Published in final edited form as:

Int Forum Allergy Rhinol. 2017 February ; 7(2): 128–134. doi:10.1002/alr.21866.

Evaluating Metrics of Responsiveness Using Patient-Reported Outcome Measures in Chronic Rhinosinusitis

Alcina K. Lidder, BA¹, Kara Y. Detwiller, MD², Caroline P.E. Price, BA², Robert C. Kern, MD^{2,3}, David B. Conley, MD², Stephanie Shintani-Smith, MD, MS², Kevin C. Welch, MD², Rakesh K. Chandra, MD⁴, Anju T. Peters, MD³, Leslie C. Grammer III, MD³, Li-Xing Man, MSc, MD, MPA^{1,5}, Robert P. Schleimer, PhD^{2,3}, and Bruce K. Tan, MD, MS¹

¹University of Rochester School of Medicine and Dentistry, University of Rochester Medical Center, Rochester, NY, USA

²Department of Otolaryngology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

³Division of Allergy-Immunology, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

⁴Department of Otolaryngology, Vanderbilt University Medical Center, Nashville, TN, USA

⁵Department of Otolaryngology Head and Neck Surgery, University of Rochester Medical Center, Rochester, NY, USA

Abstract

Background—Responsiveness, or sensitivity to clinical change, is important when selecting patient-reported outcome measures (PROMs) for research and clinical applications. This study compares responsiveness of PROMs used in chronic rhinosinusitis (CRS) to inform the future development of a highly responsive instrument that accurately portrays CRS patients' symptom experiences.

Methods—Adult CRS patients initiating medical therapy (MT: n=143) or undergoing endoscopic sinus surgery after failing MT (ESS: n=123) completed the Sino-Nasal Outcome Test (SNOT)-22, European Position Statement on Rhinosinusitis (EPOS) visual analog scale (VAS), and Patient-Reported Outcomes Measurement Information System (PROMIS)-29 at baseline and 3 months following treatment. Cohen's d and paired t statistics were used to evaluate the responsiveness of each measure.

Results—Respectively, 52 (36.4%) subjects and 42 (34.1%) subjects in the MT and ESS groups completed baseline and 3-month questionnaires. Subjects with and without 3-month data were similar with respect to baseline demographics, VAS scores, and SNOT-22 scores ($p>0.05$). In MT patients, CRS-specific measures like VAS ($d=-0.58$, $p<.01$; $t=-1.81$, $p>.05$) and SNOT-22 scores ($d=-0.70$, $p<.01$; $t=-3.29$, $p<.05$) were more responsive than PROMIS-29 general health domains ($p>0.05$ for Cohen's d). In ESS patients, VAS ($d=-1.97$; $t=-9.63$, both $p<.01$) and SNOT-22 scores

($d=-1.56$; $t=-9.99$, both $p<.01$) were similarly more responsive although changes in PROMIS-29 Fatigue ($d=-0.82$, $p=.01$; $t=-4.63$, $p<.01$); Sleep Disturbance ($d=-0.83$; $t=-3.77$, both $p<.01$); and Pain Intensity ($d=-1.0$; $t=-5.67$, both $p<.01$) domains were significant. All 22 individual SNOT-22 items differed significantly following surgery whereas only 8 items were consistently responsive after MT.

Conclusions—For both MT and ESS patients, CRS-specific PROMs are more responsive to post-treatment clinical changes than general health measures. Still, the SNOT-22 contains items that likely decrease its overall responsiveness. Our findings also indicate that existing PROMs show greater response to ESS than MT.

Keywords

Chronic rhinosinusitis; paranasal sinus diseases; medical therapy of rhinosinusitis; sinus surgery; FESS; patient reported outcome measure; SNOT-22

INTRODUCTION

Patient-reported outcomes measures (PROMs) are essential tools for evaluating patients' assessments of their health status and well-being. The strengths of these tools are assessed through reliability, validity, responsiveness, and ease of use. Responsiveness is a measure's sensitivity to clinical change and is an important consideration when selecting PROMs for clinical research purposes. There are two types of responsiveness: internal and external. Internal responsiveness is the ability of a metric to change over a specific time period (e.g.: before and after an efficacious therapy) while external responsiveness is the degree to which changes in an instrument correspond to changes in an external reference metric of health status.¹ A highly responsive instrument allows sensitive measurement of health status changes after an intervention, provides a meaningful understanding of the magnitude of that change, and reduces the sample size required for clinical trials. The responsiveness of PROMs in chronic rhinosinusitis (CRS), particularly in medically treated patients, has not been thoroughly examined, limiting our understanding of the extent to which CRS PROMs portray post-treatment changes.

