Published in final edited form as:

Laryngoscope. 2018 November ; 128(11): 2448–2454. doi:10.1002/lary.27227.

Appropriateness Criteria Predict Outcomes for Sinus Surgery and May Aid in Future Patient Selection

Daniel M. Beswick, MD¹, Jess C. Mace, MPH, CCRP¹, Zachary M. Soler, MD, MSc², Noel F. Ayoub, BS³, Luke Rudmik, MD, MSc⁴, Adam S. DeConde, MD⁵, and Timothy L. Smith, MD, MPH¹

¹Department of Otolaryngology – Head and Neck Surgery, Division of Rhinology, Oregon Health & Science University, Portland, Oregon, USA

²Department of Otolaryngology-Head and Neck Surgery, Division of Rhinology and Sinus/Skull Base Surgery, Medical University of South Carolina, Charleston, South Carolina, USA

³Department of Otolaryngology – Head and Neck Surgery/School of Medicine, Stanford University, Stanford, California, USA

⁴Department of Surgery, Division of Otolaryngology-Head and Neck Surgery; University of Calgary, Calgary, Alberta, Canada

⁵Department of Surgery, Division of Otolaryngology-Head and Neck Surgery, University of California-San Diego, San Diego, CA, USA

Abstract

OBJECTIVES—Appropriateness criteria to determine surgical candidacy for chronic rhinosinusitis (CRS) have recently been described. This study stratified patients who underwent endoscopic sinus surgery (ESS) according to these new appropriateness criteria and evaluated postoperative improvements among appropriateness categories.

METHODS—Adult patients with uncomplicated CRS electing ESS were prospectively enrolled into a multi-institutional cohort study between March, 2011 and June, 2015 to assess outcomes. Subsequently, appropriateness criteria that consider preoperative medical therapy, SinoNasal Outcome Test (SNOT-22) scores, and Lund-Mackay computed tomography scores were retrospectively applied.

RESULTS—A total of 92.6% (436/471) were categorized as “appropriate” ESS candidates, 3.8% (18/471) “uncertain”, and 3.6% (17/471) “inappropriate.” Among uncertain patients, two-thirds (12/18) had identifiable reasons for undergoing ESS, most commonly oral corticosteroid intolerance (n=6). Postoperative follow-up was available for 79% (n=372). Clinically significant SNOT-22 improvements occurred in both appropriate and uncertain groups (all $p < 0.050$) but not

Corresponding Author: Timothy L. Smith, MD, MPH, Oregon Health & Science University, Department of Otolaryngology – Head and Neck Surgery, Division of Rhinology and Sinus/Skull Base Surgery, Oregon Sinus Center, 3181 SW Sam Jackson Park Road, PV-01, Portland, OR. 97239, Phone: 503-494-7413, FAX: 503-494-4631, smithtim@ohsu.edu.

Potential Conflicts of Interest: There are no relevant financial disclosures or potential conflicts of interest for Daniel M. Beswick, Noel F. Ayoub, Luke Rudmik, or Adam S. DeConde for this investigation.

Level of Evidence: 2b

among the inappropriate group. The inappropriate group reported less mean improvement in SNOT-22 total score compared to appropriate ($p=0.008$) and uncertain ($p=0.006$) groups.

CONCLUSION—The vast majority of patients (~93%) who underwent ESS in a multi-institutional research program were identified as appropriate candidates for surgical intervention as defined by current appropriateness criteria. Valid considerations frequently exist for offering ESS to patients categorized as uncertain. Appropriate and uncertain candidates report similar, clinically significant SNOT-22 improvements following surgery. Patients classified as inappropriate reported significantly less improvement following ESS. Surgical appropriateness criteria may assist in predicting outcomes of ESS.

MeSH Key Words

Sinusitis; chronic disease; quality of life; paranasal sinuses; standards

INTRODUCTION

Numerous prospective studies have validated the utility of endoscopic sinus surgery (ESS) as a treatment option for chronic rhinosinusitis (CRS), which significantly improves patients' quality of life (QoL) and decreases societal costs associated with the disease.¹⁻⁴ Until recently however there has been a dearth of clinical guidelines for the appropriate selection of CRS patients in whom to consider ESS as an important component of treatment.

A group of experts recently developed Multi-national, Multi-specialty Appropriateness Criteria for ESS in adult patients with uncomplicated CRS.^{5, 6} Within the increasingly important context of value-driven health care, recommendations such as these are intended to help optimize appropriate use of ESS as a treatment tool. These criteria, which include defined preoperative medication regimens, thresholds of disease-specific QoL impairment using the Sino-Nasal Outcomes Test-22 (SNOT-22) instrument, and extent of disease burden quantified using the Lund-Mackay computed tomography (CT) score, categorize patients as either "appropriate," "uncertain," or "inappropriate" surgical candidates. Distinct appropriateness criteria were developed for patients with nasal polyposis (CRSwNP) and without polyposis, (CRSSNP) reflecting known differences in the management of these subtypes (Table 1). For all patients with CRS, these Appropriateness Criteria represent the minimal threshold considered for ESS to be a management option and do not suggest that patients who meet these criteria should undergo surgery. These criteria were developed according to rigorous methodology but to date have not been further validated in a patient cohort.

