

# Development of Sinonasal Outcome Test (SNOT-22) Domains in Chronic Rhinosinusitis With Nasal Polyps

Asif H. Khan, MBBS, MPH; Matthew Reaney, MSc, CPsychol; Isabelle Guillemin, PhD; Lauren Nelson, PhD; Shanshan Qin, PhD; Siddhesh Kamat, MS, MBA; Leda Mannent, MD; Nikhil Amin, MD; Diane Whalley, PhD; Claire Hopkins, BMBCh MA, FRCS, DM

**Objectives/Hypothesis:** The 22-item Sinonasal Outcome Test (SNOT-22) is a validated chronic rhinosinusitis health-related quality-of-life outcome (HRQoL) measure; however, SNOT-22 domains have not been validated specifically for chronic rhinosinusitis with nasal polyps (CRSwNP).

**Study Design:** Validation of SNOT-22 domain structure, using data from 3 randomized, placebo-controlled, double-blinded, multicenter clinical trials of dupilumab in adults with moderate-to-severe CRSwNP.

**Methods:** Preliminary dimensional structure was derived by exploratory factor analyses of SNOT-22 data from a phase 2 trial (NCT01920893) of dupilumab for the treatment of CRSwNP. Data from 2 phase 3 clinical trials (NCT02912468 and NCT02898454) were then used for confirmatory factor analysis, and evaluated for reliability, construct validity, and responsiveness. In all three trials, the SNOT-22 was administered electronically on a tablet and trial participants were required to answer all items.

**Results:** Factor analysis supported five domains: Nasal, Ear/Facial, Sleep, Function, and Emotion. Correlations between domains were moderate to high, ranging from 0.53 (Nasal–Emotion) to 0.88 (Function–Sleep). Construct validity was mostly supported; relationships with other measures were almost always in the intended direction and magnitude. Internal consistency reliability also confirmed questionnaire structure with strong Cronbach’s alpha values (all >0.80). Moderate-to-high correlations were observed between change in SNOT-22 domain scores and other study patient-reported outcome measures, along with large effect-size estimates ( $\geq 0.7$ ), demonstrating responsiveness of the Nasal, Sleep, and Function domains. Emotion and Ear/Facial domains had small-to-moderate effect sizes.

**Conclusions:** Psychometric analyses support the validity, reliability, and responsiveness of five domains of SNOT-22 (Nasal, Ear/Facial, Sleep, Function, and Emotion) for assessing symptoms and impact on HRQoL in patients with CRSwNP.

**Key Words:** Chronic rhinosinusitis, nasal polyps, psychometrics, SNOT-22, health-related quality of life.

*Laryngoscope*, 132:933–941, 2022

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

From the Global Medical Affairs (A.H.K., L.M.), Immunology and Inflammation Development, Sanofi, Chilly-Mazarin, France; Patient-Centered Outcome Assessment (M.R., I.G., L.N., S.Q.), RTI Health Solutions, Research Triangle Park, North Carolina, U.S.A.; Health Economics and Outcomes Research (S.K., N.A.), Regeneron Pharmaceuticals, Inc., Tarrytown, New York, U.S.A.; Patient-Centered Outcome Assessment (D.W.), RTI Health Solutions, Manchester, U.K.; and the ENT Department (C.H.), Guy’s and St Thomas’ NHS Foundation Trust, London, U.K.

Additional supporting information may be found in the online version of this article.

Editor’s Note: This Manuscript was accepted for publication on July 12, 2021.

Research sponsored by Sanofi and Regeneron Pharmaceuticals, Inc. C.H. has received fees from Sanofi, Smith and Nephew, Medtronic, and Chordate for participation in Advisory Boards. S.K. and N.A. are employees of Regeneron Pharmaceuticals, Inc., and own stock and/or stock options in the company. L.M. and A.H.K. are current employees of Sanofi; M.R. and I.G. are former employees of Sanofi; all may own stock and/or stock options in the company. L.N., D.W., and S.Q. are employees of RTI Health Solutions, a company that received research funding for the current study. The authors have no other funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Asif H. Khan, MBBS, MPH, Sanofi, 1 Avenue Pierre Brossolette, 91380 Chilly-Mazarin, France. E-mail: asif.khan@sanofi.com

DOI: 10.1002/lary.29766

## INTRODUCTION

Chronic rhinosinusitis with nasal polyps (CRSwNP) is a condition predominantly associated with type 2 inflammation<sup>1,2</sup> and significantly impaired health-related quality of life (HRQoL),<sup>3–5</sup> with symptoms including nasal congestion, loss of smell, and rhinorrhea, as well as sleep disturbances.<sup>6,7</sup> Given the symptomatic nature of CRSwNP, patient-reported outcome measures (PROMs) have a crucial role in informing treatment choices.<sup>8,9</sup> The 22-item Sinonasal Outcome Test (SNOT-22) is a PROM designed to evaluate the impact of chronic rhinosinusitis (CRS) on HRQoL.<sup>10</sup> The content of SNOT-22 captures symptom severity, social and emotional impact, productivity, and sleep consequences of CRS. Items are scored from 0 (no problem) to 5 (problem as bad as it can be) and summed to form a total score of 0 to 110.<sup>11</sup>

Based on the Consensus-based Standards for the Selection of Health Status Measurement Instruments checklist,<sup>12,13</sup> SNOT-22 is recognized as one of the most robust CRS-specific PROMs and has been recommended for use in routine clinical evaluation and clinical trials.<sup>14–17</sup> While the total score of SNOT-22 is an appropriate indicator of overall disease impact, greater granularity in outcome scores is required to assess the burden of CRS on patients’ HRQoL, as well as to determine which domains are

impacted, and where treatment is most efficacious when making patient-level decisions about care.<sup>16–20</sup>

Differences in symptoms, symptom severity, effects on HRQoL, and patient-perceived symptom control between patients with CRSwNP and patients with CRS without nasal polyp (CRSsNP)<sup>21</sup> also points to the need for a domain structure that is specific to CRSwNP patients.<sup>16,17</sup> Although there is considerable overlap in the clinical presentation of the CRS subtypes,<sup>3,8</sup> patients with CRSwNP experience nasal congestion, loss of smell, and rhinorrhea more frequently and with greater severity than patients with CRSsNP, who report facial pain more frequently.<sup>6,7</sup> Such differences in symptoms could lead to differences between CRSwNP and CRSsNP<sup>22</sup> in the way aspects of HRQoL are affected, thus potentially resulting in different domain structures underlying SNOT-22 items. Therefore, the objective of the present study was to evaluate SNOT-22 domain structure, using data collected in interventional studies with patients with CRSwNP, and document the cross-sectional and longitudinal psychometric properties of the identified domains. The evaluation followed US Food and Drug Administration guidance for the psychometric evaluation of PROMs.<sup>23</sup>

## METHODS

### Data Sources

Data were from three randomized, placebo-controlled, double-blinded, multicenter clinical trials of dupilumab in adults with moderate-to-severe CRSwNP; one of the studies was a phase 2 trial (NCT01920893 [ACT12340])<sup>24</sup> and two of the studies were phase 3 trials (NCT02912468 [EFC14146] and NCT02898454 [EFC14280])<sup>25</sup> (Table S1). Data from baseline, a mid-treatment time point (week 8 in phase 2, week 16 in phase 3) and end of treatment (week 16 in phase 2, week 24 in phase 3) were analyzed.

