Investigating the minimal clinically important difference for SNOT-22 symptom domains in surgically managed chronic rhinosinusitis

Naweed I. Chowdhury, *Vanderbilt University*
Jess C. Mace, *Oregon Health and Science University*
Todd E. Bodner, *Portland State University*
Jeremiah A. Alt, *University of Utah*
Adam S. Deconde, *University of California San Diego*
Joshua Levy, *Emory University*
Timothy L. Smith, *Oregon Health and Science University*

Journal Title: International Forum of Allergy and Rhinology
Volume: Volume 7, Number 12
Publisher: Wiley: 12 months | 2017-12-01, Pages 1149-1155
Type of Work: Article | Post-print: After Peer Review
Publisher DOI: 10.1002/alr.22028
Permanent URL: https://pid.emory.edu/ark:/25593/tkg9p

Final published version: http://dx.doi.org/10.1002/alr.22028

HHS Public Access

Author manuscript

Int Forum Allergy Rhinol. Author manuscript; available in PMC 2018 December 01.

Published in final edited form as:

Int Forum Allergy Rhinol. 2017 December ; 7(12): 1149–1155. doi:10.1002/alr.22028.

Investigating the Minimal Clinically Important Difference for SNOT-22 Symptom Domains in Surgically Managed Chronic Rhinosinusitis

Naweed I. Chowdhury, MD[1], **Jess C. Mace, MPH, CCRP**[2], **Todd E. Bodner, PhD**[3], **Jeremiah A. Alt, MD, PhD**[4], **Adam S. Deconde, MD**[5], **Joshua M. Levy, MD, MPH**[6], and **Timothy L. Smith, MD, MPH**[2]

[1]Vanderbilt University Medical Center, Department of Otolaryngology & Head and Neck Surgery, Nashville, TN

[2]Oregon Health & Science University, Department of Otolaryngology-Head and Neck Surgery, Portland, OR

[3]Portland State University, Department of Psychology, Portland, OR

[4]University of Utah, School of Medicine, Department of Surgery, Division of Otolaryngology – Head & Neck Surgery, Salt Lake City, UT

[5]University of California – San Diego School of Medicine, Department of Surgery, Division of Otolaryngology, San Diego, CA

[6]Emory University School of Medicine, Department of Otolaryngology-Head and Neck Surgery, Atlanta, GA

Abstract

BACKGROUND—Prior work has described five domains within the 22-item SinoNasal Outcomes Test (SNOT-22) that allow for stratification of symptoms into similar clusters that can be used to direct therapy. While the outcomes of various interventions on these symptom domains have been reported, minimal clinically important difference (MCID) values have not been previously investigated, limiting clinical interpretation of these results.

METHODS—The study was designed as secondary analysis of a prospective, multi-institutional, observational cohort. A total of 276 patients with medically refractory CRS who underwent surgical management were enrolled. Distribution-based methods (half-standard deviation, standard error of measurement, Cohen's d, and the minimum detectable change) were used to compute MCID values for both SNOT-22 total and domain scores. The Medical Outcomes Study Short-Form 6D (SF-6D) health utility score was used to operationalize anchor-based associations using receiver-operating-characteristic (ROC) curves.

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: None

The abstract for this manuscript was accepted for a podium presentation to the American Rhinologic Society during the American Academy of Otolaryngology-Head and Neck Surgery Annual Meeting & OTO Experience in Chicago, IL, September, 2017.

RESULTS—The mean MCID of several distribution-based methods for total SNOT-22 scores was 9.0, in agreement with previously published metrics. Average MCID values for the rhinologic, extra-nasal rhinologic, ear/facial, psychological, and sleep domain scores were: 3.8, 2.4, 3.2, 3.9, and 2.9, respectively. Anchor-based approaches with SF-6D did not have strong predictive accuracy across total SNOT-22 scores or domains (ROC areas under-the-curve ≤0.71), indicating weak associations between improvement in SNOT-22 scores and health utility as measured by the SF-6D.

CONCLUSIONS—This estimation of MCID values for the SNOT-22 symptom domains allows for improved clinical interpretation of results from past, present, and future rhinologic outcomes research.