There are multiple CRS-specific PROMs and significant variation exists in the PROMs used in studies evaluating the medical management of CRS. Furthermore, acceptable responsiveness has only been determined for the Rhinosinusitis Outcome Measure (RSOM-31)² and its derivatives the Sino-Nasal Outcome Test (SNOT)-16³, 20⁴, and 22⁵; Rhinosinusitis Quality of Life survey⁶; and EuroQol 5-Dimension (EQ-5D)^{7,8}. The European Position Statement on Rhinosinusitis (EPOS) guidelines recommend using a visual analog scale (VAS) to assess symptom burden.^{9,10} However, few studies have used the VAS.¹¹ Alternatively, several trials have used the SNOT-20 or 22^{4,11,12}, RhinoSinusitis Disability Index (RSDI)^{13,14}, Chronic Sinusitis Survey (CSS)¹⁵, RSOM-31,¹⁶ or unique PROMs.^{17–19}

CRS-specific PROMs are restricted in their ability to compare health-related quality of life (QOL) across diseases. Thus the degree of a symptom, such as fatigue, experienced by a patient with CRS cannot be readily compared to the fatigue experienced by a patient with

asthma or individuals in the general population. To overcome this limitation, general QOL measures such as the Short Form Health Questionnaire-6D or -36 (SF-6D, SF-36) and EuroQol 5-Dimension (EQ-5D) have been used to evaluate the medical and surgical treatment of patients with CRS.^{7,8,11,16,20–22}

Recently, a large National Institutes of Health grant helped develop the Patient-Reported Outcomes Measurement Information System (PROMIS) to evaluate general QOL and facilitate comparisons between various chronic disease states.^{23,24} The PROMIS instruments have been validated in many chronic conditions.^{24,25} These instruments can be utilized through item banks or short forms such as the PROMIS-29 questionnaire, which contains eight QOL domains. Unlike the SF-36, the PROMIS-29 provides a built-in means of measuring external responsiveness.²⁶ The flexibility of this instrument, along with its ability to compare scores across chronic diseases and thereby facilitate comparative effectiveness studies makes it an attractive PROM for studying CRS. We have previously reported PROMIS-29 data in a study comparing symptoms and QOL between CRS subtypes and endotypes in a surgical population²⁷ but we are unaware of studies evaluating its responsiveness for measuring change following medical or surgical therapy for CRS.

The current study aims to examine the internal responsiveness of SNOT-22, VAS, and PROMIS-29 QOL domains in CRS patients undergoing standard-of-care medical and surgical treatment. By evaluating the responsiveness of these PROMs, we seek to evaluate the measurement characteristics of existing PROMs used in CRS to identify responsive items and/or instruments. This information will be utilized to inform the future development of a CRS PROM that follows Food and Drug Administration (FDA) guidelines for PROMs used in clinical trials. Since there are currently no FDA-approved measures for use in CRS research, such a PROM will advance the care of CRS patients as a responsive patient-defined measure of disease status.

MATERIALS AND METHODS

Subjects

CRS patients aged 18 to 89 without endoscopic sinus surgery (ESS) in the past 18 months were recruited from a tertiary care center before initiating medical therapy (MT) or undergoing ESS. All patients were diagnosed with CRS according to American Academy of Otolaryngology guidelines.²⁸ All ESS patients previously failed a course of maximal medical therapy. Patients were excluded if they had a prior diagnosis of recurrent acute sinusitis or if they were unable or unwilling to provide informed consent for any reason. This study was approved by the institutional review board (IRB) at Northwestern University.

PROMs studied

Subjects completed the EPOS VAS (VAS)¹⁰, SNOT-22⁵, and PROMIS-29 (v1.0) at baseline and 3 months following treatment. The VAS is a 10 centimeter (cm) self-reported measure that evaluates overall symptom severity with the question, “how bothersome are your symptoms of rhinosinusitis,” where 0 represents “not bothersome” and 10 represents “worst thinkable troublesome.” The raw score from 0 to 10 cm was converted to a 0 to 100 scale

(i.e.: a score of 6.7 was converted to 67). Each PROMIS-29 domain is defined such that a higher score represents more of the construct being measured (i.e.: a higher Fatigue score signifies greater fatigue whereas a higher Physical Function score represents greater physical functioning.) The score for each domain (except Pain Intensity) is calibrated such that a score of 50 represents the census-representative general population mean and each domain has a standard deviation of 10. Subjects self-reported demographic characteristics and medical history items including, but not limited to: age, gender, race, prior ESS, and a history of asthma, allergies, aspirin intolerance, and Aspirin Exacerbated Respiratory Disease (AERD). If a subject marked “unsure” for these demographic items, answers were confirmed with a medical record review.

Treatment

MT subjects received standard-of-care treatment including antibiotics, oral steroids, and intranasal steroids tailored based on clinician and patient preferences, as is clinical practice. Similarly, the extent of surgery was determined by the treating physician and reflected disease severity on an individual patient basis. All surgical cases were followed by standard postoperative regimens including saline irrigations, topical steroids, antibiotics and/or oral steroids per the treating physician’s recommendations.