This study sought to apply these defined surgical appropriateness criteria to a large, multi-institutional patient cohort to: 1) determine the degree to which this surgical cohort met appropriateness criteria, 2) identify cofactors associated with patients classified as either uncertain or inappropriate surgical candidates, and 3) determine if patients experienced differential outcomes after ESS based on appropriateness categorizations.

MATERIALS and METHODS

Original Patient Study Population & Inclusion Criteria

Study participants were prospectively sampled from adult heterogeneous patient populations referred to tertiary, academic clinical practices for recalcitrant symptoms of uncomplicated CRS between March, 2011 and June, 2015. Prior findings from this multi-institutional research cohort have been previously described.⁷⁻⁹ Diagnosis of CRS was confirmed using guidelines currently defined by the American Academy of Otolaryngology-Head and Neck Surgery.^{10, 11} Study participants provided informed, written consent in English and completed medical history interviews, including previous medical therapy regimens, during baseline enrollment meetings and after electing ESS. Patients were assured that voluntary study participation involved minimal risk. The Institutional Review Board (IRB) at each enrollment center provided study approval and oversight of all protocols, annual reviews, and data safety monitoring. Enrollment locations included Divisions/Departments of Otolaryngology-Head and Neck Surgery within: Oregon Health & Science University (Portland, OR; eIRB #7198), Stanford University (Palo Alto, CA; IRB #4947), the Medical University of South Carolina (Charleston, SC; IRB #12409), and the University of Calgary (Calgary, AB, Canada; IRB #E-24208).

Preoperative Sinonasal Endoscopy Examination

The paranasal sinuses were evaluated using rigid endoscopes (SCB Xenon 175, Karl Storz, Tuttlingen, Germany) and quantified by each enrolling clinician using the Lund-Kennedy scoring system (score range: 0–20), which quantifies pathologic attributes within paranasal sinuses including the identification and severity of: nasal polypsis (CRSwNP), discharge, edema, scarring, and crusting.¹² Higher summarized total scores on the Lund-Kennedy staging system reflect worse overall extent and disease severity.

Preoperative Computed Tomography

High-resolution CT imaging of the sinuses was used to evaluate preoperative sinonasal disease severity using contiguous images in sagittal, axial, and coronal planes. Radiographic images were quantified by each enrolling clinician using the Lund-Mackay scoring system (score range: 0–24).¹³ Higher summarized total scores on the Lund-Mackay staging system reflect worse global disease severity.

Endoscopic Sinus Surgery

Surgical approach was dictated by each enrolling clinician using radiographic imaging and endoscopic examination of disease location and severity. Patients underwent either primary or revision ESS completed under general anesthesia. Surgical procedures included unilateral or bilateral maxillary sinus antrostomy, total/partial ethmoidectomy, sphenoidotomy, or frontal sinusotomy (*Draf* IIa, IIb, or III), when appropriate. Nasal patency was further maximized through inferior turbinate reduction and/or septoplasty, if indicated. Postoperative management included at least daily nasal saline irrigation and topical corticosteroid spray/rinses, as necessary, to facilitate optimal postoperative healing and inflammatory disease control.

Exclusion Criteria

Selected study participants were excluded from subsequent analyses if preoperative interviews identified important comorbid conditions or diagnostic definition of complicated CRS per international guidelines.^{11, 14, 15} Exclusions were also completed for those study participants without accurate record of preoperative medical therapies as outlined by current appropriateness criteria for both patients with and without nasal polypsis (CRSsNP).^{5, 6} Additional exclusions were applied to participants who did not provide complete, study-related, patient-reported outcome measures (PROMs) within 18 months postoperatively, as described below.

Patient Reported Outcome Measures

Participants were asked to provide complete responses to the SNOT-22 during preoperative enrollment meetings and postoperative follow-up appointments. The SNOT-22 is a validated instrument designed to quantify sinonasal symptom severity (©2006, Washington University, St. Louis, MO).¹⁶ Higher summarized total scores reflect worse daily functioning and/or symptom severity (score range: 0–110). Previous factor analysis of the SNOT-22 has identified five separate subdomains with minimal item cross-loading including.¹⁷

Participants were also asked to complete secondary outcome measures designed to operationalize olfactory function both before and after ESS. The Brief Smell Identification Test (BSIT; Sonosonics Inc., Haddon Heights, NJ) is a validated 12-item, non-invasive test of olfactory function designed to operationalize olfactory identification utilizing microencapsulated odorant strips. A higher number of total correct responses (score range: 0–12) reflects superior olfactory identification. Total scores can be interpreted as either 'normal olfaction' (BSIT ≥ 9), or as some severity of olfactory dysfunction (BSIT < 8).¹⁸ The last available postoperative PROMs (6-months) was used to define each postoperative time point due to statistical similarity between mean 6-month and 18-month postoperative PROMs in this patient cohort.^{19, 20}

Determinations using Surgical Appropriateness Criteria

As the primary predictor of interest, study participants pursuing ESS were categorized into surgical appropriateness determinations based on preoperative medical therapy regimens, SNOT-22 total scores, and Lund-Mackay CT scores using a retrospective application of the appropriateness criteria.^{5, 6} The presence of preoperative nasal polypsis was confirmed using sinonasal endoscopy examinations and required for appropriateness criteria defined for CRSwNP and CRSsNP subgroups (Table 1).

