Published in final edited form as:

Int Forum Allergy Rhinol. 2014 September ; 4(9): 725–733. doi:10.1002/alr.21350.

Comparative effectiveness of medical and surgical therapy on olfaction in chronic rhinosinusitis: a prospective, multi-institutional study

Adam S. DeConde, MD^{*}, Jess C. Mace, MPH, CCRP^{*}, Jeremiah A. Alt, MD PhD⁺, Rodney J. Schlosser, MD[‡], Timothy L. Smith, MD MPH^{*}, and Zachary M. Soler, MD MSc[‡]

^{*}Division of Rhinology and Sinus Surgery, Oregon Sinus Center, Department of Otolaryngology – Head and Neck Surgery; Oregon Health & Science University, Portland, Oregon, USA

⁺Division of Head and Neck Surgery, Rhinology-Sinus and Skull Base Surgery Program, Department of Surgery; University of Utah, Salt Lake City, Utah, USA

[‡]Division of Rhinology, Nose & Sinus Center, Department of Otolaryngology- Head and Neck Surgery; Medical University of South Carolina, Charleston, South Carolina, USA

Abstract

Background—Evidence comparing the impact of medical and surgical management of chronic rhinosinusitis on olfactory function is limited. This study evaluates olfactory outcomes in patients who failed initial medical management and elect either continued medical management or endoscopic sinus surgery (ESS) followed by medical management.

Methods—Adult subjects were prospectively enrolled into a non-randomized, multi-institutional cohort. Baseline characteristics, quality-of-life and objective clinical findings were collected along with two quality-of-life disease-specific measures, the Rhinosinusitis Disability Index (RSDI) and Sinonasal Outcome Test (SNOT-22). The primary outcome measure was the post-treatment change (6 months) in the Brief Smell Identification Test (B-SIT). Bivariate and multivariate analyses compared B-SIT changes by treatment type while controlling for baseline cofactors.

Results—Subjects (n=280) were enrolled between March, 2011 and May, 2013. Baseline B-SIT scores were comparable between medical and surgical treatment groups (8.8(3.2) vs 9.0(3.2); p=0.703). Subjects with baseline impaired olfaction (n=83; 29.6%) experienced mean B-SIT improvement in both the medical (n=17, 2.3(2.8), p=0.005) and surgical (n=66, 2.1(3.0), p<0.001) cohort. 38.6% of subjects with impaired olfaction return to normal olfaction at follow-up with no difference identified between treatment modalities (p=0.803). Multivariate analyses identified prior surgery as a predictor of less improvement regardless of treatment modality in patients with baseline impaired olfaction. Average changes in B-SIT scores were comparable between treatment groups (p>0.050).

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

Potential Conflicts of Interest: None

Accepted for oral presentation to the American Rhinologic Society meeting at the 117th annual Combined Otolaryngologic Spring Meeting (COSM), Las Vegas, NV, May 14–18th, 2014.

Conclusion—Subjects electing ESS experienced gains in olfaction comparable to subjects electing continued medical management. Further study with larger sample size and more sensitive measures of olfaction are needed to determine differences between treatment groups.

MeSH Key Words

Olfaction disorders; sinusitis; inflammation; smell; quality-of-life; therapeutics

INTRODUCTION

Chronic rhinosinusitis (CRS) is the most common cause of olfactory dysfunction accounting for 14–30% of such cases^{1–4} with 28.3–69.0% of CRS patients reporting olfactory dysfunction.^{5–8} Patients with CRS and olfactory dysfunction can be treated both surgically with endoscopic sinus surgery (ESS) and medically with antibiotics, oral and topical steroids. Prior studies have examined the effectiveness of endoscopic sinus surgery (ESS) in patients with CRS and olfactory dysfunction.^{7,9–11} These studies report mixed results with a variety of outcome measures, but are uniformly limited by lack of a medical control group. To date, no study has compared the olfactory outcomes in CRS of medical management to surgical management.

Study of olfactory outcomes importantly relies on various objective olfactory testing modalities;¹² however, detailed olfactory testing is not routinely done in the clinical setting and physicians must often rely upon patient self-reporting of olfactory changes. Prior studies have suggested that self-reporting of olfactory loss may underestimate measured olfactory loss,¹³ and it remain unclear the degree to which subjective changes correlate with objective findings. Furthermore, few studies have investigated the interplay between other cardinal symptoms of CRS and olfactory dysfunction.

The aims of the present study are: 1) to compare olfactory outcomes in surgically and medically managed subjects with CRS, 2) compare change in subjective measures of sinus disease and change in olfaction scores and 3) to identify baseline characteristics that forecast changes in olfaction. We hypothesize that patients treated surgically will experience greater gains in olfaction than patients treated medically, and that olfactory test scores and subjective olfaction will correlate.

MATERIALS AND METHODS

Patient Population and Data Collection

English speaking, adult patients (≥ 18 years) with CRS were enrolled into an ongoing prospective, observational cohort investigation utilizing three academic, tertiary, rhinology practices (Oregon Health & Science University {Portland, OR}, the Medical University of South Carolina {Charleston, SC}, and Stanford University {Palo Alto, CA}) after providing informed consent. Study inclusion criteria consisted of a current diagnosis of symptomatic refractory CRS as defined by the 2007 Adult Sinusitis Guidelines (Rosenfeld 2007);⁶ prior treatment with oral, broad spectrum, or culture-directed antibiotics (≥ 2-weeks); and either topical nasal corticosteroid sprays (≥ 3-weeks) or a 5-day trial of systemic steroid therapy.

