Premiere Publications from The Triological Society

Read all three of our prestigious publications, each offering high-quality content to keep you informed with the latest developments in the field.

**The Laryngoscope**
FOUNDED IN 1896
Editor-in-Chief: Michael G. Stewart, MD, MPH
The leading source for information in head and neck disorders.
Laryngoscope.com

**Investigative Otolaryngology**
Editor-in-Chief: D. Bradley Welling, MD, PhD, FACS
Rapid dissemination of the science and practice of otolaryngology-head and neck surgery.
InvestigativeOto.com

**ENTtoday**
A publication of the Triological Society
Editor-in-Chief: Alexander Chiu, MD
Must-have timely information that Otolaryngologist-head and neck surgeons can use in daily practice.
Enttoday.org
Systematic Review

Sino-Nasal Outcome Test-22 Outcomes After Sinus Surgery: A Systematic Review and Meta-Analysis

Zachary M. Soler, MD, MSc; Rabun Jones, BS; Phong Le, MD; Luke Rudmik, MD; Jose L. Mattos, MD, MPH; Shaun A. Nguyen, MD; Rodney J. Schlosser, MD

Objectives/Hypothesis: The goal of the study was to perform a systematic review with meta-analysis to determine the mean change in the 22-item Sino-Nasal Outcome Test (SNOT-22) across patients who have had endoscopic sinus surgery (ESS) for chronic rhinosinusitis (CRS) in the literature.

Methods: A literature search was performed to identify studies that assessed SNOT-22 scores before and after ESS in adult patients with CRS. A random effects model with inverse variance weighting was used to generate the mean change after surgery, along with the forest plot and 95% confidence interval (CI). The impact of patient-specific factors across studies was assessed using a mixed-effects meta-regression.

Results: The final study list included 40 unique patient cohorts published from 2008 to 2016. All studies showed a statistically significant change in mean SNOT-22 scores between baseline and postoperative time points ($P < .001$), ranging from 12.7 to 44.8, at an average follow-up of 10.6 months. The summary change in mean SNOT-22 across all studies was 24.4 (95% CI: 22.0-26.8). After forward, step-wise multivariate modeling, studies with higher mean preoperative SNOT-22 score and higher asthma prevalence were associated with greater changes in SNOT-22 score after ESS, whereas studies with longer mean follow-up had smaller changes in SNOT-22 score.

Conclusions: Studies evaluating quality-of-life outcomes after sinus surgery using the SNOT-22 instrument universally show significant improvement after ESS. Across the published literature, the magnitude of change is quite variable and appears to be influenced by a number of factors including baseline SNOT-22 score, asthma prevalence, and length of follow-up.

Key Words: Chronic rhinosinusitis, meta-analysis, quality of life, 22-item Sino-Nasal Outcome Test, quality improvement.

Laryngoscope, 128:581–592, 2018

INTRODUCTION

Improving the quality of healthcare delivery remains a major priority for all stakeholders across the healthcare landscape. This renewed emphasis can be seen at nearly every level, ranging from the solo provider who reviews their individual practice with quality improvement in mind, to the third-party payer who institutes a program that links payment to quality outcomes. Implicit in any effort that focuses on quality improvement is the understanding that quality can actually be measured and that references exist that allow one to determine if an individual or entity is underperforming, overperforming, or performing at an expected level as compared to that reference. A recent review of quality measurement for chronic rhinosinusitis (CRS) highlighted that all currently collected quality measures were process metrics and specifically lacked assessment of patient outcomes. Process metrics measure adherence to guidelines or best-practice standards, such as imaging or antibiotic utilization, but do not actually measure health outcomes. Outcome metrics, on the other hand, assess the change in health state after a healthcare activity is performed and typically focus on health states that are important to the patient.

The measurement of sinus-specific quality-of-life (QOL) is perhaps the most commonly utilized outcome measure for CRS. Although sinus-specific QOL measures are routinely employed in studies examining efficacy of treatment for CRS, their utility in quality improvement...
initiatives is less clear-cut. Currently, no regulatory bod-
ies recommend assessing the change in sinus-specific
QOL before and after an intervention for CRS. On one
hand, this is surprising, because validated sinus-specific
QOL measures are readily available, easy to implement,
and quantify outcomes in a patient-centered fashion.
However, one could argue that the biggest hurdle to
incorporating an outcome metric in any quality improve-
ment program is the lack of available reference stan-
ards. For any individual or entity to implement a
change in QOL as an outcome metric, one must have a
clear sense for the expected change in that instrument
after a given intervention, along with the various factors
that might influence the degree of change.

Roughly 300,000 endoscopic sinus surgery (ESS)
procedures are performed yearly in the United States,
primarily for the purpose of improving patient-reported
procedures are performed yearly in the United States,
582
that might influence the degree of change.

across centers.

After a given intervention, along with the various factors
on quality improvement initiatives, one must have
what might influence expected outcomes. In this
way, an individual physician or entity could compare
their SNOT-22 outcomes to a standard reference, adjust-
ing for factors that might be unique to their patient
population.