Data Management and Statistical Analyses

Retrospective data analysis of prospectively collected data was completed using commercial software (SPSS v.24; IBM Corp., Armonk, NY). Patient health information was protected using unique study identification number assignment and data storage with a HIPAA compliant, closed-environment database (Access, Microsoft Corp., Redmond, WA). Distributions of scaled measured was assessed for normality and/or linearity while patient data for the final cohort was evaluated descriptively. Differences in average PROM and

BSIT scores between appropriateness groups were compared using omnibus Kruskal-Wallis test statistics, when appropriate, due to the lack of normally distributed scaled measures. Adjustments for multiple comparisons were completed using bivariate Mann Whitney U test statistics when global differences were identified between all appropriateness groups. Within-subject improvement for both PROM and BSIT scores were evaluated for each appropriate group using Wilcoxon signed rank testing. All statistical comparisons assumed a conventional 0.050 type-I error probability (p-value).

RESULTS

A total of 530 study patients with CRS completed all inclusion criteria and were prospectively enrolled into this multi-center cohort prior to ESS. The final total cohort was comprised of 471 patients after exclusions associated with: cystic fibrosis (n=16), isolated mucocele (n=12), ciliary dyskinesia (n=5), isolate fungal ball/invasive fungal sinusitis (n=7), concern for extra-paranasal disease extension (n=3), odontogenic etiologies (n=3), silent sinus syndrome (n=2), inverted papilloma (n=1), and osteoma (n=1). Additional exclusions included patients with immunosuppression/common variable immune deficiency (n=7) or sarcoidosis (n=2).

Application of current appropriateness criteria retrospectively identified 436 (92.6%) participants as appropriate surgical candidates, 18 (3.8%) participants were categorized as uncertain, 17 (3.6%) were categorized as inappropriate. All 18 uncertain participants had CRSwNP and consisted of 17 classified due to determinations associated with uncertain-2 (oral corticosteroids omitted) and one classified as uncertain-3 (SNOT-22 score < 20). Preoperative appropriateness criteria determinations, demographics, and comorbid conditions for the final study cohort are described for study participants with CRSwNP (n=176; 37.4%) and CRSsNP (n=295; 62.6%) in Table 2. Factors associated with undergoing ESS for uncertain and inappropriate are delineated in Table 3.

Preoperative PROMs between Appropriateness Groups

The absence of normally distributed data for both uncertain and inappropriate groups, likely due to limited sample sizes, determined the use of nonparametric omnibus testing to evaluate whether differences in PROMs were evident between any two groups in the total cohort (Table 4). As predicted, the inappropriate group reported relatively lower/better preoperative measures of disease severity compared to both the appropriate and uncertain groups overall. After significance adjustments for multiple comparisons, the inappropriate group reported significant better average baseline SNOT-22 total scores compared to both the uncertain (p=0.010) and appropriate (p<0.001) groups. Similar adjusted differences for the inappropriate group were reported for the rhinologic symptom domain (both p<0.001), extra-nasal rhinologic symptom domain (both p 0.035), and psychological dysfunction (both p 0.042) scores. Mean sleep dysfunction (both p 0.037) scores were significantly better for the inappropriate group compared to the appropriate cohort (p<0.001) but not when compared to the uncertain group (p=0.063). For the ear/facial symptom domain, the inappropriate group reported significantly better scores compared to the appropriate group (p<0.001), however not with the uncertain group (p=0.449). After significance adjustments

for differences in preoperative BSIT scores, the analysis did not identify any significant differences between either the appropriate and uncertain groups ($p=0.080$), the inappropriate and uncertain groups ($p=0.079$), or the appropriate and inappropriate group ($p=0.930$).

Single pairwise comparisons in the CRSsNP subgroup were evaluated due to insufficient sample sizes in the uncertain group. As anticipated, appropriate candidates with CRSsNP reported significantly worse SNOT-22 total and domain scores compared to the inappropriate group on average (Table 4). For CRSwNP, the appropriate group reported significantly worse SNOT-22 total scores compared to the inappropriate group after adjustment for multiple comparisons ($p=0.040$). After adjustment however, no other significant differences between any two appropriateness groups were reported for any of the preoperative SNOT-22 domain scores (all $p \geq 0.054$).

Postoperative Improvement between Appropriateness Groups

Postoperative follow-up was provided by 372/471 (79%) of all participants who were observed for an average of 15.4 [± 4.9] months after ESS. A total of 140/372 (38%) were participants with CRSwNP and were observed for an average 14.8 [± 4.9] months compared to 232/372 (62%) of participants with CRSsNP who were followed for an average 15.7 [± 4.9] months after surgery. Average postoperative changes for both PROMs are described and compared in Table 5 for each appropriateness grouping.