In all three trials, SNOT-22 was administered electronically on a tablet and participants were required to answer all items. Additional PROMs data included the Total Symptom Score (TSS; weekly average [range, 0–9]), the rhinosinusitis Visual Analog Scale (VAS) assessing disease severity, the 36-Item Short-Form Health Survey version 2 (SF-36 v2), the EuroQoL-Visual Analog Scale (EQ-VAS), the Lund-Mackay score based on sinus computed tomography scan (LMK-CT), the University of Pennsylvania Smell Identification Test (UPSIT), and the Nasal Polyps Score (NPS) (Table S2).

### Analytic Approach

Analyses were performed using SAS version 9.4 (SAS Institute, Inc, Cary, North Carolina) and Mplus version 7.4 (Muthén & Muthén, Los Angeles, California).

**SNOT-22 Domain Structure.** A preliminary dimensional structure for SNOT-22 for CRSwNP was derived based on the results of an exploratory factor analysis conducted using the phase 2 trial data and further informed by the developer's knowledge of the SNOT-22 conceptual framework, as well as previous published factor structures proposed for CRS.<sup>16–20</sup> An inter-item polychoric correlation matrix with squared multiple correlations as communality estimates, weighted least-squares mean and variance adjusted (WLSMV) estimation, and quartimin rotation were used.

To evaluate the proposed preliminary structure, confirmatory factor analysis was conducted using the pooled phase 3 trial data at baseline and week 24, and WLSMV estimation.

Goodness-of-fit statistics were used to evaluate confirmatory factor analysis model fit. Specific criteria for model fit were as follows: comparative fit index (CFI)<sup>26</sup> and non-normed fit index<sup>27</sup>  $\geq 0.95$  to indicate acceptable fit; and root mean square error of approximation (RMSEA)  $< 0.06$  to indicate satisfactory fit,  $0.06$  to  $0.08$  to indicate fair fit,  $0.08$  to  $0.10$  to indicate mediocre fit, and  $> 0.10$  to indicate poor fit.<sup>28–30</sup>

**Measurement Properties of the SNOT-22 Domain Scores.** Data from both phase 2 and phase 3 trials were used to evaluate floor and ceiling effects, reliability, construct validity, and responsiveness. SNOT-22 domain scores, based on the confirmed factor solution, were computed using the mean of the corresponding SNOT-22 items comprising each of the domains.

**Floor and Ceiling Effects.** SNOT-22 item-level response distributions were examined to identify potential response biases, including floor or ceiling effects, defined as exceeding two-times the expected portion of patients with the lowest or highest scores based on a uniform distribution (e.g., 33.3% for a 6-point scale [0–5]).

**Reliability.** Cronbach's coefficient alpha values (optimal values 0.70–0.90)<sup>31</sup> of the SNOT-22 scores were computed to evaluate internal consistency reliability of the proposed domain scores. Test–retest reliability of the SNOT-22 was evaluated with intraclass correlation coefficients,<sup>32</sup> with test–retest data for a subset of stable patients defined based on no change in rhinosinusitis VAS categories (i.e., mild; 0–3; moderate; >3–7; or severe; >7–10). Week 8 (test) and week 12 (retest) data in the phase 2 trial, and week 16 (test) and week 24 (retest) data in the phase 3 trials, were used to estimate intraclass correlation coefficients from two-way mixed-effects analysis of variance (ANOVA) models for absolute agreement of single measures.<sup>32</sup> Intraclass correlation coefficients  $> 0.70$  indicated substantial agreement.<sup>33</sup>

**Construct Validity.** Convergent and divergent validities of the proposed SNOT-22 domains were assessed using correlation analyses between SNOT-22 domain scores and scores on the other PROMs and clinician-reported outcome measures. Correlations with measures of related PROM constructs (e.g., rhinosinusitis VAS and TSS) were expected to be at least moderate ( $r \geq 0.30$ ),<sup>34</sup> whereas correlations with the clinician-reported outcome measures were expected to be small to moderate ( $r < 0.30$ ).

To assess the discriminating ability of SNOT-22, known-group validity was evaluated using ANOVA to compare the baseline SNOT-22 domain scores for subgroups defined by patients' status on the rhinosinusitis VAS (mild, moderate, or severe) and by the score quartiles on the TSS, UPSIT, LMK, and NPS. Higher SNOT-22 scores (i.e., higher impact on HRQoL) were anticipated for subgroups with a worse status indicated by the known groups.

**Ability to Detect Change.** Ability to detect change, or responsiveness, of the proposed SNOT-22 domains was evaluated using correlation analyses between changes in SNOT-22 and changes in supporting PROMs and clinician-reported measures. Negative change in SNOT-22 scores indicates improvement from baseline to follow-up. The effect-size estimates of change were expressed in units of standard deviation (SD) of changes (i.e., standardized response mean); values  $\geq 0.2 \leq 0.5$  indicate small,  $\geq 0.5 < 0.8$  indicate moderate effects, and  $\geq 0.80$  indicate large effects.<sup>34</sup>

## RESULTS

### Patient Characteristics

The analysis samples comprised 60 patients from the dupilumab phase 2 study and 711 patients from the

pooled phase 3 studies (Table S3). Demographics were broadly comparable between patients from phase 2 and 3 studies with regard to mean age (48.4 and 51.4 years), race (98.3% and 87.6% White), gender (56.7% and 60.3% male), and age of NP onset (38.9 and 40.6 years), respectively. NPS (5.8 and 6.0), UPSIT (14.2 and 14.0), and LMK (18.7 and 18.4) were also similar between the phase 2 and phase 3 patient samples. The inclusion criteria for the phase 3 studies required patients to have moderate to severe symptoms at screening, thus patients from the pooled phase 3 studies had more severe symptoms (based on the rhinosinusitis VAS and TSS) than patients in the phase 2 study (Table S3).