With the above concepts in mind, we sought to per-
form a systematic review of all studies that examined
the change in SNOT-22 score after ESS for adult
patients with CRS. Our primary goal was to perform a
meta-analysis to determine mean change in SNOT-22
across all patients who have had ESS for CRS in the lit-
erature. Our secondary goal was to assess variation in
mean change across studies and to identify patient-
specific factors that might influence this variation.
Lastly, we sought to explore the feasibility of develop-
ing a reference standard for expected SNOT-22 change after
ESS, and a model that would allow adjustment for
patient-specific factors that might influence that change.

METHODS

Literature Search

A literature search was performed by two authors in par-
allel to identify studies that assessed SNOT-22 scores before
and after ESS in adult patients with CRS. Reviewers queried a
number of databases including PubMed, Embase, Cochrane
Database of Systematic Reviews, Scopus, Ebscohost, and the
York Centre for Reviews and Dissemination. The strategy com-
bined terms for chronic sinusitis (“chronic” AND “sinusitis”) and
various terms for surgery and SNOT-22 to develop an inclusive
list of possible studies. After searching all databases, duplicates
were removed and abstracts reviewed. To be included, all
patients in the study must have been ≥16 years and considered
to have CRS, or data from patients with CRS must have been
presented separate and distinct from data from patients without
CRS. Furthermore, included studies must have examined
patients undergoing surgical intervention for CRS and must
have utilized the SNOT-22 before and after surgery. Studies
that enrolled patients with acute rhinosinusitis or recurrent
acute rhinosinusitis were excluded, as were studies that utilized
other outcome metrics but not the SNOT-22. Full-text articles
were then reviewed to confirm they met eligibility require-
ments. References of all included studies were reviewed to iden-
tify any additional articles that may have been missed. Studies
from the same institution were then examined to determine
whether they are reporting data from the same cohort. If more
than one study from an institution had overlapping enrollment
dates, such that they were likely to report duplicate data, the
study with the largest sample size and/or complete data was
chosen. The reviewers then combined lists, and any discrepan-
cies were resolved through discussion with all authors.

Data Extraction and Quality Review

Data from included studies were then extracted by two
reviewers working independently and in parallel. This
included the diagnostic criteria utilized by original study
authors to determine CRS for any given study. Average patient
characteristics for each study were determined, including demo-
graphics (age and sex) and frequency of medical comorbidities
(asthma, allergic rhinitis, depression, and current tobacco use).
Disease-specific measures were also recorded, including the
frequency of polyps or prior sinus surgery for each study. If
available, the mean Lund-Mackay computed tomography (CT)
score and Lund-Kennedy endoscopy score were recorded for
each study.10,11 Although other CT and endoscopic scoring sys-
tems exist, these were specifically chosen as they are the most
commonly reported. Lastly, the preoperative (baseline) and
postoperative SNOT-22 score was recorded, including means
and standard deviations (SD). For any given study, we used
the last available SNOT-22 score and recorded the length of
follow-up in months. If not specifically provided, mean and/or
SD was calculated from provided data whenever possible.
Studies were excluded if data were unable to be obtained or
had missing values that could not be otherwise calculated. A
quality assessment was then performed using the Quality
Assessment Tool for Before-After (Pre-Post) Studies With No
Control Group provided by the National Heart, Lung, and
Blood Institute.12

Meta-analysis, Study Variation, and Model
Building

A meta-analysis was then performed determining the
mean change in SNOT-22 across all studies using Comprehen-
sive Meta Analysis Version 3.0 (Biostat, Englewood, NJ).
Given expected heterogeneity, a random effects model with
inverse variance weighting was used to generate the mean
change after surgery, along with the forest plot and 95% confi-
dence interval (CI). Studies were then arranged from highest
to lowest change and examined to determine how they com-
pare to average. The review was structured following the
We next wanted to explore how patient-specific factors might influence outcomes, focusing on the demographic, medical comorbidity, and disease-specific factors relevant to CRS. This was done in two ways. First, individual studies were queried as to whether they specifically studied the impact of this factor on SNOT-22 outcomes using patient-level data. If the factor was studied, its impact was then recorded (i.e., effect size and significance) for that individual study. However, no attempt was made to combine these data across studies, as this would require patient-level data that were neither available nor feasible. Next, we investigated the impact of these factors across studies using a mixed effects meta-regression. In the meta-regression, the mean (SD) or frequency of each variable in each study was placed into a model to determine whether it influenced outcomes across studies (i.e., using study-level data). The sensitivity of regression findings to any single study was then explored for positive findings.

Lastly, we sought to build a multivariate model that would provide the reference change in SNOT-22 after surgery for any specific population, based on published literature to date. This was done using forward, stepwise selection for any variable with a \( P \) value \(<.10\) on univariate meta-regression and robust to sensitivity analysis. This model would potentially serve as an available reference at present, and as proof of concept and starting point for future models that might better inform quality improvement initiatives in an unbiased fashion.

Fig. 1. Preferred Reporting Items for Systematic Reviews and Meta-analyses flow diagram.

DARE = Database of Abstracts of Reviews of Effects; HTA = Health Technology Assessment Database; NHSEED = National Health Service Economic Evaluation Database.