The Institutional Review Board at each practice monitored and approved all investigational protocols.

Patients were asked to provide demographic, social, and medical history data including age, gender, race, ethnicity, current tobacco use, nasal polyposis, depression, asthma, allergies (confirmed skin prick or radioallergosorbent testing), acetylsalicylic acid (ASA) sensitivity, cystic fibrosis, and history of prior sinus surgery.

Subjects elected to either undergo continued medical management or ESS as the next treatment option for refractory CRS. Patients were then observed and evaluated for the primary outcome of change in olfactory function. Patients underwent evaluation with disease severity measures and disease-specific quality of life measures at initial enrollment as well as during 6-, 12- and 18-month follow-up visits when possible. Patients without at least 6-month follow-up or who elected to crossover from continued medical management to surgical management were excluded from final analyses.

Disease Severity Measures

B-SIT Evaluations—Olfactory function was measured at the initial enrollment period and at follow-up visits and operationalized using The Brief Smell Identification Test (B-SIT; Sensonics, Inc., Haddon Heights, NJ). The B-SIT is a validated 12-item, standardized, noninvasive test of olfactory function that employs 12 microencapsulated odorant strips in a “scratch-‘n-sniff” format (range: 0–12), with higher scores indicating a better sense of smell. Complete B-SIT scores ≥ 9 are defined as “normal” for healthy males and females of all ages (Doty 2001).¹² Improvement in B-SIT scores was operationalized by calculating the difference between pre-treatment and final post-treatment score. Subjects were dichotomized by improvement status, defined by reporting a normal (≥ 9) follow-up B-SIT score.

Computed Tomography and Endoscopy Staging—Baseline computed tomography and endoscopy scores were obtained at initial enrollment for all subjects and evaluated and staged in accordance with the Lund-Mackay (range: 0–24)¹⁴ and Lund-Kennedy (range: 0–20)¹⁵ scoring systems, respectively. The enrolling physician at each site scored both exams at the time of enrollment as well as endoscopic exams during follow-up visits.

Disease-Specific Quality of Life Measures—Study subjects completed two CRS-specific QOL instruments: the Rhinosinusitis Disability Index (RSDI) and the 22-item Sinonasal Outcome Test (SNOT-22) at both baseline and follow-up visits.^{16,17} The enrolling physicians at each enrollment site were blinded to all survey responses for the study duration.

Olfactory Domain Item Scores—In an effort to investigate the correlation between subjective olfaction and objective olfactory testing changes in select questions from the SNOT-22 and RSDI related to taste and smell were investigated. Specifically, SNOT-22 item #21 asks subjects to rate their “sense of smell/taste” on a Likert scale ranging from 0 to 5 with 0 representing “No problem” and 5 representing “Problem as bad as it can be.” Similarly, RSDI item #7 of the physical subscale also investigates disease-specific impact on

smell and taste by asking subjects to report the frequency with which “Food does not taste good because of my change in smell.” using a Likert scale from 0 to 4 with 0 representing “never” and 4 representing “always.”

CRS Domain Item Scores—In addition to subjective measures of smell and taste, we investigated the association of the other three “cardinal” symptoms of CRS as defined by clinical practice guidelines: facial pain/pressure, nasal obstruction, and thick nasal discharge (Rosenfeld 2007).⁶ These symptoms were measured using the SNOT-22 in questions #6, #10 and #22 and in the RSDI physical subscale questions #1, #6, and #10. Facial pain/pressure was explicitly asked by question #6 on the SNOT-22, while question #1 of the RSDI physical subscale asks whether “the pain or pressure in my face makes it difficult for me to concentrate.” Nasal obstruction was assessed using question #22 of the SNOT-22 as well as question #10 of the RSDI physical subscale, which asks subjects to rate whether they “have difficulty with exertion due to [their] nasal obstruction.” Thick nasal discharge was operationalized by both question #6 of the SNOT-22 (“thick nasal discharge”), and question #6 of the RSDI physical subscale (“I am inconvenienced by my chronic runny nose”).

