Use of the Modified SNOT-16 in Primary Care Patients With Clinically Diagnosed Acute Rhinosinusitis

Jane Garbutt, MB, ChB; Edward Spitznagel, PhD; Jay Piccirillo, MD

Objective: To determine the reliability, validity, responsiveness, and the minimal important difference (MID) for the Sinonasal Outcome Test-16 (SNOT-16) in the measurement of disease-specific quality of life (QOL) in adults with acute rhinosinusitis.

Study Design: Randomized controlled trial to evaluate antibiotic treatment for acute rhinosinusitis.

Setting: Ten community practices in St Louis, Missouri.

Methods: The modified SNOT-16 was completed at baseline (by both face-to-face and telephone interviews) and by telephone interview at 3, 7, and 10 days by 166 adults with acute rhinosinusitis diagnosed clinically using standardized criteria (36% were male, 78% were white). Considering severity and frequency, patients rated how much they were bothered by each item using a 4-point scale. The mean SNOT-16 score (ranging from 0 [no problem] to 3 [large problem]) was compared with the patients’ global assessment of change to evaluate responsiveness and the MID.

Results: The instrument was easy to use and took less than 5 minutes to complete. The SNOT-16 score identified statistically significant differences in the hypothesized direction for those reporting more or less severe symptoms (P = .02) and more or less bother (P < .001) demonstrating construct-related validity. The Cronbach α ranged from 0.82 to 0.91, demonstrating high internal consistency. There was a statistically significant decrease in scores with time (multivariate analysis of variance, P < .001). The effect sizes at days 3, 7, and 10 were 1.45, 2.34, and 2.90, respectively, indicating high sensitivity to clinical change. The MID was 0.5 units.

Conclusion: The modified SNOT-16 is a valid instrument to assess effectiveness of interventions to improve disease-specific QOL in adults with acute rhinosinusitis.

Trial Registration: clinicaltrials.gov Identifier: NCT00377403

Arch Otolaryngol Head Neck Surg. 2011;137(8):792-797

A CUTE RHINOSINUSITIS IS AN important cause of morbidity, anxiety, lost time from work and school, and treatment costs. Treatment is directed at reducing symptoms (eg, headache, cough, and nasal obstruction) and limiting functional impairment (eg, fatigue, difficulty sleeping, and concentrating). Because there are no clinical objective measures of disease resolution for use in clinical trials, tools to assess outcomes that are meaningful for patients are needed. Improvement in quality of life (QOL) is important to patients. Several instruments to measure disease-specific QOL have been developed and evaluated in patients with chronic rhinosinusitis. These instruments measure the impact of the illness on physical and psychosocial functioning and general well-being. No disease-specific QOL instrument has been evaluated for use in studies to assess the effectiveness of interventions for patients with acute rhinosinusitis.

We report the validation of the Sinonasal Outcome Test-16 (SNOT-16), a disease-specific QOL instrument, for use in acute rhinosinusitis. The SNOT-16 is one of several Sinonasal Outcome Tests derived from the Rhinosinusitis Outcome Measure (RSOM-31). These instruments were developed using established psychometric methods including input from patients via focus groups to ensure inclusion of items they reported as important. These included functional limitations, physical problems, and emotional consequences. The SNOT-20, SNOT-16, and most recently the SNOT-22 have been evaluated for use in pa-
tients with chronic rhinosinusitis and found to be reliable, valid, and responsive to meaningful clinical change.4,5 We evaluated the reliability, concurrent validity, and responsiveness of the SNOT-16 when used in the primary care setting. This evaluation was conducted during a randomized, placebo-controlled trial to evaluate antibiotic treatment for adult patients with acute rhinosinusitis.

**METHODS**

**STUDY POPULATION**

We recruited patients from the offices of primary care physicians in St Louis, Missouri. Adult patients (18-70 years old) were eligible if they were diagnosed as having acute rhinosinusitis by their physician. Diagnostic criteria were standardized and included a report of maxillary pain or tenderness in the face or teeth and purulent nasal secretions; rhinosinusitis symptoms for 7 to 28 days that were not improving or worsening; or rhinosinusitis symptoms for less than 7 days that had worsened after initial improvement. Patients rated the severity of their current sinus symptoms on a 5-point categorical global rating scale (very mild, mild, moderate, severe, and very severe). Those who reported their symptoms as moderate, severe, or very severe were eligible to participate.