Preferred Reporting Items for Systematic Reviews and Meta-analyses guidelines.\(^{13}\)
<table>
<thead>
<tr>
<th>First Author</th>
<th>Year</th>
<th>Study Design</th>
<th>No. of Patients</th>
<th>Country</th>
<th>Diagnosis of CRS</th>
<th>Surgical Intervention</th>
<th>Length of Postoperative Follow-up (Months)</th>
<th>Select Population</th>
<th>Polyps (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ye</td>
<td>2009</td>
<td>Prospective cohort</td>
<td>113</td>
<td>China</td>
<td>Guidelines*</td>
<td>Revision ESS</td>
<td>12</td>
<td>—</td>
<td>100</td>
</tr>
<tr>
<td>Hopkins</td>
<td>2009</td>
<td>Prospective cohort</td>
<td>1,459</td>
<td>United Kingdom</td>
<td>Guidelines*</td>
<td>Primary and Revision ESS</td>
<td>3–60</td>
<td>—</td>
<td>71</td>
</tr>
<tr>
<td>Proimos</td>
<td>2010</td>
<td>Prospective cohort</td>
<td>86</td>
<td>Greece</td>
<td>Guidelines*</td>
<td>ESS</td>
<td>6–12</td>
<td>Asthma</td>
<td>100</td>
</tr>
<tr>
<td>Schaberg</td>
<td>2010</td>
<td>Retrospective cohort</td>
<td>9</td>
<td>United States</td>
<td>NR</td>
<td>Revision ESS</td>
<td>1.5</td>
<td>Hyperostotic sinusitis</td>
<td>-</td>
</tr>
<tr>
<td>Hoseini</td>
<td>2012</td>
<td>Prospective cohort</td>
<td>100</td>
<td>Iran</td>
<td>NR</td>
<td>Primary and revision ESS</td>
<td>24</td>
<td>Cystic fibrosis, diabetes</td>
<td>100</td>
</tr>
<tr>
<td>Kosugi</td>
<td>2011</td>
<td>Prospective cohort</td>
<td>89</td>
<td>Brazil</td>
<td>Guidelines*</td>
<td>ESS</td>
<td>3</td>
<td>—</td>
<td>49</td>
</tr>
<tr>
<td>Conger</td>
<td>2012</td>
<td>Prospective cohort</td>
<td>9</td>
<td>United States</td>
<td>NR</td>
<td>Draf III procedure</td>
<td>3–12</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Sadeghi</td>
<td>2013</td>
<td>Case control</td>
<td>50</td>
<td>Iran</td>
<td>Other</td>
<td>ESS ± rhinoplasty</td>
<td>12</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Snidvongs</td>
<td>2012</td>
<td>Prospective cohort</td>
<td>104</td>
<td>Australia</td>
<td>Guidelines*</td>
<td>Primary and revision ESS</td>
<td>13.975</td>
<td>—</td>
<td>56</td>
</tr>
<tr>
<td>Vaitkus</td>
<td>2013</td>
<td>Case-Control</td>
<td>36</td>
<td>Lithuania</td>
<td>Guidelines*</td>
<td>ESS</td>
<td>3</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Virgin</td>
<td>2012</td>
<td>Prospective cohort</td>
<td>22</td>
<td>United States</td>
<td>Guidelines†</td>
<td>Primary and revision ESS</td>
<td>1–12</td>
<td>Cystic fibrosis</td>
<td>—</td>
</tr>
<tr>
<td>Garetier</td>
<td>2013</td>
<td>Retrospective cohort</td>
<td>33</td>
<td>United States</td>
<td>Guidelines*</td>
<td>Primary ESS</td>
<td>3–12</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Mascarenhas</td>
<td>2013</td>
<td>Prospective cohort</td>
<td>38</td>
<td>Brazil</td>
<td>Guidelines⁵</td>
<td>ESS</td>
<td>3–24</td>
<td>—</td>
<td>45</td>
</tr>
<tr>
<td>Sacks</td>
<td>2013</td>
<td>Prospective cohort</td>
<td>53</td>
<td>Australia</td>
<td>Guidelines*</td>
<td>Primary ESS</td>
<td>12</td>
<td>—</td>
<td>44</td>
</tr>
<tr>
<td>Bizaki</td>
<td>2014</td>
<td>RCT</td>
<td>21</td>
<td>Finland</td>
<td>Guidelines⁶</td>
<td>Primary ESS</td>
<td>3</td>
<td>—</td>
<td>0</td>
</tr>
<tr>
<td>Cantone</td>
<td>2014</td>
<td>RCT</td>
<td>122</td>
<td>Italy</td>
<td>Guidelines⁶</td>
<td>Primary ESS</td>
<td>1</td>
<td>—</td>
<td>100</td>
</tr>
<tr>
<td>Cho</td>
<td>2014</td>
<td>Prospective cohort</td>
<td>21</td>
<td>South Korea</td>
<td>Guidelines⁸</td>
<td>Primary and revision ESS</td>
<td>0.25–1</td>
<td>AERD</td>
<td>100</td>
</tr>
<tr>
<td>Dautremont</td>
<td>2014</td>
<td>RCT</td>
<td>36</td>
<td>Canada</td>
<td>Guidelines¹</td>
<td>ESS with steroid-eluting spacer</td>
<td>2</td>
<td>—</td>
<td>100</td>
</tr>
<tr>
<td>DeConde</td>
<td>2014</td>
<td>Prospective cohort</td>
<td>229</td>
<td>United States</td>
<td>Guidelines⁷,¹</td>
<td>Primary and revision ESS</td>
<td>15</td>
<td>Migraine</td>
<td>50</td>
</tr>
<tr>
<td>ElBadawy</td>
<td>2014</td>
<td>Prospective cohort</td>
<td>6</td>
<td>United Kingdom</td>
<td>Other</td>
<td>Frontal recess clearance</td>
<td>12</td>
<td>—</td>
<td>33</td>
</tr>
<tr>
<td>Hashemian</td>
<td>2015</td>
<td>RCT</td>