Medical Subject Headings

Sinusitis; chronic disease; patient outcome assessment; quality-of-life; sinus surgery

INTRODUCTION

As the field of outcomes research continues to mature, there has been an increasing emphasis on placing statistical differences into a clinically meaningful context. Historically, and often incorrectly, this has been done by reporting measures of statistical significance and implying clinical significance. The concept of the minimal clinically important difference (MCID) evolved to combat this practice by defining a threshold value by which a statistically significant result may also be thought to offer a clinically meaningful result.

The validated 22-question Sinonasal Outcomes Test (SNOT-22) is a widely adopted instrument to evaluate chronic rhinosinusitis (CRS) treatment outcomes. The MCID of the total SNOT-22 score has previously been defined as 8.9 points using 3-month postoperative scores.[1] The characterization of five distinct symptom domains within the SNOT-22[2] has further refined the utility of the SNOT-22 instrument by allowing for patient responses to be measured across multiple different clinical domains, thus enabling both clinicians and researchers to examine the effects of various interventions on discrete rhinologic, extra-rhinologic, ear/facial, psychological, and sleep symptoms associated with CRS. While many studies have reported treatment outcomes across the different SNOT-22 domains[3–5], an MCID value for these domain scores has not been previously identified, representing a crucial gap in the current literature. In this investigation, we sought to use distribution-based and anchor-based methods to define the MCID of the SNOT-22 domains in order to qualify the analyses present in rhinologic outcomes research while also providing physicians with a clinical context for interpreting these scores in their patients.

MATERIALS and METHODS

Patient Enrollment and Inclusion Criteria

Findings and descriptions of this cohort investigation have been previously described.[6–8] Adult (≥18 years of age) study participants were prospectively sampled from heterogeneous patient populations referred to tertiary, academic practices in North America for recalcitrant symptoms of CRS between March, 2011 and June, 2015. Diagnoses were confirmed by

fellowship trained rhinologists using guidelines currently defined by the American Academy of Otolaryngology.[9,10] Study participants provided informed, written, consent in English during baseline enrollment meetings. An Institutional Review Board (IRB) at each enrollment center provided annual approval and regulatory oversight of all protocols, annual reviews, and safety monitoring. Clinical enrollment centers included: 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). Patients were assured that study participation involved no more than minimal risk and was voluntary per good clinical practice guidelines established by the International Conference on Harmonization.[11]

Each participant provided extensive medical history to confirm completion of prior therapeutics consisting of at least one course (≥14 days) of culture-directed or empiric antibiotics, either corticosteroid nasal sprays (≥21 days) or systemic corticosteroid therapy (≥5 days), and daily saline solution irrigation (~240ml) as needed.

Surgical intervention was not randomized or assigned for study purposes. Study participants elected endoscopic sinus surgery (ESS) due to inadequate symptom control from prior medical management. Surgical approach was dictated by each enrolling clinician using radiographic imaging and endoscopic examination findings of disease location and extent. Patients underwent either primary or revision ESS completed under general anesthesia. Surgical procedures consisted of unilateral or bilateral maxillary sinus antrostomy, total or partial ethmoidectomy, sphenoidotomy, or frontal sinusotomy when appropriate. Anatomic ventilation was further maximized by incorporating inferior turbinate reduction and/or septoplasty if indicated. Postoperative therapeutics included at least daily nasal saline irrigations QD and topical corticosteroid spray/rinses as necessary to optimize postoperative healing.

Exclusion Criteria

Potential disparities in global health as well as variability in surgical interventions and postoperative therapeutic regimen warranted the exclusion of some patient groups from this heterogeneous population. Study participants with concurrent recurrent acute rhinosinusitis (RARS), comorbid ciliary dyskinesia, or corticosteroid dependent conditions (eg. sinusitis, asthma) were excluded from final analyses. Additionally, enrolled study participants who did not return for postoperative follow-up appointments or respond to study related follow-up communication were considered lost to follow-up and excluded from final analyses.