Patients were excluded if they had an allergy to penicillin or amoxicillin, had received antibiotic therapy within the past 4 weeks, had complications of sinusitis, were pregnant, had a comorbidity that could impair their immune response, had cystic fibrosis, or required an antibiotic for a concurrent condition.

All participants provided written consent. The study was approved by the Washington University human research protection committee.

**INTERVENTION**

Participants were randomized to receive a 10-day course of either amoxicillin (500 mg orally 3 times daily) or placebo similar in appearance and taste and dispensed in the same fashion. Unless their physician felt it was contraindicated, all patients received a 5- to 7-day supply of the symptomatic treatments to be used as needed, including acetaminophen, guaifenesin, dextromethorphan hydrobromide, pseudoephedrine sustained action, and saline spray.

**MEASUREMENT**

Data were collected using structured questionnaires administered by trained research assistants (RAs) blinded to group assignment. At study enrollment (day 0), the patient completed a brief face-to-face interview with the on-site RA. The SNOT-16 was completed during this face-to-face interview and again later that day during a telephone interview. Subsequent interviews were conducted by telephone at days 3, 7, and 10.

**SNOT-16**

The patient’s disease-specific QOL was measured using a modified version of the SNOT-16, an instrument that gathers information on 16 sinus-related symptoms.3 To simplify telephone administration, we used a version of the SNOT-16 with a 4-point response scale rather than the 5-point scale used in previous studies of chronic disease.5 For our tool, respondents were instructed to consider both the severity and frequency for that item (eg, headache) and report how much they had been bothered by each item in the past few days. Response options included 0, no problem; 1, mild or slight problem; 2, moderate problem; and 3, severe problem. This modified tool was used for all study interviews. At baseline, the patients selected the items they felt were most important (≤5) from the list. The SNOT-16 score was calculated as the mean of all responses and ranged from 0 to 3.

Global rating of change in sinus symptoms was assessed using an anchor-based scale. Participants assessed status of their current sinus symptoms relative to their status at enrollment at each follow-up interview using a 6-point categorical scale (a lot worse, a little worse, the same, a little better, a lot better, or no symptoms).

**DATA ANALYSIS**

For this study, patients from the intervention and controls groups were analyzed as one group. We used the day 0 telephone-administered SNOT-16 score for all change analyses, except for 4 participants with this score missing, for whom we used the face-to-face score.

**CONSTRUCT-RELATED VALIDITY**

To assess construct-related validity (the degree to which the instrument provides results that are correlated with related measures), we examined SNOT-16 scores in subgroups expected to have higher scores. We compared day 0 scores in patients reporting more severe symptoms and those who were more bothered by their symptoms with those from patients with less severe or bothersome symptoms. We also examined the day 0 score in patients with asthma, with allergies, and who smoked.4 Statistical significance of the differences across subgroup scores was assessed using t tests or analysis of variance as appropriate.

**REPRODUCIBILITY**

We evaluated internal consistency using Cronbach’s. The minimum acceptable score for this statistic for group comparison is greater than 0.7.7 We assessed test-retest reproducibility comparing the test administered at baseline during a face-to-face interview with that administered later that same day by telephone. We calculated Pearson correlation statistic, compared the baseline SNOT-16 scores using a paired t test, and calculated the intraclass correlation coefficient (ICC).

**RESPONSIVENESS**

We defined responsiveness as the ability of the instrument to detect small but important clinical changes over time.8 We calculated the mean change in SNOT-16 scores from baseline to each time point (3, 7, and 10 days) and compared scores at baseline with those at later time points using paired t tests. We evaluated the trend in change scores using multivariate analysis of variance (MANOVA).

We calculated the effect size as a measure of responsiveness using the formula: \( (U - V) / C \), where \( U \) = the mean baseline score, \( V \) = the mean follow-up score, and \( C \) = the standard deviation of the baseline score.8 By convention, an effect size greater than 0.2 is considered to be a small improvement in health-related QOL, greater than 0.5 is a moderate change, and greater than 0.8 is a large change.7,9 Larger effect sizes indicate an instrument that is highly sensitive to change.9

**CLINICAL INTERPRETABILITY**

To aid in the interpretation of study results that show a change in the SNOT-16 score, we calculated the “minimally important difference” (MID),3 the smallest difference in score that is
clinically significant. We defined this value as the change in SNOT-16 score for those who reported a small change in their symptoms on the global rating scale. Using the patient’s global assessment of change in symptoms since baseline, we categorized “a lot worse” and “a lot better” and “no symptoms” as a large change, “a little worse” and “a little better” as a small change, and “the same” as no change.

