocal fold mucosal lesion before its removal, thus modulating its excisional biopsy and optimizing hospitalization time, costs, and undue damage to surrounding healthy structures [23].

The need for a common language to categorize and share the findings from NBI evaluation led to a number of different classification systems. The first to have widespread diffusion in the head and neck scientific community was proposed

Table 5 Clinical Utility indexes and utility grading results according to Mitchell [18]

	CU+Ve (95% CI)	CU+Ve Judgment	CU−Ve (95% CI)	CU−Ve Judgment
All				
At least SIN1	0.71 (0.67–0.75)	Good	0.54 (0.49–0.58)	Satisfactory
At least SIN2	0.70 (0.65–0.74)	Good	0.57 (0.53–0.62)	Satisfactory
At least SIN3	0.66 (0.61–0.71)	Good	0.59 (0.55–0.63)	Satisfactory
SCC	0.51 (0.45–0.57)	Satisfactory	0.53 (0.49–0.57)	Satisfactory
Untreated				
At least SIN1	0.73 (0.68–0.78)	Good	0.65 (0.60–0.70)	Good
At least SIN2	0.75 (0.69–0.80)	Good	0.70 (0.66–0.74)	Good
At least SIN3	0.73 (0.68–0.79)	Good	0.73 (0.69–0.77)	Good
SCC	0.57 (0.50–0.64)	Satisfactory	0.65 (0.61–0.69)	Good
Previous biopsy/surgery/RT				
At least SIN1	0.68 (0.61–0.75)	Good	0.23 (0.11–0.34)	Poor
At least SIN2	0.63 (0.55–0.70)	Satisfactory	0.25 (0.14–0.35)	Poor
At least SIN3	0.56 (0.47–0.64)	Satisfactory	0.27 (0.17–0.38)	Poor
SCC	0.42 (0.32–0.52)	Poor	0.26 (0.16–0.36)	Poor
Papillomatosis detection	0.69 (0.56–0.83)	Good	0.98 (0.97–0.99)	Excellent

CU+Ve positive clinical utility index, *CU−Ve* negative clinical utility index, *CI* confidence interval, *SIN* squamous intraepithelial neoplasia, *SCC* squamous cell carcinoma

by Ni et al. [11]. These authors divided the different IPCL changes in five types (I–V), judging them as benign (from types I–IV), suspected malignant, and frankly malignant (type V). However, apart from its intrinsic complexity, this classification clearly showed a lack of a clear-cut threshold between benign and malignant diseases. In fact, different authors proposed different cut-offs for the worst endoscopic feature of each lesion to be considered suspicious, ranging from type III [24], to type IV [25–27], and type V [9, 11, 28].

Therefore, in 2016 the ELS proposed a new classification system for the interpretation of glottic vascular abnormalities detected during NBI-guided endoscopies [15]. This classification considers vascular abnormalities as IPCL perpendicular to the epithelium surface as suspicious, whereas longitudinal vascular changes (e.g. dilated or tortuous vessels, increased vessels numbers) are considered as not suspicious to harbor respiratory papillomatosis, pre-malignant, or cancerous lesions. The first attempt to apply this dichotomic classification was in the study by Šifrer et al. [29] who analyzed 80 vocal cords lesions in which the identification of a perpendicular vascular pattern was diagnostic for CIS-SCC with a sensitivity of 100%, specificity of 95%, PPV of 88%, and NPV of 100%. Further analysis evaluating a larger cohort of 288 vocal cords gave similar results (sensitivity 98%, specificity 95%, PPV 88%, and NPV 99%) [30].