### ***SNOT-22 Domain Structure***

SNOT-22 item response frequencies show no evidence of extreme or unexpected distributional anomalies and indicated improvement in SNOT-22 scores. Descriptive statistics showed floor effects (<63.3%) in some items of the Ear/Facial and Emotion domains at baseline (Fig. S1; Table S4).

A preliminary exploratory factor analysis of phase 2 baseline and end-of-trial polychoric correlations yielded 4–5 factors. Several item cluster trends (i.e., item loadings  $\geq 0.30$ ) were consistently observed across the exploratory factor analysis models: Nasal domain (items 1, 2, 3, 4, 5, 6, 7, 12); Ear/Facial domain (items 8, 9, 10, 11); Sleep domain (items 13, 14, 15, 16); Function domain (items 17, 18, 19); Emotion domain (items 20, 21, 22) (Table I). The major difference between the baseline exploratory factor analysis and the end-of-trial exploratory factor analysis of the phase 2 data was in the combination or separation of the items in the Sleep and Function domains and the items in the Function and Emotion domains.

The clinical relevance of an outcome is important, and this was observed for SNOT-20, which required modification as it excluded two cardinal symptoms of CRSwNP (i.e., nasal congestion and loss of sense of smell/taste).<sup>22</sup> Therefore, the results of the exploratory factor analysis were reviewed by the developer of SNOT-22, clinicians, outcomes researchers, and psychometricians in the project team to evaluate the clinical relevance of the identified domains. Following this review, it was determined that a five-factor model separating the SNOT-22 items would be optimal from both a clinical and outcomes perspective.

The proposed five-domain structure, based on the phase 2 exploratory factor analysis results (Table I), was fitted using confirmatory factor analysis with pooled phase 3 data at baseline (Fig. 1A) and week 24 (Fig. 1B), and showed generally acceptable model fit.

Using the baseline data, all loadings were above 0.6 except for item 12 (decreased sense of smell/taste), which had a loading of 0.49 on the Nasal domain. Goodness-of-fit statistics for the five-factor confirmatory factor analysis with no cross-loading were: CFI, 0.962; Tucker-Lewis index (TLI), 0.956; RMSEA, 0.094 (95% confidence interval [CI], 0.089–0.098). Minor modification to allow item 16 (wake up tired) to load on both the Sleep and Function

Table I.  
SNOT-22 Domains and Items Based on Exploratory Factor Analysis and Confirmatory Factor Analysis.

SNOT-22 Domain	SNOT-22 Item
Nasal	1. Need to blow nose
	2. Nasal blockage
	3. Sneezing
	4. Runny nose
	5. Cough
	6. Post-nasal discharge
	7. Thick nasal discharge
	8. Decreased sense of smell/taste
Ear/Facial	9. Ear fullness
	10. Dizziness
	11. Ear pain
Sleep	12. Facial pain/pressure
	13. Difficulty falling asleep
	14. Wake up at night
	15. Lack of a good night's sleep
Function	16. Wake up tired
	17. Fatigue
	18. Reduced productivity
	19. Reduced concentration
Emotion	20. Frustrated/restless/irritable
	21. Sad
	22. Embarrassed

SNOT-22 = 22-item Sinonasal Outcome Test.

domains reduced RMSEA to 0.076 (95% CI, 0.071–0.081), with the strongest loading of this item (0.487) still on the Sleep domain. Correlations among the factors ranged from 0.53 (Nasal and Emotion domains) to 0.88 (Function and Sleep domains), providing support for the computation of a total score (Table II). Based on the week 24 data, all loadings were consistently above 0.66. Goodness-of-fit statistics for the 5-factor confirmatory factor analysis with no cross-loading were: CFI, 0.975; TLI, 0.971; and RMSEA, 0.084 (95% CI, 0.079–0.089).

### ***SNOT-22 Domain Scores***

Each domain score (of the five domains) was computed as the average score of the corresponding items in the scale. Domain scores ranged from 0 to 5, with lower scores indicating better Nasal, Ear/Facial, Sleep, Function, and Emotion status. Total scores ranged from 0 to 110, with lower scores indicating lower impact on HRQoL.

Mean domain scores (Table III) showed a similar order of magnitude across the phase 2 and pooled phase 3 trials. At baseline, mean domain scores were highest for Nasal (2.6, phase 2; 3.1, phase 3), followed by Sleep (1.8, phase 2; 2.3, phase 3), Function (1.8, phase 2; 2.1, phase 3), Emotion (1.1, phase 2; 1.7, phase 3), and Ear/Facial (1.1, phase 2; 1.4, phase 3).