Data Management and Statistical Analysis

All investigational data was collected and transferred into a central, relational database (Microsoft Office Access 2007; Microsoft Corp., Redmond, WA) by the study coordinator (JCM) using standardized clinical research forms. All statistical analyses were completed using a commercially available software application (IBM SPSS v.22; IBM Corp., Armonk, NY). A preliminary power analysis was performed. Asymmetry in enrollment between the cohorts (smaller medical than surgical cohort) was anticipated and pre-study power-analyses were estimated with a 4:1 ratio of surgical to medical patients in order to be able to detect within a single question difference on the B-SIT if 32 patients in the medical cohort and 128 in the surgical cohort are enrolled. Descriptive and graphical analysis was used to check assumptions of normality and linearity for all study variables. Two-tailed independent t-tests and Mann-Whitney U tests were used to compare all continuous scale data between treatment groups where appropriate, while Chi-square (χ^2) tests were utilized to compare prevalence of comorbid risk factors between treatment groups. Matched paired t-test or Wilcoxon signed-rank test were used to assess improvement in B-SIT scores and survey responses over time for CRS subgroups of interest. Bivariate correlation between B-SIT scores and QOL survey item scores was evaluated between subgroups using the Spearman’s rank correlation coefficient (r_s). Simple linear regression modeling was performed to adjust for independent cofactors associated with changes in B-SIT scores. Preliminary models included treatment decision as the main exposure variable of interest and independent covariates with screened univariate significance ($p < 0.250$) while controlling for clinical variables of interest including: age, gender, enrollment site, follow-up time, and nasal polyposis. Change in B-SIT scores (follow-up score minus baseline score) was the main outcome of interest. Additional covariates were introduced into preliminary models to assess confounding of the effect estimate for the treatment variable. Final models were elected with a forward selection ($p = 0.050$) and backward elimination ($p = 0.100$) manual stepwise process. Results of linear regression modeling were reported using the adjusted effect estimate (β), standard error (SE), 95% confidence intervals (CI), and corresponding p-value.

RESULTS

Baseline Characteristics

A total of 417 patients met inclusion criteria and were enrolled into the study. Of this original cohort, a total of 280 (67.1%) had both a baseline and at least a six-month follow-up B-SIT score available for analysis. Medically treated subjects were found to have a statistically similar prevalence of follow-up (63.7%) compared to surgically treated subjects (69.0%; $p=0.340$). The final study cohort with follow-up consisted of 58 (20.7%) medically treated patients and 222 (79.2%) surgically treated patients. Subjects with follow-up were significantly older compared to patients without follow-up (51.5(14.7) vs. 46.7(14.4) years; $p=0.001$). Additionally there was a higher prevalence of patients with comorbid depression in subjects with follow-up compared to subjects without follow-up (19.3% vs. 10.9%; $p=0.031$). Patients electing to undergo surgical management reported higher burden of disease on both baseline SNOT-22 (52.9(19.0) versus 46.3(18.3); $p=0.018$) and RSDI (46.2(25.0) versus 38.3(23.5); $p=0.030$) survey scores. Mean objective measures of disease (CT, endoscopy and B-SIT), the prevalence of medical comorbidities, and demographic information were otherwise comparable between treatment groups (Table 1).

Olfactory Outcome Measures

B-SIT Scoring for Total CRS Cohort—The majority of all enrolled patients (197/280; 70.4%) had normal B-SIT scores at baseline. The prevalence of normal B-SIT scores was similar for both the medical treatment group (70.7%) and the surgical treatment group (70.3%). Post-treatment B-SIT scores remained in the normal range for patients with normal baseline B-SIT scores regardless of medical or surgical treatment. Patients with baseline abnormal B-SIT (scores < 9) had significant improvement in mean B-SIT scores in both the medical group ($n=17$, 2.3(2.8), $p=0.005$; Figure 1) and the surgical treatment group ($n=66$, 2.1(3.0), $p<0.001$; Figure 2). Mean changes in B-SIT scores in patients with baseline abnormal B-SIT demonstrated no differences between medical and surgical treatment groups (all $p>0.050$).

B-SIT Scoring for CRSwNP and CRSsNP—Findings in CRS subgroups defined by polyp status closely mirrored the overall CRS findings. Patients with baseline abnormal olfaction undergoing medical therapy demonstrated improvement in both CRSwNP and CRSsNP; however, these improvements over time were not statistically significant for either CRSwNP ($n=10$, $p=0.632$) or CRSsNP ($n=7$, $p=0.391$) subgroups (Table 2). For surgically treated patients with abnormal baseline olfaction, improvement was observed both in CRSwNP ($n=41$, $p=0.001$) and CRSsNP ($n=25$, $p=0.004$) subgroups. Mean change in B-SIT score after treatment was compared between polyp subgroups undergoing surgery and medical management with no significant differences (all $p \geq 0.303$, Table 2).

SNOT-22 Olfactory Domain Item Scores—Without dichotomizing for nasal polyp status, both medical and surgical subjects demonstrated significant improvement from baseline scores in patient reported sense of smell/taste as assessed by survey item #21 of the SNOT-22 instrument ($p=0.009$ and $p<0.001$, respectively). Subset analysis of subjects with CRSsNP showed significant improvement in the surgical group ($p<0.001$), but not for the

medically treated group ($p=0.092$). Subjects with CRSwNP reported significantly improvement in olfactory function in both the medical group ($p=0.024$) and the surgical group ($p<0.001$). Mean improvements in the medical and surgical cohorts demonstrated no significant difference ($-1.0(1.7)$ vs. $-0.7(2.0)$; $p=0.181$) even when stratified by CRSwNP ($-0.9(1.8)$ vs. $-1.5(1.8)$; $p=0.181$, respectively) and CRSsNP ($-0.6(2.1)$ vs. $-0.8(1.6)$; $p=0.515$, respectively).