SinoNasal Outcome Test (SNOT-22) Instrument

Study participants were asked to complete the SNOT-22, a 22-item validated survey developed to quantify the severity of sinonasal symptoms (©2006, Washington University, St. Louis, MO).[1] The 22-items of the SNOT-22 were categorized into 5 symptom domain scores: the rhinologic symptoms domain (range: 0–30), extra-nasal rhinologic symptoms domain (range: 0–15), ear/facial symptoms domain (range: 0–25), psychological dysfunction domain (range: 0–35), and the sleep dysfunction domain (range: 0–25) as identified through previous factor analysis of this cohort.[12] A MCID value for SNOT-22 total scores has been

previously defined as within-subject postoperative improvement of at least 8.9 points in patients with CRS.[1] Study participants were observed through the standard of care up to 12-months after ESS and asked to complete the SNOT-22 survey both preoperatively and during postoperative follow-up evaluations. This time period was chosen due to consensus by the authors that this represented an appropriate follow up period by which the longitudinal effect of the intervention could be assessed. Prior work has shown that postoperative improvements in overall symptom severity are durable between 6, 12, and 18 month follow-up periods.[13]

SF-6D Health Utility Instrument

Preoperative and postoperative overall health states were also captured using the Short Form-6D (SF-6D) instrument in addition to the SNOT-22. Individual survey items of the SF-6D are extracted from the 36-item Medical Outcomes Study Short-Form (SF-36) and transformed into health utility values using a standardizing, weighted algorithm described by Brazier et al. and obtained with permission from the Department of Health Economics and Decision Science at the University of Sheffield.[14] The normalized SF-6D value represents a quantified health state that an individual assigns to themselves on a spectrum (range: 0.3–1.0) whereas 1.0 reflects a state of ‘perfect health’. A post-treatment difference of 0.03–0.05 has been previously suggested as an MCID for the SF-6D.[15] Due to the absence of a purely objective measure of sinonasal health status, a postoperative improvement of 0.03 on the SF-6D was selected as the external anchor-based criterion for improvement following ESS.

Data Management and Statistical Analyses

Patient confidentiality was achieved by unique study identification number assignment and data collection using a HIPAA compliant, closed-environment database (Access, Microsoft Corp., Redmond, WA). Secondary data analysis of this cohort was completed using commercial software (SPSS v.24; IBM Corp., Armonk, NY). Distributions of scaled data was evaluated for assumptions of normality and/or linearity while all final descriptive patient data is provided in Table 1.

There are two schools of thought with respect to the process to compute the MCID. Distribution-based methods seek to determine the spread of scores in a baseline population and use the standard deviation of the sample or the standard error of measurement to derive a threshold score that is, statistically speaking, unlikely to be due to chance. A score above this threshold then gives the user confidence (but not certainty) that the measure obtained is a tangible change as opposed to an error of measurement. In contrast, anchor-based methods seek to link changes in PROM scores to an independent clinical anchor measure. For the original SNOT-22 questionnaire, the MCID was computed to be 8.9 in a population of surgical patients with CRS using this approach. This value has now become widely adopted in the literature and used extensively for CRS outcomes research.[1]

In the present investigation, we first used distribution-based methods to compute the MCID of preoperative total SNOT-22 score and the individual domain scores. The half standard deviation estimate was computed by calculating the standard deviation (SD) of each score-value distribution and multiplying this by 0.5.[16] Note that this is also equivalent to Cohen’s medium effect estimation. To determine the standard error of measurement (SEM) MCID

score, we first calculated the internal reliability of each score distribution using Cronbach's alpha (Table 4).[17] This reliability score (R) was then used to compute the SEM using the formula: $SEM = SD * (\sqrt{1-R})$.[18,19] The minimum detectable change (MDC) MCID score was also derived from the SEM using a different formula $\mathrm{MDC} = 1.96 * \sqrt{2} * SEM$.[20] The effect size based MCID was calculated using Cohen's small effect size estimation (0.20)[21] multiplied by the baseline standard deviation of scores.[22]

Anchor-based methods were then used to determine MCID values using the SF-6D as the determinant of a true change in health status.[23] The threshold of improvement was defined at 0.03, or 1 MCID of the SF-6D.[15] Receiver-operating characteristic (ROC) curves were computed for the 12-month change in the SNOT-22 total score and the individual domain scores against the change in the SF-6D, and area-under-the-curve (AUC) values were calculated to determine the predictive accuracy of this method.[24] AUC values of 0.8 or higher were considered to be strong candidates for established MCID values as determined by the Youden Index.[25] A mean value across all methods was then computed to determine a pooled measure of the MCID across all methods.[26]

RESULTS

Study Cohort Population

Preliminary cohort data consisted of 604 study participants who met inclusion criteria, provided preoperative SNOT-22 surveys, and completed ESS between March, 2011 and June, 2015. A total of 501 subjects were considered for analysis after exclusions for RARS (n=39), ciliary dyskinesia (n=23), or steroid dependent comorbidity (n=41) and described in Table 1. Study participants providing 12-month postoperative follow-up (n=276; 55%) were evaluated for postoperative changes in SNOT-22 and SF-6D scores and comprised the final cohort selection.