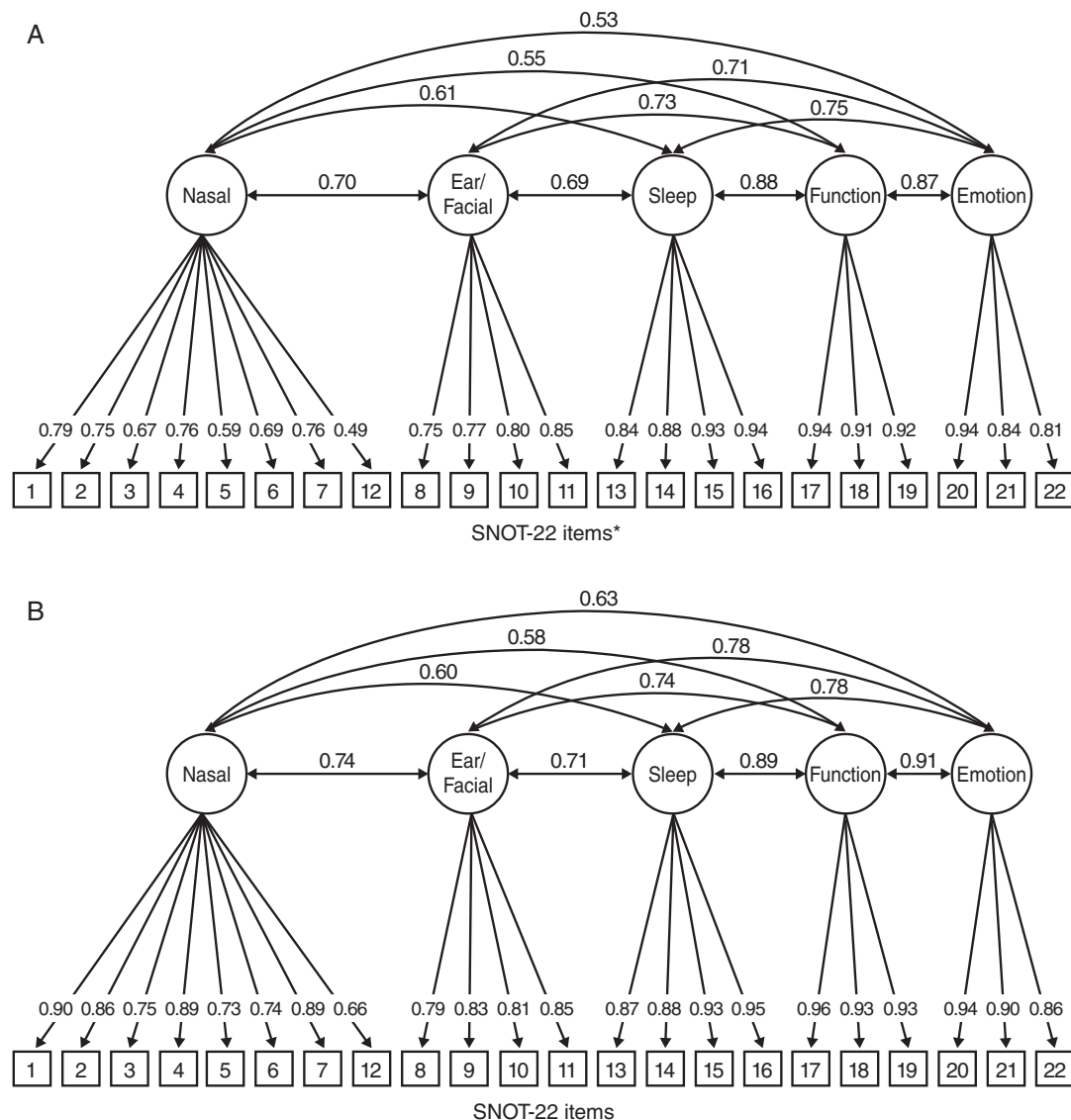


Fig. 1. Five-factor SNOT-22 confirmatory factor analysis based on pooled phase 3 data at (A) baseline; (B) week 24. (A)  $*P < .01$ . Goodness-of-fit statistics for the five-factor confirmatory factor analysis with no cross-loading are as follows: CFI, 0.962; TLI, 0.956; RMSEA, 0.094 (95% CI 0.089–0.098). Minor modification to allow item 16 to load on both the Sleep and Function domains reduced RMSEA to 0.076 (0.071–0.081), with the strongest loading of this item (0.487) still on the Sleep domain. Items: 1. need to blow nose; 2. nasal blockage; 3. sneezing; 4. runny nose; 5. cough; 6. post-nasal discharge; 7. thick nasal discharge; 8. ear fullness; 9. dizziness; 10. ear pain; 11. facial pain/pressure; 12. decreased sense of smell/taste; 13. difficulty falling asleep; 14. wake up at night; 15. lack of a good night's sleep; 16. wake up tired; 17. fatigue; 18. reduced productivity; 19. reduced concentration; 20. frustrated/restless/irritable; 21. sad; 22. embarrassed. (B) Goodness-of-fit statistics for the five-factor confirmatory factor analysis with no cross-loading are as follows: CFI, 0.975; TLI, 0.971; and RMSEA, 0.084 (95% CI, 0.079–0.089). Items: 1. need to blow nose; 2. nasal blockage; 3. sneezing; 4. runny nose; 5. cough; 6. post-nasal discharge; 7. thick nasal discharge; 8. ear fullness; 9. dizziness; 10. ear pain; 11. facial pain/pressure; 12. decreased sense of smell/taste; 13. difficulty falling asleep; 14. wake up at night; 15. lack of a good night's sleep; 16. wake up tired; 17. fatigue; 18. reduced productivity; 19. reduced concentration; 20. frustrated/restless/irritable; 21. sad; 22. embarrassed. CFI = comparative fit index; CI = confidence interval; RMSEA = root mean square error of approximation; SNOT-22 = 22-item Sinonasal Outcome Test; TLI = Tucker-Lewis index.

**Measurement Properties of the SNOT-22 Domains**  
**Floor and Ceiling Effects.** Floor effects (i.e., more than 33.3% reporting no problem) were consistently observed in phase 2 and phase 3 baseline data for the following items: dizziness (phase 2: 60%, phase 3: 50.6%), ear pain (phase 2: 63.3%, phase 3: 57.5%), facial pain/pressure (phase 2: 40%, phase 3: 38.5%), sadness (phase 2: 51.7%, phase 3: 36%), and embarrassment (phase 2: 65%, phase 3: 40.8%). None of the items were

flagged for floor or ceiling effects based on the phase 2 end-of-treatment data. Because the measure was administered electronically, incomplete SNOT-22 submissions were not allowed. The entire range of possible responses (0–5) was utilized in most items. Generally, the SNOT-22 item score means decreased and the variability increased from baseline to end of treatment. This pattern is consistent with the study including active treatment.

Table II.  
SNOT-22 Inter-Domain Correlations and Domain Reliability.

SNOT-22 Domain	Pearson Correlations at Baseline, Phase 3					Cronbach's Alpha at Baseline, Phase 2/Phase 3	Test-Retest ICC (95% CI), Phase 2/Phase 3	
	Nasal	Ear/Facial	Sleep	Function	Emotion		No Change on RS VAS*	No Change on NPS†
Nasal	1.0	–	–	–	–	0.84/0.83	0.82 (0.69, 0.90)/0.86 (0.84, 0.88)	0.83 (0.66, 0.92)/0.84 (0.80, 0.88)
Ear/facial	0.699	1.0	–	–	–	0.81/0.81	0.84 (0.72, 0.91)/0.76 (0.72, 0.80)	0.76 (0.54, 0.88)/0.76 (0.70, 0.81)
Sleep	0.609	0.686	1.0	–	–	0.89/0.91	0.88 (0.78, 0.94)/0.77 (0.74, 0.81)	0.86 (0.71, 0.93)/0.78 (0.72, 0.83)
Function	0.554	0.731	0.877	1.0	–	0.90/0.92	0.90 (0.82, 0.95)/0.75 (0.71, 0.79)	0.91 (0.81, 0.96)/0.78 (0.72, 0.83)
Emotion	0.533	0.707	0.749	0.870	1.0	0.87/0.86	0.84 (0.71, 0.91)/0.80 (0.76, 0.83)	0.85 (0.70, 0.93)/0.80 (0.75, 0.85)
Total	–	–	–	–	–	0.94/0.94	0.92 (0.85, 0.96)/0.86 (0.84, 0.88)	0.89 (0.77, 0.95)/0.85 (0.81, 0.88)

\*No change in rhinosinusitis VAS categories between Week 8 and Week 12.