RSDI Olfactory Domain Item Scores—Without dichotomizing for nasal polyp status, similar findings were seen when subjective olfaction was evaluated using survey item #7 of the RSDI physical subscale. Both medical and surgical cohorts reported significant improvement at follow-up ($p=0.036$ and $p<0.001$, respectively). After surgery, significant improvements were seen in both CRSwNP and CRSsNP subgroups ($p<0.001$). Improvements were also seen in CRSwNP and CRSsNP subgroups after ongoing medical management, but these were no longer significant ($p=0.142$ and $p=0.126$, respectively). Mean improvement for surgical patients was greater than medical patients ($-0.8(1.3)$ vs. $-0.3(1.1)$; $p=0.005$). Similarly, mean improvement was also greater in CRSsNP ($-0.7(1.3)$ vs. $-0.3(1.0)$; $p=0.026$, respectively) and CRSwNP groups ($-1.0(1.4)$ vs. $-0.4(1.1)$; $p=0.091$, respectively). Significant improvements in subjective smell and taste, as measured by the RSDI and SNOT-22 survey items, strongly correlated ($n=280$; $r_s=0.610$, $p<0.001$).

Correlation Between B-SIT Scores and QOL Survey Item Scores—Baseline B-SIT scores and baseline subjective ratings of olfaction significantly correlated for both the SNOT-22 (Q#21) and RSDI (Q#7-Physical subscale) survey score items ($r_s=-0.473$; $p<0.001$; $r_s=-0.267$; $p<0.001$, respectively) for all subjects ($n=280$). Subjects with abnormal baseline B-SIT scores ($n=83$) were found to have similar significant correlations between baseline B-SIT scores and baseline SNOT-22 item scores ($r_s=-0.384$; $p<0.001$) yet no correlation between baseline B-SIT scores and RSDI olfactory item scores exists ($r_s=-0.101$; $p=0.362$). No significant correlations were found between baseline B-SIT scores and baseline olfactory item survey scores for subjects with normal olfaction ($p=0.078$).

Improvement in B-SIT scores and subjective ratings of olfaction weakly correlated on both the SNOT-22 and RSDI survey score items relating to olfaction ($n=280$; $r_s=-0.170$; $p=0.004$; $r_s=-0.129$; $p=0.032$, respectively). Subjects with abnormal baseline B-SIT scores had a correlation between improvement in B-SIT and change in SNOT-22 olfaction item scores ($r_s=-0.433$; $p<0.001$) as well as RSDI olfaction item scores ($r_s=-0.289$; $p=0.008$). Subgroup analysis demonstrated a stronger correlation for patients that elected surgical treatment with a moderate correlation between objective and subjective measure on the SNOT-22 ($r_s=-0.549$; $p<0.001$). RSDI item score changes also demonstrated a significant correlation to B-SIT scores changes ($r_s=-0.289$; $p=0.019$). Medically treated subjects did not demonstrate significant correlation between B-SIT changes or SNOT-22 or RSDI olfaction item scores.

Clinically significant improvement of olfaction after therapy

In addition to evaluating mean olfaction changes in the cohorts, the results were interpreted as percentage of subjects with baseline abnormal B-SIT scores returned to normal by the

intervention. The frequency of B-SIT change scores were calculated for subjects with only baseline abnormal olfaction undergoing medical treatment (Figure 3) and surgical treatment (Figure 4). Of the 83 total patients with abnormal baseline olfaction, 32 (38.6%) achieved normal B-SIT scores (≥ 9) over time. No significant differences in the prevalence of patients with post-treatment normal olfactory scores between the medical treatment group (7/17) and surgical treatment (25/66) were found in patients with baseline abnormal olfaction (41.2% vs. 37.9%; $p=0.803$). After stratifying by nasal polyp status and treatment modality, comparable prevalence of post-treatment normal olfactory scores (≥ 9) were found for medical treatment group with polyps (4/10; 40.0%), medical treatment group without polyps (3/7; 42.9%), surgical treatment group with polyps (15/41; 36.6%), and surgical treatment group without polyps (10/25; 40.0%).

Correlations Between Change in Cardinal Symptoms of CRS and B-SIT Scores

All correlations between baseline and changes in B-SIT scores and patient-reported baseline and change scores for symptom-specific QOL items are listed in Table 3. Patients reporting thick nasal discharge and nasal airway obstruction correlated with lower B-SIT scores at baseline ($r_s = -0.137$; $p=0.022$ and $r_s = -0.158$; $p=0.008$, respectively) and with less improvement on B-SIT ($r_s = 0.162$; $p=0.007$ and -0.243 ; $p<0.001$, respectively). Subjects with CRSwNP demonstrated negative correlation with changes in nasal discharge on both the SNOT-22 and RSDI survey items ($r_s = -0.272$, $p<0.006$; $r_s = -0.403$, $p<0.001$, respectively) as well as nasal airway obstruction on the SNOT-22 ($r_s = -0.311$, $p=0.002$) and change in B-SIT scores. In contrast, change in B-SIT in subjects without nasal polyps did not correlation with nasal obstruction, yet there was a negative correlation with purulent discharge ($r_s = -0.208$, $p=0.005$) on the SNOT-22 and on the RSDI ($r_s = -0.184$, $p=0.015$).

