

CHEST[®]

Official publication of the American College of Chest Physicians

ONLINE FIRST

This is an Online First, unedited version of this article. The final, edited version will appear in a numbered issue of *CHEST* and may contain substantive changes. We encourage readers to check back for the final article. Online First papers are indexed in PubMed and by search engines, but the information, including the final title and author list, may be updated on final publication.

<http://journal.publications.chestnet.org/>



Online First articles are not copyedited prior to posting.

©American College of Chest Physicians.

Reproduction of this article is prohibited without written permission from the American College of Chest Physicians. See online for more details.

1 Word Count: body = 2,297 abstract = 218

2 An index to objectively score supraglottic abnormalities in refractory asthma:

3 Learning, validation, and significance

4 By

5 James T. Good Jr., M.D.¹

6 Donald R. Rollins, M.D.¹

7 Douglas Curran-Everett, Ph.D.²

8 Steven E. Lommatzsch, M.D.¹

9 Brendan J. Carolan, M.D.¹

10 Peter C. Stubenrauch, M.D.¹

11 Richard J. Martin, M.D.¹

12

13 From: Departments of Medicine, Division of Pulmonary, Critical Care, and Sleep

14 Medicine¹; and Biostatistics and Bioinformatics², National Jewish Health, Denver,

15 CO.

16 Correspondence to:

17 Richard J. Martin, M.D.

18 National Jewish Health

19 1400 Jackson Street

20 Denver, CO 80206

21 martinr@njhealth.org

22

23

24 The authors have reported to *CHEST* the following conflicts of interest:
25 Dr. Good has performed speaking activities for Merck and Genentech and
26 received a research grant from MedImmune. Dr. Martin has done consultancy
27 work and/or received travel support and/or honoraria for attendance at advisory
28 boards for Teva, AstraZeneca, MedImmune, and Merck; received research
29 grants from MedImmune and the NHLBI; received royalties from UpToDate.
30 Dr. Lommatzsch has performed speaking activities for Novartis. Dr. Carolan has
31 performed speaking activities for Novartis. Dr. Rollins, Dr. Curran-Everett, and
32 Dr. Peter Stubenrauch have reported to *CHEST* that they have no potential
33 conflicts of interest.
34

Abstract

Background

Refractory asthma patients frequently have elements of laryngopharyngeal reflux (LPR) with potential aspiration contributing to their poor control. We previously reported on a supraglottic index (SGI) scoring system that helps in the evaluation of LPR with potential aspiration. However, to further the usefulness of this SGI scoring system for bronchoscopists, a teaching system was developed that included both inter- and intra-observer reproducibility.

Methods

Five pulmonologists with expertise in fiberoptic bronchoscopy but novice to the SGI participated. A training system was developed that could be used via internet interaction in order to make this learning technique widely available.

Results

By the final testing, there was excellent inter-reader agreement (kappa of at least 0.81), thus documenting reproducibility in scoring the SGI. For the measure of intra-reader consistency, one reader was arbitrarily selected to rescore the final test four weeks later and had a kappa value of 0.93 with a 95% confidence interval of 0.79-1.00.

Conclusions

64 In this study, we demonstrate that with an organized educational approach,
65 bronchoscopists can develop skills to have highly reproducible assessment and
66 scoring of supraglottic abnormalities. The SGI can be used to determine which
67 patients need additional intervention to determine causes of LPR and
68 gastroesophageal reflux. Identification of this problem in patients with refractory
69 asthma allows for personal, individual directed therapy to improve asthma
70 control.

71

- 35 Abbreviation List
- 36 GER = gastroesophageal reflux
- 37 LPR = Laryngopharyngeal reflux
- 38 RFS = reflux finding score
- 39 SGI = supraglottic index
- 40

Introduction

Refractory¹ or severe asthma has a high associated morbidity and economic cost.^{2,3} Even with asthma guideline therapy⁴, up to 50% of patients are not well controlled or refractory to treatment.⁵ Fiberoptic bronchoscopy has been shown to be useful in evaluating both the upper and lower airways of refractory asthmatics, which allows for better phenotyping of these individuals and thus specific patient-oriented therapy.⁶ This approach has led to improved outcomes.⁶

Laryngopharyngeal reflux (LPR) can occur in association with gastroesophageal reflux (GER) or independent from GER. LPR can be injurious to the supraglottic area and lower airway with potential aspiration. To better evaluate potential supraglottic injury resulting from LPR, a supraglottic index (SGI) was developed and used to give objective, applicable and reproducible data for use in asthmatic patients.⁶ To keep consistency in that study, one individual scored the SGI. However, for clinical applicability, a teaching method for bronchoscopists to learn the SGI scoring technique was developed in concert with documentation that inter- and intra- observer reproducibility was valid. This subsequent study demonstrates these points.