Analysis

Baseline and 3-month data for mean VAS score, total SNOT-22 score, PROMIS-29 domain scores, and individual SNOT-22 items were analyzed. We additionally defined the sum of the individual SNOT-22 symptoms of “need to blow nose,” “runny nose,” “post-nasal drip,” “thick nasal discharge,” “sense of taste/smell,” “facial pain/pressure,” and “blockage/congestion of nose” as a SNOT-cardinal score because these symptoms are thematically similar to the cardinal symptoms that define CRS. Likewise, the sum of SNOT-22 items not included in the SNOT-cardinal score was used to define a SNOT-other score. CRS-specific scores and PROMIS general QOL domains were used to calculate two metrics of responsiveness: Cohen’s d and the paired t-statistic. Cohen’s d is an effect size statistic that conveys the amount of change in a variable and is defined as $(M_2 - M_1) / (SD)$ where M represents the mean and SD is the standard deviation of the baseline mean. There are conventional standards used to interpret effect sizes: a value of <0.20 is regarded as small, 0.50 is considered moderate, and 0.80 is deemed large.¹ In our study, statistical significance for the Cohen’s d was determined by the ANOVA test associated with the mean difference for each measure. All baseline data for each group was compared to available 3-month data. The paired t-statistic is also used as a metric of internal responsiveness. By definition, this analysis was restricted to patients for whom we had both baseline and 3-month data. The paired t-test assesses the null hypothesis that there was no change in the average response of a variable during a specific time period. With a sample size over 30, a paired t-statistic value >1.96 supports rejection of the null hypothesis, thereby suggesting a statistically significant change in the measure. In this case, the measure is considered responsive.¹ A $p < 0.05$ was considered statistically significant for all analyses.

RESULTS

Baseline demographics

A total of 266 patients enrolled in the study with 143 in the MT cohort and 123 in the ESS cohort. Fifty-two (36.4%) subjects and 42 (34.1%) subjects in the MT and ESS groups, respectively, completed both baseline and 3-month questionnaires. Subjects with and without 3-month data were similar with respect to baseline age, sex, polyp status, VAS scores, and SNOT-22 total scores ($p>0.05$; Supplementary table 1). Demographic characteristics of participating patients can be found in Table 1. Of those who completed the 3-month questionnaires in the MT group, MT was comprised of initiating an extended course of antibiotics only (17.3 %), antibiotics and intranasal steroids (30.8%), or antibiotics, intranasal steroids, and a course of oral steroids (51.9%).

Comparison of baseline PROM measures

Baseline VAS scores, SNOT-22 scores, and PROMIS-29 QOL domain scores for all subjects are presented in Table 2. VAS scores, SNOT-22 total scores, and SNOT-cardinal scores were significantly different between MT and ESS groups at baseline, with ESS patients reporting greater symptom burden. Of the PROMIS-29 QOL domains, ESS patients reported significantly more baseline Pain Intensity than MT patients.

Internal responsiveness, Cohen's d

In the MT group, CRS-specific measures like VAS ($d=-0.58$, $p<.01$) and SNOT-22 total scores ($d=-0.70$, $p<.01$) significantly changed and demonstrated increased responsiveness compared to PROMIS-29 general QOL domains ($p>0.05$; Table 3). Using the SNOT-cardinal score increased responsiveness further ($d=-0.83$, $p<.01$) while the SNOT-other items were less responsive ($d=-0.52$, $p<.05$). Thirteen of the 22 individual SNOT-22 items were significantly different in MT subjects with greatest responsiveness associated with "runny nose" ($d=-0.72$, $p<.01$) and "blockage/congestion of nose" ($d=-0.69$, $p<.01$; Table 4). The five most and least responsive SNOT-22 items defined by Cohen's d are shown in Table 4.

In the ESS group, VAS ($d=-1.97$, $p<.01$) and SNOT-22 scores ($d=-1.56$, $p<.01$) were highly responsive. Similar to the MT group, SNOT-cardinal score was more responsive than the SNOT-other score ($d=-1.88$; $d=-1.19$, $p<.01$ respectively). Additionally, all 22 SNOT-22 items differed significantly with "blockage/congestion of nose" ($d=-2.16$, $p<.01$) and "need to blow nose" ($d=-1.16$, $p<.01$) as the most responsive. The five most and least responsive SNOT-22 symptoms defined by Cohen's d results are included in Table 5. For ESS subjects, PROMIS-29 domains of Fatigue ($d=-0.82$, $p=.01$), Sleep Disturbance ($d=-0.83$, $p<.01$), and Pain Intensity ($d=-1.0$, $p<.01$) were also significantly different and showed large effects following surgery (Table 6). In both MT and ESS cohorts, VAS, SNOT-cardinal, and SNOT-22 total scores were more responsive than PROMIS-29 general QOL domains.