Regression Modeling For Change in B-SIT Scores

Preliminary covariate screening was conducted for three regression models in an effort to identify any variable which might forecast mean changes in B-SIT scores after intervention in all study patients, as well as patient subgroups with and without normal baseline olfaction. Without adjustment for any significant cofactors, treatment decision was not a significant predictor of improvement in B-SIT scores ($n=280$; $p=0.414$), even when evaluating subgroups with baseline normal olfaction ($n=197$; $p=0.262$) and baseline abnormal olfaction ($n=83$; $p=0.830$). After adjustment, only a history of previous sinus surgery was found to be strongly associated with less improvement in B-SIT scores for subjects with abnormal baseline olfaction while accounting for 20.3% of the model variability ($\beta(\text{SE}) = -2.0(0.7)$; 95% CI: $-3.3, -0.6$; $p=0.004$). No significant predictors were found to be associated with improvement in B-SIT scores for subjects with baseline normal olfaction. Models built to evaluate the total cohort ($n=280$) found that only ASA sensitivity ($\beta(\text{SE}) = 1.8(0.6)$; 95% CI: $0.4, 2.7$; $p=0.009$), and baseline RSDI nasal discharge survey scores ($\beta(\text{SE}) = 0.5(0.1)$; 95% CI: $0.3, 0.8$; $p<0.001$) were associated with greater improvement in B-SIT scores, whereas baseline RSDI facial/pain pressure scores ($\beta(\text{SE}) = -0.4(0.2)$; 95% CI: $-0.7, -0.1$; $p=0.007$) was associated with less improvement in B-SIT scores.

DISCUSSION

This prospective, multi-institutional, comparative effectiveness study evaluated olfactory outcomes of patients who failed a trial of medical management and were considered to be candidates for ESS. One fifth of the patients elected continuing medical therapy and four-fifths elected ESS in conjunction with medical therapy. Patients with baseline normosmia stayed normosmic regardless of treatment modality. At follow-up, subjects in both the surgical and medical cohorts perceived improved olfaction (on RSDI and SNOT-22) and demonstrated increased scores on B-SIT regardless of polyp status. Both the magnitude of improvement and percentage of patients improved on B-SIT were clinically and statistically comparable between treatment cohorts. Two-thirds of subjects with impaired baseline olfaction had interval improvement with 38.6% returning to normal olfaction. Although a lack of statistically significant differences in olfactory outcome between the medical and surgical cohorts, and subgroups of CRSwNP and CRSsNP, may be a reflection of a small sample size there was no trend toward a clinically meaningful difference between the two cohorts either. After controlling for baseline characteristics and treatment modality in patients with baseline abnormal olfaction only history of prior sinus surgery was found to be strongly associated with less improvement in olfaction.

The present study also sought to identify if gains in the cardinal symptoms of CRS were paralleled by olfaction. Cardinal symptoms are important to the clinician who must base clinical decisions on history.⁶ Furthermore, the potentially irreversible consequences of olfactory sensorineural loss may limit olfaction gains despite resolution of nasal discharge, nasal obstruction and facial pain/pressure.^{18–20} Change in nasal discharge negatively correlated to change in B-SIT changes; that is, as nasal discharge decreased, B-SIT scores tended to improve. Subjects with CRSwNP also had olfactory improvements that weakly paralleled improvement in nasal obstruction. In contrast, patients with CRSsNP only correlated gains in olfactory improvement with decrease in nasal discharge. This dichotomy of subjective symptoms may reflect an underlying conductive olfaction loss that occurs from obstructive polyps whereas CRSsNP olfactory loss is dominated by unchecked inflammation. Further study will be required to elucidate potential pathophysiologic differences between subtypes of CRS.

Prior study of CRS and olfaction is dominated by prospective cohort studies investigating the impact of ESS on olfaction.^{5,7,21,22} These studies report mixed results with a variety of outcome measures and either do not have a medical control arm^{5,7,21,22} or are limited by a lack of predefined treatment criteria (i.e., failure of “maximal” medical management).²³ Patients enrolled in the present study by definition had failed systemic therapies of antibiotics and/or corticosteroids in order to be considered a surgical candidate. This study represents the first prospective, multi-institutional, comparative effectiveness study of medical and surgical management of CRS and no difference was identified between surgical and medical cohorts regardless of polyp status.

There are two important limitations to the present study’s findings that warrant discussion: selection treatment bias and study power. Inherent in a non-randomized comparative effectiveness study is the potential for selection treatment bias, which threatens the internal

validity of the study. Potential treatment selection bias is best controlled by employing a randomized controlled trial (RCT) design, but there are many issues that preclude a RCT. Patients failing medical management may be reluctant to consent for an RCT in which chance alone would determine their impending treatment, particularly when one of the proposed treatments is a surgical procedure. Additionally, physicians and IRBs have difficulty agreeing to participate in an RCT when prospective, multi-institutional studies demonstrate clinically significant improvement following ESS in patients who have failed medical management. RCT's also provide strict control of patient variables (e.g., drug regimens, adherence, concomitant treatments, etc.), which helps to maintain internal validity but simultaneously decrease generalizability.²⁴ The present study is an effectiveness study that attempts to balance the inherent inverse relationship between internal and external validity.