Significant postoperative within-subjects improvement over time was reported for average SNOT-22 total and domain scores for only the appropriate and uncertain surgical groups, extending to both the CRSwNP and CRSsNP study participants (Table 5). Assessment of within-subject improvement of the uncertain group for CRSsNP participants was not possible due to inadequate sample size. No significant within-subject postoperative improvement on SNOT-22 scores was reported by any of the inappropriate groups on average, likely due to sample size limitations for both CRSwNP ($n=2$) and CRSsNP ($n=9$) groups. Postoperative within-subject improvement in average BSIT scores was not reported by any group, with the exception of the appropriate surgical group with CRSwNP, likely due to associated polypectomies that group received during ESS.^{20, 21}

After significance adjustments for multiple comparisons in the total cohort, the magnitude of average improvement in SNOT-22 total scores in the inappropriate group was significantly less than both appropriate ($p=0.008$) and uncertain ($p=0.006$) groups. Similar findings extended through SNOT-22 rhinologic and extra-nasal rhinologic symptom domain score improvements whereas the inappropriate group reported significantly less improvement compared to both appropriate and uncertain counterparts (all $p \leq 0.012$). Similarly, average improvement in psychological dysfunction scores was significantly less for the inappropriate group compared to the uncertain cohort after adjustment ($p=0.016$). No significant differences in average BSIT score improvement were identified between any two appropriateness groups ($p \geq 0.100$) within the total cohort with follow-up.

No differences in average postoperative improvements were reported for any PROM between the appropriate and uncertain groups in participants with CRSwNP (all $p \geq 0.050$) following multiple comparison adjustments. Interestingly, there were significant adjusted

differences in average postoperative changes in the SNOT-22 rhinologic symptom scores between the inappropriate group and uncertain group ($p=0.026$). Appropriately categorized surgical patients without polyposis did report significant greater mean improvement across both SNOT-22 total scores and extra-nasal rhinologic symptom scores compared to the inappropriate group (all $p < 0.025$).

DISCUSSION

This investigation assessed the validity of the Appropriateness Criteria for ESS in an observational cohort of patients with uncomplicated CRS. The methodology utilized in developing these Appropriateness Criteria is well established.²² The results of this study demonstrate that the vast majority of patients (~93%) who underwent ESS in a multi-institutional research program were, in fact, appropriate candidates for surgical intervention, while far smaller proportions of patients undergoing ESS were classified as either uncertain or inappropriate surgical candidates. In many instances, sound clinical determinations were identified for patients who were not classified as appropriate prior to ESS, including oral corticosteroid intolerance or the presence of unilateral sinonasal disease. This suggests that the Appropriateness Criteria are not to be interpreted without the benefit of the individual clinician's judgment, but instead act as a general guide to optimize patient selection for ESS while still providing the capacity for individual clinician nuance and discretion to manage potential clinical exceptions. Also, we concluded that no identifiable reason or exception could be found for 12 patients (Table 3) who underwent ESS were categorized as inappropriate, even following careful, retrospective audit of electronic medical records. This suggests that inappropriate surgery occurred, even in the context of a research program designed to evaluate ESS outcomes, or that incomplete case documentation occurred. Each of these potential issues would deserve further evaluation in the context of implementing a Quality Improvement program.

The findings demonstrate that the Appropriateness Criteria can help identify differences in patient reported outcomes following sinus surgery. Patients who were classified as either appropriate or uncertain demonstrated statistically and clinically significant improvements in mean SNOT-22 score following ESS, with a similar magnitude of improvement between these two patient groups. However, patients who were categorized as inappropriate experienced significantly less average improvement in QOL scores. Therefore, clinicians should be cautious offering ESS to patients who do not meet the Appropriateness Criteria if surgical goals include clinically significant improvement in QOL according to the SNOT-22 score.

A prior attempt to define appropriateness criteria for sinus surgery in CRS using the RAND/UCLA method was undertaken in the early 1990s by the American Academy of Otolaryngology-Head and Neck Surgery in cooperation with Value Health Services (VHS), a health research firm.²³ These previous VHS criteria incorporated patient history, preoperative medical therapy, and extent of disease on CT scan, and defined surgical candidacy as "appropriate," "equivocal" or "inappropriate." A report by Jones et al. applied the SNOT-20 questionnaire, and earlier version of the SNOT-22, prospectively to 49 patients categorized by the VHS criteria and did not identify any relationship between the magnitude

of SNOT-20 change and the appropriateness rating; the equivocal group demonstrated the greatest improvement followed by the inappropriate group, with the least change seen in the appropriate group.²⁴ The results of the current study differ substantially with those from the previous study and demonstrate that current Appropriateness Criteria can assist in predicting outcomes of ESS for CRS. This may be due to the current utilization of a more precise definition of CRS, evolution in surgical technique or advances in technology, the fact that the VHS criteria did not include patient reported outcome measures as part of the appropriateness selection, or potential limitations in sample size.^{5, 6}

This study was conducted within a multi-institutional prospective research program with defined, rigorous inclusion criteria developed multiple years prior to the development of the current surgical Appropriateness Criteria. However, it is still noteworthy that even within these confines, a small percentage of patients were retrospectively identified as inappropriate surgical candidates, reinforcing the notion that the guidelines may have clinical exception. This current data cannot be generalized to assume the percentage of patients who do or do not meet appropriateness criteria in other practice settings, such as community practice or alternative academic settings. Similarly, although these data suggest that patients categorized as inappropriate do not improve to a similar degree as those who are categorized as appropriate or uncertain, it remains unknown if this finding can be generalized to the diverse population of patients undergoing ESS. Future research is necessary to establish the degree to which these findings are externally valid.