†No change in NPS between Week 8 and Week 12.

CI = confidence interval; ICC = intraclass correlation coefficient; NPS = Nasal Polyps Score; RS VAS = rhinosinusitis Visual Analog Scale; SNOT-22 = 22-item Sinonasal Outcome Test.

**Reliability.** The internal consistency reliability results supported the proposed domains of SNOT-22 with strong Cronbach's alpha values (all >0.80) in both phases 2 and 3 (Table II). Test-retest reliability intraclass correlation coefficients were above the recommended level of 0.70 for all domains in both trials (Table II).

**Construct Validity.** Construct validity hypotheses were mostly supported with most correlations within the expected direction and size. As expected, the relationships between SNOT-22 scores and the rhinosinusitis VAS and TSS were greater than those between the SNOT-22 scores and UPSIT, LMK, and NPS, suggesting that the clinician-evaluated outcomes measure different concepts than the SNOT-22 domain scores (Table III).

ANOVA provided support for the discriminating ability, or known-groups validity, of SNOT-22 based on PROMs (Table S5). Higher SNOT-22 domain scores (i.e., higher impact) were observed for higher CRSwNP disease-severity groups assessed based on rhinosinusitis VAS, TSS, UPSIT, LMK, and NPS using the pooled phase 3 data. Specifically, significant mean differences in SNOT-22 scores were consistently observed between the subgroup of participants reporting mild or moderate status versus those reporting severe status on the rhinosinusitis VAS, and between the upper and lower quartiles of the TSS, UPSIT, LMK, and NPS ( $P < .05$ ). However, discrimination between subgroups was not as strong using the phase 2 data, possibly because of the small sample sizes of the subgroups. Thus, while significant differences and expected patterns of mean scores of the SNOT-22 total and domains (especially the Nasal domain) were observed across subgroups based on PROMs (rhinosinusitis VAS and TSS individual items), mean differences were not significantly different across subgroups based on clinician-evaluated measures (i.e., UPSIT, LMK, NPS).

**Ability to Detect Change.** The correlations between change in SNOT-22 domain scores and change in rhinosinusitis VAS scores (0.30 [Ear/Facial, phase 2] to 0.61 [Nasal, phase 2]) and TSS (0.36 [Emotion, phase 3] to 0.74 [Nasal and Total, phase 3]) were moderate to strong in magnitude (Table IV). Changes in all SNOT-22

domain scores correlated more strongly with changes in SF-36 mental component summary scores (–0.46 to –0.61) than with changes in SF-36 PCS scores (–0.22 to –0.32) (phase 2; change scores not available in phase 3). The change in UPSIT was strongly correlated with changes in SNOT-22 Nasal scores (–0.62) in phase 2 and moderately correlated (–0.44) in phase 3. The change in EQ-VAS was poorly correlated with most SNOT-22 domain scores, which is expected as EQ-VAS captures general health status.

Standardized response mean estimates for the SNOT-22 Emotion domain were small in phase 2 and moderate (–0.6) in phase 3 trials. Standardized response mean estimates were moderate for the Ear/Facial (–0.6), Sleep (–0.7), and Function (–0.7) domains, and large for the Nasal domain (–1.2) in all trials.

## DISCUSSION

Our study was motivated by the need for a domain structure developed for the relevant population of interest; in this case, for patients with CRSwNP, who report different symptom experience compared with CRSsNP.<sup>21</sup> The present study used data from phase 2 and phase 3 trials of dupilumab to establish five SNOT-22 domains in patients with CRSwNP, namely, Nasal, Ear/Facial, Sleep, Function, and Emotion. We found that test-retest reliability and construct validity of the identified SNOT-22 domain scores satisfied commonly accepted criteria. Specifically, test-retest reliability intraclass correlation coefficients were >0.70, the magnitudes and patterns of validity correlations were mostly consistent with *a priori* hypotheses, and the SNOT-22 domain scores adequately discriminated between levels of rhinosinusitis severity (rhinosinusitis VAS categories) and symptom severity (quartile groups of the TSS). Descriptive statistics highlighted floor effects (<63.3%) in some items of the Ear/Facial and Emotion domains at baseline; however, those items support the internal measurement structure with high item-factor loadings (>0.70) and consistently strong item-domain correlations (>0.50), and both related

Table III.  
SNOT-22 Total/Domain Descriptive Statistics, and Convergent and Divergent Validity.