Postoperative Improvement in SNOT-22 and SF-6D Scores

The distribution of all postoperative improvement scores for the SNOT-22 and SF-6D was approximately normal by visual estimation. Highly significant 12-month postoperative improvement in mean scores was reported across all patient reported outcomes (Table 2), with an average improvement of 26.7 points on total SNOT-22 scores following surgery ($p<0.001$).

Distribution-based Methods for MCID Determination

Estimations of distribution-based methods for determining the MCID values from SNOT-22 total and domain scores (n=276) were calculated and compared (Table 3). The average MCID for SNOT-22 total scores was 9.0, while the average MCID for the rhinologic domain, extra-nasal rhinologic domain, ear/facial domain, psychological domain, and sleep domain score was 3.8, 2.4, 3.2, 3.9, and 2.9, respectively. Internal consistency reliability estimates were determined using Cronbach's alpha (R) values for preoperative SNOT-22 total and domains scores (Table 4) for use in the calculation of SEM.

Anchor-based Methods for MCID Determination: ROC Curves

MCID values for SNOT-22 total and domain scores were also investigated through the use of anchor-based calculations using patient responses from the SF-6D health utility scores as a measure of global health improvement. A total of 165 (61%) participants reported improvement in SF-6D scores of at least one MCID value to represent a true improvement in global health state. Table 5 describes the AUC values representing the ability of the total SNOT-22 score and symptom sub-domains to classify patients with a minimum improvement in their overall health status as measured by the SF-6D. The AUC values ranged from as high as 0.71 for the psychological symptom domain to 0.62 for the extra-nasal rhinologic domain, with all tested domains showing statistical significance ($p<0.001$) when compared to the null AUC of 0.50.

DISCUSSION

With the increasing adoption of patient reported outcomes measures (PROM) to track clinical improvement comes the problem of how to analyze these abstract values in a clinically-relevant context. Although tests like the SNOT-22 are validated to be reliable measures of disease at a particular time point, the validation process offers very little guidance with how to interpret changes in scores over time. The ultimate objective is to identify a threshold by which a change in score is reflective of a true change in health status; in outcomes research, this threshold is commonly known as the MCID.

This investigation sought to investigate the MCID of the SNOT-22 domains using a number of commonly described methods. The first approach used the half standard deviation - a commonly cited “rule of thumb” for computing the MCID. This corresponds to Cohen’s estimate for a medium effect size calculation[21], which forms the statistical justification for this choice. In addition, psychological testing has suggested that the capability for symptom discrimination in humans is about the level of a half standard deviation of the baseline test distribution[27]. Interestingly, this measure of the MCID appears to be conserved across many different PROMs from other chronic disease states, which makes the application of this method to the SNOT-22 a reliable approach.[16,28] This was true in the present analysis as well. Using various methods, the half standard deviation was closest to the average MCID across total SNOT-22 scores and most domains as well as the closest approximation to the previously reported value of 8.9.

The SEM technique, by contrast, is a more statistically rigorous approach that uses the reliability score of the questionnaire to determine its error of measurement. In essence, the MCID is viewed as a change in score that is large enough to have a low probability of being the result of variations in scoring inherent to test sampling error. In our data, the SEM gave smaller MCIDs than the half standard deviation, reflecting a more liberal threshold for clinical significance than the half standard deviation. This is due to the high reliability of the SNOT-22 and its domains (Table 4), which minimize fluctuations due to measurement error. Cohen’s d estimate was the lowest threshold for the MCID across all scores, representing the smallest acceptable value of the MCID. Conversely, the MDC set the most conservative MCID threshold that increases the specificity of the MCID as a measure of true change at the expense of sensitivity. This higher value is a result of the MDC calculation imposing a

95% confidence interval around the SEM. All approaches are statistically valid methods, highlighting the fact that, ultimately, there is no “true” value of the MCID for all measured populations, and that the validity of the MCID is dependent on its acceptance by physicians who use the instrument routinely. For treatment outcomes research into CRS, 8.9 is generally accepted as the MCID threshold for improvement following sinus surgery. This investigation found that an equally weighted average MCID value obtained by the four methods described was 9.0, approximating that previously published result. The reported average SNOT-22 domain MCID values are therefore likely to be the best current approximations for clinically interpreting postoperative changes in patients following ESS.