Internal responsiveness, Paired t-statistic

When paired t-tests were used to evaluate responsiveness in medically-treated patients, the SNOT-22 total score ($t=-3.29$, $p<.05$), SNOT-cardinal ($t=-4.30$, $p<.01$), and SNOT-other

were significantly responsive, $t=-2.38$, $p<.05$ (Table 3). The PROMIS-29 domain of Fatigue was also responsive in the MT group, $t=-2.20$, $p<.05$ (Table 6). The most responsive individual SNOT-22 items included “runny nose” ($t=-4.17$, $p<.01$) and “blockage/congestion of nose,” $t=-3.92$, $p<.01$.

In ESS patients, VAS ($t=-9.63$), SNOT-22 total score ($t=-9.99$), SNOT-cardinal ($t=-9.72$), and SNOT-other ($t=-9.63$) were significantly responsive using the t-statistic (all $p<.01$; Table 3). All PROMIS-29 domains except Depression were responsive to post-surgical changes (Table 6). Of the general QOL domains, Pain Intensity ($t=-6.94$), Pain Interference ($t=-5.67$), and Fatigue ($t=-4.63$), demonstrated the greatest responsiveness (all $p<.01$; Table 6). The individual SNOT-22 symptoms of “blockage/congestion of nose” ($t=-9.88$) and “post-nasal discharge” ($t=-7.20$) were most responsive in surgical patients (all $p<.01$).

DISCUSSION

Both responsiveness metrics demonstrated that CRS-specific PROMs were more responsive than general QOL domains for MT and ESS patients. Although some PROMIS-29 domains showed significant post-treatment changes, none of the domains approached the magnitude of response observed with the SNOT-22 total, SNOT-cardinal, or EPOS VAS scores. The SNOT-22 total score was more responsive in MT patients compared to the VAS score but the opposite was true in ESS patients, suggesting that both are reasonably sensitive for measuring treatment changes. While this study does not seek to directly compare the effects of ESS with MT since the cohorts of patients were different, we note that all PROMs showed larger effects following ESS than MT. This is consistent with prior studies showing greater post-treatment changes in ESS patients.²⁹ Interestingly, only the ESS group had significant improvements in PROMIS-29 domains as measured by both Cohen’s d and the t-statistic, suggesting that most general QOL domains within the PROMIS-29 are not sufficiently sensitive to evaluate treatment effects in the MT group. Still, the domain of Fatigue was adequately responsive to measure post-treatment changes in medically-treated patients using the paired t-test analysis.

Our findings also suggest that using multiple items on the SNOT-22 improves responsiveness given that most individual items were less responsive than the SNOT-22 total score. However, among the 22 symptoms, not all items were equally responsive. In particular, both responsiveness metrics demonstrated that SNOT-cardinal symptoms are more responsive than SNOT-other items. For example, “blockage/congestion of nose” was consistently more responsive to treatment than non-rhinologic items within the SNOT-22, especially compared to items such as “sad,” “embarrassed,” and “dizziness. Interestingly, despite its sensitivity to treatment changes, “blockage/congestion of nose” was removed when the RSOM-31 evolved into the SNOT-20 but was reinstated in the SNOT-22. In our study, the wording of thematically similar items had effects on responsiveness with items such as “thick nasal discharge” showing lesser responsiveness to treatment changes than similar items such as “post nasal drip” and “runny nose”. Subsequently, these less responsive items showed no significant differences following MT, an intervention with a relatively small effect size, but became significant following ESS, an intervention with a large effect size.

Indeed, other studies that used existing PROMs may have encountered similar limitations. For instance, a randomized controlled trial evaluating the efficacy of omalizumab in patients with CRS with nasal polypsis found significant improvements in nasal polyp scores, “nasal congestion,” and “sense of smell” but could not detect significant differences in the nasal domain of the RSOM-31.¹¹ Another randomized placebo-controlled trial of long-term, low-dose roxithromycin found statistically significant improvements in SNOT-20 total scores with twelve weeks of antibiotic therapy but these benefits were not sustained 3 months after the conclusion of treatment despite 4 to 10 point differences (out of 100) in SNOT-20 total scores between treatment and placebo groups.¹²

General QOL instruments are essential for clinical and research applications particularly because they measure health domains relevant to many chronic diseases. Thus far, the SF-36 and EQ-5D have been the principal general QOL measures used to study CRS patients in the United States.^{7,30–33} Overall, like the SNOT-other items, we found PROMIS-29 domains less responsive than CRS-specific measures such as the VAS, SNOT-22 total, and SNOT-22-cardinal. The PROMIS-29 domains were not significantly responsive to MT except for Fatigue when measured using the paired t-statistic. In the ESS group, we consistently detected significant improvements in the PROMIS-29 domains of Fatigue, Sleep Disturbance, and Pain Intensity using ANOVA mean differences and calculated Cohen’s d values as well as paired t-tests. Using paired t-tests we additionally detected significant responsiveness in the domains of Anxiety, Satisfaction with Social Role, Pain Intensity, and Physical Function.