In an effort to control for measurable potential confounders between the cohorts we did evaluate the two treatment groups for baseline differences in QOL and other clinical factors associated with more severe disease. Comparisons between cohorts did identify a higher burden of disease as measured by baseline QOL in the surgical cohorts, but the treatment groups were otherwise similar across other potential measured confounders including baseline olfaction. The multi-institutional observational design allows for heterogeneity in medical therapy in the medical cohort as well as adjunct medical therapy in the surgical cohort. Extent and type of surgical intervention also varies across sites, but was not measured in this study's outcomes. Regression was undertaken to control for possible confounders associated with site differences, however these models did not identify treatment modality, treatment site, or baseline quality of life metrics as significant predictors of olfactory outcomes.

The present study is also limited by the small asymmetric sample sizes. Baseline normosmics were anticipated, and although there is heterogeneity in measures and reporting, prior cohort studies baseline normosmic rates range between 25% and 50%.^{5,7,10,21,22} The present study has a normosmic rate of 70.4% thus unexpectedly limiting our sample sizes to less than we anticipated with our initial power calculation. It is possible that the differential between prior studies and the current study is a product of sampling error, but measured confounders and objective measures of disease are comparable to prior cohorts.⁷ Furthermore, the B-SIT is not as effective at discriminating various levels of smell identification.²⁵ This potential lack of discrimination may lead to an under reporting of olfactory dysfunction which may bias effect estimates of change towards null hypotheses and limit our ability to detect true differences between treatment modalities. The high level of normosmics led to the evaluation of the cohort in sub-groups of normal olfaction and impaired olfaction. This subgroup analysis avoided the ceiling effect of normosmics obfuscating significant change in the impaired olfaction sub-group, but simultaneously decreased the power to detect a difference between the medical and surgical cohorts.

Despite the lack of power to detect a statistically significant difference these results and subgroup analyses still provide important findings. Mean differences in interval change between treatments and across subgroups are all within the resolution (one question) of the B-SIT instrument (Table 2). Similarly, the frequency of patients with impaired olfaction

achieving post-treatment normosmia is clinically comparable between treatment modalities (41.2% vs 37.9% for surgical vs medical therapy) although not statistically significant. Further study with larger sample sizes will be required to corroborate this finding along with CRS subtype analysis, because it is possible that sample recruitment did not occur in a random manner. The difference between the cohort means could potentially drift apart as sample size is increased.

This study raises a number of issues that could be pursued in future studies. A larger comparative effectiveness study could better account for selection bias through use of propensity scoring techniques or instrumental variable analysis. Additionally, the follow-up of six months may not be necessary and data collected at earlier time points might better elucidate the ideal follow-up necessary for clinical olfaction outcomes. Finally, study comparing the type and timing of medical management required between the post-surgical and the exclusively medical therapy groups would provide incite into the degree to which surgical therapy might deescalate and deintensify systemic therapies and the durability of medical therapy.

CONCLUSION

This prospective, multi-institutional study demonstrates improvement in olfaction in both medically and surgically treated patients. No difference in mean improvement or percentage of patients that improved was detected between treatment modalities. Multivariate analysis accounting for baseline differences between cohorts identified only history of prior surgery as a predictor of less improvement. A high prevalence of baseline normosmia, use of an olfaction measure with limited discrimination, and a medical cohort of limited size may have restricted the ability to detect a significant difference between treatment modalities.

Acknowledgments

Financial Disclosures: Timothy L. Smith, MD, MPH, Zachary M. Soler, MD, MSc, and Jess C. Mace, MPH, CCRP, are supported by a grant from the National Institute on Deafness and other Communication Disorders (NIDCD), one of the National Institutes of Health, Bethesda, Maryland. (RO1 DC005805; PI/PD: TL Smith). Timothy L. Smith, MD, MPH is also a consultant for IntersectENT, Inc. (Menlo Park, CA), which is not affiliated with this investigation.

References

1. Harris R, Davidson TM, Murphy C, Gilbert PE, Chen M. Clinical evaluation and symptoms of chemosensory impairment: one thousand consecutive cases from the Nasal Dysfunction Clinic in San Diego. *Am J Rhinol.* 2006; 20(1):101–108. [PubMed: 16539304]
2. Miwa T, Furukawa M, Tsukatani T, et al. Impact of olfactory impairment on quality of life and disability. *Arch Otolaryngol Head Neck Surg.* 2001; 127(5):497–503. [PubMed: 11346423]
3. Seiden AM, Duncan HJ. The diagnosis of a conductive olfactory loss. *Laryngoscope.* 2001; 111(1): 9–14. [PubMed: 11192906]
4. Temmel AFP, Quint C, Schickinger-Fischer B, et al. Characteristics of olfactory disorders in relation to major causes of olfactory loss. *Arch Otolaryngol Head Neck Surg.* 2002; 128(6):635–641. [PubMed: 12049556]
5. Schriever VA, Gupta N, Pade J, Szewczynska M, Hummel T. Olfactory function following nasal surgery: a 1-year follow-up. *Eur Arch Otorhinolaryngol.* 2013; 270(1):107–111. [PubMed: 22382399]