<td>53</td>
<td>Iran</td>
<td>Guidelines¹</td>
<td>ESS</td>
<td>2</td>
<td>—</td>
<td>41</td>
</tr>
<tr>
<td>Lachanas</td>
<td>2014</td>
<td>Prospective cohort</td>
<td>32</td>
<td>Greece</td>
<td>Guidelines⁵</td>
<td>ESS</td>
<td>3</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Saedi</td>
<td>2014</td>
<td>Prospective cohort</td>
<td>47</td>
<td>Iran</td>
<td>Other</td>
<td>Primary and revision ESS</td>
<td>12</td>
<td>—</td>
<td>100</td>
</tr>
<tr>
<td>Savastano</td>
<td>2014</td>
<td>Prospective cohort</td>
<td>33</td>
<td>Italy</td>
<td>Guidelines*</td>
<td>ESS</td>
<td>6–24</td>
<td>Cystic fibrosis</td>
<td>100</td>
</tr>
<tr>
<td>Zhang</td>
<td>2014</td>
<td>Retrospective cohort</td>
<td>187</td>
<td>United States</td>
<td>Guidelines†</td>
<td>Primary and revision ESS</td>
<td>1–6</td>
<td>Asthma</td>
<td>60</td>
</tr>
<tr>
<td>Amali</td>
<td>2015</td>
<td>RCT</td>
<td>66</td>
<td>Iran</td>
<td>Guidelines¹</td>
<td>Primary ESS</td>
<td>3</td>
<td>—</td>
<td>42</td>
</tr>
<tr>
<td>Divekar</td>
<td>2015</td>
<td>Retrospective cohort</td>
<td>97</td>
<td>United States</td>
<td>Guidelines¹</td>
<td>ESS</td>
<td>0.75–3</td>
<td>—</td>
<td>85</td>
</tr>
<tr>
<td>Lee</td>
<td>2015</td>
<td>Prospective cohort</td>
<td>35</td>
<td>Canada</td>
<td>Guidelines¹</td>
<td>Primary and revision ESS</td>
<td>1–6</td>
<td>—</td>
<td>85</td>
</tr>
<tr>
<td>First Author</td>
<td>Year</td>
<td>Study Design</td>
<td>No. of Patients</td>
<td>Country</td>
<td>Diagnosis of CRS</td>
<td>Surgical Intervention</td>
<td>Length of Postoperative Follow-up (Months)</td>
<td>Select Population</td>
<td>Polyps (%)</td>
</tr>
<tr>
<td>-------------------</td>
<td>------</td>
<td>---------------</td>
<td>-----------------</td>
<td>---------</td>
<td>------------------</td>
<td>-----------------------</td>
<td>-------------------------------------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Lind²⁵</td>
<td>2016</td>
<td>Prospective cohort</td>
<td>88</td>
<td>Denmark</td>
<td>Guidelines⁶</td>
<td>Primary and revision ESS</td>
<td>1–6</td>
<td>—</td>
<td>77</td>
</tr>
<tr>
<td>Ottaviano²³</td>
<td>2015</td>
<td>RCT</td>
<td>34</td>
<td>Italy</td>
<td>Guidelines⁶</td>
<td>ESS</td>
<td>0.5–3</td>
<td>—</td>
<td>59</td>
</tr>
<tr>
<td>Rotenberg²⁴</td>
<td>2015</td>
<td>Retrospective cohort</td>
<td>53</td>
<td>Canada</td>
<td>Guidelines¹</td>
<td>Primary and revision ESS</td>
<td>6</td>
<td>—</td>
<td>0</td>
</tr>
<tr>
<td>Shapira Galitz²⁶</td>
<td>2016</td>
<td>Prospective cohort</td>
<td>28</td>
<td>Israel</td>
<td>Guidelines⁶</td>
<td>ESS</td>
<td>7.59</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Adappa²⁶</td>
<td>2016</td>
<td>Prospective cohort</td>
<td>123</td>
<td>United States</td>
<td>Guidelines⁶</td>
<td>Primary and revision ESS</td>
<td>6</td>
<td>—</td>
<td>69</td>
</tr>
<tr>
<td>Bizaki²⁷</td>
<td>2016</td>
<td>RCT</td>
<td>16</td>
<td>Finland</td>
<td>Guidelines⁶</td>
<td>Primary uncincetomy</td>
<td>6</td>
<td>—</td>
<td>0</td>
</tr>
<tr>
<td>Do²⁸</td>
<td>2016</td>
<td>Case control</td>
<td>110</td>
<td>Australia</td>
<td>Guidelines⁶</td>
<td>Primary and revision ESS</td>
<td>12</td>
<td>—</td>
<td>50</td>
</tr>
<tr>
<td>Lal²⁹</td>
<td>2016</td>
<td>Retrospective cohort</td>
<td>88</td>
<td>United States</td>
<td>Guidelines⁶</td>
<td>Primary and revision ESS</td>
<td>3–24</td>
<td>—</td>
<td>43</td>
</tr>
<tr>
<td>Masterson⁸⁰</td>
<td>2016</td>
<td>Retrospective cohort</td>
<td>78</td>
<td>United Kingdom</td>
<td>Guidelines⁶</td>
<td>Primary and revision ESS</td>
<td>1–12</td>
<td>—</td>
<td>62</td>
</tr>
<tr>
<td>Meier²¹</td>
<td>2016</td>
<td>Prospective cohort</td>
<td>309</td>
<td>United States</td>
<td>Guidelines¹</td>
<td>Primary and revision ESS</td>
<td>12</td>
<td>—</td>
<td>46</td>
</tr>
<tr>
<td>Soler²²</td>
<td>2016</td>
<td>Prospective cohort</td>
<td>160</td>
<td>United States</td>
<td>Guidelines⁶</td>
<td>Primary and revision ESS</td>
<td>6–18</td>
<td>—</td>
<td>37</td>
</tr>
<tr>
<td>Wang²³</td>
<td>2016</td>
<td>Prospective cohort</td>
<td>79</td>
<td>China</td>
<td>Guidelines⁶</td>
<td>ESS</td>
<td>6</td>
<td>Bronchiectasis</td>
<td>45</td>
</tr>
</tbody>
</table>