The lesser responsiveness of general QOL measures is consistent with a study of ethmoid sinus surgery patients that used the CSS and SF-36 to assess changes 3 months post-operatively. While effect sizes for individual CSS items ranged from 0.20 ($p=.01$) to 1.02 ($p<.01$), with an effect size of 1.12 for the CSS total score ($p<.01$), effect sizes for the SF-36 domains ranged from 0.01 ($p>.05$) to 0.52 for both Physical Functioning and Bodily Pain domains (both $p=.01$).³² The greater responsiveness for CSS items compared to QOL domains contained within the SF-36 suggest that the CRS-specific instrument was more sensitive to change than the SF-36.

Our PROMIS-29 findings also suggest that not all general health domains are relevant to CRS patients. For example, CRS patients showed little baseline impairment in Physical Function, Anxiety, Depression, and Satisfaction with Social Role. Subsequently, these domains did not show large post-treatment changes although some were significant when measured by the paired t-statistic. Conversely, Fatigue, Sleep Disturbance, and Pain Intensity showed greater baseline impairment relative to the general population and improved significantly with treatment using both metrics of responsiveness. Therefore, if an investigational medication or device seeks approval for use in CRS patients, our findings suggest that a PROM containing only CRS-symptoms would enhance its responsiveness compared to PROMs that combine CRS symptoms with general QOL items. In addition, the decision to include general QOL items related to fatigue or sleep disturbance should be driven by careful considerations of potential labeling claims and whether the investigational treatment will cause sufficient change to demonstrate significant differences in less responsive health domains.

Our study has several limitations. We recognize that the MT and ESS groups are distinct and thus the values of responsiveness are not direct comparisons. Moreover, reflecting the realistic standard-of-care nature of our study, we had relatively low 3-month completion rates of 36.4% and 34.1% for the MT and ESS groups, respectively. Thus, patients who were lost to follow-up may represent a different group than those who reported 3-month outcomes. Furthermore, eleven patients in the MT group chose to undergo ESS within 3 months of initiating MT and were considered lost to follow-up. If one assumes that these patients had no response to MT, our estimate of response to MT may represent an overestimate. Nonetheless, responders and non-responders were not significantly different in terms of baseline demographics and baseline symptom burden using VAS scores and SNOT-22 total scores. Additionally, using both paired and unpaired data for responsiveness calculations showed the same items to be highly responsive. Another limitation is that MT was not standardized; still, the modalities utilized are congruent with current treatment guidelines.³⁴ Finally, it is nearly impossible to differentiate between measurement responsiveness and actual treatment efficacy. We chose MT and ESS as studied treatment modalities because these represent two commonly used regimens in CRS patients and we sought to evaluate the performance of PROMs in standard-of-care clinical settings. Nonetheless, we believe this study provides a reasonable estimate of the response to medical and surgical treatment observed in patients treated using current practices.

CONCLUSION

In evaluating changes following MT and ESS for CRS patients, the most responsive measures were CRS-specific PROMs such as the SNOT-22 total scores and EPOS VAS scores. Within the SNOT-22, certain items like the SNOT-cardinal score, “blockage/congestion of nose,” and “runny nose” were highly responsive to both MT and ESS while others, such as “sad,” “embarrassed,” and “dizziness” were not responsive. General QOL items like the SNOT-other composite score and individual PROMIS-29 domains of Fatigue, Sleep Disturbance, and Pain Intensity were generally less responsive and could detect consistent changes following ESS but not following MT. Thus the presence of non-responsive elements and general QOL items in the SNOT-22 may decrease the instrument’s overall responsiveness, detecting significant changes only after interventions with larger effects such as ESS. Further work is needed to qualitatively evaluate if the most responsive items also reflect patients’ experiences following treatment.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Triological Society/ American College of Surgeons (B.K.T); NIH grants K23DC012067 (B.K.T), Chronic Rhinosinusitis Integrative Studies Program (CRISP) U19 AI106683 (B.K.T, R.C.K, R.P.S), R01 HL078860, R01 HL068546, R01 AI072570 and R01 AI104733; and the Ernest S. Bazley Foundation.