RESULTS

STUDY PATIENTS

From November 1, 2006, through May 1, 2009, the SNOT-16 instrument was administered to 166 patients randomized in the intervention study and recruited from 10 primary care practices in St Louis, Missouri (Table 1). Follow-up interviews at days 3, 7, and 10 were completed by 94%, 94%, and 92% participants, respectively, with no difference by study group.

PATIENT CHARACTERISTICS

The median age of the study population was 32 years (range, 18-69 years), 36% were male, 78% were white, and 16% were African American (Table 1). All participants reported purulent nasal discharge and maxillary pain or tenderness in the face or teeth (the pain was bilateral in 94 patients and unilateral in 56; the laterality was unknown in 14). The mean (SD) symptom duration prior to the enrollment visit was 11.2 (5.7) days. Participants described their symptoms as moderate (48%), severe (43%), or very severe (9%).

PERFORMANCE OF THE SNOT-16

Respondent Burden

The instrument was easy to use and took less than 5 minutes to complete. It was well accepted by patients and did not require special training for the RA to administer. For each completed interview, there were no missing data.

Construct-Related Validity

The mean SNOT-16 scores assessed at day 0 for various subgroups are presented in Table 2. The score was able to identify statistically significant differences in the hypothesized direction for those reporting more or less severe symptoms (P = .02) and more or less bother (P < .001). Significant differences were also found between those with and without allergies (P = .002).

Internal Consistency

Cronbach α scores for the baseline face-to-face interview and the day 0 telephone interview and at days 3, 7,
and 10 were 0.82, 0.82, 0.87, 0.91, and 0.91, respectively. This indicates high internal consistency for this measure in this study population.

Reproducibility

We compared the SNOT-16 scores at baseline administered by face-to-face and telephone interviews in 162 patients for whom both measures were available. The mean (SD) baseline scores differed significantly: 1.84 (0.51) for the face-to-face score, and 1.70 (0.52) for the telephone measure ($P<.001$). On average, the SNOT-16 score was 0.13 (0.36) units lower when measured by telephone interview compared with the face-to-face interview. The Pearson correlation coefficient was 0.75, and the ICC was 0.73.

Responsiveness

There was a statistically significant decrease in scores with time (MANOVA, $P<.001$) (Table 3). The effect sizes at days 3, 7, and 10 were 1.45, 2.34, and 2.90, respectively, indicating a highly responsive instrument.

Clinical Interpretability

The mean 3-, 7-, and 10-day changes in SNOT-16 scores are shown in Table 4 and Table 5. In Table 4, data from patients who reported the same level of change (eg, a little better and a little worse) are combined, regardless of the direction of that change. For each day, the numbers reporting “no change” on the global rating scale were small, and the calculated mean change scores did not differ from 0. For those reporting a small change in symptoms on the global rating scale, the mean change in SNOT-16 score ranged from 0.48 unit at day 3 to 0.80 unit at day 10. For those reporting a large change in symptoms, the mean change in SNOT-16 score increased, ranging from 0.8 unit at day 3 to 1.3 units at day 10. From these data we estimate the MID for the SNOT-16 to be 0.5 unit on this 0 to 3 scale. When improvement and deterioration were examined separately (Table 5), a similar pattern was seen for those who improved.

We repeated the analyses for responsiveness and clinical interpretability using only the 5 items from the SNOT-16 considered to be most important by each patient. Again, there was a statistically significant decrease in scores over time ($P<.001$), with larger effect sizes (day 3, 1.45; day 7, 2.34; day 10, 2.90). The minimally important difference was 0.6. For those reporting a small change in symptoms (better or worse combined), the changes in SNOT-16 scores at days 3, 7, and 10 were −0.63 (95% confidence interval [CI], −0.81 to −0.45), −0.92 (95% CI, −1.11 to −0.73), and −1.10 (95% CI, −1.45 to −0.75), respectively.

Table 3. Responsiveness of the SNOT-16 Score Over Time

<table>
<thead>
<tr>
<th>Day</th>
<th>No.</th>
<th>Mean (SD)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>166</td>
<td>1.71 (0.32)</td>
<td>1.63-1.79</td>
</tr>
<tr>
<td>3</td>
<td>155</td>
<td>1.13 (0.54)</td>
<td>1.04-1.21</td>
</tr>
<tr>
<td>7</td>
<td>155</td>
<td>0.74 (0.54)</td>
<td>0.66-0.83</td>
</tr>
<tr>
<td>10</td>
<td>152</td>
<td>0.49 (0.44)</td>
<td>0.42-0.56</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; SNOT-16, Sinonasal Outcome Test-16.