Methods

National Jewish Health's Institutional Review Board approval (HS2477 and HS2639) was obtained to use these prospective clinical data for publication.

The SGI is a numeric scoring system that allows for an objective grading system to quantify the amount of supraglottic abnormalities present. The purpose of this study was to determine if the 5 individual scorers could have reproducible results grading the same photos as the control reader.

SGI

Three supraglottic structures (epiglottis, false cords, and arytenoids) are scored for the amount of edema (0 = none, 1 = mild, 2 = moderate, 3 = severe) that is present. In addition, these structures are evaluated for the amount of erythema or hyperemia by using the same numeric scale from 0 to 3. The posterior commissure (inter-arytenoid area) and the piriform recesses (piriform sinuses) are scored as normal = 0 or abnormal = 2. The score for the SGI ranges from 0 to 22 (Table 1 scoring sheet).

Teaching Program

Five pulmonologists with expertise in fiberoptic bronchoscopy, but novice to the SGI, were recruited for this study. All five were trained in bronchoscopy using ABIM guidelines and are board certified in pulmonary medicine. Three

116 pulmonologists recently completed their pulmonary fellowships. One of the two
117 senior pulmonologists was recruited from private practice; the other has a 35-
118 year academic career. We felt that learning this technique would require multiple
119 sessions before inter- and intra-observer validity would be meaningful. Prior to
120 the first scoring session, a one-hour training lecture with the use of photographs
121 and video recordings occurred. During this session the supraglottic index was
122 described in detail and each of the anatomical areas to be graded were
123 reviewed. Examples of normal, mild, moderate and severe edema and
124 erythema/hyperemia for the epiglottis, false cords and arytenoids were provided.
125 Photographs of normal and abnormal posterior commissures and piriform
126 recesses were also reviewed. Five cases were provided and each physician
127 graded them and compared their scores with the control reader.

128

129 Scoring session 1 consisted of 50 patients with 4 photographs each of the
130 supraglottic areas; these were provided to the readers on a flash drive. Readers
131 had the ability to adjust the photographic size using the zoom application on their
132 computer. Figure 1 demonstrates an example of photographs used in grading the
133 SGI. In this patient, the epiglottis received a score of 1 for edema and 1 for
134 hyperemia; the false cords: 2 for edema and 3 for hyperemia; the arytenoids: 3
135 for edema and 3 for erythema; posterior commissure = 2; and piriform recess =
136 2. The sum of these makes an SGI score of 17. It should be noted that the true
137 cords can be involved in refractory asthmatics, but this is not part of our scoring

system. After session 1 scoring, the readers were given the control reader's calculations so scores could be compared and discussed.

Session 2 consisted of the photos used in session 1 being randomly shuffled and the physicians again calculated the SGI for 50 patients. Scores were reviewed and readers again had reference to the control reader's SGI values.

In session 3, a new set of 50 different patients' supraglottic photographs were reviewed. These scores were compared to the control score. Readers were asked to rescore those SGI calculations that varied by four or more from the control reader. In addition, readers were asked to rescore if they had an $SGI \geq 10$ and control < 10 , or < 10 and control > 10 . The reason for using the SGI cut point of 10 is an $SGI \geq 10$ has been shown in a previous study to be a threshold value for the presence of LPR.⁶

Session 4 consisted of the photographs in session 3 being randomly shuffled and rescored. Individual scores were discussed and reviewed by the control reader with the other readers. Special attention was given to scores that varied by four or more points from the control reader and those discordant with the control SGI value of 10.

Inter- and intra-reader validity

After the above described sessions, 30 new patient photographs of the supraglottic area were distributed to the five learners. The distribution of SGI scores were 0-9 (9 patients), 10-16 (13 patients), and 17-22 (8 patients).