Other = diagnosis made by history, physical exam, and imaging.

 Guidelines = Sinus and Allergy Health Partnership Criteria 2004.⁶⁷
 Guidelines = Rhinosinusitis Task Force 2003.⁶⁸
 Guidelines = European Position Paper on Rhinosinusitis and Nasal Polyps 2012.⁶⁹
 Guidelines = Rhinosinusitis Task Force 1997.⁷⁰
 Guidelines = Chinese Guidelines 2012.⁷²

AERD = aspirin-exacerbated respiratory disease; ESS = endoscopic sinus surgery; RCT = randomized controlled trial; NR = not reported.
RESULTS

Initial literature search identified 420 study abstracts, of which 101 remained after eliminating duplicates and applying inclusion/exclusion criteria (Fig. 1). Full-length texts were then reviewed, and an additional 23 studies identified from references. A total of 31 studies were considered to be duplicates from the same cohort, and an additional 53 failed to meet inclusion/exclusion criteria. The final study list included 40 unique patient cohorts published from 2008 to 2016 and representing institutions from North America (n = 14), Europe (n = 12), Middle East (n = 6), Australia (n = 3), Asia (n = 3), and South America (n = 2). The majority of studies were prospective observational cohorts (n = 23), with the remainder being surgical arms from randomized clinical trials (n = 7), retrospective cohorts (n = 7), or case-control studies (n = 3). Data represented outcomes from 5,547 patients, with individual study sizes ranging from six to 1,459. Details for individual studies can be found in Table I. With regard to quality, 29 of 40 studies utilized established diagnostic criteria for CRS, with the remainder describing alternate criteria. Quality assessment findings are presented in Supplementary Table 1 in the online version of this article. Five of the studies were rated as good quality with the remainder as fair quality. Most studies failed to document whether all possible participants who met the criteria were actually enrolled and failed to comment on sufficient sample size.

Meta-analysis

All studies included in this cohort had an individual, statistically significant change in mean SNOT-22 scores between baseline and postoperative time points (P < .001), ranging from 12.7 to 44.8 at an average follow-up of 10.6 months. The summary change in mean SNOT-22 across all studies was 24.4 (95% CI: 22.0-26.8; I^2 = 13.5%). Individual study findings and the forest plot for summary measures are shown in Figure 2 arranged by year of publication. Figure 3 shows studies arranged by mean change in SNOT-22 from least to greatest. It can be seen that 19 of 40 studies have a 95% CI that crosses the mean summary change of 24.4 and thus had outcomes that might be called average. Nine studies had a mean change in SNOT-22 and 95% CI that were considered to be significant improvements.
greater than 24.4, whereas in 12 studies the mean change and 95% CI were below 24.4. No significant difference was seen between studies rated as good versus those considered fair. Additionally, no significant difference was seen between those studies that utilized established diagnostic criteria for CRS and those that used alternate criteria.