REFERENCES

1. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol.* 2000; 53(5):459–468. [PubMed: 10812317]
2. Piccirillo JF, Edwards D, Haiduk A, Yonan C, Thawley SE. Psychometric and clinimetric validity of the 31-item rhinosinusitis outcome measure (RSOM-31). *American journal of rhinology.* 1995; 9(6):297–306.
3. Anderson ER, Murphy MP, Weymuller EA Jr. Clinimetric evaluation of the Sinonasal Outcome Test-16. Student Research Award 1998. *Otolaryngol Head Neck Surg.* 1999; 121(6):702–707. [PubMed: 10580223]
4. Piccirillo JF, Merritt MG Jr, Richards ML. Psychometric and clinimetric validity of the 20-Item Sino-Nasal Outcome Test (SNOT-20). *Otolaryngol Head Neck Surg.* 2002; 126(1):41–47. [PubMed: 11821764]
5. 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]
6. Atlas SJ, Metson RB, Singer DE, Wu YA, Gliklich RE. Validity of a new health-related quality of life instrument for patients with chronic sinusitis. *Laryngoscope.* 2005; 115(5):846–854. [PubMed: 15867652]
7. Remenschneider AK, D'Amico L, Gray ST, Holbrook EH, Gliklich RE, Metson R. The EQ-5D: a new tool for studying clinical outcomes in chronic rhinosinusitis. *Laryngoscope.* 2015; 125(1):7–15. [PubMed: 24729050]
8. Rudmik L, Hopkins C, Peters A, Smith TL, Schlosser RJ, Soler ZM. Patient-reported outcome measures for adult chronic rhinosinusitis: A systematic review and quality assessment. *J Allergy Clin Immunol.* 2015; 136(6):1532–1540. e1531–1532. [PubMed: 26654198]
9. Thomas M, Yawn BP, Price D, et al. EPOS Primary Care Guidelines: European Position Paper on the Primary Care Diagnosis and Management of Rhinosinusitis and Nasal Polyps 2007 - a summary. *Prim Care Respir J.* 2008; 17(2):79–89. [PubMed: 18438594]
10. 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]
11. Videler WJ, Badia L, Harvey RJ, et al. Lack of efficacy of long-term, low-dose azithromycin in chronic rhinosinusitis: a randomized controlled trial. *Allergy.* 2011; 66(11):1457–1468. [PubMed: 21884529]
12. Wallwork B, Coman W, Mackay-Sim A, Greiff L, Cervin A. A double-blind, randomized, placebo-controlled trial of macrolide in the treatment of chronic rhinosinusitis. *Laryngoscope.* 2006; 116(2):189–193. [PubMed: 16467702]
13. 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]
14. Benninger MS, Senior BA. The development of the Rhinosinusitis Disability Index. *Arch Otolaryngol Head Neck Surg.* 1997; 123(11):1175–1179. [PubMed: 9366696]
15. Gliklich RE, Metson R. Techniques for outcomes research in chronic sinusitis. *Laryngoscope.* 1995; 105(4 Pt 1):387–390. [PubMed: 7715384]
16. Gevaert P, Calus L, Van Zele T, et al. Omalizumab is effective in allergic and nonallergic patients with nasal polyps and asthma. *J Allergy Clin Immunol.* 2013; 131(1):110–116. e111. [PubMed: 23021878]
17. Van Zele T, Gevaert P, Holtappels G, et al. Oral steroids and doxycycline: two different approaches to treat nasal polyps. *J Allergy Clin Immunol.* 2010; 125(5):1069–1076. e1064. [PubMed: 20451040]
18. Vaidyanathan S, Barnes M, Williamson P, Hopkinson P, Donnan PT, Lipworth B. Treatment of chronic rhinosinusitis with nasal polyposis with oral steroids followed by topical steroids: a randomized trial. *Ann Intern Med.* 2011; 154(5):293–302. [PubMed: 21357906]