There are several caveats to consider when interpreting these study findings. There may be some magnitude of non-differential misclassification bias using the proposed measures that constitute the Appropriateness Criteria. Patient-reported medication use is subject to recall bias while the semi-quantitative nature of CRS disease staging has historically been shown to have measurable intra-rater variability.^{25–27} Lastly, none of the significant bivariate effect estimates reported were controlled for potential confounding variables due to limited sample sizes available in both the uncertain and inappropriate populations. Future research efforts examining the predictive ability of the Appropriateness Criteria for postoperative outcomes should include applying these criteria to alternative care settings, including other academic and community practices, and prospectively categorizing study participants at the time of endoscopic sinus surgery in a larger cohort of participants to allow for multifactor analysis.

CONCLUSION

In conclusion, the majority of patients (~93%) who underwent ESS in a multi-institutional research program were appropriate candidates for surgical intervention as defined by current Multi-national, Multi-specialty Appropriateness Criteria. Valid considerations frequently exist for offering ESS to patients categorized as uncertain. Appropriate and uncertain candidates reported similar, clinically significant SNOT-22 improvements following surgery. Patients classified as inappropriate reported significantly less improvement following ESS. Surgical appropriateness criteria may assist in predicting outcomes of endoscopic sinus surgery.

Acknowledgments

Funding: Timothy L. Smith, Zachary M. Soler, and Jess C. Mace were supported by a grant for this investigation from the National Institute on Deafness and Other Communication Disorders (NIDCD), one of the National Institutes of Health, Bethesda, MD., USA (R01 DC005805; PIs: T. Smith and Z. Soler). Public clinical trial registration (www.clinicaltrials.gov) ID# NCT02720653 and NCT01332136. The NIDCD did not contribute to the design or conduct of this study; preparation, review, approval or decision to submit this manuscript for publication.

References

1. Rudmik L, Smith TL, Mace JC, et al. Productivity costs decrease after endoscopic sinus surgery for refractory chronic rhinosinusitis. *Laryngoscope*. 2016; 126(3):570–574. [PubMed: 26371457]
2. Smith TL, Kern R, Palmer JN, et al. Medical therapy vs surgery for chronic rhinosinusitis: a prospective, multi-institutional study with 1-year follow-up. *Int Forum Allergy Rhinol*. 2013; 3(1): 4–9. [PubMed: 22736422]
3. Patel ZM, Thamboo A, Rudmik L, et al. Surgical therapy vs continued medical therapy for medically refractory chronic rhinosinusitis: a systematic review and meta-analysis. *Int Forum Allergy Rhinol*. 2017; 7(2):119–127. [PubMed: 27863163]
4. Rudmik L, Soler ZM, Mace JC, et al. Economic evaluation of endoscopic sinus surgery versus continued medical therapy for refractory chronic rhinosinusitis. *Laryngoscope*. 2015; 125(1):25–32. [PubMed: 25186499]
5. Rudmik L, Soler ZM, Hopkins C, et al. Defining appropriateness criteria for endoscopic sinus surgery during management of uncomplicated adult chronic rhinosinusitis: a RAND/UCLA appropriateness study. *Rhinology*. 2016; 54(2):117–128. [PubMed: 26934470]
6. Rudmik L, Soler ZM, Hopkins C, et al. Defining appropriateness criteria for endoscopic sinus surgery during management of uncomplicated adult chronic rhinosinusitis: a RAND/UCLA appropriateness study. *Int Forum Allergy Rhinol*. 2016; 6(6):557–567. [PubMed: 26970538]
7. Alt JA, DeConde AS, Mace JC, et al. Quality of Life in Patients With Chronic Rhinosinusitis and Sleep Dysfunction Undergoing Endoscopic Sinus Surgery: A Pilot Investigation of Comorbid Obstructive Sleep Apnea. *JAMA Otolaryngol Head Neck Surg*. 2015; 141(10):873–881. [PubMed: 26356240]
8. DeConde AS, Mace JC, Levy JM, et al. Prevalence of polyp recurrence after endoscopic sinus surgery for chronic rhinosinusitis with nasal polyposis. *Laryngoscope*. 2017; 127(3):550–555. [PubMed: 27859303]
9. Levy JM, Mace JC, Bodner TE, et al. Defining the minimal clinically important difference for olfactory outcomes in the surgical treatment of chronic rhinosinusitis. *Int Forum Allergy Rhinol*. 2017; 7(8):821–826. [PubMed: 28556611]
10. Rosenfeld RM, Andes D, Bhattacharyya N, et al. Clinical practice guideline: adult sinusitis. *Otolaryngol Head Neck Surg*. 2007; 137(3 Suppl):S1–S1. [PubMed: 17761281]
11. Rosenfeld RM, Piccirillo JF, Chandrasekhar SS, et al. Clinical practice guideline (update): adult sinusitis. *Otolaryngol Head Neck Surg*. 2015; 152(2 Suppl):S1–S39.
12. Lund VJ, Kennedy DW. Staging for rhinosinusitis. *Otolaryngol Head Neck Surg*. 1997; 117(3 Pt 2):S35–40. [PubMed: 9334786]
13. Lund VJ, Mackay IS. Staging in rhinosinusitis. *Rhinology*. 1993; 31(4):183–184. [PubMed: 8140385]
14. Desrosiers M, Evans GA, Keith PK, et al. Canadian clinical practice guidelines for acute and chronic rhinosinusitis. *Allergy Asthma Clin Immunol*. 2011; 7(1):2. [PubMed: 21310056]
15. Fokkens WJ, Lund VJ, Mullol J, et al. EPOS 2012: European position paper on rhinosinusitis and nasal polyps 2012. A summary for otorhinolaryngologists. *Rhinology*. 2012; 50(1):1–12. [PubMed: 22469599]
16. Hopkins C, Gillett S, Slack R, et al. Psychometric validity of the 22-item Sinonasal Outcome Test. *Clin Otolaryngol*. 2009; 34(5):447–454. [PubMed: 19793277]