**Impact of Patient-Specific Factors**

**Within-study findings.** Few studies explicitly described the impact of patient-specific factors on SNOT-22 change after ESS (Table II). Only four studies evaluated the impact of polyp status on SNOT-22 change scores within their cohorts. One study found greater improvement in SNOT-22 in patients with polyps compared to those without. One study found lower-staged polyps had greater improvement than higher-staged polyps. One study found patients with polyps have greater improvement, but only when combined with asthma. The fourth study found no significant correlation between polyp status and SNOT-22 improvement.

Three studies examined the association of preoperative CT scores with SNOT-22 change scores, with one finding a positive correlation and the other two showing no impact. Sex, asthma, revision surgery, depression, tobacco use, preoperative SNOT-22 score, and length of follow-up. These factors were not evaluated specifically or described in only a single study; thus, none of these factors was amenable to quantitative meta-analysis.

**Across-study findings (meta-regression).** Univariate meta-regression was used to explore the association of patient-specific factors across studies on mean change in SNOT-22 scores after ESS (Table III, Fig. 4). No significant impact was found for patient age, sex, endoscopy score, CT score, polyp status, tobacco use, depression, or allergic rhinitis. The presence of asthma, prior sinus surgery, and higher preoperative SNOT-22 scores were associated with greater mean changes in SNOT-22 scores. Greater length of follow-up was associated with less change in SNOT-22 scores. Aspirin...
intolerance was associated with greater change in SNOT-22, but this association was dependent on a single study and therefore not robust to sensitivity analysis. After forward, stepwise multivariate modeling, preoperative SNOT-22 score, asthma prevalence, and length of follow-up remained significant (Tables III and IV). Studies with higher mean preoperative SNOT-22 score and higher asthma prevalence were associated with greater changes in SNOT-22 score after ESS, whereas studies with longer mean follow-up had smaller changes in SNOT-22 score.

<table>
<thead>
<tr>
<th>Study and Year</th>
<th>Type</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polyps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kosugi 201150</td>
<td>Prospective cohort</td>
<td>Greater SNOT-22 improvement in those with polyps</td>
</tr>
<tr>
<td>Mascarenhas 201357</td>
<td>Prospective cohort</td>
<td>No difference in SNOT-22 improvement between polyp and nonpolyp patients</td>
</tr>
<tr>
<td>Saedi 201466</td>
<td>Prospective cohort</td>
<td>Greater SNOT-22 improvement in those with lower-staged polyps</td>
</tr>
<tr>
<td>Zhang 201468</td>
<td>Retrospective cohort</td>
<td>Greater SNOT-22 improvement only in those with polyps and asthma</td>
</tr>
<tr>
<td>CT score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garetier 201356</td>
<td>Retrospective cohort</td>
<td>No correlation between SNOT-22 and pre-/postoperative CT score change</td>
</tr>
<tr>
<td>Savastano 201467</td>
<td>Retrospective cohort</td>
<td>No statistical correlation between SNOT-22 and CT scores</td>
</tr>
<tr>
<td>Lind 201672</td>
<td>Prospective cohort</td>
<td>Higher preoperative CT scores correlated with greater post-operative improvement in SNOT-22</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amali 201569</td>
<td>Randomized controlled trial</td>
<td>No difference observed in SNOT-22 scores between genders</td>
</tr>
<tr>
<td>Adappa 201676</td>
<td>Prospective cohort</td>
<td>No difference in SNOT-22 change between genders</td>
</tr>
<tr>
<td>Lal 201679</td>
<td>Retrospective cohort</td>
<td>No difference in trend or magnitude of SNOT-22 improvement between genders</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amali 201569</td>
<td>Randomized controlled trial</td>
<td>No significant correlation between SNOT-22 scores and patient age</td>
</tr>
<tr>
<td>Adappa 201676</td>
<td>Prospective cohort</td>
<td>No significant correlation between mean SNOT-22 change and patients aged &lt;50 years vs. patients aged &gt;50 years</td>
</tr>
<tr>
<td>Soler 201682</td>
<td>Prospective cohort</td>
<td>SNOT-22 improvement not significantly correlated with age</td>
</tr>
</tbody>
</table>