We assessed the relationship between scores for each novice reader and the control reader using a form of regression that can account for measurement error in the X variable (the scores of the control reader).^{7,8} We conducted these analyses using R.⁹ Using an SGI cutoff of 10 units,⁶ we estimated kappas between each new reader and the control reader (SAS/STAT software package, version 9.2 of the SAS System for Windows XP). We defined the critical significance level α to be 0.05.

Results

After session 1, there was expected variability among all naïve readers compared to the control reader (regression $R^2 = 51\%$, range 35-65%). Figure 2a demonstrates comparison of the naïve readers versus control of session 1 for SGI scoring. In the final scoring session, there was a marked decrease in variability between the naïve readers and control reader (Regression $R^2 = 86\%$, range 75-92%) (Figure 2b).

Table 2 demonstrates the estimate of the amount of fixed (intercept) and proportional (slope) bias at the final testing session. A perfect SGI scoring between each naïve reader and control reader would be represented by an

183 intercept of zero and a slope of 1. If the confidence interval for the intercept
184 includes zero, the intercept is consistent with zero. If the confidence interval for
185 the slope includes 1, then the intercept is consistent with 1. For naïve readers C-
186 E, both the fixed and proportional biases are tightly linked to the control reader. A
187 and B are acceptably linked to the control reader with A overall reading slightly
188 under the reader by 3.5 units and B slightly over by 2 units.

189
190 Additionally, for the final scoring for $\text{SGI} < 10$ and ≥ 10 , the inter-reader
191 agreement of the SGI with the control reader is exceptional (defined as kappa of
192 at least 0.81).¹⁰ Table 3 demonstrates agreement with kappa between 0.83 and
193 0.92. Furthermore, the five kappa values are statistically different from zero ($p <$
194 0.001), but similar to each other ($p = 0.94$).

195
196 For the measure of intra-reader consistency, Reader E was arbitrarily selected to
197 re-score the final test four weeks later. The kappa between reader E's two
198 scoring tests was 0.93 with a 95% confidence interval of 0.79-1.00.

201 Discussion

202 The purpose of this study was to develop a teaching technique that results in
203 reproducibility for evaluation and scoring of abnormalities in the supraglottic area.
204 Originally we were interested in phenotyping refractory asthma by bronchoscopic
205 evaluation of the upper and lower airways so as to be able to direct personalized

206 therapy to this difficult-to-treat group of patients.⁶ One of the largest phenotypes
207 was represented by those with LPR (perhaps incorrectly termed
208 gastroesophageal reflux [GER] in that study). We developed the supraglottic
209 index to help in the phenotyping of these refractory asthma patients as to
210 supraglottic abnormalities that would be consistent with LPR.⁶ Since classic GER
211 may be intermittent, the standard GER tests may give false positive or false
212 negative results. Additionally, GER may occur but it is not always associated with
213 LPR and thus no potential aspiration develops. The supraglottic index gives a
214 history of what occurs with intermittent or continuous irritation of this area over
215 time. Thus, we feel that LPR with potential aspiration may be better evaluated by
216 the SGI than standard GER studies.

217

218 There are several methods and scoring systems that have been used to define
219 and quantify the degree of upper airway abnormalities.^{11,12} While there are many
220 studies describing supraglottic abnormalities, an accurate, simple and
221 reproducible method to access and score supraglottic abnormalities, particularly
222 in asthma, had not existed.

223

224 In Belafsky and colleague's general reflux finding score (RFS),¹¹ eight
225 areas/items were evaluated with a variable grading scale from 0 to 4. These
226 included subglottic edema, ventricular edema, erythema-hyperemia, vocal fold
227 edema, diffuse laryngeal edema, posterior commissure hypertrophy,
228 granuloma/granulation tissue and thick endolaryngeal mucus. The RFS was

determined by two laryngologists in 40 patients with abnormal distal or proximal pH probe studies. All study patients were prospectively evaluated before treatment and at 2, 4, and 6 months after proton pump inhibitor treatment. There was an improved RFS with antireflux therapy and inter- and intra-observer reproducibility between the two observers.

Vavricka and colleague's scoring system¹² evaluated 10 supraglottic regions: posterior pharyngeal wall, interarytenoid bar, posterior commissure, posterior cricoid wall, arytenoid complex, true vocal cords, false vocal folds, anterior commissure, epiglottis, and aryepiglottic fold. In this study there was a poor correlation between abnormalities of the laryngopharyngeal area and documented esophageal reflux defined by esophageal endoscopic abnormalities. The authors questioned the diagnostic specificity of abnormal laryngopharyngeal findings attributed to reflux.