A $P$ value calculated using paired $t$ test compared with day 0; $P<.001$ for all comparisons.

To our knowledge, this is the first study to evaluate a disease-specific QOL tool for use as an outcome measure for patients with clinically diagnosed acute rhinosinusitis. We have demonstrated internal consistency, construct-related validity, and responsiveness to change of the modified SNOT-16 in this patient population and calculated the minimally important difference in score. The SNOT instruments were developed for use in patients with chronic rhinosinusitis and have been used extensively in studies to evaluate surgical and nonsurgical management for these patients. Our findings suggest that the conceptual basis used to develop the tool for patients with chronic disease is relevant for patients with acute rhinosinusitis. The instrument performed well in this patient group and can be used to assess change in disease-specific QOL in studies evaluating management of adults with acute rhinosinusitis.

There are many aspects to assessing the validity of an instrument. For the SNOT-16, criterion validity cannot be assessed because there is no gold standard measure for the QOL to use as a comparator. Content-related validity has been established during the development of the tool from the RSOM-31 and the SNOT-20.6 We assessed construct-related validity and determined that the score behaved as it would be expected to do in patients with acute rhinosinusitis who reported more or less severe symptoms and overall bother from their illness.

The internal consistency of the modified SNOT-16 used in this study when assessed with Cronbach $\alpha$ was high (≥0.82) and similar to other SNOT instruments used in patients with chronic sinusitis.6 This suggests that items included in the SNOT-16 instrument are homogeneous and that the instrument is reliable for use in intervention studies in patients with acute sinusitis. We hypothesized that baseline SNOT-16 scores assessed in the office during a face-to-face interview would not differ significantly from scores generated when the instrument was administered during a telephone interview later the same day. Unfortunately, this was not the case. We do not know if the 0.13-unit decrease in these 2 scores assessed within 12 hours of each other was due to the different mode of administration, resolution of this acute illness in the short time period involved, or regression to the mean. However, we have confirmed that if telephone follow-up is planned for outcome assessment, then the initial SNOT-16 measurement must also be completed by telephone.
Responsiveness is an important characteristic of an outcome measure for use in a clinical trial. We compared the change in the SNOT-16 score with a global rating of change provided by the patient, a metric commonly used in clinical practice and clinical research. We demonstrated that the modified SNOT-16 was highly responsive to change over time for patients with acute disease. The measured effect size increased with time but was consistently greater than 1.45. This finding compares favorably with the values of 0.69 and 0.81 calculated for other SNOT instruments for patients with chronic rhinosinusitis. We were unable to use the approach suggested by Guyatt et al to assess responsiveness (dividing the difference in mean scores by the standard deviation of the change scores in stable patients) to detect responsiveness above the normal variations in score because very few patients remained stable in the study population.

The SNOT-16 allows patients to identify the 5 items most important to them at baseline, regardless of their magnitude. Change scores for these most important items were significantly associated with treatment response in patients with chronic rhinosinusitis. The instrument developers suggest that this feature is what distinguishes the SNOT instrument as a QOL measure from a measure of health status and can be used to track response to therapies and guide clinical management. Others have argued that this feature simply adds respondent burden without providing additional information. We found that the additional respondent burden was minimal, the effect size assessed using the SNOT-16 was larger, and the MID was about the same (0.6 vs 0.5). Investigators planning to use this tool for outcome assessment will need to determine if this additional feature of the instrument is needed in their study.

The clinical meaning of a change in QOL score is difficult to assess. One approach is to calculate the smallest change that would be detected as an improvement (the MID). We calculated the MID for the modified SNOT-16 in patients with acute sinusitis to be 0.5 units. While this is the group average and may not represent meaningful change at the level of the individual, it suggests that a patient with a change in SNOT-16 score of less than 0.5 is unlikely to perceive any benefit from treatment. One way an individual could achieve a clinically meaningful difference (change of 0.5 unit) would be a change of 8 units across the 16 items in the SNOT-16.