17. DeConde AS, Bodner TE, Mace JC, et al. Response shift in quality of life after endoscopic sinus surgery for chronic rhinosinusitis. *JAMA Otolaryngol Head Neck Surg.* 2014; 140(8):712–719. [PubMed: 25074504]
18. Doty RL, Marcus A, Lee WW. Development of the 12-item Cross-Cultural Smell Identification Test (CC-SIT). *Laryngoscope.* 1996; 106(3 Pt 1):353–356. [PubMed: 8614203]
19. DeConde AS, Mace JC, Alt JA, et al. Longitudinal improvement and stability of the SNOT-22 survey in the evaluation of surgical management for chronic rhinosinusitis. *Int Forum Allergy Rhinol.* 2015; 5(3):233–239. [PubMed: 25511442]
20. Levy JM, Mace JC, Sansoni ER, et al. Longitudinal improvement and stability of olfactory function in the evaluation of surgical management for chronic rhinosinusitis. *Int Forum Allergy Rhinol.* 2016; 6(11):1188–1195. [PubMed: 27218483]
21. Litvack JR, Mace J, Smith TL. Does olfactory function improve after endoscopic sinus surgery? *Otolaryngol Head Neck Surg.* 2009; 140(3):312–319. [PubMed: 19248934]
22. Brook RH, Chassin MR, Fink A, et al. A method for the detailed assessment of the appropriateness of medical technologies. *Int J Technol Assess Health Care.* 1986; 2(1):53–63. [PubMed: 10300718]
23. Dana S. VHS Sinus Panel Report: Appropriateness Criteria for Sinus Surgery. *Bulletin of the AAO-HNS.* 1994; 13:12–14.
24. Jones ML, Piccirillo JF, Haiduk A, et al. Functional endoscopic sinus surgery: do ratings of appropriateness predict patient outcomes? *Am J Rhinol.* 1998; 12(4):249–255. [PubMed: 9740917]
25. Colley P, Mace JC, Schaberg MR, et al. Impact of educational intervention on the interrater agreement of nasal endoscopy interpretation. *Laryngoscope.* 2015; 125(10):2259–2265. [PubMed: 25781864]
26. McCoul ED, Smith TL, Mace JC, et al. Interrater agreement of nasal endoscopy in patients with a prior history of endoscopic sinus surgery. *Int Forum Allergy Rhinol.* 2012; 2(6):453–459. [PubMed: 22696506]
27. Raithatha R, Anand VK, Mace JC, et al. Interrater agreement of nasal endoscopy for chronic rhinosinusitis. *Int Forum Allergy Rhinol.* 2012; 2(2):144–150. [PubMed: 22253129]

Table 1Determination of appropriateness criteria for study participants electing endoscopic sinus surgery^{5, 6}

Subgroup:	Determinations:	Surgical Appropriateness Criteria Requirements		
		Minimum previous medical therapies:	SNOT-22 total score	Lund-Mackay CT total score
CRSwNP	Appropriate	Dual therapy including: <ol style="list-style-type: none"> Topical intranasal corticosteroids (8 weeks), and Short-course of systemic corticosteroids (1–3 weeks) 	20	1
	Uncertain (1)	Single therapy including: <ol style="list-style-type: none"> Short-course of systemic corticosteroids (1–3 weeks) only 	20	1
	Uncertain (2)	Single therapy including: <ol style="list-style-type: none"> Topical intranasal corticosteroids (8 weeks) only 	20	>5
	Uncertain (3)	Dual therapy including: <ol style="list-style-type: none"> Topical intranasal corticosteroids (8 weeks), and Short-course of systemic corticosteroids (1–3 weeks) 	<20	1
	Inappropriate	Not classified using criteria for appropriate or uncertain		
CRSsNP	Appropriate	Dual therapy including: <ol style="list-style-type: none"> Topical intranasal corticosteroids (8 weeks), and either: Short-course of broad-spectrum or culture-directed systemic antibiotic (2–3 weeks), or Extended-course of systemic low-dose anti-inflammatory antibiotic (12 weeks) 	20	1
	Uncertain	Single therapy including: <ol style="list-style-type: none"> Topical intranasal corticosteroids (8 weeks) 	20	1
	Inappropriate	Not classified using criteria for appropriate or uncertain		

SNOT-22, 22-item SinoNasal Outcome Test; CT, computed tomography; CRSwNP, chronic rhinosinusitis with nasal polyposis; CRSsNP, chronic rhinosinusitis without nasal polyposis.