We previously noted in 58 patients with refractory asthma that 44 (78%) had an SGI ≥ 10 .⁶ Forty-three of these had GER testing and 34 had documented GER. Fourteen patients had an SGI < 10 with nine having GER testing, eight of whom had negative results. The SGI was significantly higher in the patients with GER (15.8 [SD 3.6]) than those without GER (8.9 [SD 5.5], $P < .0001$). Additionally, two patients with an SGI > 10 and negative GER studies were placed in the non-specific phenotype as they did not meet the study criteria (including GI reflux studies) for any phenotypes. Several months post-study, they returned still not

improved. Both now had a positive GER study. Directed reflux treatment improved their asthma control and lung function. Because of the strong correlation between the SGI and GER we thought this would be an excellent diagnostic tool to identify patients with LPR contribution to their pulmonary problems while avoiding potential false negatives or positives from GI studies.

Study limitations include the limited number of “naïve” SGI pulmonologists (all faculty from National Jewish Health) that represented the learning group. However, they did represent junior and senior pulmonologists, as well as academic and recent private practice physicians. Additional studies are needed to evaluate non-pulmonologists with expertise in laryngoscopy (both physicians and other healthcare providers) as to the success of this educational program in those populations. It should be noted that the SGI has only been studied in patients with refractory asthma and its use in diagnosing LPR has not been demonstrated in other conditions. Other limitations are not including additional supraglottic and subglottic areas in the scoring. However, we wanted to produce a relatively simple and functional scoring system.

The purpose of this current study was to demonstrate the SGI’s teaching technique and reproducibility utilizing five naïve-SGI-scoring physicians who perform bronchoscopy. The teaching technique was prospectively developed, but modified between sessions with feedback from the naïve readers. Modifications were made mainly to focus discussion on areas of “over-reading” and “under-

reading.” This simple focus change improved understanding of the scoring system. From the initial to final SGI scoring there was marked overall improvement in variability between the naïve and control readers and, in particular, the important SGI cut point of 10. The intra-observer consistency was also very tight as demonstrated by reader E.

The importance of LPR contributing to a multitude of pulmonary conditions is significant. Most previous publications have focused on classic GER, an esophageal disorder, not a pulmonary condition. The criteria used to establish a diagnosis of GER does not include the multiple potential causes for pulmonary complications related to LPR with potential aspiration. One aspiration event can produce lung irritation and subsequent inflammation. Recurrent LPR with aspiration may manifest itself with many different pulmonary problems such as refractory asthma,⁶ ground glass infiltrates,¹³ interstitial lung disease,¹⁴ bronchiectasis,^{15,16} and aspiration pneumonitis.¹⁷

While it is important to utilize well-established gastrointestinal studies to document and elucidate specific pathophysiologic problems, false positive and false negatives may occur. Dual channel pH-impedance probes, Bravo studies, modified and regular barium swallows, and esophageal manometry evaluate potential abnormalities at one point in time. In contrast careful endoscopy of the supraglottic area and calculating a supraglottic index allows the bronchoscopist

297 to have an anatomical picture of events that have occurred over time, not just a
298 single snapshot or a recording of isolated events.

299

300 The challenge that remains is training and educating caregivers to accurately
301 assess the supraglottic area with a tool that is not too cumbersome. In this study
302 it took 4 separate reading sessions of different still and video recordings of
303 multiple different patients for the naïve readers to have consistent and
304 reproducible results compared with the control reader. Although an initial labor-
305 intensive period is needed for any naïve reader to become proficient in the SGI
306 scoring system, the effort pays off in better classification and thus treatment of
307 refractory asthma patients. We now have an extensive online educational
308 teaching program for learning to score the SGI.¹⁸

309

310 Acknowledgements

311 **Author contributions:** Dr. Martin: had input into study design, data collection,
312 analyses, and manuscript writing. He takes responsibility for the integrity of the
313 data and the accuracy of the data analyses.