Our results, obtained in two of the European pioneer centers applying NBI for evaluation of the UADT since 2007, herein confirm the intrinsic value of the ELS classification for laryngeal vascular changes in the identification of lesions harboring pre-cancerous or frankly neoplastic alterations. In particular, we applied this diagnostic tool to demonstrate its possible role in performing a so-called optical biopsy. In fact, our policy has always been, for early glottic lesions, a one-stage modulated excisional biopsy based on a number of pre- and intraoperative diagnostic tests in which WL and NBI rigid endoscopy under general anesthesia has always played a paramount role [23]. Moreover, as asserted by many authors, NBI is capable of enhancing small lesions that are undetectable by WL alone, thus ameliorating the treatment of laryngeal SSC, as well as assessing the potential multifocality of the disease and correct evaluation of intraoperative margins [31], as well as early identification of small recurrences during follow-up that may still allow application of minimally invasive treatments such as laser office-based procedures or second-look microlaryngoscopic operations [32–35]. Of note, the present study demonstrated a lower diagnostic accuracy of NBI in the previously treated patients compared to the untreated ones, thus confirming the potential confounding factor played by invasive sampling procedures when not directed to the full removal (excisional biopsy) of the entire visible lesion within safe margins.

The excellent interobserver reliability of the ELS classification with a $k > 0.81$ in all scenarios tested and reproducible in two independent centers confirms the reproducibility of the operators' findings in applying this classification tool. The high interobserver reliability of the ELS classification can be explained by its intrinsic simple application and dichotomic arrangement, providing better performance compared to other proposed classification systems such as that by Ni, which is complicated by a 5-tier structure and

associated with moderate/substantial interrater accordance, with a *k* ranging from 0.55 to 0.69 [36, 37].

On the other hand, it has to be noted that all the observers involved in this study had a minimum experience of 3 years in the use of NBI technology. Even though application of the ELS Classification on vocal fold vascular changes as observed by NBI is no more subjective than any other diagnostic performance, evaluation of certain subtle and sometimes ambiguous neoangiogenic patterns still may require a higher level of expertise, for which a learning curve is inevitably necessary. However, data derived from the gastrointestinal field show that less than a year of training evaluating 200 cases is enough to guarantee an accurate evaluation of NBI frames and that the motivation of the trainer itself can significantly improve the overall performance [38].

Investigating the diagnostic test, having as a target all the possible grades of pre-malignant or malignant transformation, allowed us to depict the capability of the ELS classification in helping to correctly identify pre-malignant cases with the highest PPV and specificity for at least SIN1 diagnosis. The lower performance of such parameters observed for the final diagnosis of glottic SCC can be explained by the presence, and progressively increase, of perpendicular vascular changes at early stages of pre-malignant transformation (SIN1-SIN2). By contrast, for diagnosis of laryngeal SCC, the ELS classification had good performance in terms of sensitivity and NPV, with a low rate of false negative cases and good confidence in a negative result (absence of perpendicular vascular changes).

Furthermore, several authors have underlined the utility of NBI for detection of recurrent respiratory papillomatosis and its ability to increase the detection rate of small lesions that invisible by WL alone [29, 33, 39–42]. The excellent performance in terms of CU + Ve and CU−Ve searching for wide angle IPCLs in the identification of respiratory papillomatosis mandates, as previously suggested by the recent literature [33, 39–44], the use of biologic endoscopy tools like NBI, and should be considered the endoscopic gold standard for optical biopsy and follow-up of patients affected by laryngeal papillomas.

The main limits of our study are represented by its retrospective design, balanced by analyzing a broad bicentric cohort. Nevertheless, among the estimator analyzed, the suboptimal performance in terms of specificity, negative predictive value, and CU−Ve could has been underestimated having chosen among the inclusive criteria the need for a histopathological diagnosis: several patients without any suspicious lesion at the first evaluation and along time could be considered as true negatives too, thus improving the values of such estimators.

Further developments in this field might include the analysis of a prospective cohort of patients, implementing the enrollment of true negative cases and developing a real-time

software applicable in the head and neck, based on artificial intelligence algorithms already tested on retrospective studies [45, 46], thus improving the objectivity and detection rate of these diagnostic tools, as already devised for gastrointestinal tract tumors [47, 48].