Recently, the SNOT-20 was modified for use in studies of chronic rhinosinusitis. On the recommendation of a focus group of experienced endoscopic sinus surgeons from the United Kingdom, 2 items were added owing to concerns about the validity of the tool for patients with chronic disease (nasal blockage and altered taste or smell). Our findings suggest that these additions to the

---

**Table 4. Changes in SNOT-16 Scores vs Global Rating of Change: Better and Worse Data Combined**

<table>
<thead>
<tr>
<th>Level of Change</th>
<th>Day 3</th>
<th>Day 7</th>
<th>Day 10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients, No.</td>
<td>Change in SNOT-16 Score</td>
<td>Patients, No.</td>
</tr>
<tr>
<td>Same</td>
<td>14</td>
<td>−0.098 (0.33) [−0.29 to 0.09]</td>
<td>7</td>
</tr>
<tr>
<td>Small change</td>
<td>83</td>
<td>−0.48 (0.52) [−0.60 to −0.37]</td>
<td>46</td>
</tr>
<tr>
<td>Large change</td>
<td>58</td>
<td>−0.83 (0.51) [−0.96 to −0.69]</td>
<td>102</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; SNOT-16, Sinonasal Outcome Test-16.

a Data are given as mean (SD) [95% confidence interval].

**Table 5. Change in SNOT-16 Scores vs Global Rating of Change Over Time**

<table>
<thead>
<tr>
<th>Global Rating</th>
<th>Day 3</th>
<th>Day 7</th>
<th>Day 10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients, No.</td>
<td>Change in SNOT-16 Score</td>
<td>Patients, No.</td>
</tr>
<tr>
<td>No symptoms</td>
<td>0</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td>A lot better</td>
<td>55</td>
<td>−0.86 (0.49)</td>
<td>95</td>
</tr>
<tr>
<td>A little better</td>
<td>75</td>
<td>−0.53 (0.49)</td>
<td>44</td>
</tr>
<tr>
<td>The same</td>
<td>14</td>
<td>−0.10 (0.33)</td>
<td>7</td>
</tr>
<tr>
<td>A little worse</td>
<td>8</td>
<td>0.03 (0.59)</td>
<td>2</td>
</tr>
<tr>
<td>A lot worse</td>
<td>3</td>
<td>0.19 (0.47)</td>
<td>1</td>
</tr>
</tbody>
</table>

Abbreviation: SNOT-16, Sinonasal Outcome Test-16.

a Data are given as mean (SD).

b No SD is given because there is only 1 patient in this group.
instrument are unnecessary for patients with acute rhinosinusitis because the 16-item score was responsive and easy to use.

There are several limitations to our work. We did not repeat previous work that was part of the development of the SNOT instruments. For example, we did not conduct focus groups with patients with acute rhinosinusitis to ensure that the conceptual model developed for use in the assessment of chronic sinusitis still applied. Nor did we evaluate content-related or discriminant validity for the SNOT-16 in patients with acute sinusitis. These metrics of instrument performance have been evaluated for the SNOT-16, SNOT-20, and SNOT-22 in previous studies in patients with chronic sinusitis,4,6 and we did not plan to use the instrument as a diagnostic tool for acute rhinosinusitis. Because our interest is in an acute disease, it is difficult to assess test-test reliability and identify a group expected to remain stable as a comparator to evaluate responsiveness.

In conclusion, we found the modified SNOT-16 to be valid, responsive, and easy to use, and we recommend its use to assess change in disease-related QOL in interventions targeting patients with acute rhinosinusitis. It is important to note that if the planned mode of administration is by telephone interview, then the baseline assessment must also be by telephone interview.

Submitted for Publication: February 22, 2011; final revision received May 6, 2011; accepted May 24, 2011.

Author Contributions: Dr Garbutt had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Garbutt and Spitznagel. Analysis and interpretation of data: Garbutt, Spitznagel, and Piccirillo. Drafting of the manuscript: Garbutt and Spitznagel. Critical revision of the manuscript for important intellectual content: Garbutt, Spitznagel, and Piccirillo. Statistical analysis: Spitznagel. Obtained funding: Garbutt. Study supervision: Garbutt and Piccirillo.

Correspondence: Jane Garbutt, MB, ChB, Department of Pediatrics, Washington University School of Medicine, Campus Box 8116, 660 S Euclid Ave, St Louis, MO 63110 (jgarbutt@dom.wustl.edu).

Financial Disclosure: The SNOT-16 is intellectual property of and copyright protected by Washington University. Dr Piccirillo receives royalty income as a result of licensing contacts negotiated by Washington University.

Funding/Support: This study was funded by grant IU01 AI0645501A1 from the National Institute of Allergy and Infectious Diseases.

REFERENCES