SNOT-22 Domain	Time	Mean ± SD, Median	Pearson Correlations							
			RS VAS	TSS	SF-36 PCS	SF-36 MCS	EQ-VAS	UPSIT	LMK	NPS
Nasal										
Phase 2	BL	2.6 ± 0.86, 2.6	0.50*	0.69*	-0.45*	-0.31*	-0.45*	-0.35*	0.09	0.07
	End of treatment	1.4 ± 1.04, 1.1	0.73*	0.86*	-0.36*	-0.25	-0.40*	-0.74*	0.68*	0.31*
Phase 3	BL	3.1 ± 0.82, 3.1	0.48*	0.67*	-0.42*	-0.26*	-0.29*	-0.28*	0.25*	0.16*
	Week 24	1.67 ± 1.11, 1.50	0.69*	0.78*	-	-	-0.45*	-0.50*	0.52*	0.42*
Ear/facial										
Phase 2	BL	1.1 ± 1.04, 0.8	0.43*	0.44*	-0.36*	-0.48*	-0.38*	-0.03	-0.24	-0.08
	End of treatment	0.5 ± 0.86, 8.3	0.39*	0.39*	-0.46*	-0.50*	-0.37*	-0.40*	0.19	-0.02
Phase 3	BL	1.4 ± 1.15, 1.3	0.33*	0.37*	-0.39*	-0.42*	-0.29*	-0.09*	0.11*	0.08*
	Week 24	0.69 ± 0.89, 0.25	0.49*	0.48*	-	-	-0.38*	-0.29*	0.27*	0.23*
Sleep										
Phase 2	BL	1.8 ± 1.24, 1.8	0.24*	0.47*	-0.28*	-0.17	-0.25	-0.09	-0.00	0.13
	End of treatment	0.8 ± 1.01, 0.5	0.41*	0.61*	-0.50*	-0.32*	-0.45*	-0.43*	0.34*	0.13
Phase 3	BL	2.3 ± 1.40, 2.3	0.35*	0.38*	-0.34*	-0.34*	-0.31*	-0.13*	0.14*	0.17*
	Week 24	1.30 ± 1.20, 1.0	0.45*	0.48*	-	-	-0.43*	-0.27*	0.20*	0.26*
Function										
Phase 2	BL	1.8 ± 1.29, 1.7	0.28*	0.46*	-0.55*	-0.50*	-0.44*	-0.05	-0.17	-0.06
	End of treatment	0.8 ± 1.03, 0.3	0.35*	0.50*	-0.56*	-0.62*	-0.42*	-0.36*	0.16	-0.03
Phase 3	BL	2.1 ± 1.39, 2.3	0.33*	0.32*	-0.39*	-0.49*	-0.36*	-0.12*	0.11*	0.09*
	Week 24	1.19 ± 1.17, 1.0	0.45*	0.46*	-	-	-0.47*	-0.27*	0.20*	0.23*
Emotion										
Phase 2	BL	1.1 ± 1.12, 0.7	0.35*	0.41*	-0.30*	-0.60*	-0.48*	-0.10	-0.11	0.10
	End of treatment	0.5 ± 0.84, 0.0	0.41*	0.45*	-0.20	-0.63*	-0.22	-0.31*	0.11	-0.03
Phase 3	BL	1.7 ± 1.36, 1.7	0.31*	0.32*	-0.28*	-0.64*	-0.37*	-0.10*	0.08*	0.13*
	Week 24	0.9 ± 1.08, 0.33	0.47*	0.46*	-	-	-0.42*	-0.26*	0.20*	0.24*
Total score										
Phase 2	BL	41.0 ± 18.92, 40.5	0.46*	0.64*	-0.49*	-0.48*	-0.49*	-0.19	-0.08	0.04
	End of treatment	20.5 ± 17.55, 17.0	0.64*	0.78*	-0.50*	-0.49*	-0.47*	-0.64*	0.51*	0.16
Phase 3	BL	50.9 ± 20.67, 50.0	0.45*	0.53*	-0.45*	-0.48*	-0.39*	-0.19*	0.18*	0.16*
	Week 24	27.62 ± 20.17, 23.0	0.65*	0.69*	-	-	-0.52*	-0.42*	0.39*	0.37*

Magnitudes of the correlations are evaluated as weak ( $r < 0.30$ ), moderate ( $r = 0.30-0.49$ ), or strong ( $r \geq 0.50$ ).

\* $P < .05$ .

BL = baseline; EQ-VAS = EuroQoL-Visual Analog Scale; LMK = Lund-Mackay score; MCS = mental component summary; NPS = Nasal Polyps Score; PCS = physical component summary; RS VAS = rhinosinusitis Visual Analog Scale; SD = standard deviation; SF-36 = 36-Item Short-Form Health Survey; SNOT-22 = 22-item Sinonasal Outcome Test; TSS = Total Symptom Score (of the symptoms e-diary; computed as the sum of the weekly averages of nasal congestion/obstruction item, the average of the 2 anterior/posterior rhinorrhea items, and the loss of sense of smell item); UPSIT = University of Pennsylvania Smell Identification Test.

domains were sensitive to change when defined by other outcome measures. Previous studies have shown a wide spread of symptoms in patients with CRSwNP prior to surgical intervention, in which all symptoms had a mean score above 1 and showed improvement after surgery.<sup>22</sup>

The SNOT-22 domain scores also demonstrated ability to detect change, as evidenced by the moderate-to-strong responsiveness correlations with changes in other PROMs and clinician-reported measures. The Emotion and Ear/Facial domains were less responsive than the other SNOT-22 domains and had small-to-moderate sizes of standardized response mean, whereas the other domains had moderate-to-large sizes of standardized response mean. Insofar as previous studies have also identified five domains of SNOT-22,<sup>16,17</sup> these evaluations were conducted in patients with mixed types of CRS. Within previous studies,

three domains (Ear/Facial, Sleep dysfunction, and Psychological dysfunction) contained the same or similar SNOT-22 items to the Ear/Facial, Sleep, and Emotion domains identified in our study. However, the other two domains identified (Rhiniologic symptoms and Extra-nasal rhiniologic symptoms)<sup>16,17</sup> were different to those identified in our study (Nasal and Function). Such differences in identified domains might have arisen from variation in patient characteristics in our study compared with previous studies assessing SNOT-22 domains; for example, inclusion of patients with medically refractory CRS, of whom only 39% had nasal polyposis. Conversely, our study comprised a homogenous sample of patients with CRSwNP. Furthermore, the previous studies<sup>16,17</sup> included patients from the United States only, while our study was based on data from three international clinical trials.<sup>24,25</sup> Similarly, a previously published four-

Table IV.  
Change in SNOT-22 Scores and Responsiveness Correlations.

SNOT-22 Domain	Phase	Mean Change ± SD (n = 52,674)	Standardized Response Mean	Responsiveness Pearson Correlations Between Change Scores							
				RS VAS	TSS	SF-36 PCS	SF-36 MCS	EQ-5D VAS	UPSIT	LMK	NPS
Nasal	2	-1.2 ± 1.04	-1.2	0.61*	0.74*	-0.32*	-0.39*	-0.43*	-0.62*	0.40*	0.35*
	3	-1.4 ± 1.14	-1.2	0.57*	0.71*	NA	NA	-0.29*	-0.44*	0.43*	0.47*
Ear/facial	2	-0.6 ± 1.05	-0.6	0.30*	0.41*	-0.23	-0.54*	-0.17	-0.37*	0.14	0.18
	3	-0.7 ± 1.09	-0.6	0.35*	0.40*	NA	NA	-0.19*	-0.21*	0.20*	0.26*
Sleep	2	-0.9 ± 1.26	-0.7	0.54*	0.71*	-0.22	-0.46*	-0.17	-0.42*	0.30*	0.26
	3	-1.0 ± 1.41	-0.7	0.35*	0.43*	NA	NA	-0.24*	-0.22*	0.22*	0.29*
Function	2	-0.9 ± 1.26	-0.7	0.42*	0.60*	-0.32*	-0.61*	-0.25	-0.33*	0.23	0.24
	3	-0.9 ± 1.34	-0.7	0.36*	0.39*	NA	NA	-0.25*	-0.21*	0.21*	0.23*
Emotion	2	-0.5 ± 1.34	-0.4	0.40*	0.51*	-0.32*	-0.51*	-0.31*	-0.33*	0.26	0.04
	3	-0.8 ± 1.23	-0.6	0.31*	0.36*	NA	NA	-0.24*	-0.20*	0.18*	0.18*
Total score	2	-19.9 ± 21.46	-0.9	0.57*	0.74*	-0.33*	-0.56*	-0.34*	-0.53*	0.35*	0.28*
	3	-22.63 ± 21.90	-1.0	0.51*	0.62*	NA	NA	-0.30*	-0.35*	0.35*	0.40*

Magnitudes of the correlations are evaluated as weak ( $r < 0.30$ ), moderate ( $r = 0.30-0.49$ ), or strong ( $r \geq 0.50$ ). Magnitudes of standardized response mean are evaluated as small (0.20-0.49), moderate (0.50-0.79), or large ( $\geq 0.80$ ).