Conclusion

The ELS classification for NBI vascular changes of laryngeal lesions, herein validated in a large multicenter cohort, is a highly reliable tool with good diagnostic performance in the optical biopsy setting, confirming its overall value. The systematic use of this classification seems to allow better (and purely endoscopic) diagnostic capability of suspicious glottic lesions, reliably distinguishing benign ones from those with a diagnosis of papillomatosis, SIN, or invasive SCC. The excellent performance of NBI for correct identification of respiratory papillomatosis also confirms its usefulness in this clinical setting.

Declarations

References

1. Sun C, Han X, Li X et al (2017) Diagnostic performance of narrow band imaging for laryngeal cancer: a systematic review and meta-analysis. Otolaryngol Head Neck Surg 156:589–597. https://doi.org/10.1177/0194599816685701

2. Watanabe A, Taniguchi M, Tsujie H et al (2009) The value of narrow band imaging for early detection of laryngeal cancer. Eur Arch Otorhinolaryngol. https://doi.org/10.1007/s00405-008-0835-1

3. Chu EA, Kim YJ (2008) Laryngeal cancer: diagnosis and preoperative work-up. OtolaryngolClin North Am 41:673–695

4. Green B, Cobb ARM, Brennan PA, Hopper C (2014) Optical diagnostic techniques for use in lesions of the head and neck: review of the latest developments. Br J Oral MaxillofacSurg 52:675–680. https://doi.org/10.1016/j.bjoms.2014.06.010

5. Muto M, Nakane M, Katada C et al (2004) Squamous cell carcinoma in situ at oropharyngeal and hypopharyngeal mucosal sites. Cancer 101:1375–1381. https://doi.org/10.1002/cncr.20482

6. Piazza C, Dessouky O, Peretti G et al (2008) Narrow-band imaging: a new tool for evaluation of head and neck squamous cell carcinomas. Review of the literature. ActaOtorhinolaryngolItalorganoUffdellaSocItal di Otorinolaringol e ChirCerv-facc 28:49–54

7. Piazza C, Del Bon F, Peretti G, Nicolai P (2012) Narrow band imaging in endoscopic evaluation of the larynx. CurrOpinOtolaryngol Head Neck Surg 20:472–476

8. Cohen J (2008) Comprehensive atlas of high resolution endoscopy and narrowband imaging (1st edn)

9. Kraft M, Fostiropoulos K, Gurtler N et al (2016) Value of narrow band imaging in the early diagnosis of laryngeal cancer. Head Neck 38:15–20. https://doi.org/10.1002/hed.23838

10. Lin YC, Wang WH, Lee KF et al (2012) Value of narrow band imaging endoscopy in early mucosal head and neck cancer. Head Neck 34:1574–1579

11. Ni X-G, He S, Xu Z-G et al (2011) Endoscopic diagnosis of laryngeal cancer and precancerous lesions by narrow band imaging. J LaryngolOtol. https://doi.org/10.1017/S0022215110002033

12. Takano JH, Yakushiji T, Kamiyama I et al (2010) Detecting early oral cancer: narrowband imaging system observation of the oral mucosa microvasculature. Int J Oral MaxillofacSurg 39:208–213. https://doi.org/10.1016/j.ijom.2010.01.007

13. Barnes L, Eveson JW, Reichart P, Sidransky D (2005) World Health Organization Classification of Head and Neck Tumours. WHO Classif Tumour

14. Remacle M, Eckel HE, Antonelli A et al (2000) Endoscopic cordectomy. A proposal for a classification by the working committee, European Laryngological Society. Eur Arch Otorhinolaryngol 257:227–231. https://doi.org/10.1007/s004050050228

15. Arens C, Piazza C, Andrea M et al (2016) Proposal for a descriptive guideline of vascular changes in lesions of the vocal folds by the committee on endoscopic laryngeal imaging of the European Laryngological Society. Eur Arch Otorhinolaryngol 273:1207–1214. https://doi.org/10.1007/s00405-015-3851-y

16. Altman DG (1991) Practical statistics for medical research. Chapman and Hall. Stat Med. https://doi.org/10.1002/sim.4780101015

17. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics. https://doi.org/10.2307/2529310