It may seem somewhat atypical to use statistical determinations alone to define clinically important differences, and one could argue that reaching a minimum statistical post-treatment difference is a necessary but not sufficient condition of a clinically important difference. An alternative methodology attempts to address these concerns by comparing a change in PROM to a “gold standard” measure of health status. This is known as an anchor-based MCID calculation.[23,29] This investigation attempted to use the SF-6D as an anchor by which to compare changes in SNOT-22 total and domain scores. While statistically significant relationships were noted between improvements in the SF-6D and the SNOT-22, the ROC analysis showed poor diagnostic accuracy across SNOT-22 total and domain scores, thus limiting its use as an anchor measure for MCID calculation. Fundamentally, the poor diagnostic accuracy of the AUC analysis is a sign that, while the SNOT-22 and its domain scores are appropriate measures of disease-specific symptom severity, the questionnaire does not necessarily predict overall health utility well. Put another way, an improvement in SNOT-22 scores alone may not change health utility as captured by the SF-6D. In a clinical context, this makes sense, as many of the questions on utility questionnaires are designed to capture systemic quality-of-life changes that may not be measured by a disease-specific questionnaire. A separate consideration is that even a successful anchor-based ROC analysis with greater AUC values would provide a wide range of possible MCID values with respective sensitivities and specificities, and several valid definitions for optimal MCID values. Future work within our group will continue to define these relationships between PROMs to determine if alternative disease-specific questionnaires are better equipped to reflect improvements in global health utility than others.

Having estimated the MCID for the five SNOT-22 subdomains, we next turn our attention to the practical utility and interpretation of these values. They are of particular relevance to the field of rhinologic outcomes research, as they enable researchers to compare post-treatment changes in domain scores against a minimum threshold to see if statistically significant improvements are actually clinically impactful. For example, a recent article[3] investigating improvements in psychological dysfunction following ESS found that the mean improvement was 7.4 points. While this was a statistically significant result, it is difficult to determine if 7.4 points on a 35-point continuum is a discernible result for an individual patient or just statistical “noise” within the range of patient variation. With the findings of this investigation estimating the psychological domain MCID as 3.9, we have some guidance that, in properly selected patients with CRS, ESS can actually provide clinically meaningful improvements in discernible psychological function. This in turn allows clinicians to

properly counsel patients about the potential benefits of surgery in the preoperative setting. In addition, as the breadth of clinical treatment options in rhinology grows, having MCID values for the SNOT-22 domains will allow for a better assessment of how a particular intervention impacts quality-of-life. To illustrate this point, consider a hypothetical scenario where two competing interventions for CRS have similar total SNOT-22 improvements, but only one achieves an MCID of improvement in the sleep dysfunction domain. When selecting the appropriate clinical option for a CRS patient who has significant impairment in sleep dysfunction, the clinician may be more inclined to select the one that will achieve an MCID in the sleep domain, even if it may not do so in the other symptom domains.

There are several limitations to discuss within the context of these findings. First, this analysis is based off SNOT-22 score distributions from a sample of patients with medically refractory CRS enrolled from academic, tertiary care centers. The MCID values presented here may not be generalizable to SNOT-22 score improvements obtained in patients with other sinonasal diagnoses or in other clinical environments. Additionally, this analysis included only patients who underwent surgical management of CRS. This was a conscious decision made to mirror the approach of previous studies which define the MCID for SNOT-22 total scores. In reality, due to the nature of the distribution based methodology, there may be alternative MCID values for medical management of CRS just as there may be different values for other sinonasal diagnoses and even subgroups of patients within our cohort. This represents another gap within the current literature that we are aiming to address separately. Of course, it may be that there is some level of discriminative capability that is inherent within the SNOT-22 that makes these values close enough to each other that a general value is obtained for all disease states and treatment options for CRS. This transitions to a general limitation of the MCID as a metric of minimum change – ultimately, it is a useful heuristic rather than an absolute value applicable in all instances. In our analysis health utility as measured by the SF-6D did not provide sufficient diagnostic utility to perform a robust anchor-based analysis, in contrast to prior work by our group investigating the Brief Smell Identification Test (BSIT), in which a successful anchor-based approach was used to confirm the distribution-based findings.[30] Nevertheless, having a frame of reference to interpret the clinical significance of changes in PROM scores has significant utility in outcomes research, and the previously discussed agreement of our mean scores with widely accepted anchor-based measures of the MCID lends credibility to the distribution approach in this particular case and offers a useful benchmark to interpret domain score changes.