\* $P < .05$ .

EQ-5D = EuroQol 5-dimensional; LMK = Lund-Mackay score; MCS = mental component summary; NA = not available; NPS = Nasal Polyps Score; PCS = physical component summary; RS VAS = rhinosinusitis Visual Analog Scale; SD = standard deviation; SF-36 = 36-Item Short-Form Health Survey; SNOT-22 = 22-item Sinonasal Outcome Test; TSS = Total Symptom Score (of the symptoms e-diary; computed as the sum of the weekly averages of nasal congestion/obstruction item, the average of the 2 anterior/posterior rhinorrhea items, and the loss of sense of smell item); UPSIT = University of Pennsylvania Smell Identification Test; VAS = visual analog scale.

domain model of SNOT-22<sup>20</sup> was also developed in a heterogeneous sample of patients with CRS. To determine the best-fitting domain structure, further research is warranted using an independent dataset and comparing those models against the domain structure identified in our study.

Although clinically relevant CRS phenotypes can be generally defined by observable characteristic (including the presence or absence of nasal polyps), recognition of CRS heterogeneity has promoted the concept of multiple “endotypes”, which have distinct underlying pathophysiologic mechanisms.<sup>1,35</sup> This improved understanding has aided development of biologic agents for CRS management.<sup>36</sup> Treatment recommendations are now tailored to this broader classification of CRS phenotypes beyond just the presence of NP.<sup>37</sup> The distinct symptomatology across the different CRS phenotypes and endotypes highlights the need for phenotype-specific PROMs, in addition to correlating endotype physiology with existing and/or specific scales. We identified five SNOT-22 domains in our sample of patients with CRSwNP, which is characterized in Western countries by dominance of the type 2 endotype (>80% of patients),<sup>38-40</sup> many of whom would be considered to have type 2 inflammation. We hypothesize that the domain structure would apply to all CRS patients with type 2 inflammation, but further work would be needed to confirm this assumption. The SNOT-22 domain scores can be used alongside the SNOT-22 total score to provide more granular, empirically derived data in clinical research and practice, which allows physicians to tailor treatment options, and can support a more detailed understanding of the impact and burden of CRSwNP across the domains in different patients, for which the total score does not discriminate. For example, patients who present with primary complaints on the Nasal

domain but no impairment on Sleep may require a different treatment approach than patients who present with both nasal complaints and sleep impairment. The ability to obtain clinically relevant and interpretable domain scores from SNOT-22 provides valuable information on the HRQoL impact of CRSwNP to inform treatment decision-making and is especially pertinent to personalized medicine.

### Study Strengths and Limitations

The strengths of this study include the independent and large size of the sample provided by the phase 3 studies and the interventional, placebo-controlled study design, allowing for evaluation of the ability to detect change. This approach of using phase 3 data for the confirmatory analysis was followed because of the timelines of the studies, with phase 2 data used to inform the conduct of the phase 3 trials. Although the phase 2 sample was relatively small ( $n = 60$ ), the confirmatory analysis conducted in the larger phase 3 sample ( $n = 711$ ) should mitigate any limitations with using the phase 2 sample. In addition, as the studies were conducted globally, the racial homogeneity of the sample should be noted; however, there is no reason to suggest that scores or domains of the SNOT-22 would be culturally variable, although further study would be needed to confirm this assumption.

The findings of our study must be considered in light of a few limitations. First, the trials comprised only patients with moderate or severe CRSwNP; hence, until further psychometric validation is conducted in a wider patient group to include patients with mild CRSwNP, the results may have limited external validity for patients

with less severe disease. However, the use of end-of-treatment data from phase 2 and phase 3 trials, at which point many patients' symptoms would have become mild-to-moderate, provides indirect support for observations in a population with mild CRSwNP. It is also noteworthy that further research is needed to identify changes in scores that imply the clinical meaningful change in these domains. Further investigation of the validity of identified domain scores in patients with other phenotypes of CRS would also be of interest, particularly, given the focus on improving diagnostic approaches and treatment strategies for each CRS phenotype.<sup>37</sup> Finally, although the clinical relevance the proposed domains was ensured through input from clinicians during identification of these domains, further testing would be required to demonstrate their relevance to patients with CRSwNP.

Despite these limitations, our psychometric analyses support the validity, reliability, and responsiveness of five domains of SNOT-22 (Nasal, Ear/Facial, Sleep, Function, and Emotion) suitable for assessing symptoms and the impact of CRSwNP on HRQoL.

## Acknowledgments

Medical writing support was provided by Gauri Saal, MA Economics, and Abby Armitt, MSc, and Editorial Support was provided by Ian Norton, PhD, all of Prime, UK, supported by Sanofi and Regeneron Pharmaceuticals, Inc., according to Good Publication Practice guidelines.

## Author Contributions

The authors were responsible for all content and editorial decisions, and received no honoraria related to the development of this publication. All authors had full access to all of the data in this study and take complete responsibility for the integrity of the data and accuracy of the data analysis. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this manuscript, take responsibility for the integrity of the manuscript as a whole, and have given final approval for the version to be published.