18. Mitchell AJ (2008) The clinical significance of subjective memory complaints in the diagnosis of mild cognitive impairment and dementia: a meta-analysis. Int J Geriatr Psychiatry 23:1191–1202. https://doi.org/10.1002/gps.2053

19. Mitchell AJ (2011) Sensitivity x PPV is a recognized test called the clinical utility index (CUI+). Eur J Epidemiol 26:251–252. https://doi.org/10.1007/s10654-011-9561-x

20. Hawkshaw MJ, Sataloff JB, Sataloff RT (2013) New concepts in vocal fold imaging: a review. J Voice 27:738–743. https://doi.org/10.1016/j.jvoice.2013.05.011

21. Deganello A, Paderno A, Morello R et al (2020) Diagnostic accuracy of narrow band imaging in patients with oral lichen planus: a prospective study. Laryngoscope. https://doi.org/10.1002/lary.29035

22. Staníková L, Walderová R, Jančatová D et al (2018) Comparison of narrow band imaging and the Storz Professional Image Enhancement System for detection of laryngeal and hypopharyngeal pathologies. Eur Arch Otorhinolaryngol 275:1819–1825. https://doi.org/10.1007/s00405-018-4987-3

23. Mora F, Carta F, Missale F et al (2020) Laryngeal mid-cord erythroleukoplakias: how to modulate the transoral$CO(2)$ laser excisional biopsy. Cancers (Basel). https://doi.org/10.3390/cancers12082165

24. Rzepakowska A, Sielska-Badurek E, Cruz R et al (2018) Narrow band imaging versus laryngovideostroboscopy in precancerous and malignant vocal fold lesions. Head Neck 40:927–936. https://doi.org/10.1002/hed.25047

25. Ni XG, Zhu JQ, Zhang QQ et al (2019) Diagnosis of vocal cord leukoplakia: the role of a novel narrow band imaging endoscopic classification. Laryngoscope 129:429–434. https://doi.org/10.1002/lary.27346

26. Vilaseca I, Valls-Mateus M, Nogués A et al (2017) Usefulness of office examination with narrow band imaging for the diagnosis of head and neck squamous cell carcinoma and follow-up of premalignant lesions. Head Neck 39:1854–1863. https://doi.org/10.1002/hed.24849

27. De Vito A, Meccariello G, Vicini C (2017) Narrow band imaging as screening test for early detection of laryngeal cancer: a prospective study. ClinOtolaryngol 42:347–353. https://doi.org/10.1111/coa.12728

28. Rzepakowska A, Sielska-Badurek E, Żurek M et al (2018) Narrow band imaging for risk stratification of glottic cancer within leukoplakia. Head Neck 40:2149–2154. https://doi.org/10.1002/hed.25201

29. Šifrer R, Rijken JA, Leemans CR et al (2018) Evaluation of vascular features of vocal cords proposed by the European Laryngological Society. Eur Arch Otorhinolaryngol 275:147–151. https://doi.org/10.1007/s00405-017-4791-5

30. Šifrer R, Šereg-Bahar M, Gale N, Hočevar-Boltežar I (2020) The diagnostic value of perpendicular vascular patterns of vocal cords defined by narrow-band imaging. Eur Arch Otorhinolaryngol. https://doi.org/10.1007/s00405-020-05864-5

31. Fiz I, Mazzola F, Fiz F et al (2017) Impact of close and positive margins in transoral laser microsurgery for Tis-T2 glottic cancer. Front Oncol 7:245. https://doi.org/10.3389/fonc.2017.00245

32. Ivancic R, Iqbal H, deSilva B et al (2018) Current and future management of recurrent respiratory papillomatosis. Laryngoscope InvestigOtolaryngol 3:22–34. https://doi.org/10.1002/lio2.132

33. Ochsner MC, Klein AM (2015) The utility of narrow band imaging in the treatment of laryngeal papillomatosis in awake patients. J Voice 29:349–351. https://doi.org/10.1016/j.jvoice.2014.08.002