CONCLUSION

Similar to currently published thresholds, the mean MCID of the SNOT-22 total score in this cohort using distribution-based methods was approximately 9.0. The average MCID for the rhinologic domain, extra-nasal rhinologic domain, ear/facial domain, psychological domain, and sleep domain score was 3.8, 2.4, 3.2, 3.9, and 2.9, respectively. Anchor-based methods using the SF-6D did not have strong predictive accuracy across SNOT-22 total scores or individual domain scores.

Acknowledgments

Financial Disclosures: Timothy L. Smith, Jeremiah A. Alt, Todd E. Bodner, 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; PI: T. Smith). Public clinical trial registration (www.clinicaltrials.gov) ID# NCT02720653. This funding organization did not contribute to the design or conduct of this study; preparation, review, approval or decision to submit this manuscript for publication. There are no relevant financial disclosures for Naweed I. Chowdhury or Joshua M. Levy. Adam S. DeConde is a consultant for Intersect ENT, Stryker Endoscopy and Olympus, none of which are affiliated with this investigation.

References

1. Hopkins C, Gillett S, Slack R, Lund VJ, Browne JP. Psychometric validity of the 22-item Sinonasal Outcome Test. Clin Otolaryngol. 2009; 34(5):447–454. [PubMed: 19793277]
2. DeConde AS, Bodner TE, Mace JC, Smith TL. 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]
3. Levy JM, Mace JC, DeConde AS, Steele TO, Smith TL. Improvements in psychological dysfunction after endoscopic sinus surgery for patients with chronic rhinosinusitis. Int Forum Allergy Rhinol. 2016; 6(9):906–913. [PubMed: 27080574]
4. Alt JA, Mace JC, Smith TL, Soler ZM. Endoscopic sinus surgery improves cognitive dysfunction in patients with chronic rhinosinusitis. Int Forum Allergy Rhinol. 2016; 6(12):1264–1272. [PubMed: 27384037]
5. Chowdhury NI, Mace JC, Smith TL, Rudmik L. What drives productivity loss in chronic rhinosinusitis? A SNOT-22 subdomain analysis. Laryngoscope. Jun.2017 doi: 10.1002/lary.26723
6. Smith TL, Mace JC, Rudmik L, et al. Comparing surgeon outcomes in endoscopic sinus surgery for chronic rhinosinusitis. Laryngoscope. 2017; 127(1):14–21. [PubMed: 27298069]
7. Levy JM, Mace JC, Rudmik L, Soler ZM, Smith TL. Low 22-item sinonasal outcome test scores in chronic rhinosinusitis: Why do patients seek treatment? Laryngoscope. 2017; 127(1):22–28. [PubMed: 27377575]
8. DeConde AS, Mace JC, Levy JM, Rudmik L, Alt JA, Smith TL. Prevalence of polyp recurrence after endoscopic sinus surgery for chronic rhinosinusitis with nasal polyposis. Laryngoscope. 2017; 127(3):550–555. [PubMed: 27859303]
9. Rosenfeld RM, Piccirillo JF, Chandrasekhar SS, et al. Clinical Practice Guideline (Update): Adult Sinusitis. Otolaryngol Head Neck Surg. 2015; 152(2 Suppl):S1–S39.
10. Fokkens WJ, Lund VJ, Mullol J, et al. European Position Paper on Rhinosinusitis and Nasal Polyps 2012. Rhinol Suppl. 2012; (23):1–298. 3 p preceding table of contents.
11. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline. Guideline for Good Clinical Practice E6 (R1). Available at: http://www.ich.org/products/guidelines/
12. DeConde AS, Mace JC, Bodner T, et al. SNOT-22 quality of life domains differentially predict treatment modality selection in chronic rhinosinusitis. Int Forum Allergy Rhinol. 2014; 4(12):972–979. [PubMed: 25323055]
13. DeConde AS, Mace JC, Alt JA, Rudmik L, Soler ZM, Smith TL. 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]
14. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ. 2002; 21(2):271–292. [PubMed: 11939242]
15. Walters SJ, Brazier JE. What is the relationship between the minimally important difference and health state utility values? The case of the SF-6D. Health Qual Life Outcomes. 2003; 1:4. [PubMed: 12737635]
16. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care. 2003; 41(5):582–592. [PubMed: 12719681]

17. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951; 16(3):297–334.
18. Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. Med Care. 1999; 37(5):469–478. [PubMed: 10335749]
19. Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. J Clin Epidemiol. 1999; 52(9):861–873. [PubMed: 10529027]
20. Beaton DE. Understanding the relevance of measured change through studies of responsiveness. Spine (Phila Pa 1976). 2000; 25(24):3192–3199. [PubMed: 11124736]
21. Cohen, J. Statistical Power Analysis for the Behavioral Sciences. New York, NY: Routledge Academic; 1988.
22. Samsa G, Edelman D, Rothman ML, Williams GR, Lipscomb J, Matchar D. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. Pharmacoeconomics. 1999; 15(2):141–155. [PubMed: 10351188]
23. Copay AG, Subach BR, Glassman SD, Polly DW, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. Spine J. 2007; 7(5):541–546. [PubMed: 17448732]
24. van der Roer N, Ostelo RWJG, Bekkering GE, van Tulder MW, de Vet HCW. Minimal Clinically Important Change for Pain Intensity, Functional Status, and General Health Status in Patients With Nonspecific Low Back Pain. Spine (Phila Pa 1976). 2006; 31(5):578–582. [PubMed: 16508555]
25. YOUDEN WJ. Index for rating diagnostic tests. Cancer. 1950; 3(1):32–35. [PubMed: 15405679]
26. Schünemann HJ, Puhan M, Goldstein R, Jaeschke R, Guyatt GH. Measurement properties and interpretability of the Chronic respiratory disease questionnaire (CRQ). COPD. 2005; 2(1):81–89. [PubMed: 17136967]
27. MILLER GA. The magical number seven plus or minus two: some limits on our capacity for processing information. Psychol Rev. 1956; 63(2):81–97. [PubMed: 13310704]
28. Farivar SS, Liu H, Hays RD. Half standard deviation estimate of the minimally important difference in HRQOL scores? Expert Rev Pharmacoecon Outcomes Res. 2004; 4(5):515–523. [PubMed: 19807545]
29. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989; 10(4):407–415. [PubMed: 2691207]
30. Levy JM, Mace JC, Bodner TE, Alt JA, Smith TL. 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]

Table 1

Preoperative characteristics and comorbidity in patients with CRS (n=276)

Characteristics:	Mean [±SD]	Range	N (%)
Age at enrollment (years)	52.6 [±14.8]	(18 - 86)	-----
Males *	-----	-----	129 (47%)
White/Caucasian	-----	-----	236 (86%)
African American	-----	-----	10 (4%)
Asian	-----	-----	10 (4%)
Hispanic/Latino	-----	-----	17 (6%)
Nasal polyposis	-----	-----	96 (35%)
Turbinate hypertrophy	-----	-----	43 (16%)
Asthma	-----	-----	95 (35%)
AERD/ASA intolerance	-----	-----	18 (7%)
Allergic rhinitis ^	-----	-----	123 (45%)
Depression *	-----	-----	44 (16%)
Current tobacco use/smoking	-----	-----	7 (3%)
Current alcohol use	-----	-----	119 (43%)
Diabetes mellitus (Type I / II)	-----	-----	16 (6%)
SNOT-22 total score	52.7 [±19.5]	(4 – 106)	-----
Rhinologic symptoms domain	16.4 [±6.3]	(0 – 30)	-----
Extra-nasal rhinologic symptoms domain	8.6 [±3.6]	(0 – 15)	-----
Ear/facial symptoms domain	9.0 [±4.8]	(0 – 23)	-----
Psychological dysfunction domain	15.8 [±8.2]	(0 – 35)	-----
Sleep dysfunction domain	13.5 [±6.8]	(0 – 25)	-----
SF-6D Health Utility score	0.69 [±0.15]	(0.35 – 1.00)	-----