**TABLE II.** Within-Study Findings: Polyps, CT Score, Gender, and Age.

**TABLE III.** Univariate Meta-Regression Results.

<table>
<thead>
<tr>
<th>Variable</th>
<th>No. of Studies</th>
<th>Coefficient</th>
<th>SE</th>
<th>95% CI</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>40</td>
<td>0.057</td>
<td>0.558</td>
<td>-1.05 to 1.15</td>
<td>.919</td>
</tr>
<tr>
<td>Age</td>
<td>37</td>
<td>-0.006</td>
<td>0.186</td>
<td>-0.37 to 0.36</td>
<td>.974</td>
</tr>
<tr>
<td>Gender (% female)</td>
<td>37</td>
<td>-0.019</td>
<td>0.123</td>
<td>-0.26 to 0.22</td>
<td>.880</td>
</tr>
<tr>
<td>Allergy (%)</td>
<td>17</td>
<td>0.147</td>
<td>0.086</td>
<td>-0.02 to 0.32</td>
<td>.090</td>
</tr>
<tr>
<td>Asthma (%)</td>
<td>23</td>
<td>0.131</td>
<td>0.041</td>
<td>0.05 to 0.21</td>
<td>.002</td>
</tr>
<tr>
<td>Polyps (%)</td>
<td>32</td>
<td>-0.067</td>
<td>0.039</td>
<td>-0.14 to 0.01</td>
<td>.081</td>
</tr>
<tr>
<td>Prior sinus surgery (%)</td>
<td>27</td>
<td>0.109</td>
<td>0.054</td>
<td>0.00 to 0.22</td>
<td>.045</td>
</tr>
<tr>
<td>ASA intolerance (%)</td>
<td>20</td>
<td>0.220</td>
<td>0.056</td>
<td>0.11 to 0.33</td>
<td>.001</td>
</tr>
<tr>
<td>Depression (%)</td>
<td>7</td>
<td>-0.203</td>
<td>0.670</td>
<td>-1.52 to 1.11</td>
<td>.762</td>
</tr>
<tr>
<td>Current tobacco use (%)</td>
<td>15</td>
<td>-0.183</td>
<td>0.215</td>
<td>-0.60 to 0.24</td>
<td>.395</td>
</tr>
<tr>
<td>Baseline CT score</td>
<td>24</td>
<td>0.574</td>
<td>0.321</td>
<td>-0.05 to 1.20</td>
<td>.074</td>
</tr>
<tr>
<td>Baseline endoscopy score</td>
<td>12</td>
<td>1.628</td>
<td>1.451</td>
<td>-1.22 to 4.47</td>
<td>.262</td>
</tr>
<tr>
<td>Length of follow-up (mo)</td>
<td>40</td>
<td>-0.280</td>
<td>0.102</td>
<td>-0.48 to 0.08</td>
<td>.006</td>
</tr>
<tr>
<td>Preop SNOT-22 score</td>
<td>40</td>
<td>0.443</td>
<td>0.092</td>
<td>0.26 to 0.62</td>
<td>.001</td>
</tr>
</tbody>
</table>

ASA = acetylsalicylic acid; CI = confidence interval; CT = computed tomography; Preop = preoperative; SE = standard error; SNOT-22 = 22-item Sino-Nasal Outcome Test.
DISCUSSION

Currently, individual surgeons or group practices that wish to evaluate their patient outcomes are able to collect SNOT-22 data on their CRS patients before and after surgery, along with basic patient-specific data. However, if they wish to compare their outcomes to a benchmark to understand if they are achieving expected improvements, they would have to choose a reference point. Selecting reference points based upon individual studies is problematic due to wide variability in over 40 published series. Thus, a summary measure for expected SNOT-22 change generated across all published studies would be an improved reference point as opposed to any individual study. Data from this study will allow providers to compare their individual data with the summary mean and 95% CI. Outcomes that fall within the expected range would provide confidence, whereas those falling below this range might prompt further investigation and perhaps quality improvement initiatives. If desired, one could go a step further and use the regression model to adjust for baseline differences in that individual surgeon’s practice such as preoperative SNOT-22, asthma prevalence, or length of follow-up, which appear to impact published outcomes across studies.

Summary measures such as this provide a starting point based on current available data but should not be

<table>
<thead>
<tr>
<th>Variable</th>
<th>Univariate Regression P Value</th>
<th>No. of Studies</th>
<th>Coefficient</th>
<th>SE</th>
<th>95% CI</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preop SNOT-22</td>
<td>.001</td>
<td>40</td>
<td>0.324</td>
<td>0.077</td>
<td>0.17 to 0.47</td>
<td>.001</td>
</tr>
<tr>
<td>Asthma (%)</td>
<td>.002</td>
<td>23</td>
<td>0.064</td>
<td>0.031</td>
<td>0.004 to 0.124</td>
<td>.036</td>
</tr>
<tr>
<td>Length of follow-up (mo)</td>
<td>.006</td>
<td>23</td>
<td>−0.191</td>
<td>0.055</td>
<td>−0.30 to −0.08</td>
<td>.001</td>
</tr>
</tbody>
</table>

I² = 84.41%.
CI = confidence interval; SE = standard error; SNOT-22 = 22-item Sino-Nasal Outcome Test.

Fig. 4. Univariate meta-regression bubble plots for length of follow-up (months), preoperative SNOT-22, asthma prevalence (%), and age (years). Pre-op = preoperative; SNOT-22 = 22-item Sino-Nasal Outcome Test. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]
mislabeled as the ideal. Data generated by combining all of the published studies have the potential to be biased in a number of ways. It is possible that published studies come from centers that are more experienced and higher performing than might be expected across all providers. Studies are likely to come from academic centers, and thus study populations may not reflect what is typical in community practices. Alternatively, centers with poorer results may be less inclined to publish their findings, biasing the results further. Overall, there was high heterogeneity ($I^2 = 84.4\%$) seen on meta-regression, and visual inspection of the overall funnel plot did have some asymmetry with Egger's test of the intercept being significant ($P = .003$) (see Supporting Figure 1 in the online version of this article). These findings suggest there may be some degree of publication bias. Certainly, expected results should shift over time as techniques are refined, indications change, and postoperative medications improve, and this may not be reflected in published literature. Although this review only included patients with CRS undergoing ESS, there certainly is variability within this population with regard to underlying pathophysiology, extent and technique of surgery, adjuvant procedures, and postoperative medical management that is not reflected in the variables commonly reported across studies and used in this analysis. There is also variability in study quality, with some data coming from prospective clinical trials and others representing retrospective reviews, but none providing robust controlled comparisons to nonsurgical interventions. Lastly, meta-regression is a study-level analysis, which should best be thought of as hypothesis generating.