34. Zeitels SM, Burns JA, Franco RA et al (2004) Office-based treatment of glottal dysplasia and papillomatosis with the 585-nm pulsed dye laser and local anesthesia. Ann OtolRhinolLaryngol 113:265–276. https://doi.org/10.1177/000348940411300403

35. Motz KM, Hillel AT (2016) Office-based management of recurrent respiratory papilloma. CurrOtorhinolaryngol Rep 4:90–98. https://doi.org/10.1007/s40136-016-0118-0

36. Nogués-Sabaté A, Aviles-Jurado FX, Ruiz-Sevilla L et al (2018) Intra and interobserver agreement of narrow band imaging for the detection of head and neck tumors. Eur Arch Otorhinolaryngol 275:2349–2354. https://doi.org/10.1007/s00405-018-5063-8

37. Zwakenberg MA, Dikkers FG, Wedman J et al (2016) Narrow band imaging improves observer reliability in evaluation of upper aerodigestive tract lesions. Laryngoscope 126:2276–2281. https://doi.org/10.1002/lary.26008

38. Dias-Silva D, Pimentel-Nunes P, Magalhães J et al (2014) The learning curve for narrow-band imaging in the diagnosis of precancerous gastric lesions by using Web-based video. GastrointestEndosc 79:910–920. https://doi.org/10.1016/j.gie.2013.10.020 (**quiz 983-e1, 983.e4**)

39. TjonPianGi REA, Halmos GB, Van Hemel BM et al (2012) Narrow band imaging is a new technique in visualization of recurrent respiratory papillomatosis. Laryngoscope 122:1826–1830. https://doi.org/10.1002/lary.23344

40. Lukes P, Zabrodsky M, Lukesova E et al (2014) The role of NBI HDTV magnifying endoscopy in the prehistologic diagnosis of laryngeal papillomatosis and spinocellular cancer. Biomed Res Int. https://doi.org/10.1155/2014/285486

41. Dippold S, Becker C, Nusseck M et al (2015) Narrow band imaging: a tool for endoscopic examination of patients with laryngeal papillomatosis. Ann OtolRhinolLaryngol 124:886–892. https://doi.org/10.1177/0003489415590656

42. Jackowska J, Klimza H, Winiarski P et al (2018) The usefulness of narrow band imaging in the assessment of laryngeal papillomatosis. PLoS ONE 13:1–9. https://doi.org/10.1371/journal.pone.0205554

43. Imaizumi M, Okano W, Tada Y, Omori K (2012) Surgical treatment of laryngeal papillomatosis using narrow band imaging. Otolaryngol Head Neck Surg (United States) 147:522–524. https://doi.org/10.1177/0194599812448162

44. Adachi K, Umezaki T, Kiyohara H, Komune S (2015) New technique for laryngomicrosurgery: narrow band imaging-assisted video-laryngomicrosurgery for laryngeal papillomatosis. J LaryngolOtol 129:S74–S76. https://doi.org/10.1017/S0022215114002436

45. Moccia S, De Momi E, Guarnaschelli M et al (2017) Confident texture-based laryngeal tissue classification for early stage diagnosis support. J Med Imaging (Bellingham, Wash) 4:34502. https://doi.org/10.1117/1.JMI.4.3.034502

46. Moccia S, Vanone GO, De ME et al (2018) Learning-based classification of informative laryngoscopic frames. Comput Methods Programs Biomed 158:21–30. https://doi.org/10.1016/j.cmpb.2018.01.030

47. Mori Y, Kudo S-E, Misawa M et al (2018) Real-time use of artificial intelligence in identification of diminutive polyps during colonoscopy: a prospective study. Ann Intern Med 169:357–366. https://doi.org/10.7326/M18-0249

48. Guo L, Xiao X, Wu C et al (2020) Real-time automated diagnosis of precancerous lesions and early esophageal squamous cell carcinoma using a deep learning model (with videos). GastrointestEndosc 91:41–51. https://doi.org/10.1016/j.gie.2019.08.018