* identified through self-report;

^ confirmed via modified radioallergosorbent or skin prick testing; CRS, chronic rhinosinusitis; SD, standard deviation; AERD, aspirin exacerbated respiratory disease; ASA, acetylsalicylic acid; SNOT-22, 22-item SinoNasal Outcome Test. SF-6D, Medical Outcomes Study Short Form 6D survey

Table 2

Mean postoperative improvement in SNOT-22 and SF-6D scores (n=276)

	12-month postoperative	Δ value		
	Mean [±SD]	Mean [±SD]	Paired t-testing statistic	p-value
SNOT-22 total score	26.0 [±20.2]	−26.7 [±22.3]	19.8	<0.001
Rhinologic symptoms domain	8.0 [±6.9]	−8.4 [±8.1]	17.0	<0.001
Extra-nasal rhinologic symptoms domain	4.4 [±3.6]	−4.2 [±4.0]	17.3	<0.001
Ear/facial symptoms domain	4.1 [±4.2]	−4.9 [±5.0]	16.2	<0.001
Psychological dysfunction domain	7.8 [±7.9]	−8.0 [±8.3]	15.9	<0.001
Sleep dysfunction domain	7.2 [±6.7]	−6.4 [±7.0]	15.0	<0.001
SF-6D Health Utility score	0.77 [±0.15]	0.08 [±0.14]	−9.4	<0.001

SNOT-22, 22-item SinoNasal Outcome Test; SF-6D, Medical Outcomes Study Short Form 6D survey; SD, standard deviation.

Table 3

Results of distribution-based methods for determining MCID values for SNOT-22 scores

	MCID Value Determinations				Mean MCID
Preoperative measures:	**0.5 standard deviation**	**Standard error of measurement (SEM)**	**Cohens effect size (d)***	**Minimum detectable change (MDC)**	
SNOT-22 total score	9.8	5.9	3.9	16.4	**−9.0**
Rhinologic symptoms	3.2	2.8	1.3	7.8	**−3.8**
Extra-nasal rhinologic symptoms	1.8	1.9	0.7	5.3	**−2.4**
Ear/facial symptoms	2.4	2.5	1.0	6.9	**−3.2**
Psychological dysfunction	4.1	2.6	1.6	7.2	**−3.9**
Sleep dysfunction	3.4	1.8	1.4	5.0	**−2.9**

MCID, minimal clinically important difference; SNOT-22, 22-item SinoNasal Outcome Test

*Cohens effect size (d) determined by multiplying preoperative standard deviation value by 0.20 to reflect an MCID equal to a minimal effect size.

Table 4

Reliability estimations for preoperative SNOT-22 scores.

Preoperative SNOT-22 total and domains:	Cronbach's alpha reliability estimate (R)
SNOT-22 total score	0.91
Rhinologic symptoms	0.80
Extra-nasal rhinologic symptoms	0.71
Ear/facial symptoms	0.72
Psychological dysfunction	0.90
Sleep dysfunction	0.93

SNOT-22, 22-item SinoNasal Outcome Test

Table 5

Results of receiver operating characteristic (ROC) curve methods for determining MCID values for SNOT-22 scores

SNOT-22 total and domains:	AUC [±SE]	95% CI	Z-statistic
SNOT-22 total score	0.70 [±0.03] *	(0.64 – 0.75)	6.09
Rhinologic symptoms	0.64 [±0.03] *	(0.58 – 0.70)	3.97
Extra-nasal rhinologic symptoms	0.62 [±0.03] *	(0.56 – 0.68)	3.55
Ear/facial symptoms	0.63 [±0.03] *	(0.57 – 0.69)	3.76
Psychological dysfunction	0.71 [±0.03] *	(0.65 – 0.76)	6.60
Sleep dysfunction	0.68 [±0.03] *	(0.62 – 0.73)	5.36

* reflects AUC values significantly different from 0.5 ($p<0.001$); AUC, area under the ROC curve; ROC, receiver operating characteristic; SNOT-22, 22-item SinoNasal Outcome Test; SE, standard error; CI, confidence inteval