6. Rosenfeld RM, Andes D, Neil B, et al. Clinical practice guideline: Adult sinusitis. *Otolaryngol Head Neck Surg.* 2007; 137(3 suppl):S1–S31. [PubMed: 17761281]
7. Litvack JR, Mace J, Smith TL. Does olfactory function improve after endoscopic sinus surgery? *Otolaryngol Head Neck Surg.* 2009; 140(3):312–319. [PubMed: 19248934]
8. Alt JA, Mace JC, Buniel MC, Soler ZM, Smith TL. Predictors of olfactory dysfunction in rhinosinusitis using the brief smell identification test. *Laryngoscope.* 2014; 124(7):E259–E266. [PubMed: 24402746]
9. Minovi A, Hummel T, Ural A, Draf W, Bockmuhl U. Predictors of the outcome of nasal surgery in terms of olfactory function. *Eur Arch Otorhinolaryngol.* 2008; 265(1):57–61. [PubMed: 17676329]
10. Sugiyama K, Hasegawa Y, Sugiyama N, et al. Smoking-induced olfactory dysfunction in chronic sinusitis and assessment of brief University of Pennsylvania Smell Identification Test and T&T methods. *Am J Rhinol.* 2006; 20(5):439–444. [PubMed: 17063734]
11. Pade J, Hummel T. Olfactory function following nasal surgery. *Laryngoscope.* 2008; 118(7):1260–1264. [PubMed: 18438263]
12. Doty RL, Mishra A. Olfaction and Its Alteration by Nasal Obstruction, Rhinitis, and Rhinosinusitis. *Laryngoscope.* 2001; 111(3):409–423. [PubMed: 11224769]
13. Murphy C, Schubert CR, Cruickshanks KJ, et al. Prevalence of olfactory impairment in older adults. *JAMA.* 2002; 288(18):2307–2312. [PubMed: 12425708]
14. Lund VJ, Mackay IS. Staging in rhinosinusitis. *Rhinology.* 1993; 31(4):183–184. [PubMed: 8140385]
15. Lund VJ, Kennedy DW. Staging for rhinosinusitis. *Otolaryngol Head Neck Surg.* 1997; 117(3 Pt 2):S35–S40. [PubMed: 9334786]
16. Benninger MS, Senior BA. The development of the Rhinosinusitis Disability Index. *Arch Otolaryngol Head Neck Surg.* 1997; 123(11):1175–1179. [PubMed: 9366696]
17. 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]
18. Stevens MH. Steroid-dependent anosmia. *Laryngoscope.* 2001; 111(2):200–203. [PubMed: 11210860]
19. Raviv JR, Kern RC. Chronic sinusitis and olfactory dysfunction. *Otolaryngol Clin North Am.* 2004; 37(6):1143–1157. [PubMed: 15563907]
20. Kern RC. Candidate's Thesis: Chronic Sinusitis and Anosmia: Pathologic Changes in the Olfactory Mucosa. *Laryngoscope.* 2000; 110(7):1071–1077. [PubMed: 10892672]
21. Katotomichelakis M, Simopoulos E, Tripsianis G, et al. Improvement of olfactory function for quality of life recovery. *Laryngoscope.* 2013; 123(11):E10–E16. [PubMed: 23686475]
22. Oka H, Tsuzuki K, Takebayashi H, et al. Olfactory changes after endoscopic sinus surgery in patients with chronic rhinosinusitis. *Auris Nasus Larynx.* 2013; 40(5):452–457. [PubMed: 23312885]
23. Baradaranfar MH, Ahmadi ZS, Dadgarnia MH, et al. Comparison of the effect of endoscopic sinus surgery versus medical therapy on olfaction in nasal polyposis. *Eur Arch Otorhinolaryngol.* 2014; 271(2):311–316. [PubMed: 23700266]
24. Rothwell PM. External validity of randomised controlled trials: “to whom do the results of this trial apply?”. *Lancet.* 2005; 365(9453):82–93. [PubMed: 15639683]
25. Menon C, Westervelt HJ, Jahn DR, Dressel JA, O'Bryant SE. Normative Performance on the Brief Smell Identification Test (BSIT) in a Multi-Ethnic Bilingual Cohort: A Project FRONTIER Study. *Clinical Neuropsychol.* 2013; 27(6):946–961.

Figure 1.
Average B-SIT scores for the medical treatment group

Figure 2.
Average B-SIT scores for the surgical treatment group

Figure 3.
Distribution of change in B-SIT scores for the medical treatment group (n=17) with baseline abnormal olfaction

Figure 4.
Distribution of Change in B-SIT Scores for the Surgical Treatment Group (n=66) with
Baseline Abnormal Olfaction