314 Dr. Good: had input into study design, data collection, analyses, and manuscript
315 writing.

316 Dr. Rollins: had input into study design, data collection, analyses, and manuscript
317 writing.

318 Dr. Curran-Everett: had input into study design, data collection, analyses, and
319 manuscript writing.

320 Dr. Carolan: had input into study design, data collection, analyses, and
321 manuscript writing.

322 Dr. Stubenrauch: had input into study design, data collection, analyses, and
323 manuscript writing.

324

325 **Other contributions:** The authors thank Elizabeth Kellermeyer and Sarah
326 Murrell for manuscript transcription and editing.

327

References

1. Proceedings of the ATS workshop on refractory asthma: current understanding, recommendations, and unanswered questions. American Thoracic Society. *Am J Respir Crit Care Med*. 2000;162(6):2341-2351.
2. Mannino DM, Homa DM, Akinbami LJ, Moorman JE, Gwynn C, Redd SC. Surveillance for asthma--United States, 1980-1999. *MMWR Surveill Summ*. 2002;51(1):1-13.
3. National Heart L, and Blood Institute. Morbidity and Mortality: 2002 Chart Book on Cardiovascular, Lung, and Blood Diseases. National Institutes of Health. Bethesda, MD, 2002
4. National Asthma Education and Prevention Program (National Heart Lung and Blood Institute) Third Expert Panel on the Management of Asthma., National Center for Biotechnology Information (U.S.). Expert Panel report 3 guidelines for the diagnosis and management of asthma. NIH publication no. 07-4051. Bethesda, Md.: National Institutes of Health National Heart Lung and Blood Institute, 2007
5. Bateman ED, Boushey HA, Bousquet J, et al. Can guideline-defined asthma control be achieved? The Gaining Optimal Asthma Control study. *Am J Respir Crit Care Med*. 2004;170(8):836-844.
6. Good JT, Jr., Kolakowski CA, Groshong SD, Murphy JR, Martin RJ. Refractory asthma: importance of bronchoscopy to identify phenotypes and direct therapy. *Chest*. 2012;141(3):599-606.

- 350 7. Curran-Everett D. Explorations in statistics: regression. *Adv Physiol Educ.*
351 2011;35(4):347-352.
- 352 8. Ludbrook J. Linear regression analysis for comparing two measurers or
353 methods of measurement: but which regression? *Clin Exp Pharmacol*
354 *Physiol.* 2010;37(7):692-699.
- 355 9. Team RDC. R: A language and environment for statistical computing. *R*
356 *Foundation for Statistical Computing, Vienna, Austria.* 2012
- 357 10. Altman D. Practical Statistics for Medical Research. New York: Chapman and
358 Hall, 1991; 404
- 359 11. Belafsky PC, Postma GN, Koufman JA. The validity and reliability of the
360 reflux finding score (RFS). *Laryngoscope.* 2001;111(8):1313-1317.
- 361 12. Vavricka SR, Storck CA, Wildi SM, et al. Limited diagnostic value of
362 laryngopharyngeal lesions in patients with gastroesophageal reflux during
363 routine upper gastrointestinal endoscopy. *The American journal of*
364 *gastroenterology.* 2007;102(4):716-722.
- 365 13. Naidich DP, Zerhouni EA, Hutchins GM, Genieser NB, McCauley DI,
366 Siegelman SS. Computed tomography of the pulmonary parenchyma. Part
367 1: Distal air-space disease. *Journal of thoracic imaging.* 1985;1(1):39-53.
- 368 14. Tobin RW, Pope CE, 2nd, Pellegrini CA, Emond MJ, Sillery J, Raghu G.
369 Increased prevalence of gastroesophageal reflux in patients with
370 idiopathic pulmonary fibrosis. *American journal of respiratory and critical*
371 *care medicine.* 1998;158(6):1804-1808.

- 372 15. Moulton BC, Barker AF. Pathogenesis of bronchiectasis. *Clinics in chest*
373 *medicine*. 2012;33(2):211-217.
- 374 16. Ilowite J, Spiegler P, Chawla S. Bronchiectasis: new findings in the
375 pathogenesis and treatment of this disease. *Current opinion in infectious*
376 *diseases*. 2008;21(2):163-167.
- 377 17. Field SK, Field TS, Cowie RL. Extraesophageal manifestations of
378 gastroesophageal reflux. *Minerva gastroenterologica e dietologica*.
379 2001;47(3):137-150.
- 380 18. Supraglottic Index Learning Program. *National Jewish Health*.
381 <http://www.njhealth.org/SGL>. Accessed: September 2013.