Table 2

Preoperative cofactors for study participants with and without nasal polyposis (n=471)

Multinational, Multispecialty Criteria:		CRSwNP (n=176)	CRSsNP (n=295)
Appropriate	n (%)	156 (89%)	280 (95%)
Uncertain	n (%)	18 (10%)	0 (0%)
Inappropriate	n (%)	2 (1%)	15 (5%)
Cofactors:			
Age (years)	mean [$\pm$ SD]	50.4 [$\pm$ 15.2]	50.1 [$\pm$ 15.2]
Male	n (%)	98 (56%)	127 (43%)
Female	n (%)	78 (44%)	168 (57%)
White/Caucasian	n (%)	142 (81%)	258 (88%)
Hispanic/Latino	n (%)	14 (8%)	14 (5%)
Revision ESS	n (%)	116 (66%)	143 (49%)
Deviated septum	n (%)	51 (29%)	137 (46%)
Asthma	n (%)	104 (59%)	80 (27%)
ASA sensitivity	n (%)	32 (18%)	4 (1%)
Allergy	n (%)	94 (53%)	121 (41%)
Depression *	n (%)	16 (9%)	54 (18%)
Current tobacco use *	n (%)	9 (5%)	20 (7%)
Current alcohol use *	n (%)	68 (39%)	135 (46%)
Corticosteroid dependent	n (%)	21 (12%)	16 (5%)
Diabetes mellitus (Type I/II)	n (%)	7 (4%)	19 (6%)
Lund-Kennedy endoscopy score	mean [$\pm$ SD]	9.0 [$\pm$ 3.3]	4.4 [$\pm$ 2.9]
Lund-Mackay CT score	mean [$\pm$ SD]	16.6 [$\pm$ 4.8]	9.6 [$\pm$ 5.1]

CRSwNP, chronic rhinosinusitis with nasal polyposis; CRSsNP, chronic rhinosinusitis without nasal polyposis; ESS, endoscopic sinus surgery; ASA, acetylsalicylic acid; SD, standard deviation; CT, computed tomography;

* self-reported.

Table 3

Factors associated with being offered endoscopic sinus surgery for uncertain and inappropriate patients.

Chronic Rhinosinusitis with Nasal Polyposis (CRSwNP)	
<i>Uncertain (n=18)</i>	
Factors:	N
- Patient declined oral corticosteroids after counseling of risk/benefit *	6
- None identified	6
- Limited preoperative oral corticosteroids (5–6 days) **	3
- Severe comorbid asthma/AERD with anticipated lack of benefit from medical therapy	2
- Unilateral disease	1
<i>Inappropriate (n=2)</i>	
Factors:	
- Prior sinus surgery aborted due to bleeding, referred for completion surgery	1
- None identified	1
Chronic Rhinosinusitis without Nasal Polyposis (CRSsNP)	
<i>Inappropriate (n=15)</i>	
Factors:	
- None identified	11
- Unilateral disease	3
- Chronic oral corticosteroid use required for symptom control	1

* Indicates: poorly controlled diabetes mellitus (n=2), patient preference (n=2), psychiatric comorbidity concerns (n=1), corneal disease with prior corneal transplant (n=1);

** appropriateness criteria recommends at least 7 days; AERD, aspirin exacerbated respiratory disease; CRSwNP, chronic rhinosinusitis with nasal polyposis; CRSsNP, chronic rhinosinusitis without nasal polyposis.

Table 4

Comparison of average preoperative PROM scores between appropriateness groups

Total Cohort (n=471)	Appropriate (n=436)		Uncertain (n=18)		Inappropriate (n=17)		Omnibus test statistic	Omnibus p-value
	Mean [±SD]		Mean [±SD]		Mean [±SD]			
SNOT-22 total	55.8 [±18.8]		46.9 [±18.4]		20.8 [±19.6]		KW=34.0	<0.001
Rhinologic symptoms	17.1 [±5.9]		17.0 [±6.9]		6.5 [±5.1]		KW=32.0	<0.001
Extra-nasal rhinologic symptoms	8.8 [±3.4]		6.8 [±3.5]		3.2 [±2.4]		KW=35.8	<0.001
Ear/facial symptoms	9.7 [±5.1]		6.8 [±5.3]		4.6 [±5.6]		KW=22.0	<0.001
Psychological dysfunction	16.8 [±8.2]		13.7 [±5.8]		5.2 [±7.1]		KW=27.5	<0.001
Sleep dysfunction	14.5 [±6.5]		12.1 [±6.0]		5.1 [±6.5]		KW=25.4	<0.001
BSIT total	8.8 [±3.1]		6.9 [±3.7]		10.2 [±1.4]		KW=6.1	0.047
CRSsNP (n=176)	n=156		n=18		n=2			
SNOT-22 total	57.6 [±19.7]		46.9 [±18.4]		13.0 [±0.0]		KW=9.7	0.008
Rhinologic symptoms	19.8 [±5.5]		17.0 [±6.9]		6.5 [±0.7]		KW=7.6	0.023
Extra-nasal rhinologic symptoms	8.9 [±3.4]		6.8 [±3.5]		1.5 [±0.7]		KW=9.6	0.008
Ear/facial symptoms	9.5 [±5.4]		6.8 [±5.3]		4.5 [±2.1]		KW=5.9	0.051
Psychological dysfunction	16.3 [±8.4]		13.7 [±5.8]		1.5 [±0.7]		KW=6.6	0.037
Sleep dysfunction	14.5 [±6.8]		12.1 [±6.0]		1.0 [±1.4]		KW=7.8	0.020
BSIT total	7.0 [±3.7]		6.9 [±3.7]		11.5 [±0.7]		KW=3.7	0.160
CRSsNP (n=295)	n=280				n=15			
SNOT-22 total	54.8 [±18.2]	-----	-----		21.9 [±20.8]		MWU=476.0	<0.001
Rhinologic symptoms	15.6 [±5.6]	-----	-----		6.5 [±5.5]		MWU=523.0	<0.001
Extra-nasal rhinologic symptoms	8.8 [±3.4]	-----	-----		3.5 [±2.5]		MWU=476.0	<0.001
Ear/facial symptoms	9.8 [±4.9]	-----	-----		4.6 [±6.0]		MWU=840.5	<0.001
Psychological dysfunction	17.1 [±8.1]	-----	-----		5.7 [±7.5]		MWU=661.5	<0.001
Sleep dysfunction	14.5 [±6.4]	-----	-----		5.6 [±6.7]		MWU=729.0	<0.001
BSIT total	9.7 [±2.2]	-----	-----		9.9 [±1.3]		MWU=675.0	0.679