19. Gevaert P, Van Bruaene N, Cattaert T, et al. Mepolizumab, a humanized anti-IL-5 mAb, as a treatment option for severe nasal polyposis. *J Allergy Clin Immunol*. 2011; 128(5):989–995. e981–988. [PubMed: 21958585]
20. Smith TL, Litvack JR, Hwang PH, et al. Determinants of outcomes of sinus surgery: a multi-institutional prospective cohort study. *Otolaryngol Head Neck Surg*. 2010; 142(1):55–63. [PubMed: 20096224]
21. Soler ZM, Wittenberg E, Schlosser RJ, Mace JC, Smith TL. Health state utility values in patients undergoing endoscopic sinus surgery. *Laryngoscope*. 2011; 121(12):2672–2678. [PubMed: 22034223]
22. Soler ZM, Sauer D, Mace J, Smith TL. Impact of mucosal eosinophilia and nasal polyposis on quality-of-life outcomes after sinus surgery. *Otolaryngol Head Neck Surg*. 2010; 142(1):64–71. [PubMed: 20096225]
23. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J Clin Epidemiol*. 2010; 63(11):1179–1194. [PubMed: 20685078]
24. Gershon RC, Rothrock N, Hanrahan R, Bass M, Cella D. The use of PROMIS and assessment center to deliver patient-reported outcome measures in clinical research. *J Appl Meas*. 2010; 11(3):304–314. [PubMed: 20847477]
25. NIH. [Accessed February 1, 2016] PROMIS Validity Studies. <http://www.nihpromis.org/science/validitystudies.aspx>
26. Veehof MM, ten Klooster PM, Taal E, van Riel PL, van de Laar MA. Comparison of internal and external responsiveness of the generic Medical Outcome Study Short Form-36 (SF-36) with disease-specific measures in rheumatoid arthritis. *J Rheumatol*. 2008; 35(4):610–617. [PubMed: 18322989]
27. Thompson CF, Price CP, Huang JH, et al. A pilot study of symptom profiles from a polyp vs an eosinophilic-based classification of chronic rhinosinusitis. *Int Forum Allergy Rhinol*. 2015
28. Rosenfeld RM, Piccirillo JF, Chandrasekhar SS, et al. Clinical practice guideline (update): adult sinusitis. *Otolaryngol Head Neck Surg*. 2015; 152(2 Suppl):S1–S39.
29. 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]
30. Remenschneider AK, D'Amico L, Litvack JR, et al. Long-Term Outcomes in Sinus Surgery: A New Tool for Measuring Health-Related Quality of Life. *Otolaryngol Head Neck Surg*. 2014; 151(1):164–170. [PubMed: 24705224]
31. Gliklich RE, Metson R. The health impact of chronic sinusitis in patients seeking otolaryngologic care. *Otolaryngol Head Neck Surg*. 1995; 113(1):104–109. [PubMed: 7603703]
32. Gliklich RE, Hilinski JM. Longitudinal sensitivity of generic and specific health measures in chronic sinusitis. *Qual Life Res*. 1995; 4(1):27–32. [PubMed: 7711687]
33. Metson RB, Gliklich RE. Clinical outcomes in patients with chronic sinusitis. *Laryngoscope*. 2000; 110(3 Pt 3):24–28. [PubMed: 10718411]
34. Meltzer EO, Hamilos DL. Rhinosinusitis diagnosis and management for the clinician: a synopsis of recent consensus guidelines. *Mayo Clin Proc*. 2011; 86(5):427–443. [PubMed: 21490181]

Table 1

Demographic characteristics of study patients.

	MT (N=143)	ESS (N=123)
Age (Mean, SD)	45.9, 15.3	43.9, 14.2
Sex (% female)	49	53.7
Polyp status (% wNP)	30.8	45.5
Hispanic Ethnicity N (%)	5 (3.5)	10 (8.1)
White N (%)	114 (79.7)	102 (82.9)
Black N (%)	13 (9.1)	11 (8.9)
Asian N (%)	10 (7.0)	4 (3.3)
Native American N (%)	0	1 (0.8)
Other N (%)	0	1 (0.8)
Atopy N (%)	67 (46.9)	69 (56.1)
Active asthma N (%)	39 (27.3)	46 (37.4)
Prior ESS N (%)	29 (20.3)	40 (32.5)
AERD N (%)	6 (4.2)	12 (9.8)

MT=Patients initiating medical therapy, ESS=Patients undergoing endoscopic sinus surgery, wNP=with nasal polyposis, AERD= Aspirin Exacerbated Respiratory Disease

Patients were able to report >1 race and ethnicity options.

Table 2Baseline scores of CRS-specific measures and PROMIS-29 domains^x in MT and ESS groups

Baseline	MT (N=143)	ESS (N=123)	P value
	Mean (SD)	Mean (SD)	
VAS	60.1 (24.0)	69.5 (22.9)	<.01
SNOT-22 Total	39.1 (16.8)	44.7 (19.1)	0.01
SNOT-22 Cardinal	16.4 (5.8)	18.9 (6.7)	<.01
SNOT-22 Other	22.8 (13.3)	25.8 (14.4)	NS
Physical Functioning	51.0 (7.4)	50.3 (8.1)	NS
Anxiety	48.7 (9.9)	49.4 (9.8)	NS
Depression	47.4 (8.8)	47.3 (8.5)	NS
Fatigue	51.3 (11.7)	52.1 (11.5)	NS
Sleep Disturbance	52.3 (8.5)	53.1 (7.4)	NS
Satisfaction with Social Role	49.6 (9.4)	49.6 (9.2)	NS
Pain Interference	50.2 (8.9)	52.0 (9.0)	NS
Pain Intensity	2.9 (2.5)	3.5 (2.6)	0.04

MT=Patients initiating medical therapy, ESS=Patients undergoing endoscopic sinus surgery

^xPROMIS-29 domain scores are T-scores calibrated such that a T-score score of 50 represents the census-representative general population mean and each domain has a SD of 10.