382

383

384

385

Table 1: Supraglottic Index (SGI) Scoring Sheet

	0 = Normal 1 = Mild	2 = Moderate 3 = Severe
	Edema	Erythema/Hyperemia
Epiglottis		
False Cords		
Arytenoid Cartilage		

	0 = Normal 2 = Abnormal
Piriform Recess – pooled secretions/mucosal abnormalities	
Posterior Commissure (Interarytenoid area) – scarring/tissue hypertrophy	

Supraglottic Index (SGI) _____

395 Table 2. Relationships between scores from 5 naïve readers and the control
396 scores
397
398

Reader	Intercept*	Slope†
A	−3.5 −5.6 to −1.7	1.2 1.1 to 1.4
B	2.1 0.6 to 3.4	1.0 0.9 to 1.1
C	0.6 −2.0 to 2.7	1.0 0.8 to 1.2
D	−0.3 −2.2 to 1.3	1.0 0.8 to 1.1
E	−0.8 −2.5 to 0.6	1.0 0.9–1.2

399
400 *The intercept represents the amount of fixed bias, the naïve reader score when
401 the control score is zero.

402 †The slope represents the amount of proportional bias, the change in naïve
403 reader score when the control score changes by 1 unit.

404
405

406 Table 3. Agreement of naïve readers with control readers for SGI score < 10 and
407 ≥ 10 .

408

Reader	Kappa	95% Conf Int
A	0.92	0.78–1.00
B	0.83	0.61–1.00
C	0.83	0.61–1.00
D	0.85	0.65–1.00
E	0.85	0.65–1.00