PROM, patient reported outcome measure; SD, standard deviation; SNOT-22, 22-item SinoNasal Outcome Test; BSIT, Brief Smell Identification Test; CRSsNP, chronic rhinosinusitis with nasal polyposis; CRSsNP, chronic rhinosinusitis without nasal polyposis; KW, Kruskal-Wallis omnibus testing; MWU, Mann Whitney U bivariate testing.

Table 5

Comparison of average postoperative changes in PROM scores between pre-surgical appropriateness groups (n=372)

Total Cohort (n=372)	Appropriate n=348		Uncertain n=13		Inappropriate n=11		Omnibus test statistic	Omnibus p-value
	Mean [±SD]		Mean [±SD]		Mean [±SD]			
SNOT-22 total	-26.4 [±21.9] *		-34.2 [±18.3] *		-6.6 [±19.9]		KW=10.9	0.004
Rhinologic symptoms	-8.3 [±7.3] *		-12.9 [±6.1] *		-0.8 [±7.5]		KW=14.5	0.001
Extra-nasal rhinologic symptoms	-4.1 [±3.8] *		-4.7 [±3.0] *		0.4 [±3.9]		KW=11.3	0.003
Ear/facial symptoms	-4.9 [±5.0] *		-5.1 [±4.5] *		-2.3 [±4.8]		KW=5.5	0.065
Psychological dysfunction	-7.8 [±8.8] *		-11.4 [±6.5] *		-3.3 [±6.0]		KW=7.9	0.019
Sleep dysfunction	-6.3 [±6.9] *		-7.5 [±6.9] *		-2.0 [±6.0]		KW=4.3	0.115
BSIT total	0.2 [±2.8]		0.4 [±3.6]		-1.2 [±2.3]		KW=1.9	0.379
CRSsNP (n=140)		n=125		n=13		n=2		
SNOT-22 total	-29.0 [±23.6] *		-34.2 [±18.3] *		6.0 [±1.4]		KW=5.2	0.076
Rhinologic symptoms	-10.6 [±7.6] *		-12.9 [±6.1] *		5.5 [±0.7]		KW=7.0	0.031
Extra-nasal rhinologic symptoms	-4.5 [±3.8] *		-4.7 [±3.0] *		3.5 [±0.7]		KW=5.7	0.058
Ear/facial symptoms	-5.3 [±5.3] *		-5.1 [±4.5] *		-0.5 [±0.7]		KW=2.4	0.305
Psychological dysfunction	-7.6 [±9.4] *		-11.4 [±6.5] *		-1.0 [±0.0]		KW=5.4	0.066
Sleep dysfunction	-6.3 [±7.1] *		-7.5 [±6.9] *		0.0 [±0.0]		KW=2.5	0.290
BSIT total	0.9 [±3.7] *		0.4 [±3.6]		-1.0 [±1.4]		KW=1.4	0.503
CRSsNP (n=232)		n=223				n=9		
SNOT-22 total	-25.0 [±20.9] *		----		-9.4 [±21.1]		MWU=1445.5	0.025
Rhinologic symptoms	-7.1 [±6.7] *		----		-2.2 [±7.6]		MWU=1359.0	0.071
Extra-nasal rhinologic symptoms	-3.9 [±3.8] *		----		-0.3 [±4.0]		MWU=1470.5	0.018
Ear/facial symptoms	-4.7 [±4.9] *		----		-2.7 [±5.3]		MWU=1357.5	0.072
Psychological dysfunction	-8.0 [±8.5] *		----		-3.8 [±6.5]		MWU=1347.5	0.081
Sleep dysfunction	-6.3 [±6.8] *		----		-2.4 [±6.6]		MWU=1308.5	0.122

Total Cohort (n=372)	Appropriate n=348	Uncertain n=13	Inappropriate n=11		Omnibus test statistic	Omnibus p-value
	Mean [±SD]	Mean [±SD]	Mean	[±SD]		
BSIT total	-0.1 [±2.3]	---	-1.3	[±3.1]	MWU=179.0	0.362

PROM, patient reported outcome measure; SD, standard deviation; SNOT-22, 22-item SinoNasal Outcome Test; BSIT, Brief Smell Identification Test; CRSwNP, chronic rhinosinusitis with nasal polypsis; CRSsNP, chronic rhinosinusitis without nasal polypsis; KW, Kruskal-Wallis omnibus testing; MWU, Mann Whitney U bivariate testing.

* indicates significant (p<0.050) within-subjects improvements over time using Wilcoxon signed rank testing.