Table 1

Comparison of Baseline Characteristics of Medical and Surgical Cohorts

Demographics:	Medical treatment (n=58)		Surgical treatment (n=222)		p-value
	Mean (SD)	N(%)	Mean (SD)	N (%)	
Follow-up (mos.)	11.6 (5.3)		12.1 (5.6)		0.543
Age (yrs)	50.3 (15.0)		51.9 (14.6)		0.487
Males		23 (39.7)		119 (53.6)	
Females		35 (60.3)		103 (46.4)	0.358
White/Caucasian		51 (87.9)		187 (84.2)	0.438
African American		3 (5.2)		12 (5.4)	>0.999
Asian		1 (1.7)		9 (4.1)	0.693
American Indian/Native Alaskan		2 (3.4)		1 (0.5)	0.110
Hispanic/Latino		0 (0.0)		13 (5.9)	0.077
Medical comorbidity:					
Asthma		15 (25.9)		84 (37.8)	0.089
Nasal polyps		20 (34.5)		80 (36.0)	0.826
Allergies (test confirmed)		20 (34.5)		88 (39.6)	0.472
ASA sensitivity		6 (10.3)		16 (7.2)	0.429
Depression		12 (20.7)		42 (18.9)	0.761
Tobacco use		1 (1.7)		13 (5.9)	0.314
Previous sinus surgery		34 (58.6)		120 (54.1)	0.534
Baseline disease severity:					
SNOT-22 score	46.3 (18.3)		52.9 (19.0)		0.018
RSDI score	38.3 (23.5)		46.2 (25.0)		0.030
RSDI physical score	16.6 (9.5)		18.2 (9.0)		0.248
RSDI functional score	12.3 (7.9)		15.1 (9.0)		0.030
RSDI emotional score	9.4 (8.1)		12.9 (9.0)		0.007
Computed tomography score	13.4 (6.0)		12.1 (5.8)		0.132
Endoscopy score	6.9 (4.3)		6.0 (3.8)		0.128
B-SIT olfaction score	8.8 (3.2)		9.0 (3.2)		0.703

SD, standard deviation; mos., months; yrs, years; ASA, acetylsalicylic acid, SNOT-22, 22-item SinoNasal Outcome Test; RSDI, Rhinosinusitis Disability Index, B-SIT, Brief Smell Identification Test

Table 2

Comparison of change in B-SIT scores by polyp and treatment status

	Surgical Treatment Without Polyposis (n=142)			Medical Treatment Without Polyposis (n=38)			Difference in change between treatment groups
	Baseline B-SIT scores	Follow-up B-SIT Scores	Change	Baseline B-SIT scores	Follow-up B-SIT Scores	Change	
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	p-value
Normal Baseline B-SIT Score	10.8 (1.0)	10.1 (1.9)	-0.7 (2.0)	10.7 (1.1)	10.4 (1.7)	-0.2 (1.7)	0.303
Abnormal Baseline B-SIT Score	5.8 (2.0)	7.2 (2.8)	1.4 (2.1)	5.4 (1.5)	6.1 (3.4)	0.7 (2.1)	0.420
	Surgical Treatment With Polyposis (n=80)			Medical Treatment With Polyposis (n=20)			Difference in change between treatment groups
	Baseline B-SIT scores	Follow-up B-SIT Scores	Change	Baseline B-SIT scores	Follow-up B-SIT Scores	Change	
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	p-value
Normal Baseline B-SIT Score	10.9 (1.1)	9.7 (2.2)	-1.2 (2.2)	10.7 (1.2)	9.7 (2.6)	-1.0 (2.9)	0.517
Abnormal Baseline B-SIT Score	4.0 (1.8)	6.5 (3.4)	2.6 (3.4)	3.7 (2.0)	7.1 (3.1)	3.4 (2.8)	0.344

B-SIT, Brief Smell Identification Test; SD, standard deviation

Table 3

Correlations between B-SIT baseline and change scores with baseline and change in CRS cardinal symptoms scores as measured by the SNOT-22 and RSDI

B-SIT Scores				
QOL / Symptom Severity:	Correlation coefficient (r_s) Baseline Scores	p-value	Correlation coefficient (r_s) Change Scores	p-value
Total cohort (n=280):				
Facial pain/pressure				
SNOT-22	0.068	0.255	-0.050	0.405
RSDI Physical	0.083	0.167	-0.007	0.906
Nasal obstruction				
SNOT-22	-0.158	0.008	-0.162	0.007
RSDI Physical	-0.043	0.478	-0.076	0.208
Thick nasal discharge				
SNOT-22	-0.137	0.022	-0.243	<0.001
RSDI Physical	-0.190	0.001	-0.294	<0.001
CRSwNP (n=100):				
Facial pain/pressure				
SNOT-22	0.057	0.572	0.001	0.989
RSDI Physical	0.114	0.258	-0.007	0.942
Nasal obstruction				
SNOT-22	-0.214	0.032	-0.311	0.002
RSDI Physical	-0.020	0.845	-0.162	0.108
Thick nasal discharge				
SNOT-22	-0.138	0.172	-0.272	0.006
RSDI Physical	-0.188	0.061	-0.403	<0.001
CRSsNP (n=180):				
Facial pain/pressure				
SNOT-22	0.049	0.515	-0.093	0.213
RSDI Physical	0.026	0.732	-0.020	0.791
Nasal obstruction				
SNOT-22	-0.038	0.614	-0.049	0.516
RSDI Physical	-0.061	0.420	-0.002	0.979
Thick nasal discharge				
SNOT-22	-0.143	0.056	-0.208	0.005
RSDI Physical	-0.127	0.092	-0.184	0.015

r_s , Spearman's correlation coefficient; QOL, quality of life; SNOT-22, 22-item SinoNasal Outcome Test; RSDI, Rhinosinusitis Disability Index; CRSwNP, chronic rhinosinusitis with nasal polyposis; CRSsNP, chronic rhinosinusitis without nasal polyposis.