409

410 A kappa value of ≥ 0.81 denotes that the strength of agreement is very good.¹⁰

411

412 Figure Legends

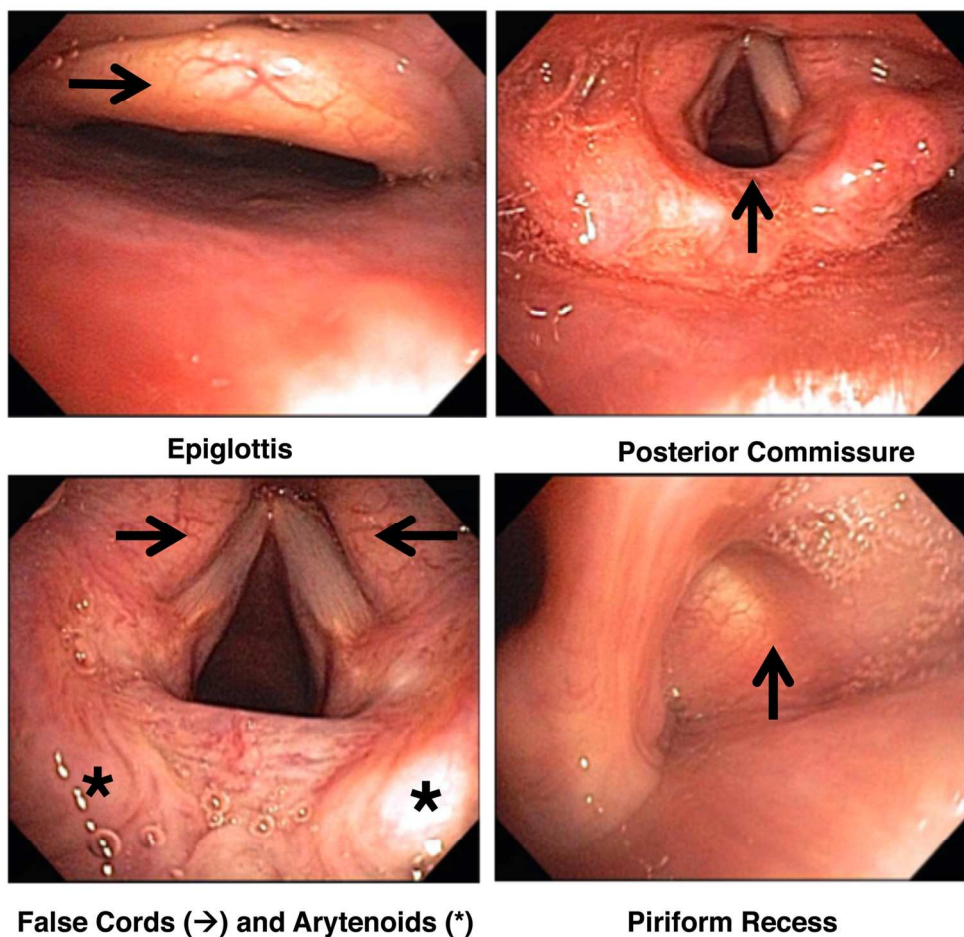
413

414 Figure 1. Example of the supraglottic areas to score. See text.

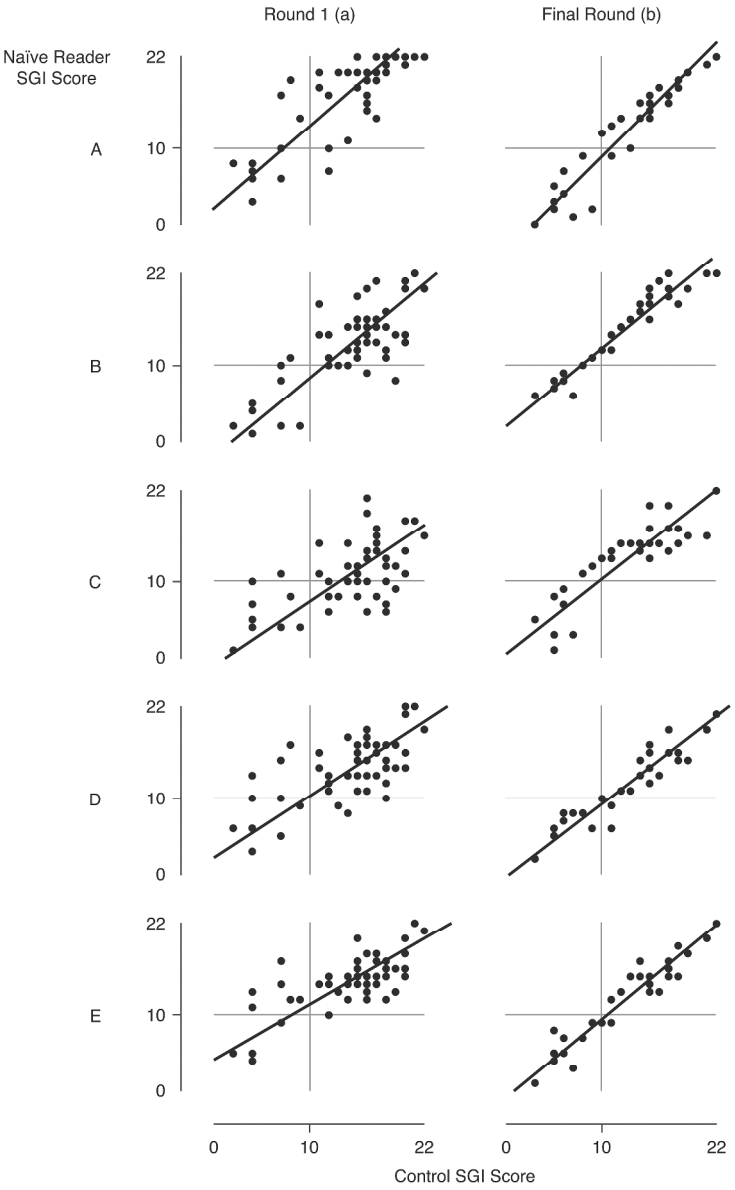
415

416 Figure 2a, b. The vertical axis represents the readers' SGI scoring and the
417 horizontal axis represents the control scores. After the first training session there
418 was discordance between readers and control (regression $R^2 = 51\%$, range 35-
419 65%) (2a). By the final session very good concordance was achieved (2b)
420 (regression $R^2 = 86\%$, range 75-92%).¹¹

421



Example of the supraglottic areas to score. See text.
144x137mm (300 x 300 DPI)



The vertical axis represents the readers' SGI scoring and the horizontal axis represents the control scores. After the first training session there was discordance between readers and control (regression $R^2 = 51\%$, range 35-65%) (2a). By the final session very good concordance was achieved (2b) (regression $R^2 = 86\%$, range 75-92%).¹¹
292x466mm (300 x 300 DPI)