Some associations seen across studies may not be replicated on an individual patient level. For example, individual studies have reported that changes in QOL are lower in patients undergoing revision ESS compared to primary ESS, however, in this analysis, studies that reported a higher percentage of revision surgeries had greater improvement in SNOT-22 scores. This apparent discrepancy could be explained if one hypothesized that surgeons who perform more revision surgery are more experienced and thus on average have better outcomes across their entire patient population, even if patients undergoing revision surgery have worse outcomes compared to those undergoing primary surgery. Discerning these relationships requires individual, patient-level data.

One might theorize how to develop an ideal reference point for quality improvement programs interested in expected outcomes after ESS for patients with CRS. This would require establishment of a registry and care taken to eliminate potential biases. Ideally, all patients undergoing ESS would be enrolled, as opposed to a select group, and relevant patient-specific factors would be recorded. This would allow patient-level data to be evaluated as opposed to study-level data as reported in the above meta-regressions. Outcomes from all patients would be queried, as opposed to just those who do well, eliminating problems with follow-up bias and reporting bias. Lastly, care would have to be taken that SNOT-22 values are not influenced for the purposes of artificially enhancing results. Ways in which a surgeon could game the results would be to ask patients to give them good scores, as is often done in the service industry when customer surveys are tied to reimbursement. Another way would be to increase medical treatment (i.e., burst of oral steroids) just prior to postoperative SNOT-22, giving an artificially improved snapshot rather than a more accurate representation of their long-term condition. It is only with unbiased, patient-level data across a large range of patients, providers, healthcare systems, and countries that an ideal reference and accurate estimate of variability could be established and used to guide quality improvement initiatives.

**CONCLUSION**

Studies evaluating QOL outcomes after sinus surgery using the SNOT-22 instrument universally show significant improvement after ESS. Across the published literature, the magnitude of change is quite variable, and appears to be influenced by a number of factors including baseline SNOT-22 score, asthma prevalence, and length of follow-up. Findings from this study provide a point estimate and range of expected changes after surgery that could inform quality improvement initiatives. Future efforts to report unbiased data and patient level metrics across a wide spectrum of patients and providers will allow future analyses to improve the accuracy and precision of this estimate and range.

**BIBLIOGRAPHY**

Schlosser RJ, Hyer JM, Smith TL, et al. Depression-specific outcomes
29. Steele TO, Mace JC, DeConde AS, et al. Does comorbid obesity impact
22. Sedaghat AR, Metson R, Gray ST. Impact of day of week on outcomes of
34. El Rassi E, Mace JC, Steele TO, Alt JA, Smith TL. Improvements in
42. Zhang Z, Palmer JN, Morales KH, et al. Culture-inappropriate antibiotic
41. Lal D, Rounds AB, Divekar R. Gender-specific differences in chronic rhino-
57. Mascarenhas JG, da Fonseca VM, Chen VG, Fonseca VM, Cursino MM, Mendes Neto JA, Grecoroo LC. Translation, cross-cultural adaptation and validation of
50. Kosugi EM, Chen VG, Fonseca VM, Cursino MM, Mendes Neto JA, Grecoroo LC. Translation, cross-cultural adaptation and validation of
53. Snidvongs K, Pratt E, Chin D, Sacks R, Earls P, Harvey RJ. The impact of diabe-
57. Mascarenhas JG, da Fonseca VM, Chen VG, Fonseca VM, Cursino MM, Mendes Neto JA, Grecoroo LC. Translation, cross-cultural adaptation and validation of
58. Sacks PL, Snidvongs K, Rom D, Earls P, Sacks R, Harvey RJ. The impact of neo-ostegeneis on disease control in chronic rhinosinusitis after pri-
59. Sacks PL, Snidvongs K, Rom D, Earls P, Harvey RJ. The impact of neo-ostegeneis on disease control in chronic rhinosinusitis after pri-
11. Sacks PL, Snidvongs K, Rom D, Earls P, Harvey RJ. The impact of neo-ostegeneis on disease control in chronic rhinosinusitis after pri-
63. Hashemian F, Baghbanian N, Majd Z, Rouini MR, Jahanshahi J, Hashemian F. The effect of thyme honey nasal spray on chronic rhinosi-
64. Hashemian F, Baghbanian N, Majd Z, Rouini MR, Jahanshahi J, Hashemian F. The effect of thyme honey nasal spray on chronic rhinosi-
66. Saedi B, Sadeghi M, Akhavan-Khalghi N, Seifmahz H. Impact of endo-
69. Amali A, Saedi B, Rahavi-Ezabadi S, Ghazavi H, Hassanpour N. Long-
71. Devekar B, Patel N, Jin J, et al. Symptom-based clustering in chronic rhino-