## REFERENCES

- Dennis SK, Lam K, Luong A. A review of classification schemes for chronic rhinosinusitis with nasal polyposis endotypes. *Laryngoscope Investig Otolaryngol* 2016;1:130–134.
- Groger M, Bernt A, Wolf M, et al. Eosinophils and mast cells: a comparison of nasal mucosa histology and cytology to markers in nasal discharge in patients with chronic sino-nasal diseases. *Eur Arch Otorhinolaryngol* 2013;270:2667–2676.
- Stevens WW, Schleimer RP, Kern RC. Chronic rhinosinusitis with nasal polyps. *J Allergy Clin Immunol Pract* 2016;4:565–572.
- Naclerio RM, Baroody FM, Pinto JM. Nasal polyps and biomarkers. *J Allergy Clin Immunol Pract* 2017;5:1589–1590.
- Khan A, Vandeplas G, Huynh TMT, et al. The Global Allergy and Asthma European Network (GALEN) rhinosinusitis cohort: a large European cross-sectional study of chronic rhinosinusitis patients with and without nasal polyps. *Rhinology* 2019;57:32–42.
- Huvenne W, van Bruaene N, Zhang N, et al. Chronic rhinosinusitis with and without nasal polyps: what is the difference? *Curr Allergy Asthma Rep* 2009;9:213–220.
- Dietz de Loos DA, Hopkins C, Fokkens WJ. Symptoms in chronic rhinosinusitis with and without nasal polyps. *Laryngoscope* 2013;123:57–63.
- Fokkens WJ, Lund VJ, Mullol J, et al. European position paper on rhinosinusitis and nasal polyps 2012. A summary for otorhinolaryngologists. *Rhinology* 2012;50:1–12.
- Rosenfeld RM, Andes D, Bhattacharyya N, et al. Clinical practice guideline: adult sinusitis. *Otolaryngol Head Neck Surg* 2007;137:S1–S31.
- Hopkins C, Gillett S, Slack R, Lund VJ, Browne JP. Psychometric validity of the 22-item Sinonasal Outcome Test. *Clin Otolaryngol* 2009;34:447–454.
- Levy JM, Mace JC, Rudmik L, Soler ZM, Smith TL. Low 22-item sinonasal outcome test scores in chronic rhinosinusitis: why do patients seek treatment? *Laryngoscope* 2017;127:22–28.
- Rudmik L, Hopkins C, Peters A, Smith TL, Schlosser RJ, Soler ZM. Patient-reported outcome measures for adult chronic rhinosinusitis: a systematic review and quality assessment. *J Allergy Clin Immunol* 2015;136:1532–1540.e1532.
- Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539–549.
- Rudmik L, Hopkins C, Mace JC, DeConde AS, Schlosser RJ, Smith TL. Using preoperative SNOT-22 score to inform patient decision for endoscopic sinus surgery. *Laryngoscope* 2015;125:1517–1522.
- Kennedy JL, Hubbard MA, Huyett P, Patrie JT, Borish L, Payne SC. Sinonasal outcome test (SNOT-22): a predictor of postsurgical improvement in patients with chronic sinusitis. *Ann Allergy Asthma Immunol* 2013;111:246–251.e242.
- DeConde AS, Bodner TE, Mace JC, Smith TL. Response shift in quality of life after endoscopic sinus surgery for chronic rhinosinusitis. *JAMA Otolaryngol Head Neck Surg* 2014;140:712–719.
- DeConde AS, Mace JC, Bodner T, et al. SNOT-22 quality of life domains differentially predict treatment modality selection in chronic rhinosinusitis. *Int Forum Allergy Rhinol* 2014;4:972–979.
- Dejaco D, Riedl D, Huber A, et al. The SNOT-22 factorial structure in European patients with chronic rhinosinusitis: new clinical insights. *Eur Arch Otorhinolaryngol* 2019;276:1355–1365.
- Sedaghat AR, Gray ST, Caradonna SD, Caradonna DS. Clustering of chronic rhinosinusitis symptomatology reveals novel associations with objective clinical and demographic characteristics. *Am J Rhinol Allergy* 2015;29:100–105.
- Feng AL, Wesely NC, Hoehle LP, et al. A validated model for the 22-item Sino-Nasal Outcome Test subdomain structure in chronic rhinosinusitis. *Int Forum Allergy Rhinol* 2017;7:1140–1148.
- Talat R, Speth MM, Gengler I, et al. Chronic rhinosinusitis patients with and without polyps experience different symptom perception and quality of life burdens. *Am J Rhinol Allergy* 2020;34:742–750.
- Abdalla S, Alreefy H, Hopkins C. Prevalence of sinonasal outcome test (SNOT-22) symptoms in patients undergoing surgery for chronic rhinosinusitis in the England and Wales National prospective audit. *Clin Otolaryngol* 2012;37:276–282.
- U.S. Food and Drug Administration. Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims>. Accessed September 14, 2020.
- Jonstam K, Swanson BN, Mannent LP, et al. Dupilumab reduces local type 2 pro-inflammatory biomarkers in chronic rhinosinusitis with nasal polyposis. *Allergy* 2019;74:743–752.
- Bachert C, Han JK, Desrosiers M, et al. Efficacy and safety of dupilumab in patients with severe chronic rhinosinusitis with nasal polyps (LIBERTY NP SINUS-24 and LIBERTY NP SINUS-52): results from two multicentre, randomised, double-blind, placebo-controlled, parallel-group phase 3 trials. *Lancet* 2019;394:1638–1650.
- Bentler PM. *EQS Structural Equations Program Manual*. Los Angeles, CA: BMDP Statistical Software; 1989.
- Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika* 1973;38:11.
- Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, eds. *Testing Structural Equation Models*. Newbury Park, CA: Sage; 1993.
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model* 1999;6:56.
- Schumacker RE, Lomax RG. *A Beginner's Guide to Structural Equation Modeling*. Mahwah NJ: Lawrence Erlbaum Associates, Inc.; 1996.
- Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297–334.
- McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods* 1996;1:30–46.
- Nunnally JC, Bernstein IH. *Psychometric Theory*. New York, NY: McGraw-Hill; 1994.
- Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Routledge; 1988.



35. Lötvall J, Akdis CA, Bacharier LB, et al. Asthma endotypes: a new approach to classification of disease entities within the asthma syndrome. *J Allergy Clin Immunol* 2011;127:355–360.
36. Wang C, Zhang L. Use of biologics in chronic sinusitis with nasal polyps. *Curr Opin Allergy Clin Immunol* 2019;19:365–372.
37. Cho SH, Hamilos DL, Han DH, Laidlaw TM. Phenotypes of chronic rhinosinusitis. *J Allergy Clin Immunol Pract* 2020;8:1505–1511.
38. Fokkens WJ, Lund VJ, Hopkins C, et al. European position paper on rhinosinusitis and nasal polyps 2020. *Rhinology* 2020;58:1–464.
39. Bachert C, Han JK, Wagenmann M, et al. EUFOREA expert board meeting on uncontrolled severe chronic rhinosinusitis with nasal polyps (CRSwNP) and biologics: definitions and management. *J Allergy Clin Immunol* 2021;147:29–36.
40. Naclerio R, Baroody F, Bachert C, et al. Clinical research needs for the management of chronic rhinosinusitis with nasal polyps in the new era of biologics: a National Institute of Allergy and Infectious Diseases workshop. *J Allergy Clin Immunol Pract* 2020;8:1532–1549. e1531.