Table 3

Responsiveness of CRS-specific measures in MT and ESS groups

	MT				ESS			
	Mean	#	d [#]	t [‡]	Mean	#	d [#]	t [‡]
VAS	-13.8		-0.58 ^{**}	-1.81	-45.1		-1.97 ^{**}	-9.63 ^{**}
SNOT-22 Total	-11.7		-0.70 ^{**}	-3.29 ^{**}	-29.6		-1.56 ^{**}	-9.99 ^{**}
SNOT-Cardinal	-4.8		-0.83 ^{**}	-4.30 ^{**}	-12.6		-1.88 ^{**}	-9.72 ^{**}
SNOT-Other	-6.9		-0.52 [*]	-2.38 [*]	-17.1		-1.19 ^{**}	-7.24 ^{**}

* p < .05,

** p < .01

MT=Patients initiating medical therapy, ESS=Patients undergoing endoscopic sinus surgery

[#]Mean changes and Cohen's d (effect sizes) were calculated using the mean and standard deviations of all baseline patients and patients with 3-month data.

[‡]Paired t-statistics were calculated for patients with baseline and 3-month data.

Table 4

Most and least responsive individual SNOT-22 items in MT group

	Mean	#	d [#]	t [‡]
Most responsive:				
Runny nose	-1.0		-0.72 ^{**}	4.17 ^{**}
Blockage/congestion of nose	-0.9		-0.69 ^{**}	3.92 ^{**}
PND	-0.7		-0.55 ^{**}	2.70 ^{**}
Need to blow nose	-0.7		-0.49 ^{**}	2.31 [*]
Facial pain/pressure	-0.7		-0.48 ^{**}	3.44 ^{**}
Least responsive:				
Dizziness	-0.31		-0.29	2.04 [*]
Reduced concentration	-0.4		-0.29	1.44
Embarrassed	-0.31		-0.29	0.6
Sad	-0.36		-0.27	2.11 [*]
Ear fullness	-0.3		-0.26	0.58

* p < .05,

** p < .01

MT=Patients initiating medical therapy

[#] Mean changes and Cohen's d (effect sizes) were calculated using the mean and standard deviations of all baseline patients and patients with 3-month data.

[‡] Paired t-statistics were calculated for the patients with baseline and 3-month data.

Table 5

Most and least responsive individual SNOT-22 items in ESS group

	Mean	#	d [#]	t [‡]
Most responsive:				
Blockage/congestion of nose	-2.60		-2.16 ^{**}	-9.88 ^{**}
Need to blow nose	-1.70		-1.16 ^{**}	-6.34 ^{**}
Frustrated/restless/irritable	-1.60		-1.10 ^{**}	-4.91 ^{**}
Wake up tired	-1.60		-1.10 ^{**}	-6.58 ^{**}
Runny nose	-1.65		-1.08 ^{**}	-5.73 ^{**}
Least responsive:				
Sad	-0.82		-0.64 ^{**}	-2.98 ^{**}
Embarrassed	-0.64		-0.64 ^{**}	-2.99 ^{**}
Sneezing	-0.81		-0.62 ^{**}	-3.64 ^{**}
Dizziness	-0.67		-0.59 ^{**}	-3.95 ^{**}
Cough	-0.90		-0.59 ^{**}	-4.78 ^{**}

*
p < .05,**
p < .01

ESS=Patients undergoing endoscopic sinus surgery

Mean changes and Cohen's d (effect sizes) were calculated using the mean and standard deviations of all baseline patients and patients with 3-month data.‡
Paired t-statistics were calculated for patients with baseline and 3-month data.

Table 6

Responsiveness of PROMIS-29 domains^x in MT and ESS groups

	MT				ESS			
	Mean	#	d [#]	t [‡]	Mean	#	d [#]	t [‡]
Anxiety	-1.50		-0.15	-1.08	-5.0		-0.51	-2.63 [*]
Depression	-1.70		-0.19	-1.06	-3.2		-0.37	-1.06
Sleep Disturbance	-2.40		-0.28	-1.18	-6.2		-0.83 ^{**}	-3.77 ^{**}
Fatigue	-2.40		-0.28	-2.20 [*]	-9.4		-0.82 [*]	-4.63 ^{**}
Physical Function	2.10		0.28	1.0	2.7		0.33	2.26 [*]
Satisfaction with Social Role	1.70		0.18	1.14	6.1		0.66	3.82 ^{**}
Pain Interference	-3.0		-0.34	-1.71	-7.8		-0.86	-5.67 ^{**}
Pain Intensity	-0.8		-0.33	-1.77	-2.6		-1.0 ^{**}	-6.94 ^{**}

^{*}p < .05,

^{**}p < .01

MT=patients initiating medical therapy, ESS=patients undergoing endoscopic sinus surgery

^xPROMIS-29 domain scores are T-scores calibrated such that a T-score score of 50 represents the census-representative general population mean and each domain has a SD of 10.

[#]Mean changes and Cohen's d (effect sizes) were calculated using the mean and standard deviations of all baseline patients and patients with 3-month data.

[‡]Paired t-statistics were calculated for patients with baseline and 3-month data.